



**Universidade Estadual de Campinas**

**LCSI**

Faculdade de Engenharia Elétrica e Computação

Departamento de Máquinas, Componentes e Sistemas Inteligentes

Laboratório de Controle e Sistemas Inteligentes

**Modelagem Computacional Distribuída e Paralela de Sistemas e de Séries Temporais Multivariáveis no Espaço de Estado**

Tese apresentada à Faculdade de Engenharia Elétrica e Computação da Universidade Estadual de Campinas, como parte dos requisitos exigidos para a obtenção do título de

Doutor em Engenharia Elétrica  
por

**Gilmar Barreto**

Mestre em Engenharia Elétrica - UNICAMP/SP

**Prof. Dr. Celso Pascoli Bottura**

Orientador - FEEC/UNICAMP/SP

12 de junho de 2002

FICHA CATALOGRÁFICA ELABORADA PELA  
BIBLIOTECA DA ÁREA DE ENGENHARIA - BAE - UNICAMP

Barreto, Gilmar

B275m

Modelagem computacional distribuída e paralela de sistemas e de séries temporais multivariáveis no espaço de estado. / Gilmar Barreto. -- Campinas, SP: [s.n.], 2002.

Orientador: Celso Pascoli Bottura.

Tese (doutorado) - Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Análise de séries temporais - Processamento de dados.  
2. Identificação. 3. Processamento paralelo ( Computadores ).  
4 . Processamento eletrônico de dados - Processamento distribuído. 5 Sistemas lineares. 6. Métodos de espaço de estados. 7. Teoria da realização. 8 . Métodos de subespaços.  
I. Bottura, Celso Pascoli. II. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. III. Título.



## **Universidade Estadual de Campinas**

**LCSI**

Faculdade de Engenharia Elétrica e Computação

Departamento de Máquinas, Componentes e Sistemas Inteligentes

**Laboratório de Controle e Sistemas Inteligentes**

Tese : **Modelagem Computacional Distribuída e Paralela de Sistemas e de Séries Temporais Multivariáveis no Espaço de Estado**

Autor : **Gilmar Barreto**

Orientador : **Prof. Dr. Celso Pascoli Bottura**

Aprovada em 12 de junho de 2002 pela banca examinadora:

- Prof. Dr. Celso Pascoli Bottura (Presidente)
- Prof. Dr. Luiz Carlos Sandoval Góes - ITA
- Prof. Dr. João Viana da Fonseca Neto - UFMA
- Prof. Dr. Wagner Caradori do Amaral - UNICAMP
- Prof. Dr. Luis Gimeno Latre - UNICAMP
- Prof. Dr. André Luiz Morelato França - UNICAMP
- Prof. Dr. Carlos Alberto de Castro Jr. - UNICAMP
- Prof. Dr. Marconi Kolm Madrid - UNICAMP

*“ Ando devagar porque já tive pressa  
e levo esse sorriso, porque já chorei demais.  
Hoje me sinto mais forte, mais feliz quem sabe  
eu só levo a certeza de que muito pouco eu sei, eu nada sei ... ”*

Almir Sater e Renato Teixeira

À minha Esposa, Dulcinéa  
Aos meus Pais, Antonio e Iolanda  
Aos meus Irmãos, Gilson e Giovana  
Aos meus Sobrinhos, Rodrigo e Lucas

# Agradecimentos

Gostaria de demonstrar, aqui, minha gratidão a todas as pessoas, que me ajudaram na realização deste trabalho. Em especial,

- Ao Prof. Celso Pascoli Bottura, pela fé e coragem de acreditar sempre na possibilidade da realização deste trabalho, pela dedicação e honestidade na orientação, pelo companheirismo e amizade e pela oportunidade.
- Aos Professores Yaro Burian Jr., Ioshiaki Doi, César José Bonjuani Pagan, Ana Cristina Cavalcanti Lyra, Carlos Alberto de Castro Jr., Carlos Alberto Favarin Murari, Celso Munaro, Rogério Quirino, João Viana da Fonseca Neto, Angel Fernando Torrico Caceres, José Tarcísio Costa Filho, Maurício José Bordon e Annabell Del Real Tamariz pela convivência e pelas didáticas intervenções.
- Aos colegas do DMCSI, Sérgio, Gustavo, Tetsu, Leandro, Rosana, Márcio e Trettel pelos inúmeros auxílios.
- Aos colegas do CENAPAD, Ana Drumond, Ana Seixas e Ricardo Kusel.
- Às colegas da FEEC, Edna Servidone e Alaíde da Silva Ramos.
- Às colegas da BAE - Biblioteca da Área de Engenharia, Ana Paula de Moraes, Raquel Cocatto Ribeiro e Rose Meire da Silva pelo auxílio na obtenção de diversas referências utilizadas nesta tese.
- Finalmente, à UNICAMP.

# Sumário

<b>RESUMO</b>	<b>xi</b>
<b>PREFÁCIO</b>	<b>xii</b>
<b>1 INTRODUÇÃO À MODELAGEM DE DADOS</b>	<b>1</b>
1.1 Modelagem de Dados . . . . .	1
1.2 Sistema Dinâmico . . . . .	3
1.2.1 Conceitos Sobre Modelagem de Dados . . . . .	4
1.2.2 Tipos de Identificação de Sistemas . . . . .	16
1.2.3 Critérios Para Escolha do Estimador . . . . .	19
1.2.4 Alguns Métodos Para Minimizar o Erro do Estimador	22
1.2.5 Mínimos Quadrados . . . . .	23
1.2.6 Mínimos Quadrados Recursivo . . . . .	26
1.3 Teoria da Realização . . . . .	30
1.4 Métodos de Subespaços . . . . .	32
<b>2 ALGUNS ELEMENTOS DE ÁLGEBRA LINEAR E ÁLGEBRA LINEAR NUMÉRICA</b>	<b>37</b>
2.1 Introdução . . . . .	37
2.2 Alguns Conceitos Fundamentais . . . . .	37
2.2.1 Representação de Um Vetor em Relação a Uma Base e Mudança de base . . . . .	42
2.2.2 Operadores Lineares . . . . .	44
2.2.3 Soma Direta de Subespaços . . . . .	47
2.2.4 Subespaço Invariante . . . . .	48
2.2.5 Representação de Um Operador Linear Por Uma Matriz	49
2.2.6 Exemplos de Operadores Lineares . . . . .	51
2.2.7 Particularidade de Um Operador Linear . . . . .	55
2.2.8 Distância, Norma e Produto Interno em Um Espaço Vetorial . . . . .	57
2.2.9 Matriz Ortogonal . . . . .	60
2.2.10 Projeção . . . . .	61

2.2.11	Projeção Ortogonal . . . . .	62
2.2.12	Reflexão . . . . .	64
2.3	Transformações Ortogonais de Householder . . . . .	67
2.3.1	Propriedades da Matriz de Householder . . . . .	71
2.4	Fatoração <b>QR</b> . . . . .	72
2.4.1	Fatoração <b>QR</b> Através de Matrizes de Householder . . . . .	74
2.4.2	Fatoração <b>QR</b> Através de Rotações de Givens . . . . .	77
2.5	O método <b>QR</b> . . . . .	83
<b>3</b>	<b>PROCESSAMENTO DE ALTO DESEMPENHO DE ALGORITMOS NUMÉRICOS</b>	<b>87</b>
3.1	Introdução . . . . .	87
3.1.1	Medidas de Desempenho para o Processamento Paralelo	93
3.1.2	Arquiteturas Distribuídas e Paralelas . . . . .	98
3.1.3	Praticabilidade de Algoritmos Distribuídos e Paralelos	108
3.2	Ambientes de Programação Distribuída e Paralela . . . . .	109
3.2.1	<i>PVM - Parallel Virtual Machine</i> . . . . .	109
3.2.2	<i>MPI - Message Passing Interface</i> . . . . .	112
3.2.3	ScaLAPACK - Scalable Linear Algebra PACKage . . . . .	113
3.2.4	Algoritmos por Blocos . . . . .	114
3.2.5	Utilização de Algoritmos por Blocos . . . . .	115
3.3	Algoritmos Numéricos Paralelos . . . . .	117
3.3.1	Solução Paralela do Sistema $\mathbf{Ax} = \mathbf{b}$ . . . . .	117
3.3.2	Um Procedimento Recursivo Para Computação de Matriz Pseudo-Inversa . . . . .	120
3.4	Uma Proposta de Método de Identificação Paralela Recursiva de Mínimos Quadrados . . . . .	128
3.4.1	Identificação Paralela Via Fatoração Triangular . . . . .	128
3.5	Computação Paralela da Decomposição em Valores Singulares	130
3.5.1	Experimentos Com a Computação da Decomposição em Valores Singulares . . . . .	132
3.6	Proposta de Procedimento Para Avaliação de Desempenho Numérico de Algoritmos Matriciais . . . . .	136
3.6.1	Cálculo do Número de Operações de Algoritmos Matriciais . . . . .	137
3.6.2	Aplicação do Método de Avaliação Proposto . . . . .	138
<b>4</b>	<b>DE PSEUDO-INVERSA, PROJEÇÕES, DECOMPOSIÇÃO EM VALORES SINGULARES E SOLUÇÃO DE MÍNIMOS QUADRADOS EM SUBESPAÇOS</b>	<b>141</b>
4.1	Introdução . . . . .	141



4.2	Estrutura da Transformação Linear:	
	Subespaços Fundamentais . . . . .	143
4.3	Diferentes Tipos de Solução de $\mathbf{Ax} = \mathbf{b}$ . . . . .	149
4.4	Matrizes Projção Ortogonal e Pseudo Inversa de $\mathbf{A}$ à Esquerda	156
4.5	Solução do Problema $\mathbf{Ax} = \mathbf{b}$ nos Subespaços Fundamentais .	160
4.6	Solução de Mínimos Quadrados de Sistema Sobredeterminado	
	Através de Matriz Pseudo-Inversa . . . . .	167
4.7	Matrizes Projção Ortogonal e Pseudo-Inversa à Direita . . . .	176
4.8	Solução de Mínimos Quadrados de Sistema Subdeterminado	
	Através de Matriz Pseudo-Inversa . . . . .	179
4.9	Princípios para Decomposição em Valores Singulares . . . . .	187
4.9.1	Diagonalização de Uma Matriz Quadrada . . . . .	188
4.9.2	Diagonalização de Uma Matriz Qualquer . . . . .	190
4.9.3	Decomposição em Valores Singulares e Pseudo-Inversa	
	de Matriz . . . . .	203
4.9.4	Interpretações Geométricas da Decomposição em Va-	
	lores Singulares . . . . .	205
4.9.5	Propriedades da Decomposição em Valores Singulares .	210
4.10	Projção Oblíqua . . . . .	211
4.11	Ângulo e Direções Principais Entre Subespaços . . . . .	220
4.12	Ferramentas Estatísticas e Ferramentas Geométricas . . . . .	227
<b>5</b>	<b>REALIZAÇÃO, OPERADOR DE HANKEL E FUNDAMEN-</b>	
	<b>TOS PARA MODELAGEM DE DADOS NO ESPAÇO DE</b>	
	<b>ESTADO</b>	<b>231</b>
5.1	Introdução . . . . .	231
5.1.1	Resposta ao Impulso de Sistema Monovariável . . . .	233
5.1.2	Resposta ao Impulso de Sistema Multivariável . . . .	234
5.2	Realização Mínima da Resposta ao Impulso . . . . .	236
5.3	Realização e Operador de Hankel . . . . .	238
5.4	Gênese de Operador de Hankel . . . . .	242
5.5	Propriedades da Matriz de Hankel . . . . .	245
5.6	Modelos Estendidos no Espaço de Estado . . . . .	250
5.7	Métodos de Subespaço para Modelagem de Dados: Idéias	
	Básicas. . . . .	257
5.7.1	Projeções Ortogonais para um Modelo de Estado Es-	
	tendido . . . . .	258
5.7.2	Algumas Propriedades Geométricas de Um Modelo de	
	Estado Estendido . . . . .	270
5.8	Realização Mínima Estocástica de Sequência de Covariâncias	276

5.9	Sistemas com Entradas Exógenas: Os Subsistemas Determinístico e Estocástico . . . . .	278
<b>6</b>	<b>IDENTIFICAÇÃO MULTIVARIADA PARALELA E DISTRIBUÍDA VIA TRATAMENTO MOESP</b>	<b>293</b>
6.1	Introdução . . . . .	293
6.2	Método de Subespaços . . . . .	295
6.3	Implementação de Alto Desempenho De Método de Subespaço	297
6.3.1	Cálculo das Matrizes <b>A</b> e <b>C</b> . . . . .	297
6.3.2	Cálculo das Matrizes <b>B</b> e <b>D</b> . . . . .	298
6.3.3	Proposta de Implementação Multiprocessada . . . . .	299
6.4	Ambiente de Programação . . . . .	299
6.4.1	Aspectos Paralelos Para Processamento Computacional	301
6.4.2	Estrutura da rede de estações de trabalho . . . . .	302
6.4.3	Subrotina para Fatoração <b>QR</b> . . . . .	303
6.4.4	Subrotina para Decomposição em Valores Singulares .	304
6.4.5	Especificações para Subrotinas em FORTRAN . . . . .	304
6.5	Resultados e Comentários . . . . .	305
<b>7</b>	<b>MODELAGEM PARALELA E DISTRIBUÍDA DE SÉRIES TEMPORAIS EM SUBESPAÇOS DO ESPAÇO DE ESTADO</b>	<b>311</b>
7.1	Introdução . . . . .	311
7.2	Modelagem Paralela de Séries Temporais no Espaço de Estado	312
7.2.1	Solução Paralela da Decomposição em Valores Singulares	319
7.2.2	Solução Paralela da Equação Algébrica de Ricatti . .	319
7.3	Resultados Experimentais . . . . .	324
7.3.1	Caso A . . . . .	324
7.3.2	Caso B . . . . .	325
7.3.3	Resultados Computacionais: Equação Algébrica de Ricatti . . . . .	327
7.4	Avaliação da Modelagem de Séries Temporais Multivariadas no Espaço de Estado . . . . .	328
7.5	Conclusões . . . . .	331
<b>8</b>	<b>CONCLUSÕES E COMENTÁRIOS</b>	<b>333</b>
	<b>BIBLIOGRAFIA</b>	<b>335</b>

# Modelagem Computacional Distribuída e Paralela de Sistemas e de Séries Temporais Multivariáveis no Espaço de Estado

## Resumo

Este estudo primeiramente investiga fundamentos teóricos para análise, desenvolvimento e implementação de algoritmos para modelagem de dados de sistemas dinâmicos e de séries temporais multivariáveis no espaço de estado, através de métodos de subespaço. Tem como segundo objetivo o desenvolvimento e implementação de algoritmos para modelagem computacional distribuída e paralela destes tipos de dados multivariados. A modelagem computacional de dados no espaço de estado é apresentada, comentada e avaliada sobre “*benchmarks*”. Desta forma esperamos viabilizar uma metodologia original e eficiente que contribuirá de forma direta para a modelagem de sistemas multivariáveis e de formas direta e ou indireta para o controle de sistemas multivariáveis.

## Abstract

This study investigates firstly theoretical foundations in analysis, development and implementation of algorithms for state space modelling of time series and dynamic systems data. The second objective is the development and implementation of parallel and distributed computational modelling algorithms for such types of multivariate data. State space computational data modelling is presented, commented upon and evaluated against benchmarks. This procedure leads to the expectation of assured feasibility of an original and efficient methodology that will contribute in a direct way to multivariable systems modelling and, both in direct and indirect ways, to the control of multivariable systems.

# PREFÁCIO

Sistemas de grande porte estão sendo construídos pelo homem em diversas áreas. Sua modelagem e ou identificação são cada vez mais necessários e urgentes, particularmente em sistemas energéticos, de comunicações, de transportes, aeroespaciais e ambientais. A modelagem computacional de dados de sistemas dinâmicos lineares invariantes no tempo é um problema central em processamento de sinais e em análise e projeto de sistemas de controle. Resumindo, este problema é equivalente a encontrar uma realização para um sistema dinâmico que represente uma sequência de dados de entrada - saída no espaço de estado. Desta maneira uma realização é uma escolha de uma quádrupla de matrizes que pode representar dados de entrada - saída dentro de um erro aceitável. Mais ainda, isto leva a uma ferramenta adequada para aplicações de engenharia. Com este propósito métodos computacionais foram recentemente desenvolvidos com alto grau de confiança e podem modelar sistemas dinâmicos com dados de entrada-saída com ótima exatidão estatística. Estes métodos são aplicados a uma classe geral de sistemas lineares incluindo sistemas com múltiplas entradas - múltiplas saídas, com ruídos de entrada, ordem desconhecida e dinâmicas instáveis. Em contraste com os métodos existentes para modelagem de dados que acabam sendo imprecisos e exigem um elevado desempenho computacional, os métodos utilizados neste trabalho são fundamentalmente diferentes e permitem a determinação direta dos estados de sistema através de técnicas de subespaços. As técnicas de subespaços envolvem princípios e conceitos que tornam mais simples a modelagem multivariada. Neste trabalho apresentamos uma metodologia usando computação distribuída e paralela para processamento de algoritmos para modelagem computacional de dados no espaço de estado. Este avanço em modelagem de dados tem implicações importantes para análise e projeto de sistemas de controle para muitas aplicações em processos industriais, químicos, espaciais entre outras.

Tradicionalmente, principalmente na área de engenharia elétrica, estuda-se na graduação e na pós-graduação, temas como teoria de sistemas lineares, álgebra linear, representação de estado, probabilidades, processos estocásticos, análise linear, princípios de controle e servomecanismos, métodos

computacionais, dentre outros, que, de certa forma, chamaremos clássicos e que suporemos conhecidos; contudo parece-nos que certos temas e pontos de vista são mais enfatizados e talvez até repetidos excessivamente enquanto que outros ou não são vistos ou são tratados apenas superficialmente; estes são os casos da decomposição em valores singulares e da matriz pseudo inversa dentre outros. Esta tese não tem como objetivo sanar todas as lacunas nem repetir temas frequentemente estudados em sua forma tradicional, contudo ela procura além de dar sua contribuição específica, completar algumas destas lacunas cujo preenchimento é essencial para o atingimento de seus objetivos. Este ponto de vista é fundamental neste trabalho. A matemática aqui tratada e a forma utilizada para tanto não enfatiza o rigor matemático; ela é feita para viabilizar o que consideramos objetivos importantes na visão de engenheiros preocupados com a adequada aplicação da matemática aos problemas de modelagem de sistemas e de séries temporais, com ênfase em aspectos determinísticos, embora aspectos estocásticos também sejam tratados porém com menor ênfase teórica.

Neste trabalho utilizamos álgebra linear, análise numérica, bem como a teoria de sistemas lineares com o enfoque de apresentar contribuições para a modelagem computacional distribuída e paralela de dados. Na Figura i destacamos uma visão possível para a área de atuação da modelagem de dados.

A identificação de sistemas é um campo de modelagem de sistemas dinâmicos através de dados experimentais. Normalmente o objetivo da identificação de sistemas é a utilização do modelo com a finalidade de controlar o processo; deste modo o modelo obtido deve representar de forma adequada a saída do processo; isto pode ser obtido minimizando, através de algum critério, a diferença entre a saída verdadeira e a saída estimada do processo.

Para as aplicações em sistemas de tempo real onde o controle e a identificação são separados em tarefas independentes e cooperantes e que podem ser executadas sobre computadores paralelos e ou distribuídos é possível aumentar o desempenho computacional das tarefas de identificação e controle, possibilitando a utilização eficiente de recursos de *hardware* e *software*, através do emprego de esquemas de fatoração matricial na estrutura algébrica de métodos de estimação de parâmetros, visando a geração de algoritmos paralelos e/ou vetoriais de alto nível. Uma idéia básica pode ser a obtenção de decomposições matriciais baseadas nas transformações ortonormais de Householder e de Givens para paralelizar algoritmos de identificação, Costa Filho, Bottura & Barreto, [222].

Com os comentários anteriores nota-se que esta tese envolve o estudo de três grandes áreas do conhecimento:

- Modelagem de Dados.

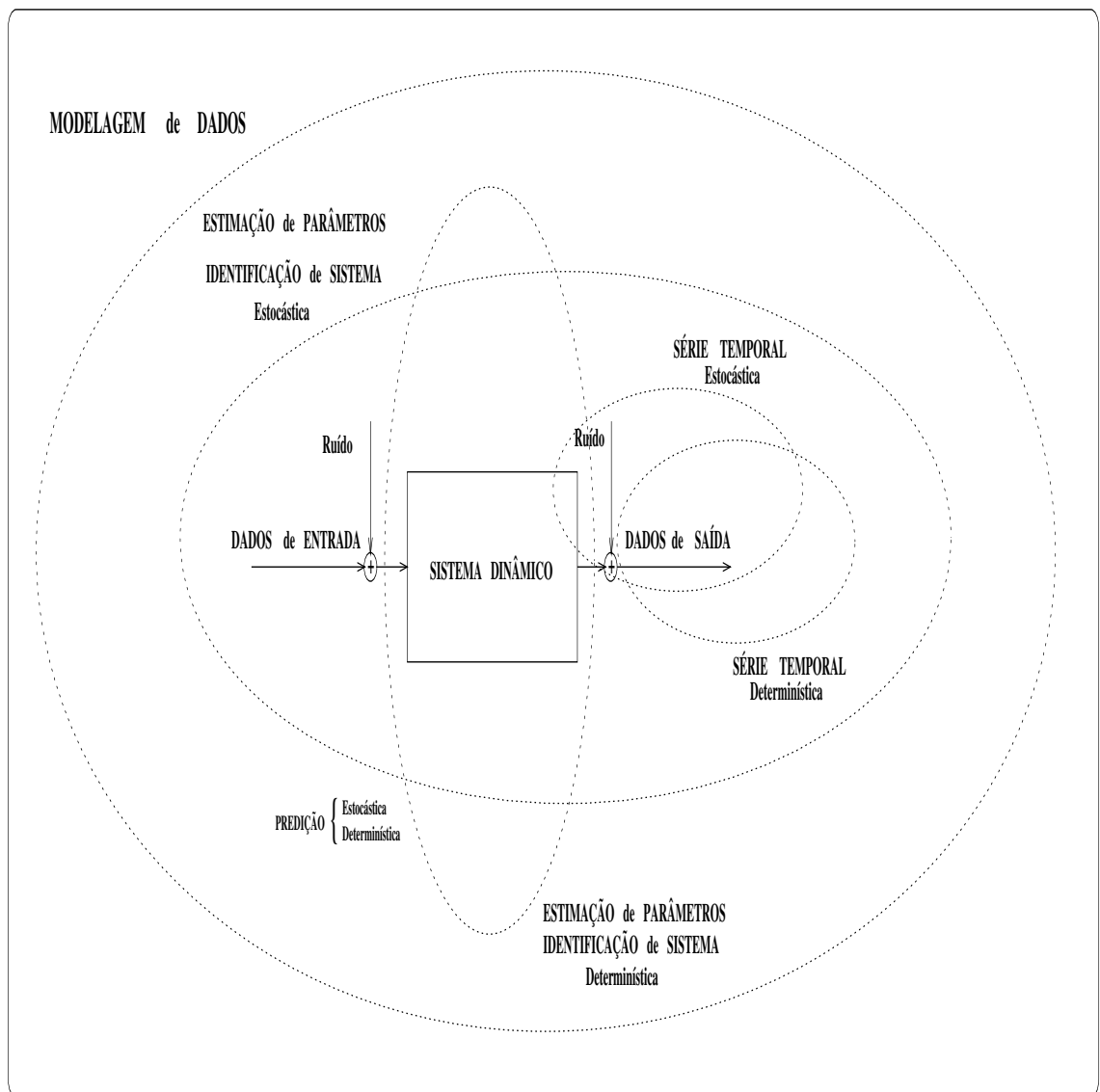


Figura i: Campo de Atuação da Modelagem de Dados.

- Álgebra Linear Computacional.
- Processamento Distribuído e Paralelo.

Na Figura ii apresentamos um mapa organizacional deste trabalho.

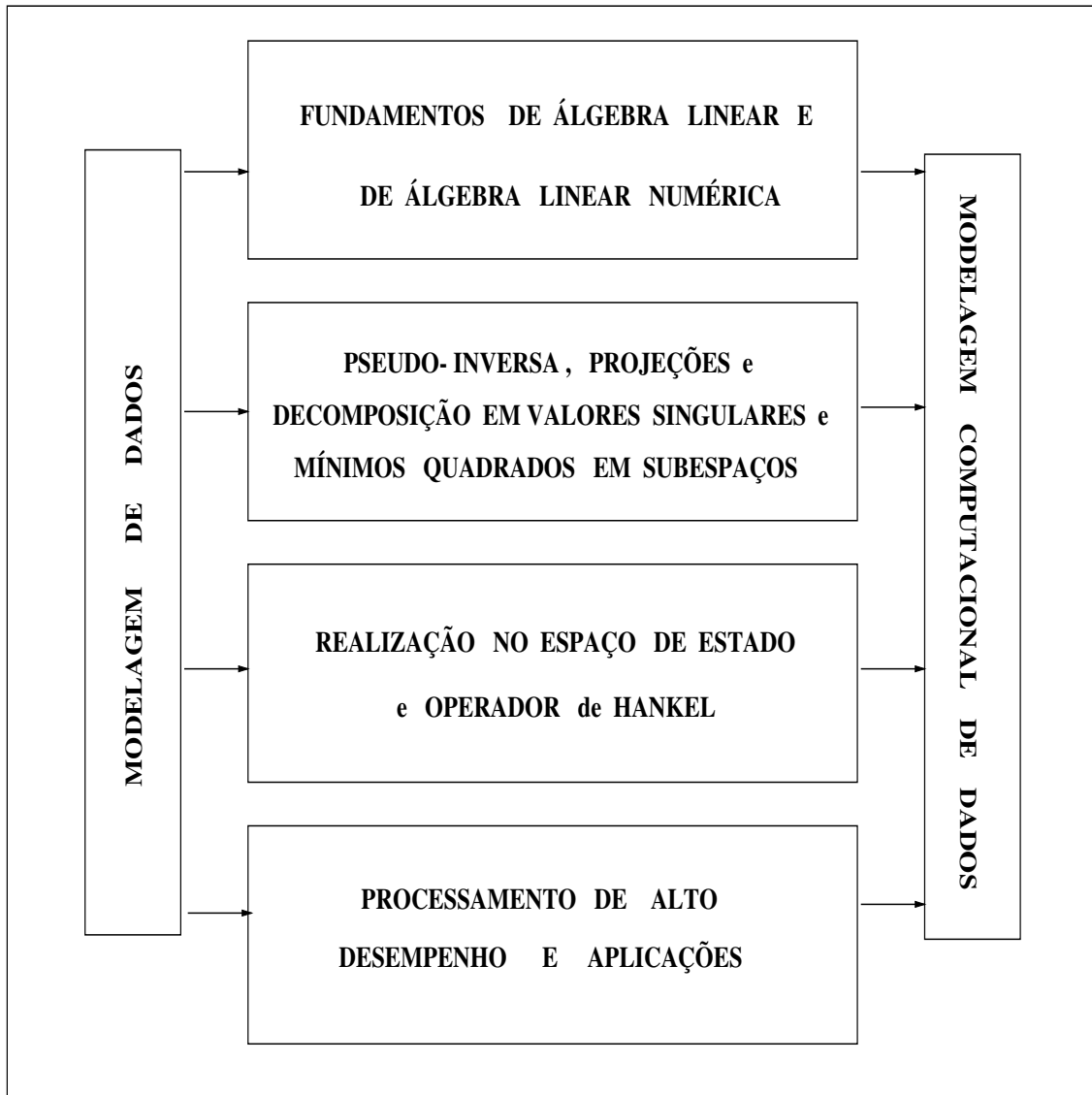


Figura ii: Mapa Organizacional da Tese.

No Capítulo 1 introduzimos conceitos básicos de modelagem de dados, onde damos destaque para a identificação de sistemas, tema tratado em inúmeras referências, das quais destacamos: Amaral [23, 24], Franklin [356], Söderström [989], Ljung [678] e Landau, [636].

A motivação para este estudo é que esta área ainda está completamente aberta para estudos e investigações com relação a computação paralela e distribuída, bem como para um estudo aprofundado dos seus fundamentos teóricos em uma linguagem matemática de alto nível porém adequada à área de engenharia e a outras das matemáticas aplicada e computacional.

Dois aspectos importantes da era computacional são a obtenção de maior velocidade de processamento e a computação em rede. Os projetistas de computadores têm desenvolvido arquiteturas paralelas com o objetivo de obter maiores velocidades através do aumento do número de processadores unitários, bem como redes de computadores e de estações de trabalho que permitem a computação distribuída e paralela, bem como a aquisição distribuída de dados e o controle paralelo e distribuído. Estes desenvolvimentos têm impactos significativos na escolha dos algoritmos e de como eles serão implementados. Uma tendência é buscar maior desempenho utilizando o maior número de unidades de processamento simultaneamente em um mesmo local, outra tendência é utilizar computação distribuída.

No Capítulo 2 fazemos uma revisão de álgebra linear, destacamos transformações ortogonais e abordamos as matrizes de Householder e Givens. As referências que tratam deste tema são inúmeras, mais devemos destacar: Householder [504, 505, 506, 507], Stewart [1000], Strang [1016, 1017], Noble [823], Goldberg [394], Gill [388] Golub & Van Loan [399], Fraleigh [353], Dongarra [304] entre outros. Nosso objetivo neste Capítulo é apresentar algumas noções fundamentais de álgebra linear e de álgebra linear numérica tendo em vista suas aplicações nesta tese, tanto nos seus aspectos teóricos como nos seus aspectos computacionais. Uma parte é tratada neste Capítulo, enquanto que a outra é tratada sobretudo no Capítulo 4.

No Capítulo 3 estuda-se apenas de forma essencial alguns aspectos de processamento distribuído e paralelo dando ênfase aos tipos de topologias existentes e destacando o *hardware* e o *software* para utilização nesta tese. Como referências fundamentais para este tema destacamos: Costa Filho, [219, 220]; Schendel, [945]; Hwang, [512]; Baase, [59]; Obermeier, [827]; Geist, [375, 376] e Van De Velde, [1059].

No Capítulo 4 utilizando-se dos fundamentos de álgebra linear, Capítulo 2, e motivados pela solução do chamado problema de mínimos quadrados, visto no Capítulo 1, elaboramos um estudo personalizado e aprofundado sobre matrizes pseudo-inversas, projeções, decomposição em valores singulares e solução de mínimos quadrados do problema  $\mathbf{Ax} = \mathbf{b}$  com o objetivo de criar estrutura adequada à elaboração de fundamentos para modelagem de dados em subespaços do espaço de estado, que será estudada no Capítulo 5 e muito utilizada nos Capítulos 6 e 7.

No Capítulo 5 apresentamos fundamentos para modelagem multivariável



de dados com base na teoria da realização de sistemas lineares. Acreditamos que este Capítulo constitui uma contribuição importante desta tese, não apenas para modelagem de dados, como para um estudo teórico de sistemas lineares de forma mais avançada e atual.

No Capítulo 6, o problema de identificação de sistemas multivariáveis pelo emprego da teoria da realização de sistemas lineares no espaço de estado é o objetivo central. Através da metodologia aqui abordada problemas de modelagem da dinâmica de sistemas aero-espaciais complexos, de dinâmica de estruturas, de vibrações mecânicas dentre outros, podem ser resolvidos de forma original e computacionalmente avançada. Problemas deste tipo são e serão em número cada vez maior estudados e resolvidos pela teoria de identificação em subespaços do espaço de estado, especialmente quando estiverem envolvidos com a solução de problemas de controle automático de sistemas dinâmicos. Neste Capítulo, uma nova metodologia usando computação distribuída e paralela é apresentada para processar um algoritmo previamente desenvolvido por Verhaegen & Dewilde [1113] para modelagem de sistemas no espaço de estados.

Capítulo 7 tratamos de séries temporais multivariadas. Uma sequência de dados observados de um sinal resultante de um teste sobre um sistema mecânico, por exemplo um veículo, obtidos a intervalos constantes e ordenados temporalmente, constitui uma série temporal monovariada. Este sinal mecânico ou série temporal discreta pode ser processado por razões várias, dentre elas para se obter o seu modelo matemático discreto no tempo, com suas múltiplas utilidades. Raciocínio semelhante nos leva à séries temporais multivariadas obtidas por exemplo de múltiplos testes de uma grande e complexa estrutura aeroespacial. Podemos processar e ou modelar estes sinais multivariados por inúmeras razões seja para realizar regressões, seja para realizar predições, seja para detectar falhas, seja para tomar decisões, seja para seu monitoramento, seja para alguma forma de análise de dados e ou de controle, dentre outras. Quando é possível modelar com acuidade um sinal vetorial torna-se possível executar importantes tarefas de processamento de sinais, dentre elas extrapolações e interpolações.

O problema de modelagem de séries temporais é muito importante nas mais variadas áreas do conhecimento. Gauss dispondo de observações sobre o movimento de um asteróide elaborou o famoso e importante método dos mínimos quadrados. Norbert Wiener, [1155] realizou contribuição importante para a solução do problema de predição de trajetória futura de avião a partir de séries temporais e consequentemente de problemas de estimação e controle associados. Problemas de processamento de sinais e de modelagem de séries temporais estão intimamente relacionados e em muitas situações se confundem entre si. Algoritmos para modelagem de sinais multivariados

discretos no tempo para múltiplas finalidades e aplicações constituem uma área muito fundamental e útil neste campo.

Uma série temporal multivariada constitui um processo estocástico vetorial e modelos destes sinais, que variam discretamente e aleatoriamente no tempo, podem ser tratados no espaço de estado, pelo emprego da teoria de realizações estocásticas de sistemas lineares, Akaike, [8] , Caines [171].

De um ponto de vista metodológico, no Capítulo 7 desenvolvemos e implementamos um procedimento computacional paralelo e distribuído original para a modelagem de séries temporais multivariadas no espaço de estado de um algoritmo originalmente proposto por Aoki, [40], e o aplicamos a um sistema dinâmico que permite comprovar a qualidade e a utilidade do trabalho que realizamos sobre tal algoritmo, bem como abre muitas possibilidades para sua aplicação e para muitas outras pesquisas teóricas e práticas nesta área. O algoritmo de Aoki para modelagem multivariada de séries temporais em subespaços do espaço de estado é abordado em aspectos numéricos e computacionais distribuídos e paralelos. A solução paralela da equação de Ricatti, por ser uma etapa de maior demanda de processamento, é destacada. Resultados de predições de um passo à frente pelo modelo no espaço de estado utilizando Filtro de Kalman para realizações calculadas são apresentados para avaliação da modelagem.

Os algoritmos distribuídos e paralelos propostos nos Capítulos 6 e 7 são contribuições muito relevantes desta tese e as implementações computacionais mostram a praticabilidade destes algoritmos. Algumas conclusões e resultados também são apresentadas nestes Capítulos.

A excepcional coleção de referências desta tese constitui banco bibliográfico que também consideramos uma contribuição deste trabalho. Como o tema modelagem de dados está se expandindo rapidamente, naturalmente esta coleção não é completa, mas é abrangente e atual. Esta bibliografia aborda o tema métodos numéricos para modelagem de dados, bem como inclui referências de programação, álgebra linear, álgebra linear computacional, controle, computação e outros tópicos de interesse científico relacionados com o tema modelagem de dados. Supomos que tais referências coletadas durante vários anos sejam de valor aos pesquisadores da área.

# Capítulo 1

## INTRODUÇÃO À MODELAGEM DE DADOS

### 1.1 Modelagem de Dados

Neste trabalho entendemos por *Modelagem de Dados* a Modelagem de Séries Temporais e a Identificação de Sistemas e consideramos isto suficientemente abrangente para nossos propósitos. O estudo de Séries Temporais tem uma longa história; a área da Identificação de Sistemas, embora mais recente, tem recebido uma grande atenção nas três últimas décadas. Esta área constitui um campo maduro e de grande interesse para os engenheiros de controle.

A modelagem de Séries Temporais e a Identificação de Sistemas podem beneficiar-se dos avanços recentes em Teoria de Sistemas. A Teoria de Sistemas, especialmente no século XX, teve um grande desenvolvimento. No final da década de 1950, a abordagem de Kalman no espaço de estado teve grande impacto sobre ela, Kailath, [566]; Caines, [171]. A modelagem de Séries Temporais, a Identificação, a Análise de Sistemas podem e obtém benefícios desta, Bottura, [140, 143], de diferentes formas, em algum sentido. A computação de alto desempenho é mais recente, e a modelagem no espaço de estado e a computação de séries temporais multivariáveis em todas as áreas estão sendo beneficiadas com estes avanços.

O campo da Teoria de Sistemas Dinâmicos trata problemas de modelagem de dados, predição e controle onde as variáveis interagem com o tempo. A caracterização e modelagem de dados são problemas fundamentais em Teoria de Sistemas. O problema de caracterização está relacionado com o tipo da representação matemática do sistema.

A teoria da realização visando a modelagem de sistemas dinâmicos no espaço de estado e utilizando algoritmos baseados em subespaços, gerados por entradas e saídas passadas, tem sido recentemente estudada e implementada,

sendo uma alternativa em relação aos tradicionais algoritmos de identificação de sistemas que utilizam métodos baseados na predição de erros; Viberg, Wahlberg & Ottersten, [1125]; Van Overschee & De Moor, [1094]; Swindlehurst, Roy, Ottersten, & Kailath, [1026] e Verhaegen & Dewilde, [1113].

A modelagem de dados é um problema fundamental em quase todas disciplinas científicas; particularmente em engenharia e em economia, dados multivariados de entrada - saída, chamados de sinais ou de séries temporais, e sua análise, geralmente servem a pelo menos um de dois possíveis propósitos, dentre outros:

- Modelagem de sinais que necessitam ser reconhecidos ou recuperados por procedimento de análise, como em aplicações em comunicações; ou em predições econômicas de séries temporais.
- Como sinais armazenam informações pertinentes aos sistemas dinâmicos que os produziram, ou a sistemas dinâmicos hipotéticos que os poderiam ter produzido, a análise de sinais proporciona determinar os parâmetros desconhecidos do modelo do sistema.

Os métodos de modelagem de dados baseados em decomposição em valores singulares são classificados como esquemas de subespaços. A decomposição em valores singulares de uma matriz obtida a partir de uma estrutura de Hankel com sinais de dados observados multivariados de entrada,  $\mathbf{u}$ , e de saída,  $\mathbf{y}$ , proporciona o passo chave a uma determinação do modelo no espaço de estado utilizando subespaços que contém informações necessárias e suficientes sobre os parâmetros do sistema.

Tradicionalmente o sistema dinâmico é descrito pelo seguinte conjunto de equações no espaço de estado:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k \end{cases} \quad (1.1)$$

onde os vetores  $\mathbf{u}_k \in \mathbb{R}^m$  e  $\mathbf{y}_k \in \mathbb{R}^l$  são as medidas no  $k$ -ésimo instante de amostragem, respectivamente,  $m$  entradas e  $l$  saídas do processo. O vetor  $\mathbf{x}_k \in \mathbb{R}^n$  é o vetor de estado do processo e contém os  $n$  valores numéricos do estado e as matrizes do sistema são:  $\mathbf{A} \in \mathbb{R}^{n \times n}$ ,  $\mathbf{B} \in \mathbb{R}^{n \times m}$ ,  $\mathbf{C} \in \mathbb{R}^{l \times n}$  e  $\mathbf{D} \in \mathbb{R}^{l \times m}$ .

Para um dado operador de transferência  $\mathbf{T}$  podemos relacionar os dados multivariados de entrada,  $\mathbf{u}$ , e de saída,  $\mathbf{y}$ , tal que:

$$\mathbf{y} = \mathbf{T} \mathbf{u} \quad (1.2)$$

O problema de realização consiste em encontrar uma representação no espaço de estado:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix} \quad (1.3)$$

tal que as representações das equações (1.2) e (1.3) sejam equivalentes.

O problema investigado neste trabalho é o de determinar modelo no espaço de estado, incluindo sua ordem e a quádrupla de matrizes,  $(\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D})$  a partir de algoritmo que explora as propriedades de subespaços para os dados fornecidos por sistemas e/ou séries temporais multivariáveis.

Em princípio irá existir um número infinito de realizações no espaço de estado para modelar um conjunto de dados; portanto nós iremos limitar nosso trabalho para realizações mínimas no espaço de estado, isto é modelos para os quais a dimensão do estado descreva o subsistema observável e atingível e seja igual a  $n$ . Mesmo para realizações mínimas, a representação de estado também não é única e uma representação de sistema equivalente – produz a mesma relação entrada - saída – é obtida aplicando a transformação de estado  $\mathbf{S}$  (uma matriz inversível  $n \times n$ ) para um novo vetor de estado definido como:  $\bar{\mathbf{x}}_k = \mathbf{S}\mathbf{x}_k$ . Os autovalores da matriz  $\mathbf{A}$  obtidos através desta transformação são idênticos aos da matriz  $\mathbf{A}$  anterior desde que  $\mathbf{S}^{-1}\mathbf{A}\mathbf{S}$  seja uma transformação de similaridade.

## 1.2 Sistema Dinâmico

Um sistema dinâmico pode ser conceitualmente descrito como na Figura 1.1. O usuário pode controlar o sinal de entrada,  $\mathbf{u}_k$ , mas não a perturbação,  $\mathbf{e}_k$ . Em algumas aplicações,  $\mathbf{u}_k$  e  $\mathbf{e}_k$  podem ser confundidos. As medidas dos sinais de saída e de entrada fornecem informações úteis sobre o sistema não conhecido. Neste trabalho estamos interessados na modelagem de dados entrada - saída de sistemas e de séries temporais discretas. Evidentemente uma série temporal contínua, por exemplo, pode ser transformada em discreta através de amostragem em intervalos de tempo equiespaçados. O tratamento adequado de dados nos fornece modelos com capacidade de predição de valores futuros da série sob consideração.

Um problema significativo em teoria de sistemas é a determinação da dinâmica deste sistema através dos dados de entrada e de saída que foram anteriormente medidos. O comportamento de um sistema dinâmico pode ser descrito convenientemente e sucintamente pela introdução de noções de estados e vetores de estado.

O estado de um sistema no  $k$ -ésimo instante é a menor coleção de números, também chamados de variáveis de estado, que deve ser especificada naquele instante para que se possa determinar de forma única o comportamento do

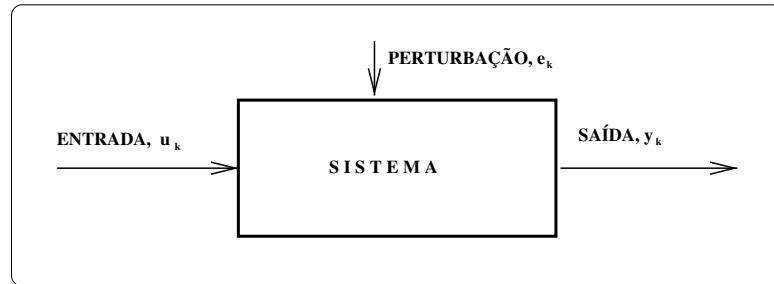


Figura 1.1: Um sistema dinâmico determinístico onde as setas representam os vetores de sinais de entradas  $\mathbf{u}_k$ , saídas  $\mathbf{y}_k$  e distúrbios  $\mathbf{e}_k$  no  $k$ -ésimo instante.

sistema para qualquer instante, maior ou igual ao tempo inicial, para qualquer entrada pertencente ao conjunto de entradas, Kalman [573]. Se esta coleção de números for ordenada em um vetor, este vetor receberá o nome de vetor de estado. O vetor de estado de um sistema determinístico é a mínima coleção de informação necessária para determinar de forma única a evolução futura de um sistema dinâmico, mostrando a evolução das variáveis que afetam o sistema (variáveis de controle ou decisão). O espaço de estado será formado por  $n$  vetores de estado, formando um espaço  $n$ -dimensional, e qualquer estado do sistema poderá ser representado por um ponto no espaço de estado. Quando um espaço de estado de dimensão finita não for suficiente para capturar o comportamento do sistema então a dinâmica terá dimensão infinita. Nesta tese apenas a dimensão finita e modelos discretos no espaço de estado são tratados.

### 1.2.1 Conceitos Sobre Modelagem de Dados

Neste trabalho usaremos definições que na sua maioria podem ser encontradas em referências tais como Eykhoff [320], Harvey [461, 462], Isermann [519], Ljung [678], Verhaegen e Dewilde [1113] e Young [1185]. A seguir apresentamos algumas destas definições:

- **Coletas de dados central e distribuída**

O controle inteligente ou o controle adaptativo de um processo sujeito a grande variações paramétricas será consideravelmente simplificado se um método para identificação rápida do modelo matemático do processo puder ser implantado. Normalmente, tais métodos exigem uma considerável carga computacional, em adição ao fato de que é necessário esperar por uma acumulação de um conjunto de dados à disposição. Apesar da existência de algoritmos numéricos paralelos, por exemplo, para identificação recursiva, Ljung e Söderström [678], que possibilitam iniciar uma estimativa de parâmetros do modelo sempre que um primeiro conjunto de dados esteja disponível, e do emprego da computação digital em sistemas de controle, a implantação prática de tais algoritmos em uma estrutura conectada, (*on-line*), para identificação de sistemas, em tempo crítico e determinado exige um grande desempenho computacional. Este desempenho pode ser obtido a um baixo custo com a utilização de sistemas de processamento distribuído e paralelo, Leonhardt [661], Jainandunsing [522].

Na Figura 1.2 apresentamos diversas situações para a coleta de dados, resultando em possíveis estruturas de identificação.

Sobre as diversas situações apresentadas na Figura 1.2 comentamos:

- I - Neste caso a coleta de dados de entrada e saída de um processo é feita manualmente e depois eles são inseridos em um computador. Após o processamento destes dados e de posse dos resultados obtidos o homem irá atuar no processo. A linha tracejada mostra que o fluxo de informações é feito manualmente (estrutura não conectada central).
- II - A coleta de dados é feita sistematicamente sem a presença do homem. Por exemplo poderia existir um instrumento com registrador para efetuar a amostragem. Tanto neste caso como no anterior o computador só atua quando da entrada manual de dados do processo. Pela definição ambas estruturas são não conectadas, embora centralizadas.
- III - Neste exemplo o operador entra com os dados do processo diretamente no computador, sendo que a atuação ou não no processo se dará manualmente após a obtenção dos resultados. É evidente, em todas as situações abordadas até o momento que uma entrada incorreta de dados poderá levar a situações não previstas.
- IV - Medidores com interfaces analógicas - digitais estão conectados ao computador. A amostragem do processo neste caso poderá ocorrer

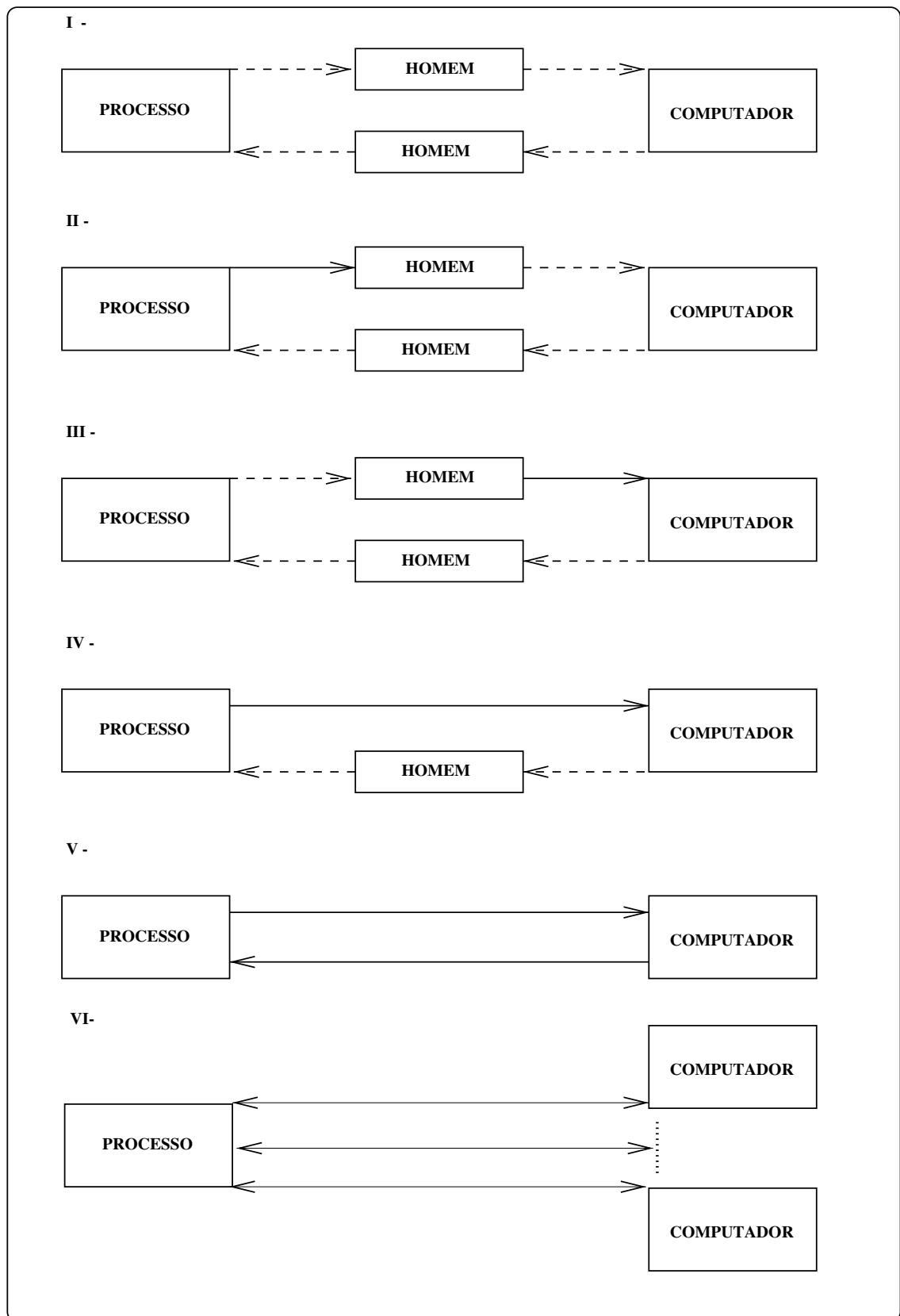


Figura 1.2: Possíveis Estruturas de Identificação e Controle de Processo.



de modo rápido, frequente, exato e preciso. Os resultados gerados serão utilizados para a gerência manual do processo.

- V - Como no caso anterior a coleta e tratamento de dados são efetuados através de uma conexão direta entre o processo e o computador. Os resultados obtidos (por exemplo, parâmetros estimados) serão utilizados pelo computador para gerenciar o processo sem interferência do homem. Temos uma estrutura conectada, porém centralizada.
- VI - Diferentemente do caso anterior a coleta e tratamento de dados são efetuados através de uma rede de computadores. Temos uma estrutura conectada, porém distribuída.

As situações anteriores de certo modo retratam o desenvolvimento da engenharia de controle nas últimas décadas. Contudo, sobretudo na última década, redes de computadores foram desenvolvidas e implantadas em grande quantidade, e esta tendência é dominante. Em decorrência, estrutura conectada e distribuída para coleta de dados e consequente modelagem computacional de dados, é possível, e desejável em muitas situações, desde que existam algoritmos adequados para modelagem computacional distribuída de dados, que por ser distribuída, também é paralela, Figura 1.2, item VI.

A proposta dos algoritmos desenvolvidos tem como objetivo auxiliar a execução de tarefas a serem implantadas nas Situações V-VI, não descartando a possibilidade da sua utilização nas situações anteriores.

- **Conjunto de dados:** Uma seleção de informações que podem ser obtidas e transmitidas com características tais como: quantidade analógica cuja amostra tenha um significado específico que possa ser determinado. A modelagem de dados de um sistema terá sucesso quando a coleção do conjunto de dados incluir informações significativas sobre o sistema. Este fato requer que a aquisição do conjunto de dados seja bem planejada. Então algumas decisões sobre este fato devem ser tomadas:
  - Quais sinais no processo devem ser medidos?
  - Como as entradas devem ser escolhidas?
  - Qual intervalo de amostragem deve ser usado?
  - Quantos dados necessitam ser coletados?

As questões seguintes referem-se ao pós-tratamento do conjunto de dados e também estão relacionadas a qualidade do conjunto de dados

- Qual a frequência de amostragem adequada?
- Qual o critério para remover um dado do conjunto de dados?
- Existe a necessidade de filtrar o conjunto de dados?
- Existe a necessidade de remover frequências de ruído?
- Todo o conjunto de dados é confiável?
- A sequência do conjunto de dados deve ser fracionada? (isso é, a amostragem deve ser efetuada em uma frequência mais baixa?)

Respostas adequadas a todas estas perguntas exigem uma amplo tratado que extrapola o âmbito deste trabalho.

- **Série Temporal:** Em alguns sistemas não temos acesso a nenhum sinal de entrada, então um conjunto de dados de saída ordenado no tempo é chamado de série temporal. As séries podem ser contínuas ou discretas dependendo do tipo de amostragem efetuada. Consideraremos apenas as discretas. Neste caso as observações são tomadas em instantes de tempo especificados, normalmente a intervalos iguais, e o problema de modelagem de dados é o de descrever matematicamente as propriedades deste sinal.

Uma possibilidade de realizar este tratamento é através da decomposição desta série temporal,  $\mathbf{y}_k$ , em seus movimentos básicos tais como: tendência, T ; movimento cíclico, C ; movimento sazonal, S e ou movimento irregular ou aleatório, I , então  $\mathbf{y}_k$  pode ser decomposta como:

$$\mathbf{y}_k = (T + C + S + I)_k \quad \text{ou} \quad \mathbf{y}_k = (TCSI)_k$$

A tendência pode ser estimada, por exemplo, utilizando um gráfico onde plota-se os pontos e traça-se uma curva ou o método dos mínimos quadrados ou o método das semi-médias onde separamos os dados em duas partes e calcula-se a média para cada parte. Uma linha de tendência é traçada entre os dois pontos. A escolha do método utilizado terá uma eficiência adequada ou não para o tipo de resultado desejado.

- **Sistema :** Um conjunto integrado de uma ou diversas estruturas que interagem entre si gerando um conjunto de dados possibilitando uma relação com um outro conjunto de dados. Tal representação foi apresentada de modo genérico na Figura 1.1, e pode ser apresentada de modo mais específico através da Figura 1.3, que representa um sistema multivariável, ou sistema MIMO (*Multiple Input - Multiple Output*)



Figura 1.3: Sistema Multivariável.

como mostrado na Figura 1.3, ou seja um sistema de múltiplas entradas e múltiplas saídas. O sistema com uma única entrada e saída é conhecido com SISO (*Single Input - Single Output*).

- **Modelo:** Uma representação compacta dos aspectos essenciais de um sistema existente que descreverá como ele em atividade irá se comportar. Os modelos paramétricos precisam ser quantificados.
- **Modelo Matemático:** Uma representação matemática com aspectos essenciais de um sistema existente, apresentando conhecimento do sistema em forma compacta.

Uma forma de representação de sistemas multivariáveis, lineares, invariante no tempo é a no espaço de estado, dada por:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \end{cases} \quad (1.4)$$

onde temos que  $\mathbf{u}_k \in \mathcal{R}^m$ ,  $\mathbf{y}_k \in \mathcal{R}^l$  e  $\mathbf{x}_k \in \mathcal{R}^n$  representam o vetor de entrada, o vetor de saída e o vetor de estado respectivamente, no  $k$ -ésimo instante de tempo. Sabemos que somente podemos fazer uso dos dados de entrada  $\mathbf{u}_k, \mathbf{u}_{k+1}, \dots$  e de saída  $\mathbf{y}_k, \mathbf{y}_{k+1}, \dots$  e como as matrizes  $\mathbf{A}_{n \times n}$ ,  $\mathbf{B}_{n \times m}$ ,  $\mathbf{C}_{l \times n}$  e  $\mathbf{D}_{l \times m}$  não são conhecidas será um dos objetivos deste trabalho mostrar um procedimento eficiente para encontrá-las. Estas matrizes podem também ser variáveis no tempo.  $\mathbf{w}_k$  e  $\mathbf{v}_k$  são sinais de ruído que podem levar em consideração ruídos

nas medidas e no processo, entre outros. Os ruídos ou perturbações originários de um processo podem ser semelhantes aos resíduos de cálculo, de um conjunto de equações. Caso sejam omitidos para simplificação de cálculo, o conjunto de equações que representa um sistema multivariável é dado por:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k \end{cases} \quad (1.5)$$

Além disso podemos supor que a representação do modelo é mínima, Kailath [566], isto é, o sistema é completamente observável e atingível. Caso o sistema não tenha entradas medidas teremos uma série temporal e o modelo no espaço de estado será:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \end{cases} \quad (1.6)$$

A Figura 1.4 mostra três diferentes tipos de modelagem para tratamento de dados multivariados com seus respectivos modelos no espaço de estado, onde  $q$  é o operador deslocamento para trás:  $q\mathbf{x}_{k+1} = \mathbf{x}_k$ .

- **Conjunto de Modelos:** Uma coleção de modelos que representam de maneira equivalente um sistema.
- **Dimensão de Um Modelo:** O número mínimo de componentes do vetor de estado que o descreve.
- **Processo:** A parte principal de um sistema físico que será modelado.
- **Erro:** Um conjunto de dados que representa a diferença entre o conjunto de dados obtidos através do processo e o correspondente conjunto de dados gerados pelo modelo selecionado para representar este processo.
- **Método de Estimação:** Uma função matemática aplicada ao erro, ou a um conjunto de dados usados para gerar o erro, permitindo que este assuma alguma característica que permita a seleção de um modelo do conjunto de modelos, Figura 1.5.
- **Intervalo de Estimação:** Tamanho do conjunto de dados usado para a estimação de parâmetros do modelo.
- **Função Objetivo ou Critério:** Função matemática do erro, baseada no método de estimação selecionado, que quantifica os possíveis modelos para um conjunto de modelos.

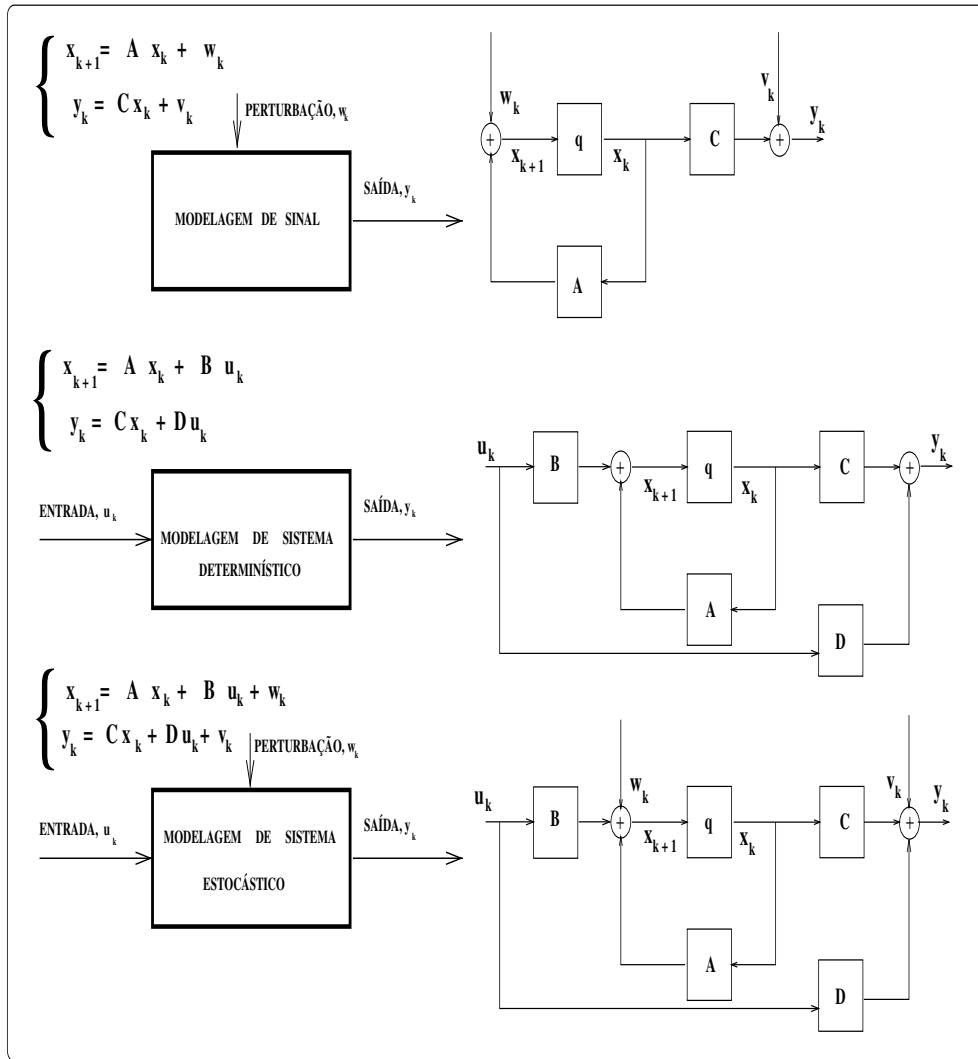


Figura 1.4: Tipos de Modelagem de Dados no Espaço de Estado.

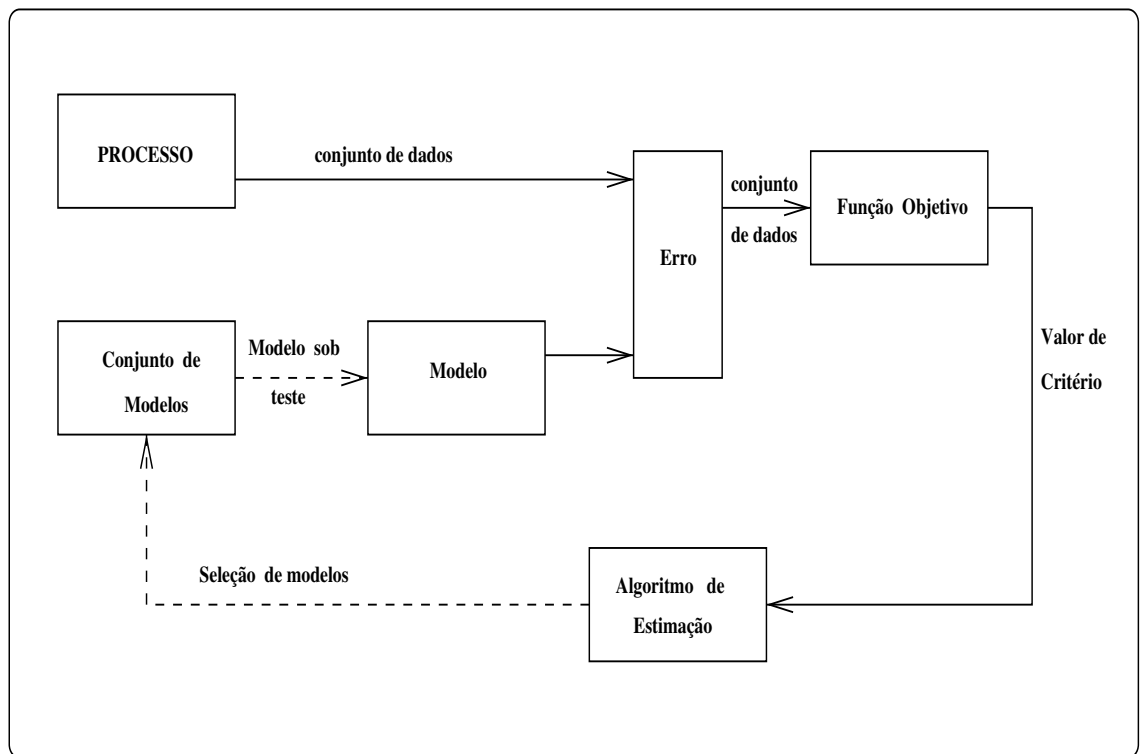


Figura 1.5: Método de Estimação.

- **Algoritmo de Estimação:** Um método computacional usado para determinar um modelo, dentro de um conjunto específico de modelos, com base na seleção dos tipos de erro, na seleção do método de estimação, bem como no conjunto de dados e informações *a-priori* obtidas do processo que foi representado por um modelo.
- **Identificação:** O procedimento usado para determinação, por meio de um algoritmo de estimação, de um modelo matemático de um processo, tendo como orientação o conhecimento de engenharia experimental disponível, das leis químicas e físicas e informações experimentais do processo, o que permite a elaboração de um conjunto de modelos, de um tipo de erro, de um método de estimação e a preparação do conjunto de dados obtidos através do processo.

Para obter os parâmetros de um modelo baseado nos sinais que contém as informações relevantes do comportamento de um processo dinâmico operando em um ponto determinado (funcionando sob certas restrições) a princípio quatro escolhas precisam ser feitas:

- seleção do conjunto de modelos e seus parâmetros.

- seleção do tipo de erro usado para estimar os parâmetros do modelo.
- seleção do método de estimação usado para determinar os parâmetros do modelo.
- seleção do critério de validação do modelo.

Caso os parâmetros estimados do modelo escolhido representem o sistema de maneira satisfatória, o processo de identificação poderá ser finalizado ou atualizado a medida que novos dados estiverem sendo gerados.

### Exemplo

Considere um tanque de mistura como na Figura 1.6 onde 2 fluídos são misturados. A concentração de cada fluído pode ser variada. Os fluxos  $F_1$  e  $F_2$  podem ser controlados com válvulas. Os sinais  $F_1(k)$  e  $F_2(k)$  são entradas do sistema. Os fluxos de saída  $F(k)$  e a concentração  $c(k)$  do tanque constituem variáveis de saída. As concentrações  $c_1(k)$  e  $c_2(k)$  não podem ser controladas e são consideradas perturbações.

Suponha que se deseja projetar um regulador que atue em  $F_1(k)$  e  $F_2(k)$  usando as medidas  $F(k)$  e a concentração  $c(k)$ . O regulador irá assegurar que  $F(k)$  ou a concentração  $c(k)$  permaneçam constantes com as variações de  $c_1(k)$  e  $c_2(k)$ . Para tal projeto é necessário algum modelo matemático que descreva a relação das concentrações de entrada e de saída.

Muitos processos industriais, por exemplo para produção de papel, ferro, vidro ou compostos químicos, poderão ser facilmente e eficientemente controlados. De um modo geral a modelagem é utilizada em muitos ramos da ciência para descrever e entender a realidade.

Alguns sistemas reais podem ainda não existir mas modelos da sua dinâmica poderão ser de interesse quando de sua existência no futuro. Por exemplo, a dinâmica de uma estrutura espacial pode não ser deduzível através de ensaios aqui na Terra, devido a gravidade e efeitos atmosféricos. Neste caso a modelagem baseada em teoria trará o conhecimento que a modelagem de dados experimentais ainda não obtidos não permitiria.

Quando do projeto de um sistema de controle para o conhecimento do modelo da dinâmica do processo ou sistema a controlar pode-se utilizar informações das seguintes fontes:

#### 1. Leis Físicas, Químicas, Biológicas, Econômicas

Por exemplo: Equações de movimento, circuitos elétricos, dinâmica dos fluídos.

A construção de um modelo de um processo pode ser obtida através de leis químicas e físicas e suposições de que o comportamento deste

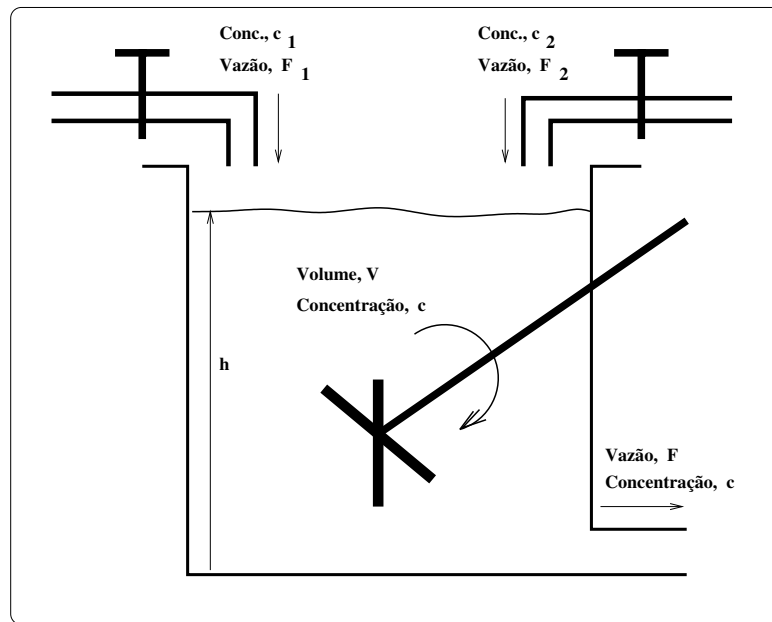


Figura 1.6: Tanque de Mistura.

processo é representado satisfatoriamente por estas leis. Este método é muito usado para a modelagem de sistemas mecânicos (por exemplo, trajetória de satélites, aeronaves, robôs), sistemas elétricos (por exemplo, sistema de transmissão de energia elétrica, circuito eletrônico) e processos químicos (por exemplo, reações químicas). Em geral, se estas leis possibilitam a criação de um modelo que atenda determinadas necessidades, por exemplo para o controle de um determinado processo, este modelo é satisfatório. Em situações onde um processo industrial complexo precisa ser modelado, a abordagem através das leis citadas acima conduz a modelos complexos com muitas variáveis desconhecidas. Devido a complexidade dos modelos obtidos através das leis físicas e químicas torna-se difícil encontrar relações apropriadas para estimar grandezas não conhecidas. Os modelos obtidos desta maneira geralmente apresentam não linearidades e são de dimensão elevada o que torna este método muito trabalhoso quando da sua aplicação em diversos processos industriais, sendo que frequentemente não conduzem a resultados completamente satisfatórios. Além disso, modelos baseados em leis físicas e químicas podem nos dar uma complexa descrição do processo onde muitas vezes podemos somente estar interessados nos dados de entrada e saída deste processo.

Um exemplo é a modelagem de comportamento de um circuito  $RL$  alimentado por uma fonte de tensão e a aplicação da lei das tensões



de Kirchhoff para o circuito permite obter o valor da queda de tensão dado por:

$$v_R + v_L = e$$

esta equação física pode ser reescrita como:

$$Ri + L \frac{di}{dt} = e$$

Considerando a corrente como a variável de estado do sistema então:

$$\dot{\mathbf{x}} = -\frac{R}{L}\mathbf{x} + \frac{1}{L}e$$

Esta equação descreve o comportamento do circuito  $RL$  no espaço de estado unidimensional, supondo conhecidos  $R$  e  $L$ . Neste procedimento não tivemos necessidade de obter um conjunto de dados experimentais para caracterizar o sistema, mas este tipo de procedimento nem sempre é possível.

## 2. Experimentos

Por exemplo: Através da obtenção de dados, pela excitação do processo a controlar.

Esta técnica de modelagem é baseada nas observações do comportamento do processo em diferentes pontos de operação. Este tipo de modelagem é conhecido como técnica da “caixa-preta”. Backx, [60], apresenta um tipo de modelagem denominada de “caixa - cinza”, devido ao fato de que para a elaboração do modelo do processo além dos dados de entrada e saída são utilizadas as leis físicas e químicas envolvidas. As leis auxiliam na escolha do modelo a ser adotado para representar um dado processo. Nas Figuras 1.7 e 1.8 ilustra-se representações para propostas de modelagem e/ou controle respectivamente centralizado e distribuído e paralelo.

A modelagem de dados usada para a determinação do modelo de um sistema representando seus aspectos essenciais, a partir dos dados de entrada e de saída como a fonte de informações recebe o nome de identificação de sistemas. Então, identificação de sistemas é uma abordagem experimental para modelagem de processos. Ela pode tanto ser centralizada como distribuída e paralela. Na identificação de sistemas estão as seguintes etapas de abordagem:

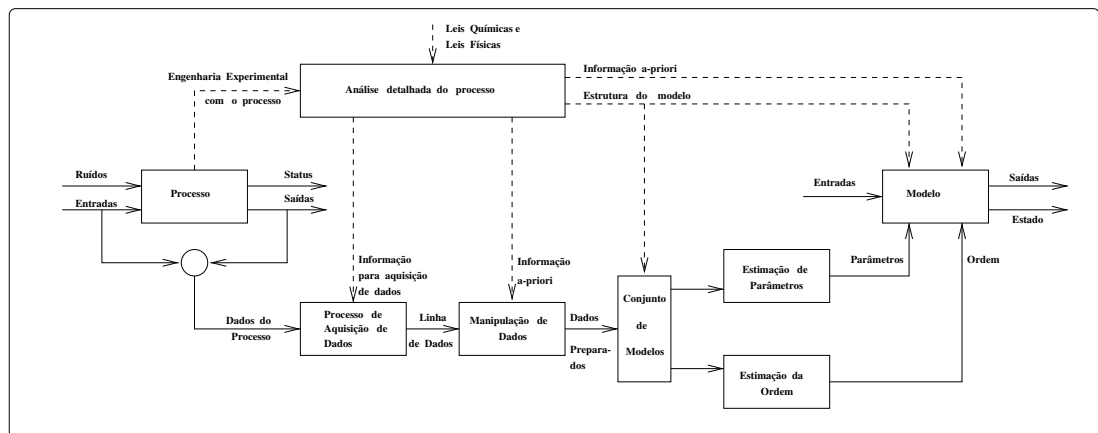


Figura 1.7: Proposta Para Modelamento Centralizado de um Processo.

- Planejamento experimental.
- Validação de Dados.
- Planejamento computacional.
- Seleção da estruturas do modelo.
- Estimação de parâmetros.
- Validação do modelo.

Na prática o procedimento de identificação de sistemas é iterativo. Quando investigamos um processo cujo conhecimento prévio não é suficiente para modelar o processo com transitórios e distúrbios, podemos realizar uma estimativa grosseira da dinâmica deste processo. Então, os resultados obtidos podem ser usados para estimar parâmetros não conhecidos do modelo. Baseado nestes resultados, a estrutura do modelo pode ser aperfeiçoada e novos experimentos poderão ser realizados.

Este procedimento esta ilustrado na Figura 1.9.

### 1.2.2 Tipos de Identificação de Sistemas

A teoria para identificação de sistemas pode ser dividida em:

- Paramétrica
- Não Paramétrica

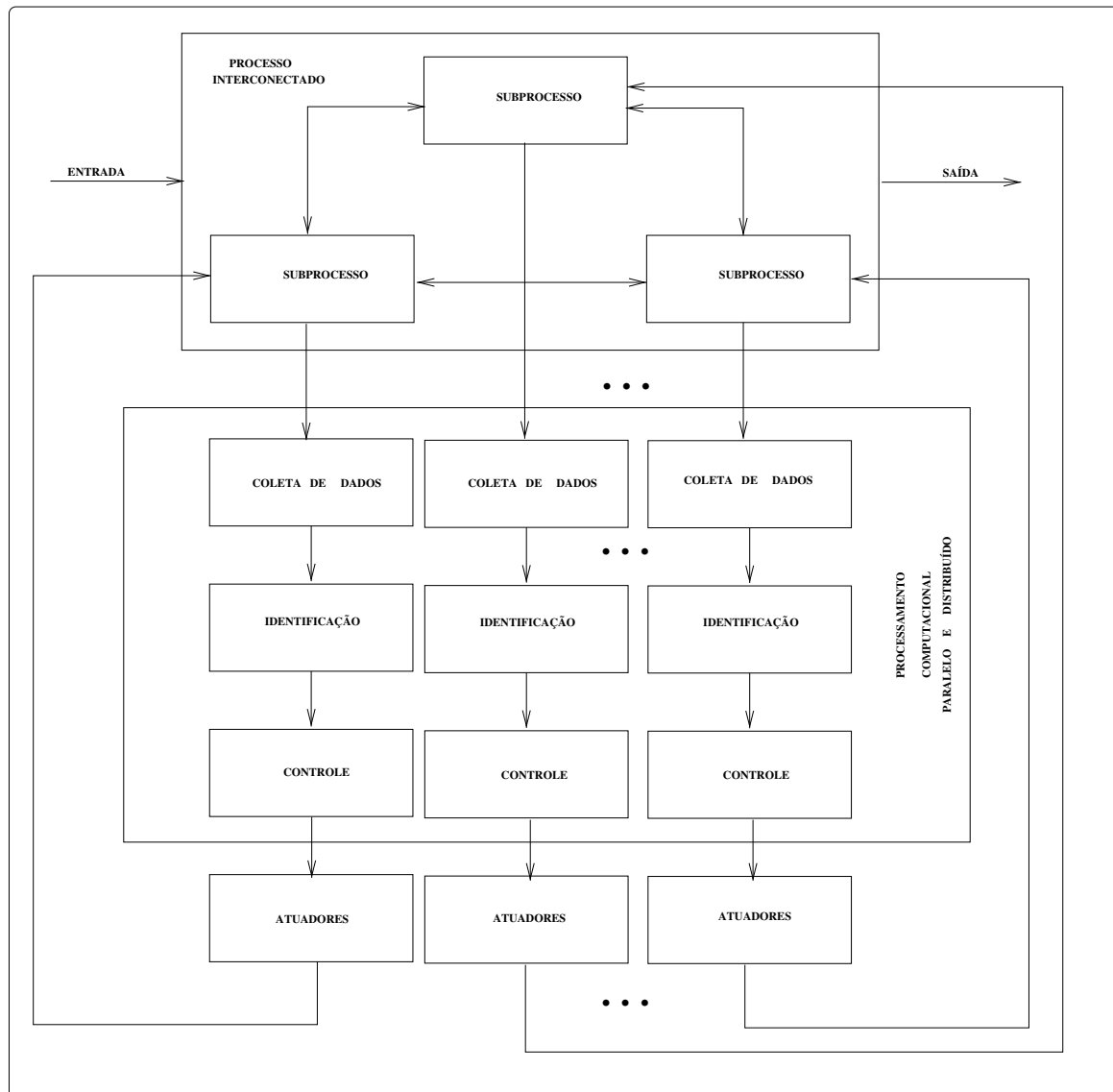


Figura 1.8: Possível Sistema de Controle Distribuído e Paralelo de Um Processo.

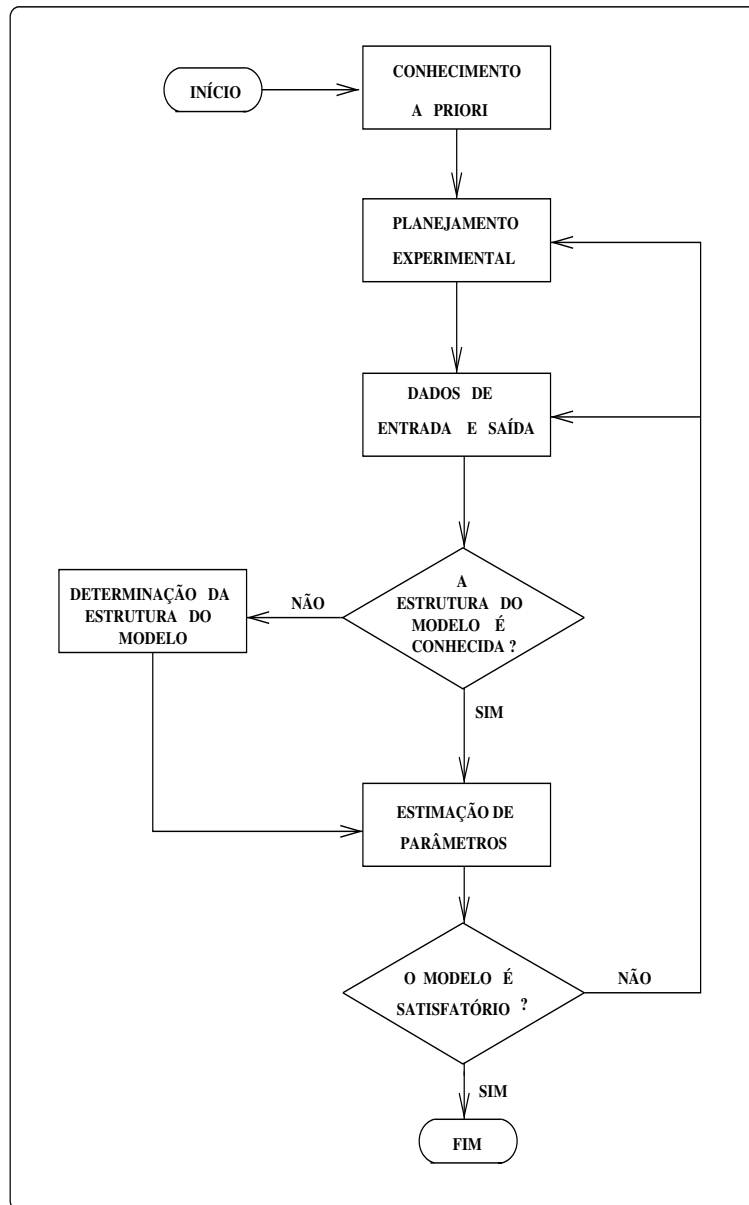


Figura 1.9: Um Provável Diagrama Para Identificação Centralizada de Sistemas.

Em modelos paramétricos a representação do sistema se dá, através de funções de transferência ou de variáveis de estado, na forma matricial, sendo portanto representada por números.

Em modelos não paramétricos a representação do sistema se dá através de curvas de resposta em frequência, por exemplo. Não existe um número finito de parâmetros.

Neste trabalho nos limitaremos a estudar os modelos paramétricos.

### 1.2.3 Critérios Para Escolha do Estimador

Um objetivo para a identificação de sistemas é o de obter um modelo matemático que auxilie no projeto do controlador do processo; deste modo o modelo obtido deve representar de forma adequada o processo; isto pode ser obtido, por exemplo, minimizando, através de algum critério, a diferença entre a saída verdadeira e a saída estimada do processo. A seguir veremos alguns destes critérios.

#### Erro de equação

Seja um processo discreto no tempo descrito por:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k; \Theta)^1 \quad (1.7)$$

É suposto que o vetor  $\mathbf{f}$  é conhecido, mas que não temos os parâmetros  $\Theta$ , que descrevem o processo. Também é suposto que é possível medir tanto o controle  $\mathbf{u}_k$ , como os estados,  $\mathbf{x}_k$  e  $\mathbf{x}_{k+1}$ .

Se temos  $\mathbf{x}_k$ ,  $\mathbf{x}_{k+1}$  e  $\mathbf{u}_k$  para um valor especificado de  $\Theta$ , então:

$$\mathbf{x}_{k+1} - \mathbf{f}(\mathbf{x}_k, \mathbf{u}_k, \Theta) = \varepsilon(k; \Theta) \quad (1.8)$$

onde  $\varepsilon(k; \Theta) = 0$  quando  $\Theta$  tiver idealmente os valores dos parâmetros que exatamente representam o processo. O vetor  $\varepsilon(k; \Theta)$  representa o **erro de equação**. O problema de identificação poderá ser resolvido através da determinação de  $\hat{\Theta}$  que minimiza o critério:

$$J(\Theta) = \sum_{k=1}^N \varepsilon^T(k; \Theta) \varepsilon(k; \Theta) \quad (1.9)$$

então  $\Theta = \hat{\Theta}$  será o vetor de parâmetros que melhor representa o processo.

Este critério minimiza o erro quadrático.

---

<sup>1</sup>Convenciona-se separar por “;” as grandezas de diferentes ordens em relação as grandezas inicialmente citadas e separadas por “,”.

A Figura 1.10 mostra como o erro de equação pode ser gerado. Neste exemplo, cada parte do modelo trata separadamente os sinais de entrada e de saída de um sistema como ilustrado, [60].

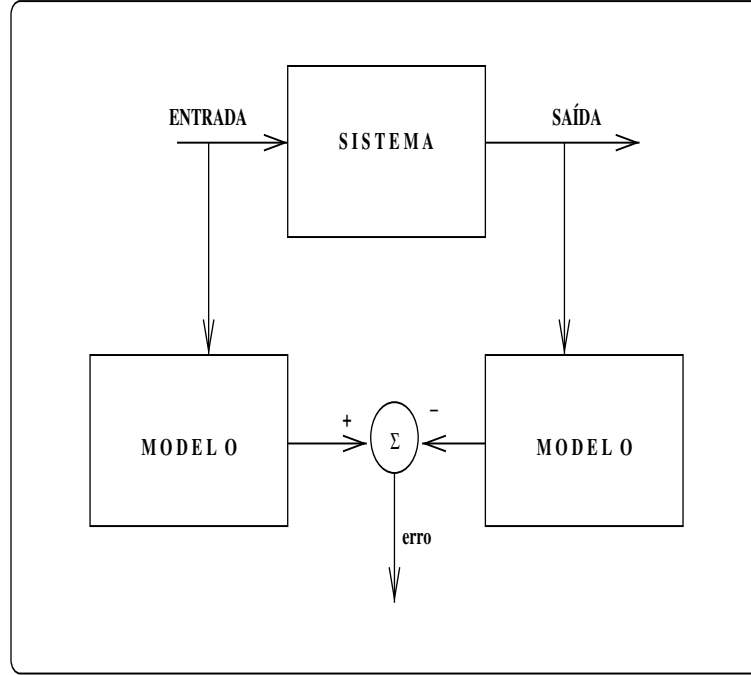


Figura 1.10: Erro de Equação.

### Erro de Saída

O conhecimento de todo o vetor de estado de um sistema pode não ser elementar; um critério onde não necessitamos do estado do sistema para obtenção dos parâmetros é o **erro de saída**, Figura 1.11.

Neste caso o estimador é usado para gerar a saída do modelo.

O critério a minimizar para um sistema monovariável a ser identificado é:

$$\min_{\Theta} J(\Theta) = \sum_{k=1}^N e^2(k, \Theta) \quad (1.10)$$

Por exemplo, para modelo ARMA (auto regressivo média móvel) monovariável temos:

$$\begin{aligned} \mathbf{y}_m(k) = & a_1 \mathbf{y}_m(k-1) + a_2 \mathbf{y}_m(k-2) + \dots + a_m \mathbf{y}_m(k-m) + \\ & b_1 \mathbf{u}(k-1) + b_2 \mathbf{u}(k-2) + \dots + b_m \mathbf{u}(k-m) \end{aligned} \quad (1.11)$$

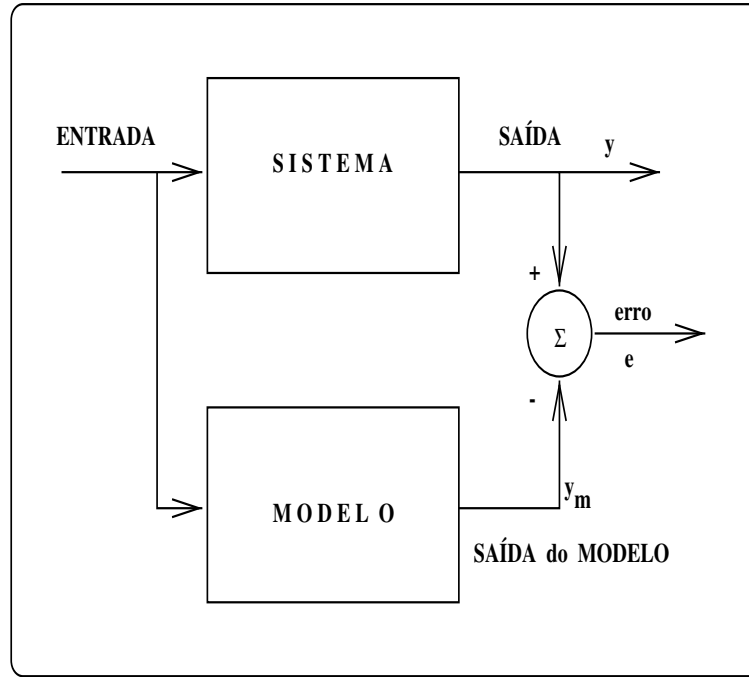


Figura 1.11: Erro de Saída.

onde  $y_m$  é o valor da saída do modelo e  $u$  é a entrada do processo. Se  $y(k)$  é o valor real da saída do processo ao erro dado por:

$$e(k, \Theta) = y(k) - y_m(k) \quad (1.12)$$

chamamos erro de saída:

$$\begin{aligned} e(k, \Theta) &= y_k - y_m(k) = \\ &= y_k - a_1 y_m(k-1) + a_2 y_m(k-2) + \dots + a_m y_m(k-n) + \\ &\quad b_1 u(k-1) + b_2 u(k-2) + \dots + b_m u(k-m) \end{aligned} \quad (1.13)$$

onde  $y_k$  é a saída medida e  $y_m(k)$  é a saída calculada.

Portanto a diferença entre os dois critérios apresentados é que o primeiro usa valores passados das medidas reais enquanto que o segundo usa valores passados das medidas da saída do modelo.

### Erro de Predição de Saída

Uma alternativa mais sensata ao modelo anterior, dado que o erro não pode ser controlado, e eventualmente teremos um sinal de ruído é utilizar o critério

de erro de predição de saída, Figura 1.12.

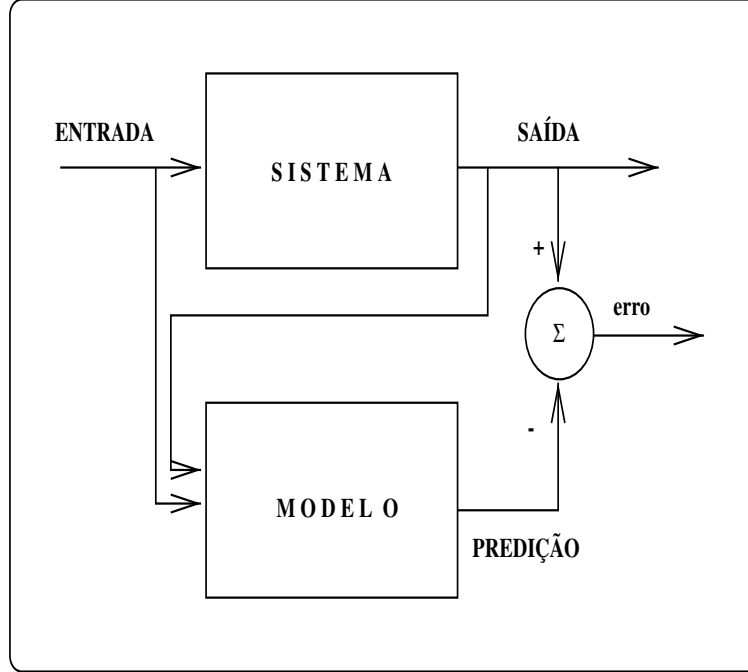


Figura 1.12: Erro de Predição de Saída.

O objetivo deste **preditor** é obter o menor valor de erro entre as saídas  $y_m(k)$  e  $y(k)$ .

Agora a definição de critério que busque  $\Theta$  que torne o erro entre o valor estimado e o valor real o menor possível se faz necessária. Veremos a seguir alguns destes critérios.

#### 1.2.4 Alguns Métodos Para Minimizar o Erro do Estimador

Métodos clássicos para minimizar o erro entre o valor real e o valor estimado da saída de um processo são vários e muito estudados. Dentre os métodos existentes destacamos dois métodos determinísticos <sup>2</sup>:

- Mínimos Quadrados.
- Mínimos Quadrados Recursivo.

<sup>2</sup>Para o estudo de métodos estocásticos sugerimos as referências: El-Sherief & Sinha, [312]; Faure, [338]; Katayama & Picci, [586]; Luenberger, [697]; Maybeck, [730]; Saridis, [933]; Caines, [171]; Sen & Sinha, [959]; Söderström, [990] entre outras para maiores detalhes, pois neste trabalho nossos objetivos prioritários são outros.



### 1.2.5 Mínimos Quadrados

Estudos de astronomia que tiveram origem na Babilônia em 300 AC é que estimularam a invenção e o desenvolvimento do método dos mínimos quadrados, termo este utilizado pela primeira vez em 1795 e publicado em 1806 no livro *Nouvelles Méthodes pour la Détermination des Orbites des Comètes* por Legendre. Mas o conceito básico da técnica dos mínimos quadrados foi elaborado por Carl Friedrich Gauss, matemático e astrônomo alemão (1777-1855), que enunciou: “...a linha reta da forma  $y = a + bx$  poderá representar adequadamente as medidas ponderadas  $(x_1, y_1) \dots (x_n, y_n)$  desde que a soma dos quadrados das distâncias destes pontos a esta linha reta representativa destes pontos seja mínima, quando esta distância for medida na vertical...”, publicado em 1809 no famoso “*Theoria Motus Corporum Coelestium*”; Seal [955]; Stigler [1007]; Plackett [880].

Este conceito foi aplicado originalmente com grande êxito no primeiro dia do século dezenove quando um novo corpo celeste, batizado de Ceres, que mais tarde saber-se-ia tratar-se de um asteroide, fora descoberto, e poucas semanas depois fora perdido de vista. Gauss que tinha habilidade excepcional para cálculos, além da vantagem adicional de estar elaborando a teoria do método dos mínimos quadrados, aceitou o desafio de calcular, a partir das poucas observações do asteroide registradas, a órbita em que ele se movia; o número de observações disponíveis não era superior a seis. Para a tarefa de calcular a órbita a partir de um número limitado de observações ele inventou um processo, hoje chamado método de Gauss (método de observações elípticas a partir de três observações), que ainda hoje é usado para acompanhar satélites. Como resultado da aplicação de seu método, Ceres foi redescoberto no fim do ano de 1800, muito próximo da posição indicada por seus cálculos. Este fato provavelmente pode ser considerado como a primeira identificação de sistemas realizada. Devido a esta identificação, Gauss teve sua reputação firmada na astronomia e na matemática e em 1807 Gauss foi nomeado diretor do Observatório de Göttingen, posto que conservou durante 40 anos.

Seja um sistema discreto cuja relação entrada-saída, supondo um erro  $e_k$  no  $k$ -ésimo instante, é:

$$\begin{aligned} y_k = & b_0 u_k + b_1 u_{k-1} + b_2 u_{k-2} + \dots + b_n u_{k-n} + \\ & - a_1 y_{k-1} - a_2 y_{k-2} - \dots - a_n y_{k-n} + e_k \end{aligned} \quad (1.14)$$

onde o conjunto de entradas,  $u_k$  e saídas  $y_k$  é:

$$u_0, u_1, \dots, u_N, y_0, y_1, \dots, y_N$$

Para  $N$  medidas da saída temos:

$$\begin{aligned}
\mathbf{y}_1 &= b_0 \mathbf{u}_{1-0} + b_1 \mathbf{u}_{1-1} + b_2 \mathbf{u}_{1-2} + \dots + b_k \mathbf{u}_{1-n} + \\
&\quad - a_1 \mathbf{y}_{1-1} - a_2 \mathbf{y}_{1-2} - \dots - a_n \mathbf{y}_{1-n} + e_1 \\
\mathbf{y}_2 &= b_0 \mathbf{u}_{2-0} + b_1 \mathbf{u}_{2-1} + b_2 \mathbf{u}_{2-2} + \dots + b_k \mathbf{u}_{2-n} + \\
&\quad - a_1 \mathbf{y}_{2-1} - a_2 \mathbf{y}_{2-2} - \dots - a_n \mathbf{y}_{2-n} + e_2 \\
&\vdots \quad \quad \quad \vdots \\
\mathbf{y}_N &= b_0 \mathbf{u}_{N-0} + b_1 \mathbf{u}_{N-1} + b_2 \mathbf{u}_{N-2} + \dots + b_k \mathbf{u}_{N-n} + \\
&\quad - a_1 \mathbf{y}_{N-1} - a_2 \mathbf{y}_{N-2} - \dots - a_n \mathbf{y}_{N-n} + e_N
\end{aligned} \tag{1.15}$$

onde  $N$  refere-se aos dados atuais.

Fazendo

$$\mathbf{Y} = [\mathbf{y}_1 \dots \mathbf{y}_N]^T, \tag{1.16}$$

o vetor de  $N$  medidas da saída; a matriz de dados:

$$\mathbf{\Psi} = \begin{bmatrix} \mathbf{u}_{1-0} & \mathbf{u}_{1-1} & \dots & \mathbf{u}_{1-n} & -\mathbf{y}_{1-1} & -\mathbf{y}_{1-2} & \dots & \mathbf{y}_{1-n} \\ \mathbf{u}_{2-0} & \mathbf{u}_{2-1} & \dots & \mathbf{u}_{2-n} & -\mathbf{y}_{2-1} & -\mathbf{y}_{2-2} & \dots & \mathbf{y}_{2-n} \\ \vdots & & & & & & & \\ \mathbf{u}_{N-0} & \mathbf{u}_{N-1} & \dots & \mathbf{u}_{N-n} & -\mathbf{y}_{N-1} & -\mathbf{y}_{N-2} & \dots & \mathbf{y}_{N-n} \end{bmatrix} \tag{1.17}$$

o vetor de parâmetros a ser estimado:

$$\mathbf{\Theta} = \begin{bmatrix} b_0 \\ b_1 \\ \vdots \\ b_n \\ a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \tag{1.18}$$

e o vetor de erros:

$$\mathbf{\varepsilon} = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix} \tag{1.19}$$

pode-se reescrever a equação (1.15) como:

$$\mathbf{Y} = \mathbf{\Psi} \mathbf{\Theta} + \mathbf{\varepsilon} \tag{1.20}$$

Se há erros,  $\varepsilon = \mathbf{Y} - \Psi\Theta$  podemos enunciar o problema de mínimos quadrados como:

Dados  $\mathbf{Y}$  e  $\Psi$ , determine o estimador  $\hat{\Theta}$  de  $\Theta$  que minimize o critério  $\mathbf{J}$ :

$$\mathbf{J} = \frac{1}{2} \|\varepsilon\|^2 = \frac{1}{2} (\mathbf{Y} - \Psi\Theta)^T (\mathbf{Y} - \Psi\Theta) \quad (1.21)$$

Da condição necessária de mínimo,  $\frac{\partial \mathbf{J}}{\partial \Theta} = 0$  tem-se para a minimização da forma quadrática <sup>3</sup>:

$$\begin{aligned} \mathbf{J} &= (\mathbf{Y} - \Psi\Theta)^T (\mathbf{Y} - \Psi\Theta) \\ &= (\mathbf{Y} - \Psi\Theta)^T (\mathbf{Y} - \Psi\Theta) \\ &= (\mathbf{Y}^T - (\Psi\Theta)^T) (\mathbf{Y} - \Psi\Theta) \\ &= \mathbf{Y}^T \mathbf{Y} - \Theta^T \Psi^T \mathbf{Y} - \mathbf{Y}^T \Psi \Theta + \Theta^T \Psi^T \Psi \Theta \end{aligned} \quad (1.22)$$

que a derivada de  $\mathbf{J}$  em relação a  $\Theta$  é <sup>4</sup>:

$$\frac{\partial \mathbf{J}}{\partial \Theta} = -\mathbf{Y}^T \Psi - \mathbf{Y}^T \Psi + 2\Theta^T \Psi^T \Psi = 0 \quad (1.23)$$

Deste modo podemos escrever a equação normal:

$$\Psi^T \Psi \hat{\Theta} = \Psi^T \mathbf{Y} \quad (1.24)$$

donde

$$\hat{\Theta} = (\Psi^T \Psi)^{-1} \Psi^T \mathbf{Y} \equiv \Psi^\dagger \mathbf{Y} \quad (1.25)$$

é o estimador ótimo de mínimos quadrados dos parâmetros do sistema e  $\Psi^\dagger = (\Psi^T \Psi)^{-1} \Psi^T$  é a matriz pseudo-inversa de  $\Psi$ .

Para a solução da equação (1.25) diversos métodos foram desenvolvidos, dentre os quais destacamos: Golub [394, 395]. O desempenho deste método é baseado nas medidas efetuadas, deste modo podemos ter diferentes erros para cada medida efetuada. Para melhorar o desempenho do método usa-se um tratamento conhecido como **mínimos quadrados ponderado**, onde

---

3

$$(\mathbf{A} + \mathbf{B})^T = \mathbf{A}^T + \mathbf{B}^T \quad \text{e} \quad (\mathbf{AB})^T = \mathbf{B}^T \mathbf{A}^T$$

4

$$\frac{\partial(\mathbf{A}^T \mathbf{Q} \mathbf{A})}{\partial \mathbf{A}} = 2\mathbf{A}^T \mathbf{Q}$$

dados melhores ou mais recentes terão uma influência maior no cálculo dos parâmetros; para tanto define-se o critério:

$$\mathbf{J}(\Theta) = \sum_{k=1}^N \omega(k) \varepsilon^2(k; \Theta) \quad (1.26)$$

sendo  $\omega$  o valor da influência de cada medida para o cálculo dos parâmetros.

Neste caso o estimador  $\hat{\Theta}$  de mínimos quadrados é:

$$\hat{\Theta} = (\Psi^T \omega \Psi)^{-1} \omega \Psi^T \mathbf{Y} \quad (1.27)$$

### 1.2.6 Mínimos Quadrados Recursivo

No item anterior analisamos o tratamento em bloco das observações efetuadas; tal procedimento não será adequado quando se desejar elevado desempenho computacional “*on-line*” com novos dados em cada instante de amostragem. A cada nova medida é necessário efetuar inversão de matriz o que pode não ser trivial. O método de mínimos quadrados recursivo, Goodwin & Payne, [403]; Ljung & Söderström, [678], que leva em consideração somente a chegada da nova observação para atualizar os parâmetros estimados foi desenvolvido para superar tal dificuldade e é utilizado por exemplo quando:

- O objetivo é realizar controle adaptativo, quando a ação (controle, filtragem, entre outros) for baseada no comportamento recente do processo, Figura 1.13.
- O processador de controle tem memória limitada e nem todos os dados de entrada e saída podem ser armazenados.
- Algoritmos para controle em tempo-real são utilizados e é necessário variar os parâmetros a cada momento. Os parâmetros  $\Theta$  não convergem quando o tempo,  $k$ , tende a infinito.
- Algoritmos de detecção de falha são necessários. Quando um dado algoritmo de detecção de falhas atuar, o algoritmo de identificação recursiva irá utilizar os últimos dados para estimar os parâmetros, para retorno do processo a normalidade.

As equações que apresentamos a seguir mostram como tornar recursiva a equação (1.25). Por exemplo, se desejamos estimar o valor de uma constante:

$$\mathbf{y}_k = b$$

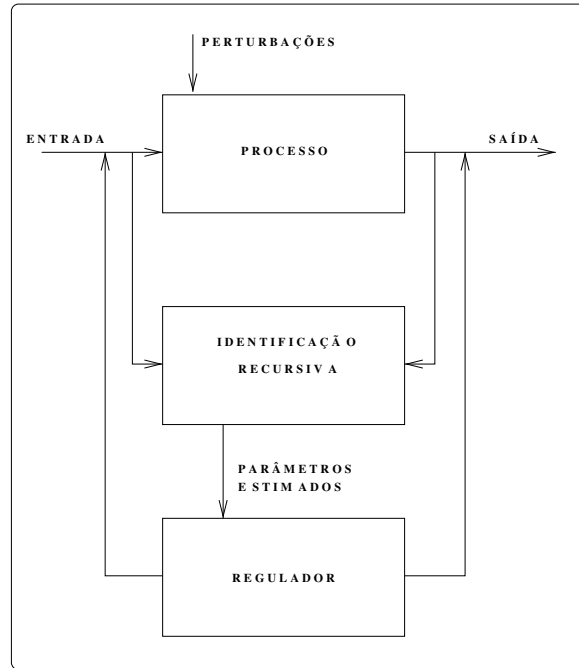


Figura 1.13: Um Esquema Geral Para Controle Adaptativo.

de forma recursiva, podemos efetuar uma série de medidas. Neste caso tem-se:

$$\psi_k = 1 \quad , \quad \Theta = b$$

donde,

$$\Psi = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix}$$

Fazendo as substituições necessárias na equação (1.25) temos:

$$\hat{\Theta} = \frac{1}{N} [ \mathbf{y}_1 + \mathbf{y}_2 + \dots + \mathbf{y}_N ]$$

Então podemos generalizar o exemplo acima para:

$$\hat{\Theta} = \frac{1}{k} \sum_{s=1}^k \mathbf{y}_s \quad (1.28)$$

Um algoritmo recursivo é obtido a partir da equação (1.28) do seguinte modo:

$$\begin{aligned}
\hat{\Theta}_k &= \frac{1}{k} \left[ \sum_{s=1}^{k-1} \mathbf{y}_s + \mathbf{y}_k \right] \\
&= \frac{1}{k} \left[ (k-1) \hat{\Theta}_{k-1} + \mathbf{y}_k \right] \\
&= \frac{1}{k} \left[ (k \hat{\Theta}_{k-1} - \hat{\Theta}_{k-1} + \mathbf{y}_k \right] \\
&= \hat{\Theta}_{k-1} + \frac{1}{k} \left[ \mathbf{y}_k - \hat{\Theta}_{k-1} \right]
\end{aligned} \tag{1.29}$$

Portanto a estimativa de  $\hat{\Theta}_k$  no  $k$ -ésimo instante é dependente de  $\hat{\Theta}_{k-1}$  e de um termo de correção. Note que neste caso o fator  $\frac{1}{k}$  pode ser visto como um fator de esquecimento pois quanto maior for  $k$  menor será o efeito do termo de correção. Para o exemplo façamos:

$$\mathbf{P}_k = \Psi^T \Psi = \frac{1}{k}$$

onde ,

$$\Psi^T = (1, 1, \dots, 1, 1) \quad , \quad k = 1, 2, \dots, N$$

ou seja:

$$\mathbf{P}_k^{-1} = k = \mathbf{P}_{k-1}^{-1} + 1$$

logo:

$$\mathbf{P}_k = \frac{1}{\mathbf{P}_{k-1}^{-1} + 1} = \frac{\mathbf{P}_{k-1}}{1 + \mathbf{P}_{k-1}}$$

Para uma discussão formal do procedimento de mínimos quadrados recursivos reescrevamos a equação (1.25) como:

$$\hat{\Theta}_k = [\Psi_k^T \Psi_k]^{-1} \Psi_k^T \mathbf{Y}_k$$

Quando da adição de novas medidas,  $\Psi_{k+1}$  pode ser reescrita como:

$$\Psi_{k+1} = \begin{bmatrix} \Psi_k \\ \psi_{k+1} \end{bmatrix} \quad , \quad \mathbf{Y}_{k+1} = \begin{bmatrix} \mathbf{Y}_k \\ \mathbf{y}_{k+1} \end{bmatrix} \tag{1.30}$$

onde  $\psi_{k+1}$  é um vetor com as novas medidas de entradas e saídas no sistema.

A estimativa de  $\hat{\Theta}_{k+1}$  pode ser reescrita como:

$$\begin{aligned}
\hat{\Theta}_{k+1} &= [\Psi_{k+1}^T \Psi_{k+1}]^{-1} \Psi_{k+1}^T \mathbf{Y}_{k+1} \\
&= \left[ \begin{bmatrix} \Psi_k^T & \psi_k^T \end{bmatrix} \begin{bmatrix} \Psi_{k+1} \\ \psi_{k+1} \end{bmatrix} \right]^{-1} \left[ \begin{bmatrix} \Psi_k^T & \psi_{k+1}^T \end{bmatrix} \begin{bmatrix} \mathbf{Y}_k \\ \mathbf{y}_{k+1} \end{bmatrix} \right]
\end{aligned} \tag{1.31}$$

que pode ser reescrita como:

$$\hat{\Theta}_{k+1} = \left[ \sum_{s=1}^{k+1} \psi_s \psi_s^T \right]^{-1} \left[ \sum_{s=1}^{k+1} \psi_s \mathbf{y}_s \right] \quad (1.32)$$

Introduzindo a notação:

$$\mathbf{P}_{k+1} = \left[ \sum_{s=1}^{k+1} \psi_s \psi_s^T \right]^{-1} \quad (1.33)$$

podemos reescrever a equação (1.33) como:

$$\mathbf{P}_{k+1}^{-1} = \mathbf{P}_k^{-1} + \psi_{k+1} \psi_{k+1}^T \quad (1.34)$$

Substituindo em (1.32), (1.33) e (1.34) temos:

$$\begin{aligned} \hat{\Theta}_{k+1} &= \mathbf{P}_{k+1} \left[ \sum_{s=1}^k \psi_s y_s + \psi_{k+1} \mathbf{y}_{k+1} \right] \\ &= \mathbf{P}_{k+1} [\mathbf{P}_k^{-1} \hat{\Theta}_k + \psi_{k+1} \mathbf{y}_{k+1}] \\ &= \hat{\Theta}_k + \mathbf{P}_{k+1} \psi_{k+1} [\mathbf{y}_{k+1} - \psi_{k+1}^T \hat{\Theta}_k] \end{aligned} \quad (1.35)$$

Assim podemos escrever:

$$\hat{\Theta}_{k+1} = \hat{\Theta}_k + \mathbf{K}_{k+1} \varepsilon_{k+1} \quad (1.36)$$

$$\mathbf{K}_{k+1} = \mathbf{P}_{k+1} \psi_{k+1} \quad (1.37)$$

$$\varepsilon_{k+1} = \mathbf{y}_{k+1} - \psi_{k+1}^T \hat{\Theta}_k \quad (1.38)$$

O termo  $\varepsilon_{k+1}$  pode ser interpretado como um preditor do erro de saída, isto é da diferença entre a saída  $\mathbf{y}_{k+1}$  e o valor estimado desta saída dado por  $\psi_{k+1}^T \hat{\Theta}_k$ . O vetor  $\mathbf{K}_{k+1}$  pode ser interpretado como um fator de compensação ou ganho das medidas. Para o cálculo de (1.37) precisamos em (1.34) inverter uma matriz a cada nova medida; este procedimento poderá consumir um tempo de processamento desnecessário; o melhor é utilizar o lema de inversão matricial<sup>5</sup>:

$$(\mathbf{A} + \mathbf{BCD})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{B} (\mathbf{C}^{-1} + \mathbf{DA}^{-1} \mathbf{B})^{-1} \mathbf{DA}^{-1}$$

---

<sup>5</sup>O histórico do desenvolvimento do lema de inversão matricial é dado por Henderson & Searle, [484].

A prova deste lema é dada por substituição direta:

$$\begin{aligned}
& (\mathbf{A} + \mathbf{BCD}) [\mathbf{A}^{-1} - \mathbf{A}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}] = \\
& = \mathbf{I} + \mathbf{BCDA}^{-1} - \mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} - \mathbf{BCDA}^{-1}\mathbf{B}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} \\
& = \mathbf{I} + \mathbf{BCDA}^{-1} - (\mathbf{B} + \mathbf{BCDA}^{-1}\mathbf{B})(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} \\
& = \mathbf{I} + \mathbf{BCDA}^{-1} - \mathbf{B}(\mathbf{I} + \mathbf{CDA}^{-1}\mathbf{B})(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1}
\end{aligned}$$

Substituindo  $\mathbf{I} = \mathbf{CC}^{-1}$  e colocando  $\mathbf{C}$  em evidência teremos:

$$\begin{aligned}
& = \mathbf{I} + \mathbf{BCDA}^{-1} - \mathbf{BC}(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})(\mathbf{C}^{-1} + \mathbf{DA}^{-1}\mathbf{B})^{-1}\mathbf{DA}^{-1} \\
& = \mathbf{I} + \mathbf{BCDA}^{-1} - \mathbf{BCDA}^{-1} = \mathbf{I}
\end{aligned}$$

Então, aplicando o lema 1.2.6 na equação (1.34) com:  $\mathbf{P}_{k-1}^{-1} = \mathbf{A}$ ,  $\psi_k = \mathbf{B}$ ,  $1 = \mathbf{C}$  e finalmente  $\psi_k^T = \mathbf{D}$  temos:

$$\mathbf{P}_{k+1} = \mathbf{P}_k - \mathbf{P}_k \psi_{k+1} \psi_{k+1}^T \mathbf{P}_k / [1 + \psi_{k+1}^T \mathbf{P}_k \psi_{k+1}] \quad (1.39)$$

Então podemos reescrever  $\mathbf{K}_{k+1}$  dado em (1.37) como:

$$\begin{aligned}
\mathbf{K}_{k+1} &= \mathbf{P}_k \psi_{k+1} - \mathbf{P}_k \psi_{k+1} \psi_{k+1}^T \mathbf{P}_k \psi_k / [1 + \psi_{k+1}^T \mathbf{P}_k \psi_{k+1}] \\
&= \mathbf{P}_k \psi_{k+1} / [1 + \psi_{k+1}^T \mathbf{P}_k \psi_{k+1}]
\end{aligned} \quad (1.40)$$

O método dos mínimos quadrados recursivo está completo.

### 1.3 Teoria da Realização

A teoria de realização diz respeito à determinação do modelo no espaço de estado, a partir dos parâmetros de Markov do sistema. O tratamento dado para a identificação de sistemas até o momento, é aplicável apenas a sistemas tendo uma única entrada e uma única saída, não sendo aplicável com a mesma simplicidade a sistemas de múltiplas entradas e múltiplas saídas. Quando usamos a teoria da realização a identificação de sistemas multivariáveis faz-se com a mesma simplicidade que a de sistemas monovariáveis e isto é certamente muito importante, além de sê-lo de forma não recursiva e sem exigir otimização como é o caso do método de predição de erro.

Uma classe particular de soluções, discutida por Ljung [679] ou Söderström e Stoica [989], mostra uma versão para solucionar o problema de estimação de parâmetros usando esquemas iterativos de otimização. Uma dificuldade encontrada neste tipo de abordagem do problema está na seleção de uma classe de modelos e na parametrização do modelo escolhido.



Uma classe alternativa de métodos para solucionar o problema MIMO são os modelos de identificação nos subespaços; Van Overschee e De Moor [1094]. A principal característica destes métodos é a aproximação de um subespaço, definido pelo espaço coluna, ou pelo espaço linha das matrizes determinadas pelo modelo aos dados de entrada-saída do processo. O modelo paramétrico invariante no tempo, neste caso, é obtido através desta aproximação de subespaços.

O modelo a ser obtido representa o comportamento entrada-saída da parte determinística (ou ideal) do sistema a ser identificado.

A identificação de sistemas de múltiplas entradas - múltiplas saídas lineares e invariantes no tempo via modelos de espaço de estado, através de medidas de entrada-saída, é um problema de central importância na análise e projeto de sistemas de controle.

Em termos gerais pode-se ver o problema como o de mapear a relação entre as seqüências de dados de entrada-saída e os parâmetros não conhecidos usados para definir o modelo no espaço de estado.

O modelo linear multivariável mais geral no espaço de estado pode ser escrito como:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k + \mathbf{v}_k \end{cases} \quad (1.41)$$

onde  $\mathbf{w}_k$  e  $\mathbf{v}_k$  são sinais que levam em consideração ruídos nas medidas e no processo entre outros. Os ruídos ou perturbações originários de um processo podem ser semelhantes aos resíduos do cálculo de um conjunto de equações. Para o caso determinístico, o conjunto de equações que representam um sistema multivariável é dado por:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k \\ \mathbf{y}_k &= \mathbf{C}_k \mathbf{x}_k + \mathbf{D}_k \mathbf{u}_k \end{cases} \quad (1.42)$$

que também pode ser reescrita como:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix}$$

Reescrevendo tal equação para  $n$  instantes teremos:

$$\begin{bmatrix} \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \dots & \mathbf{x}_{n+1} \\ \mathbf{y}_k & \mathbf{y}_{k+1} & \dots & \mathbf{y}_n \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k & \mathbf{x}_{k+1} & \dots & \mathbf{x}_n \\ \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \mathbf{u}_n \end{bmatrix} \quad (1.43)$$

Se as quantidades  $\mathbf{u}_k$ ,  $\mathbf{y}_k$  para  $k = 0, 1, 2, \dots, n$  e  $\mathbf{x}_k$  para  $k = 0, 1, 2, \dots, (n+1)$  são conhecidas, então  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  e  $\mathbf{D}$  podem ser determinadas através da equação (1.43) com técnicas de mínimos quadrados. Infelizmente, a seqüência dos vetores de estado  $\mathbf{x}_k$  para  $k = 0, 1, 2, \dots, (n+1)$  é

geralmente não conhecida e é necessário estimá-la e ou medi-la, se possível. Na prática, usualmente somente as seqüências de entradas e de saídas  $\mathbf{u}_k$  e  $\mathbf{y}_k$  para  $k = 0, 1, 2, \dots, n+1$ , respectivamente, são disponíveis. Então, outras formas precisam ser encontradas para uso em identificação. Veremos que os métodos de subespaços constituem uma destas formas.

Na Figura 1.14 apresentamos uma representação das equações (1.42).

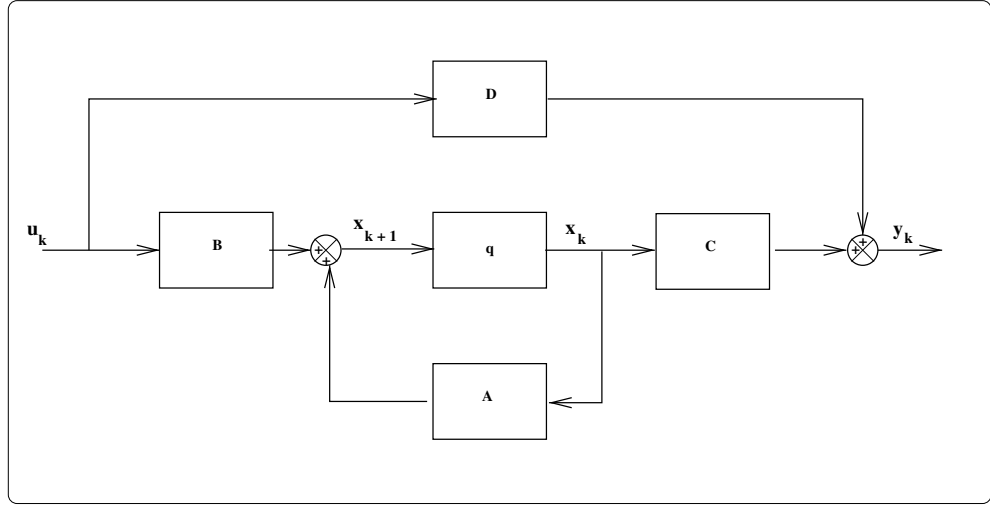


Figura 1.14: Representação Genérica de Sistema Multivariável.

O modelo apresentado para representar um dado sistema é único, a menos de uma transformação de similaridade,  $\mathbf{S}$  não-singular, das equações (1.42):

$$\begin{cases} \underbrace{\mathbf{S}\mathbf{x}_{k+1}}_{\mathbf{z}_{k+1}} = \underbrace{\mathbf{S}\mathbf{A}_k\mathbf{S}^{-1}}_{\mathbf{\bar{A}}}\underbrace{\mathbf{S}\mathbf{x}_k}_{\mathbf{z}_k} + \underbrace{\mathbf{S}\mathbf{B}_k}_{\mathbf{\bar{B}}}\mathbf{u}_k \\ \mathbf{y}_k = \underbrace{\mathbf{C}_k\mathbf{S}^{-1}}_{\mathbf{\bar{C}}}\underbrace{\mathbf{S}\mathbf{x}_k}_{\mathbf{z}_k} + \mathbf{D}_k\mathbf{u}_k \end{cases}$$

Através deste exemplo vemos que somente o sistema de matrizes e o estado são alterados e que os dados de entrada e saída não sofrem modificações. Isto implica que somente para uma escolha muito específica de um sistema de coordenadas para o espaço de estado (escolha da matriz  $\mathbf{S}$ ), os componentes do vetor de estado poderão ter um significado físico.

## 1.4 Métodos de Subespaços

O desenvolvimento de métodos de subespaços pode ser encontrado em diversos trabalhos dentre os quais destacamos: De Moor, [271]; De Moor &

Overschee , [277]; De Moor & Overschee & Favoreel, [279]; Aoki, [40, 43]; Verhaegen & Dewilde, [1113] Viberg, [1124]; McKelvey, [739] entre outros.

Os modelos matemáticos de sistemas dinâmicos são usados para a análise, a simulação, a predição, a otimização, a monitoração, a deteção de falha, o treinamento e o controle entre outras inúmeras aplicações. Há diversas aproximações para gerar um modelo de um sistema. Por exemplo, poder-se-ia gerar o modelo que descreve o comportamento do sistema a partir de leis básicas da física ou da química. Aproximação deste tipo é chamada “caixa-preta” e é adequada para exemplos simples, no entanto modelos do mundo real tem maior complexidade.

Em alguns casos, as equações do sistema são conhecidas, e parâmetros desconhecidos são estimados usando algum método de estimação de parâmetros.

Noutros casos a modelagem é realizada pela identificação do sistema: primeiramente medidas do sistema são coletadas e então ele é modelado usando-se a chamada identificação por “caixa-preta”. Tal método consiste basicamente em definir um modelo paramétrico e então determinar os parâmetros do modelo de tal maneira que as medidas sejam explicadas com a maior acuidade possível pelo modelo.

Classicamente, isto é realizado formulando o problema de identificação como um problema de otimização, em que as variáveis são os parâmetros desconhecidos do modelo, as restrições são as equações do modelo e a função objetivo é uma medida do desvio entre as observações e as predições ou as simulações que resultam do modelo.

Os modelos e/ou os sistemas podem ser divididos em classes, tais como linear e não-linear, invariante ou variante no tempo, discreto ou contínuo no tempo, com parâmetros distribuídos ou concentrados, etc.

A classe de modelos linear invariante no tempo com parâmetros concentrados, que parece ser muito restrita para comportamentos reais, pode aproximar muito bem dados de entrada-saída de processos práticos e ou industriais.

O campo de identificação linear de sistemas certamente não é recente embora possamos com segurança dizer que começou a ter maior desenvolvimento a partir de 1970.

Neste período as pesquisas geraram muitos resultados e experiências práticas. Podemos mencionar algumas referências clássicas tais como: Åström & Eykhoff, [51]; Eykhoff, [320]; Ljung, [679] e Söderström & Stoica, [989] - para identificação de sistemas - ; Box & Jenkins, [150] - para análise de séries temporais.

Em meados da década de 80 ocorreu o nascimento de um novo tipo de algoritmos para modelagem de séries temporais e para identificação de sistemas lineares no espaço de estado, os chamados métodos de subespaços que

são baseados na teoria da realização de sistemas.

Os métodos de identificação de sistema linear em subespaços consideram sistemas e modelos da seguinte forma:

$$\begin{cases} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \end{cases} \quad (1.44)$$

com

$$\mathbf{E} \left[ \begin{pmatrix} \mathbf{w}_p \\ \mathbf{v}_p \end{pmatrix} \begin{pmatrix} \mathbf{w}_q \\ \mathbf{v}_q \end{pmatrix}^T \right] = \begin{pmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{pmatrix} \delta_{p,q} \geq 0. \quad (1.45)$$

onde  $\mathbf{E}$  representa o operador esperança e  $\delta_{p,q}$  o delta de Kronecker.

Os vetores  $\mathbf{u}_k \in \Re^{m \times 1}$  e  $\mathbf{y}_k \in \Re^{l \times 1}$  são medidos no tempo  $k$  respectivamente com as  $m$  entradas e as  $l$  saídas do processo.

O vetor  $\mathbf{x}_k$  representa o estado do processo no instante de tempo  $k$ ;  $\mathbf{v}_k \in \Re^{l \times 1}$  e  $\mathbf{w}_k \in \Re^{n \times 1}$  são vetores de sinais não observáveis,  $\mathbf{v}_k$  é o vetor de ruído nas medidas e  $\mathbf{w}_k$  é o vetor de ruído no processo. Supõe-se estes vetores com média zero, ruído branco, estacionários e não correlatos com as entradas,  $\mathbf{u}_k$ .

$\mathbf{A} \in \Re^{n \times n}$  é a matriz do sistema;  $\mathbf{B} \in \Re^{n \times m}$  é a matriz de entrada;  $\mathbf{C} \in \Re^{l \times n}$  é a matriz de saída enquanto  $\mathbf{D} \in \Re^{l \times m}$  é a matriz de realimentação.

As matrizes  $\mathbf{Q} \in \Re^{n \times n}$ ;  $\mathbf{S} \in \Re^{n \times l}$  e  $\mathbf{R} \in \Re^{l \times l}$  são matrizes de covariância das seqüências de ruídos  $\mathbf{w}_k$  e  $\mathbf{v}_k$ . Na identificação por subespaços supõe-se tipicamente que o número de dados de medidas disponível tende a infinito e que os dados são ergódicos.

O problema principal tratado em identificação de sistema em subespaços é:

*Dado um grande número de medidas de entradas,  $\mathbf{u}_k$ , e de saídas  $\mathbf{y}_k$ , geradas por um sistema desconhecido representado pelas equações (1.44) e (1.45), determine a ordem  $n$  do sistema desconhecido, as matrizes do sistema,  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  e  $\mathbf{D}$ , a menos de uma transformação de similaridade e as matrizes de covariância do ruído,  $\mathbf{Q}$ ,  $\mathbf{R}$  e  $\mathbf{S}$ .*

Os métodos de subespaços originam basicamente de uma união entre três áreas: a teoria de sistemas, a geometria e a álgebra linear numérica. Algumas referências históricas e básicas sobre a metodologia de identificação de sistema em subespaços são:

- a) teoria da realização - Ho & Kalman, [490] e [491]; Zeiger & McEwen, [1195]; Kung, [622] e [626] e Moore, [783].
- b) realização estocástica - Van Overschee & De Moor, [1089]; Akaike, [8] e [10]; Faurre, [338] e Peternell, [865].

- c) identificação determinística - De Moor & Vandewalle & Vanderberghe & Miegheem, [260]; De Moor, [259] e Moonen & De Moor & Vandenberghe & Vandewalle [763].
- d) combinação entre identificação determinística e estocástica - Larimore, [643] e [640]; Verhaegen & Dewilde, [1113]; De Moor & Van Overschee, [1090] e [1094] e Verhaegen [1116] .

Trabalhos e livros enfatizam diferentes aspectos da identificação de sistemas e processamento de sinais em subespaços dos quais destacamos: Aoki, [40, 43]; De Moor, [271]; Van Der Veen & Deprettere & Swindlehurst, [1070]; Viberg, [1124]; De Moor & Van Overschee, [277] e Van Overschee, [1091], além de números especiais das revistas *Automatica: Special Issue On Trends in System Identification* Volume 31, Number 12, december, 1995 e *Special Issue On Statistical Signal Processing and Control* Volume 30, Number 1, January, 1994 e *Signal Processing* , com uma edição especial sobre *Signal Processing* no Volume: 52, Issue: 2 July, 1996.

Uma das idéias importantes na origem do desenvolvimento de algoritmos de subespaços foi a reintrodução do conceito do estado  $\mathbf{x}_k$  de um sistema dinâmico dentro do contexto de identificação de sistema.

Em contraste com tradicionais algoritmos de identificação, os algoritmos de subespaços calculam (explícita ou implicitamente) o estado, para depois determinar o modelo. Esta é a principal diferença entre os métodos de subespaços e os métodos tradicionais de predição de erro.

Por que primeiro obter os estados, diretamente dos dados de entrada e saída, e somente depois obter o modelo de espaço de estado? A resposta é dada em duas partes:

- a) Quando os estados  $\mathbf{x}_k$  estão disponíveis pode-se obter as matrizes  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  e  $\mathbf{D}$  do modelo no espaço de estado do sistema resolvendo um problema de mínimos quadrados a partir das equações (1.44) e (1.45) porque são os elementos desconhecidos nas equações do estado de espaço. Também as matrizes de covariância  $\mathbf{Q}$ ,  $\mathbf{R}$  e  $\mathbf{S}$  podem ser obtidas através dos resíduos do problema de mínimos quadrados. Assim o problema de identificação é “ linearizado ” no sentido de que foi reduzido a um problema de mínimos quadrados, diferentemente dos tradicionais métodos de predição de erro onde problemas de otimização não-linear precisam ser resolvidos.
- b) Na maioria dos métodos modernos de controle, o conhecimento do estado é necessário para realizar realimentação de estado. O fato de que os estados são encontrados diretamente dos dados, isto é, mesmo antes que o modelo da planta esteja calculado, característica dos métodos de

subespaços, os torna muito adequados para finalidades de controle e de controle ótimo.

## Capítulo 2

# ALGUNS ELEMENTOS DE ÁLGEBRA LINEAR E ÁLGEBRA LINEAR NUMÉRICA

### 2.1 Introdução

Este capítulo trata de álgebra linear, matrizes e suas aplicações. No campo da álgebra linear numérica, já há um tratamento enciclopédico dado por Golub & Van Loan, [399]. Este capítulo não é uma tentativa de resumir o livro citado ou outros excelentes livros dentre os quais destacamos Strang, [1016, 1017]. Nosso objetivo é apresentar algumas noções fundamentais de álgebra linear e de álgebra linear numérica tendo em vista suas aplicações nesta tese, tanto nos seus aspectos teóricos como nos seus aspectos computacionais. Uma parte é tratada neste Capítulo, enquanto que a outra é tratada sobretudo no Capítulo 4. Em especial neste último a organização e o conteúdo têm aspectos e resultados que reputamos originais.

### 2.2 Alguns Conceitos Fundamentais

Iniciamos este capítulo com alguns conceitos fundamentais de álgebra linear essenciais para o ponto de vista que desejamos dar a todo este trabalho, particularmente aos Capítulos 4 e 5, pois para o bom entendimento de sistemas lineares, os conceitos básicos de álgebra linear são indispensáveis, em especial os conceitos de espaços vetoriais, operadores lineares e projeções. A nosso ver alguns conceitos como autovalores, autovetores e similaridade são bem conhecidos, enquanto outros não o são; eles serão tratados neste

e no Capítulo 4, onde uma abordagem personalizada e dirigida aos nossos propósitos é adotada.

Nos restringimos aos espaços vetoriais com dimensões finitas para os quais as representações matriciais de operadores lineares existem e são muito úteis.

Este tema pode ser encontrado em diversas referências clássicas das quais destacamos: Chen, [189]; Birkhoff & MacLane, [108]; Noble & Daniel, [823] Strang, [1016, 1017]; Zadeh & Desoer, [1191]; Halmos, [445]; Bottura, [148] e Horn & Johnson [500].

**Definição 2.1** *Um corpo consiste de um conjunto  $\mathcal{F}$  de elementos chamados escalares e de duas operações chamadas adição e multiplicação definidas sobre  $\mathcal{F}$  e satisfazendo os seguintes axiomas:*

**Axioma 2.1** *A todo par de escalares  $\alpha$  e  $\beta$  em  $\mathcal{F}$  corresponde um escalar  $\alpha + \beta$  em  $\mathcal{F}$  chamado soma de  $\alpha$  e  $\beta$  tal que:*

*A1) A adição é comutativa: para todo  $\alpha$  e  $\beta \in \mathcal{F}$ :*

$$\alpha + \beta = \beta + \alpha$$

*A2) A adição é associativa: para todo  $\alpha$ ,  $\beta$  e  $\gamma \in \mathcal{F}$ :*

$$(\alpha + \beta) + \gamma = \alpha + (\beta + \gamma)$$

*A3)  $\mathcal{F}$  contém um único elemento chamado zero e denotado por 0 (zero), tal que para todo  $\alpha \in \mathcal{F}$ :*

$$\alpha + 0 = \alpha$$

*A4) Para todo  $\alpha \in \mathcal{F}$  corresponde um único elemento  $\beta = -\alpha \in \mathcal{F}$  tal que:*

$$\alpha + \beta = 0$$

**Axioma 2.2** *Para qualquer par de escalares  $\alpha$  e  $\beta \in \mathcal{F}$  corresponde um escalar  $\alpha\beta \in \mathcal{F}$  chamado produto de  $\alpha$  e  $\beta$  tal que:*

*M1) A multiplicação é comutativa: para todo  $\alpha$  e  $\beta \in \mathcal{F}$ :*

$$\alpha\beta = \beta\alpha$$



M2) A multiplicação é associativa: para todo  $\alpha, \beta, \gamma \in \mathcal{F}$ :

$$(\alpha\beta)\gamma = \alpha(\beta\gamma)$$

M3)  $\mathcal{F}$  contém um único elemento não nulo 1 tal que:

$$1\alpha = \alpha$$

para todo  $\alpha \in \mathcal{F}$ .

M4) Para todo  $\alpha \in \mathcal{F}$ , que não é o elemento 0 (zero), há um único elemento  $\beta \in \mathcal{F}$  tal que:

$$\beta = \alpha^{-1} = \frac{1}{\alpha}$$

correspondentemente  $\alpha\beta = 1$

M5) A multiplicação é distributiva em relação à adição:

$$\alpha(\beta + \gamma) = \alpha\beta + \alpha\gamma$$

para todo  $\alpha, \beta, \gamma \in \mathcal{F}$ .

Neste trabalho serão muito importantes o corpo dos números reais  $\mathcal{F} = \mathbb{R}$  e o corpo dos números complexos  $\mathcal{F} = \mathbb{C}$ .

**Definição 2.2** Um espaço vetorial ou espaço linear ou espaço vetorial linear sobre um corpo  $\mathcal{F}$  consiste de um conjunto não vazio,  $\mathcal{V}$ , de elementos chamados vetores e de duas operações chamadas adição vetorial e multiplicação por um escalar definidas sobre  $\mathcal{V}$  e  $\mathcal{F}$  e satisfazendo os seguintes axiomas:

**Axioma 2.3** Para todo par de vetores  $\mathbf{x}$  e  $\mathbf{y} \in \mathcal{V}$  corresponde um vetor  $\mathbf{x} + \mathbf{y} \in \mathcal{V}$  chamado soma de  $\mathbf{x}$  e  $\mathbf{y}$  tal que:

a) A adição é comutativa.

$$\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x} \quad , \quad \forall \mathbf{x}, \mathbf{y} \in \mathcal{V}$$

b) A adição é associativa.

$$(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z}) \quad , \quad \forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathcal{V}$$

- c)  $\mathcal{V}$  contém um único elemento chamado vetor nulo, ou origem, e denotado por  $0$  (zero) tal que para todo  $\mathbf{x} \in \mathcal{V}$ :

$$\mathbf{x} + 0 = \mathbf{x}$$

- d) Para todo  $\mathbf{x} \in \mathcal{V}$  corresponde um único elemento  $\mathbf{y} = -\mathbf{x} \in \mathcal{V}$  tal que:

$$\mathbf{x} + \mathbf{y} = 0$$

**Axioma 2.4** Para qualquer escalar  $\alpha \in \mathcal{F}$ , à qualquer vetor  $\mathbf{x} \in \mathcal{V}$  corresponde um vetor  $\alpha\mathbf{x} \in \mathcal{V}$  chamado produto de um escalar por um vetor, tal que:

- a) A multiplicação por escalares é associativa:

$$\alpha (\beta \mathbf{x}) = (\alpha \beta) \mathbf{x} \quad , \quad \forall \quad \alpha, \beta \in \mathcal{F} \text{ e } \forall \mathbf{x} \in \mathcal{V}$$

- b) A multiplicação pelo elemento  $1 \in \mathcal{F}$  de todo  $\mathbf{x} \in \mathcal{V}$  é igual ao  $\mathbf{x} \in \mathcal{V}$ .

**Axioma 2.5** A multiplicação por escalar é:

- a) Distributiva com respeito à adição escalar:

$$(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x} \quad , \quad \forall \alpha, \beta \in \mathcal{F} \text{ e } \forall \mathbf{x} \in \mathcal{V}$$

- b) Distributiva com respeito à adição vetorial:

$$\alpha (\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y} \quad , \quad \forall \alpha \in \mathcal{F} \text{ e } \mathbf{x}, \mathbf{y} \in \mathcal{V}$$

Um espaço vetorial  $\mathcal{V}$  sobre um corpo  $\mathcal{F}$  será denotado por  $(\mathcal{V}, \mathcal{F})$  ou simplesmente por  $\mathcal{V}$ , sempre que possível.

**Definição 2.3** Seja  $\mathcal{V}$  um espaço vetorial sobre o corpo  $\mathcal{F}$ . Um subespaço de  $\mathcal{V}$  é um subconjunto  $\mathcal{Y}$  de  $\mathcal{V}$  que é um espaço vetorial sobre  $\mathcal{F}$  com as duas operações: adição vetorial e multiplicação por um escalar definidas sobre  $\mathcal{V}$  e  $\mathcal{F}$ .

Equivalentemente  $\mathcal{Y}$  é um subespaço de  $\mathcal{V}$  sempre que quaisquer  $\mathbf{y}_1$  e  $\mathbf{y}_2 \in \mathcal{Y}$  e  $\alpha_1, \alpha_2 \in \mathcal{F}$  implicarem em  $\alpha_1 \mathbf{y}_1 + \alpha_2 \mathbf{y}_2 \in \mathcal{Y}$ .

**Definição 2.4** Um vetor  $\mathbf{x} \in (\mathcal{V}, \mathcal{F})$  é chamado uma combinação linear sobre  $\mathcal{F}$  dos vetores  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n \in (\mathcal{V}, \mathcal{F})$  se existirem escalares  $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathcal{F}$  tais que:

$$\begin{aligned} \mathbf{x} &= \sum_{i=1}^n \alpha_i \mathbf{x}_i = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_n \mathbf{x}_n = \\ &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} \end{aligned}$$

**Definição 2.5** Um conjunto de vetores  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  em um espaço vetorial  $(\mathcal{V}, \mathcal{F})$  é chamado **linearmente dependente** sobre o corpo  $\mathcal{F}$  se e somente se existirem escalares  $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathcal{F}$ , não todos nulos tais que:

$$\sum_{i=1}^n \alpha_i \mathbf{x}_i = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \alpha = 0 \quad (2.1)$$

onde,

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix}.$$

Agora se a equação (2.1) for válida apenas para o conjunto de valores  $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$  ou seja  $\alpha = 0$  então o conjunto de vetores  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  é chamado **linearmente independente** em  $(\mathcal{V}, \mathcal{F})$ .

**Definição 2.6** Um conjunto de vetores **linearmente independentes** de um espaço vetorial  $(\mathcal{V}, \mathcal{F})$  é chamado **base** de  $\mathcal{V}$  se todo vetor em  $\mathcal{V}$  puder ser expresso como uma única combinação linear destes vetores.

Desta definição concluímos:

- A base de um espaço vetorial é um conjunto de vetores que possuem duas propriedades:
  - a - Os vetores são linearmente independentes.
  - b - Os vetores representam um espaço.

**Definição 2.7** *A dimensão de um espaço vetorial é igual ao número de vetores linearmente independentes nele contido.*

Um espaço vetorial terá dimensão finita se tiver uma base finita. Por exemplo, no espaço vetorial real  $n$ -dimensional  $(\mathcal{R}^n, \mathcal{R})$  há no máximo  $n$  vetores linearmente independentes sobre  $\mathcal{R}$ .

**Teorema 2.1** *Se  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  é uma base de um espaço vetorial  $n$ -dimensional  $(\mathcal{V}, \mathcal{F})$ , então todo vetor  $\mathbf{x} \in \mathcal{V}$  pode ser escrito de uma única maneira na forma:*

$$\mathbf{x} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \dots + \beta_n \mathbf{x}_n = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \beta \quad (2.2)$$

onde  $\beta_1, \beta_2, \dots, \beta_n \in \mathcal{F}$

**Prova:** Por contradição suponhamos que exista uma outra combinação linear dos vetores da base tal que:

$$\mathbf{x} = \alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_n \mathbf{x}_n \quad (2.3)$$

então poderemos escrever das equações (2.2) e (2.3):

$$(\beta_1 - \alpha_1) \mathbf{x}_1 + (\beta_2 - \alpha_2) \mathbf{x}_2 + \dots + (\beta_n - \alpha_n) \mathbf{x}_n = 0$$

como os vetores de uma base são linearmente independentes isto implica que:  $\beta_i = \alpha_i \forall i$ .

**C.Q.D.**

### 2.2.1 Representação de Um Vetor em Relação a Uma Base e Mudança de base

Escolhida uma base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ , que supomos ordenada, para um espaço vetorial  $n$ -dimensional, então qualquer vetor  $\mathbf{x}$  pode ser representado de maneira única na forma:

$$\mathbf{x} = \sum_{i=1}^n \beta_i \mathbf{x}_i = \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_n \end{bmatrix} \beta = \mathbf{R} \beta \quad (2.4)$$

onde

$$\beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{bmatrix}$$

é chamada representação do vetor  $\mathbf{x}$  em relação à base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  e  $\beta_i$  é a  $i$ -ésima coordenada do vetor  $\mathbf{x}$  em relação a tal base e onde  $\mathbf{R} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]$  é uma matriz  $n \times n$  tendo como colunas os vetores da base.

Para outra base  $\{\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \bar{\mathbf{x}}_3, \dots, \bar{\mathbf{x}}_n\}$ , seja  $\bar{\beta}$  a representação do vetor  $\mathbf{x}$ ,

$$\mathbf{x} = \sum_{i=1}^n \bar{\beta}_i \bar{\mathbf{x}}_i = [\bar{\mathbf{x}}_1 \ \bar{\mathbf{x}}_2 \ \dots \ \bar{\mathbf{x}}_n] \bar{\beta} = \bar{\mathbf{R}} \bar{\beta}$$

onde,

$$\bar{\mathbf{R}} = [\bar{\mathbf{x}}_1 \ \bar{\mathbf{x}}_2 \ \dots \ \bar{\mathbf{x}}_n] \quad (2.5)$$

é uma matriz  $n \times n$ .

Conhecendo a representação de cada vetor de uma base em relação à outra base, podemos passar de uma representação em relação a uma base para a representação em relação a outra base de  $(\mathcal{V}, \mathcal{F})$ . Observe que  $\mathbf{x}_i$  é um vetor da equação (2.4); ou seja  $\mathbf{x}_i$  é um vetor da base  $\mathbf{R}$ , e  $\mathbf{s}_i$  a sua representação com relação a  $\{\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n\}$  então:

$$\mathbf{x}_i = [\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n] \mathbf{s}_i = \bar{\mathbf{R}} \mathbf{s}_i$$

Portanto podemos escrever:

$$\mathbf{R} = \bar{\mathbf{R}} \mathbf{S} \quad (2.6)$$

onde  $\mathbf{S} = [\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_n]_{n \times n}$ , tem como  $i$ -ésima coluna a representação de  $\mathbf{x}_i$  com relação à base  $\{\bar{\mathbf{x}}_1, \bar{\mathbf{x}}_2, \dots, \bar{\mathbf{x}}_n\}$ . Das equações (2.4) e (2.6) tem-se:

$$\mathbf{x} = \bar{\mathbf{R}} \mathbf{S} \beta = \bar{\mathbf{R}} \bar{\beta}$$

onde:

$$\bar{\beta} = \mathbf{S} \beta$$

Se por outro lado representamos  $\bar{\mathbf{x}}_i$  em relação a base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ :

$$\bar{\mathbf{x}}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \mathbf{t}_i$$

obtemos:

$$\beta = \mathbf{T}\bar{\beta}$$

onde  $\mathbf{T}$  tem como  $i$ -ésima coluna  $\mathbf{t}_i$  a representação de  $\bar{\mathbf{x}}_i$  com relação a base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$ :

$$\bar{\mathbf{x}}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n]\mathbf{t}_i$$

Como

$$\bar{\beta} = \mathbf{S}\beta$$

e

$$\beta = \mathbf{T}\bar{\beta}$$

então,

$$\mathbf{T} = \mathbf{S}^{-1}.$$

### 2.2.2 Operadores Lineares

Uma transformação ou função  $f$  de  $\mathcal{X}$  em  $\mathcal{Y}$ , que representamos por  $f : \mathcal{X} \rightarrow \mathcal{Y}$  é constituída por:

- a) Um conjunto  $\mathcal{X}$  denominado domínio da função.
- b) Um conjunto  $\mathcal{Y}$  denominado contradomínio da função.
- c) Uma regra ou correspondência  $f$ , que associa a cada elemento  $\mathbf{x}$  em  $\mathcal{X}$  um único elemento  $f(\mathbf{x})$  em  $\mathcal{Y}$ .

Se  $f$  é uma função de  $\mathcal{X}$  em  $\mathcal{Y}$ , a imagem de  $f$  é o conjunto dos  $f(\mathbf{x})$  com  $\mathbf{x} \in \mathcal{X}$ , ou seja, a imagem de  $f$  é constituída pelos elementos  $\mathbf{y}$  em  $\mathcal{Y}$ , tais que  $\mathbf{y} = f(\mathbf{x})$  para algum  $\mathbf{x}$  em  $\mathcal{X}$ .

Sejam  $(\mathcal{X}, \mathcal{F})$  e  $(\mathcal{Y}, \mathcal{F})$  espaços vetoriais  $n$  e  $m$  dimensionais sobre o mesmo corpo. Uma função  $\mathbf{L}$  que mapeia o domínio  $(\mathcal{X}, \mathcal{F})$  no contradomínio  $(\mathcal{Y}, \mathcal{F})$  é um operador linear se e somente se:

$$\mathbf{L}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{L}\mathbf{x}_1 + \alpha_2 \mathbf{L}\mathbf{x}_2$$

para quaisquer vetores  $\mathbf{x}_1$  e  $\mathbf{x}_2$  em  $\mathcal{X}$  e quaisquer escalares  $\alpha_1$  e  $\alpha_2$  em  $\mathcal{F}$ .

A imagem  $\Re(\mathbf{L})$  de um operador linear  $\mathbf{L}$  é definida como o conjunto de todos os elementos  $\mathbf{y}$  de  $(\mathcal{Y}, \mathcal{F})$ , para os quais existe pelo menos um vetor  $\mathbf{x}$  em  $(\mathcal{X}, \mathcal{F})$  tal que  $\mathbf{y} = \mathbf{L}\mathbf{x}$ .

**Teorema 2.2** *A imagem  $\mathfrak{R}(\mathbf{L})$  de um operador linear  $\mathbf{L} : (\mathcal{X}, \mathcal{F}) \rightarrow (\mathcal{Y}, \mathcal{F})$  é um subespaço do contradomínio.*

**Prova:** Sejam  $\mathbf{y}_1$  e  $\mathbf{y}_2$  elementos de  $\mathfrak{R}(\mathbf{L})$ , então existem vetores  $\mathbf{x}_1$  e  $\mathbf{x}_2$  em  $(\mathcal{X}, \mathcal{F})$ , tais que  $\mathbf{y}_1 = \mathbf{L}\mathbf{x}_1$  e  $\mathbf{y}_2 = \mathbf{L}\mathbf{x}_2$ . Então, para quaisquer  $\alpha_1$  e  $\alpha_2$  em  $\mathcal{F}$  temos:

$$\mathbf{L}(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\mathbf{L}\mathbf{x}_1 + \alpha_2\mathbf{L}\mathbf{x}_2 = \alpha_1\mathbf{y}_1 + \alpha_2\mathbf{y}_2 \text{ em } \mathfrak{R}(\mathbf{L})$$

donde  $\mathfrak{R}(\mathbf{L})$  é um subespaço de  $(\mathcal{Y}, \mathcal{F})$ .

**C.Q.D.**

**Teorema 2.3** *O conjunto de todos os elementos  $\mathbf{x}$  de  $(\mathcal{X}, \mathcal{F})$  para o qual  $\mathbf{L}(\mathbf{x}) = 0$ , chamado de **espaço nulo** de  $\mathbf{L}$  e que representamos por  $\mathcal{N}(\mathbf{L})$  é um subespaço de  $\mathcal{X}$ .*

**Prova:** Sejam  $\mathbf{x}_1$  e  $\mathbf{x}_2$  em  $\mathcal{N}(\mathbf{L})$  e  $\alpha_1$  e  $\alpha_2$  em  $\mathcal{F}$ . Como  $\mathbf{L}\mathbf{x}_1 = 0$  e  $\mathbf{L}\mathbf{x}_2 = 0$  então:

$$\mathbf{L}(\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2) = \alpha_1\mathbf{L}\mathbf{x}_1 + \alpha_2\mathbf{L}\mathbf{x}_2 = \alpha_1 0 + \alpha_2 0 = 0$$

donde  $\alpha_1\mathbf{x}_1 + \alpha_2\mathbf{x}_2$  pertence a  $\mathcal{N}(\mathbf{L})$  sendo este, consequentemente, um subespaço de  $\mathcal{X}$ .

**C.Q.D.**

**Teorema 2.4** *Seja  $\mathbf{L}$  um operador linear  $\mathbf{L} : (\mathcal{X}, \mathcal{F}) \rightarrow (\mathcal{Y}, \mathcal{F})$  sendo  $\mathcal{X}$  de dimensão finita  $n$  e sejam  $r(\mathbf{L})$  a dimensão de  $\mathfrak{R}(\mathbf{L})$ , que chamaremos de posto de  $\mathbf{L}$  e  $\nu(\mathbf{L})$  a dimensão de  $\mathcal{N}(\mathbf{L})$ , que chamaremos de nulidade de  $\mathbf{L}$ <sup>1</sup>. Então,*

$$r(\mathbf{L}) + \nu(\mathbf{L}) = n$$

<sup>1</sup>Usaremos indistintamente as notações:

$$\text{posto}(\mathbf{L}) = r(\mathbf{L}) = r \quad \text{e} \quad \text{nulidade}(\mathbf{L}) = \nu(\mathbf{L}) = \nu$$

**Prova:** No espaço nulo escolhemos uma base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\nu\}$ . Para completar esta base para o espaço  $\mathcal{X}$ , basta incluir os vetores  $\{\mathbf{x}_{\nu+1}, \mathbf{x}_{\nu+2}, \dots, \mathbf{x}_n\}$ . Demonstramos que  $\{\mathbf{Lx}_{\nu+1}, \mathbf{Lx}_{\nu+2}, \dots, \mathbf{Lx}_n\}$  é uma base da imagem de  $\mathbf{L}$ .

- **Primeira Parte:**  $\{\mathbf{Lx}_{\nu+1}, \mathbf{Lx}_{\nu+2}, \dots, \mathbf{Lx}_n\}$  gera  $\mathfrak{R}(\mathbf{L})$ .

Como qualquer vetor  $\mathbf{x} \in \mathcal{X}$  pode ser escrito na forma:

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i$$

então

$$\mathbf{Lx} = \mathbf{L} \left( \sum_{i=1}^n \alpha_i \mathbf{x}_i \right) = \sum_{i=1}^n \alpha_i \mathbf{Lx}_i = \sum_{i=\nu+1}^n \alpha_i \mathbf{Lx}_i$$

pois  $\sum_{i=1}^{\nu} \alpha_i \mathbf{Lx}_i = 0$ , donde os vetores  $[\mathbf{Lx}_{\nu+1}, \mathbf{Lx}_{\nu+2}, \dots, \mathbf{Lx}_n]$  geram a imagem de  $\mathbf{L}$ .

- **Segunda Parte:** São linearmente independentes.

Suponhamos que existam escalares  $\alpha_j$ , tais que:

$$\sum_{j=\nu+1}^n \alpha_j \mathbf{x}_j = 0$$

donde

$$\mathbf{L} \left( \sum_{j=\nu+1}^n \alpha_j \mathbf{x}_j \right) = 0$$

e conseqüentemente o vetor  $\mathbf{x} = \sum_{j=\nu+1}^n \alpha_j \mathbf{x}_j$  está no espaço nulo de  $\mathbf{L}$ .

Como  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_\nu\}$  constitui uma base do espaço nulo, existem necessariamente escalares  $\beta$ , tais que:

$$\mathbf{x} = \sum_{j=1}^{\nu} \beta_j \mathbf{x}_j$$

portanto

$$\sum_{j=1}^{\nu} \beta_j \mathbf{x}_j - \sum_{j=\nu+1}^n \alpha_j \mathbf{x}_j = 0$$



e como  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  são linearmente independentes, devemos ter:

$$\beta_1 = \beta_2 = \dots = \beta_\nu = \alpha_{\nu+1} = \alpha_{\nu+2} = \dots = \alpha_n = 0$$

Portanto  $\{\mathbf{Lx}_{\nu+1}, \mathbf{Lx}_{\nu+2}, \dots, \mathbf{Lx}_n\}$  são linearmente independentes, donde constituem uma base para  $\mathfrak{R}(\mathbf{L})$ , cuja dimensão é  $r = n - \nu$ .

**C.Q.D.**

### 2.2.3 Soma Direta de Subespaços

O espaço linear  $\mathcal{X}$  será chamado soma de dois subespaços  $\mathcal{Y}_1$  e  $\mathcal{Y}_2$ , e representado por  $\mathcal{X} = \mathcal{Y}_1 + \mathcal{Y}_2$ , se todo vetor  $\mathbf{x}$  puder ser escrito como a soma de um vetor em  $\mathcal{Y}_1$  e de um vetor em  $\mathcal{Y}_2$ .

Se todo vetor  $\mathbf{x}$  em  $\mathcal{X}$  puder ser escrito de uma única maneira como a soma de um vetor  $\mathbf{x}_1$  em  $\mathcal{Y}_1$  e de um vetor  $\mathbf{x}_2$  em  $\mathcal{Y}_2$ , sendo  $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ , então o espaço vetorial  $\mathcal{X}$  será chamado a soma direta dos subespaços  $\mathcal{Y}_1$  e  $\mathcal{Y}_2$ , e representado por:

$$\mathcal{X} = \mathcal{Y}_1 \oplus \mathcal{Y}_2.$$

Por exemplo, o espaço vetorial  $\mathcal{R}^2$  é a soma direta de duas retas linearmente independentes que se cruzam na sua origem.

**Teorema 2.5** *O espaço vetorial  $\mathcal{X}$  é a soma direta de seus subespaços  $\mathcal{Y}_1$  e  $\mathcal{Y}_2$  se e somente se:*

- a)  $\mathcal{Y}_1$  e  $\mathcal{Y}_2$  têm em comum o vetor 0, ou seja,  $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \{0\}$ .
- b)  $\mathcal{X} = \mathcal{Y}_1 + \mathcal{Y}_2$

**Prova:**

- **Necessidade:** Seja  $\mathbf{x} \in \mathcal{X}$ . Se  $\mathcal{X} = \mathcal{Y}_1 + \mathcal{Y}_2$ , existem  $\mathbf{x}_1 \in \mathcal{Y}_1$  e  $\mathbf{x}_2 \in \mathcal{Y}_2$ , tais que  $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ . Mostremos que esta soma é única. Para isto suponhamos que  $\mathbf{x} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2$  onde  $\bar{\mathbf{x}}_1 \in \mathcal{Y}_1$  e  $\bar{\mathbf{x}}_2 \in \mathcal{Y}_2$ . Então:

$$\mathbf{x}_1 + \mathbf{x}_2 = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2$$

$$\mathbf{x}_1 - \bar{\mathbf{x}}_1 = \bar{\mathbf{x}}_2 - \mathbf{x}_2$$

mas,  $\mathbf{x}_1 - \bar{\mathbf{x}}_1 \in \mathcal{Y}_1$  e  $\bar{\mathbf{x}}_2 - \mathbf{x}_2 \in \mathcal{Y}_2$ , donde

$$\mathcal{Y}_1 \cap \mathcal{Y}_2 = \{0\}$$

Portanto:

$$\mathbf{x}_1 - \bar{\mathbf{x}}_1 = 0$$

$$\mathbf{x}_2 - \bar{\mathbf{x}}_2 = 0$$

Então  $\mathbf{x}_1 = \bar{\mathbf{x}}_1$  e  $\mathbf{x}_2 = \bar{\mathbf{x}}_2$ .

- **Suficiência:** Se  $\mathcal{X} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$ , então qualquer  $\mathbf{x} \in \mathcal{X}$  pode ser escrito de maneira única como  $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$ , onde  $\mathbf{x}_1 \in \mathcal{Y}_1$  e  $\mathbf{x}_2 \in \mathcal{Y}_2$ . Portanto, podemos também escrever  $\mathcal{X} = \mathcal{Y}_1 + \mathcal{Y}_2$ . Se agora supusermos que  $\mathcal{X} \in \mathcal{Y}_1 \cap \mathcal{Y}_2$ , teremos:

a)  $\mathbf{x} = \mathbf{x} + 0$  onde  $\mathbf{x} \in \mathcal{Y}_1$  e  $0 \in \mathcal{Y}_2$ .

b)  $\mathbf{x} = 0 + \mathbf{x}$  onde  $0 \in \mathcal{Y}_1$  e  $\mathbf{x} \in \mathcal{Y}_2$ .

Como a representação da soma deve ser única para  $\mathbf{x}$ ,  $\mathbf{x} = 0$ , portanto  $\mathcal{Y}_1 \cap \mathcal{Y}_2 = \{0\}$ .

**C.Q.D.**

## 2.2.4 Subespaço Invariante

Seja uma transformação linear  $\mathbf{L} : (\mathcal{X}, \mathcal{F}) \rightarrow (\mathcal{Y}, \mathcal{F})$ . Um subespaço  $\mathcal{Y}$  de  $\mathcal{X}$  é chamado invariante sob  $\mathbf{L}$  se para todo  $\mathbf{y}$  em  $\mathcal{Y}$  o vetor  $\mathbf{L}\mathbf{y}$  está em  $\mathcal{Y}$ , isto é, se  $\mathbf{L}(\mathbf{y})$  está contido em  $\mathcal{Y}$ .

### Isomorfismo de Espaços Vetoriais

Se a imagem de  $f$  é todo o conjunto  $\mathcal{Y}$ ,  $f$  é chamada função sobrejetiva de  $\mathcal{X}$  em  $\mathcal{Y}$ . Por outro lado, se:

$$f(\mathbf{x}) = f(\bar{\mathbf{x}})$$

com  $\mathbf{x}$  e  $\bar{\mathbf{x}}$  em  $\mathcal{X}$  implica que  $\mathbf{x} = \bar{\mathbf{x}}$ , a função é chamada injetiva.

Se  $\mathcal{X}$  e  $\mathcal{Y}$  são dois espaços vetoriais sobre o mesmo corpo  $\mathcal{F}$ , uma transformação linear bijetiva (injetiva e sobrejetiva)  $\mathbf{S}$  de  $\mathcal{X}$  em  $\mathcal{Y}$  é denominada isomorfismo de  $\mathcal{X}$  em  $\mathcal{Y}$ , se houver uma correspondência biunívoca entre os vetores  $\mathbf{x}$  de  $\mathcal{X}$  e  $\mathbf{v}$  de  $\mathcal{Y}$ , digamos  $\mathbf{v} = \mathbf{S}(\mathbf{x})$ , tal que

$$\mathbf{S}(\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2) = \alpha_1 \mathbf{S}(\mathbf{x}_1) + \alpha_2 \mathbf{S}(\mathbf{x}_2).$$

Um isomorfismo  $\mathbf{S} : \mathcal{X} \rightarrow \mathcal{V}$  transforma toda base de  $\mathcal{X}$  em base de  $\mathcal{V}$ .

Se existir um isomorfismo de  $\mathcal{X}$  em  $\mathcal{V}$  diremos que  $\mathcal{X}$  é isomorfo a  $\mathcal{V}$ , ou que os espaços vetoriais  $\mathcal{X}$  e  $\mathcal{V}$  são isomorfos.

**Teorema 2.6** *Todo o espaço vetorial  $n$ -dimensional  $(\mathcal{X}, \mathcal{F})$  é isomorfo a  $\mathcal{F}^n$ .*

**Prova:** Seja  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  uma base de  $\mathcal{X}$ , então cada  $\mathbf{x}$  em  $\mathcal{X}$ , pode ser escrito na forma  $\alpha_1 \mathbf{x}_1 + \alpha_2 \mathbf{x}_2 + \dots + \alpha_n \mathbf{x}_n$ , e sabemos que  $\mathbf{x}$  determina os  $n$  escalares  $\alpha_1, \alpha_2, \dots, \alpha_n$  de forma única.

Consideremos a correspondência um a um:

$$\mathbf{x} \rightleftharpoons (\alpha_1, \alpha_2, \dots, \alpha_n) \text{ entre } \mathcal{X} \text{ e } \mathcal{F}^n$$

Se  $\bar{\mathbf{x}} = \beta_1 \mathbf{x}_1 + \beta_2 \mathbf{x}_2 + \dots + \beta_n \mathbf{x}_n$ , então para  $\gamma$  e  $\delta$  em  $\mathcal{F}$ :

$$\gamma \mathbf{x} + \delta \bar{\mathbf{x}} = (\gamma \alpha_1 + \delta \beta_1) \mathbf{x}_1 + \dots + (\gamma \alpha_n + \delta \beta_n) \mathbf{x}_n$$

Com isto estabelecemos o isomorfismo desejado e temos a correspondência biunívoca:

$$\gamma \mathbf{x} + \delta \bar{\mathbf{x}} \rightleftharpoons \gamma(\alpha_1, \alpha_2, \dots, \alpha_n) + \delta(\beta_1, \beta_2, \dots, \beta_n).$$

**C.Q.D.**

### 2.2.5 Representação de Um Operador Linear Por Uma Matriz

**Teorema 2.7** *Se  $(\mathcal{X}, \mathcal{F})$  e  $(\mathcal{Y}, \mathcal{F})$  são espaços vetoriais  $n$  e  $m$  dimensionais respectivamente e se  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  é um conjunto de vetores linearmente independentes em  $\mathcal{X}$ , o operador linear  $\mathbf{L}$ , cujo domínio é  $\mathcal{X}$  e cujo contradomínio é  $\mathcal{Y}$ , é determinado de forma única pelos  $n$  pares de correspondências  $\mathbf{y}_i = \mathbf{L}\mathbf{x}_i$ ,  $i = 1, 2, \dots, n$ .*

**Prova:** Seja  $\mathbf{x}$  em  $\mathcal{X}$ . Como  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  são linearmente independentes,

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \alpha$$

onde  $\alpha = [\alpha_1 \ \alpha_2 \ \alpha_3 \ \dots \ \alpha_n]^T$  é única.

Se  $\mathbf{L}$  é um operador linear então:

$$\mathbf{Lx} = \alpha_1 \mathbf{Lx}_1 + \alpha_2 \mathbf{Lx}_2 + \alpha_3 \mathbf{Lx}_3 + \dots + \alpha_n \mathbf{Lx}_n = \sum_{i=1}^n \alpha_i \mathbf{Lx}_i = \sum_{i=1}^n \alpha_i \mathbf{y}_i$$

o que implica que, para qualquer  $\mathbf{x}$  em  $\mathcal{X}$ ,  $\mathbf{Lx}$  é determinado de forma única pelos pares:

$$\mathbf{y}_i = \mathbf{Lx}_i, \ i = 1, 2, \dots, n.$$

C.Q.D.

**Teorema 2.8** *Sejam  $(\mathcal{X}, \mathcal{F})$  e  $(\mathcal{Y}, \mathcal{F})$  espaços vetoriais com bases  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  e  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ , respectivamente. O operador linear  $\mathbf{L} : (\mathcal{X}, \mathcal{F}) \rightarrow (\mathcal{Y}, \mathcal{F})$  pode ser representado por uma matriz  $\mathbf{M}_{m \times n}$  com elementos no corpo  $\mathcal{F}$ , cuja  $i$ -ésima coluna constitui-se das componentes de  $\mathbf{Lx}_i$ , relativamente à base  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ .*

**Prova:** Para  $\mathbf{x}$  em  $\mathcal{X}$  temos:

$$\mathbf{x} = \sum_{i=1}^n \alpha_i \mathbf{x}_i = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n] \alpha$$

Portanto

$$\mathbf{Lx} = \sum_{i=1}^n \alpha_i \mathbf{Lx}_i$$

Como  $\mathbf{Lx}_i$  é um vetor em  $\mathcal{Y}$ , sua representação relativa à base  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$  é dada por:

$$\mathbf{Lx}_i = \sum_{j=1}^m m_{ji} \mathbf{y}_j = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_m \end{bmatrix} \begin{bmatrix} m_{1i} \\ m_{2i} \\ m_{3i} \\ \vdots \\ m_{mi} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_m \end{bmatrix} m_i$$

para  $i = 1, 2, 3, \dots, n$  e  $m_{ij} \in \mathcal{F}$ .

Seja  $\beta$  a representação de  $\mathbf{y}$  relativa à base  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m\}$ ,

$$\mathbf{y} = \sum_{j=1}^m \beta_j \mathbf{y}_j = [\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_m] \beta$$

Como

$$\mathbf{y} = \mathbf{L}\mathbf{x} = \mathbf{L} \sum_{i=1}^n \alpha_i \mathbf{x}_i = \sum_{i=1}^n \alpha_i \mathbf{L}\mathbf{x}_i = \sum_{i=1}^n \alpha_i \sum_{j=1}^m m_{ji} \mathbf{y}_j$$

Portanto,

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_m \end{bmatrix} \sum_{i=1}^n \alpha_i m_i = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \dots & \mathbf{y}_m \end{bmatrix} \mathbf{M}\alpha$$

onde  $\mathbf{M} = [m_1, m_2, \dots, m_n]$ . Então temos que  $\beta = \mathbf{M}\alpha$ .

Após a escolha das bases não há diferenças entre  $\mathbf{y}$  e  $\mathbf{x}$  e suas respectivas representações  $\alpha$  e  $\beta$ , donde no estudo de  $\mathbf{y} = \mathbf{L}\mathbf{x}$  podemos estudar simplesmente as relações entre  $\alpha$  e  $\beta$ .

Observemos que os elementos de  $\mathbf{M}$  estão no corpo  $\mathcal{F}$  e que a  $i$ -ésima coluna de  $\mathbf{M}$  é a representação de  $\mathbf{L}\mathbf{x}_i$  em relação a base de  $\mathcal{Y}$ .

Portanto, se as bases de  $(\mathcal{X}, \mathcal{F})$  e  $(\mathcal{Y}, \mathcal{F})$  são escolhidas, o operador  $\mathbf{L}$  pode ser representado por uma matriz de coeficientes em  $\mathcal{F}$ .

A equação  $\beta = \mathbf{M}\alpha$  evidencia que a matriz  $\mathbf{M}$  relaciona as representações  $\alpha$  e  $\beta$  e não os vetores  $\mathbf{y}$  e  $\mathbf{x}$ , bem como evidencia que  $\mathbf{M}$  depende das bases escolhidas.

**C.Q.D.**

### 2.2.6 Exemplos de Operadores Lineares

A seguir, apresentamos exemplos de operadores lineares que irão produzir rotação, permutação e reflexão em matrizes.

**Exemplo 2.1** *Se dois vetores  $\mathbf{x}$  e  $\mathbf{y}$  satisfazem à relação:*

$$\mathbf{y} = \mathbf{L}\mathbf{x} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix} \mathbf{x}$$

*$\mathbf{y}$  é o vetor  $\mathbf{x}$  girado de um ângulo  $\alpha$  e  $\mathbf{L}$  é um operador rotação e a matriz que representa este operador é a matriz de rotação:*

$$\mathbf{Q} = \begin{bmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{bmatrix}$$

*Para esta matriz temos:*

$$\mathbf{Q}^T = \mathbf{Q}^{-1} = \begin{bmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{bmatrix}$$

A matriz  $\mathbf{Q}^{-1}$  é igual a  $\mathbf{Q}^T$  pois  $\cos(-\alpha) = \cos \alpha$  e  $\sin(-\alpha) = -\sin \alpha$ ; assim temos:

$$\mathbf{Q}\mathbf{Q}^T = \mathbf{I} \quad e \quad \mathbf{Q}^T\mathbf{Q} = \mathbf{I}.$$

As colunas de  $\mathbf{Q}$  são vetores ortogonais unitários pois:  $\sin^2 \alpha + \cos^2 \alpha = 1$ . Estes vetores coluna de  $\mathbf{Q}$  fornecem uma base ortonormal dada por:

$$\mathbf{i} = [1, 0]^T \quad e \quad \mathbf{j} = [0, 1]^T$$

pois:

$$\mathbf{Q}\mathbf{i} = [\cos \alpha, \sin \alpha]^T \quad e \quad \mathbf{Q}\mathbf{j} = [-\sin \alpha, \cos \alpha]^T.$$

Na Figura 2.1 vemos que os vetores  $\mathbf{i}, \mathbf{j}$  são girados de um ângulo  $\alpha$ .

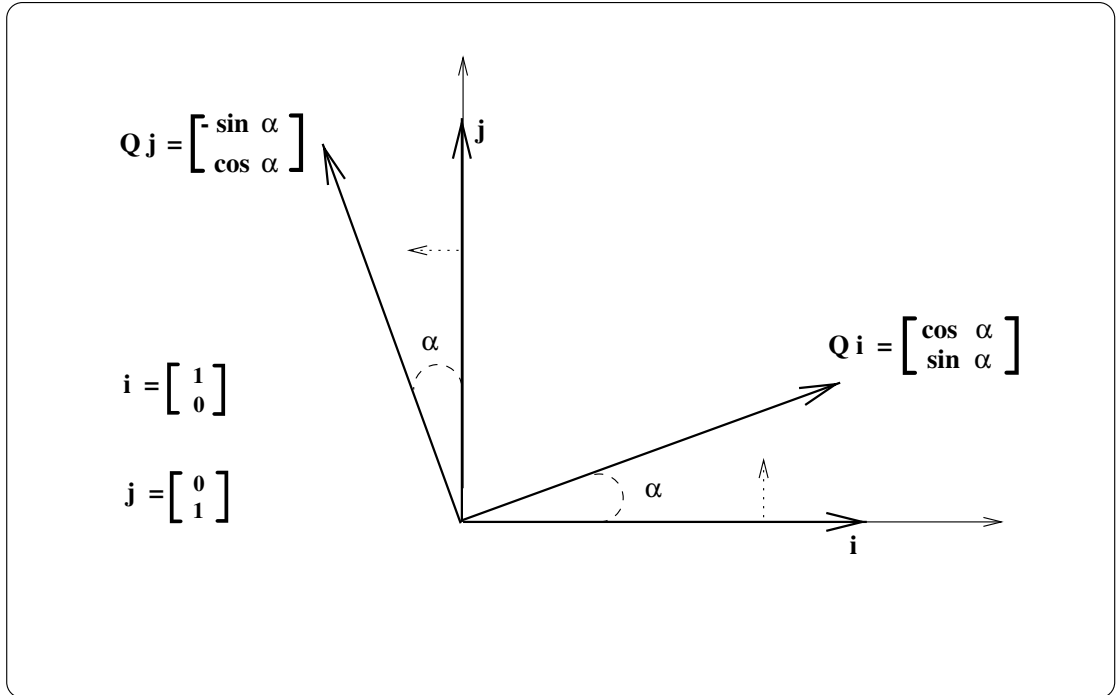


Figura 2.1: Matriz de Rotação.

Observe que  $\mathbf{Q}^{-1}$  é uma matriz que produz uma rotação de  $-\alpha$ .

- Uma sequência de rotações no plano pode ser utilizada para introduzir vários zeros em posições especificadas de um vetor ou matriz.
- Se as colunas de uma matriz  $\mathbf{Q}$  são vetores unitários, ortogonais e a inversa desta é sua transposta, ela é chamada matriz de permutação:

$$\mathbf{Q}^T = \mathbf{Q}^{-1}$$

A seguir apresentamos exemplos com matrizes de permutação:

**Exemplo 2.2**

$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} y \\ z \\ x \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ z \\ x \end{bmatrix}$$

**Exemplo 2.3**

$$\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} \Rightarrow \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} y \\ x \end{bmatrix}$$

Neste exemplo em particular a matriz de permutação também é uma matriz de reflexão.

**Exemplo 2.4** Uma representação esquemática do mapa do Brasil no plano  $\mathcal{R}^2$ , pode ser obtida com uma sequência de vetores coluna,  $\mathbf{b}$ , de uma matriz  $\mathbf{B}$ ; em nossa representação utilizamos 66 vetores onde cada vetor coluna  $\mathbf{b}$ , em  $\mathcal{R}^2$  teve o seu ponto final conectado, em uma sequência crescente das colunas.

$$\mathbf{B} = \begin{bmatrix} \begin{matrix} -0.5000 & -8.5000 \\ 0.5000 & -7.5000 \\ 1.0000 & -6.5000 \\ 2.5000 & -5.5000 \\ 3.5000 & -5.5000 \\ 4.0000 & -5.0000 \\ 4.5000 & -4.0000 \\ 5.5000 & -3.0000 \\ 5.5000 & -2.0000 \\ 7.0000 & 0 \\ 7.5000 & 1.0000 \end{matrix} & \begin{matrix} \downarrow & \downarrow \\ 7.0000 & 1.5000 \\ 5.5000 & 2.5000 \\ 5.5000 & 3.0000 \\ 4.5000 & 3.0000 \\ 4.5000 & 2.5000 \\ 4.0000 & 3.0000 \\ 4.0000 & 3.5000 \\ 2.5000 & 4.0000 \\ 2.0000 & 4.0000 \\ 2.0000 & 3.5000 \\ 1.5000 & 3.5000 \end{matrix} & \begin{matrix} \downarrow & \downarrow \\ 1.5000 & 4.0000 \\ 2.0000 & 4.5000 \\ 1.5000 & 4.0000 \\ 1.5000 & 4.0000 \\ 1.0000 & 4.5000 \\ 1.0000 & 4.5000 \\ 1.5000 & 5.0000 \\ 1.0000 & 5.0000 \\ 1.0000 & 5.5000 \\ 0.5000 & 5.5000 \\ 0 & 5.0000 \end{matrix} & \begin{matrix} \downarrow & \downarrow \\ -0.5000 & 5.5000 \\ -2.0000 & 5.0000 \\ -2.5000 & 5.5000 \\ -2.5000 & 6.5000 \\ -3.0000 & 6.5000 \\ -4.0000 & 6.0000 \\ -4.5000 & 6.5000 \\ -4.5000 & 6.0000 \\ -4.5000 & 5.5000 \\ -5.5000 & 5.0000 \\ -6.0000 & 5.5000 \end{matrix} & \begin{matrix} \downarrow & \downarrow \\ -7.0000 & 5.5000 \\ -6.5000 & 5.0000 \\ -6.5000 & 4.5000 \\ -7.0000 & 4.0000 \\ -7.0000 & 3.0000 \\ -8.0000 & 3.0000 \\ -8.5000 & 1.5000 \\ -7.5000 & 1.0000 \\ -7.0000 & 0 \\ -5.5000 & 0.5000 \\ -5.5000 & -0.5000 \end{matrix} & \begin{matrix} \downarrow & \downarrow \\ -0.5000 & -8.5000 \\ -0.5000 & -9.0000 \\ -1.5000 & -9.5000 \\ -1.0000 & -9.0000 \\ -2.5000 & -8.0000 \\ -1.0000 & -6.0000 \\ -2.5000 & -5.5000 \\ -2.0000 & -4.0000 \\ -3.5000 & -2.5000 \\ -3.5000 & -1.5000 \\ -4.5000 & -1.0000 \end{matrix} \end{bmatrix}$$

Note na Figura 2.2 as diferentes projeções do mapa do Brasil obtidas a partir da matriz original  $\mathbf{B}$  através de transformações do tipo  $\mathbf{LB}$ . As matrizes  $\mathbf{L}$  responsáveis pelas seis transformações apresentadas, na Figura 2.2, são as seguintes:

–  $I$

$$\mathbf{L} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

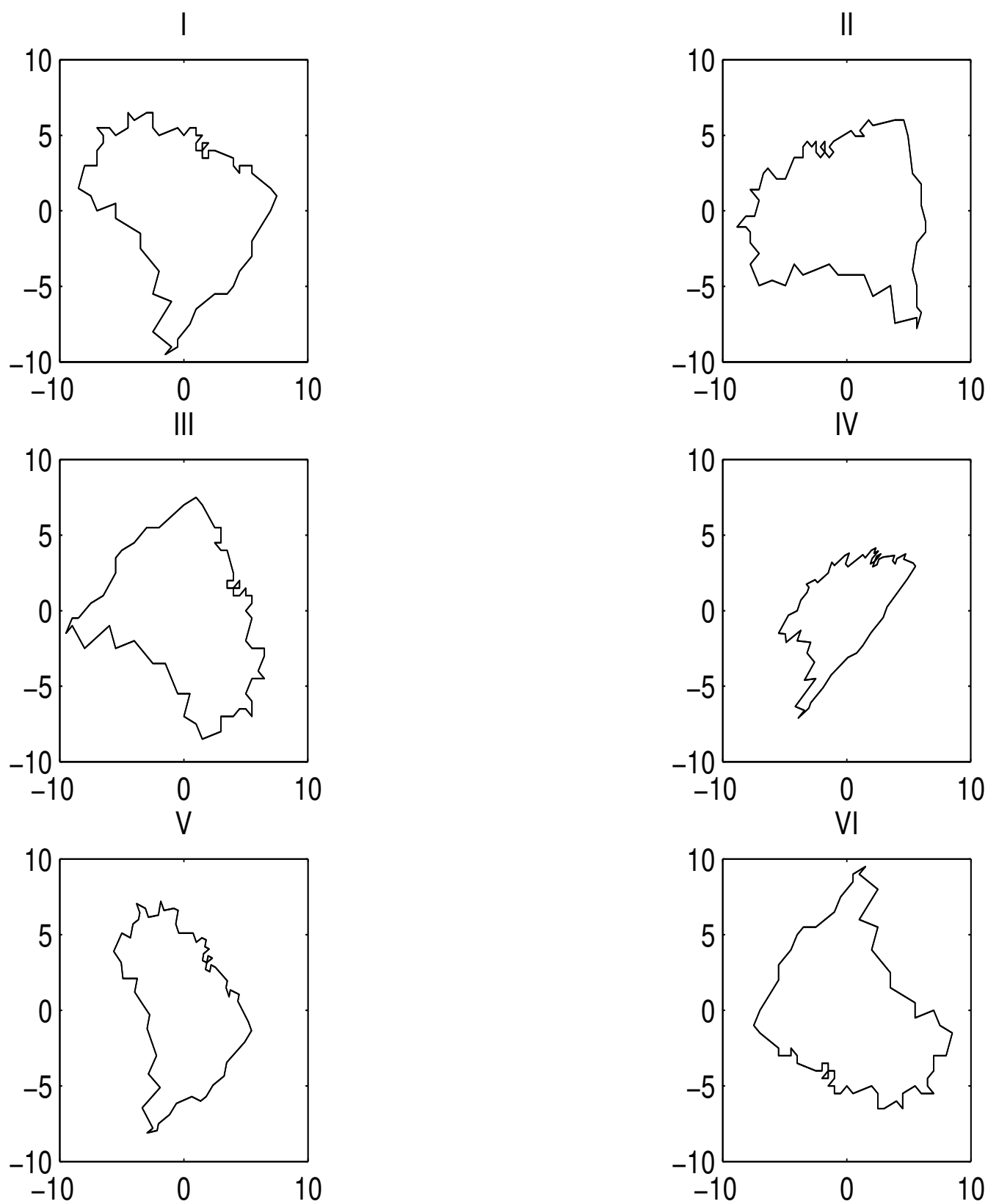


Figura 2.2: Transformações realizadas por operadores lineares.



– II

$$\mathbf{L} = \begin{bmatrix} \cos(\pi/4) & -\sin(\pi/4) \\ \sin(\pi/4) & \cos(\pi/4) \end{bmatrix}$$

– III

$$\mathbf{L} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

– IV

$$\mathbf{L} = \begin{bmatrix} 0.7 & 0.3 \\ 0.3 & 0.7 \end{bmatrix}$$

– V

$$\mathbf{L} = \begin{bmatrix} 0.7 & 0.2 \\ -0.3 & 0.9 \end{bmatrix}$$

– VI

$$\mathbf{L} = \begin{bmatrix} \cos(\pi) & -\sin(\pi) \\ \sin(\pi) & \cos(\pi) \end{bmatrix}$$

*Este exemplo não deve ser visto como um simples recurso visual gráfico mas como a possibilidade de obter uma forma diferenciada para solucionar o mesmo problema. Apesar do contorno do Brasil não ter se alterado em nenhuma das situações as transformações lineares no entanto criam situações aonde a busca de solução para um determinado tipo de problema pode ser obtida de forma diferenciada. Por exemplo no cálculo do perímetro as situações I e VI apresentariam o mesmo resultado, o que não iria ocorrer nas outras situações.*

## 2.2.7 Particularidade de Um Operador Linear

Consideremos o caso particular e muito importante dos operadores lineares que mapeiam um espaço linear  $(\mathcal{X}, \mathcal{F})$  n-dimensional em si mesmo. Neste caso podemos obter dois resultados muito úteis para o nosso trabalho:

- **Primeiro Resultado:** Dada uma matriz quadrada  $\mathbf{M}_{n \times n}$ , sem qualquer especificação de base, é evidente que podemos interpretar  $\mathbf{M}$  como sendo uma transformação linear que leva o i-ésimo vetor na base padrão para o vetor representado na base padrão pela i-ésima coluna de  $\mathbf{M}$ , ou, dito de outra forma, para a base padrão de  $(\mathcal{X}, \mathcal{F})$ , a representação do operador  $\mathbf{L} \equiv \mathbf{M}$  é a própria matriz  $\mathbf{M}$ .

Logo, se  $m_i$  é a i-ésima coluna de  $\mathbf{M}$ , a transformação linear  $\mathbf{y} = \mathbf{M}\mathbf{x}$  com domínio  $\mathcal{X}$ , contradomínio  $\mathcal{Y}$  e imagem  $\mathfrak{R}(\mathbf{M})$ , onde esta é por definição o conjunto dos elementos  $\mathbf{y}$  para os quais existe pelo menos um  $\mathbf{x}$  tal que  $\mathbf{y} = \mathbf{M}\mathbf{x}$ , pode ser escrita como:

$$\mathbf{y} = \mathbf{x}_1 m_1 + \mathbf{x}_2 m_2 + \dots + \mathbf{x}_n m_n$$

onde  $\mathbf{x}_i, i = 1, 2, 3, \dots, n$  são as componentes de  $\mathbf{x}$ , e elementos de  $\mathcal{F}$ , então  $\mathfrak{R}(\mathbf{M})$  é o conjunto dos elementos  $\mathbf{y}$  com  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$  tomando todos os valores possíveis em  $\mathcal{F}$ , o que nos permite concluir que:  $\mathfrak{R}(\mathbf{M})$  é o conjunto de todas as combinações lineares possíveis das colunas de  $\mathbf{M}$  e, conseqüentemente, que a dimensão de  $\mathfrak{R}(\mathbf{M})$  é o número máximo de colunas linearmente independentes da matriz  $\mathbf{M}$ , o que na teoria de matrizes é definida como posto da matriz  $\mathbf{M}$ .

Como concluímos anteriormente que:

$$r(\mathbf{M}) + \nu(\mathbf{M}) = n$$

a dimensão de  $\nu(\mathbf{M})$  é igual ao número de colunas linearmente dependentes da matriz  $\mathbf{M}$ .

- **Segundo Resultado:** Se  $\mathbf{M}$  é a representação de um operador linear em relação a uma base  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  e  $\overline{\mathbf{M}}$  é a representação do mesmo operador em relação  $\{\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \dots, \overline{\mathbf{x}}_n\}$  qual a relação entre as representações  $\mathbf{M}$  e  $\overline{\mathbf{M}}$ ?

Sejam  $\alpha$  e  $\overline{\alpha}$  as representações de  $\mathbf{x}$  em  $\mathcal{X}$  em relação às bases  $\{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n\}$  e  $\{\overline{\mathbf{x}}_1, \overline{\mathbf{x}}_2, \dots, \overline{\mathbf{x}}_n\}$ , respectivamente. Como o vetor  $\mathbf{y} = \mathbf{Lx}$  também está no espaço vetorial, sejam  $\beta$  e  $\overline{\beta}$  suas representações em relação àquelas mesmas respectivas bases, que desejamos determinar.

Como vimos anteriormente  $\overline{\alpha} = \mathbf{S}\alpha$ ,  $\overline{\beta} = \mathbf{S}\beta$  e  $\mathbf{S} = \mathbf{T}^{-1}$ , bem como  $\beta = \mathbf{M}\alpha$  e  $\overline{\beta} = \overline{\mathbf{M}}\overline{\alpha}$ , então  $\overline{\beta} = \mathbf{S}\beta = \mathbf{SM}\alpha = \mathbf{SMS}^{-1}\overline{\alpha}$  donde, pela unicidade da representação em relação a uma base, temos:

$$\overline{\mathbf{M}}\overline{\alpha} = \mathbf{SMS}^{-1}\overline{\alpha}$$

para todo  $\overline{\alpha}$ .

Portanto

$$\overline{\mathbf{M}} = \mathbf{SMS}^{-1} = \mathbf{T}^{-1}\mathbf{MT} \quad (2.7)$$

Se existe uma matriz quadrada, não-singular,  $\mathbf{T}$  sobre  $\mathcal{F}$  dizemos que  $\mathbf{M}$  é similar à  $\overline{\mathbf{M}}$  e que ela é obtida por meio de uma transformação de similaridade, equação (2.7).

Como vimos, a transformação de similaridade descreve o efeito de uma mudança de base em subespaços invariantes. Portanto todas as representações de um mesmo operador, relativas a diferentes bases, são similares.

### 2.2.8 Distância, Norma e Produto Interno em Um Espaço Vetorial

**Definição 2.8** Uma métrica  $d(\mathbf{x}, \mathbf{y})$  ou distância entre dois vetores  $\mathbf{x}$  e  $\mathbf{y} \in \mathcal{X}$ , é qualquer função escalar real com as propriedades:

- a) A distância é sempre não negativa:  $d(\mathbf{x}, \mathbf{y}) \geq 0$ .
- b) A distância é nula se e somente se os pontos coincidem:  $d(\mathbf{x}, \mathbf{y}) = 0$  se e somente  $\mathbf{x} \equiv \mathbf{y}$ .
- c) A distância de  $\mathbf{x}$  a  $\mathbf{y}$  é idêntica à distância de  $\mathbf{y}$  a  $\mathbf{x}$ :

$$d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x}).$$

- d) Desigualdade do triângulo:

$$d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z}) \equiv d(\mathbf{x}, \mathbf{z}) \quad \forall \quad \mathbf{x}, \mathbf{y} \text{ e } \mathbf{z} \in \mathcal{X}$$

A partir deste ponto, neste Capítulo, nos restringiremos ao espaço complexo  $\mathcal{X} = (\mathcal{C}^n, \mathcal{C})$ .

Introduzindo a idéia de origem (fazendo  $\mathbf{y} = 0$ ) e fazendo a métrica ter um comprimento definido, independentemente de  $\mathbf{x}$ , definimos norma ou comprimento de um vetor  $\mathbf{x} \in (\mathcal{C}^n, \mathcal{C})$  representada por  $\|\mathbf{x}\|$  a seguir:

**Definição 2.9** Qualquer função escalar de valor real de um vetor  $\mathbf{x}$  pode ser definida como uma norma de um vetor se tiver as seguintes propriedades para qualquer  $a \in \mathcal{C}$  e  $\mathbf{x} \in (\mathcal{C}^n, \mathcal{C})$ .

- a)  $\|\mathbf{x}\| \geq 0$ .
- b)  $\|\mathbf{x}\| = 0$  se e somente se  $\mathbf{x} = 0$ .
- c)  $\|\mathbf{x}_1 + \mathbf{x}_2\| \leq \|\mathbf{x}_1\| + \|\mathbf{x}_2\|$
- d)  $\|a\mathbf{x}\| = |a| \|\mathbf{x}\|$  onde  $|a|$  é o valor absoluto de  $a$ .

Uma norma de um vetor  $\mathbf{x}$  pode ser escolhida dentre as seguintes:

a)

$$\| \mathbf{x} \|_1 \stackrel{\text{def}}{=} \sum_{i=1}^n | \mathbf{x}_i |$$

b) Norma euclidiana

$$\| \mathbf{x} \|_2 \stackrel{\text{def}}{=} \left( \sum_{i=1}^n | \mathbf{x}_i |^2 \right)^{1/2}$$

c) Norma de Holder

$$\| \mathbf{x} \|_p \stackrel{\text{def}}{=} \left( \sum_{i=1}^n | \mathbf{x}_i |^p \right)^{1/p}, \quad 1 \leq p \leq \infty$$

d)

$$\| \mathbf{x} \|_\infty \stackrel{\text{def}}{=} \max_{1 \leq i \leq n} | \mathbf{x}_i |$$

Um espaço vetorial onde definimos uma norma é chamado de espaço vetorial normado.

**Definição 2.10** *O produto interno ou escalar de dois vetores  $\mathbf{x}$  e  $\mathbf{y} \in (\mathcal{C}^n, \mathcal{C})$  é um número complexo denotado por  $\langle \mathbf{x}, \mathbf{y} \rangle$  com as seguintes propriedades:*

a)

$$\langle \overline{\mathbf{x}}, \overline{\mathbf{y}} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle$$

onde a barra denota complexo conjugado.

b)

$$\langle a_1 \mathbf{x}_1 + a_2 \mathbf{x}_2, \mathbf{y} \rangle = \overline{a_1} \langle \mathbf{x}_1, \mathbf{y} \rangle + \overline{a_2} \langle \mathbf{x}_2, \mathbf{y} \rangle,$$

$$a_1 \in \mathcal{C}, \quad \mathbf{x}_1, \mathbf{x}_2 \text{ e } \mathbf{y} \in (\mathcal{C}^n, \mathcal{C})$$

c)

$$\langle \mathbf{x}, \mathbf{x} \rangle \geq 0 \quad \forall \quad \mathbf{x} \neq 0.$$

De a) e b) temos:

$$\begin{aligned} \langle \mathbf{x}, a\mathbf{y} \rangle &= \langle \overline{a\mathbf{y}}, \overline{\mathbf{x}} \rangle = \overline{\overline{a} \langle \mathbf{y}, \mathbf{x} \rangle} = \\ &= a \langle \overline{\mathbf{y}}, \overline{\mathbf{x}} \rangle = a \langle \mathbf{x}, \mathbf{y} \rangle \end{aligned}$$

Em princípio adotaremos como produto escalar:

$$\langle \mathbf{x}, \mathbf{y} \rangle \equiv \mathbf{x}^* \bullet \mathbf{y} = \mathbf{x}^* \mathbf{y} = \sum_{i=1}^n \overline{x_i} y_i = \|\mathbf{x}\| \|\mathbf{y}\| \cos \theta$$

onde  $*$  denota transposta do complexo conjugado; portanto para qualquer matriz  $\mathbf{M}_{n \times n}$ , temos:

a)

$$\langle \mathbf{x}, \mathbf{M}\mathbf{y} \rangle = \mathbf{x}^* \mathbf{M}\mathbf{y}$$

b)

$$\langle \mathbf{M}^* \mathbf{x}, \mathbf{y} \rangle = (\mathbf{M}^* \mathbf{x})^* \mathbf{y} = \mathbf{x}^* \mathbf{M}\mathbf{y}$$

Portanto de a) e b) temos:

$$\langle \mathbf{x}, \mathbf{M}\mathbf{y} \rangle = \langle \mathbf{M}^* \mathbf{x}, \mathbf{y} \rangle$$

Em particular no  $(\mathbb{R}^n, \mathbb{R})$  temos:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \mathbf{x}^T \mathbf{y} \quad \text{e} \quad \langle \mathbf{x}, \mathbf{M}\mathbf{y} \rangle = \langle \mathbf{M}^T \mathbf{x}, \mathbf{y} \rangle$$

Notemos que a norma euclidiana de um vetor é exprimível como:

$$\|\mathbf{x}\| = \sqrt{\langle \mathbf{x}, \mathbf{x} \rangle}$$

Um espaço vetorial onde definimos o produto interno é chamado espaço vetorial com produto interno.

**Definição 2.11** *Dois vetores são ortogonais se  $\langle \mathbf{x}, \mathbf{y} \rangle = 0$*

**Definição 2.12** *Um vetor está normalizado se  $\langle \mathbf{x}, \mathbf{x} \rangle = 1$*

**Definição 2.13** *Uma base ortonormal de  $(\mathbb{C}^n, \mathbb{C})$  é o conjunto de vetores  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_n\}$ , tal que:*

$$\langle \mathbf{u}_i, \mathbf{u}_j \rangle = \delta_{i,j} = \begin{cases} 0 & \text{para } i \neq j, \\ 1 & \text{para } i = j \end{cases} \quad i, j = 1, 2, \dots, n$$

**Definição 2.14** *O ângulo  $\Theta$  entre dois vetores  $\mathbf{x}$  e  $\mathbf{y}$  é definido por:*

$$\cos \Theta = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{x}\| \|\mathbf{y}\|}$$

### 2.2.9 Matriz Ortogonal

Seja  $\mathbf{A}_{n \times n}$  uma matriz com vetores coluna linearmente independentes,  $\{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_n\}$ , estes vetores constituem uma base ortonormal se:

$$\mathbf{a}_i^T \mathbf{a}_j = \begin{cases} 0 & \text{para } i \neq j \\ 1 & \text{para } i = j \end{cases} \quad \text{ou} \quad \begin{cases} \mathbf{a}_i \perp \mathbf{a}_j \\ \|\mathbf{a}_i\| = 1 \end{cases}$$

Neste caso temos:

$$\mathbf{A}^T \mathbf{A} = \begin{bmatrix} - & \mathbf{a}_1^T & - \\ - & \mathbf{a}_2^T & - \\ & \vdots & \\ - & \mathbf{a}_n^T & - \end{bmatrix} \begin{bmatrix} | & | & & | \\ \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & & 0 \\ \vdots & & \ddots & \\ 0 & & & 1 \end{bmatrix} = \mathbf{I}$$

**Teorema 2.9** A matriz  $\mathbf{A}_{n \times n}$  é ortonormal se e somente se  $\mathbf{A}^T \mathbf{A} = \mathbf{I}$  e:

$$\|\mathbf{Ax}\| = \|\mathbf{x}\|$$

para qualquer vetor  $\mathbf{x}$ .  $\mathbf{A}$  também preserva o produto interno:

$$(\mathbf{Ax})^T (\mathbf{Ay}) = \mathbf{x}^T \mathbf{A}^T \mathbf{Ay} = \mathbf{x}^T \mathbf{y}$$

**Prova:**

$$\|\mathbf{Ax}\|^2 = \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = \mathbf{x}^T \mathbf{Ix} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2$$

pois,

$$(\mathbf{Ax})^T (\mathbf{Ax}) = \mathbf{x}^T \mathbf{A}^T \mathbf{Ax} = \mathbf{x}^T \mathbf{Ix} = \mathbf{x}^T \mathbf{x} = \|\mathbf{x}\|^2$$

**C.Q.D.**

Algumas características de uma matriz ortonormal são:

- a) As linhas de  $\mathbf{A}$  formam uma base ortonormal.
- b) As colunas de  $\mathbf{A}$  formam uma base ortonormal.
- c)  $\mathbf{A}^{-1} = \mathbf{A}^T$ .
- d) Se  $\mathbf{A}_1$  e  $\mathbf{A}_2$  são matrizes ortonormais,  $\mathbf{A}_1 \mathbf{A}_2$  também o será.

### 2.2.10 Projecção

Seja  $\mathcal{X} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$  sendo  $\mathbf{x} \in \mathcal{X}$ ,  $\mathbf{y}_1 \in \mathcal{Y}_1$ ,  $\mathbf{y}_2 \in \mathcal{Y}_2$ , . Então , a transformação  $\mathbf{S}$  , tal que  $\mathbf{S}\mathbf{x} = \mathbf{y}_1$  é chamada a projecção de  $\mathcal{X}$  sobre  $\mathcal{Y}_1$  ao longo de  $\mathcal{Y}_2$  .

**Teorema 2.10** *Se  $\mathcal{Y}$  é um subespaço de  $\mathcal{X}$  e  $\mathbf{x}$  é um vetor não em  $\mathcal{Y}$ , podemos representá-lo de uma única forma como  $\mathbf{x} = \mathbf{y} + \mathbf{y}_p$  , onde  $\mathbf{y} \in \mathcal{Y}$  e  $\mathbf{y}_p = \mathbf{x} - \mathbf{y}$  é ortogonal a todos os vetores de  $\mathcal{Y}$ .  $\mathbf{y}$  é definido como a projecção ortogonal do vetor  $\mathbf{x}$  sobre  $\mathcal{Y}$ .*

**Prova:** Seja  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s, \mathbf{u}_{s+1}, \dots, \mathbf{u}_n\}$  uma base ortogonal de  $\mathcal{X}$ , e seja  $\mathcal{Y}$  o subespaço gerado por  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s\}$  . Então qualquer vetor  $\mathbf{x}$  em  $\mathcal{C}^n$ :

$$\mathbf{x} = \sum_{i=1}^s a_i \mathbf{u}_i + \sum_{i=s+1}^n a_i \mathbf{u}_i$$

Seja:

$$\mathbf{y} = \sum_{i=1}^s a_i \mathbf{u}_i \in \mathcal{Y}$$

e

$$\mathbf{y}_p = \mathbf{x} - \mathbf{y} = \sum_{i=s+1}^n a_i \mathbf{u}_i$$

como a base é ortogonal,  $\mathbf{y}_p$  é ortogonal a  $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_s\}$ , e portanto também ao vetor  $\mathbf{y}$ .

Mostremos que a representação é única. Por contradição, suponhamos que exista uma outra representação:

$$\mathbf{x} = \bar{\mathbf{y}} + \bar{\mathbf{y}}_p, \bar{\mathbf{y}} \in \mathcal{Y}$$

e  $\bar{\mathbf{y}}_p$  é ortogonal a todos os vetores de  $\mathcal{Y}$ , com  $\mathbf{y} \neq \bar{\mathbf{y}}$ :

$$(\mathbf{y} - \bar{\mathbf{y}}) + (\mathbf{y}_p - \bar{\mathbf{y}}_p) = 0$$

fazendo o produto interno com  $\mathbf{y} - \bar{\mathbf{y}}$  temos:

$$\langle \mathbf{y} - \bar{\mathbf{y}}, \mathbf{y} - \bar{\mathbf{y}} \rangle + \langle \mathbf{y} - \bar{\mathbf{y}}, \mathbf{y}_p - \bar{\mathbf{y}}_p \rangle = 0$$

Como

$$\langle \mathbf{y} - \bar{\mathbf{y}}, \mathbf{y}_p - \bar{\mathbf{y}}_p \rangle = 0$$

e

$$\langle \mathbf{y} - \bar{\mathbf{y}}, \mathbf{y} - \bar{\mathbf{y}} \rangle \neq 0$$

concluimos que é uma contradição; portanto:

$$\mathbf{y} - \bar{\mathbf{y}} = 0 \Rightarrow \mathbf{y} = \bar{\mathbf{y}}$$

Logo a representação é única.

C.Q.D.

**Definição 2.15** O espaço gerado por  $\{\mathbf{u}_{s+1}, \dots, \mathbf{u}_n\}$ , representado por  $\mathcal{Y}^\perp$  contém todos os vetores de  $\mathcal{X}$  que são ortogonais a todos os vetores de  $\mathcal{Y}$  e por isto,  $\mathcal{Y}^\perp$  é chamado complemento ortogonal de  $\mathcal{Y}$ .

### 2.2.11 Projeção Ortogonal

A projeção ortogonal,  $\mathbf{p}$ , de  $\mathbf{x}$  sobre  $\mathbf{y}$  :  $\mathbf{p} = P_{\mathbf{y}}\mathbf{x} \equiv \mathbf{x}|\mathbf{y}$  é tal que o vetor  $\mathbf{v} = \mathbf{x} - P_{\mathbf{y}}\mathbf{x}$  é ortogonal à  $\mathbf{y}$ . Na Figura 2.3 ilustramos este tipo de projeção.

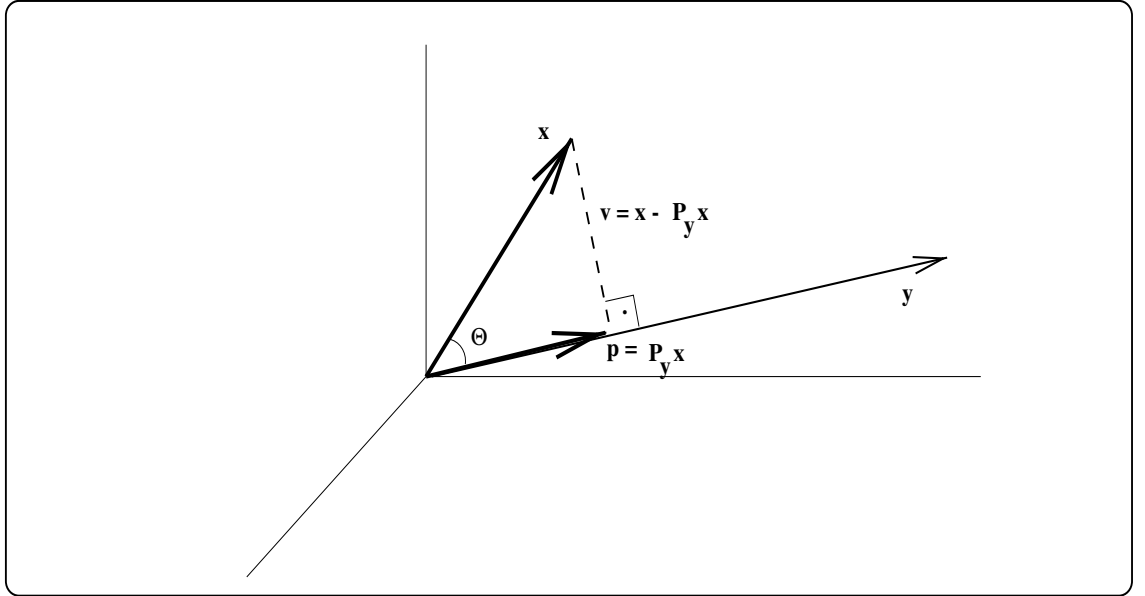


Figura 2.3: Projeção ortogonal de vetor  $\mathbf{x}$  sobre vetor  $\mathbf{y}$ .

**Teorema 2.11** Se  $P_{\mathbf{y}}\mathbf{x}$  é a projeção ortogonal do vetor  $\mathbf{x}$  sobre vetor  $\mathbf{y}$  em  $(\mathbb{R}^n, \mathbb{R})$ , então:

$$P_{\mathbf{y}}\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|} \frac{\mathbf{y}}{\|\mathbf{y}\|} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2} \mathbf{y} = \frac{\mathbf{y}\mathbf{y}^T}{\|\mathbf{y}\|^2} \mathbf{x} \quad (2.8)$$



**Prova:**

A projeção ortogonal de  $\mathbf{x}$  sobre  $\mathbf{y}$  pode ser escrita como:

$$P_{\mathbf{y}}\mathbf{x} = \lambda\mathbf{y} \quad (2.9)$$

com  $\lambda$  constante escalar. Portanto:

$$\begin{aligned} \mathbf{x} &= \lambda\mathbf{y} + \mathbf{v} \\ \langle \mathbf{x}, \mathbf{y} \rangle &= \lambda \|\mathbf{y}\|^2 + \langle \mathbf{v}, \mathbf{y} \rangle \end{aligned}$$

como  $\langle \mathbf{v}, \mathbf{y} \rangle = 0$  tem-se:

$$\langle \mathbf{x}, \mathbf{y} \rangle = \lambda \|\mathbf{y}\|^2$$

então,

$$\lambda = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2}$$

substituindo o valor de  $\lambda$  em (2.9) tem-se:

$$P_{\mathbf{y}}\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2} \mathbf{y} \quad (2.10)$$

Qual a matriz  $\mathbf{M}$  que representa o operador  $P_{\mathbf{y}}$  na equação (2.11) ?

$$P_{\mathbf{y}}\mathbf{x} = \mathbf{M}\mathbf{x} = \frac{\langle \mathbf{x}, \mathbf{y} \rangle}{\|\mathbf{y}\|^2} \mathbf{y} \quad (2.11)$$

Da equação (2.11) temos:

$$\mathbf{y} = \frac{\|\mathbf{y}\|^2}{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{M}\mathbf{x} \quad \Rightarrow \quad \mathbf{y}\mathbf{y}^T = \frac{\|\mathbf{y}\|^2}{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{M}\mathbf{x}\mathbf{y}^T$$

Portanto:

$$\frac{\mathbf{y}\mathbf{y}^T}{\|\mathbf{y}\|^2} = \frac{1}{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{M}\mathbf{x}\mathbf{y}^T$$

logo,

$$\begin{aligned} \frac{\mathbf{y}\mathbf{y}^T}{\|\mathbf{y}\|^2} \mathbf{x} &= \frac{1}{\langle \mathbf{x}, \mathbf{y} \rangle} \mathbf{M}\mathbf{x}\mathbf{y}^T \mathbf{x} \\ \frac{\mathbf{y}\mathbf{y}^T}{\|\mathbf{y}\|^2} \mathbf{x} &= \mathbf{M}\mathbf{x} \end{aligned} \quad (2.12)$$

Das equações (2.11) e (2.12) concluímos que o operador projetor ortogonal sobre o vetor  $\mathbf{y}$  é dado por:

$$P_{\mathbf{y}} = \frac{\mathbf{y}\mathbf{y}^T}{\|\mathbf{y}\|^2}$$

**C.Q.D.**

### 2.2.12 Reflexão

Se  $P_y \mathbf{x}$  é a projeção ortogonal de  $\mathbf{x}$  sobre o vetor  $\mathbf{y}$ , então a reflexão de  $\mathbf{x}$  em relação à  $\mathbf{y}$ ,  $\mathbf{R}_y \mathbf{x}$ , é um vetor tal que:

$$\mathbf{R}_y \mathbf{x} = 2P_y \mathbf{x} - \mathbf{x} \quad (2.13)$$

Na Figura 2.4 apresentamos duas possíveis reflexões do vetor  $\mathbf{x}$ , uma reflexão em relação a  $\mathbf{y}$  outra em relação a  $\mathbf{y}^\perp$ . Observe que a reflexão de  $\mathbf{x}$  em relação a  $\mathbf{y}$  associa o vetor simétrico  $\mathbf{R}_y \mathbf{x}$  à  $\mathbf{x}$ , de tal modo que  $\mathbf{y}$  seja a mediatriz entre  $\mathbf{x}$  e  $\mathbf{R}_y \mathbf{x}$ .

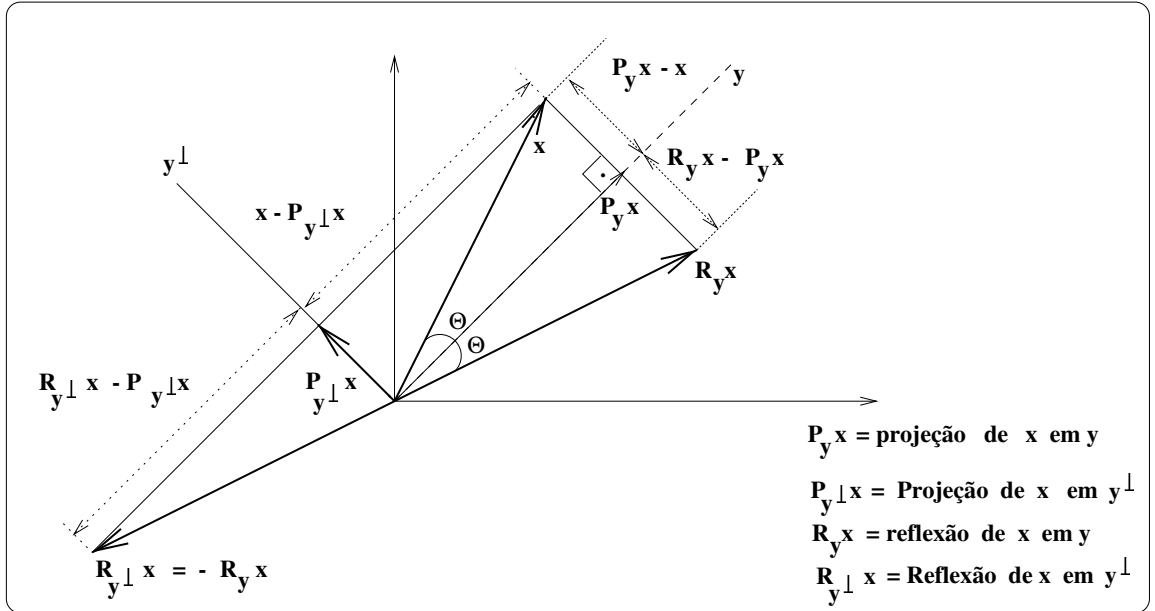


Figura 2.4: Reflexões de um vetor  $\mathbf{x}$  em relação à  $\mathbf{y}$  e  $\mathbf{y}^\perp$ .

Se  $P_y \mathbf{x}$  é a projeção sobre  $\mathbf{y}$  do vetor  $\mathbf{x}$ , então:

$$P_y \mathbf{x} - \mathbf{x} = \mathbf{R}_y \mathbf{x} - P_y \mathbf{x} \Rightarrow \mathbf{R}_y \mathbf{x} = -\mathbf{x} + 2P_y \mathbf{x}$$

$$\mathbf{R}_y \mathbf{x} = -\mathbf{x} + 2P_y \mathbf{x} = (2P_y - \mathbf{I})\mathbf{x}$$

logo o operador reflexão,  $\mathbf{R}_y$  é:

$$\mathbf{R}_y = 2P_y - \mathbf{I} \quad (2.14)$$

Como consequência de uma transformação de reflexão podemos tecer os seguintes comentários:

- A reflexão também pode ser vista como uma transformação que preserva a norma do vetor; podemos observar isto na Figura 2.4 onde as normas dos vetores  $\mathbf{x}$  e  $\mathbf{R}_y \mathbf{x}$  e  $\mathbf{R}_{y^\perp} \mathbf{x}$  são tais que:

$$\|\mathbf{x}\| = \|\mathbf{R}_y \mathbf{x}\| = \|\mathbf{R}_{y^\perp} \mathbf{x}\|$$

- Reflexão de  $\mathbf{x}$  em relação à  $y^\perp$

Da equação (2.14) tem-se que para reflexão de  $\mathbf{x}$  em relação à  $y^\perp$ :

$$\mathbf{R}_{y^\perp} \mathbf{x} = 2P_{y^\perp} \mathbf{x} - \mathbf{x} \quad (2.15)$$

O vetor  $\mathbf{x}$  pode ser decomposto como:

$$\mathbf{x} = P_y \mathbf{x} + P_{y^\perp} \mathbf{x}$$

logo

$$P_{y^\perp} \mathbf{x} = \mathbf{x} - P_y \mathbf{x} \quad (2.16)$$

substituindo a equação (2.16) em (2.15) tem-se:

$$\mathbf{R}_{y^\perp} \mathbf{x} = 2(\mathbf{x} - P_y \mathbf{x}) - \mathbf{x}$$

donde

$$\mathbf{R}_{y^\perp} \mathbf{x} = \mathbf{x} - 2P_y \mathbf{x} \quad (2.17)$$

Fazendo que  $\mathbf{w}$  seja a projeção ortogonal de  $\mathbf{x}$  em  $y$  e substituindo a equação (2.8) em (2.17) tem-se:

$$\mathbf{R}_{y^\perp} \mathbf{x} = \mathbf{x} - 2P_y \mathbf{x} = \mathbf{x} - 2 \frac{\mathbf{w}\mathbf{w}^T}{\|\mathbf{w}\|^2} \mathbf{x}$$

e o vetor  $\mathbf{R}_{y^\perp} \mathbf{x}$  pode ser representado na forma:

$$\mathbf{R}_{y^\perp} \mathbf{x} = \left( \mathbf{I} - 2 \frac{\mathbf{w}\mathbf{w}^T}{\|\mathbf{w}\|^2} \right) \mathbf{x} = \mathbf{H} \mathbf{x} \quad (2.18)$$

onde o operador  $\mathbf{R}_{y^\perp}$  é:

$$\mathbf{R}_{y^\perp} = \mathbf{I} - 2 \frac{\mathbf{w}\mathbf{w}^T}{\|\mathbf{w}\|^2} \quad (2.19)$$

e  $\mathbf{H}$  é a matriz de Householder.

- Se o vetor  $\mathbf{w}$  for unitário, a reflexão em relação à  $\mathbf{y}^\perp$  pode ser obtida através da matriz:

$$\mathbf{Q} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T \quad (2.20)$$

que é ortogonal e simétrica:

$$\mathbf{Q}^T = \mathbf{Q}^{-1} = \mathbf{Q}$$

pois:

$$\mathbf{Q}^T = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T = \mathbf{Q}$$

e

$$\mathbf{Q}^T \mathbf{Q} = \mathbf{I} - 4\mathbf{w}\mathbf{w}^T + 4\mathbf{w}\mathbf{w}^T = \mathbf{I}$$

**Exemplo 2.5** Seja o vetor unitário  $\mathbf{w}_1 = [1, 0]^T$ , calculando a matriz de reflexão associada a este vetor através da equação (2.20) temos:

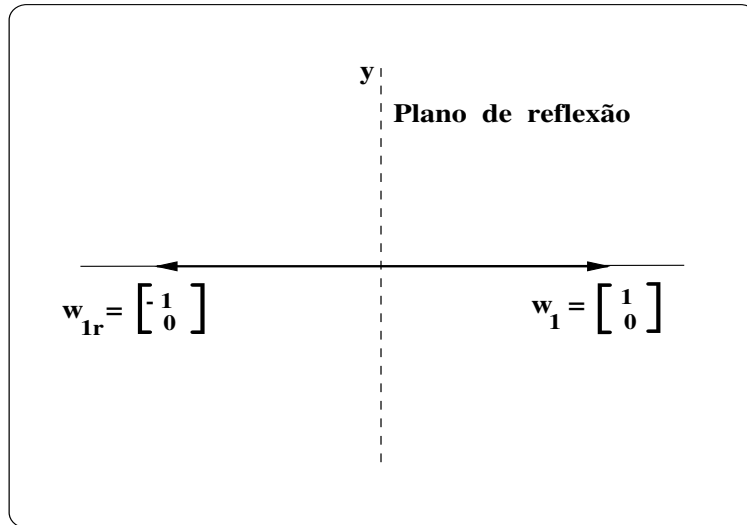


Figura 2.5: Reflexão de um vetor.

$$\mathbf{Q}_1 = \mathbf{I} - 2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}$$

Como podemos ver  $\mathbf{Q}_1$  reflete  $\mathbf{w}_1$  para  $\mathbf{w}_{1r} = [-1, 0]^T$ .

Neste caso  $\mathbf{Q}_1 * \mathbf{w} = \mathbf{w}_r$  reflete em relação ao eixo  $y$  qualquer vetor, Figura 2.5.

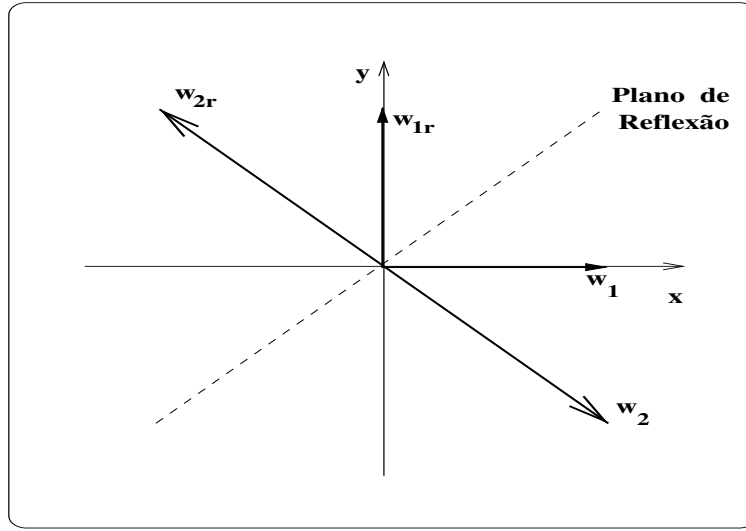


Figura 2.6: Reflexão de um vetor.

**Exemplo 2.6** Seja o vetor unitário  $\mathbf{w}_2 = [1/\sqrt{2}, -1/\sqrt{2}]^T$ , calculando a matriz de reflexão associada a este vetor através da equação (2.20) temos:

$$\mathbf{Q}_2 = \mathbf{I} - 2 \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

Como podemos ver, Figura 2.6,  $\mathbf{Q}_2$  reflete  $\mathbf{w}_2$  para  $\mathbf{w}_{2r} = [-1/\sqrt{2}, 1/\sqrt{2}]$ .

Neste caso  $\mathbf{Q}_2 * \mathbf{w} = \mathbf{w}_r$  reflete qualquer vetor sobre o plano cujo ângulo é de  $45^\circ$ . Na Figura 2.6 também está ilustrada a reflexão do vetor  $\mathbf{w}_1$ , exemplo 2.5, utilizando  $\mathbf{Q}_2$ . Note que um vetor sobre o plano de reflexão, por exemplo  $[1, 1]^T$  não sofrerá reflexão.

## 2.3 Transformações Ortogonais de Householder

Vários problemas algébricos envolvendo matrizes podem ser resolvidos deixando-as em alguma forma especial (diagonal, triangular, entre outras); isto é possível através de uma sequência de transformações ortogonais elementares. A implementação computacional das transformações ortogonais elementares deverá ser realizada por meio de métodos que garantam a estabilidade numérica.

Três tipos de transformações geralmente são utilizados para resolução de problemas algébricos: rotações no plano, reflexões elementares e permutações, Householder [508, 507].

Para a construção de uma matriz triangular a partir de uma matriz  $\mathbf{A}$ , o método de Householder envolve uma classe de matrizes que são simultaneamente simétricas, elementares<sup>2</sup> e ortogonais. Esta técnica é apresentada a seguir.

Qualquer vetor refletido sobre um vetor da base produz zeros na sua reflexão. Uma sequência de operações para refletir um vetor  $\mathbf{x}$  sobre um vetor  $\mathbf{Q}$  da base é apresentada na Figura 2.7.

Sejam  $P_{\mathbf{w}}\mathbf{x}$  a projeção de  $\mathbf{x}$  em  $\mathbf{w}$  e  $P_{\mathbf{w}^\perp}\mathbf{x}$  a projeção de  $\mathbf{x}$  em  $\mathbf{w}^\perp$ , Figura 2.7: A. Portanto:

$$\mathbf{x} = P_{\mathbf{w}^\perp}\mathbf{x} + P_{\mathbf{w}}\mathbf{x}$$

A reflexão em relação à  $\mathbf{w}^\perp$  de  $P_{\mathbf{w}}\mathbf{x}$  :  $-P_{\mathbf{w}}\mathbf{x}$  é mostrada na Figura 2.7: B.

A reflexão de  $\mathbf{x}$  sobre  $\mathbf{w}^\perp$  é dada:

$$\mathbf{Q} = P_{\mathbf{w}^\perp}\mathbf{x} - P_{\mathbf{w}}\mathbf{x}$$

$\mathbf{Q}$  difere de  $\mathbf{x}$  somente pelo sinal de  $P_{\mathbf{w}}\mathbf{x}$ , Figura 2.7: C.

Se o vetor  $\mathbf{w}$  for a diagonal de um paralelogramo equilátero formado pelo vetor  $\mathbf{x}$  e por um vetor de igual comprimento sobre a base, Figura 2.7: D, a reflexão  $\mathbf{Q}$  do vetor  $\mathbf{x}$  em relação à  $\mathbf{w}^\perp$  terá uma coordenada nula.

Se a escolha do vetor  $\mathbf{w}$  for tal que o vetor  $\mathbf{Q}$ , vetor de Householder, é também a reflexão do vetor sobre a base, Figura 2.7: D, tem-se que:

$$\mathbf{w} = \mathbf{x} + \|\mathbf{x}\| \mathbf{e}_i$$

onde  $\mathbf{e}_i$  é vetor unitário da base. Como a reflexão é uma operação que preserva o comprimento do vetor original,  $\mathbf{x}$ , e  $\mathbf{e}_i$  tem norma unitária, a norma euclidiana  $\|\mathbf{x}\|$  é utilizada sobre  $\mathbf{e}_i$ .

Desta forma o vetor de Householder,  $\mathbf{Q} = \mathbf{H}\mathbf{x}$ , deve estar sobre um vetor da base para que sejam introduzidos zeros no vetor  $\mathbf{x}$ :

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_m \end{bmatrix} \quad \text{e} \quad \mathbf{Q} = \begin{bmatrix} \|\mathbf{x}\| \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.21)$$

A operação  $\mathbf{Q} = \mathbf{H}\mathbf{x}$  provoca uma reflexão que irá preservar a norma:  $\|\mathbf{x}\| = \|\mathbf{Q}\|$ .

---

<sup>2</sup>Matriz elementar é obtida aplicando-se uma operação elementar na matriz identidade.

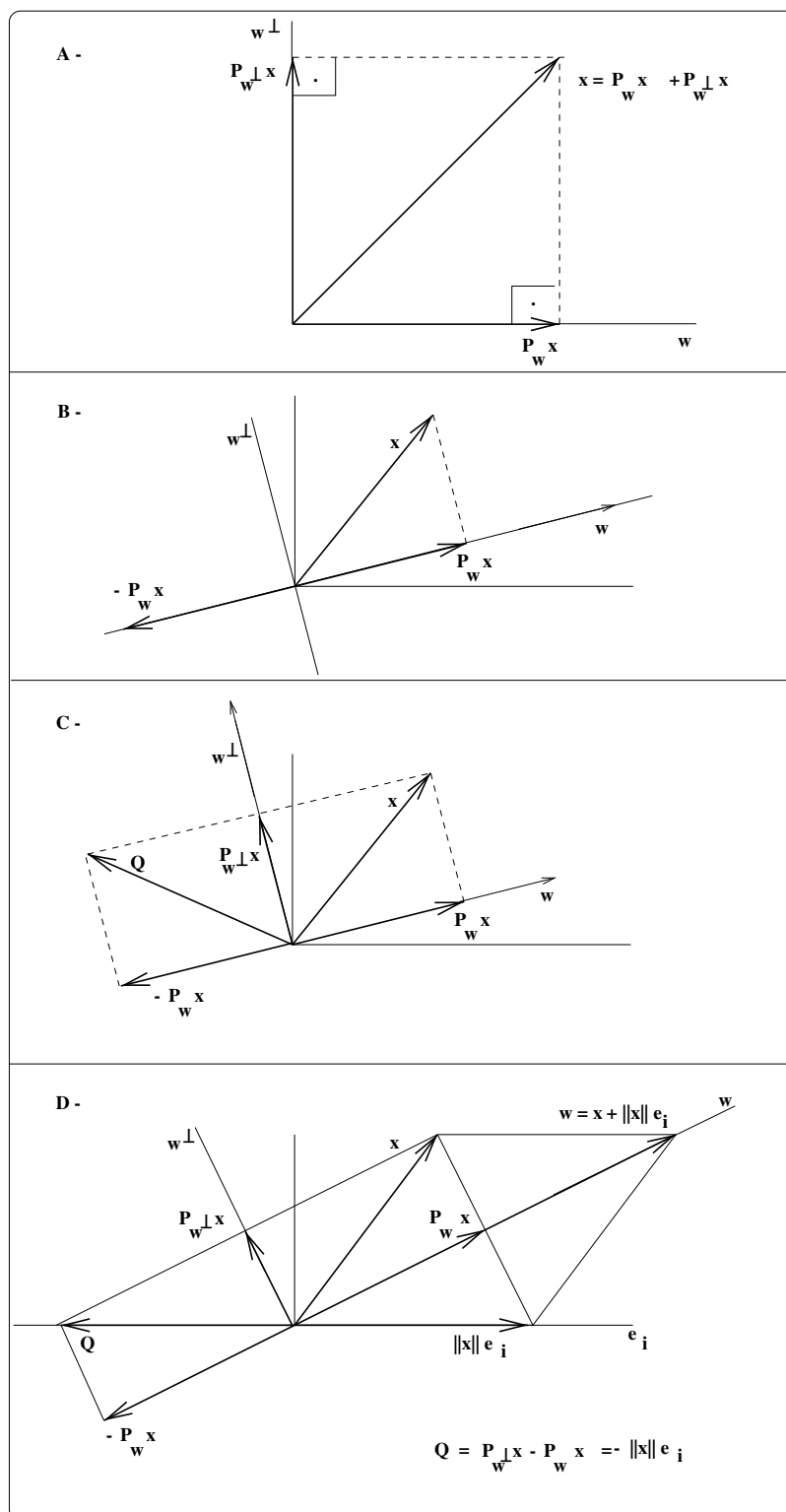


Figura 2.7: A - Decomposição de um vetor em componentes ortogonais;  
 B - Reflexão de  $P_w x$  em relação a  $w^\perp$ ; C - Reflexão de  $x$  em relação a  $w^\perp$ ;  
 D - Transformação de Householder do vetor  $x$  sobre um vetor da base.

Observe que se  $\mathbf{w} = \mathbf{x}$  então  $\mathbf{H}\mathbf{x} = -\mathbf{x}$ ; isto mostra que a matriz de transformação  $\mathbf{H}$  é um operador que reflete a imagem de qualquer vetor  $\mathbf{x}$  com respeito ao subespaço ortogonal a  $\mathbf{w}$ . A interpretação geométrica desta observação esta na Figura 2.8. O eixo de reflexão para um vetor  $\mathbf{x}$  é dado por:  $\mathbf{w} = \mathbf{x} + \|\mathbf{x}\| \mathbf{e}_i$  desde que o sinal de  $x_i$  seja positivo. Caso  $x_i$  seja negativo o eixo de reflexão de  $\mathbf{x}$  será dado por:  $\mathbf{w} = \mathbf{x} - \|\mathbf{x}\| \mathbf{e}_i$ .

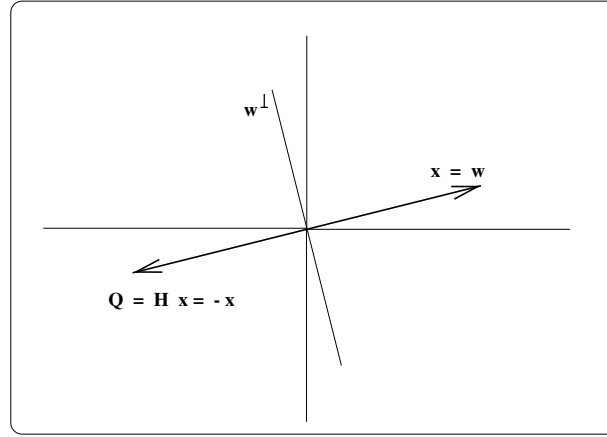


Figura 2.8: Reflexão de Householder.

Portanto uma transformação de Householder realizada pela equação (2.19) terá um vetor  $\mathbf{w}$  dado por:

$$\mathbf{w} = \mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i \quad (2.22)$$

onde:

$$\sigma = \frac{x_i}{|x_i|} \quad (2.23)$$

$x_i$  é a  $i$ -ésima componente de  $\mathbf{x}$  e o vetor  $\mathbf{e}_i$  é um vetor  $n$ -dimensional com  $n - 1$  elementos zero e um elemento igual a um na  $i$ -ésima posição.

Aplicando a transformação de Householder em  $\mathbf{x}$  tem-se:

$$\mathbf{Q} = \mathbf{H}\mathbf{x} = \mathbf{x} - \frac{2\mathbf{w}\mathbf{w}^T\mathbf{x}}{\mathbf{w}^T\mathbf{w}} = \mathbf{x} - 2(\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) \frac{(\mathbf{x}^T + \sigma \|\mathbf{x}\| \mathbf{e}_i^T)\mathbf{x}}{\|\mathbf{w}\|^2}$$

Para a norma do vetor  $\mathbf{w}$  tem-se:

$$\begin{aligned} \|\mathbf{w}\|^2 &= \mathbf{w}^T\mathbf{w} = (\mathbf{x}^T + \sigma \|\mathbf{x}\| \mathbf{e}_i^T) (\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) = \\ &= \mathbf{x}^T\mathbf{x} + 2[\sigma \|\mathbf{x}\| \mathbf{e}_i^T\mathbf{x}] + |\sigma|^2 \|\mathbf{x}\|^2 \mathbf{e}_i^T\mathbf{e}_i, \end{aligned}$$

Como  $\mathbf{e}_i^T\mathbf{x} = x_i$  tem-se:



$$\sigma \mathbf{e}_i^T \mathbf{x} = \frac{\mathbf{x}_i}{|\mathbf{x}_i|} \mathbf{x}_i = \frac{|\mathbf{x}_i|^2}{|\mathbf{x}_i|} = |\mathbf{x}_i|$$

e como  $\mathbf{e}_i^T \mathbf{e}_i = 1$ ,  $\mathbf{e}_i^T \mathbf{x} = x_i$  e

$$|\sigma|^2 = \frac{|\mathbf{x}_i|^2}{|\mathbf{x}_i|^2} = 1$$

tem-se:

$$\|\mathbf{w}\|^2 = \mathbf{w}^T \mathbf{w} = 2 \|\mathbf{x}\|^2 + 2 \|\mathbf{x}\| [\sigma \mathbf{e}_i^T \mathbf{x}] = 2(\|\mathbf{x}\|^2 + |\mathbf{x}_i| \|\mathbf{x}\|)$$

Portanto:

$$\begin{aligned} \mathbf{H}\mathbf{x} &= \mathbf{x} - \frac{2\mathbf{w}\mathbf{w}^T \mathbf{x}}{\|\mathbf{w}\|^2} = \mathbf{x} - 2(\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) \frac{(\mathbf{x}^T + \sigma \|\mathbf{x}\| \mathbf{e}_i^T) \mathbf{x}}{\|\mathbf{w}\|^2} = \\ &= \mathbf{x} - 2(\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) \left[ \frac{\|\mathbf{x}\|^2 + \sigma \|\mathbf{x}\| \mathbf{e}_i^T \mathbf{x}}{\|\mathbf{w}\|^2} \right] = \\ &= \mathbf{x} - 2(\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) \frac{\|\mathbf{x}\|^2 + |\mathbf{x}_i| \|\mathbf{x}\|}{2(\|\mathbf{x}\|^2 + |\mathbf{x}_i| \|\mathbf{x}\|)} = \\ &= \mathbf{x} - (\mathbf{x} + \sigma \|\mathbf{x}\| \mathbf{e}_i) = -\sigma \|\mathbf{x}\| \mathbf{e}_i = -\frac{\mathbf{x}_i}{|\mathbf{x}_i|} \|\mathbf{x}\| \mathbf{e}_i \end{aligned} \quad (2.24)$$

donde

$$\mathbf{Q} = -\frac{\mathbf{x}_i}{|\mathbf{x}_i|} \|\mathbf{x}\| \mathbf{e}_i \quad (2.25)$$

A equação (2.25) mostra que a transformação de Householder reflete o vetor  $\mathbf{x}$  em um vetor  $\mathbf{Q}$  com apenas um elemento não nulo na posição  $i$ .

### 2.3.1 Propriedades da Matriz de Householder

A matriz  $\mathbf{H}$  tem as seguintes propriedades:

a) Simetria:  $\mathbf{H} = \mathbf{H}^T$

b) Ortogonalidade:

$$\mathbf{H}\mathbf{H}^T = \left[ \mathbf{I} - \frac{2\mathbf{w}\mathbf{w}^T}{\mathbf{w}^T \mathbf{w}} \right] \left[ \mathbf{I} - \frac{2\mathbf{w}\mathbf{w}^T}{\mathbf{w}^T \mathbf{w}} \right] = \mathbf{I} - \frac{4\mathbf{w}\mathbf{w}^T}{\mathbf{w}^T \mathbf{w}} + \frac{4\mathbf{w}\mathbf{w}^T \mathbf{w}\mathbf{w}^T}{\mathbf{w}^T \mathbf{w} \mathbf{w}^T \mathbf{w}} = \mathbf{I}$$

c) Se  $\mathbf{w} = \mathbf{x}_i + \text{sign}(a_i) \|\mathbf{x}_i\|_2 \mathbf{e}_i$ , sendo  $\mathbf{e}_i$  um vetor unitário e  $\mathbf{x}_i$  o  $i$ -ésimo vetor coluna de  $\mathbf{A}$ , a cada transformação de Householder serão criados zeros na coluna  $i$ , abaixo da diagonal principal da matriz  $\mathbf{A}$ .

As transformações ortogonais de Householder são adequadas para a fatoração  $\mathbf{QR}$  de matriz.

## 2.4 Fatoração QR

Seja uma matriz  $\mathbf{A}$  cujas colunas são respectivamente  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  e  $\mathbf{a}_3$  e uma matriz,  $\mathbf{Q}$ , construída com vetores coluna ortonormais às colunas de  $\mathbf{A}$ , cujas colunas são respectivamente  $\mathbf{q}_1$ ,  $\mathbf{q}_2$  e  $\mathbf{q}_3$ . Se os vetores  $\mathbf{q}_i$  são combinações dos vetores  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  e  $\mathbf{a}_3$ , então existe uma terceira matrix  $\mathbf{R}$  que relaciona  $\mathbf{A}$  e  $\mathbf{Q}$ . Esta fatoração matricial é chamada decomposição **QR** da matriz  $\mathbf{A}$ , Francis, [354, 355], e é realizada pela redução de Gram-Schmidt, pela obtenção de vetores ortonormais  $\mathbf{q}_1$ ,  $\mathbf{q}_2$  e  $\mathbf{q}_3$  utilizando  $\mathbf{a}_1$ ,  $\mathbf{a}_2$  e  $\mathbf{a}_3$ :

Seja:

$$\mathbf{q}_1 = \frac{\mathbf{a}_1}{\|\mathbf{a}_1\|}$$

$\mathbf{q}_2$  é escrito da mesma forma em relação ao vetor  $\mathbf{a}_2$ , no entanto temos que encontrar  $\mathbf{b}$  ortogonal a  $\mathbf{a}_2$ :

$$\mathbf{b} = \mathbf{a}_2 - (\mathbf{q}_1^T \mathbf{a}_2) \mathbf{q}_1$$

logo,

$$\mathbf{q}_2 = \frac{\mathbf{b}}{\|\mathbf{b}\|}$$

$\mathbf{q}_3$  é obtido a partir de  $\mathbf{a}_3$ , que por sua vez é obtido de:

$$\hat{\mathbf{c}} = \mathbf{a}_3 - (\mathbf{q}_1^T \mathbf{a}_3) \mathbf{q}_1 \quad \Rightarrow \quad \mathbf{c} = \hat{\mathbf{c}} - (\mathbf{q}_2^T \hat{\mathbf{c}}) \mathbf{q}_2$$

e

$$\mathbf{q}_3 = \frac{\mathbf{c}}{\|\mathbf{c}\|}$$

O procedimento adotado foi subtrair de cada novo vetor suas projecções nas direções existentes. O resultado final é que os vetores  $\mathbf{q}_1$ ,  $\mathbf{q}_2$  e  $\mathbf{q}_3$  serão ortonormais.

Desta forma podemos relacionar  $\mathbf{A}$  com  $\mathbf{Q}$  através de  $\mathbf{R}$ :

$$\mathbf{A} = \left[ \begin{array}{c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 \end{array} \right] = \left[ \begin{array}{c|c|c} \mathbf{q}_1 & \mathbf{q}_2 & \mathbf{q}_3 \end{array} \right] \left[ \begin{array}{ccc} \mathbf{q}_1^T \mathbf{a}_1 & \mathbf{q}_1^T \mathbf{a}_2 & \mathbf{q}_1^T \mathbf{a}_3 \\ 0 & \mathbf{q}_2^T \mathbf{a}_2 & \mathbf{q}_2^T \mathbf{a}_3 \\ 0 & 0 & \mathbf{q}_3^T \mathbf{a}_3 \end{array} \right] = \mathbf{QR}. \quad (2.26)$$

como  $\mathbf{Q}^T \mathbf{Q} = \mathbf{I}$  temos  $\mathbf{Q}^T \mathbf{A} = \mathbf{R}$ .  $\mathbf{R}$  é triangular superior porque os vetores  $\mathbf{q}_i$  são ortogonais aos vetores  $\mathbf{a}_i$  e  $\mathbf{r}_{ij} = \mathbf{q}_i^T \mathbf{a}_j = 0$ .

Para formalizar a decomposição **QR** de  $\mathbf{A}$  apresentamos o seguinte resultado:

**Teorema 2.12** *Dada uma matriz  $\mathbf{A}_{m \times n}$  de posto  $r$ , tal que  $m \geq n$ , existem matrizes  $\mathbf{Q}_{m \times m}$  ortogonal e  $\mathbf{R}_{n \times n}$  triangular superior com elementos na diagonal não nulos tal que:*

$$\mathbf{A} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ 0 \end{bmatrix}, \quad (2.27)$$

**Prova:**

Para  $m = 1$ ,  $\mathbf{A}$  é um escalar e podemos ter  $\mathbf{Q} = \pm 1$  de acordo com  $\mathbf{A}$  ser negativo ou positivo. Para  $m > 1$ ,  $\mathbf{A}$  pode ser particionado como:

$$\mathbf{A} = \begin{bmatrix} a_1 & \mathbf{A}_2 \end{bmatrix}, \quad a_1 \in \mathbb{R}^m$$

Criando uma matriz ortogonal  $\mathbf{U} = [y, \mathbf{U}_1]$  com  $\mathbf{U}_1^T y = 0$  e

$$y = \begin{cases} a_1 / \|a_1\|, & a_1 \neq 0 \\ e_1, & a_1 = 0 \end{cases}$$

onde  $e_1$  representa a primeira coluna de uma matriz unitária  $\mathbf{I}$  tem-se:

$$\mathbf{U}^T \mathbf{A} = \begin{bmatrix} \rho & r^T \\ 0 & \mathbf{B} \end{bmatrix},$$

onde,

$$\rho = \|a_1\|, \quad r = \mathbf{A}_2^T y, \quad \text{e} \quad \mathbf{B} = \mathbf{U}_1^T \mathbf{A}_2 \in \mathbb{R}^{(m-1) \times (n-1)}.$$

Donde:

$$\mathbf{A} = \mathbf{U} \begin{bmatrix} \rho & r^T \\ 0 & \mathbf{B} \end{bmatrix} = \mathbf{Q} \begin{bmatrix} \mathbf{R} \\ 0 \end{bmatrix}$$

com

$$\mathbf{Q} = \mathbf{U} \begin{bmatrix} 1 & 0 \\ 0 & \bar{\mathbf{Q}} \end{bmatrix}, \quad \mathbf{R} = \begin{bmatrix} \rho & r^T \\ 0 & \bar{\mathbf{R}} \end{bmatrix}$$

e existe uma matriz ortogonal  $\bar{\mathbf{Q}}$  tal que:

$$\bar{\mathbf{Q}}^T \mathbf{B} = \begin{bmatrix} \bar{\mathbf{R}} \\ 0 \end{bmatrix}$$

**C.Q.D.**

### 2.4.1 Fatoração QR Através de Matrizes de Householder

O objetivo desta seção é obter a decomposição **QR** de **A** utilizando transformações de Householder.

Seja:

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}$$

- Passo 1 - Transforme **A** em **A**<sup>(1)</sup> com zeros na primeira coluna:

$$\mathbf{A}^{(1)} = \mathbf{H}_1 \mathbf{A} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \dots & \vdots \\ 0 & a_{m2}^{(1)} & \dots & a_{mn}^{(1)} \end{bmatrix}$$

A obtenção de **H**<sub>1</sub> requer somente a primeira coluna de **A**:

$$\mathbf{H}_1 = \left[ \mathbf{I} - \frac{2\mathbf{w}_1 \mathbf{w}_1^T}{\mathbf{w}_1^T \mathbf{w}_1} \right],$$

e **w**<sub>1</sub> é dado por:

$$\mathbf{w}_1 = \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{bmatrix} + \left( \frac{a_{11}}{\|\mathbf{a}_{11}\|} \right) \sqrt{a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2} \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad (2.28)$$

**H**<sub>1</sub> multiplicado pela primeira coluna da matriz **A** produz:

$$\mathbf{H}_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

onde  $x = \sqrt{a_{11}^2 + a_{21}^2 + \dots + a_{n1}^2}$ .

Então  $\mathbf{A}^{(1)} = \mathbf{H}_1 \mathbf{A}$  é da forma:

$$\mathbf{H}_1 \mathbf{A} = \mathbf{A}^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & a_{32}^{(1)} & \dots & a_{3n}^{(1)} \\ \vdots & \vdots & \dots & \vdots \\ 0 & a_{m2}^{(1)} & \dots & a_{mn}^{(1)} \end{bmatrix}$$

- Passo 2 - Transforme em zeros os componentes abaixo da diagonal principal da segunda coluna da matriz  $\mathbf{A}^{(1)}$  através de  $\mathbf{H}_2$ :

$$\mathbf{A}^{(2)} = \mathbf{H}_2 \mathbf{A}^{(1)} = \begin{bmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ 0 & 0 & \dots & a_{3n}^{(2)} \\ \vdots & \vdots & \dots & \vdots \\ 0 & 0 & \dots & a_{mn}^{(2)} \end{bmatrix}$$

Como o objetivo é zerar as componentes da segunda coluna de  $\mathbf{A}^{(1)}$  a partir da diagonal principal, devemos encontrar uma matriz  $\mathbf{H}_2$ , da forma:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 0 \\ 0 & \mathbf{H}_2^{\otimes} \end{bmatrix}$$

tal que

$$\mathbf{H}_2^{\otimes} \begin{bmatrix} a_{22}^{(1)} \\ a_{32}^{(1)} \\ \vdots \\ a_{m2}^{(1)} \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Portanto a obtenção de  $\mathbf{H}_2$  requer somente a segunda coluna de  $\mathbf{A}^{(1)}$  a partir da diagonal principal. Então  $\mathbf{H}_2^{\otimes}$  é obtido de forma semelhante ao passo 1.

Logo  $\mathbf{H}_2$  terá a seguinte forma:

$$\mathbf{H}_2 = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & x & \dots & x \\ 0 & x & \dots & x \\ \vdots & \vdots & & \vdots \\ 0 & x & \dots & x \end{bmatrix}$$

Então  $\mathbf{A}^{(2)} = \mathbf{H}_2 \mathbf{A}^{(1)}$ .

Note que  $\mathbf{H}_2$  tem zeros na primeira coluna abaixo da diagonal e a multiplicação de  $\mathbf{A}^{(1)}$  por  $\mathbf{H}_2$  irá preservar os zeros criados em  $\mathbf{A}^{(1)}$ .

- Passo  $i$  - Criar uma matriz de Householder  $\mathbf{H}_i^{\oplus}$  tal que:

$$\mathbf{H}_i^{\oplus} \begin{bmatrix} a_{ii} \\ a_{(i+1)i} \\ \vdots \\ a_{mi} \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

definindo  $\mathbf{H}_i$  como:

$$\mathbf{H}_i = \begin{bmatrix} \mathbf{I}_{(n-i) \times (n-i)} & 0 \\ 0 & \mathbf{H}_i^{\oplus} \end{bmatrix}$$

- Passo  $n-1$  -  $\mathbf{A}^{(n-1)} = \mathbf{H}_{n-1} \mathbf{A}^{(n-2)}$  é uma matriz que possui zeros nas colunas abaixo da diagonal principal sendo assim uma matriz triangular superior. Se  $\mathbf{A}^{(n-1)} = \mathbf{R}$ , definindo  $\mathbf{Q}^T$  por:

$$\mathbf{Q}^T = \mathbf{H}_{(n-1)} \mathbf{H}_{(n-2)} \dots \mathbf{H}_2 \mathbf{H}_1 \quad (2.29)$$

tem-se  $\mathbf{R} = \mathbf{Q}^T \mathbf{A}$ .

Note que  $\mathbf{H}_{n-1}^T = \mathbf{H}_{n-1}$ ,  $\mathbf{H}_{n-1}^{-1} = \mathbf{H}_{n-1}$  e que  $\mathbf{H}_{n-1}^T \mathbf{H}_{n-1} = \mathbf{I}$  então:

$$\begin{aligned} (\mathbf{Q}^T)^{-1} &= (\mathbf{H}_{(n-1)} \mathbf{H}_{(n-2)} \dots \mathbf{H}_2 \mathbf{H}_1)^{-1} = \\ &= \mathbf{H}_{(n-1)}^{-1} \mathbf{H}_{(n-2)}^{-1} \dots \mathbf{H}_2^{-1} \mathbf{H}_1^{-1} = \\ &= \mathbf{H}_1 \mathbf{H}_2 \mathbf{H}_{(n-2)} \dots \mathbf{H}_{(n-1)} = \mathbf{Q} \end{aligned}$$

Portanto  $\mathbf{A} = \mathbf{QR}$ . Note que a matriz  $\mathbf{Q}$  é ortogonal,  $(\mathbf{Q}^T \mathbf{Q} = \mathbf{I})$ .

**Exemplo 2.7** Seja a matriz ,

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 7 & 6 \end{bmatrix}$$

deseja-se obter a decomposição QR de  $\mathbf{A}$  utilizando transformações de Householder.

O primeiro passo será zerar a primeira coluna abaixo da diagonal.

$$\mathbf{H}_1 \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} = \begin{bmatrix} x \\ 0 \\ 0 \end{bmatrix}$$

Então definimos a matriz:

$$\mathbf{H}_1 = \mathbf{I} - \frac{2\mathbf{w}_1\mathbf{w}_1^T}{\mathbf{w}_1^T\mathbf{w}_1}$$

onde  $\mathbf{w}_1$  é dado por:

$$\mathbf{w}_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix} + \left( \frac{a_{11}}{\|\mathbf{a}_{11}\|} \right) \sqrt{1^2 + 2^2 + 2^2} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ 2 \\ 2 \end{bmatrix}$$

Substituindo  $\mathbf{w}_1$  em  $\mathbf{H}_1$  teremos:

$$\mathbf{H}_1 = \begin{bmatrix} -0.3333 & -0.6667 & -0.6667 \\ -0.6667 & 0.6667 & -0.3333 \\ -0.6667 & -0.3333 & -0.6667 \end{bmatrix}$$

Então obtemos uma matriz:

$$\mathbf{A}^{(1)} = \mathbf{H}_1\mathbf{A} = \begin{bmatrix} -3.000 & -9.0 & -8.6667 \\ 0 & 0 & -1.3333 \\ 0 & 0 & -0.3333 \end{bmatrix}$$

que é uma matriz triangular superior, então  $\mathbf{R} = \mathbf{A}^{(1)}$  e  $\mathbf{Q} = \mathbf{H}_1$  fatoram a matriz  $\mathbf{A}$  como  $\mathbf{A} = \mathbf{QR}$ .

### 2.4.2 Fatoração QR Através de Rotações de Givens

Considere uma matriz  $\mathbf{A}$  com elementos em  $\mathcal{R}$  e operações elementares tais como:

- a) multiplicação de uma linha ou coluna por um número,
- b) troca de posição de duas linhas ou colunas,
- c) adição do produto de uma linha ou coluna pelo produto de outra linha ou coluna.

efetuadas usando matrizes elementares.

**Exemplo 2.8** *Sejam a matriz  $\mathbf{A}_{5 \times 5}$  qualquer e três matrizes elementares,  $\mathbf{E}_{5 \times 5}$ :*

$$\mathbf{E}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad \mathbf{E}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad \mathbf{E}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & d & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

onde  $c$  e  $d$  são reais e diferentes de zero.

A multiplicação à esquerda de  $\mathbf{A}$  pelas matrizes elementares produz as seguintes alterações em  $\mathbf{A}$ :

- a)  $\mathbf{E}_1\mathbf{A}$ : multiplica por  $c$  a quarta linha de  $\mathbf{A}$ .
- b)  $\mathbf{E}_2\mathbf{A}$ : troca de posição a segunda e quinta linhas de  $\mathbf{A}$ .
- c)  $\mathbf{E}_3\mathbf{A}$ : adiciona a quarta linha de  $\mathbf{A}$  à segunda linha multiplicada por  $d$ .

se  $\mathbf{A}$  for multiplicada à direita teremos:

- a)  $\mathbf{A}\mathbf{E}_1$ : multiplica por  $c$  a quarta coluna de  $\mathbf{A}$ .
- b)  $\mathbf{A}\mathbf{E}_2$ : troca de posição a segunda e quinta colunas de  $\mathbf{A}$ .
- c)  $\mathbf{A}\mathbf{E}_3$ : adiciona a segunda coluna de  $\mathbf{A}$  à quarta coluna de  $\mathbf{A}$  multiplicada por  $d$ .

As matrizes elementares são matrizes quadradas e não singulares e suas inversas são:

$$\mathbf{E}_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1/c & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad \mathbf{E}_2^{-1} = \mathbf{E}_2 \quad e \quad \mathbf{E}_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -d & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

também matrizes elementares.

A matriz apresentada em (2.30) é chamada matriz de Givens, [389].



$$\mathbf{J}(i, j, c, s) = \begin{bmatrix} 1 & 0 & 0 & & \dots & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & c_{ii} & 0 & \dots & 0 & s_{ij} \\ & & & 0 & 1 & \dots & & 0 \\ & & & \vdots & & \ddots & & \vdots \\ & & & 0 & & \dots & 1 & 0 \\ & & & -s_{ji} & 0 & \dots & 0 & c_{jj} \\ & & & & & & & 1 \\ & & & & & & & & \ddots & \\ 0 & 0 & 0 & & \dots & & & & & 0 & 1 \end{bmatrix} \quad (2.30)$$

onde  $c^2 + s^2 = 1$ ; desde que a escolha seja  $c = \cos \alpha$  e  $s = \sin \alpha$  para algum  $\alpha$ , a matriz de Givens pode receber a seguinte notação  $\mathbf{J}(i, j, \alpha)$ . Geometricamente, a matriz  $\mathbf{J}(i, j, \alpha)$  irá rodar uma matriz  $\mathbf{A}$  no plano de coordenadas  $(i, j)$  de um ângulo  $\alpha$  e é conhecida como rotação de Givens no plano  $(i, j)$ .

Desta forma, quando um vetor com  $n$  componentes for multiplicado por uma matriz de Givens  $\mathbf{J}(i, j, \alpha)$  somente as componentes  $i$ -ésima e  $j$ -ésima serão alteradas. Como  $c^2 + s^2 = 1$ ,  $\mathbf{J}(i, j, \alpha)\mathbf{J}(i, j, \alpha)^T = \mathbf{I}$ , então a matriz de rotação é ortogonal.

**Exemplo 2.9** Seja  $\mathbf{x} = (x_1, x_2)^T$ , definindo:

$$c = \frac{x_1}{\sqrt{x_1^2 + x_2^2}}, \quad s = \frac{x_2}{\sqrt{x_1^2 + x_2^2}}$$

temos a seguinte matriz de rotação de Givens:

$$\mathbf{J}(1, 2, \alpha) = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}$$

que é ortogonal pois:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix}^T \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c & -s \\ s & c \end{bmatrix} \begin{bmatrix} c & s \\ -s & c \end{bmatrix} = \begin{bmatrix} c^2 + s^2 & 0 \\ 0 & c^2 + s^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \mathbf{I}$$

**Exemplo 2.10** Rotações de Givens são especialmente utilizadas para criar zeros em posições específicas de um vetor, ou abaixo da diagonal de uma matriz. Deseja-se criar um zero na terceira linha de um dado vetor:

$$\mathbf{x} = [1, 3, 4]^T$$

se considerarmos  $i$  como a segunda e  $j$  como a terceira linha de  $\mathbf{x}$  tem-se:

$$\mathbf{J}(i, j, \alpha)\mathbf{x} = \begin{bmatrix} ? \\ 0 \end{bmatrix}$$

donde tem-se que:

$$c^2 + s^2 = 1$$

e

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad \text{portanto,} \quad c = \frac{3}{\sqrt{25}}, \quad s = \frac{4}{\sqrt{25}}$$

e  $x = \sqrt{3^2 + 4^2} = 5$ . Então resulta a matriz de Givens:

$$\mathbf{J}(i, j, \alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{3}{\sqrt{25}} & \frac{4}{\sqrt{25}} \\ 0 & \frac{-4}{\sqrt{25}} & \frac{3}{\sqrt{25}} \end{bmatrix}$$

que realiza:

$$\mathbf{J}(i, j, \alpha)\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.6 & 0.8 \\ 0 & -0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}$$

A idéia da criação de zeros em uma dada posição de um vetor pode ser estendida para criar zeros em uma posição específica de uma matriz.

**Exemplo 2.11** Usando matriz de Givens deseja-se criar um zero na posição  $(2, 1)$  da matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 6 & 6 \end{bmatrix}$$

Para o vetor  $[1 \ 2]^T$  da primeira coluna de  $\mathbf{A}$  formamos a seguinte matriz de rotação:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad \text{donde} \quad c = \frac{1}{\sqrt{5}}, \quad s = \frac{2}{\sqrt{5}} \quad \text{e} \quad x = \sqrt{1^2 + 2^2}$$

Colocando o termo  $-s$  da matriz de rotação na posição do elemento que desejamos zerar em  $\mathbf{A}$  realizamos nosso objetivo sem o uso explícito da matriz de Givens. Para este exemplo resulta a matriz de Givens:

$$\mathbf{J}(1, 2, \alpha) = \begin{bmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} & 0 \\ \frac{-2}{\sqrt{5}} & \frac{1}{\sqrt{5}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

que realiza:

$$\mathbf{J}(1, 2, \alpha)\mathbf{A} = \begin{bmatrix} 2.2361 & 6.7082 & 6.2610 \\ 0 & 0 & -1.3416 \\ 2 & 6 & 6 \end{bmatrix}$$

Como a primeira e segunda colunas de  $\mathbf{A}$  são linearmente dependentes também foi criado um zero na posição  $a_{(2,2)}$  da matriz original neste caso.

**Exemplo 2.12** Desejamos criar um zero na posição  $(3,1)$  da matriz  $\mathbf{A}$  usando rotações de Givens:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 6 & 6 \end{bmatrix}$$

Usando  $\mathbf{J}(2, 3, \alpha)$  tem-se:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} x \\ 0 \end{bmatrix}, \quad c = \frac{2}{\sqrt{8}}, \quad s = \frac{2}{\sqrt{8}}, \quad e \quad x = \sqrt{2^2 + 2^2}$$

Portanto:

$$\mathbf{J}(2, 3, \alpha) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \\ 0 & \frac{-2}{\sqrt{8}} & \frac{2}{\sqrt{8}} \end{bmatrix}$$

donde:

$$\mathbf{J}(1, 2, \alpha)\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2.8284 & 8.4853 & 7.7782 \\ 0 & 0 & 0.7071 \end{bmatrix}$$

A fatoração **QR** de uma matriz  $\mathbf{A}$  pode ocorrer de diferentes modos, e um deles é através de Rotações de Givens, exemplificado a seguir.

**Exemplo 2.13** Para a matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 6 & 6 \end{bmatrix}$$

Colocamos um zero na posição  $(2, 1)$  fazendo:

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_{(1,1)} \\ a_{(2,1)} \end{bmatrix} = \begin{bmatrix} ? \\ 0 \end{bmatrix}$$

Como  $a_{(1,1)} = 1$  e  $a_{(2,1)} = 2$  e:

$$c = \frac{a_{(1,1)}}{\sqrt{a_{(1,1)}^2 + a_{(2,1)}^2}} \quad , \quad s = \frac{a_{(2,1)}}{\sqrt{a_{(1,1)}^2 + a_{(2,1)}^2}}$$

tem-se  $c = 0.4472$  e  $s = 0.8944$ . Onde:

$$\mathbf{J}(1, 2, \alpha) = \begin{bmatrix} 0.4472 & 0.8944 & 0 \\ -0.8944 & 0.4472 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{A}^{(1)} = \mathbf{J}(1, 2, \alpha) \mathbf{A} = \begin{bmatrix} 0.4472 & 0.8944 & 0 \\ -0.8944 & 0.4472 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 2.2361 & 6.7082 & 6.2610 \\ 0 & 0 & -1.3416 \\ 2 & 6 & 6 \end{bmatrix}$$

No passo seguinte colocaremos um zero na posição  $(3, 1)$  da matriz  $\mathbf{A}^{(1)}$ ; para isto trabalhamos com as componentes  $a_{(1,1)}^1$  e  $a_{(3,1)}^1$  da matriz  $\mathbf{A}^{(1)}$ :

$$\begin{bmatrix} c & s \\ -s & c \end{bmatrix} \begin{bmatrix} a_{(1,1)}^1 \\ a_{(3,1)}^1 \end{bmatrix} = \begin{bmatrix} ? \\ 0 \end{bmatrix}$$

como  $a_{(1,1)}^1 = 2.2361$  e  $a_{(3,1)}^1 = 2$  e:

$$c = \frac{a_{(1,1)}^1}{\sqrt{(a_{(1,1)}^1)^2 + (a_{(3,1)}^1)^2}} = 0.7454 \quad , \quad s = \frac{a_{(3,1)}^1}{\sqrt{(a_{(1,1)}^1)^2 + (a_{(3,1)}^1)^2}} = 0.6667$$

temos a matriz de Givens:

$$\mathbf{J}(1, 3, \alpha) = \begin{bmatrix} 0.7454 & 0 & 0.6667 \\ 0 & 1 & 0 \\ -0.6667 & 0 & 0.7454 \end{bmatrix}$$

multiplicando  $\mathbf{J}(1, 3, \alpha)$  pela matriz  $\mathbf{A}^{(1)}$  tem-se:

$$\begin{aligned} \mathbf{A}^{(2)} &= \mathbf{J}(1, 3, \alpha) \mathbf{A}^{(1)} = \\ &= \begin{bmatrix} 0.7454 & 0 & 0.6667 \\ 0 & 1 & 0 \\ -0.6667 & 0 & 0.7454 \end{bmatrix} \begin{bmatrix} 2.2361 & 6.7082 & 6.2610 \\ 0 & 0 & -1.3416 \\ 2 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 3 & 9 & 8.6667 \\ 0 & 0 & -1.3416 \\ 0 & 0 & 0.2982 \end{bmatrix} \end{aligned}$$

O próximo passo que seria obter um zero na posição  $(3, 2)$  da matriz  $\mathbf{A}^{(2)}$ , não é necessário neste exemplo.

O resultado obtido por  $\mathbf{A}^{(2)} = \mathbf{J}(1, 3, \alpha)\mathbf{A}^{(1)}$  corresponde a matriz  $\mathbf{R}$  da decomposição **QR**. A matriz  $\mathbf{Q}$  pode ser obtida através do produto das diversas matrizes  $\mathbf{J}(i, j, \alpha)$ :

$$\mathbf{Q} = \mathbf{Q}_1^T \mathbf{Q}_2^T \dots \mathbf{Q}_{(m-1)}^T$$

Observe que cada matriz  $\mathbf{Q}_i$  é formada por diversas matrizes  $\mathbf{J}(i, m, \alpha)$  correspondentes as rotações de Givens que ocorreram na coluna  $m$ :

$$\mathbf{Q}_i = \mathbf{J}(i, m, \alpha) \mathbf{J}(i, m-1, \alpha) \dots \mathbf{J}(i, i+1, \alpha)$$

De modo similar ao que ocorre com as rotações de Householder, as rotações de Givens fazem com que o sistema  $\mathbf{Ax} = \mathbf{b}$  seja reescrito como  $\mathbf{Rx} = \mathbf{Q}^T \mathbf{b}$ .

Sobre os processos de fatoração matricial Givens e Householder podemos tecer os seguintes comentários:

- A decomposição **QR** da matriz  $\mathbf{A}$  pode ser realizada através dos métodos de Givens ou Householder. As variações entre os dois métodos estão relacionadas principalmente com a velocidade da obtenção do resultado (Givens utiliza duas vezes mais operações do que Householder), logo a capacidade de memória da máquina e o tipo de aplicação entre outros deve ser levado em consideração quando da opção para escolha do método mais adequado.
- A diferença principal no processo de fatoração é que o método de Householder anula elementos de uma coluna em uma única operação, enquanto o método de Givens anula um elemento por vez.
- Em estudos de redes de potência elétrica ocorrem matrizes esparsas de grande porte. Para a redução do esforço computacional a esparsidade destas matrizes deve ser explorada, e para tanto o método de Givens é amplamente utilizado, Murari, [797].

## 2.5 O método **QR**

O método **QR** utilizado para a obtenção dos autovalores de uma matriz se origina da fatoração **QR**, Francis [354, 355].

Seja:

$$\mathbf{A}_k = \mathbf{Q}_k \mathbf{R}_k \tag{2.31}$$

onde  $\mathbf{Q}_k$  é ortogonal e  $\mathbf{R}_k$  é triangular superior; fazendo:

$$\mathbf{A}_{k+1} = \mathbf{R}_k \mathbf{Q}_k \quad (2.32)$$

tem-se:

$$\mathbf{A}_{k+1} = \mathbf{Q}_k^T \mathbf{A}_k \mathbf{Q}_k = \mathbf{Q}_k^T \mathbf{Q}_k \mathbf{R}_k \mathbf{Q}_k = \mathbf{R}_k \mathbf{Q}_k, \quad (2.33)$$

que sendo uma transformação de similaridade preserva os autovalores. Na prática a equação (2.31) é reescrita como:

$$\mathbf{Q}_k^T \mathbf{A}_k = \mathbf{R}_k \quad (2.34)$$

onde  $\mathbf{Q}_k^T$  é o produto de matrizes ortogonais de Givens ou de Householder.

A cada passo da obtenção das matrizes  $\mathbf{Q}_k \mathbf{R}_k$  a partir da matriz  $\mathbf{A}_k$  uma nova matriz  $\mathbf{A}_{k+1} = \mathbf{R}_k \mathbf{Q}_k$  é obtida. Fazendo  $\mathbf{A}_{k+1} = \mathbf{Q}_{k+1} \mathbf{R}_{k+1}$  e gerando uma nova matriz  $\mathbf{A}_{k+2} = \mathbf{R}_{k+1} \mathbf{Q}_{k+1}$ , e assim sucessivamente, teremos a convergência da matriz  $\mathbf{A}_{k+n} = \mathbf{R}_{k+n-1} \mathbf{Q}_{k+n-1}$  para uma matriz triangular cuja diagonal contém os autovalores da matriz  $\mathbf{A}_k$  inicial.

**Exemplo 2.14** *Seja a matriz:*

$$\mathbf{A}_1 = \begin{bmatrix} 1 & 3 & 4 \\ 2 & 6 & 5 \\ 2 & 6 & 6 \end{bmatrix} = \mathbf{Q}_1 \mathbf{R}_1 = \begin{bmatrix} -0.3333 & 0.9428 & 0 \\ -0.6667 & -0.2357 & -0.7071 \\ -0.6667 & -0.2357 & -0.7071 \end{bmatrix} \begin{bmatrix} -3 & -9 & -8.6667 \\ 0 & 0 & 1.1785 \\ 0 & 0 & 0.7071 \end{bmatrix}$$

*Calculando  $\mathbf{A}_2$ :*

$$\mathbf{A}_2 = \mathbf{R}_1 \mathbf{Q}_1 = \begin{bmatrix} -12.7778 & 1.3356 & 0.2357 \\ -0.7857 & -0.2778 & 0.8333 \\ -0.4714 & -0.1667 & 0.5000 \end{bmatrix}$$

*Fatorando  $\mathbf{A}_2$  temos:*

$$\mathbf{A}_2 = \mathbf{Q}_2 \mathbf{R}_2 = \begin{bmatrix} -0.9974 & -0.0715 & 0 \\ -0.0613 & -0.8553 & -0.5145 \\ 0.0368 & -0.5132 & 0.8575 \end{bmatrix} \begin{bmatrix} -12.8106 & -1.3554 & -0.1656 \\ 0 & 0.2276 & -0.9862 \\ 0 & 0 & 0 \end{bmatrix}$$

*Calculando  $\mathbf{A}_3$ :*

$$\mathbf{A}_3 = \mathbf{R}_2 \mathbf{Q}_2 = \begin{bmatrix} 12.6886 & 2.1605 & 0.5554 \\ -0.0223 & 0.3114 & -0.9627 \\ 0 & 0 & 0 \end{bmatrix}$$

*Fatorando  $\mathbf{A}_3$  temos:*

$$\mathbf{A}_3 = \mathbf{Q}_3 \mathbf{R}_3 = \begin{bmatrix} -1 & -0.0018 & 0 \\ 0.0018 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -12.6886 & -2.1599 & -0.5570 \\ 0 & -0.3152 & 0.9618 \\ 0 & 0 & 0 \end{bmatrix}$$

Calculando  $\mathbf{A}_4$ :

$$\mathbf{A}_4 = \mathbf{R}_3 \mathbf{Q}_3 = \begin{bmatrix} 12.6848 & 2.1823 & -0.5570 \\ -0.0006 & 0.3152 & 0.9618 \\ 0 & 0 & 0 \end{bmatrix}$$

Fatorando  $\mathbf{A}_4$  temos:

$$\mathbf{A}_4 = \mathbf{Q}_4 \mathbf{R}_4 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -12.6848 & -2.1823 & 0.5571 \\ 0 & -0.3153 & -0.9617 \\ 0 & 0 & 0 \end{bmatrix}$$

Calculando  $\mathbf{A}_5$ :

$$\mathbf{A}_5 = \mathbf{R}_4 \mathbf{Q}_4 = \begin{bmatrix} 12.6847 & 2.1828 & 0.5571 \\ 0 & 0.3153 & -0.9617 \\ 0 & 0 & 0 \end{bmatrix}$$

Fatorando  $\mathbf{A}_5$  temos:

$$\mathbf{A}_5 = \mathbf{Q}_5 \mathbf{R}_5 = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} -12.6847 & -2.1828 & -0.5571 \\ 0 & -0.3153 & 0.9617 \\ 0 & 0 & 0 \end{bmatrix}$$

Calculando  $\mathbf{A}_6$ :

$$\mathbf{A}_6 = \mathbf{R}_5 \mathbf{Q}_5 = \begin{bmatrix} 12.6847 & 2.1828 & 0.5571 \\ 0 & 0.3153 & -0.9617 \\ 0 & 0 & 0 \end{bmatrix}$$

Vemos que as matrizes  $\mathbf{A}_5$  e  $\mathbf{A}_6$  são iguais e que os valores 12.6847 , 0.3153 e zero que estão na diagonal principal da matriz  $\mathbf{A}_5$  são os autovalores da matriz  $\mathbf{A}_1$  inicial.





## Capítulo 3

# PROCESSAMENTO DE ALTO DESEMPENHO DE ALGORITMOS NUMÉRICOS

### 3.1 Introdução

A computação de alto desempenho é uma área de intensa atividade de pesquisa motivada por uma grande variedade de aplicações tais como métodos numéricos, modelagem de dados, otimização, computação gráfica, controle e automação dentre outras que demandam velocidades de processamento cada vez maiores. A cada ano, a indústria de computadores investe em busca de computadores mais poderosos. Basicamente, o investimento em novas tecnologias atua em dois campos:

1. Expansão dos limites físicos que determinam a velocidade máxima em que os circuitos elétricos conseguem operar.
2. Desenvolvimento de novas arquiteturas que permitam obter maior desempenho, dentro dos limites físicos atuais.

As máquinas desenvolvidas a cada ano sempre trazem avanços nestes dois aspectos. No primeiro, o surgimento de processadores já conhecidos, porém com velocidade de processamento mais elevada. No segundo aspecto as diferenças estruturais podem ir desde as mudanças na montagem e conexão de dispositivos de um computador às relacionadas ao uso de vários computadores.

Uma das principais estratégias utilizadas para obtenção de maior desempenho computacional é a paralelização, seja através de arquiteturas paralelas, seja através da paralelização e ou distribuição de algoritmos matriciais. Outra

estratégia importante é a distribuição da computação em redes de estações de trabalho ou de computadores.

Diversos livros apresentam introdução ao tema, dentre os quais ressaltamos: Bertsekas & Tsitsiklis [101]; Carey , [176]; Golub & Ortega [400]; Hwang & Briggs [512]; Hwang & Xu [514]; Modi [759]; Petitet & Casanova & Dongarra & Robert & Whaley [869]; Schendel [945]; Amorim & Barbosa & Fernandes [26]; Van De Velde [1059] dentre outros.

Um bom desempenho de processamento paralelo começa a ser determinado pela qualidade do programa que está sendo executado. Para se ter um processamento computacional eficiente de um algoritmo, deve-se estruturá-lo convenientemente, bem como elaborar um programa que leve em conta a linguagem utilizada.

Por exemplo quando multiplicamos as matrizes **A** e **B** para obter a matriz **C** podemos fazer esta operação de seis maneiras distintas, Figura 3.1; na linguagem FORTRAN a seqüência *jki* é mais eficiente enquanto que na linguagem C a seqüência *ikj* apresenta melhores resultados.

O desenvolvimento de um programa distribuído e paralelo deve começar com a implementação eficiente de um programa sequencial.

Um computador de processamento paralelo pode ser simplesmente definido como uma máquina que utiliza mais de um processador simultaneamente para obter um resultado. Por necessidade temática divide-se processamento paralelo em três grandes áreas de pesquisa:

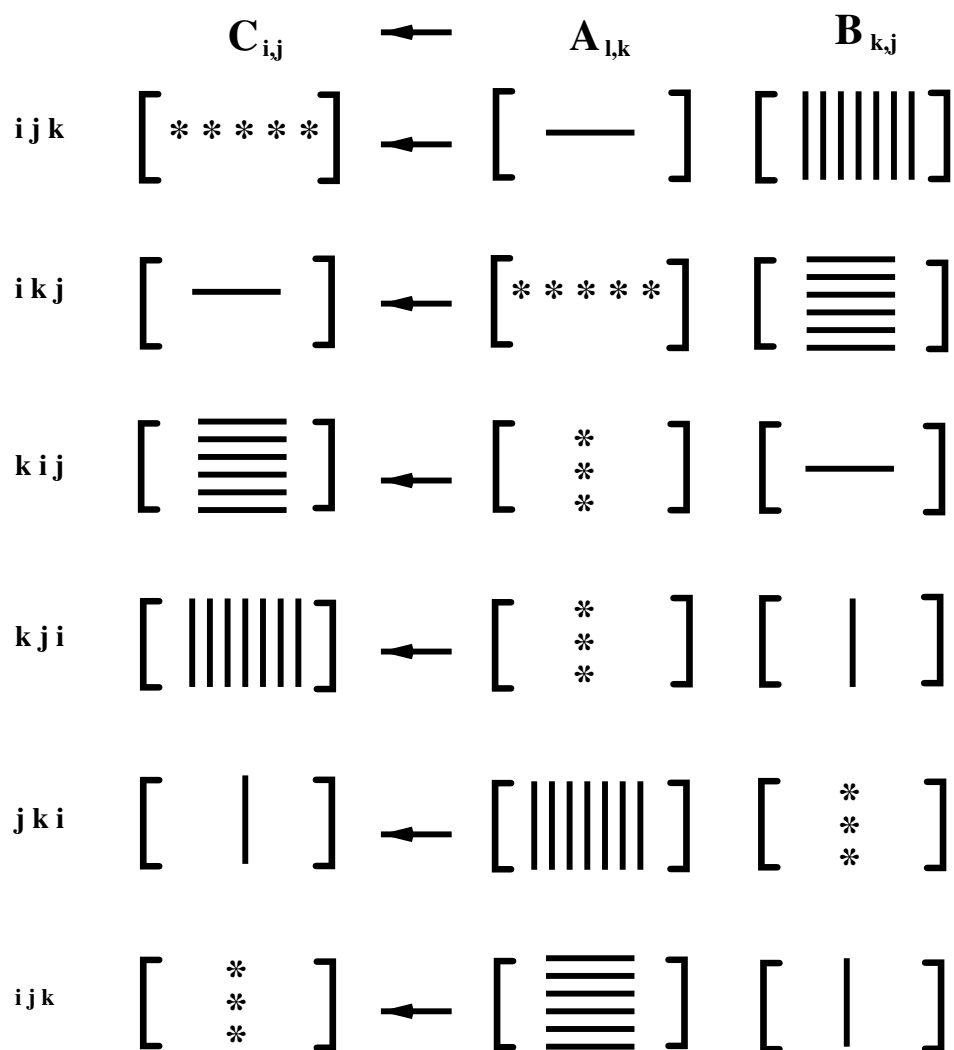
1. Algoritmos e suas aplicações,
2. Linguagens de programação,
3. Arquiteturas (incluindo sistemas operacionais e equipamentos (*hardware*)).

O modelo clássico usado para descrever métodos numéricos no contexto de computadores digitais sequenciais é conhecido universalmente como modelo de von Neumann <sup>1</sup>; Schendel [945]. O computador de von Neumann é composto dos módulos:

---

<sup>1</sup>Aqui surgem polêmicas sobre o pai dos computadores digitais no entanto aqui estamos nos referindo a “ arquitetura da máquina de Neumann ”:

1. Konrad Zuse, um dos pioneiros da ciência da computação, morreu aos 85 anos de idade, no dia 18 de dezembro de 1995. No Brasil, pouca gente ouviu falar dele, mas as suas conquistas foram tão impressionantes (e pouco reconhecidas pela história “ oficial ” da Informática). Zuse inventou, produziu e fez funcionar o primeiro computador digital binário de programa armazenado da história. Isso não é pouco, pois o que a maioria dos historiadores americanos considera como sendo o primeiro computador a funcionar, o ENIAC, inventado na Escola de Engenharia da Universidade de Pensilvânia, nos EUA, usava o sistema decimal e não tinha programa



```

para ? = 1 até n;
  para ? = 1 até n;
    para ? = 1 até n;
       $C_{i,j} \leftarrow C_{i,j} + A_{l,k} B_{k,j}$ 
    fim
  fim
fim

```

? : i, j ou k da sequência  
 \* : elemento da matriz  
 — : linha da matriz  
 | : coluna da matriz

Figura 3.1: Diferentes Modos de Multiplicação de Matrizes.

1. Processador: uma unidade lógica e aritmética e uma unidade de controle com um registro de controle de seqüência, RCS, que contém o endereço da próxima instrução e um registro de controle, RC, que contém uma cópia da instrução atual.
2. Programa: versão codificada de um algoritmo à ser implementada.
3. Memória: local onde os dados e as instruções do programa são mantidos.

Uma representação esquemática deste tipo de computador está ilustrada na Figura 3.2.

O computador universal utilizando um único processador obedece a seguinte seqüência de instruções:

1. Representação digital das variáveis.
2. Processamento sequencial de operações aritméticas e lógicas.

---

armazenado em memória (a programação era realizada com milhares de conectores, fios e seletores giratórios). Acontece que Zuse conseguiu fazer funcionar o seu primeiro computador em 1941, enquanto que o ENIAC realizou as suas primeiras computações somente em 1947, seis anos depois. No entanto, Zuse permaneceu desconhecido fora da Alemanha por muitos anos, por causa da II Guerra Mundial. O reconhecimento mundial ao seu pioneirismo veio somente nos anos sessenta.

2. J. Presper Eckert ajudou a inventar o primeiro computador, Eckert era instrutor na Universidade de Pensilvânia quando projetou e construiu, junto com John Mauchly, o primeiro computador eletrônico digital, pesando 30 toneladas. O computador em forma de U foi batizado de Eniac - calculadora e integradora eletrônica numérica - e inaugurado dia 14 de fevereiro de 1946 na Escola de Eletrônica de Penns Moore. O computador continha 18 mil válvulas a vácuo, 70 mil resistores e 10 mil capacitores. O Eniac podia resolver 5 mil problemas em um segundo e fazer cálculos que atualmente podem ser desenvolvidos por um computador do tamanho de um chaveiro. Em 1943, com 150 mil dólares de ajuda governamental, Eckert e seu companheiro da Universidade da Pensilvânia, John Mauchly, começaram a trabalhar numa calculadora de grande porte, da qual o exército norte-americano faria uso nos cálculos de distância necessários para a artilharia pesada. Os dois foram demitidos da universidade um ano após o Eniac entrar em uso, um fato seguido por uma acirrada disputa pelos direitos comerciais sobre a invenção. Os dois resolveram formar a sua própria firma, a Corporação de Computadores Eckert-Mauchly, que mais tarde foi comprada pela Remington Rand. Entre os projetos da dupla estava o Univac, um computador desenvolvido para o escritório do censo norte-americano, em 1951.
3. John Von Neumann propõe que um programa seja armazenado em um computador da mesma forma que são armazenados os dados. A proposta dele, chamada de "Arquitetura de Neumann", é a base para os computadores modernos.

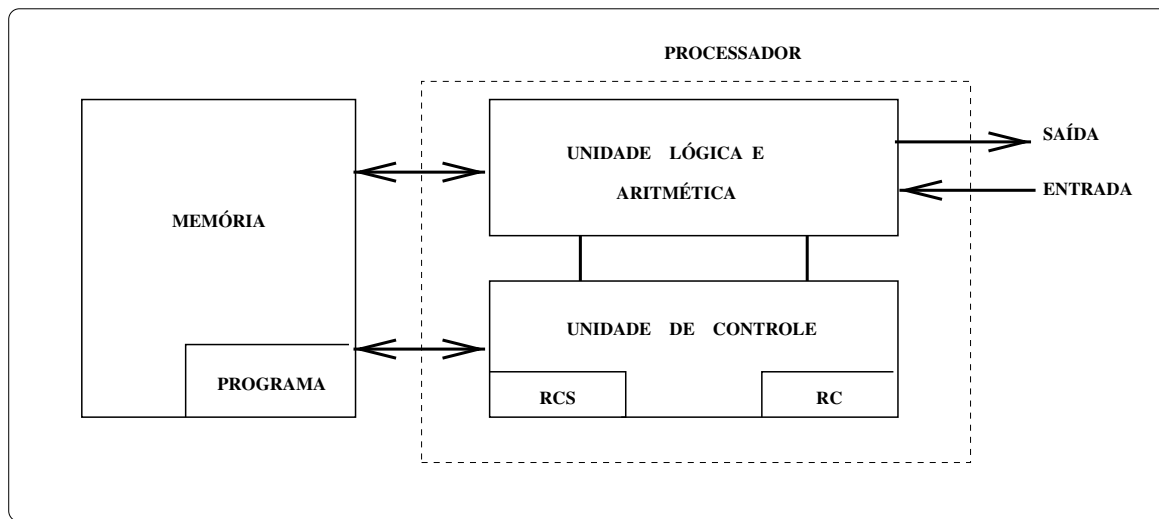


Figura 3.2: Representação Esquemática de Um Computador Universal.

Todo desenvolvimento computacional tem como ponto de partida este modelo básico. Em particular, os algoritmos de análise numérica são baseados nestes tipo de computador e exigem um grande número de operações elementares. Problemas relacionados com a geração de erros e a sua propagação, e questões de estabilidade numérica <sup>2</sup> ainda são temas de pesquisa.

Durante os últimos quarenta anos, os computadores sequenciais tiveram um aumento de desempenho computacional devido às novas tecnologias e ao desenvolvimento de projetos que incorporam características de paralelismo dentre as quais destacam-se:

1. A organização dos canais de entrada e saída de dados.
2. A sobreposição na execução das instruções.
3. As técnicas de acesso aos dados compartilhados.

No início dos anos 60 a possibilidade de utilização de máquinas paralelas começou a ser discutida. Computadores com características realmente paralelas foram desenvolvidos (CDC STAR 100, ILLIAC IV, STARAN, CRAY-1, Texas Instruments TI-ASC, IBM 2938 e CDC 7600). Uma análise mostrando a vantagem do processamento paralelo era necessária e ainda é de interesse

---

<sup>2</sup>Um algoritmo para determinar  $f(d)$  é numericamente estável se não for sensível a perturbações que são inerentes ao problema. A estabilidade assegura que a solução calculada está próxima da solução do problema que teve seus dados levemente perturbados; Klema & Laub, [602]; Stewart, [1000]; Kahaner, Moler & Nash, [560].

teórico a verificação, para um dado conjunto de problemas, de se existe um paralelismo máximo. Questões desta natureza estão no domínio da teoria da complexidade <sup>3</sup>.

A seguir apresentamos alguns aspectos favoráveis e outros contrários à utilização de computadores paralelos.

**Aspectos favoráveis:**

1. Aumento da velocidade computacional; por exemplo, estruturas convencionais têm limites físicos.
2. Possibilidade de solução de problemas complexos para máquinas seriais; por exemplo, previsão meteorológica.
3. Possibilidade de solução de problemas que têm paralelismo natural; por exemplo, operação com vetores.
4. Possibilidade de solução de problemas de sistemas de tempo-real <sup>4</sup>; por exemplo, processamento de imagens.
5. Possibilidade de operação contínua dada a falha em um dos processadores.

**Aspectos contrários:**

1. Pouca utilização da capacidade da máquina (problema gerencial). Um computador paralelo é um conjunto de processadores capazes de trabalhar cooperativamente, na solução de um problema computacional. Esta definição é ampla o suficiente para incluir supercomputadores com centenas ou milhares de processadores, redes de estações de trabalho, estações de trabalho multiprocessadas e sistemas acoplados ou interligados tipo hipercubo. Todos estes sistemas computacionais requerem

---

<sup>3</sup>Admite-se que complexidade se refere a sistemas com elevado número de componentes. Uma definição simples e meramente qualitativa (e também restritiva) para complexidade é a caracterização como comportamento rico e inesperado, com aspectos de coerência desobedecendo o princípio da superposição. Uma pergunta básica desta Teoria é “ Que tipos de problemas são intrinsecamente difíceis ? ” Esta é uma questão fundamental que tem preocupado um grande número de pesquisadores nos últimos anos. Para alguns dos problemas, algoritmos eficientes já foram descobertos, mas para muitos outros, a procura por algoritmos eficientes continua. Suspeita-se que para muitos dos problemas não existem algoritmos eficientes. A fundamentação desta suspeita é fornecida pela Teoria da Complexidade Computacional.

<sup>4</sup>Um sistema de tempo real é caracterizado por situações onde a atividade de processar uma dada informação, ou de responder a um estímulo de entrada externo, deve ocorrer num intervalo de tempo finito e determinado.

altos custos de manutenção e operação que são justificados por uma utilização eficiente da capacidade computacional disponível. Desta maneira existe a necessidade de verificar qual o tipo de processamento ideal e suas restrições para solução em um sistema paralelo e/ou distribuído. Esta característica permite um questionamento sobre vantagens e desvantagens do processamento paralelo utilizando um único processador ou um único computador paralelo ou um sistema distribuído de computadores.

2. Organização dos dados é complexa (acesso paralelo). O processamento paralelo requer uma interação entre processadores o que exige altas taxas de comunicação. Além disso programas paralelos reagem a eventos assíncronos, como a chegada de um “ pacote ” pela rede, ou o apertar de uma tecla; este fato introduz características de não-determinismo, e de difícil reprodução da entrada.

Estas dificuldades tornam necessários um formalismo mais rigoroso para o tratamento de programas paralelos e o desenvolvimento de métodos numéricos paralelos e distribuídos.

### 3.1.1 Medidas de Desempenho para o Processamento Paralelo

Um ponto importante em análise numérica paralela é avaliar o ganho de velocidade quando da utilização simultânea de  $p$  processadores. Para este propósito foi sugerida uma medida  $S_p(n)$ , chamada “ *speedup* ” que denominaremos ganho de velocidade de processamento ou razão de aceleração:

$$S_p(n) = \frac{T(n)}{T_p(n)} \quad (3.1)$$

a variável  $n$  é o número de processadores que pode ser função da dimensão da matriz do sistema que o algoritmo irá processar;  $T(n)$  será o tempo de computação ótimo em uma máquina sequencial e  $T_p(n)$  o tempo de computação ótimo em uma máquina paralela. Para termos um resultado aceitável é necessário comparar o “ *melhor* ” algoritmo sequencial com o “ *melhor* ” algoritmo paralelo. Esta é uma deficiência desta definição, pois o melhor tempo de processamento em uma máquina sequencial não é necessariamente ótimo, daí termos na literatura, Bertsekas [101], três possíveis definições para a estimação de  $T(n)$ :

1. Tempo requerido pelo melhor algoritmo conhecido.

2. Tempo requerido por um algoritmo padrão.
3. Tempo requerido por um único processador para executar o algoritmo paralelo em estudo. Nesta definição um único processador irá simular a operação de  $n$  processadores.

G.M. Amdahl propôs em 1967 um limite para o valor de  $S_p(n)$ . A lei de Amdahl, como é conhecida, diz o seguinte:

“ Se  $n$  é o número de processadores,  $s$  é o tempo despendido por um processador na parte sequencial de um programa (parte esta que não pode ser paralelizada) e  $p$  é o tempo gasto por um processador na parte do programa que pode ser completamente paralelizada, então a melhor razão de aceleração  $S_p(n)$ , que pode ser obtida é dada por  $\frac{1}{s}$ , onde o tempo total é normalizado para:  $s + p = 1$ ” .

Então:

$$S_p(n) \stackrel{\text{def}}{=} \frac{s + p}{s + \frac{p}{n}} = \frac{1}{s + \frac{p}{n}} \quad (3.2)$$

A razão de aceleração pode também ser expressa em função de  $f_s$ : fração sequencial do tempo de execução de um programa paralelo; então para a proposta de Amdahl  $f_s = s$ , e da equação (3.2) tem-se:

$$S_p(n) = \frac{1}{f_s + \frac{1-f_s}{n}} \quad (3.3)$$

Observe que o tempo de processamento paralelo,  $p$ , é independente do número de processadores  $n$ . Para criar uma relação entre o tempo de processamento paralelo,  $p$ , e o número de processadores  $n$ , uma outra proposta de definição para razão de aceleração foi feita por Gustafson , [428], e é representada na Figura 3.3, permitindo-nos escrever:

$$S_p(n) \stackrel{\text{def}}{=} \frac{(s + p n)}{(s + p)} = s + p n = n + (1 - n) s = n + (1 - n) f_s \quad (3.4)$$

onde o tempo total  $s + p = 1$  e  $f_s = s$ .

**Exemplo 3.1** Para 1024 processadores compara-se o comportamento da razão de aceleração de Amdahl com a de Gustafson. Observe que para uma pequena variação da fração sequencial do algoritmo, a lei de Amdahl produz uma grande variação na razão de aceleração, o que não ocorre com a de Gustafson, Figura 3.4, que produz resultados mais adequados.



Uma representação esquemática do que ocorre com as propostas de Amdahl e de Gustafson é representada na Figura 3.3; os tempos  $p$  de processamento são interpretados de forma diferentes por Amdahl e Gustafson; enquanto Amdahl considera que  $p/n$  é o valor mínimo do tempo obtido devido a utilização de  $n$  processadores, Gustafson considera que o tempo total do algoritmo será dado por  $np + s$  e a unidade padrão de tempo é dada por  $p + s$ .

Note que o tempo de execução no processador  $i$  é composto de duas componentes:

1. Tempo de processamento:  $T_{process.}(i)$
2. Tempo de comunicação:  $T_{comunic.}(i)$

Normalmente,  $\sum_{i=1}^n T_{process.}(i)$  é um valor que decresce com aumento de  $n$ , enquanto que  $\sum_{i=1}^n T_{comunic.}(i)$  aumenta.

Outras grandezas utilizadas, além da razão de aceleração, são a eficiência e a eficácia, para avaliação do desempenho de processamento paralelo.

A eficiência  $\varepsilon(n)$  do processamento paralelo é definida como a utilização média dos  $n$  processadores alocados para uma determinada tarefa. Esta razão é dada por:

$$\varepsilon(n) \stackrel{\text{def}}{=} \frac{S_p(n)}{n} \quad (3.5)$$

Observe que para o caso de termos apenas um processador,  $\varepsilon(1) = 1$ . Portanto  $\varepsilon(n)$  fornece uma medida do ganho de tempo de processamento obtido com a utilização de mais de um processador em paralelo.

Uma função custo para processamento paralelo pode ser definida como:

$$C(n) \stackrel{\text{def}}{=} n \frac{T(n)}{T(1)} = \frac{n}{S_p(n)} \quad (3.6)$$

onde  $C(n)$  relaciona o número  $n$  de processadores utilizados em paralelo no processamento, com a razão de aceleração obtida com estes  $n$  processadores. A motivação para a apresentação da função custo é a seguinte: se a razão de aceleração for  $S_p(n) \geq n$ , então o custo do uso de mais processadores será menor ou igual a 1. Por outro lado, se  $S_p(n) = 1$ , então o custo será igual ao número de processadores usados; Ghosal & Serazzi & Tripathi, [384].

Baseado na função custo podemos avaliar o grau de utilização dos processadores pela eficácia,  $\eta_p(n)$ , definida como:

$$\eta_p(n) \stackrel{\text{def}}{=} \frac{S_p(n)}{C(n)} = \frac{S_p^2(n)}{n} = \varepsilon(n) S_p(n) \quad (3.7)$$

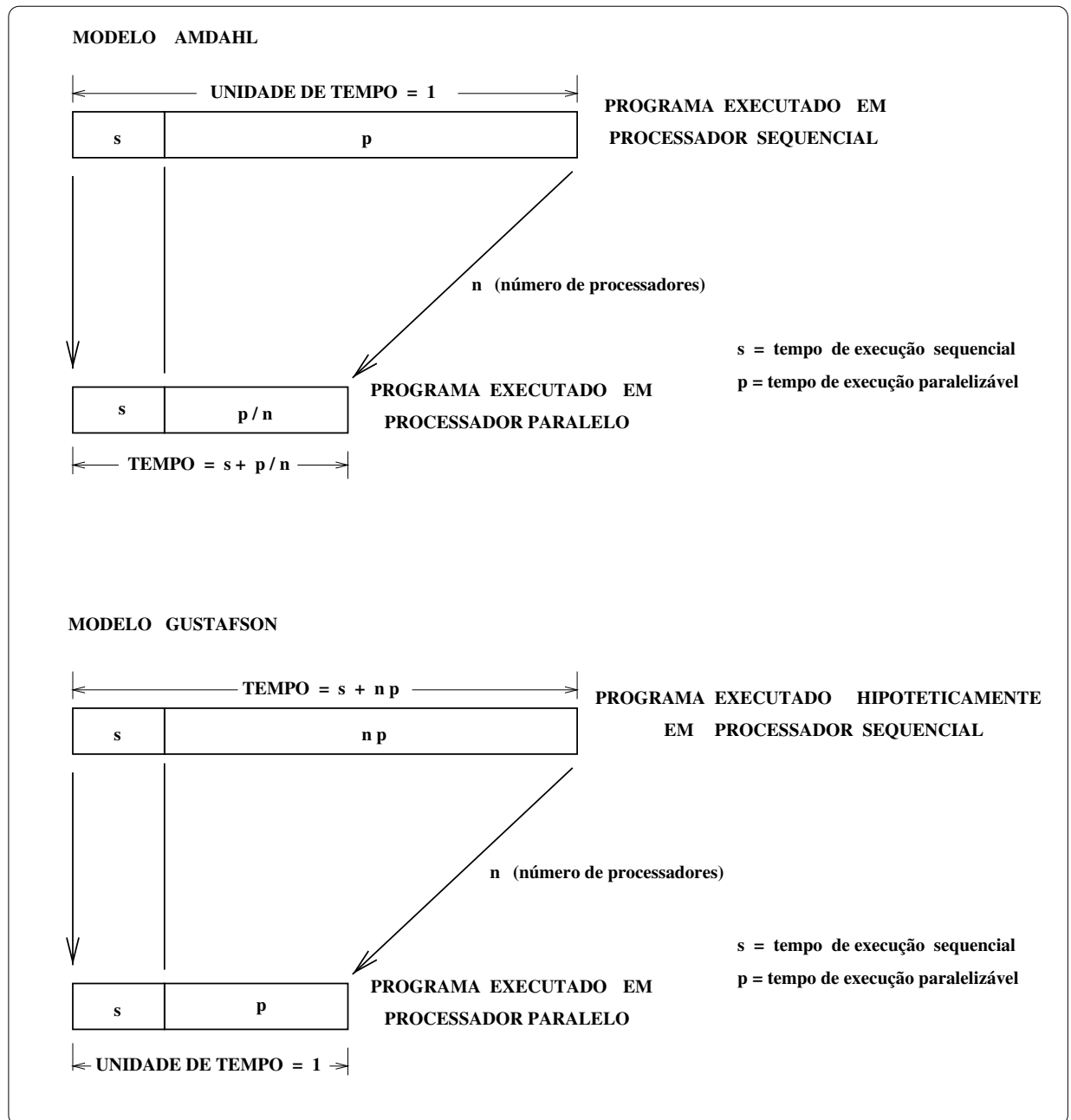


Figura 3.3: Modos de Avaliar a Razão de Aceleração de um Programa Computacional.

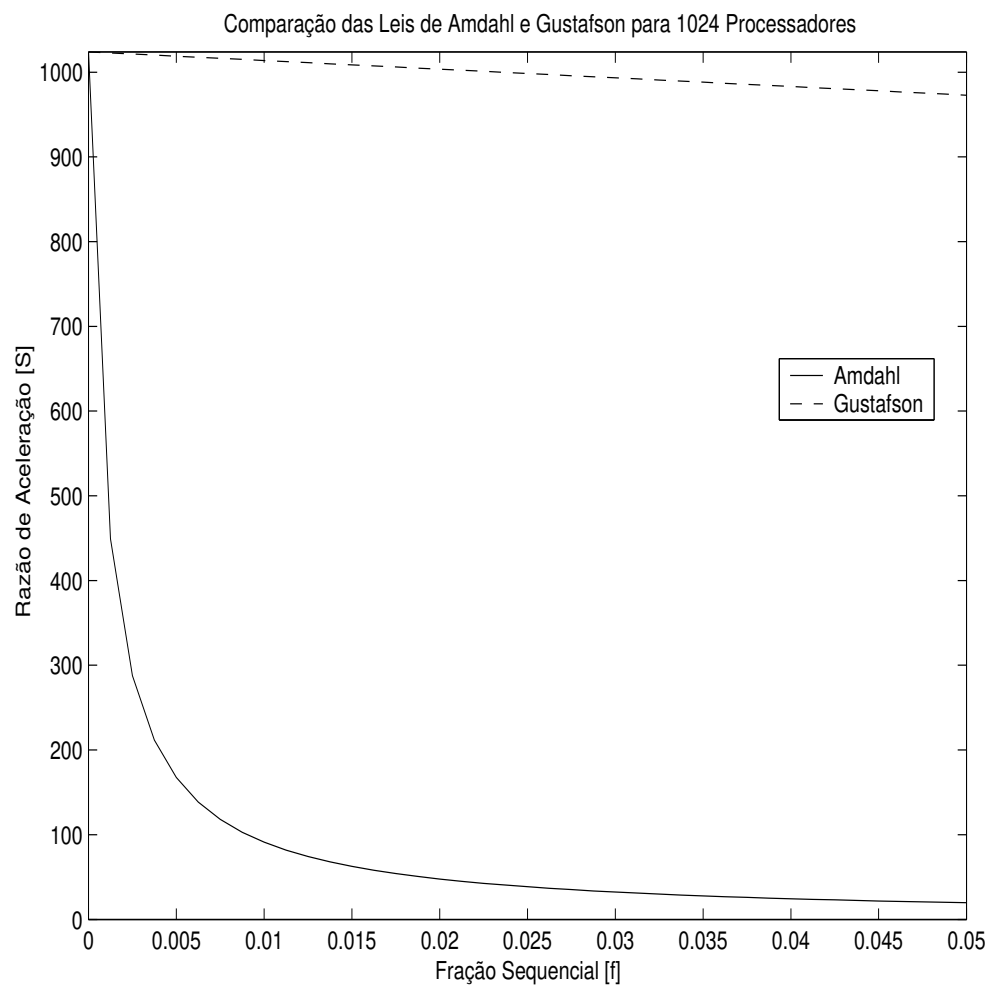


Figura 3.4: Comparação das Razões de Aceleração de Amdahl e de Gustafson.

A eficácia aumenta com  $n$  até um valor máximo e depois decresce, tendo portanto um valor máximo. Os valores máximos para a eficácia e para a razão de aceleração são obtidos para diferentes valores de  $n$ . A eficácia determina o número de processadores para o qual a relação entre o aumento da razão de aceleração e o aumento da eficiência é máxima.

Trabalhos sobre métodos de medidas de desempenho de computação paralela podem ser encontrados em Calzarossa & Serazzi [173] e Jean-Marie & Lefebvre-Barbaroux [534].

### 3.1.2 Arquiteturas Distribuídas e Paralelas

Arquiteturas, topologias de rede e modos de comunicação entre processadores paralelos e distribuídos são abordados nesta seção. A influência destes tópicos no desempenho computacional justifica o estudo das diversas possibilidades físicas de configurar processadores.

Das alternativas de arquitetura para processamento paralelo e distribuído, duas estão representadas na Figura 3.5. Outras possibilidades de arranjos de sistemas computacionais paralelos podem ser encontradas em Hwang e Xu [514].

Uma das classificações mais utilizadas para arquiteturas de máquinas paralelas foi apresentada por Flynn, [347]. Esta classificação para máquinas paralelas esta baseada em fluxos de dados e instruções:

1. SI ( *Single Instruction stream* ) - fluxo único de instruções.
2. MI ( *Multiple Instruction stream* ) - fluxos múltiplos de instruções.
3. SD ( *Single Data stream* ) - fluxo único de dados.
4. MD ( *Multiple Data stream* ) - fluxos múltiplos de dados.

A partir desta classificação quatro tipos de arquiteturas são desenvolvidos, Figura 3.6:

1. SISD - é o clássico “ modelo de von Neumann ”. Na arquitetura SISD ( *Single - Instruction Stream, Single-Data Stream* ) (Único Fluxo de Instruções- Único Fluxo de Dados) a sequência de instruções é executada em um único processador. Esta arquitetura compreende os computadores convencionais monoprocessadores. O termo sequencial é empregado para designar esta classe de computadores.
2. SIMD - Esta arquitetura é utilizada pelos computadores vetoriais. Todos os processadores executam a mesma instrução, fornecida pela unidade de controle, sobre vetores de dados, geralmente diferentes, como

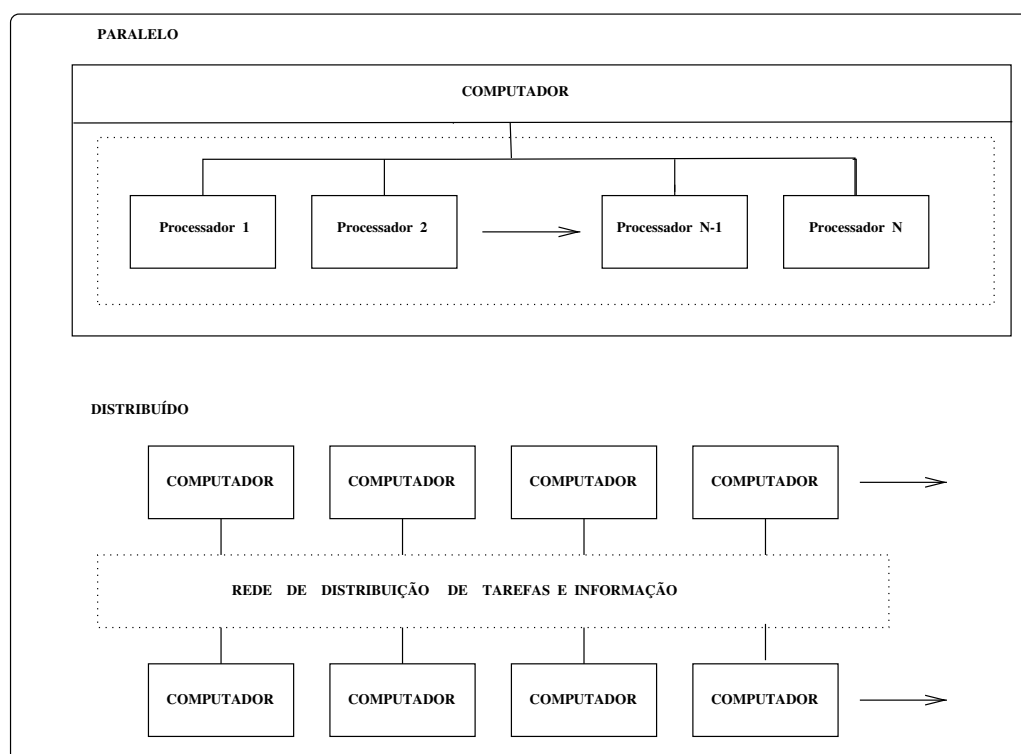


Figura 3.5: Alguns Tipos de Processamento Paralelo e Distribuído.

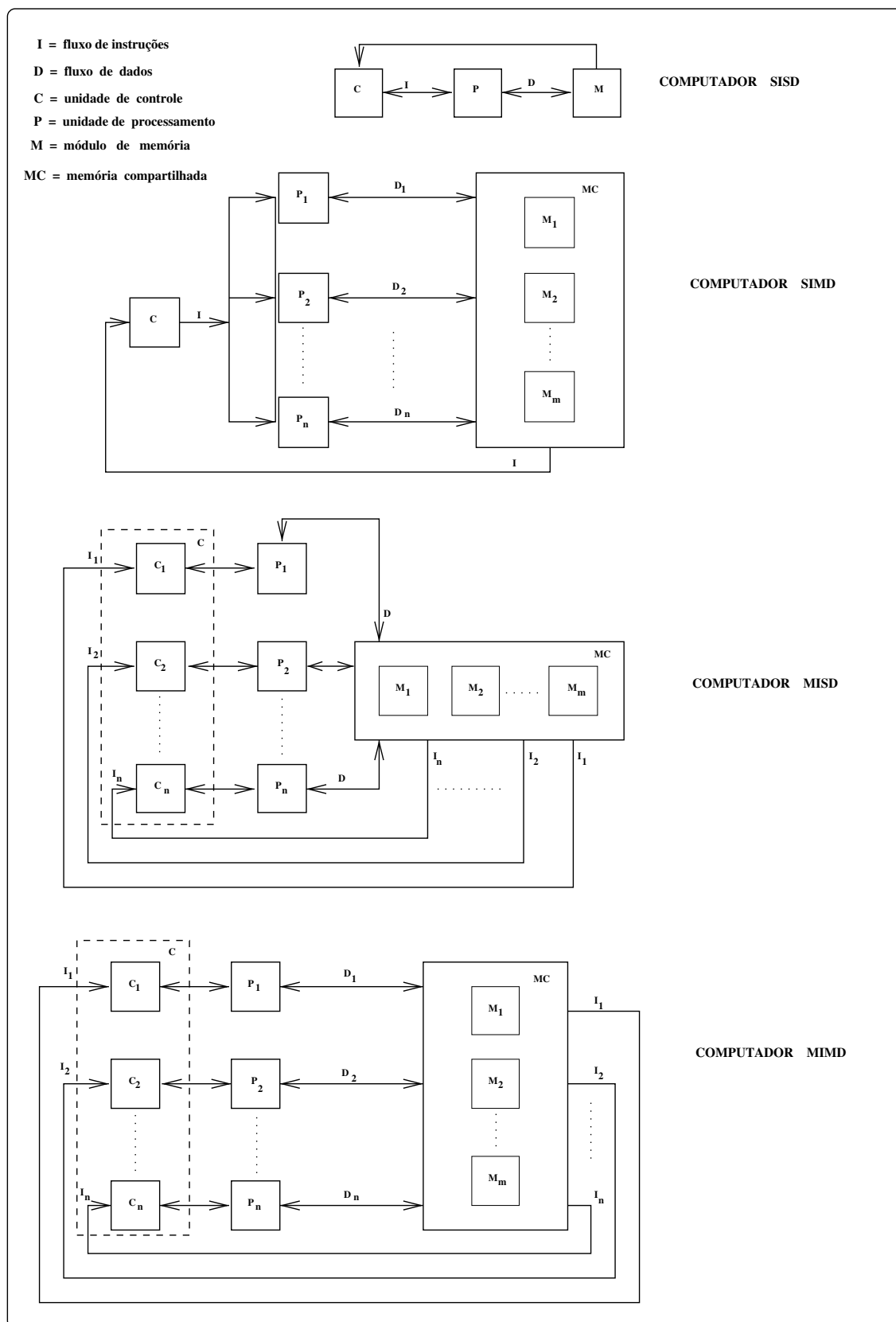


Figura 3.6: Classificação de Flynn para Organização de Computadores.

unidade de informação. A execução das instruções é síncrona no sentido de que cada processador executando uma instrução em paralelo deve terminá-la antes que a próxima instrução seja fornecida para execução.

3. MISD - Nesta arquitetura diversas instruções estão operando o mesmo conjunto de dados, daí tem-se duas maneiras diferentes de interpretar a organização de máquinas do tipo MISD. A primeira maneira considera uma classe de máquinas que requer que unidades de processamento distintas recebam instruções distintas e operem os mesmos dados. Esta classe de máquinas foi considerada por muitos arquitetos de computadores como pouco prática ou impossível sua implementação devido a não eficiência na transmissão de dados, e no presente não existe nenhum exemplo de computador trabalhando com este tipo de arquitetura. A outra maneira considera uma classe de computadores em que os dados circulam através de uma série de unidades de processamento. As arquiteturas do tipo *pipeline* tais como processadores dispostos de forma sistólica e vetorial são classificadas frequentemente sob este tipo. As arquiteturas *pipeline* processam um vetor em uma série de estágios, cada um executando uma função particular e produzindo um resultado intermediário. Por esta razão tais arquiteturas são classificadas como MISD; pode-se considerar que os elementos de um vetor pertencem ao mesmo conjunto de dados, e todos os estágios do *pipeline* representam as múltiplas instruções que estão sendo aplicadas ao vetor, Zargham, [1193].
4. MIMD - Esta arquitetura combina paralelismo nos fluxos de dados e de instruções; é composta de unidades de processamento capazes de realizar operações lógicas e aritméticas padrões. As unidades operam assincronamente sob o controle de fluxo de instruções individuais. Durante a computação as unidades comunicam resultados a outras unidades. Vários modelos de comunicação entre os processadores para MIMD têm sido propostos, Figura 3.7, Hockney e Jesshope, [492]. Nesta classificação de sistemas de computação estão incluídos os sistemas de multiprocessadores. Nos sistemas de rede, a memória é distribuída e os processadores são interconectados através de uma topologia definida.

Outra classificação de arquitetura de processamento paralelo que tem por base a independência entre os processadores é:

1. SMP - (*Symmetric Multi Processing*): em um único computador mais de um processador compartilham os recursos de memória e disco existentes, segundo uma política de controle de concorrência adotada pelo

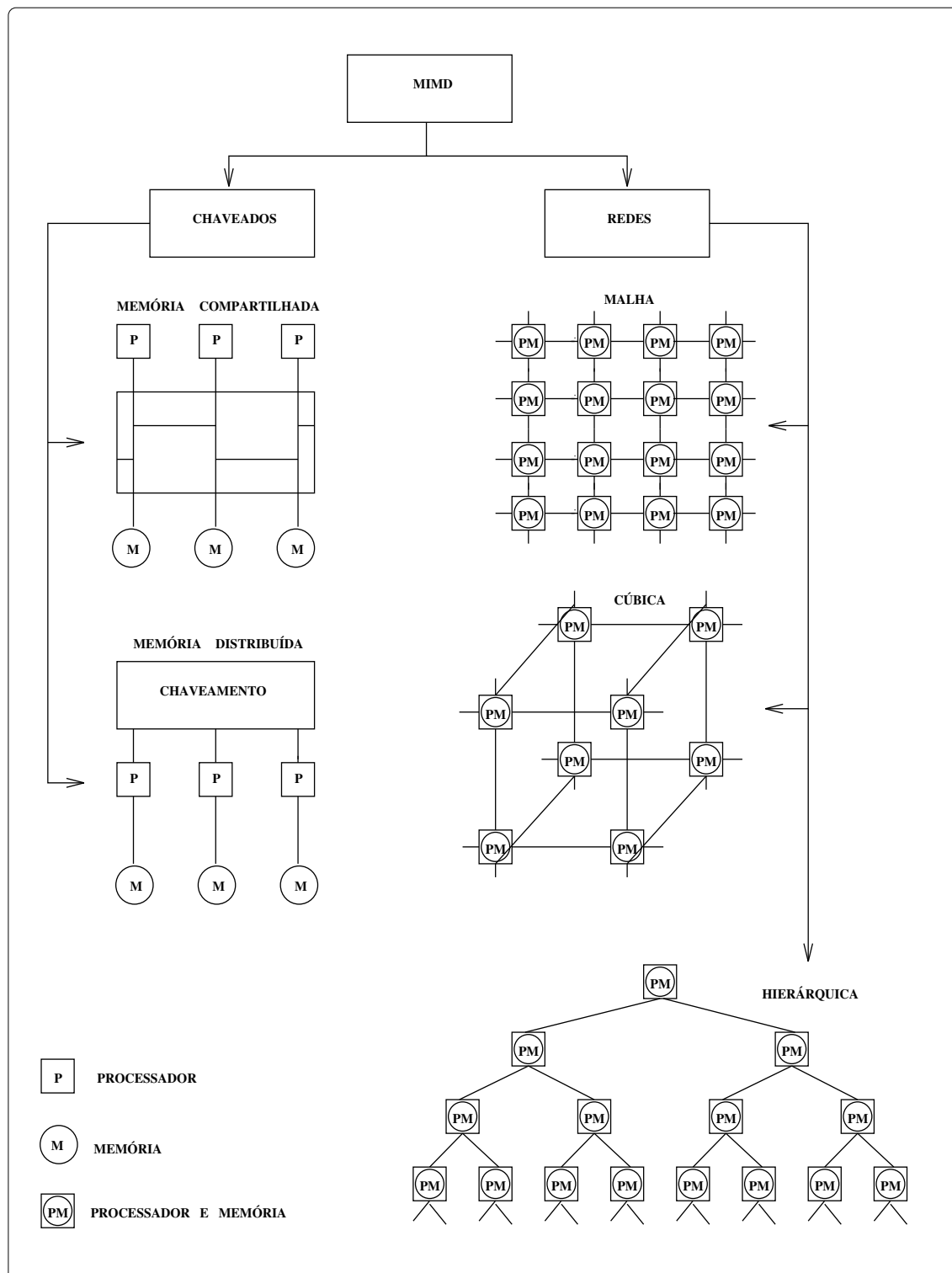


Figura 3.7: Classificação de Hockney para Comunicação Entre Processadores em Computadores Tipo MIMD.



sistema operacional. Esta complexa arquitetura é bastante transparente para o desenvolvimento de aplicações, pois a maior parte da complexidade fica a cargo do sistema operacional.

2. MPP - ( *Massive Parallel Processing* ): possui processadores com maior independência entre si, havendo pouco ou nenhum compartilhamento de recursos. Tipicamente cada nó <sup>5</sup> de um sistema MPP é um computador independente, com memória e discos próprios. O controle do paralelismo é feito pela aplicação, que deve coordenar a distribuição de tarefas e a coerência entre os diversos nós.

Uma classificação generalizada de arquiteturas para processamento paralelo é apresentada na Figura 3.8, Zargham, [1193]. Observe que estão representadas arquiteturas de rede neural artificial e de lógica nebulosa; este fato caracteriza a dinâmica de criação de novas arquiteturas.

Uma outra classificação de computadores foi proposta por Gajski [363]. Neste referência comenta-se que o desenvolvimento nas arquiteturas tem um relacionamento entre a aplicação e as limitações das máquinas. A arquitetura para processamento paralelo MIMD-multiprocessador deve refletir soluções particulares dos seguintes problemas:

1. Modelo computacional,
2. Particionamento dos programas,
3. Mecanismos de sincronização,
4. Acesso à memória,
5. Escanolamento de tarefas.

Estes problemas são básicos em ambientes de processamento paralelo.

A topologia de uma rede é o padrão de conexão utilizado para interligar um conjunto de entradas a um conjunto de saídas. Nos multicomputadores do tipo SIMD ou MIMD a rede define a topologia de conexão dos processadores. Em função da topologia, a rede de computadores pode ser classificada como:

1. Anel.
2. Árvore binária.
3. Hipercubo.

---

<sup>5</sup> “ nó ” : ligação ou cruzamento ou ponto que capta e transfere sinais de transmissão de diferentes direções em uma rede de computadores , por exemplo na Internet.

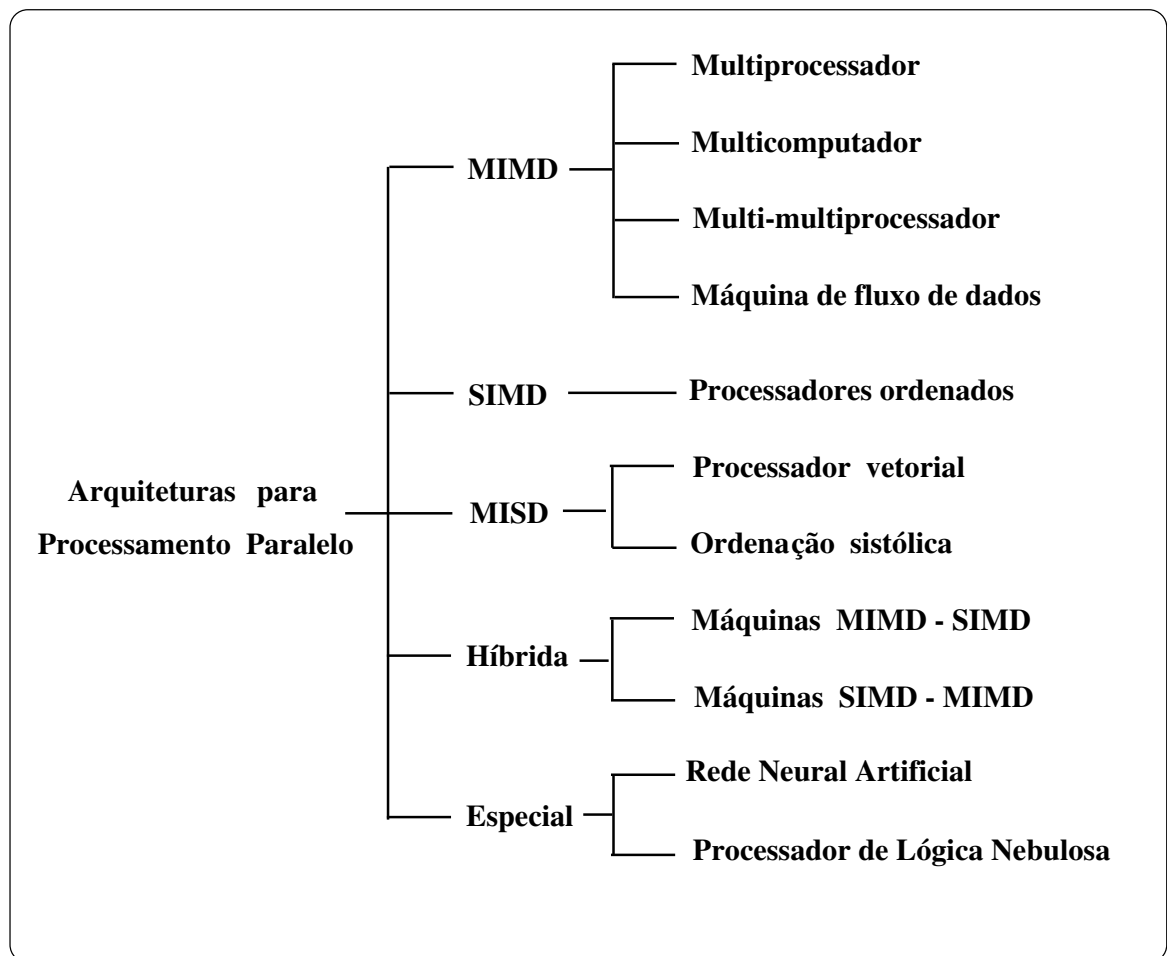


Figura 3.8: Classificação de Arquiteturas Para Processamento Paralelo.

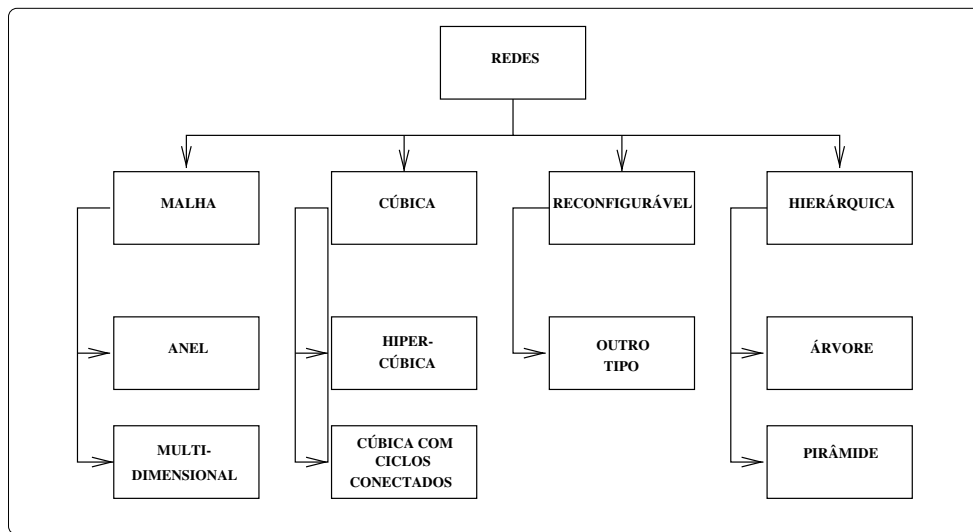


Figura 3.9: Topologias de Rede.

4. Cubo com ciclos conectados.
5. Sistemas multiprocessadores com memória múltipla, entre outros.

Exemplos de topologias de rede são apresentados na Figura 3.9, Hwang & Xu ,[514].

É importante notar que a seleção da topologia será influenciada pelos seguintes fatores:

1. Dimensão da topologia.
2. Capacidade de expansão.
3. Faixa de passagem (relacionada com a velocidade de transmissão).
4. Tecnologia de construção.
5. Distribuição espacial.
6. Tolerância à falhas.
7. Aplicações.

Para um aprofundamento em cada um destes itens descritos recomendamos Costa Filho [220].

### Tipos de Comunicação Entre Processadores

Uma vez que dois ou mais processadores operam simultaneamente, precisamos de meios para sincronizá-los e estabelecer comunicação entre eles; daí o surgimento dos seguintes conceitos:

- 1) Concorrência - Refere-se ao fato da existência de diversas atividades ocorrendo de forma paralela. Quando existir um mesmo número de processos e processadores, pode ficar garantida a execução concorrente dos processos; no entanto podemos ter concorrência aparente ou pseudo-concorrência (o tempo de processamento é compartilhado pelos processos: multiprogramação).
- 2) Comunicação - Sua existência permite a interação entre os processadores paralelos. A interação pode ser:
  - a) Assíncrona - é o tipo mais comum de comunicação sequencial ou por *modem*. Cada caracter vem entre *bits* de início e de fim, e a temporização entre os caracteres pode ser desigual. Na comunicação assíncrona, o processo que envia a mensagem não espera que haja um sinal de recebimento da mensagem pelo destinatário.
  - b) Síncrona - neste modo de transmissão, o receptor e o transmissor estão sincronizados quase que permanentemente, no entanto podem ocorrer perdas de sincronismo durante a transmissão. Na comunicação síncrona, o processo que envia a mensagem não retorna à execução normal enquanto não houver um sinal do recebimento da mensagem pelo destinatário.

Em relação a comunicação síncrona, o receptor conhecendo os intervalos de tempo representativos dos *bits*, identifica a seqüência de *bits* transmitida, fazendo amostragem do sinal recebido em intervalos regulares de T segundos. Essa temporização básica corresponde ao *clock* de relógio de período T segundos que estabelece a taxa ou velocidade de transmissão  $1/T$ , expressa normalmente em *bits* por segundo (bps). Para que ocorra a sincronização e esta amostragem possa ser feita nos instantes apropriados - mesma cadência de emissão dos *bits* - deve ocorrer o envio de uma configuração de *bits* de sincronização no início do bloco a ser transmitido. Contudo, se o bloco é muito extenso, algumas máquinas costumam ressincronizar seus osciladores, enviando, no interior do bloco, caracteres de sincronização. Este caracter deve ser precisamente diferenciado dos demais para não haver ressincronização em momento errado.

Observa-se que o modo de transmissão síncrono é mais utilizado quando as máquinas transmitem sua informação continuamente na rede, oferecendo uma utilização mais eficiente. Neste modo, os *bits* de um caracteres são seguidos imediatamente pelos próximos, não havendo delimitadores de caracter (*start/stop bits*) como na transmissão assíncrona. A montagem desses blocos de transmissão (com tamanho fixo ou variável) exige o uso de *buffers* para acumular as informações e armazená-las antes e depois da comunicação.

Com relação à comunicação síncrona tem-se:

Vantagens:

1. Em relação à transmissão assíncrona, é mais eficiente, pois a proporção de informação para sinais de controle (sincronização) é bem maior, não necessitando de sinais de início e fim de caracter (*start/stop bits*).
2. Facilita o uso de algoritmos de compactação devido ao armazenamento em *buffer*. Isto permite aumentar a velocidade de transmissão.
3. A transmissão síncrona oferece melhor proteção contra erros, pois existe no final da transmissão um conjunto de caracteres para verificação de erros.

Desvantagens:

1. Se há um erro de sincronização, todo o bloco é perdido, pois até a ressincronização a amostragem será realizada em instantes incorretos.
2. Exige o uso de *buffers*, o que encarece o custo do equipamento, pois os caracteres devem ser enviados em blocos e não conforme sua disponibilidade.

A importância do ponto de sincronização deve-se ao fato de que em um determinado instante várias tarefas deverão encerrar suas atividades, dado que uma nova tarefa dependente das tarefas anteriores deverá iniciar suas atividades. Deste modo é conveniente que as tarefas anteriores terminem sua execução em tempos próximos, maximizando a utilização dos processadores utilizados.

O tipo de comunicação é dependente da arquitetura de memória utilizada no ambiente computacional, que por sua vez, irá influenciar na maneira de escrever um programa paralelo.

### 3.1.3 Praticabilidade de Algoritmos Distribuídos e Paralelos

Nesta seção descreve-se diretrizes para o desenvolvimento e implementação de algoritmos distribuídos e paralelos, a isto entende-se por praticabilidade. Para termos um algoritmo paralelo adequado é necessário satisfazer uma série de pré-requisitos, tais como:

- a) O algoritmo precisa ser numericamente estável.
- b) O algoritmo precisa ser eficiente, isto é necessita convergir para a solução com um número de operações computacionais o menor possível. O número de operações de ponto flutuante que o mesmo irá exigir para se obter uma solução é conhecido também como o desempenho computacional de um algoritmo. O número de operações de ponto flutuante é usualmente conhecido por *flops*. Alguns autores consideram de modo aproximado o número de adições igual ao número de multiplicações e não consideram na totalização do número de *flops* operações envolvendo zeros.
- c) O espaço de memória requerido para sua execução deve ser o menor possível.
- d) A organização do algoritmo precisa ter um fluxo de informação adequado, por exemplo a razão de aceleração, equações (3.2) e (3.4), a eficiência, equação (3.5), a eficácia, equação (3.7) nem sempre terão um valor fixo, dado que o fluxo de informação existente em uma rede de computadores será dependente da sua taxa de utilização.
- e) A organização dos processos deve permitir um eficiente balanceamento nas atividades paralelas e distribuídas a serem executadas por diferentes processadores; ou seja, as tarefas devem demandar tempos de utilização de mesma ordem de grandeza.
- f) O programa precisa ter uma certa portabilidade, isto é, ele deve ser executado em uma variedade de máquinas sem nenhuma alteração e produzir resultados idênticos; isto nem sempre é possível.
- g) O algoritmo precisa ser programado de forma estruturada e documentada.
- h) A programação do algoritmo deve levar em conta a linguagem utilizada, Figura 3.1.

Uma aplicação paralela é *grosseira* se o tempo de transmissão de dados é muito pequeno se comparado com o tempo de computação, caso contrário ela é considerada *refinada*. Esta comparação só deverá ser feita após efetuarmos uma série de procedimentos com o objetivo de minimizar o tempo de comunicação entre os diferentes processadores. Para uma implementação paralela recomenda-se os seguintes passos:

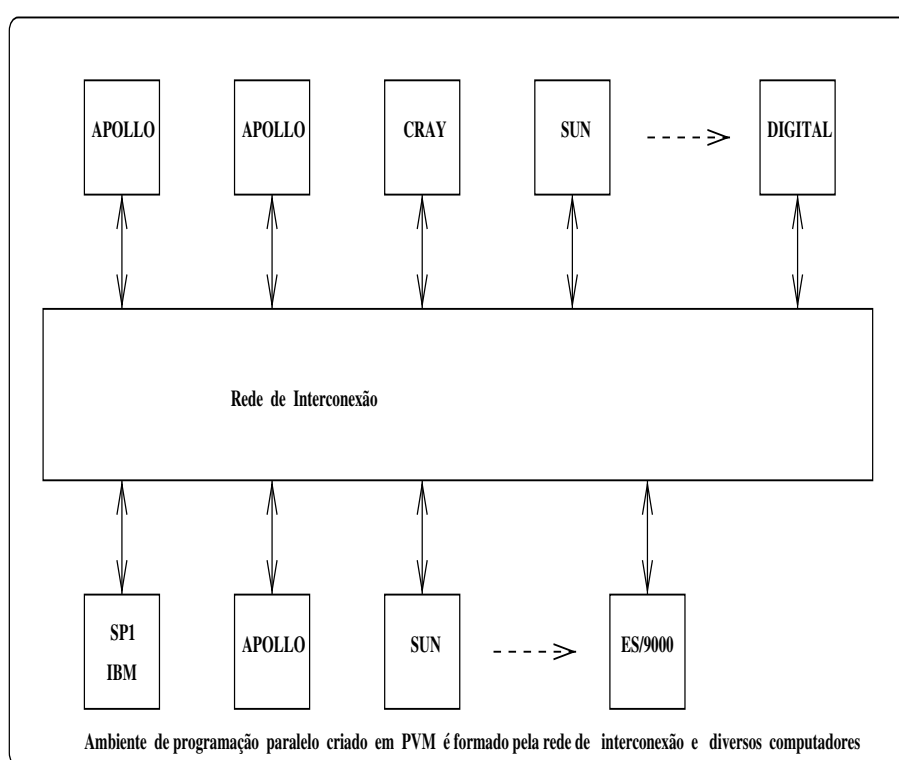
- a) Inicie as tarefas em processadores diferentes uma única vez.
- b) Maximize o trabalho computacional de cada processador.
- c) Quando possível inicie um número de tarefas maior que o número de processadores.
- d) Envie o mínimo de dados possível para cada processador.
- e) Quando possível utilize transmissão assíncrona.

## 3.2 Ambientes de Programação Distribuída e Paralela

A pesquisa e o desenvolvimento atuais em computação distribuída e paralela tem enfatizado modelos computacionais, versões paralelas de algoritmos, arquiteturas de máquinas e aplicações. O desenvolvimento de ambientes padrão de programação para execução paralela e distribuída dentro de uma dada arquitetura de máquina consolidou-se devido a um esforço conjunto de fabricantes e pesquisadores em ambientes tais como: *PVM* - (*Parallel Virtual Machine*) e *MPI* (*Message Passing Interface*). Nesta tese os algoritmos propostos foram implementados utilizando estes ambientes de programação paralela e distribuída, *PVM* e *MPI*, que são de domínio público. Estes ambientes consistem basicamente de rotinas para o controle de sequência de tarefas, informativas, para manipulação com *buffers* de memória e de envio e recebimento de mensagens. Em conjunto com o ambiente *MPI* utilizou-se também rotinas paralelas de álgebra linear numérica *ScaLAPACK* - *Scalable Linear Algebra PACKage* com o objetivo de obter maior eficiência nos algoritmos paralelos propostos e implementados.

### 3.2.1 *PVM* - *Parallel Virtual Machine*

O *PVM* - *Parallel Virtual Machine* é ambiente de programação que fornece a possibilidade para que uma rede heterogênea de computadores sequenciais

Figura 3.10: Possível Arquitetura Paralela Criada em Ambiente *PVM*.



ou paralelos tenha comportamento de um único recurso computacional concorrente, Geist e outros [375].

Uma possível arquitetura em um ambiente *PVM* é apresentada na Figura 3.10, onde os recursos computacionais de diversos tipos de máquinas ( *Sun*, *Apollo*, *Cray*, *SP1 - IBM* (ambiente central paralelo), *ES/9000* (ambiente vetorial) ) podem ser conectados através do *PVM* possibilitando que um conjunto de máquinas heterogêneas interagindo e cooperando entre si possa ser visto como um único recurso computacional. Para tornar isto possível o *PVM* consiste de uma coleção de protocolos para implementar de maneira confiável e em uma dada sequência, a transferência de dados entre as diversas máquinas. O ambiente *PVM* permite a implementação de programas segundo o paradigma “ pai ” e “ filho ”; isto significa que diferentes computadores podem estar processando diferentes programas.

O paradigma “ pai ” e “ filho ” tem as seguintes atividades:

**Processo “ pai ”:**

1. Gerar o número de processos “filhos” necessários para execução das atividades paralelas e distribuídas.
2. Identificar cada um desses “filhos”.
3. Empacotar dados.
4. Enviar dados aos processos “filho”.
5. Aguardar chegada dos dados elaborados pelos processos “filho”.
6. Desempacotar dados que chegaram dos “filhos”.
7. Desativar os “filhos” que executaram as tarefas.
8. Finalizar as atividades, após a geração dos resultados solicitados.

**Processo “filho”:**

1. Receber a sua identidade.
2. Receber a identidade de seus irmãos, se eles existirem.
3. Receber dados do processo “pai”.
4. Desempacotar dados do processo “pai”.
5. Verificar se necessita dados de processos “irmão”; caso necessite:
  - (a) Esperar receber dados do processo “irmão”.

- (b) Desempacotar os dados do processo “irmão”.
  - (c) Utilizar os dados recebidos.
  - (d) Voltar ao item 5.
6. Efetuar sua tarefa.
  7. Empacotar os dados gerados.
  8. Enviar os dados gerados para os processos “irmão” que estão aguardando.
  9. Enviar os dados para o processo “pai”.
  10. Encerrar a sua atividade.

### 3.2.2 MPI - Message Passing Interface

O ambiente de programação *MPI - Message Passing Interface* é um conjunto de especificações para comunicação entre processadores, com objetivos similares aos do *PVM*, [986]. O MPI não tem suporte para a criação e destruição de tarefas. O número de tarefas é fixo desde o início do processamento, assim como o número de máquinas participantes. O sistema não inclui suporte para o conceito de máquina virtual. O *MPI* permite implementação do paradigma *SPMD - Single Program Multiple Data*, ao contrário do *PVM*, o que significa a existência de um único programa em todo o processamento paralelo; Snir, Otto, Huss-Lederman; Walker, e Dongarra, [986].

O MPI é uma biblioteca de subrotinas padronizada, cada fabricante de computadores é responsável por desenvolver e otimizar uma biblioteca *MPI* para o seu ambiente paralelo de processamento em linguagem que lhe for mais conveniente, por exemplo FORTRAN, C ou C++.

O *MPI* oferece um conjunto flexível de funções para a troca de mensagens ponto a ponto. Tanto as funções para enviar mensagens quanto as para receber podem ser bloqueantes ou não bloqueantes. A semântica das funções para comunicação em grupo é sempre bloqueante, para simplificar a implementação e é escolhida pelo programador dentre as seguintes possibilidades:

- a) *Standard* - o sistema decide se armazena a mensagem em um *buffer* ou se a envia diretamente.
- b) *Buffered* - o sistema sempre armazena a mensagem em um *buffer*.
- c) *Synchronous* - a função não retorna até a mensagem ser recebida.

- d) *Ready* - a função retorna com sucesso somente se alguma outra tarefa já anunciou que vai receber a mensagem.

Além das funções para a comunicação ponto a ponto, o *MPI* oferece também um grande conjunto de funções para as comunicações em grupo:

- a) *Broadcast*: um processo envia uma mesma mensagem para todos os processos do grupo.
- b) *Scatter*: um processo envia uma mensagem diferente a cada um dos outros processos do grupo.
- c) *Gather*: um processo recebe uma mensagem diferente de cada um dos outros processos do grupo.
- d) *Allgather*: todos os processos enviam uma única e diferente mensagem para todos os outros processos do grupo.
- e) *Alltoall*: todos os processos enviam diferentes mensagens para todos os processos do grupo.
- f) Rotinas de operações globais de redução: são utilizadas para calcular um resultado, que envolve dados distribuídos, para um grupo de processos.

O *MPI* suporta diferentes topologias de comunicação entre os processos; o sistema escolhe a melhor representação para enviar a mensagem através da rede. O programador pode, por exemplo, organizar seus processos em forma de árvore. A partir deste ponto, cada processo pode se comunicar com a hierarquia de rede sem especificar explicitamente o identificador de tarefa.

### 3.2.3 ScaLAPACK - Scalable Linear Algebra PACKage

A biblioteca de álgebra linear numérica paralelizada *ScaLAPACK* - *Scalable Linear Algebra PACKage* foi projetada especificamente para obter uma elevada eficiência para computadores que utilizam memória distribuída, Blackford, [120]. O *ScaLAPACK* é uma extensão da biblioteca *LAPACK* - *Linear Algebra PACKage* para computadores MIMD; nesta arquitetura, a hierarquia de memória inclui a memória de processamento dos outros processadores, além da hierarquia de registros, *cache* e memória local em cada processador. De forma similar à *LAPACK*, as rotinas da *ScaLAPACK* são baseadas em algoritmos bloco-particionados de modo a reduzir a frequência de movimento de dados entre diferentes níveis da hierarquia de memória.

A estrutura hierárquica do ambiente computacional *ScaLAPACK* contém os seguintes blocos:

1. *BLAS - Basic Linear Algebra Subprograms*: é uma biblioteca de subrotinas para álgebra linear que permite operações tais como multiplicação de vetor-matriz, multiplicação de matriz-matriz e produto interno.
2. *PBLAS - Parallel BLAS*: é uma biblioteca paralela cuja interface é similar a *BLAS*. A única diferença importante em relação a *BLAS* é a rotina para transposição matricial, pois esta operação é bastante complicada em ambientes de memória distribuída, Choi & Dongarra & Walker, [201].
3. *BLACS - Basic Linear Algebra Communication Subprograms*: é uma biblioteca de *message passing* projetada para álgebra linear com ênfase na multiplicação matriz-matriz.
4. *LAPACK - Linear Algebra PACKage*: é uma coleção de rotinas para solução de problemas lineares, problemas de mínimos quadrados e problemas de autovalores e autovetores.
5. *ScaLAPACK - Scalable Linear Algebra PACKage*: é uma biblioteca com rotinas paralelas para solução de equações de sistemas lineares, fatorações ortogonais, inversão de matrizes, solução de problemas de auto valores e autovetores, solução de problemas de mínimos quadrados e outros problemas de álgebra linear numérica.
6. *Primitivas de troca de mensagens*: ambientes de programação paralelo, por exemplo *PVM* ou *MPI*.

Uma representação da estrutura hierárquica da *ScaLAPACK* é apresentada na Figura 3.11. Os componentes abaixo da linha tracejada, chamados locais, são sequenciais, deste modo utilizam apenas um processador. Os componentes acima da linha tracejada, designados por global, são rotinas paralelas síncronas cujos argumentos incluem matrizes e vetores distribuídos. Nas rotinas da *ScaLAPACK*, toda comunicação entre os processadores ocorre dentro do PBLAS e do BLACS, de tal modo que o código de fonte da camada superior da *ScaLAPACK* funciona de forma muito semelhante à do *LAPACK*.

### 3.2.4 Algoritmos por Blocos

As rotinas do *ScaLAPACK* utilizam algoritmos particionados em blocos, isto significa que em cada passo importante do algoritmo um bloco de linhas

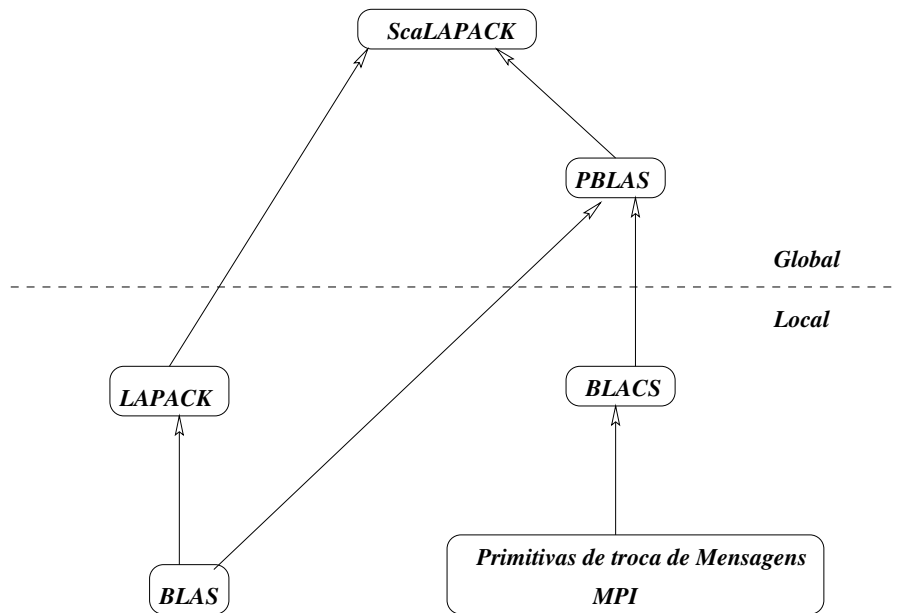


Figura 3.11: Estrutura do Ambiente Computacional Scalapack.

ou de colunas é atualizado, e a maioria da computação é executada por operações matriz-matriz nestes blocos. Os blocos são distribuídos entre os processadores que participam do processo numa faixa de blocos cíclicos de duas dimensões. As operações matriz-matriz são executadas por chamadas ao nível-3 *PBLAS* (*BLAS* Paralelo), que conta com as primitivas de comunicação fornecidas pelo *BLACS*. O desempenho de um algoritmo particionado em bloco varia até certo ponto com o tamanho do bloco, dito pelo seu número de linhas ou colunas.

### 3.2.5 Utilização de Algoritmos por Blocos

As rotinas do *ScaLAPACK* utilizam uma distribuição de blocos cíclicos de duas dimensões para todas as matrizes e vetores, para reduzir a comunicação de dados. Esta distribuição é tal que os blocos linha da matriz são distribuídos “empacotados” à linha associada a grade lógica de processadores e, do mesmo modo, blocos coluna são distribuídos “empacotados” à coluna associada à grade lógica de processadores. Aqui os termos bloco linha e bloco coluna referem-se a uma ou mais linhas ou colunas contínuas de uma matriz que são tratadas como uma única entidade do ponto de vista do algoritmo.

A Tabela 3.1 mostra como os blocos de uma matriz são distribuídos sobre uma grade lógica de processadores ( $2 \times 3$ ): a matriz tem 10 blocos linha e 10 blocos coluna; os índices linha e coluna na Tabela 3.1 indicam blocos linha

	1	2	3	4	5	6	7	8	9	10
1	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}
2	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}
3	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}
4	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}
5	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}
6	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}
7	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}
8	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}
9	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}	{0,1}	{0,2}	{0,0}
10	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}	{1,1}	{1,2}	{1,0}

Tabela 3.1: Distribuição de blocos sobre uma grade lógica de  $2 \times 3$  processadores.

e coluna, respectivamente. Cada caixa contém o índice do processador que armazena o bloco particular.

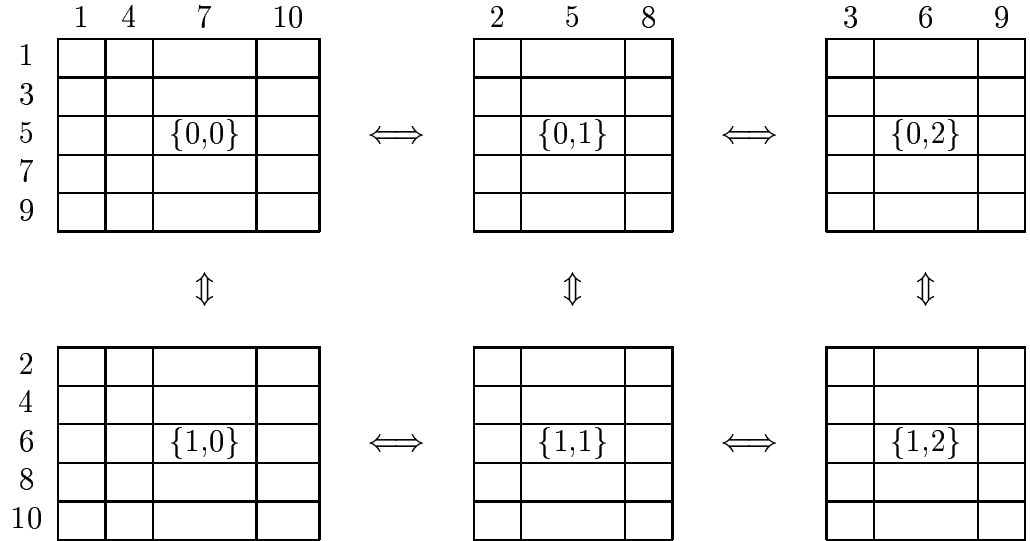


Tabela 3.2: Distribuição de Dados do Ponto de Vista do Processador.

A Tabela 3.2 mostra a mesma distribuição do ponto de vista do processador. Cada uma das caixas grandes na Tabela 3.2 é nomeada pelo índice do processador que armazena os blocos que a caixa contém. Os índices das linhas e colunas na Tabela 3.2 denotam os índices dos blocos linha e coluna, respectivamente MB e NB. Nesta terminologia, diremos que MB e NB são os fatores bloqueantes linha e coluna da distribuição da matriz.

### 3.3 Algoritmos Numéricos Paralelos

Nesta seção aborda-se alguns algoritmos básicos de álgebra linear numérica úteis para a implementação de métodos de modelagem de dados. São apresentados algoritmos paralelos básicos para solução do problema  $\mathbf{Ax} = \mathbf{b}$  utilizando rotações de Givens, reflexões de Householder e pseudo-inversa.

Características importantes para elaboração de algoritmo paralelo são obtidas através da análise da estrutura algébrica do respectivo algoritmo sequencial; este procedimento é utilizado nas propostas de paralelização apresentadas nesta tese.

#### 3.3.1 Solução Paralela do Sistema $\mathbf{Ax} = \mathbf{b}$

A solução de sistemas de equações lineares do tipo  $\mathbf{Ax} = \mathbf{b}$  é fundamental em muitos problemas.

Quando o problema envolve amostragem (caso de identificação de sistemas) é difícil a sua solução em um curto espaço de tempo. Então técnicas de computação paralela e distribuída estão sendo adotadas promissoramente para obtenção de solução eficiente, no tempo, deste tipo de problema, desde que estruturas paralelas adequadas sejam utilizadas. Para atingir este objetivo supomos que a série de pré-requisitos para que um programa paralelo seja elaborado, vista na seção 3.1.3, é adotada.

**Exemplo 3.2** Considere um sistema  $\mathbf{Ax} = \mathbf{b}$ , com  $\mathbf{A}_{3 \times 3}$  triangular superior, para o qual deseja-se estimar o valor do vetor  $\mathbf{x}$  utilizando um sistema paralelo de 3 processadores e o algoritmo paralelo de substituição para trás (back substitution), Bertsekas e Tsitsiklis [101].

Por hipótese, na Figura 3.12, o processador 1, gerenciador ou pai, para uma maior eficiência computacional, trabalha com uma linha da matriz  $\mathbf{A}$ . O processador 1 cria  $(n - 1)$  processos filho onde  $n$  é o número de linhas da matriz  $\mathbf{A}$ . O algoritmo para trás na forma paralela exige comunicação entre os processadores conforme ilustrado. Nota-se que vários intervalos de tempo não são utilizados, regiões sombreadas; por exemplo: intervalos de cálculo e transmissão de dados, constituindo tempos de espera; o tempo final de processamento é limitado pelo processador 1, que tem uma sobrecarga computacional no final do processamento, enquanto que os processadores 2 e 3 ficam ociosos. A atualização do vetor  $\mathbf{b}$  permite retirar a dependência da variável  $\mathbf{x}_i$  nas linhas 1, 2,  $\dots (i - 1)$ . Se considerarmos um sistema com  $n$  incógnitas, talvez não haja  $n$  processadores disponíveis e a diferença de “carga” computacional entre os processadores torna-se indesejável, daí a busca de uma solução que proporcione maior eficiência na distribuição de tarefas entre os processadores. A utilização eficiente dos processadores não ocorre neste

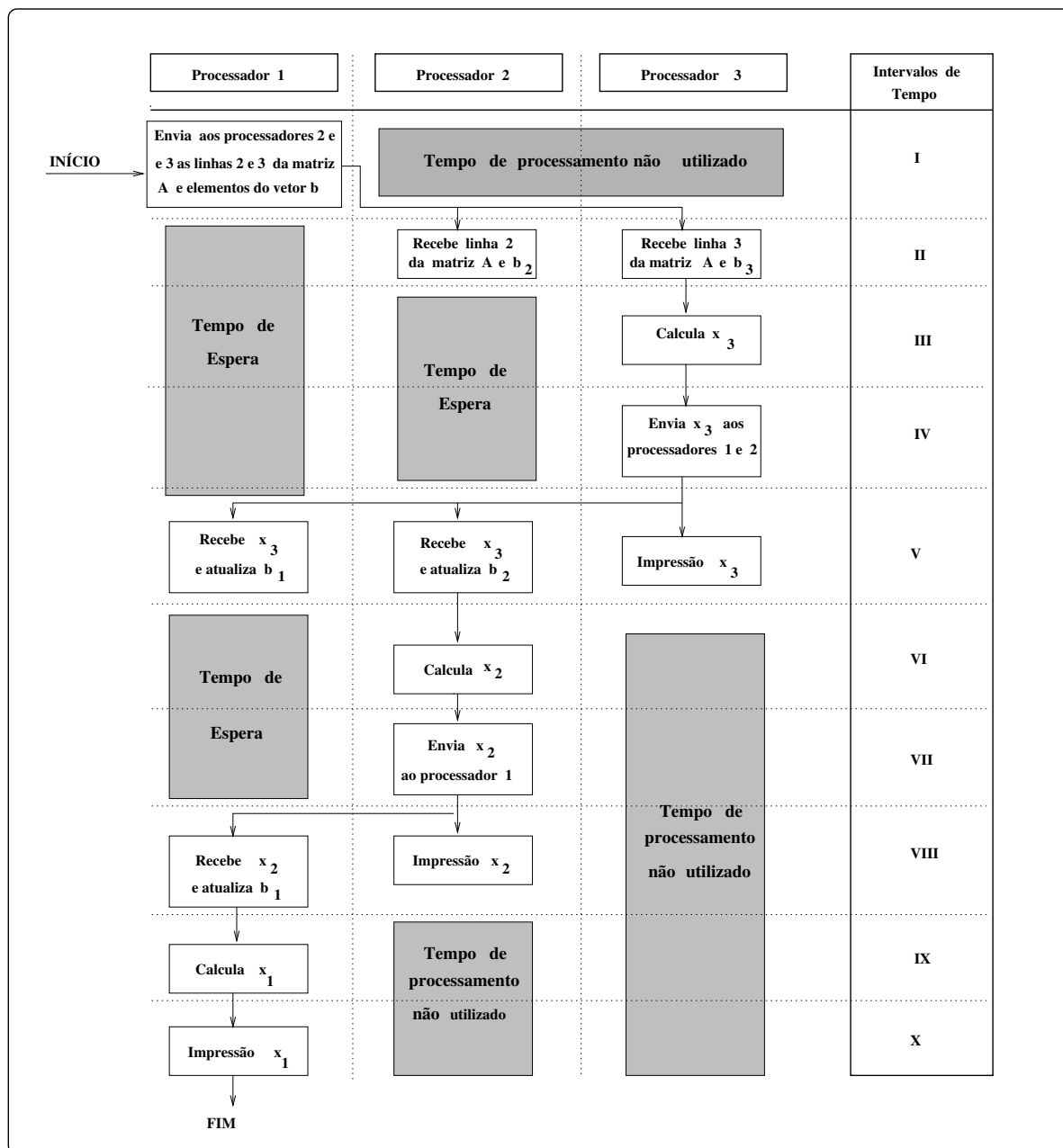


Figura 3.12: Solução Paralela Para Matriz  $3 \times 3$  Utilizando Algoritmo Substituição Para Trás.



*exemplo.* A solução do sistema  $\mathbf{Ax} = \mathbf{b}$  por substituição para trás, embora simples, não é conveniente por gerar uma estrutura paralela não balanceada. Além deste fato, em exemplos reais muitas vezes novos dados estão sendo gerados a cada instante, o que torna necessário que o procedimento numérico adotado trate os novos dados de forma contínua, sem necessidade de reiniciar cálculos a cada momento; por esta razão estratégias algébricas utilizando rotações de Householder, de Givens ou a pseudo-inversa, são utilizadas para superar as deficiências que o algoritmo de substituição para trás apresenta, permitindo a elaboração de algoritmos recursivos eficientes para implementação em ambientes paralelos.

**Exemplo 3.3** Como dito no exemplo anterior, uma motivação para a inserção do procedimento de Givens na solução de  $\mathbf{Ax} = \mathbf{b}$  é a chegada de um novo vetor com novas medidas a cada instante de amostragem.

A cada nova medida de entrada  $\mathbf{ne}$  teremos uma nova saída,  $\mathbf{ns}$ , e o sistema anterior poderá ser reescrito como:

$$\begin{bmatrix} \mathbf{R} \\ \mathbf{ne} \end{bmatrix} \begin{bmatrix} \mathbf{x} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{b}} \\ \mathbf{ns} \end{bmatrix}$$

Portanto para cada nova linha adicionada é necessário atualizar os parâmetros  $\mathbf{x}$  estimados anteriormente. A Figura 3.13 mostra a estrutura matricial deste conjunto de dados. Empregando rotações de Givens pode-se obter a solução do sistema de equações a cada instante sem necessidade de reinicialização do algoritmo.

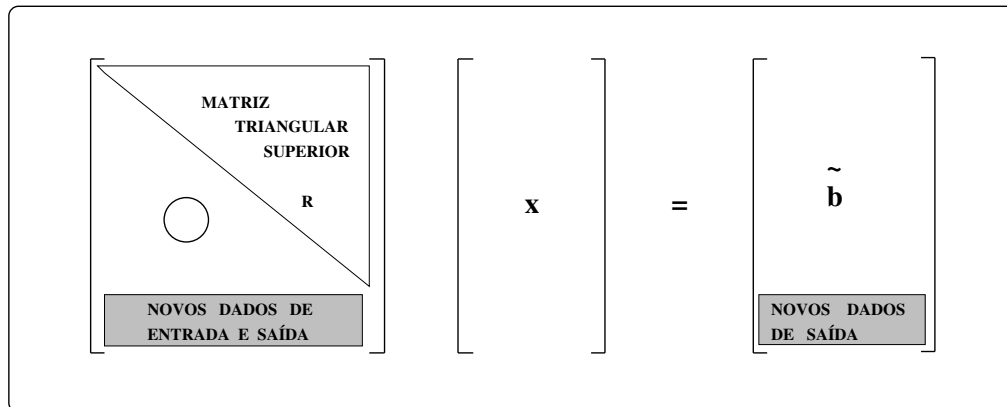


Figura 3.13: Estrutura Matricial Com Chegada de Um Novo Conjunto de Dados.

**Exemplo 3.4** A partir da utilização de transformações de Householder processar o sistema  $\mathbf{Ax} = \mathbf{b}$  de modo a ter  $\mathbf{Rx} = \tilde{\mathbf{b}}$  com  $\mathbf{R}$  triangular superior e  $\tilde{\mathbf{b}}$  o vetor  $\mathbf{b}$  original atualizado pelas matrizes de Householder,  $\mathbf{H}_k$ .

Fazendo  $\mathbf{H}_1 \mathbf{b} = \mathbf{b}_1$  e  $\mathbf{H}_1 \mathbf{A} = \mathbf{A}_1$  tem-se uma nova equação dada por:

$$\mathbf{A}_1 \mathbf{x} = \mathbf{b}_1$$

e assim sucessivamente até que  $\mathbf{A}_{n-1} = \mathbf{R}$ . A matriz triangular superior,  $\mathbf{R}$ , é obtida a partir da multiplicação das várias matrizes  $\mathbf{H}_k$  pelas respectivas matrizes  $\mathbf{A}_k$ . Ressaltamos que as colunas  $1, 2, \dots$  da matriz  $\mathbf{A}$  não se alteram após as transformações obtidas através das matrizes  $\mathbf{H}_1, \mathbf{H}_2, \dots$  respectivamente e que  $\tilde{\mathbf{b}}$  é o vetor obtido após diversas transformações efetuadas pelas matrizes de Householder em  $\mathbf{b}$ . Deste modo o sistema original  $\mathbf{Ax} = \mathbf{b}$  se torna  $\mathbf{Rx} = \tilde{\mathbf{b}}$ .

A estrutura paralela de um algoritmo de Householder é apresentada na Figura 3.14. Note que o processo pai é subutilizado, sendo usado apenas para: difusão da matriz  $\mathbf{A}$  inicial, atualização da matriz  $\mathbf{A}$  e do vetor  $\mathbf{b}$ . O desempenho computacional paralelo pode ser melhorado quando o número de processos filho criados é igual ao número de colunas menos um, sendo que o processo pai é responsável pela última coluna da matriz e pelo cálculo do vetor  $\tilde{\mathbf{b}}$ . Desta forma, os processos filho não precisam conhecer o vetor  $\mathbf{b}$ . A cada cálculo dos vetores de Householder  $\mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_{n-1}$ , equação (2.28), pelos processos filho, o processo pai recebe esta informação: vetor  $\mathbf{w}_i$ , para atualizar o vetor  $\tilde{\mathbf{b}}$  e a matriz  $\mathbf{A}_i$ .

**Exemplo 3.5** Obter a solução de um sistema  $\mathbf{Ax} = \mathbf{b}$ , transformado em  $\mathbf{Rx} = \tilde{\mathbf{b}}$ , utilizando a pseudo-inversa da matriz  $\mathbf{R}$ .

A solução é obtida aplicando um algoritmo recursivo para o cálculo da pseudo-inversa,  $\mathbf{R}^\dagger$ , da matriz  $\mathbf{R}$ , que é uma matriz triangular superior. Deste modo o sistema se torna  $\mathbf{x} = \mathbf{R}^\dagger \tilde{\mathbf{b}}$ , onde no vetor  $\mathbf{x}$  estão os parâmetros que deseja-se estimar; este algoritmo é apresentado na seção 3.3.2.

A obtenção da matriz pseudo-inversa de  $\mathbf{R}$  pode ser inicializada a partir do momento em que o procedimento de Householder conclui sua primeira coluna. A Figura 3.15 dá uma noção da utilização dos processadores quando do desenvolvimento do algoritmo recursivo para cálculo da pseudo-inversa.

Embora, o método de substituição para trás exija um desempenho computacional de  $\mathcal{O}(n^2)$  operações e o método recursivo da pseudo-inversa exija de  $\mathcal{O}(n^6)$  operações, esta diferença é compensada quando da utilização de processamento distribuído e paralelo.

### 3.3.2 Um Procedimento Recursivo Para Computação de Matriz Pseudo-Inversa

Para obter pseudo-inversa da matriz  $\mathbf{A}$  em forma recursiva um procedimento é apresentado a seguir inspirado em Janakiraman e Renganathan, [524].

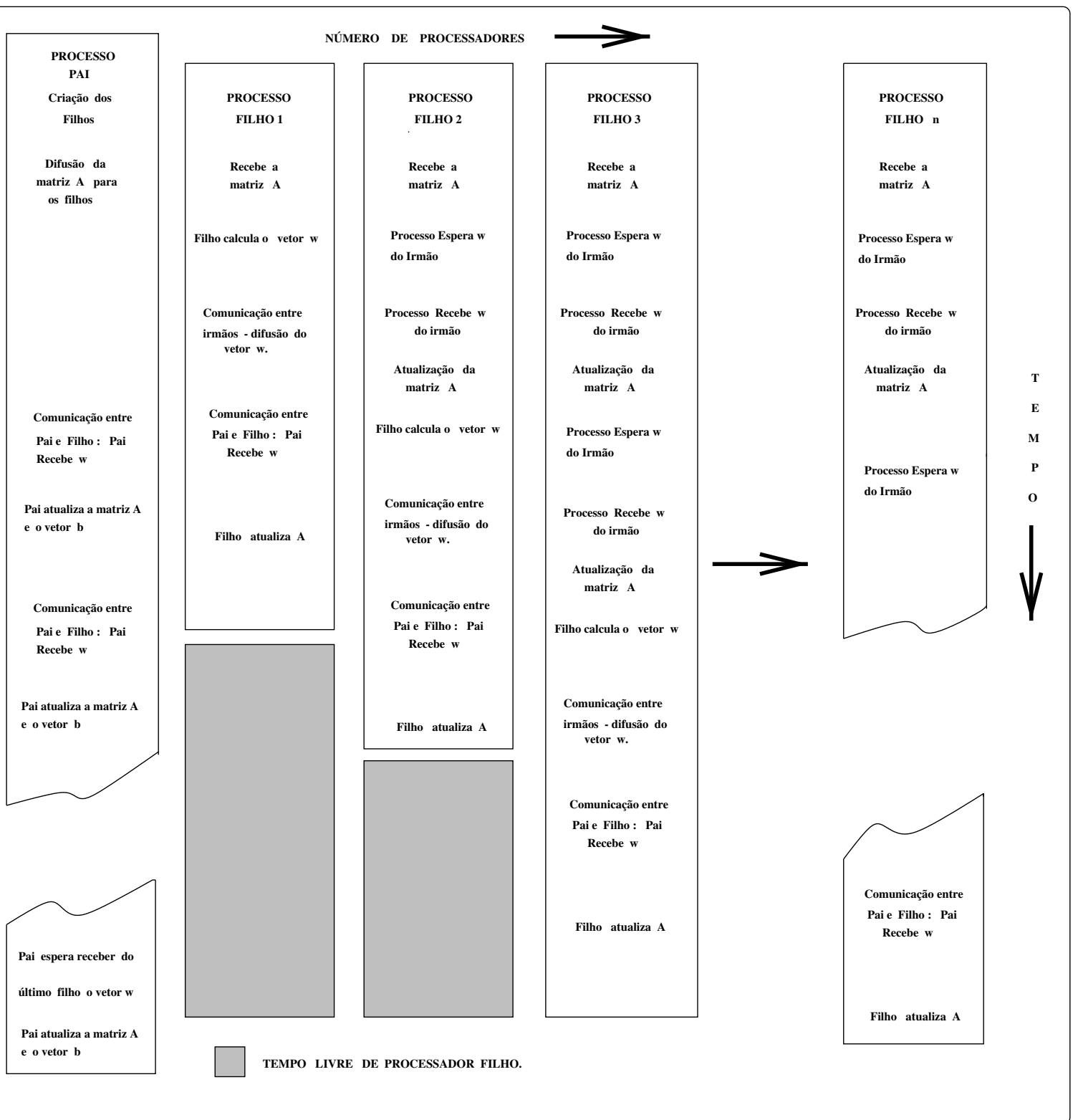


Figura 3.14: Estrutura Paralela de Um Algoritmo de Householder.



Seja uma matriz retangular ,  $\mathbf{A}_1$ ,  $m \times n$  de posto  $n < m$  , cuja pseudo-inversa é  $\mathbf{A}_1^\dagger$ :

$$\mathbf{A}_1^\dagger = [\mathbf{A}_1^T \mathbf{A}_1]^{-1} \mathbf{A}_1^T \quad (3.8)$$

Existe uma matriz  $\mathbf{A}_2$ ,  $m \times n$  de posto  $n$  que difere de  $\mathbf{A}_1$  por apenas uma linha:

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{c}_1^T \\ \mathbf{c}_2^T \\ \vdots \\ \mathbf{c}_{m-1}^T \\ \mathbf{c}_m^T \end{bmatrix} \quad \text{e} \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{c}_2^T \\ \mathbf{c}_3^T \\ \vdots \\ \mathbf{c}_m^T \\ \mathbf{c}_{m+1}^T \end{bmatrix} \quad (3.9)$$

onde  $\mathbf{c}_1^T, \mathbf{c}_2^T, \dots, \mathbf{c}_m^T, \mathbf{c}_{m+1}^T$  são vetores linha de dimensão  $n$ . As matrizes  $\mathbf{A}_1$  e  $\mathbf{A}_2$  têm  $(m-1)$  linhas idênticas e de modo análogo podemos escrever  $\mathbf{A}_2^\dagger$  como:

$$\mathbf{A}_2^\dagger = [\mathbf{A}_2^T \mathbf{A}_2]^{-1} \mathbf{A}_2^T \quad (3.10)$$

Um procedimento de cálculo, com o objetivo de obter a matriz  $\mathbf{A}_2^\dagger$  de modo recursivo, dado que em aplicações pode ocorrer a geração de uma nova linha de dados de entrada e saída de um processo a cada instante de amostragem é apresentado a seguir. Reescrevendo as equações (3.9) como:

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{a}^T \\ - - - \\ \mathbf{D}^T \end{bmatrix} \quad \text{e} \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{D}^T \\ - - - \\ \mathbf{b}^T \end{bmatrix} \quad (3.11)$$

onde  $\mathbf{a} = \mathbf{c}_1$  e  $\mathbf{b} = \mathbf{c}_{m+1}$ , um procedimento de cálculo, com o objetivo de obter  $\mathbf{A}_2^\dagger$  de modo recursivo, a partir de  $\mathbf{A}_1$  e  $\mathbf{A}_1^\dagger$  é desenvolvido a seguir.

Desde que  $m > n$ , a remoção ou adição de uma linha não irá mudar o posto  $n$  da matriz.

Seja

$$\mathbf{P}_1 = [\mathbf{A}_1^T \mathbf{A}_1]^{-1} \quad , \quad m \times m \quad (3.12)$$

Logo,

$$\mathbf{A}_1^\dagger = \mathbf{P}_1 \mathbf{A}_1^T \quad (3.13)$$

tem dimensão  $n \times m$ ; do mesmo modo tem-se:

$$\mathbf{P}_2 = [\mathbf{A}_2^T \mathbf{A}_2]^{-1} \quad (3.14)$$

com dimensão  $m \times m$ . Consequentemente,

$$\mathbf{A}_2^\dagger = \mathbf{P}_2 \mathbf{A}_2^T \quad (3.15)$$

com dimensão  $n \times m$ .

A seguir mostra-se que  $\mathbf{P}_2$  pode ser calculado a partir de  $\mathbf{P}_1$ , e que então pode-se calcular  $\mathbf{A}_2^\dagger$ . Das equações iniciais de  $\mathbf{A}_1^\dagger$  e  $\mathbf{P}_1$ , sabendo que  $\mathbf{P}_1$  é simétrica, pode-se escrever:

$$\mathbf{P}_1 = \mathbf{A}_1^\dagger \mathbf{A}_1^{\dagger T} \quad (3.16)$$

Das equações anteriores pode-se reescrever  $\mathbf{P}_1$  como:

$$\mathbf{P}_1^{-1} = \mathbf{A}_1^T \mathbf{A}_1 = \mathbf{a} \mathbf{a}^T + \mathbf{D} \mathbf{D}^T \quad (3.17)$$

De modo similar reescreve-se  $\mathbf{P}_2$ :

$$\mathbf{P}_2^{-1} = \mathbf{b} \mathbf{b}^T + \mathbf{D} \mathbf{D}^T = \mathbf{b} \mathbf{b}^T + \mathbf{P}_1^{-1} - \mathbf{a} \mathbf{a}^T \quad (3.18)$$

Fazendo,

$$\mathbf{R}^{-1} = \mathbf{P}_1^{-1} + \mathbf{b} \mathbf{b}^T \quad (3.19)$$

então

$$\mathbf{P}_2^{-1} = \mathbf{R}^{-1} + \mathbf{a} \mathbf{a}^T \quad (3.20)$$

Aplicando o lema de inversão matricial tem-se:

$$\mathbf{R} = \mathbf{P}_1 - \frac{\mathbf{P}_1 \mathbf{b} \mathbf{b}^T \mathbf{P}_1}{1 + \mathbf{b}^T \mathbf{P}_1 \mathbf{b}} \quad (3.21)$$

Então  $\mathbf{R}$  pode ser computada a partir de  $\mathbf{P}_1$  e  $\mathbf{b}$ .

Aplicando o lema da inversão matricial obtém-se:

$$\mathbf{P}_2 = \mathbf{R} + \frac{\mathbf{R} \mathbf{a} \mathbf{a}^T \mathbf{R}}{1 - \mathbf{a}^T \mathbf{R} \mathbf{a}} \quad (3.22)$$

As equações com as matrizes  $\mathbf{R}$  e  $\mathbf{P}_2$  nos dão a recursividade para o cálculo da matriz pseudo inversa a cada chegada de uma nova linha.

Como a forma da pseudo-inversa para este o problema de mínimos quadrados é  $\mathbf{A}^\dagger = [\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T$ , então após a convergência do cálculo teremos a pseudo-inversa desejada como:

$$\mathbf{A}^\dagger = \mathbf{A}_2^\dagger = \mathbf{P}_2 \mathbf{A}_2^T = \mathbf{P}_1 \mathbf{A}_1^T = \mathbf{A}_1^\dagger \quad (3.23)$$

Até aqui estudamos a obtenção de maneira recursiva de uma matriz pseudo-inversa quando da chegada de uma nova linha na matriz original;

vejamos agora como o método se comporta quando da chegada de uma nova coluna. Seja uma matriz retangular  $m \times n$ , de posto  $m < n$  cuja pseudo-inversa é  $\mathbf{A}_1^\dagger$ :

$$\mathbf{A}_1^\dagger = \mathbf{A}_1^T [\mathbf{A}_1 \mathbf{A}_1^T]^{-1} \quad (3.24)$$

Fazendo  $\mathbf{A}_2$ , uma nova matriz  $m \times n$ , de posto  $m$ , que difere de  $\mathbf{A}_1$  apenas por uma coluna, pode-se escrever:

$$\mathbf{A}_1 = [\mathbf{c}_1 \quad \mathbf{c}_2 \quad \dots \quad \mathbf{c}_{n-1} \quad \mathbf{c}_n] \quad \text{e} \quad \mathbf{A}_2 = [\mathbf{c}_2 \quad \mathbf{c}_3 \quad \dots \quad \mathbf{c}_n \quad \mathbf{c}_{n+1}] \quad (3.25)$$

onde  $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_n, \mathbf{c}_{n+1}$  são vetores coluna de dimensão  $m$ . As matrizes  $\mathbf{A}_1$  e  $\mathbf{A}_2$  tem  $(n-1)$  colunas idênticas e de modo análogo:

$$\mathbf{A}_2^\dagger = \mathbf{A}_2^T [\mathbf{A}_2 \mathbf{A}_2^T]^{-1} \quad (3.26)$$

Um procedimento de cálculo, com o objetivo de obter  $\mathbf{A}_2^\dagger$  de modo recursivo, a partir de  $\mathbf{A}_1$  e  $\mathbf{A}_1^\dagger$  é desenvolvido a seguir.

Desde que  $m < n$ , a remoção ou adição de uma coluna não irá mudar o posto  $m$  da matriz.

Fazendo,

$$\mathbf{A}_1 = \begin{bmatrix} \mathbf{a}^T \\ \vdots \\ \mathbf{D}^T \end{bmatrix} \quad \text{e} \quad \mathbf{A}_2 = \begin{bmatrix} \mathbf{D}^T \\ \vdots \\ \mathbf{b}^T \end{bmatrix} \quad (3.27)$$

onde  $\mathbf{a}^T = \mathbf{c}_1$  é a primeira coluna de dados e  $\mathbf{b}^T = \mathbf{c}_{n+1}$  é a última coluna de dados, pode-se calcular:

$$\mathbf{H}_1 = \mathbf{A}_1 \mathbf{A}_1^T = \begin{bmatrix} \mathbf{r}^2 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{R}^{-1} \end{bmatrix} \quad (3.28)$$

onde  $\mathbf{r}^2 = \mathbf{a}^T \mathbf{a}$ ,  $\mathbf{x} = \mathbf{D}^T \mathbf{a}$  e  $\mathbf{R}^{-1} = \mathbf{D}^T \mathbf{D}$ .

Fazendo,

$$\mathbf{H}^{-1} = \mathbf{P} = \begin{bmatrix} \mathbf{p}^1 & \mathbf{p}_2^T \\ \mathbf{p}_2 & \mathbf{P}_3 \end{bmatrix} \quad (3.29)$$

Como  $\mathbf{H}_1 \mathbf{H}_1^{-1} = \mathbf{I}$  tem-se:

$$\mathbf{H}_1 \mathbf{H}_1^{-1} = \begin{bmatrix} \mathbf{r}^2 & \mathbf{x}^T \\ \mathbf{x} & \mathbf{R}^{-1} \end{bmatrix} \begin{bmatrix} \mathbf{p}^1 & \mathbf{p}_2^T \\ \mathbf{p}_2 & \mathbf{P}_3 \end{bmatrix} = \mathbf{I} \quad (3.30)$$

e pode-se escrever as seguintes equações:

$$\mathbf{r}^2 \mathbf{p}_1 + \mathbf{x}^T \mathbf{p}_2 = 1 \quad (3.31)$$

$$\mathbf{r}^2 \mathbf{p}_2^T + \mathbf{x}^T \mathbf{P}_3 = 0 \quad (3.32)$$

$$\mathbf{x} \mathbf{p}_1 + \mathbf{R}^{-1} \mathbf{p}_2 = 0 \quad (3.33)$$

$$\mathbf{x} \mathbf{p}_2^T + \mathbf{R}^{-1} \mathbf{P}_3 = \mathbf{I} \quad (3.34)$$

Isolando  $\mathbf{R}^{-1}$  com auxílio das equações (3.31) a (3.34) tem-se:

$$\mathbf{R}^{-1} = \mathbf{P}_3^{-1} + \frac{\mathbf{x} \mathbf{x}^T}{\mathbf{r}^2} \quad (3.35)$$

Aplicando o lema de inversão matricial tem-se:

$$\mathbf{R} = \mathbf{P}_3 - \frac{\mathbf{P}_3 \mathbf{x} \mathbf{x}^T \mathbf{P}_3}{\mathbf{r}^2 + \mathbf{x}^T \mathbf{P}_3 \mathbf{x}} \quad (3.36)$$

Esta equação pode ser simplificada através de substituições adequadas das equações (3.31) a (3.34). Deste modo resulta:

$$\mathbf{R} = \mathbf{P}_3 - \frac{\mathbf{p}_2 \mathbf{p}_2^T}{\mathbf{p}_1} \quad (3.37)$$

De maneira similar:

$$\mathbf{H}_2 = \mathbf{A}_2 \mathbf{A}_2^T = \begin{bmatrix} \mathbf{R}^{-1} & \mathbf{y} \\ \mathbf{y}^T & \mathbf{s}^2 \end{bmatrix} \quad (3.38)$$

onde  $\mathbf{y} = \mathbf{D}^T \mathbf{b}$  e  $\mathbf{s}^2 = \mathbf{b}^T \mathbf{b}$ . De modo análogo ao anterior tem-se:

$$\mathbf{Q} = \mathbf{H}_2^{-1} = \begin{bmatrix} \mathbf{Q}_1 & \mathbf{q}_2 \\ \mathbf{q}_2^T & \mathbf{q}_m \end{bmatrix} \quad (3.39)$$

Como  $\mathbf{H}_2 \mathbf{H}_2^{-1} = \mathbf{I}$ , novamente monta-se um conjunto de equações tal que:

$$\mathbf{R}^{-1} \mathbf{Q}_1 + \mathbf{y} \mathbf{q}_2^T = \mathbf{I} \quad (3.40)$$

$$\mathbf{R}^{-1} \mathbf{q}_2 + \mathbf{y} \mathbf{q}_m = 0 \quad (3.41)$$

$$\mathbf{y}^T \mathbf{Q}_1 + \mathbf{s}^2 \mathbf{q}_2^T = 0 \quad (3.42)$$

$$\mathbf{y}^T \mathbf{q}_2 + \mathbf{s}^2 \mathbf{q}_m = 1 \quad (3.43)$$



Com auxílio das equações (3.40) a (3.43) obtém-se:

$$\mathbf{Q}_1^{-1} = \mathbf{R}^{-1} - \frac{\mathbf{y}\mathbf{y}^T}{s^2} \quad (3.44)$$

Aplicando o lema de inversão matricial resulta:

$$\mathbf{Q}_1 = \mathbf{R} + \frac{\mathbf{R}\mathbf{y}\mathbf{y}^T\mathbf{R}}{s^2 - \mathbf{y}^T\mathbf{R}\mathbf{y}} \quad (3.45)$$

desde que  $s^2 - \mathbf{y}^T\mathbf{R}\mathbf{y} \neq 0$ .

Esta equação pode ser simplificada através de substituições adequadas das equações (3.40) a (3.43). Deste modo tem-se:

$$\mathbf{Q}_1 = \mathbf{R} + \frac{\mathbf{q}_2\mathbf{q}_2^T}{\mathbf{q}_m} \quad (3.46)$$

onde  $\mathbf{q}_m$  é dado por:

$$\mathbf{q}_m = \frac{1}{s^2 - \mathbf{y}^T\mathbf{R}\mathbf{y}} \quad (3.47)$$

e

$$\mathbf{q}_2 = -\mathbf{q}_m\mathbf{R}\mathbf{y} \quad (3.48)$$

$$\mathbf{A}_2^\dagger = \mathbf{A}_2^T\mathbf{Q} \quad (3.49)$$

Portanto o algoritmo recursivo é dado por:

$$\mathbf{R} = \mathbf{P}_3 - \frac{\mathbf{p}_2\mathbf{p}_2^T}{\mathbf{p}_1} \quad (3.50)$$

$$\mathbf{y} = \mathbf{D}^T\mathbf{b} \quad (3.51)$$

$$s^2 = \mathbf{b}^T\mathbf{b} \quad (3.52)$$

$$\mathbf{q}_m = \frac{1}{s^2 - \mathbf{y}^T\mathbf{R}\mathbf{y}} \quad (3.53)$$

$$\mathbf{q}_2 = -\mathbf{q}_m\mathbf{R}\mathbf{y} \quad (3.54)$$

$$\mathbf{Q}_1 = \mathbf{R} + \frac{\mathbf{q}_2\mathbf{q}_2^T}{\mathbf{q}_m} \quad (3.55)$$

### 3.4 Uma Proposta de Método de Identificação Paralela Recursiva de Mínimos Quadrados

O emprego de arquiteturas de processamento paralelo e vetorial para o controle em tempo real de sistemas dinâmicos estocásticos ainda apresenta muitas dificuldades de implantação prática, devido a variedade das aplicações e de situações de contexto.

A estimação de parâmetros em tempo real é particularmente difícil de resolver uma vez que envolve grande quantidade de dados e equações mal condicionadas. Para as aplicações onde o controle e a identificação são separados em tarefas independentes e cooperantes e que podem ser executadas sobre computadores paralelos, é possível aumentar o desempenho computacional das tarefas de identificação e controle, possibilitando a utilização eficiente de recursos de *hardware* e *software*.

A seguir descreve-se esquemas de fatoração matricial na estrutura algébrica de métodos de estimação recursiva de parâmetros, visando a geração de algoritmos paralelos e vetoriais de alto nível. A idéia básica é a obtenção de decomposições matriciais baseadas nas transformações ortonormais de Householder para paralelizar algoritmos de identificação, Costa Filho, Bottura & Barreto, [222].

O algoritmo de mínimos quadrados, visto na seção 1.2.5, estima  $\Theta$ , baseado nas observações de  $y_1, y_2, \dots, y_N$ , e pode ser aqui descrito como:

$$\Theta_N = (\Psi_N^T \Psi_N)^{-1} \Psi_N^T \mathbf{Y}_N \quad (3.56)$$

onde:  $\Psi_N = [\Psi_1^T, \Psi_2^T, \dots, \Psi_N^T]$ ;  $\mathbf{Y}_N = [y_1, y_2, \dots, y_N]$ . A versão recursiva clássica do algoritmo de mínimos quadrados, vista na seção 1.2.6, pode aqui ser descrita pelas equações:

$$\hat{\Theta}_{N+1} = \hat{\Theta}_N + K_{N+1}(y_{N+1} - \Psi_{N+1}^T \hat{\Theta}_N) \quad (3.57)$$

onde  $K_N$  é a matriz de ganho variante satisfazendo:

$$K_{N+1} = P_N \Psi_{N+1} (1 + \Psi_{N+1}^T P_N \Psi_{N+1}) \quad (3.58)$$

$$P_{N+1} = P_N - K_{N+1} \Psi_{N+1}^T P_N \quad (3.59)$$

#### 3.4.1 Identificação Paralela Via Fatoração Triangular

Um melhor desempenho computacional pode ser obtido se o programa for desenvolvido com a possibilidade de organizar seus dados de forma a tornar

eficiente o uso de memórias: virtual, principal e *cache*, Costa Filho, [220] e Ortega, [833], e registradores. Uma forma de se conseguir maior velocidade de processamento sobre arquiteturas paralelas e vetoriais é através do emprego de algoritmos-bloco; Costa Filho & Bottura, [221]. Estes algoritmos realizam as operações matriciais na forma particionada, utilizando submatrizes em vez de escalares. Há, portanto, um aumento na velocidade das operações, explorando-se a capacidade vetorial de cada processador, e uma redução da quantidade de dados transferidos da memória principal para a *cache* e da *cache* para os registradores, gerenciando-se eficientemente o compartilhamento da *cache* e dos registradores vetoriais. Na maioria das arquiteturas paralelas e vetoriais existentes, há um conjunto de subrotinas de álgebra linear básica que implementam eficientemente operações bloco-matriciais, por exemplo *ScalaPACK*, seção 3.2.3 .

Para obter uma formulação bloco para processamento paralelo e vetorial do algoritmo de identificação recursiva, emprega-se uma sequência de transformações de Householder:

$$\mathbf{H} = \mathbf{I} - 2\mathbf{w}\mathbf{w}^T, \quad \|\mathbf{w}\|_2 = 1 \quad (3.60)$$

onde  $\mathbf{w}$  é um vetor de comprimento unitário, para formar a seguinte equação:

$$\hat{\Theta}_{N+1} = \mathbf{R}_{N+1}^{-1} \eta_{N+1} \quad (3.61)$$

onde  $\mathbf{R}_{N+1}$  é uma matriz triangular superior  $p \times p$  e  $\eta_{N+1}$  é um vetor de dimensão  $p$ . A sequência de transformações de Householder deve ser escolhida de forma conveniente para que  $\mathbf{R}_{N+1}$  e  $\eta_{N+1}$  possam ser computados como:

$$\begin{pmatrix} \mathbf{R}_{N+1} \\ 0 \end{pmatrix} = \Psi \begin{pmatrix} \mathbf{R}_N \\ x_{N+1}^T \end{pmatrix} \quad (3.62)$$

onde  $\Psi$  é uma matriz ortonormal, ( $\Psi^T \Psi = I$ ) dada pelo produto de transformações de Householder:

$$\Psi = \mathbf{H}_p \mathbf{H}_{p-1} \dots \mathbf{H}_1 \quad (3.63)$$

onde:

$$\mathbf{H}_i = \begin{bmatrix} I_{i-1} & 0 & & \\ & c_i & 0 & s_i \\ 0 & 0 & I_{p-1} & 0 \\ & s_i & 0 & c_i \end{bmatrix}, \quad i = 1, \dots, p \quad (3.64)$$

com:  $c_i = (\mathbf{R}_N)_{ii}/\lambda_i$ ;  $s_i = (x_{N+1})_i/\lambda_i$ ;  $\lambda_i = [(\mathbf{R}_N)_i^2 + (x_{N+1})_i^2]^{1/2}$  e

$$\begin{pmatrix} \eta_{N+1} \\ v_{N+1} \end{pmatrix} = \Psi \begin{pmatrix} \eta_N \\ y_{N+1} \end{pmatrix} \quad (3.65)$$

A aplicação de  $\mathbf{H}$  à matriz  $\Psi$ , equação (3.56), envolve somente simples operações matriciais do tipo matriz-vetor:  $\mathbf{z} \leftarrow \Psi^T \mathbf{u}$  e a atualização:  $\Psi \leftarrow \Psi - 2\mathbf{u}\mathbf{z}^T$ . Cada uma destas operações requer  $\mathcal{O}(n^2)$  operações de ponto de flutuante em  $\mathcal{O}(n^2)$  dados. Normalmente, as dimensões matriciais são muito maiores do que o número de processadores disponíveis, implicando na necessidade de se realizar multiplicação e adição na forma particionada. Uma análise do procedimento de paralelização e vetorização dessas operações sobre blocos matriciais é dada em Costa Filho, [220].

Já a equação (3.61) pode ser resolvida em paralelo sem dificuldades, uma vez que  $\mathbf{R}_{N+1}$  é uma matriz triangular superior e  $\hat{\Theta}_{N+1}$  pode ser obtido por substituição para trás:

$$\begin{bmatrix} \mathbf{R}_{11} & \dots & \mathbf{R}_{1p} \\ & & \vdots \\ & & \mathbf{R}_{pp} \end{bmatrix}_{N+1} \hat{\Theta}_{N+1} = \begin{bmatrix} \eta_1 \\ \vdots \\ \eta_n \end{bmatrix}_{N+1} \quad (3.66)$$

por simplicidade de notação, consideremos  $\hat{\Theta}_{N+1} = [\hat{\Theta}_{N+1}, \dots, \hat{\Theta}_p]^T$ . O algoritmo de substituição para trás mais comum é:

$$\hat{\Theta}_j = \frac{(\eta_j - \mathbf{R}_{jj+1}\hat{\Theta}_{j+1} - \dots - \mathbf{R}_{jp}\hat{\Theta}_p)}{\mathbf{R}_{jj}}, \quad j = p, \dots, 1. \quad (3.67)$$

Se  $\mathbf{R}_{N+1}$  é armazenado por linha, então (3.66) pode ser resolvida pelo produto interno com vetores de comprimento variando de 1 a  $(p-1)$ , juntamente com  $n$  divisões escalares. Ignorando as divisões, o grau médio de vetorização é proporcional a  $p/2$ .

Portanto a exploração do paralelismo intrínseco das estruturas de equações algébricas de um algoritmo de identificação recursiva através da obtenção de decomposições matriciais, utilizando, convenientemente seqüências de transformações de Householder, possibilitou o desenvolvimento de um algoritmo recursivo paralelo para identificação de sistemas. Esta estratégia de paralelização e vetorização investigada pode ser facilmente estendida para outros processos de identificação recursiva, porém este não é um objetivo desta tese.

### 3.5 Computação Paralela da Decomposição em Valores Singulares

Um importante e não trivial problema tanto na área teórica da álgebra linear como na da álgebra linear numérica é o da determinação da decomposição em valores singulares, *SVD*, de matriz que é estudado em detalhes no Capítulo 4, aplicado nos Capítulos 5, 6 e 7 e cuja computação aqui discutimos.

A decomposição em valores singulares de  $\mathbf{A}_{(m \times n)}$  consiste na fatoração  $\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T$ , sendo  $\mathbf{U}$  e  $\mathbf{V}$  matrizes ortogonais e  $\Sigma$  uma matriz diagonal tal que  $(\sigma_1, \dots, \sigma_r)$ ,  $r = \min(m, n)$ , com  $\sigma_1 \geq \dots \geq \sigma_r \geq 0$ .  $\sigma_i$  são os valores singulares, e as primeiras  $r$  colunas da matriz  $\mathbf{V}$  são os vetores singulares à direita, e as primeiras  $r$  colunas de  $\mathbf{U}$  são os vetores singulares à esquerda da matriz  $\mathbf{A}$ .

- Em princípio, a decomposição em valores singulares poderia ser obtida através da computação de  $\mathbf{A}^T\mathbf{A}$  ou  $\mathbf{A}\mathbf{A}^T$ . No entanto este procedimento gera algoritmo não estável como exemplificado a seguir:

**Exemplo 3.6** Considere o caso  $n = 2$ :  $[a_1, a_2] \in \mathbb{R}^{m \times n}$ , onde  $a_1^T a_2 = \cos\gamma$ , com  $\|a_1\| = \|a_2\| = 1$ ; logo:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & \cos\gamma \\ \cos\gamma & 1 \end{bmatrix}$$

terá autovalores:  $\lambda_1 = 2\cos^2(\gamma/2)$  e  $\lambda_2 = 2\sin^2(\gamma/2)$ , donde:

$$\sigma_1 = \sqrt{2}\cos\frac{\gamma}{2} \quad e \quad \sigma_2 = \sqrt{2}\sin\frac{\gamma}{2}$$

Os autovetores de  $\mathbf{A}^T\mathbf{A}$ :

$$\mathbf{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad e \quad \mathbf{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

são os vetores singulares de  $\mathbf{A}$  à direita. Se  $\gamma \ll 1$ , então  $\sigma_1 \approx \sqrt{2}$  e  $\sigma_2 \approx \gamma/\sqrt{2}$  e deste modo:

$$\mathbf{u}_1 \approx \frac{1}{2}(a_1 + a_2) \quad e \quad \mathbf{u}_2 \approx \frac{(a_1 - a_2)}{\gamma}$$

Supondo que  $\gamma$  é menor que a raiz quadrada da precisão de ponto flutuante, computa-se  $\cos\gamma = 1$  em  $\mathbf{A}^T\mathbf{A}$ . Desta forma a computação de  $\mathbf{A}^T\mathbf{A}$  será singular com autovalores 2 e 0 e não será possível obter o valor singular  $\gamma/\sqrt{2}$ .

Este exemplo mostra como informações podem ser perdidas por problemas de precisão na computação da SVD utilizando  $\mathbf{A}^T\mathbf{A}$ .

A seguir descrevemos um procedimento computacional, comum às bibliotecas *LAPACK* e *ScaLAPACK*, para calcular a SVD em duas etapas:

Etapa do Algoritmo	1 Processador	2 Processadores	3 Processadores
SVD de $\mathbf{A}_{8 \times 20}$	1.0	0.27	0.50
SVD de $\mathbf{A}_{15 \times 15}$	1.0	0.43	0.79
SVD de $\mathbf{A}_{42 \times 40}$	1.0	0.61	0.97
SVD de $\mathbf{A}_{239 \times 120}$	1.0	1.92	2.38

Tabela 3.3: Ganho de Velocidade de Processamento na Decomposição em Valores Singulares.

1. A matriz  $\mathbf{A}$  é reduzida para a forma bidiagonal:  $\mathbf{A} = \mathbf{U}_1 \mathbf{B} \mathbf{V}_1^T$  se  $\mathbf{A}$  é real, sendo  $\mathbf{U}_1$  e  $\mathbf{V}_1$  ortogonais, e  $\mathbf{B}$  é real e bidiagonal superior quando  $m \geq n$  ou bidiagonal inferior quando  $m < n$ , e tal que  $\mathbf{B}$  é diferente de zero só na diagonal principal e/ou na primeira superdiagonal, se  $m \geq n$ , ou na primeira subdiagonal, se  $m < n$ .
2. A SVD da matriz bidiagonal  $\mathbf{B}$  é calculada:  $\mathbf{B} = \mathbf{U}_2 \Sigma \mathbf{V}_2^T$ , sendo  $\mathbf{U}_2$  e  $\mathbf{V}_2$  ortogonais e  $\Sigma$  diagonal como anteriormente. Os vetores singulares da matriz  $\mathbf{A}$  são  $\mathbf{U} = \mathbf{U}_1 \mathbf{U}_2$  e  $\mathbf{V} = \mathbf{V}_1 \mathbf{V}_2$ .

**Exemplo 3.7** Utilizando processamento distribuído e paralelo, variando o número de processadores, em uma rede de estações de trabalho, realizou-se a SVD de 4 matrizes  $\mathbf{A}$  de dimensões diferentes.

Os resultados obtidos com computação distribuída e paralela da decomposição em valores singulares das matrizes são apresentados na Tabela 3.3 em termos do ganho de velocidade de processamento.

Observa-se que o desempenho computacional com 2 ou 3 processadores melhora, ( $S_p(n) > 1$ ), na situação onde ocorre o processamento da matriz  $\mathbf{A}_{239 \times 120}$ , o que não ocorre nas situações anteriores. Esta característica do processamento distribuído e paralelo da decomposição em valores singulares nos motivou a elaboração da próxima seção onde variamos os tamanhos dos blocos de matriz alocados para cada processador.

### 3.5.1 Experimentos Com a Computação da Decomposição em Valores Singulares

Experimentos para avaliação de desempenho para a computação da SVD, utilizando processamento distribuído e paralelo, são realizados nesta seção.

As rotinas da biblioteca *ScaLAPACK* para resolver sistemas lineares densos e problemas de autovalores consideram que todos os dados globais foram

	1	2	3	4	5
1					
2		{0,0}		{0,1}	{0,2}
3					
4		{1,0}		{1,1}	{1,2}
5		{0,0}		{0,1}	{0,2}

Tabela 3.4: Distribuição em blocos sobre uma grade com processadores ( $2 \times 3$ ).

alocados aos processos com uma distribuição de dados em blocos cíclicos com até duas dimensões, o que minimiza a frequência de movimento de dados entre os diferentes níveis da memória hierárquica. Esta alocação é uma expressão natural dos algoritmos particionados em blocos presente na *ScaLAPACK*.

Assim, uma matriz global, com dimensão  $(m \times n)$ , é particionada em blocos de tamanho  $(m_i \times n_i)$ , denominados matrizes locais, distribuídos entre os diferentes processadores que participam do processamento. Uma representação para melhor compreensão do processo de distribuição de uma matriz global sobre um processo grade de duas dimensões é apresentada a seguir.

A Tabela 3.4 apresenta uma distribuição em blocos cíclicos de duas dimensões de uma matriz global  $(5 \times 5)$  com blocos de tamanho  $(2 \times 2)$  sobre uma grade com processadores  $(2 \times 3)$ , ou seja 2 processadores por linha e 3 processadores por coluna, tendo 6 processadores no total.

Os índices linha/coluna indicam, respectivamente, os blocos linha/coluna. Cada caixa contém o índice  $\{i, j\}$  do processador que contém a informação específica do bloco. A seguir apresentamos a mesma distribuição da matriz, mas do ponto de vista do processador, Tabela 3.5.

Neste caso, cada caixa representa a informação que contém cada processador. Por exemplo, o processador  $\{0, 0\}$  contém os elementos  $A(1, 1)$ ,  $A(1, 2)$ ,  $A(2, 1)$ ,  $A(2, 2)$ ,  $A(5, 1)$  e  $A(5, 2)$  da matriz global.

A distribuição em blocos cíclicos de duas dimensões garante um bom balanceamento de carga do qual resulta bom desempenho e escalabilidade do algoritmo.

Para execução do programa paralelo para o cálculo da *SVD* define-se as dimensões das matrizes e o tamanho dos blocos em que será dividida a matriz global. A saída do programa nos fornece o tempo de execução respectivo para cada caso.

Com base nos resultados de tempo de execução, em segundos, apresentados nas Tabelas 3.6, 3.7 e 3.8, pode-se observar que, em geral, é melhor definir tamanhos de blocos grandes pois desta maneira os melhores tempos

	1	2	3	4	5
1					
2		{0,0}		{0,1}	{0,2}
5					
3					
4		{1,0}		{1,1}	{1,2}

Tabela 3.5: Distribuição de Dados do ponto de vista do processador.

Tamanho do bloco	Dimensões das matrizes	1Pr.	2Pr.	3Pr.
$(8 \times 8)$	$(8 \times 20)$	0.0457	0.1909	0.2897
$(4 \times 4)$	$(8 \times 20)$	0.0578	0.2312	0.3291
$(10 \times 10)$	$(8 \times 20)$	0.0459	0.1870	0.2743
$(20 \times 20)$	$(20 \times 8)$	0.0454	0.2002	0.2903
$(10 \times 10)$	$(20 \times 8)$	0.0462	0.2070	0.3160
$(8 \times 8)$	$(20 \times 8)$	0.0456	0.2241	0.3178
$(4 \times 4)$	$(8 \times 40)$	0.0611	0.2714	0.3505
$(8 \times 8)$	$(8 \times 40)$	0.0502	0.1969	0.3159
$(10 \times 10)$	$(8 \times 40)$	0.0460	0.1845	0.2807
$(40 \times 40)$	$(40 \times 8)$	0.0453	0.2209	0.2968
$(10 \times 10)$	$(40 \times 8)$	0.0457	0.2038	0.3856
$(8 \times 8)$	$(40 \times 8)$	0.0457	0.2291	0.3249
$(15 \times 15)$	$(15 \times 20)$	0.0605	0.2390	0.4200
$(10 \times 10)$	$(15 \times 20)$	0.0864	0.3682	0.6114
$(5 \times 5)$	$(15 \times 20)$	0.0863	0.3603	0.5315
$(20 \times 20)$	$(20 \times 15)$	0.0573	0.2443	0.3901
$(10 \times 10)$	$(20 \times 15)$	0.0965	0.3541	0.5537
$(32 \times 32)$	$(20 \times 15)$	0.0611	0.2413	0.3847
$(15 \times 15)$	$(15 \times 40)$	0.0644	0.2405	0.3524
$(7 \times 7)$	$(15 \times 40)$	0.1044	0.4355	0.5662
$(5 \times 5)$	$(15 \times 20)$	0.0952	0.3855	0.5080
$(64 \times 64)$	$(40 \times 15)$	0.0633	0.2474	0.3412
$(32 \times 32)$	$(40 \times 15)$	0.0616	0.2756	0.4356
$(10 \times 10)$	$(40 \times 15)$	0.0937	0.3717	0.5603

Tabela 3.6: Tempo de Execução: 1, 2 e 3 Processadores.



Tamanho do bloco	Dimensões das matrizes	1Pr.	2Pr.	3Pr.
$(42 \times 42)$	$(42 \times 20)$	0.0824	0.2975	0.4307
$(6 \times 6)$	$(42 \times 20)$	0.1469	0.4759	0.6455
$(32 \times 32)$	$(42 \times 20)$	0.0799	0.3318	0.5337
$(20 \times 20)$	$(20 \times 42)$	0.0797	0.2839	0.4380
$(10 \times 10)$	$(20 \times 42)$	0.1122	0.4312	0.5849
$(7 \times 7)$	$(20 \times 42)$	0.1281	0.4979	0.6158
$(42 \times 42)$	$(42 \times 40)$	0.1780	0.5070	0.8452
$(20 \times 20)$	$(42 \times 40)$	0.2554	0.7899	1.3172
$(6 \times 6)$	$(42 \times 40)$	0.3043	0.8703	1.1879
$(40 \times 40)$	$(40 \times 42)$	0.1817	0.4875	0.7511
$(20 \times 20)$	$(40 \times 42)$	0.2653	0.7320	1.0910
$(13 \times 13)$	$(40 \times 42)$	0.3090	0.9496	1.1987

Tabela 3.7: *Tempo de Execução: 1, 2 e 3 Processadores.*

de execução são obtidos. Quando temos só 1 processador, o melhor tempo é obtido se o número de linhas é maior do que o de colunas e o tamanho do bloco é suficientemente grande. Nestes experimentos a distribuição de dados foi feita por linhas. Quanto maior é a quantidade de elementos na matriz, menor é o tempo de execução. À medida que aumentamos o número de processadores que participam do processamento, o tempo de execução também aumenta, devido à comunicação que existe entre os diferentes processadores para realizar o envio de dados.

Para matrizes com maior número de linhas do que de colunas, sempre que a distribuição for feita por linhas, é melhor definir o tamanho do bloco, com valores grandes, por exemplo  $(32 \times 32)$ . Quando comparamos matrizes com dimensões  $(m \times n)$  e  $(n \times m)$ , é muito melhor escolher matrizes onde o número de linhas seja maior que o número de colunas, sempre que utilizarmos distribuição por linhas, como ocorre nos exemplos utilizados para gerar a Tabela 3.6.

Os resultados apresentados na Tabela 3.8, referem-se a matrizes com dimensões bem maiores, com a finalidade de encontrar dimensões matriciais onde começa a ocorrer melhor desempenho do programa paralelo e distribuído em termos de tempo de execução. Pode-se observar que o tempo de execução, diminui quando o tamanho do bloco aumenta para matrizes na faixa de 1000 a 2000 elementos.

A partir dos experimentos realizados podemos concluir que para a paralelização deste algoritmo melhores resultados são obtidos quando o problema envolve cálculos com matrizes de ordens bem maiores e é feita uma distri-

Tamanho do bloco	Dimensões das matrizes	1Pr.	2Pr.	3Pr.
(120 × 120)	(240 × 20)	0.1688	0.3877	0.5630
(70 × 70)	(240 × 20)	0.1632	0.4189	0.5386
(13 × 13)	(240 × 20)	0.2246	0.5453	0.7659
(64 × 64)	(20 × 240)	0.1832	0.4513	0.5770
(120 × 120)	(20 × 240)	0.1831	0.4175	0.5604
(32 × 32)	(20 × 240)	0.1830	0.4320	0.5526
(120 × 120)	(240 × 40)	0.4495	0.6932	1.0045
(70 × 70)	(240 × 40)	0.4521	0.7305	1.1950
(13 × 13)	(240 × 40)	0.6047	1.0444	1.2757
(120 × 120)	(240 × 120)	3.9898	3.6668	4.2159
(64 × 64)	(240 × 120)	4.7554	4.5040	6.3719

Tabela 3.8: *Tempo de Execução: 1, 2 e 3 Processadores.*

buição de dados em blocos de tamanho superior a 32.

Os resultados da computação paralela em que o tempo gasto para a computação com 2 ou 3 processadores supera o com apenas um processador devem-se à transmissão de dados e a complexidade da rede.

Estes experimentos foram realizadas num computador paralelo virtual formado por computadores conectados numa rede local com estações de trabalho *Sun* de tipos *SPARC-20*, *SPARC-4*, *SPARC-Classic* e *SPARC-2*.

### 3.6 Proposta de Procedimento Para Avaliação de Desempenho Numérico de Algoritmos Matriciais

Considerando que:

- A avaliação do desempenho computacional de algoritmo matricial pelos métodos tradicionais, muitas vezes se torna tediosa.
- Se o algoritmo desenvolvido for extenso e complexo, a avaliação pode não corresponder à realidade, dado que em muitos casos o algoritmo é solucionado iterativamente.
- Alguns critérios analíticos de desempenho consideram como número total de operações de ponto flutuante as que não envolvem zeros.

- Outros critérios para estimar desempenho de algoritmo levam em consideração o tempo de execução do mesmo, fato que precisa ser melhor considerado quando da utilização de computação paralela e distribuída realizada em um conjunto de máquinas interligadas em rede.

com o objetivo de contribuir para a avaliação de desempenho de algoritmos matriciais paralelos e distribuídos apresenta-se um procedimento computacional experimental para tanto, Bottura, Barreto, Bordon & Costa Filho, [131].

A elaboração deste procedimento tem como ponto inicial a praticabilidade computacional paralela, seção 3.1.3 onde de todos os itens citados a maioria diz respeito a análise da eficiência do algoritmo, ou seja ao número de operações computacionais necessário para se obter a solução, que pode ser determinado através da contagem linha a linha do número de operações do programa desenvolvido para implementar o algoritmo.

Para algoritmo de mínimos quadrados, Tabela 3.9, são apresentados resultados onde  $m, n$  são respectivamente o número de linhas e colunas da matriz  $\mathbf{A}$ , Golub & Van Loan, [399].

Algoritmos de Mínimos Quadrados	Número de Operações
Equação Normal	$mn^2 + n^3/3$
Ortogonalização de Householder	$2mn^2 - 2n^3/3$
Ortogonalização de Givens	$3mn^2 - n^3$
Gram Schmidt Modificado	$2mn^2$
Bidiagonalização de Householder	$4mn^2 - 4n^3/2$
R Bidiagonalização	$2mn^2 + 2n^3$
Golub-Reinsch SVD	$4mn^2 + 8n^3$
Chan SVD	$2mn^2 + 11n^3$

Tabela 3.9: Número de Operações Para Solução de Problema de Mínimos Quadrados.

A contagem de número de operações é conceituada como medida de complexidade de algoritmo.

### 3.6.1 Cálculo do Número de Operações de Algoritmos Matriciais

Da Tabela 3.9 podemos observar que para todos os algoritmos citados os números de operações são funções semelhantes às obtidas para operações elementares em matrizes, Tabela 3.10.

Tipo de Operação	Número de Operações
$\mathbf{A}_{n \times n} + \mathbf{B}_{n \times n}$	$n^2$
$\mathbf{A}_{n \times r} \mathbf{B}_{r \times m}$	$2nrm$
$\mathbf{A}_{n \times n}^{100}$	$(99)2n^3$
$\log_e \mathbf{A}_{n \times n}$	$(2/3)n^3$

Tabela 3.10: Números de Operações Elementares Para Matrizes.

Se a contagem do número de operações de um algoritmo que envolve uma matriz  $\mathbf{A}_{m \times n}$  for dada na forma:

$$\text{Número de Operações} = K(mn)^i \quad (3.68)$$

onde  $K$  e  $i$  são dependentes do algoritmo, então através da determinação experimental de tais parâmetros poderemos obter uma avaliação da complexidade do algoritmo pelo levantamento do número de operações envolvidas, “*flops*”, para diferentes dimensões matriciais. Este levantamento também pode, por exemplo, ser efetuado com o comando “*flops*” do *MATLAB - Matrix Laboratory* - [761]. Após tal tipo de levantamento traça-se uma curva em escala log-log relacionando o número de operações executadas pela dimensão matricial, por exemplo  $m \times n$ ; o coeficiente angular desta reta é o valor de  $i$  e o coeficiente linear é o valor de  $K$ .

### 3.6.2 Aplicação do Método de Avaliação Proposto

Para a determinação dos parâmetros  $K$  e  $i$ :

- Executa-se o algoritmos com diversas dimensões matriciais.
- Para cada execução tabela-se as dimensões,  $mn$ , e o número de operações envolvidas.
- Traça-se o gráfico log-log relacionando o número de operações envolvidas com a dimensão matricial,  $mn$ .

Alguns Tipos de Solução de $\mathbf{Ax} = \mathbf{b}$	
Casos	Solução
$m = n = \text{posto}(\mathbf{A})$	$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$
$n > m = \text{posto}(\mathbf{A})$	$\mathbf{x} = \mathbf{A}^\dagger \mathbf{b}$ onde $\mathbf{A}^\dagger = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$
$m > n = \text{posto}(\mathbf{A})$	$\mathbf{x} = \mathbf{A}^\dagger \mathbf{b}$ onde $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$

Tabela 3.11: Alguns Tipos de Solução de Sistema de Equações Lineares.

Para avaliar experimentalmente o número de operações na computação de alguns casos da solução de  $\mathbf{Ax} = \mathbf{b}$ ,  $\mathbf{A}_{m \times n}$ ,  $\mathbf{b} \in \Re^m$   $\mathbf{x} \in \Re^n$  consideramos os casos indicados na Tabela 3.11.

Para a avaliação de desempenho computacional pelo método proposto, levando em conta os algoritmos da Tabela 3.10, e os casos indicados na Tabela 3.11, produzimos a Tabela 3.12. Observe que para  $n = m$ , algoritmos diferentes exigem o mesmo número de operações, fato este não perceptível na Tabela 3.9.

Algoritmos de Mínimos Quadrados	Dimensão da Matriz <b>A</b>		
	$n = m$	$n > m$	$n < m$
Equação Normal	$1.4493(mn)^{1.4621}$	$1.7370(mn)^{0.9345}$	$0.8765(mn)^{1.5850}$
Ortogonalização de Householder	$1.4493(mn)^{1.4621}$	$1.0518(mn)^{1.5850}$	$1.2768(mn)^{1.5070}$
Ortogonalização de Givens	$2.0000(mn)^{1.5000}$	$1.6292(mn)^{1.5670}$	$1.7974(mn)^{1.5325}$
Gram Schmidt Modificado	$2.0000(mn)^{2.5000}$	$2.8523(mn)^{1.4150}$	$1.4024(mn)^{1.5850}$
Bidiagonalização de Householder	$2.0000(mn)^{1.5000}$	$0.6779(mn)^{1.8301}$	$2.2093(mn)^{1.4919}$
R Bidiagonalização	$2.0000(mn)^{1.5000}$	$8.5228(mn)^{1.3155}$	$12.3622(mn)^{0.6554}$
Golub-Reinsch SVD	$12.0000(mn)^{1.5000}$	$28.5734(mn)^{1.2895}$	$5.4755(mn)^{1.6845}$
Chan SVD	$11.9348(mn)^{1.5617}$	$34.5533(mn)^{1.2645}$	$5.1369(mn)^{1.7201}$

Tabela 3.12: Avaliação de Desempenho de Métodos para Solução de Problema de Mínimos Quadrados.

## Capítulo 4

# DE PSEUDO-INVERSA, PROJEÇÕES, DECOMPOSIÇÃO EM VALORES SINGULARES E SOLUÇÃO DE MÍNIMOS QUADRADOS EM SUBESPAÇOS

### 4.1 Introdução

Utilizando os fundamentos de álgebra linear, Capítulo 2, e motivados pela solução do chamado problema de mínimos quadrados, visto no Capítulo 1, elaboramos neste capítulo um estudo personalizado e aprofundado sobre matrizes pseudo-inversas, projeções, decomposição em valores singulares e solução de mínimos quadrados do problema  $\mathbf{Ax} = \mathbf{b}$  com o objetivo de criar estrutura adequada à elaboração de fundamentos para modelagem de dados em subespaços do espaço de estado, que será estudada no Capítulo 5 e muito utilizada nos Capítulos 6 e 7.

Considerando o sistema da equação (1.20):

$$\mathbf{Y} = \mathbf{\Psi}\mathbf{\Theta} + \varepsilon \quad (4.1)$$

pode-se representá-lo na forma  $\mathbf{b} = \mathbf{Ax}$  onde  $\mathbf{A} = \mathbf{\Psi}$ ,  $\mathbf{b} = \mathbf{Y}$  e  $\mathbf{x} = \mathbf{\Theta}$  (observe que o erro nesta representação foi desprezado).

A solução da equação  $\mathbf{Ax} = \mathbf{b}$ , no problema de identificação em pauta, constitui-se na determinação do vetor de parâmetros,  $\mathbf{x}$ .

Se a solução não é exata pode-se definir um erro dado por:

$$\mathbf{e} = \mathbf{Ax} - \mathbf{b} \quad (4.2)$$

e desta forma considerar o erro da equação (4.1).

Quando  $\mathbf{e}$  for zero a solução será exata. O “problema dos mínimos quadrados” busca obter o  $\mathbf{e}$  menor possível, daí o motivo para a representação  $\cong$  utilizada na Figura 4.2.

Com base nos conceitos de álgebra linear a solução de mínimos quadrados do problema  $\mathbf{Ax} = \mathbf{b}$  nos faz refletir em pelo menos quatro questões fundamentais:

- O problema tem solução ?
- Quando da existência de solução ela é única ?
- Quais os efeitos de perturbações <sup>1</sup> na solução ?
- Como obter computacionalmente a melhor solução ?

Podemos interpretar o “problema de mínimos quadrados” como:

*Dada uma matriz  $\mathbf{A}_{(m \times n)}$  real de posto  $r \leq \min(m, n)$  e dado um vetor real  $\mathbf{b}$  de dimensão  $m$ , obter um vetor real  $\hat{\mathbf{x}}$  que minimize a norma euclidiana de  $\mathbf{Ax} - \mathbf{b}$ .*

Esta representação do problema dos mínimos quadrados tem como ponto de partida um conjunto de  $n$  vetores  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n$ . Quando supomos que estes vetores são linearmente independentes, ao encontrar a combinação linear de  $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n\}$  com pesos  $\{\hat{\mathbf{x}}_1, \hat{\mathbf{x}}_2, \hat{\mathbf{x}}_3, \dots, \hat{\mathbf{x}}_n\} = \hat{\mathbf{x}}^T$  que gera o vetor  $\mathbf{b}$ , resolve-se o problema. Contudo há diversas situações além desta e o restante deste capítulo procura abordá-las.

A solução de  $\mathbf{Ax} = \mathbf{b}$  poderá ser complicada ou simples dependendo das bases escolhidas. A decomposição em valores singulares de uma matriz e sua utilização para cálculo da matriz pseudo-inversa é uma eficaz ferramenta para encontrar tal combinação linear.

---

<sup>1</sup>Este item é tratado quando da análise do problema de mínimos quadrados estocástico para o qual sugerimos como referências Saridis [933], Moustafa [791], El-Sherief e Sinha [312] entre outras, todas abordando o problema de identificação de sistemas.



## 4.2 Estrutura da Transformação Linear: Subespaços Fundamentais

Com a motivação acima e com base no capítulo anterior temos alguns comentários sobre a solução da equação  $\mathbf{Ax} = \mathbf{b}$  cujo estudo com mais detalhes é importante para os objetivos desta tese e será feito neste capítulo. Para tanto seja a transformação linear  $\mathbf{Ax} = \mathbf{y}$ ,  $\mathbf{A}_{m \times n}$  de posto  $r$ .

- A matriz  $\mathbf{A}$  representa um operador linear.
- O produto de matriz  $\mathbf{A}$  por  $\mathbf{x}$  é o vetor coluna  $m$ -dimensional de coordenadas:

$$y_i = \sum_{j=1}^n a_{ij}x_j \quad i = 1, 2, \dots, m \quad (4.3)$$

Por simplicidade suporemos  $y_i$ ,  $a_{ij}$  e  $x_j$  reais.

Se  $\mathbf{a}_j$  denota a  $j$ -ésima coluna de  $\mathbf{A}$ , um vetor  $m \times 1$ , então pode-se reescrever (4.3) como:

$$\mathbf{y} = \mathbf{Ax} = \sum_{j=1}^n x_j \mathbf{a}_j \quad (4.4)$$

que podemos representar pelo seguinte esquema:

$$\begin{aligned} \begin{bmatrix} \mathbf{y} \end{bmatrix} &= \mathbf{Ax} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \\ &= x_1 \begin{bmatrix} \mathbf{a}_1 \end{bmatrix} + x_2 \begin{bmatrix} \mathbf{a}_2 \end{bmatrix} + \dots + x_n \begin{bmatrix} \mathbf{a}_n \end{bmatrix} \end{aligned}$$

Em (4.4) exprimimos  $\mathbf{y}$  como uma combinação linear das colunas  $\mathbf{a}_i$  e podemos interpretar a fórmula  $\mathbf{Ax} = \mathbf{b}$  de duas maneiras:

- (4.3) afirma que  $\mathbf{A}$  atua sobre  $\mathbf{x}$  para produzir  $\mathbf{y}$ .
- (4.4) afirma que  $\mathbf{x}$  atua sobre  $\mathbf{A}$  para produzir  $\mathbf{y}$ .

•

$$\text{posto}(\mathbf{A}) + \text{nulidade}(\mathbf{A}) = n \quad (4.5)$$

- Um subespaço com dimensão  $r$  pode ser expresso como uma combinação linear do conjunto de vetores linha linearmente independentes ou dos vetores coluna linearmente independentes da matriz  $\mathbf{A}$ . O número  $r$  de linhas ou de colunas linearmente independentes da matriz  $\mathbf{A}$  constitui seu posto.
- O conjunto de todos os vetores que são combinações lineares das colunas da matriz  $\mathbf{A}$  gera o **espaço coluna**, ou espaço imagem da matriz  $\mathbf{A}$ ,  $\mathfrak{R}(\mathbf{A})$ , e tem dimensão  $r$ . O espaço coluna de  $\mathbf{A}$  é um subespaço de  $\mathbf{R}^m$ .
- O conjunto de todos os vetores que são combinações lineares das linhas da matriz  $\mathbf{A}$  gera o **espaço linha** da matriz  $\mathbf{A}$ , ou **espaço coluna** da matriz  $\mathbf{A}^T$  e tem dimensão  $r$ . O espaço linha de  $\mathbf{A}$  é um subespaço de  $\mathbf{R}^n$ .

•

$$r = \text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}^T) \quad (4.6)$$

- O espaço nulo de  $\mathbf{A}$ ,  $\mathcal{N}(\mathbf{A})$ , é formado por todos os vetores  $\mathbf{x}$  que são as soluções para  $\mathbf{A}\mathbf{x} = \mathbf{0}$ .
- O espaço nulo de  $\mathbf{A}$ ,  $\mathcal{N}(\mathbf{A})$ , é ortogonal ao espaço linha em  $\mathbf{R}^n$ , logo cada vetor  $\mathbf{x}$  no espaço nulo de  $\mathbf{A}$  é perpendicular a cada linha de  $\mathbf{A}$ .

Para verificar esta propriedade recorreremos ao produto interno de vetores e ao esquema da equação (4.7). Para todo  $\mathbf{x} \in \mathcal{N}(\mathbf{A})$  tem-se+:

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} \text{linha 1 de } \mathbf{A} \\ \text{linha 2 de } \mathbf{A} \\ \vdots \\ \text{linha } m \text{ de } \mathbf{A} \end{bmatrix} \mathbf{x} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad (4.7)$$

Como o produto interno de cada linha de  $\mathbf{A}$  por  $\mathbf{x}$  é nulo, então cada linha da matriz  $\mathbf{A}$  é perpendicular à  $\mathbf{x}$ . Dado que as linhas de  $\mathbf{A}$  são as colunas de  $\mathbf{A}^T$ , podemos afirmar que o espaço coluna de  $\mathbf{A}^T$  é ortogonal ao espaço nulo,  $\mathcal{N}(\mathbf{A})$ .

- O espaço nulo de  $\mathbf{A}^T$  é ortogonal ao espaço coluna em  $\mathbf{R}^m$ , logo cada vetor  $\mathbf{z}$  do espaço nulo de  $\mathbf{A}^T$  é perpendicular a cada coluna de  $\mathbf{A}$ .

Analogamente verificamos esta propriedade através do produto interno de vetores:

Para todo  $\mathbf{z} \in \mathcal{N}(\mathbf{A}^T)$  tem-se:

$$\mathbf{A}^T \mathbf{z} = \begin{bmatrix} \text{linha 1 de } \mathbf{A}^T \\ \text{linha 2 de } \mathbf{A}^T \\ \vdots \\ \text{linha } n \text{ de } \mathbf{A}^T \end{bmatrix} \quad \mathbf{z} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

Como o produto interno de cada linha  $\mathbf{A}^T$  por  $\mathbf{z}$  é nulo, então cada linha da matriz  $\mathbf{A}^T$  é perpendicular à  $\mathbf{z}$ . Dado que as linhas de  $\mathbf{A}^T$  são as colunas de  $\mathbf{A}$ , podemos afirmar que o espaço coluna de  $\mathbf{A}$  é ortogonal ao espaço  $\mathcal{N}(\mathbf{A}^T)$ .

- O espaço nulo de  $\mathbf{A}$ ,  $\mathcal{N}(\mathbf{A})$ , subespaço de  $\mathbf{R}^n$  tem dimensão  $n - r$ .
- O espaço nulo de  $\mathbf{A}^T$ ,  $\mathcal{N}(\mathbf{A})^T$ , subespaço de  $\mathbf{R}^m$  tem dimensão  $m - r$ .

O espaço nulo  $\mathbf{A}^T$ ,  $\mathcal{N}(\mathbf{A}^T)$ , e o espaço coluna de  $\mathbf{A}$ ,  $\mathfrak{R}(\mathbf{A})$ , em  $\mathbf{R}^m$  além de ortogonais, são ortogonalmente complementares. Do mesmo modo o espaço nulo de  $\mathbf{A}$  e o espaço linha de  $\mathbf{A}$  em  $\mathbf{R}^n$  além de serem espaços ortogonais são ortogonalmente complementares. Cada vetor  $\mathbf{x}$  que for perpendicular ao espaço linha de  $\mathbf{A}$  satisfaz  $\mathbf{A}\mathbf{x} = 0$  e está no seu espaço nulo. O complemento ortogonal do espaço nulo de  $\mathbf{A}$ ,  $\mathcal{N}(\mathbf{A})^\perp$ , é o espaço linha de  $\mathbf{A}$ .

Em síntese, a Figura 4.1 representa a estrutura da transformação linear de  $\mathbf{A}$  através de 4 subespaços fundamentais. Nesta estrutura há dois pares de espaços que são ortogonalmente complementares.

**Lema 4.1** Para a transformação linear:  $\mathbf{A} : \mathbf{R}^n \rightarrow \mathbf{R}^m$

$$\mathfrak{R}(\mathbf{A}) = \mathfrak{R}(\mathbf{A}\mathbf{A}^T) \quad (4.8)$$

$$\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}\mathbf{A}^T) \quad (4.9)$$

onde  $\mathbf{A}\mathbf{A}^T : \mathbf{R}^m \rightarrow (\mathbf{A})\mathbf{R}^m$  e  $\mathbf{A}^T\mathbf{A} : \mathbf{R}^n \rightarrow (\mathbf{A})\mathbf{R}^n$ .

**Prova:**

- a) Se  $\mathbf{y} \in \mathfrak{R}(\mathbf{A}\mathbf{A}^T)$  então:

$$\mathbf{y} = \mathbf{A}\mathbf{A}^T \mathbf{z}$$

para  $\mathbf{z} \in \mathbf{R}^m$ . Seja  $\mathbf{A}^T \mathbf{z} = \mathbf{x} \in \mathbf{R}^n$  então  $\mathbf{y} = \mathbf{A}\mathbf{x}$  e concluímos que  $\mathbf{y} \in \mathfrak{R}(\mathbf{A})$ .

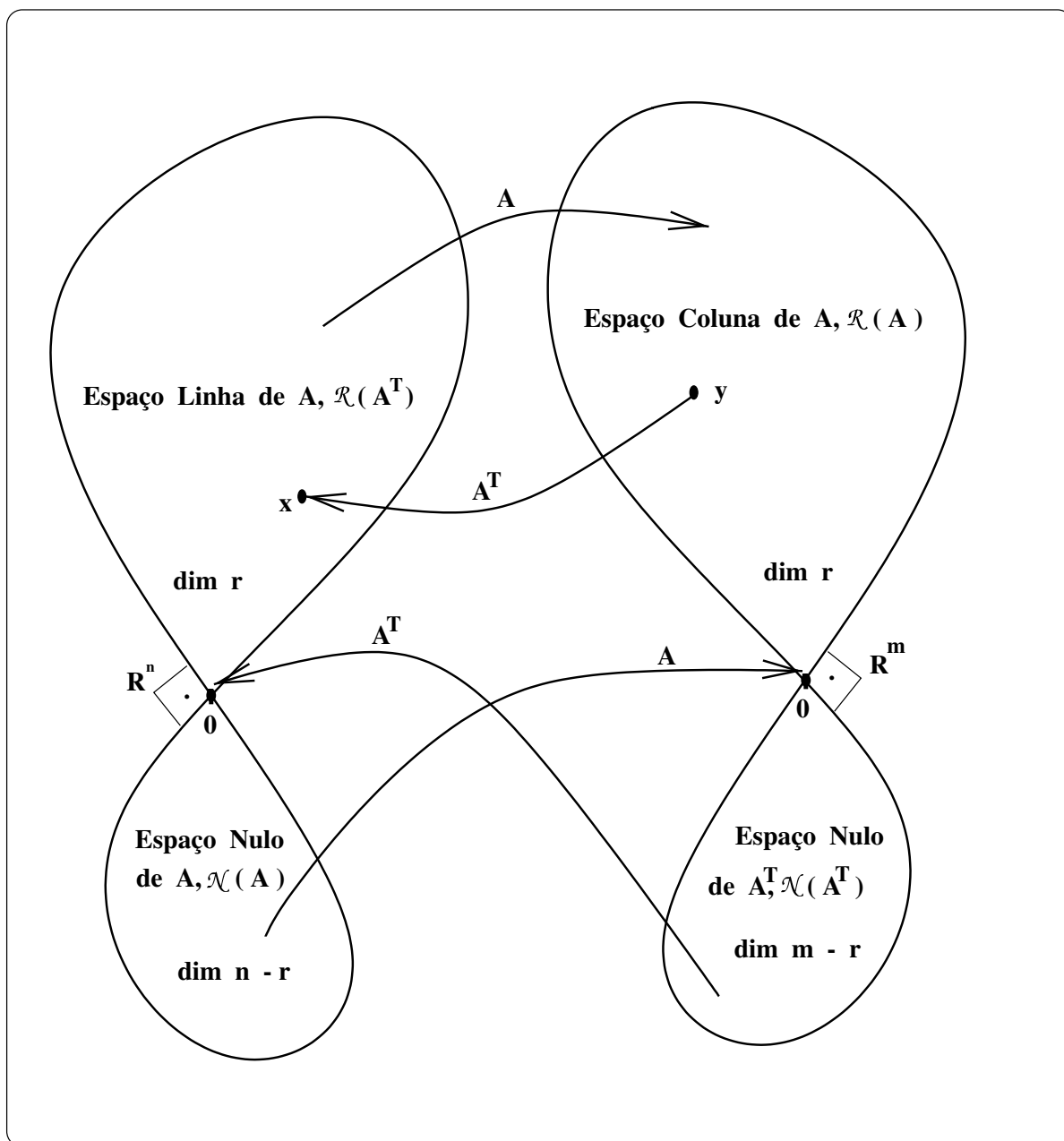


Figura 4.1: Representação Geométrica dos Subespaços Fundamentais da Matriz  $A_{m \times n}$ .

- b) Se  $\mathbf{y} \in \mathfrak{R}(\mathbf{A})$  então:

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{para} \quad \mathbf{x} \in \mathbf{R}^n$$

podemos fazer  $\mathbf{x} = \mathbf{x}_1 + \mathbf{x}_2$  com  $\mathbf{x}_1 \in \mathfrak{R}(\mathbf{A})^T \Rightarrow \mathbf{x}_1 = \mathbf{A}^T \mathbf{z}$  para  $\mathbf{z} \in \mathbf{R}^m$  e  $\mathbf{x}_2 \in \mathcal{N}(\mathbf{A}) \Rightarrow \mathbf{A}\mathbf{x}_2 = 0$ .

Então  $\mathbf{y} = \mathbf{A}\mathbf{x} = \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2) = \mathbf{A}\mathbf{A}^T \mathbf{z} + 0$  para  $\mathbf{z} \in \mathbf{R}^m$  donde  $\mathbf{y} \in \mathfrak{R}(\mathbf{A}\mathbf{A}^T)$ .

Provamos (4.8); (4.9) resulta porque:

$$\mathcal{N}(\mathbf{A}) = \mathfrak{R}(\mathbf{A}^T)^\perp = \mathfrak{R}(\mathbf{A}^T \mathbf{A})^\perp = \mathcal{N}(\mathbf{A}^T \mathbf{A})$$

**C.Q.D.**

Deste lema e de (4.8) decorrem:

$$\text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}\mathbf{A}^T) = \text{posto}(\mathbf{A}^T \mathbf{A}) = \text{posto}(\mathbf{A}^T) \quad (4.10)$$

e

$$\text{nulidade}(\mathbf{A}) = \text{nulidade}(\mathbf{A}^T \mathbf{A}). \quad (4.11)$$

Para ilustrar a estrutura da transformação linear seja a matriz:

$$\mathbf{A} = \begin{bmatrix} 4 & 6 & 6 & -2 & 2 \\ 1 & 7 & -4 & 5 & 6 \\ 3 & 8 & 1 & 2 & 5 \end{bmatrix}$$

O produto  $\mathbf{A}\mathbf{x}$  é uma combinação linear das colunas de  $\mathbf{A}$ . As colunas da matriz  $\mathbf{A}$  são multiplicadas por componentes de  $\mathbf{x}$ . Então  $\mathbf{A}\mathbf{x}$  é  $x_1$  vezes [coluna 1] + ... +  $x_n$  vezes [coluna  $n$ ]. A  $i$ -ésima componente de  $\mathbf{A}\mathbf{x}$  é:  $a_{i1}x_1 + a_{i2}x_2 + \dots + a_{in}x_n$ , que será igual a  $\sum_{j=1}^n a_{ij}x_j$ . Retornando ao exemplo, as combinações lineares das colunas da matriz  $\mathbf{A}$  serão dadas por:

$$\mathbf{A}\mathbf{x} = x_1 \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 6 \\ 7 \\ 8 \end{bmatrix} + x_3 \begin{bmatrix} 6 \\ -4 \\ 1 \end{bmatrix} + x_4 \begin{bmatrix} -2 \\ 5 \\ 2 \end{bmatrix} + x_5 \begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix}$$

Os vetores  $\mathbf{x}$  que fazem  $\mathbf{A}\mathbf{x} = \mathbf{0}$  formam o espaço nulo de  $\mathbf{A}$ ,  $\mathcal{N}(\mathbf{A})$ ; desta forma as combinações lineares que geram o espaço nulo serão dadas por:

$$(x_1 + x_2 + 2x_3 - x_4) \begin{bmatrix} 4 \\ 1 \\ 3 \end{bmatrix} + (x_2 - x_3 + x_4 + x_5) \begin{bmatrix} 2 \\ 6 \\ 5 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

Como os vetores linearmente independentes  $[4, 1, 3]^T$  e  $[2, 6, 5]^T$  formam  $\mathfrak{R}(\mathbf{A})$  e são não nulos devemos ter:

$$x_1 + x_2 + 2x_3 - x_4 = 0$$

e

$$x_2 - x_3 + x_4 + x_5 = 0$$

Observemos que o número mínimo de equações que devem ser iguais a zero é igual ao posto da matriz  $\mathbf{A}$ :  $r(\mathbf{A}) = 2$ .

A solução  $\mathbf{x}$  para  $\mathbf{Ax} = 0$  será tal que  $x_1$  e  $x_5$  terão que ser determinados por  $x_2$ ,  $x_3$  e  $x_4$  escolhidos arbitrariamente:

Fazendo as escolhas:

- a)  $x_2 = 1$ ,  $x_3 = 0$  e  $x_4 = 0$  teremos  $x_1 = -1$  e  $x_5 = -1$
- b)  $x_2 = 0$ ,  $x_3 = 1$  e  $x_4 = 0$  teremos  $x_1 = -2$  e  $x_5 = 1$
- c)  $x_2 = 0$ ,  $x_3 = 0$  e  $x_4 = 1$  teremos  $x_1 = 1$  e  $x_5 = -1$

teremos três vetores linearmente independentes:

$$\begin{bmatrix} -1 \\ 1 \\ 0 \\ 0 \\ -1 \end{bmatrix}, \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} \text{ e } \begin{bmatrix} 1 \\ 0 \\ 0 \\ 1 \\ -1 \end{bmatrix}$$

Estes vetores que são solução para  $\mathbf{Ax} = 0$ , formam uma base para  $\mathcal{N}(\mathbf{A})$  que tem nulidade  $\nu(\mathbf{A}) = 3$ .

O espaço linha,  $\mathfrak{R}(\mathbf{A}^T)$  e o espaço nulo  $\mathcal{N}(\mathbf{A}^T)$ , serão obtidos da matriz:

$$\mathbf{A}^T = \begin{bmatrix} 4 & 1 & 3 \\ 6 & 7 & 8 \\ 6 & -4 & 1 \\ -2 & 5 & 2 \\ 2 & 6 & 5 \end{bmatrix}$$

As combinações lineares das colunas da matriz  $\mathbf{A}^T$  serão dadas por:

$$\mathbf{A}^T \mathbf{x} = x_1 \begin{bmatrix} 4 \\ 6 \\ 6 \\ -2 \\ 2 \end{bmatrix} + x_2 \begin{bmatrix} 1 \\ 7 \\ -4 \\ 5 \\ 6 \end{bmatrix} + x_3 \begin{bmatrix} 3 \\ 8 \\ 1 \\ 2 \\ 5 \end{bmatrix}$$

Uma base do espaço nulo de  $\mathbf{A}^T$  é obtida de forma semelhante à do espaço nulo de  $\mathbf{A}$ :

$$\begin{cases} 4x_1 + x_2 + 3x_3 = 0 \\ 6x_1 + 7x_2 + 8x_3 = 0 \\ 6x_1 - 4x_2 + x_3 = 0 \\ -2x_1 + 5x_2 + 2x_3 = 0 \\ 2x_1 + 6x_2 + 5x_3 = 0 \end{cases}$$

Fazendo diversas transformações reduzimos as 5 equações <sup>2</sup> para:

$$\begin{cases} 22x_1 + 13x_3 = 0 \\ 11x_2 + 7x_3 = 0 \end{cases}$$

Os vetores linearmente independentes  $[22, 0, 13]$  e  $[0, 11, 7]$  formam  $\mathcal{R}(\mathbf{A}^T)$  e o vetor  $[-13/22, -7/11, 1]$  é uma base de  $\mathcal{N}(\mathbf{A}^T)$ .

### 4.3 Diferentes Tipos de Solução de $\mathbf{Ax} = \mathbf{b}$

Usando conhecimentos elementares e alguma intuição vejamos como resolver possíveis problemas  $\mathbf{Ax} = \mathbf{b}$ , Figura 4.2.

- **Tipo I** : Sistema consistente: solução única e exata.

**Exemplo 4.1** *Seja o sistema linear:*

$$\begin{cases} -0.7071\mathbf{x}_1 - 0.7071\mathbf{x}_2 = -3.5355 \\ -0.7071\mathbf{x}_1 + 0.7071\mathbf{x}_2 = 0.7071 \end{cases}$$

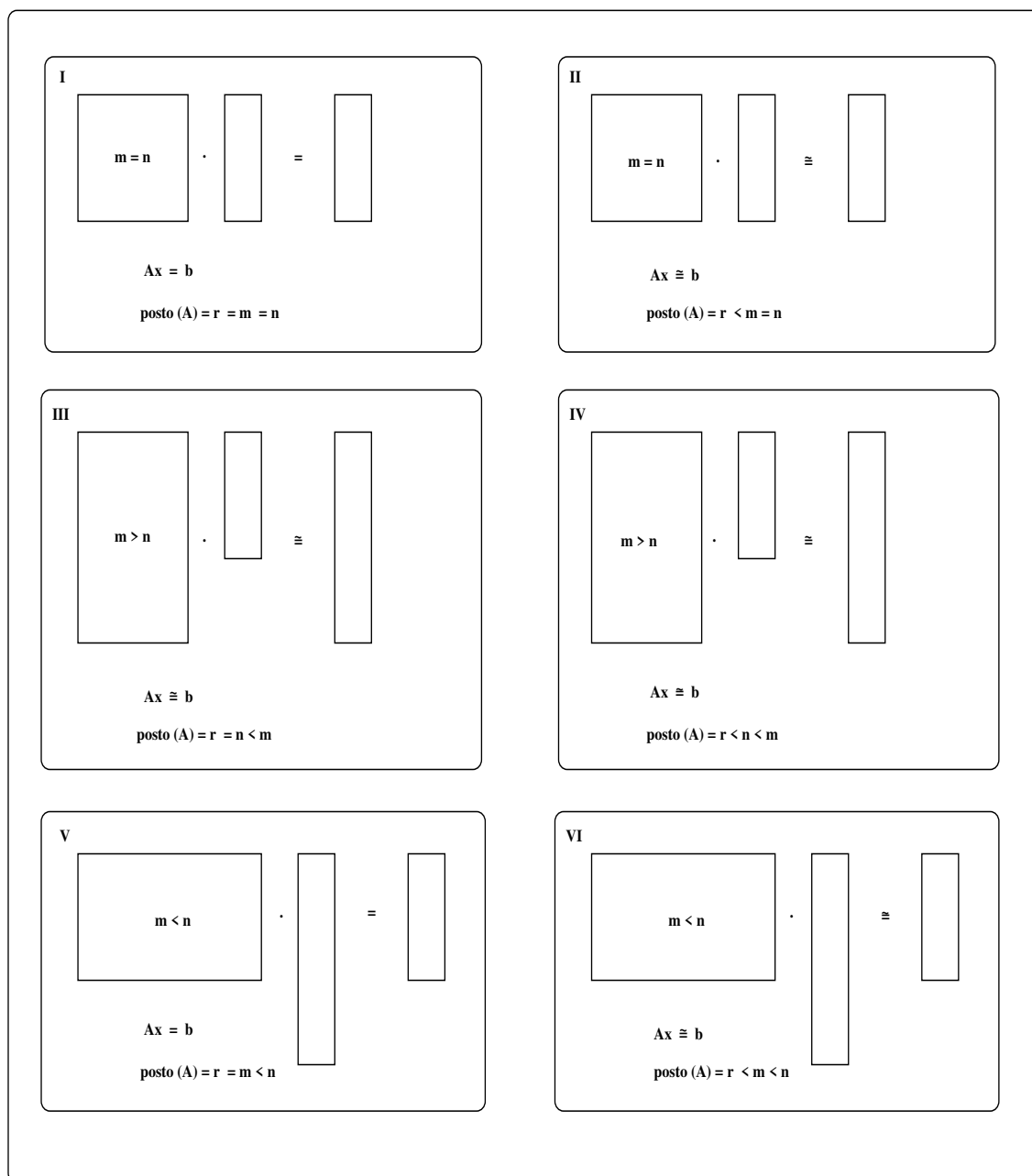
*representado matricialmente por:*

$$\mathbf{Ax} = \begin{bmatrix} -0.7071 & -0.7071 \\ -0.7071 & 0.7071 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} -3.5355 \\ 0.7071 \end{bmatrix} = \mathbf{b}$$

*com  $m = n = r = 2$ ; ele pode ser representado geometricamente por duas retas como mostrado na Figura 4.3, onde vemos que a solução pode ser encontrada através da combinação linear dos vetores coluna da matriz do sistema linear, e é única e exata porque o vetor  $\mathbf{b}$  do sistema  $\mathbf{Ax} = \mathbf{b}$  está no espaço coluna da matriz  $\mathbf{A}$ .*

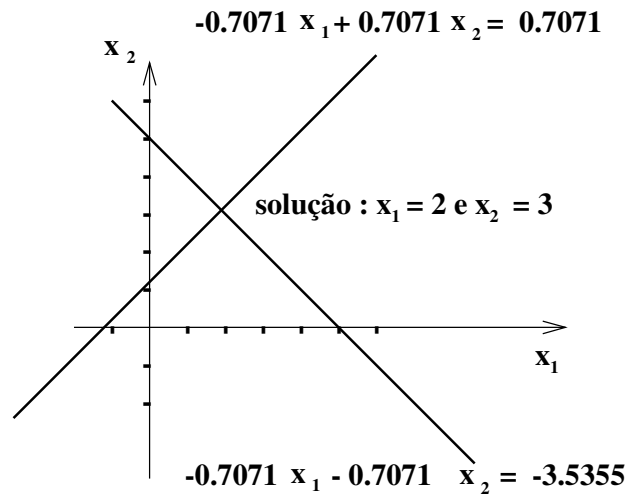
---

<sup>2</sup>A matriz  $\mathbf{R}$  da fatoração  $\mathbf{QR}$  também representa uma base dos espaços linha ou coluna.

Figura 4.2: Possíveis Problemas  $\mathbf{Ax} = \mathbf{b}$ .



**Obtenção da solução  $x$  a partir das linhas do sistema linear.**



**Podemos considerar que o mesmo problema pode ser encontrar a combinação linear dos vetores coluna que conduzem ao vetor  $b$  dado.**

$$A x = [a_1 \ a_2] x = b$$

$$\begin{bmatrix} -0.7071 & -0.7071 \\ -0.7071 & 0.7071 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -3.5355 \\ 0.7071 \end{bmatrix}$$

**O vetor  $b$  será igual a : 2 vezes o vetor  $a_1^T$  mais 3 vezes o vetor  $a_2^T$ . Os pesos 2 e 3 dão a solução  $x = [2 \ 3]$  anteriormente obtida.**

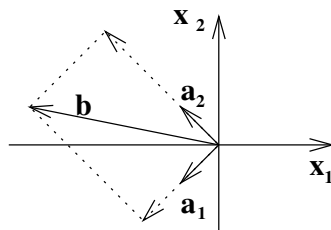


Figura 4.3: Solução Exata e Única.

- **Tipo II** : Matriz  $\mathbf{A}$  singular de posto  $r < n$  .

A solução pode não existir e se existir não ser única.

**Exemplo 4.2** *Seja o sistema  $\mathbf{Ax} = \mathbf{b}$  representado matricialmente por:*

$$\begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

com  $m = n = 2 > r = 1$ . Sua solução não é exata, mas podemos determinar uma solução aproximada.

Como  $\mathbf{A}$  é singular, a inversa de  $\mathbf{A}$ ,  $\mathbf{A}^{-1}$ , não existe e um primeiro passo é multiplicar à esquerda ambos os membros da equação pela transposta da matriz  $\mathbf{A}$  para obter:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \quad (4.12)$$

Se o termo  $\mathbf{A}^T \mathbf{A}$ , quadrado, for regular (não singular) terá uma inversa pela qual multiplicaremos ambos os membros da equação (4.12) obtendo:

$$(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{Ax} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b}$$

donde:

$$\mathbf{Ix} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A}^\dagger \mathbf{b}$$

onde  $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$  é a pseudo-inversa da matriz  $\mathbf{A}$  à esquerda.

Retornando ao exemplo e aplicando o procedimento acima exposto temos:

$$\begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 & 2 \\ 3 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

donde:

$$\mathbf{A}^\dagger = \begin{bmatrix} 0.12 & 0.04 \\ 0.06 & 0.02 \end{bmatrix}$$

e a solução aproximada é dada por:

$$\mathbf{x} = \mathbf{A}^\dagger \mathbf{b} = \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.336 \\ 0.168 \end{bmatrix}$$

donde  $\mathbf{x}$  é a melhor solução aproximada para o sistema.

- **Tipo III:** Sistema sobredeterminado de posto completo:  $m > n = r$ .

Sua solução pode ser exata ou aproximada.

Caso 1: Sistema com solução exata:

**Exemplo 4.3** *Seja o sistema:*

$$\begin{bmatrix} 3 & -1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 10 \end{bmatrix}$$

*Como a terceira linha é linearmente dependente da segunda, a solução é igual à do exemplo do tipo I.*

Caso 2: Sistema com solução aproximada:

Neste caso a resolução é semelhante à do Tipo II, conforme evidencia o exemplo:

**Exemplo 4.4**

$$\mathbf{Ax} = \begin{bmatrix} 3 & -1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix} = \mathbf{b}$$

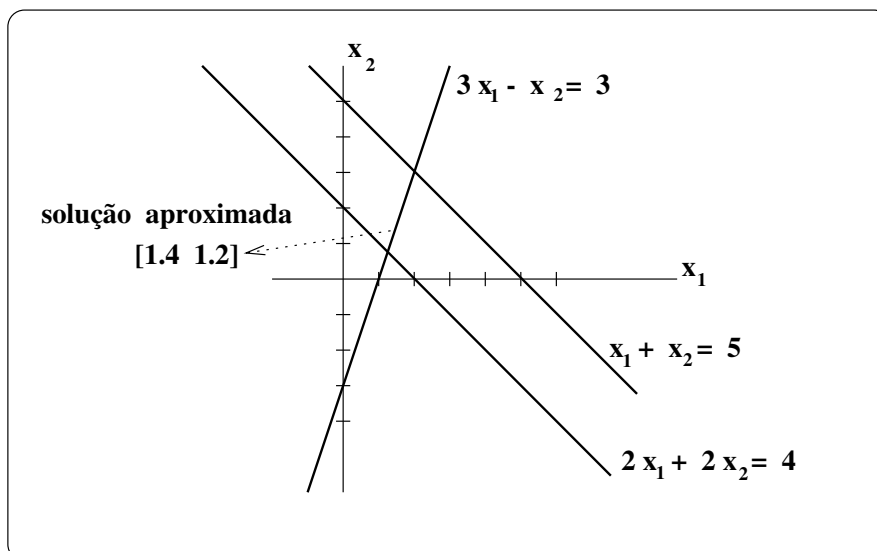


Figura 4.4: Solução Aproximada,  $m > n = r$ .

Como  $\mathbf{A}$  é retangular e singular  $\mathbf{A}^{-1}$  não existe e um primeiro passo é multiplicar à esquerda ambos os termos da equação pela transposta da matriz  $\mathbf{A}$  para obter:

$$\begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 & -1 \\ 1 & 1 \\ 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 & 1 & 2 \\ -1 & 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix}$$

donde

$$\mathbf{A}^\dagger = \begin{bmatrix} 0.2500 & 0.0500 & 0.1000 \\ -0.2500 & 0.1500 & 0.3000 \end{bmatrix}$$

e a solução aproximada é dada por:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 1.4 \\ 1.2 \end{bmatrix}$$

e representada na Figura 4.4.

- **Tipo IV** : Sistema sobredeterminado de posto incompleto:  $r < n < m$ .  
Somente uma solução aproximada pode ser obtida. Sua resolução é semelhante à do tipo III, conforme evidencia o exemplo:

#### Exemplo 4.5

$$\begin{bmatrix} 2 & 1 \\ 2 & 1 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 6 \\ 2 \\ -2 \end{bmatrix}$$

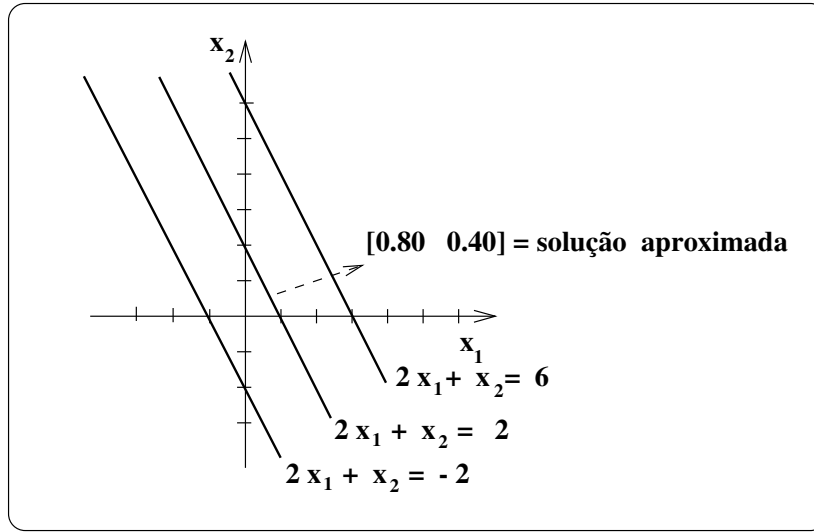
para o qual obtemos, Figura 4.5:

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.80 \\ 0.40 \end{bmatrix}$$

- **Tipo V** : Sistema subdeterminado de posto completo:  $r = m < n$   
A solução se existir não será única.

#### Exemplo 4.6 Seja o sistema:

$$\mathbf{Ax} = \begin{bmatrix} 6 & 3 & 1 \\ 2 & 2 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \mathbf{b}$$

Figura 4.5: Solução Aproximada ,  $r < n < m$ .

Como  $\mathbf{A}$  é retangular e singular, um primeiro passo é multiplicar à esquerda por  $\mathbf{A}^T$  ambos os membros de  $\mathbf{Ax} = \mathbf{b}$  para obter:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T \mathbf{b} \quad (4.13)$$

Porém  $(\mathbf{A}^T \mathbf{A})_{n \times n}$  tem posto  $r = m < n$  e não é invertível; por outro lado  $(\mathbf{AA}^T)_{m \times m}$ , tem posto  $r = m$  e é regular. Multiplicando o segundo membro da equação (4.13) por  $(\mathbf{AA}^T)^{-1}(\mathbf{AA}^T) = \mathbf{I}_m$  tem-se:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^T (\mathbf{AA}^T)^{-1} (\mathbf{AA}^T) \mathbf{b} \quad (4.14)$$

Fazendo  $\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{AA}^T)^{-1}$  chamada pseudo-inversa de  $\mathbf{A}$  à direita pois:

$$\mathbf{A}_{m \times n} \mathbf{A}_{n \times m}^\dagger = \mathbf{I}_m$$

tem-se que:

$$\mathbf{A}^T \mathbf{Ax} = \mathbf{A}^\dagger \mathbf{AA}^T \mathbf{b} \quad (4.15)$$

Fazendo  $\mathbf{x} = \mathbf{A}^\dagger \mathbf{b}$  na equação (4.15) tem-se:

$$\mathbf{A}^T \underbrace{\mathbf{AA}^\dagger}_{\mathbf{I}_m} \mathbf{b} = \mathbf{A}^\dagger \mathbf{AA}^T \mathbf{b} = \mathbf{A}^T \underbrace{(\mathbf{AA}^T)^{-1} \mathbf{AA}^T}_{\mathbf{I}_m} \mathbf{b} = \mathbf{A}^T \mathbf{b}$$

donde  $\mathbf{A}^T \mathbf{b} = \mathbf{A}^T \mathbf{b}$  e portanto  $\mathbf{x} = \mathbf{A}^\dagger \mathbf{b}$  é solução aproximada para este tipo de problema.

*Para o exemplo tem-se a solução aproximada:*

$$\mathbf{x} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.0526 \\ 0.3684 \\ 0.5789 \end{bmatrix}$$

- **Tipo VI:** Sistema subdeterminado de posto incompleto:  $r < m < n$

A solução pode ser obtida pela utilização da matriz pseudo-inversa à direita.

**Exemplo 4.7** *Seja o sistema:*

$$\begin{bmatrix} 6 & 3 & 1 \\ 12 & 6 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

*Portanto a solução aproximada será:*

$$\bar{\mathbf{x}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.1565 \\ 0.0783 \\ 0.0261 \end{bmatrix}$$

Destes exemplos tratados de forma elementar vemos que através de pseudo-inversa os tipos III a VI nos remetem aos tipos I e II, para os quais pode-se obter uma solução exata ou aproximada conforme conhecimento usual, que consideramos insatisfatório para os objetivos deste trabalho. Um estudo mais adequado aos nossos objetivos, que reputamos essencial além de fundamental, é a seguir apresentado.

## 4.4 Matrizes Projeção Ortogonal e Pseudo Inversa de $\mathbf{A}$ à Esquerda

Retornamos à definição anterior de projeção ortogonal de um vetor em uma reta pela origem, mas com o objetivo de obter expressão de projeção para matrizes.

A matriz  $\mathbf{A}$  tem  $n$  colunas:  $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \dots, \mathbf{a}_n$ . As combinações destas colunas em  $\mathcal{R}^m$  são os vetores  $\mathbf{Ax}$  no espaço coluna de  $\mathbf{A}$ . Inicialmente recapitulemos o teorema 2.11 de uma forma diferente e consideremos a projeção de um vetor  $\mathbf{b}$  sobre um vetor  $\mathbf{a}_i$  do espaço coluna de  $\mathbf{A}$ , Figura 4.6.

Ao projetarmos o vetor  $\mathbf{b}$  sobre o ponto  $\mathbf{p}$  em  $\mathbf{a}_i$ , a junção de  $\mathbf{p}$  com  $\mathbf{b}$  forma o vetor  $\mathbf{e} = \mathbf{b} - \mathbf{p}$  que será mínimo quando  $\mathbf{p}$  for a projeção ortogonal de  $\mathbf{b}$ . Como a projeção  $\mathbf{p}$  é algum múltiplo de  $\mathbf{a}_i$ , pode-se escrever:

$$\mathbf{p} = x\mathbf{a}_i \quad \text{e} \quad \bar{x} = \frac{\mathbf{a}_i^T \mathbf{b}}{\mathbf{a}_i^T \mathbf{a}_i}$$

é o valor de  $x$  que minimiza  $\mathbf{e}$ , cujo mínimo é  $\bar{\mathbf{e}}$ .

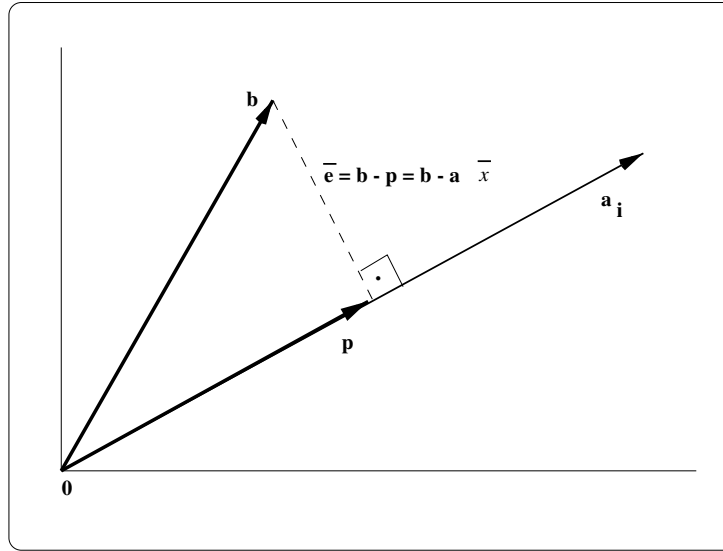


Figura 4.6: Projeção de um vetor  $\mathbf{b}$  sobre um vetor  $\mathbf{a}_i$ .

Para a projeção ortogonal de  $\mathbf{b}$  sobre  $\mathbf{a}_i$  tem-se  $\mathbf{b} - \mathbf{p} = \mathbf{b} - \mathbf{a}_i \bar{x}$  e:

$$\mathbf{p} = \mathbf{a}_i \bar{x} = \frac{\mathbf{a}_i \mathbf{a}_i^T \mathbf{b}}{\mathbf{a}_i^T \mathbf{a}_i}$$

Consideremos a projeção ortogonal do vetor  $\mathbf{b}$  sobre o espaço coluna de  $\mathbf{A}_{m \times n}$  com  $m > n = r$ , Figura 4.7.

O vetor de erro  $\bar{\mathbf{e}} = \mathbf{b} - \mathbf{A}\bar{\mathbf{x}}$  é perpendicular ao subespaço gerado pela combinação linear dos vetores,  $[\mathbf{a}_1^T, \mathbf{a}_2^T, \dots, \mathbf{a}_n^T]$  do espaço coluna de  $\mathbf{A}$ .

No espaço nulo de  $\mathbf{A}^T$  está  $\bar{\mathbf{e}}$ , que pode ser descrito por  $n$  equações na forma:

$$\begin{aligned} \mathbf{a}_1^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) &= 0 \\ &\vdots \\ \mathbf{a}_n^T (\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) &= 0 \end{aligned}$$

e representado pelo esquema da equação (4.16):

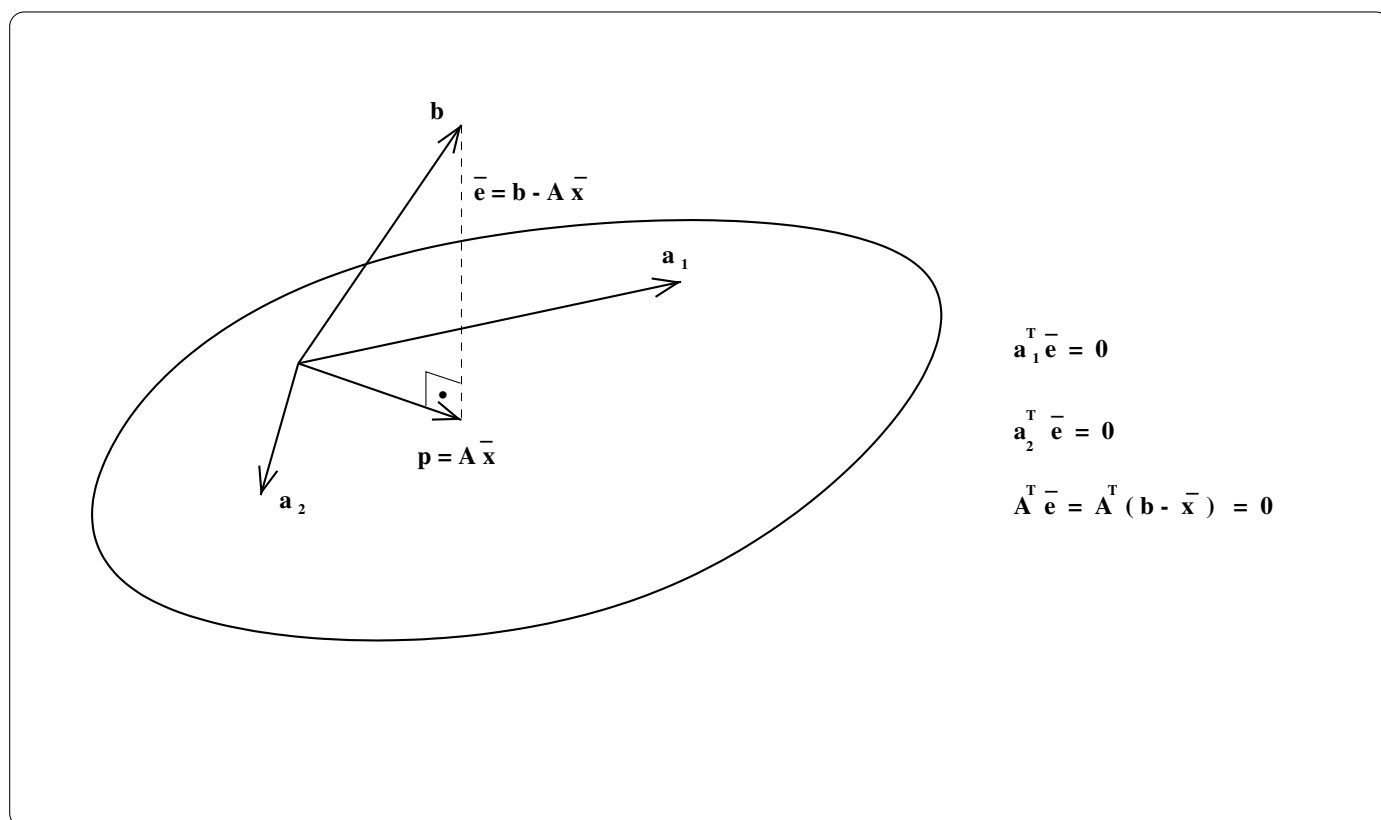


Figura 4.7: Projeção ortogonal de um vetor  $\mathbf{b}$  sobre um subespaço gerado por  $\mathbf{A}$ .



$$\begin{bmatrix} \leftarrow & \mathbf{a}_1^T & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{a}_n^T & \rightarrow \end{bmatrix} \begin{bmatrix} \mathbf{b} - \mathbf{A}\bar{\mathbf{x}} \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix} \quad (4.16)$$

Portanto:  $\mathbf{A}^T(\mathbf{b} - \mathbf{A}\bar{\mathbf{x}}) = 0$ , donde:

$$\mathbf{A}^T \mathbf{A} \bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$$

O valor ótimo de  $\mathbf{x}$  será:

$$\bar{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A}^\dagger \mathbf{b}$$

onde a matriz  $\mathbf{A}^\dagger$ :

$$\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

é a pseudo-inversa da matriz  $\mathbf{A}$  à esquerda.

Observe que o estimador ótimo  $\bar{\mathbf{x}}$  corresponde à projeção ortogonal  $\mathbf{p}$  de  $\mathbf{b}$  no espaço coluna da matriz  $\mathbf{A}$  e que:

$$\mathbf{p} = \mathbf{A}\bar{\mathbf{x}} = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A}\mathbf{A}^\dagger \mathbf{b} \quad (4.17)$$

Portanto, a matriz de projeção ortogonal ou meramente projetor ortogonal no espaço coluna da matriz  $\mathbf{A}$  é dada por:

$$P = \mathbf{A}(\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T = \mathbf{A}\mathbf{A}^\dagger \quad (4.18)$$

e

$$\mathbf{p} = P\mathbf{b} \quad (4.19)$$

Comentários:

- $r = \text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}^T \mathbf{A}) = n$ , logo  $\mathbf{A}^T \mathbf{A}$  é regular.
- $\mathbf{A}^T \mathbf{A}$  é regular e a pseudo-inversa à esquerda é:

$$\mathbf{A}_e^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

A Figura 4.8 ilustra a pseudo-inversa à esquerda.

- Os autovalores da matriz  $P$  são iguais a 0 ou 1. Contudo uma matriz que tenha autovalores iguais a 0 ou 1 não é necessariamente uma matriz de projeção.
- A matriz projeção é quadrada e tem como propriedades:

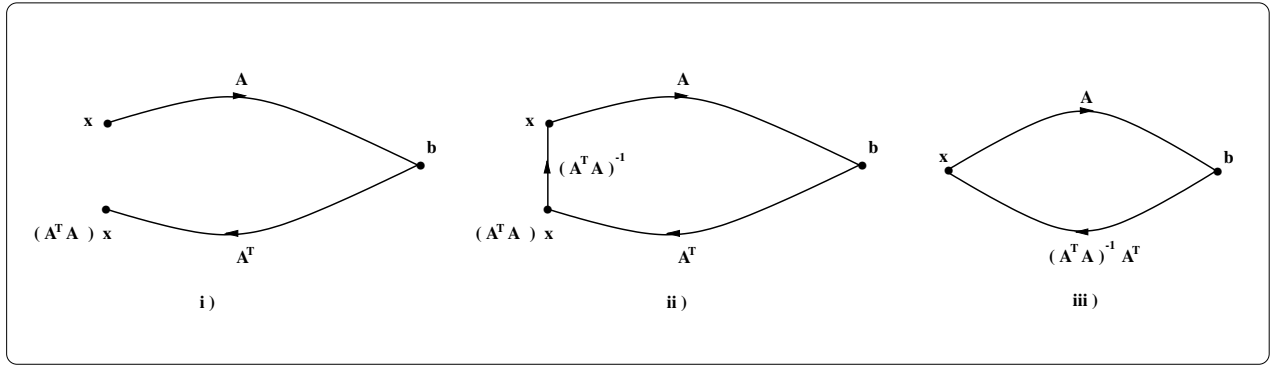


Figura 4.8: Pseudo-inversa à esquerda.

a) Idempotência:

$$P^2 = P$$

b) Para matriz projeção ortogonal:

$$P^T = P$$

c) Se  $P$  é um projetor,  $\mathbf{I} - P$  também é um projetor:

$$(\mathbf{I} - P)^2 = \mathbf{I} - 2P + P^2 = \mathbf{I} - P$$

A matriz  $\mathbf{I} - P$  é chamada projetor complementar de  $P$ .

## 4.5 Solução do Problema $\mathbf{Ax} = \mathbf{b}$ nos Subespaços Fundamentais

Neste ponto estabelecemos algumas questões fundamentais sobre a solução do sistema de equações lineares  $\mathbf{Ax} = \mathbf{b}$  tais como:

- a) Quando o sistema tem uma solução, e se ele a tem, ela é única ?
- b) Se ele não tem solução, o que fazer ?

Para responder questões deste tipo primeiramente consideramos o teorema 4.1 e então analisamos suas consequências, tendo por base a estrutura da transformação linear  $\mathbf{A}_{m \times n}$ .

**Teorema 4.1** A equação:

$$\mathbf{Ax} = \mathbf{b} \tag{4.20}$$

com  $\mathbf{A}_{m \times n}$  e  $\mathbf{b} \in \mathbf{R}^m$  dados e  $\mathbf{x} \in \mathbf{R}^n$  como incógnita:

a) *tem solução  $\mathbf{x}$  se e somente se:*

$$\mathbf{b} \in \mathfrak{R}(\mathbf{A}) = \mathfrak{R}(\mathbf{AA}^T) \quad (4.21)$$

*ou equivalentemente:*

$$\mathbf{b} \perp \mathcal{N}(\mathbf{A}^T) = \mathcal{N}(\mathbf{AA}^T) \quad (4.22)$$

*ou equivalentemente:*

$$\text{posto}(\mathbf{A}) = \text{posto}([\mathbf{A} \ \mathbf{b}]) \quad (4.23)$$

b) *Há no máximo uma solução se e somente se:*

$$\mathcal{N}(\mathbf{A}) = \mathcal{N}(\mathbf{A}^T \mathbf{A}) = \{0\} \quad (4.24)$$

*ou equivalentemente:*

$$\text{posto}(\mathbf{A}) = n \quad (4.25)$$

c) *Há uma solução para cada  $\mathbf{b} \in \mathbf{R}^m$  se e somente se:*

$$\mathfrak{R}(\mathbf{A}) = \mathfrak{R}(\mathbf{AA}^T) = \mathbf{R}^m \quad (4.26)$$

*ou equivalentemente:*

$$\text{posto}(\mathbf{A}) = m \quad (4.27)$$

d) *Há uma e somente uma solução para cada  $\mathbf{b} \in \mathbf{R}^m$  se e somente se  $\mathbf{A}$  for quadrada e regular, caso em que a relação é:*

$$\mathbf{x} = \mathbf{A}^{-1} \mathbf{b} \quad (4.28)$$

*e temos  $\text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}^T) = m = n = r$*

### Prova:

- a) A condição (4.21) re-enuncia a definição de espaço coluna de  $\mathbf{A}$ :  $\mathbf{Ax}$  deve pertencer a  $\mathfrak{R}(\mathbf{A})$ ,  $\forall \mathbf{x} \in \mathbf{R}^n$ . A condição (4.22) decorre da aplicação do conceito de espaço nulo. As relações (4.26) e (4.27) simplesmente repetem o Lema 4.1. A relação (4.23) reconhece que  $\mathbf{b} \in \mathfrak{R}(\mathbf{A})$  se e somente se  $\mathbf{b}$  for linearmente dependente das colunas de  $\mathbf{A}$ .

- b) Suponha que  $\mathcal{N}(\mathbf{A})$  contenha somente o vetor zero e que há múltiplas soluções de (4.20), isto é:

$$\mathbf{Ax}_1 = \mathbf{Ax}_2 = \mathbf{b} \quad (4.29)$$

então

$$\mathbf{Ax}_1 - \mathbf{Ax}_2 = \mathbf{A}(\mathbf{x}_1 - \mathbf{x}_2) = \mathbf{0} \quad (4.30)$$

o que significa que  $(\mathbf{x}_1 - \mathbf{x}_2) \in \mathcal{N}(\mathbf{A})$  e concluímos que  $\mathbf{x}_1 = \mathbf{x}_2$ .

Por outro lado, se  $\mathbf{x}$  é uma solução (4.20) e  $\mathcal{N}(\mathbf{A})$  contém um vetor  $\mathbf{x} \neq \mathbf{0}$ , então:

$$\mathbf{A}(\mathbf{x} + \mathbf{x}_h) = \mathbf{Ax} + \mathbf{Ax}_h = \mathbf{b} + \mathbf{0} \quad (4.31)$$

donde  $(\mathbf{x} + \mathbf{x}_h)$  também é uma solução de (4.20).

A condição (4.25) segue de (4.5).

- c) Segue trivialmente de a).  
 d) Neste caso ambos b) e c) devem ocorrer, isto é,  $\text{posto}(\mathbf{A}) = n = m$ , implicando em  $\mathbf{A}$  quadrada e regular.

**C.Q.D.**

Em consequência deste teorema e das discussões anteriores podemos tecer os comentários a seguir:

O sistema  $\mathbf{Ax} = \mathbf{b}$  é solucionável se e somente se o vetor  $\mathbf{b}$  puder ser expresso como uma combinação linear das colunas da matriz  $\mathbf{A}$ . Como o vetor  $\mathbf{b}$  nem sempre estará no espaço coluna de  $\mathbf{A}$  então o problema a ser resolvido é o de determinar a melhor projeção (projeção ortogonal) do vetor  $\mathbf{b}$  que está no espaço coluna de  $\mathbf{A}$ .

A representação geométrica da resolução do problema  $\mathbf{Ax} = \mathbf{b}$  nos subespaços fundamentais é apresentada na Figura 4.9 e podemos observar que:

- Os vetores  $\mathbf{x}_h \in \mathcal{N}(\mathbf{A})$  fazem  $\mathbf{Ax}_h = \mathbf{0}$
- Os vetores  $\mathbf{x}_p \in \mathcal{R}(\mathbf{A})$  fazem  $\mathbf{Ax}_p = \mathbf{b}$ .
- Os vetores do espaço linha de  $\mathbf{A}$  ou espaço coluna de  $\mathbf{A}^T$  são:  
 $\mathbf{x}_p = \mathbf{A}^T \mathbf{b}$ .

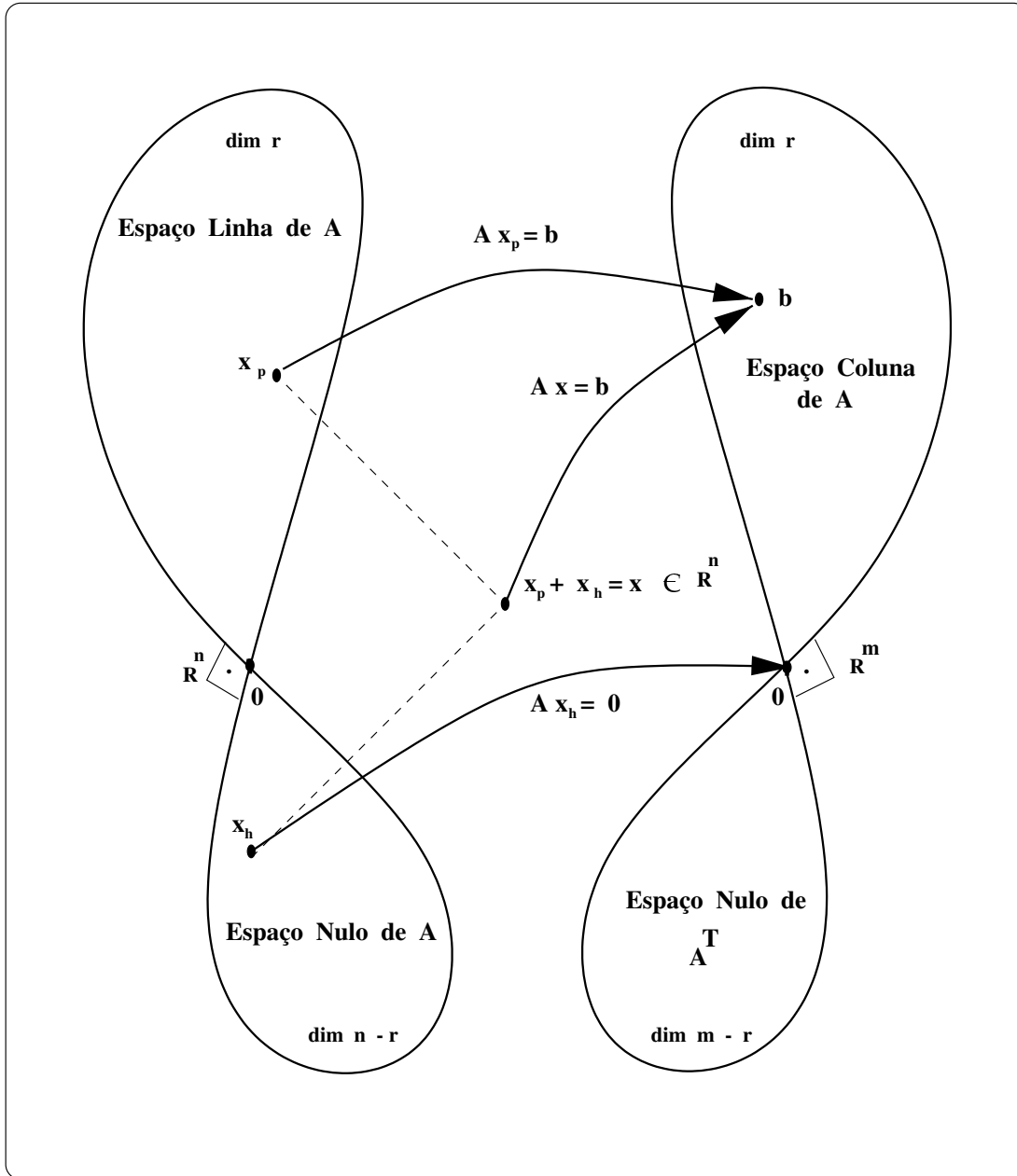


Figura 4.9: Representação Geométrica nos Subespaços Fundamentais da Solução de  $\mathbf{Ax} = \mathbf{b}$ .

- Cada vetor no espaço coluna de  $\mathbf{A}$  corresponde à um e somente um vetor  $\mathbf{x}_p$  do espaço linha. Para demonstrar este fato recorreremos a dois vetores,  $\mathbf{x}_p$  e  $\mathbf{x}'_p$  do espaço linha. Se  $\mathbf{Ax}_p = \mathbf{Ax}'_p$ , a diferença  $\mathbf{x}_p - \mathbf{x}'_p$  deve estar no espaço nulo. Como por suposição inicial  $\mathbf{x}_p$  e  $\mathbf{x}'_p$  estão no espaço linha, e este é perpendicular ao espaço nulo, então teremos que  $\mathbf{x}_p = \mathbf{x}'_p$ .
- Como os espaços linha de  $\mathbf{A}$  e nulo de  $\mathbf{A}$  são ortogonalmente complementares, um vetor  $\mathbf{x}$  qualquer pode ser decomposto nos vetores componentes  $\mathbf{x}_p$  do espaço linha e  $\mathbf{x}_h$  do espaço nulo, Figura 4.9. Em decorrência temos  $\mathbf{Ax}_p = \mathbf{Ax}$ .
- O conjunto de soluções do sistema  $\mathbf{Ax} = \mathbf{b}$  pode ser:
  - vazio,  $\mathbf{Ax} = \mathbf{b}$  não tem solução.
  - um ponto,  $\mathbf{Ax} = \mathbf{b}$  tem solução única, neste caso diz-se que o sistema de equações é consistente.
  - infinito,  $\mathbf{Ax} = \mathbf{b}$  tem infinitas soluções.
- Se  $\mathbf{x}_h$  é solução para o sistema homogêneo  $\mathbf{Ax}_h = 0$ , e  $\mathbf{x}_p$  é uma solução particular para  $\mathbf{Ax} = \mathbf{b}$ , então:

$$\mathbf{Ax}_p + \mathbf{Ax}_h = \mathbf{A}(\mathbf{x}_p + \mathbf{x}_h) = \mathbf{b} + 0 = \mathbf{b}$$

portanto  $\mathbf{x}_p + \mathbf{x}_h$  é outra solução para  $\mathbf{Ax} = \mathbf{b}$ .

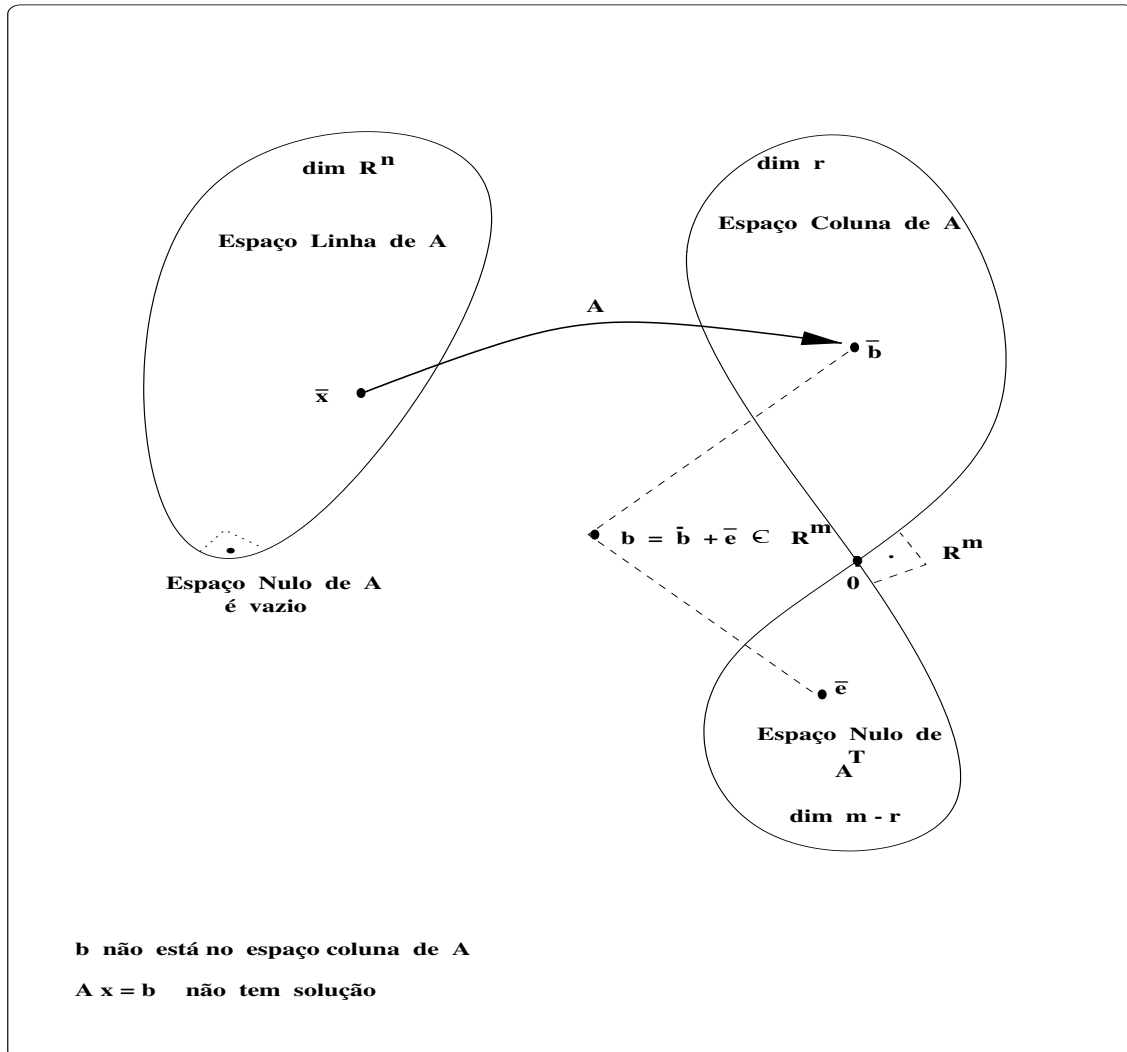
- Se os vetores  $\mathbf{x}$  e  $\mathbf{x}_p$  são solução para  $\mathbf{Ax} = \mathbf{b}$  tal que  $\mathbf{Ax} = \mathbf{b}$  e  $\mathbf{Ax}_p = \mathbf{b}$ , então:

$$\mathbf{Ax} - \mathbf{Ax}_p = \mathbf{A}(\mathbf{x} - \mathbf{x}_p) = \mathbf{b} - \mathbf{b} = 0$$

portanto  $\mathbf{x}_h = \mathbf{x} - \mathbf{x}_p$  é solução do sistema homogêneo.

- Suponha que  $\mathbf{x}_p$  é alguma solução particular para o sistema  $\mathbf{Ax} = \mathbf{b}$ . Então o conjunto de todas soluções  $\mathbf{x}$  para o sistema  $\mathbf{Ax} = \mathbf{b}$  será idêntico ao conjunto de soluções  $\mathbf{x}_p + \mathbf{x}_h$ , onde  $\mathbf{x}_h$  é o conjunto das soluções do sistema homogêneo,  $\mathbf{Ax}_h = 0$ . Isto mostra que se existir mais de uma solução para o sistema  $\mathbf{Ax} = \mathbf{b}$ , existirá um  $\mathbf{x}_h$  não nulo tal que  $\mathbf{Ax}_h = 0$ .
- A solução completa do sistema  $\mathbf{Ax} = \mathbf{b}$  é constituída por uma solução particular mais todas as soluções homogêneas, onde:

- Solução particular  $\mathbf{x}_p$  satisfaz  $\mathbf{Ax}_p = \mathbf{b}$ .
- Solução geral da homogênea,  $\mathbf{x}_h$ , satisfaz  $\mathbf{Ax}_h = \mathbf{0}$ .
- A solução de  $\mathbf{Ax} = \mathbf{b}$  será única se e somente se o espaço nulo  $\mathcal{N}(\mathbf{A})$  for vazio, ou seja, se a solução para  $\mathbf{Ax} = \mathbf{0}$  for  $\mathbf{x} = \mathbf{0}$ , Figura 4.10.
- O sistema  $\mathbf{Ax} = \mathbf{b}$  terá solução exata se e somente se  $\mathbf{b}$  estiver no espaço coluna de  $\mathbf{A}$ , Figura 4.10.


 Figura 4.10: Representação Geométrica de  $\bar{\mathbf{e}} = \mathbf{b} - \bar{\mathbf{b}}$ .

- Se a solução de  $\mathbf{Ax} = \mathbf{b}$  não existe então busca-se a solução  $\bar{\mathbf{b}} = \mathbf{A}\bar{\mathbf{x}}$  de mínimos quadrados da equação  $\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$ , onde  $\bar{\mathbf{b}}$  é a projeção

ortogonal de  $\mathbf{b}$  no espaço coluna de  $\mathbf{A}$ , tal que  $\bar{\mathbf{x}}$  produza o menor erro possível,  $\bar{\mathbf{e}} = \mathbf{b} - \bar{\mathbf{b}}$ , Figura 4.10.

- O valor ótimo de  $\mathbf{x}$ ,  $\bar{\mathbf{x}}$ , que minimiza  $\|\mathbf{Ax} - \mathbf{b}\|^2$  constitui a solução de mínimos quadrados do problema  $\mathbf{Ax} = \mathbf{b}$ . Como já demonstrado tal solução é dada por:

$$\bar{\mathbf{x}} = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T \mathbf{b} = \mathbf{A}^\dagger \mathbf{b}$$

onde se supõe que  $\mathbf{A}^T \mathbf{A}$  seja regular.

- A fatoração  $\mathbf{QR}$  é útil na solução de mínimos quadrados do problema  $\mathbf{Ax} = \mathbf{b}$ , como visto a seguir.

Dado que:

$$\mathbf{A}^T \mathbf{A} = \mathbf{R}^T \mathbf{Q}^T \mathbf{Q} \mathbf{R} = \mathbf{R}^T \mathbf{R}$$

a solução de  $\mathbf{A}^T \mathbf{A} \bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$  pode ser simplificada para:

$$\mathbf{R}^T \mathbf{R} \bar{\mathbf{x}} = \mathbf{R}^T \mathbf{Q}^T \mathbf{b}$$

donde,

$$\mathbf{R} \bar{\mathbf{x}} = \mathbf{Q}^T \mathbf{b}$$

Deste modo em lugar de resolver  $\mathbf{Ax} = \mathbf{b}$ , o que as vezes não é trivial, resolve-se  $\mathbf{R} \bar{\mathbf{x}} = \mathbf{Q}^T \mathbf{b}$  por substituição para trás.

- Observe no Exemplo 4.1 que os vetores coluna da matriz  $\mathbf{A}$  são ortonormais, o que implica que:  $\mathbf{A}^T \mathbf{A} = \mathbf{I}$ . A solução de  $\mathbf{Ax} = \mathbf{b}$  neste caso é  $\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$ .

A matriz de projeção é simplificada para  $\mathbf{P} = \mathbf{A} \mathbf{A}^T = \mathbf{I}$ , logo não existe matriz para ser invertida e a solução do problema é imediata, daí uma importante razão para o emprego de bases ortonormais.

As componentes de  $\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b}$  são produtos internos de  $\mathbf{b}$  com as linhas de  $\mathbf{A}^T$ :

$$\bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b} = \begin{bmatrix} \text{linha 1 de } \mathbf{A}^T \\ \text{linha 2 de } \mathbf{A}^T \\ \vdots \\ \text{linha } n \text{ de } \mathbf{A}^T \end{bmatrix} \mathbf{b} = \begin{bmatrix} \mathbf{a}_1^T \mathbf{b} \\ \mathbf{a}_2^T \mathbf{b} \\ \vdots \\ \mathbf{a}_n^T \mathbf{b} \end{bmatrix}$$



$$\begin{aligned}
\mathbf{p} &= \mathbf{A}\bar{\mathbf{x}} = \left[ \begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{array} \right] \left[ \begin{array}{c} \mathbf{a}_1^T \mathbf{b} \\ \mathbf{a}_2^T \mathbf{b} \\ \vdots \\ \mathbf{a}_n^T \mathbf{b} \end{array} \right] = \\
&= \mathbf{a}_1(\mathbf{a}_1^T \mathbf{b}) + \mathbf{a}_2(\mathbf{a}_2^T \mathbf{b}) + \dots + \mathbf{a}_n(\mathbf{a}_n^T \mathbf{b})
\end{aligned}$$

Quando  $\mathbf{A}$  for quadrada o subespaço é o espaço completo, então:

$$\mathbf{A}^T = \mathbf{A}^{-1} \quad \text{e} \quad \bar{\mathbf{x}} = \mathbf{A}^T \mathbf{b} = \mathbf{A}^{-1} \mathbf{b}.$$

A solução é exata. A projeção de  $\mathbf{b}$  ocorre sobre ele mesmo. Neste caso  $\mathbf{p} = \mathbf{b}$  e  $\mathbf{P} = \mathbf{A}\mathbf{A}^T = \mathbf{I}$ .

## 4.6 Solução de Mínimos Quadrados de Sistema Sobredeterminado Através de Matriz Pseudo-Inversa

Na Figura 4.11 temos uma representação geométrica da pseudo-inversa na solução de problema  $\mathbf{Ax} = \mathbf{b}$  por mínimos quadrados. Tal problema dado que o vetor  $\mathbf{b}$  não está no espaço coluna da matriz  $\mathbf{A}$ , terá uma solução aproximada. Pela projeção ortogonal do vetor  $\mathbf{b}$  no espaço coluna de  $\mathbf{A}$ :

$$\bar{\mathbf{b}} = \mathbf{P}\mathbf{b} \quad (4.32)$$

tem-se:

$$\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}} \quad (4.33)$$

e

$$\bar{\mathbf{x}} = \mathbf{A}^\dagger \bar{\mathbf{b}} \quad (4.34)$$

constitui a solução ótima de mínimos quadrados.

**Teorema 4.2** *Seja  $\mathbf{A}_{m \times n}$ ,  $n < m$ , com posto completo  $r = n$ . Então:*

- $\mathbf{Ax} = \mathbf{b}$  tem no máximo uma solução e pode ter nenhuma. Se  $\mathbf{b} \in \mathcal{R}(\mathbf{A})$  existe uma única solução; se  $\mathbf{x}_h \neq 0$ , não existe solução.*
- Nos casos em que  $\mathbf{Ax} = \mathbf{b}$  não tem solução,  $\bar{\mathbf{x}} = \mathbf{A}_e^\dagger \mathbf{b}$  é a melhor aproximação à uma solução no sentido que  $\|\mathbf{A}\bar{\mathbf{x}} - \mathbf{b}\| < \|\mathbf{Ax} - \mathbf{b}\|$  para todos os demais  $\mathbf{x} \in \mathbf{R}^n$ .*

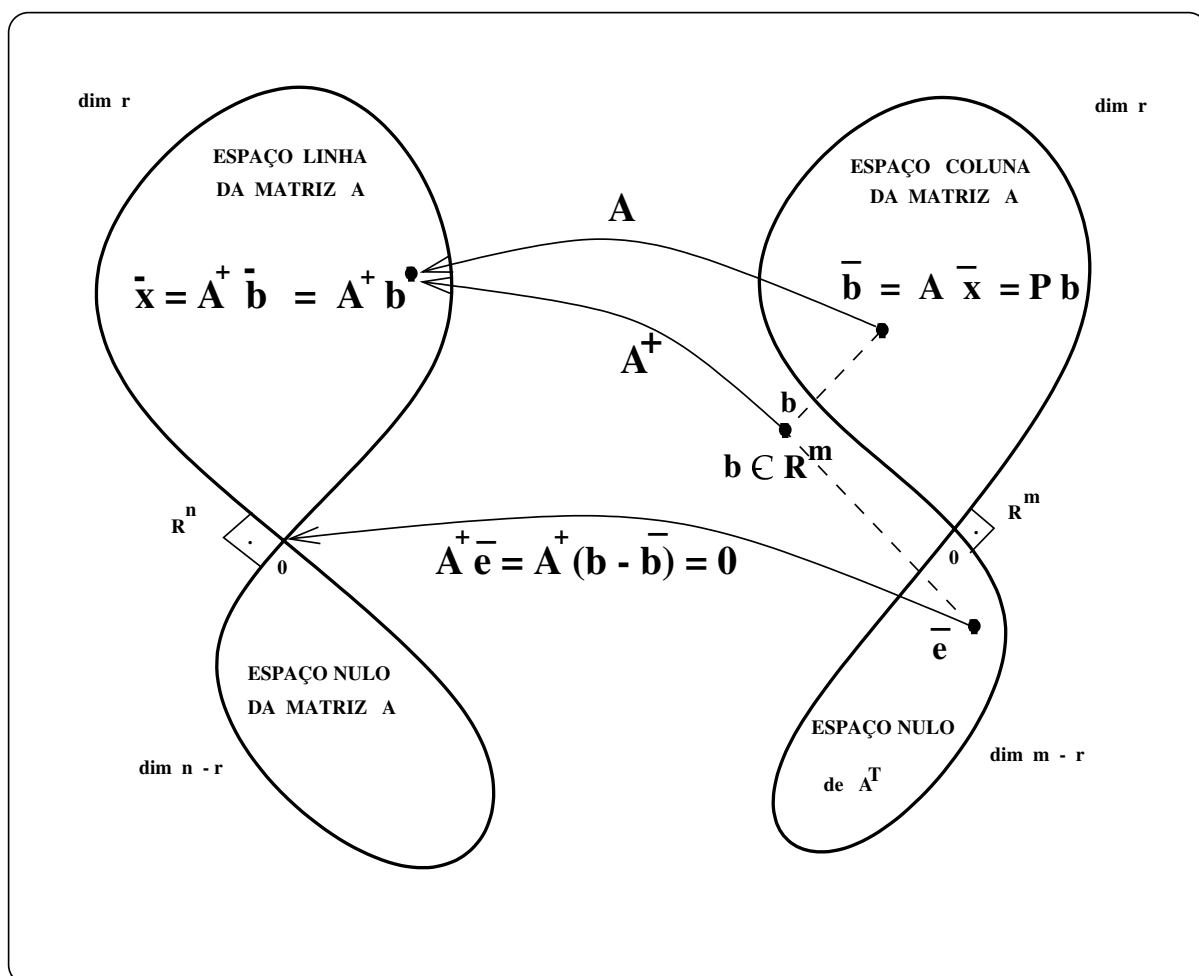


Figura 4.11: Representação Geométrica de Solução por Mínimos Quadrados Através da Pseudo-Inversa.

**Prova:**

a) Óbvio, pela Figura 4.11: estrutura da transformação e projeção ortogonal: como  $\mathbf{b} \in \mathbf{R}^m$  e  $\mathbf{b} = \bar{\mathbf{b}} + \bar{\mathbf{e}}$  onde  $\bar{\mathbf{b}} \in \mathcal{R}(\mathbf{A})$  e  $\bar{\mathbf{e}} \in \mathcal{N}(\mathbf{A}^T)$  :

- se  $\bar{\mathbf{e}} = 0$ , existe uma única solução, pois  $\mathbf{b} = \bar{\mathbf{b}} \in \mathcal{R}(\mathbf{A})$ .
- se  $\bar{\mathbf{e}} \neq 0$ , não existe solução.

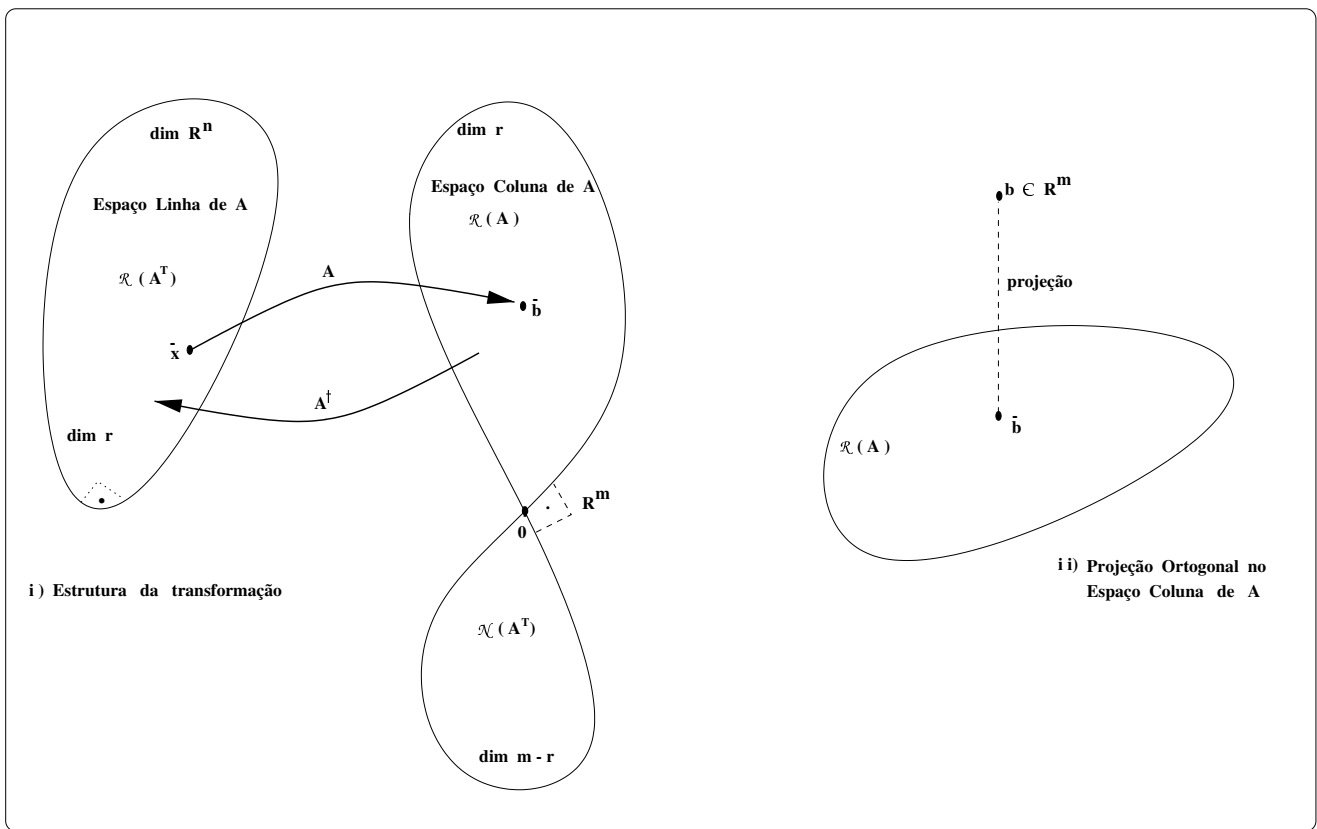


Figura 4.12: Pseudo-inversa à esquerda: i) estrutura da transformação e ii) projeção ortogonal no espaço coluna de  $\mathbf{A}$  .

b) Pode-se escrever  $\mathbf{e} = \bar{\mathbf{e}} + \hat{\mathbf{e}}$  onde tem-se:

- i) o erro:  $\mathbf{e} = \mathbf{b} - \mathbf{Ax}$  ,
- ii) o erro:  $\bar{\mathbf{e}} = \mathbf{b} - \bar{\mathbf{b}} \in \mathcal{N}(\mathbf{A}^T)$  ,
- iii) o resíduo:  $\hat{\mathbf{e}} = \bar{\mathbf{b}} - \mathbf{Ax} \in \mathcal{R}(\mathbf{A})$

$$\mathbf{e} = \mathbf{b} - \mathbf{A}\mathbf{x} = \bar{\mathbf{e}} + \hat{\mathbf{e}} = (\mathbf{b} - \bar{\mathbf{b}}) + (\bar{\mathbf{b}} - \mathbf{A}\mathbf{x})$$
$$\| \mathbf{b} - \mathbf{Ax} \|^2 = \| \mathbf{b} - \bar{\mathbf{b}} \|^2 + \| \bar{\mathbf{b}} - \mathbf{Ax} \|^2$$
$$\| \mathbf{b} - \mathbf{A}\bar{\mathbf{x}} \|^2 < \| \mathbf{b} - \mathbf{A}\mathbf{x} \|^2 \quad \text{para} \quad \mathbf{x} \neq \bar{\mathbf{x}}$$

A Figura 4.13 ilustra a solução aproximada e os respectivos erros .

**C.Q.D.**

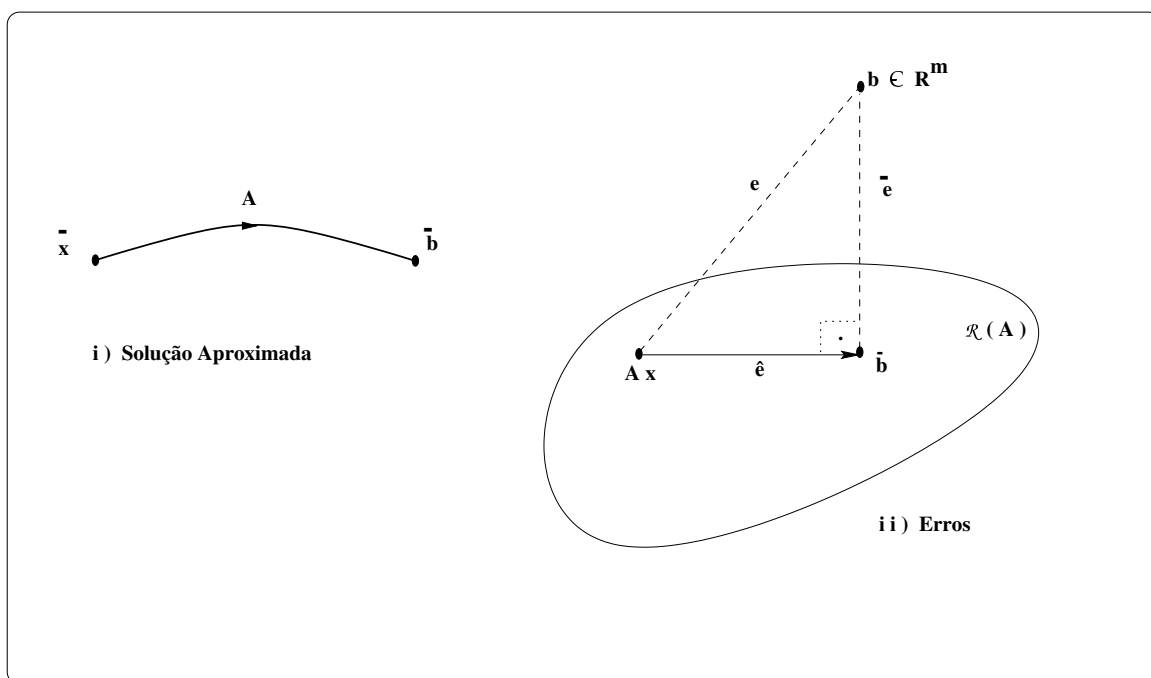


Figura 4.13: Solução Aproximada e Erros.

Observemos que para o sistema sobredeterminado podemos tecer os seguintes comentários:

- Se  $\mathbf{A}_{m \times n}$  tem posto  $r = n$  então:

$$\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$

A equação  $Ax = b$ ,  $m > n$  :

$$\begin{array}{c}
 \text{n colunas} \\
 \boxed{A_{m \times n}} \\
 \text{m linhas}
 \end{array}
 \begin{array}{c}
 \text{desconhecido} \\
 \boxed{x_{n \times 1}}
 \end{array}
 =
 \begin{array}{c}
 \text{conhecido} \\
 \boxed{b_{m \times 1}}
 \end{array}$$

se a matriz  $A$  tem posto  $n$ , existe uma matriz  $A_{n \times m}^+$  tal que :

$$\boxed{A_{n \times m}^+} \boxed{A_{m \times n}} = \boxed{I_{n \times n}}$$

Deste modo podemos resolver a equação inicial como :

$$\boxed{x_{n \times 1}} = \boxed{A_{n \times m}^+} \boxed{b_{m \times 1}}$$

Onde  $A_{n \times m}^+$  é a pseudo-inversa à esquerda dada por :

$$\boxed{A_{n \times m}^+} = \boxed{(A^T A)^{-1}} \boxed{A^T}$$

Figura 4.14: Sistema sobredeterminado: Pseudo-inversa à esquerda.

e

$$\mathbf{A}^\dagger \mathbf{A} = \mathbf{I}_{n \times n}$$

Esta forma de pseudo-inversa, pseudo-inversa à esquerda, é esquematizada na Figura 4.14.

- A pseudo-inversa à esquerda corresponde à projeção ortogonal do vetor  $\mathbf{b}$  sobre o espaço coluna da matriz  $\mathbf{A}$ , para a qual observamos que:  $\mathbf{b} \in \mathbf{R}^m$ ,  $\bar{\mathbf{b}} \in \mathfrak{R}(\mathbf{A})$ :

$$\bar{\mathbf{b}} = P\mathbf{b}$$

$$P = \mathbf{A}\mathbf{A}^\dagger$$

$$\bar{\mathbf{e}} = \mathbf{b} - \bar{\mathbf{b}} = \mathbf{b} - P\mathbf{b} = (\mathbf{I} - P)\mathbf{b}$$

onde  $(\mathbf{I} - P)$  é o chamado projetor complementar; ele projeta  $\mathbf{b}$  no espaço nulo de  $\mathbf{A}$ .

**Exemplo 4.8** *Seja o sistema  $\mathbf{A}\mathbf{x} = \mathbf{b}$  representado matricialmente por:*

$$\begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

com  $m = n = 2 > r = 1$ . Sua solução não é exata, mas podemos determinar uma solução aproximada, que seja a melhor possível. O espaço coluna da matriz  $\mathbf{A}$  é dado pelo vetor  $[3 \ 1]^T$ . Se projetarmos o vetor  $\mathbf{b}$  sobre  $\bar{\mathbf{b}}$  no espaço coluna de  $\mathbf{A}$ ,  $\bar{\mathbf{e}} = \mathbf{b} - \bar{\mathbf{b}}$  será o valor mínimo de  $\mathbf{e}$  quando  $\bar{\mathbf{b}}$  for a projeção ortogonal de  $\mathbf{b}$ .

Deste modo:

$$\bar{\mathbf{b}} = k[3 \ 1]^T$$

onde  $k$  é um escalar constante.

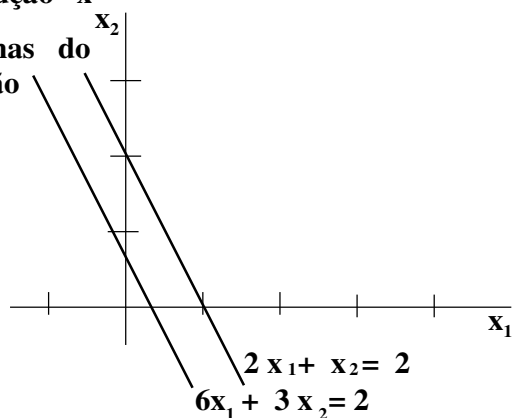
A melhor projeção obtida está na direção do vetor  $[3 \ 1]^T$  que é ortogonal ao vetor  $[3 \ 1]^T$ . Observe que o produto escalar destes vetores ortogonais é  $[3 \ 1]^T \bullet [3 \ 1]^\perp = 0$ .

Deste modo:  $[3 \ 1]^\perp = [-1 \ 3]$ . Então podemos supor que  $\mathbf{b}$  resulta da combinação linear com pesos  $\lambda_1$  e  $\lambda_2$  dos vetores acima:

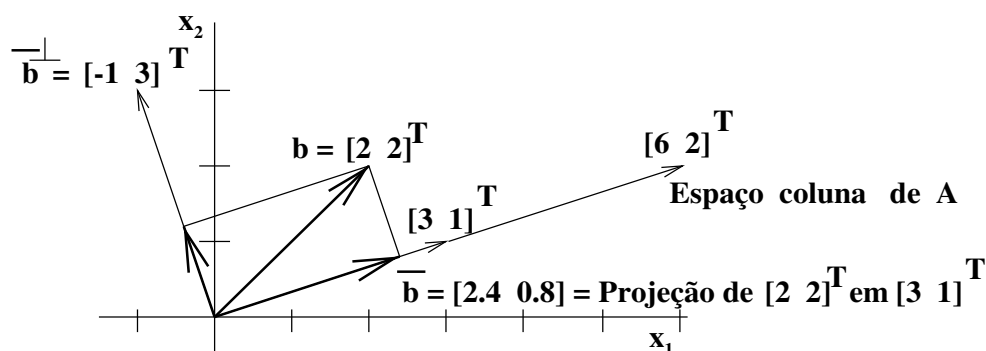
$$\mathbf{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \lambda_1 \begin{bmatrix} 3 \\ 1 \end{bmatrix} + \lambda_2 \begin{bmatrix} -1 \\ 3 \end{bmatrix}$$

Resolvendo este sistema determinamos  $\lambda_1 = 0.8$  e  $\lambda_2 = 0.4$ ; como  $\bar{\mathbf{b}}$  é a projeção ortogonal do vetor  $\mathbf{b}$  no espaço coluna da matriz, tem-se:

Obtenção da solução  $\bar{x}$  a partir das linhas do sistema linear não é exata. Pode ser aproximada.



Através da projeção ortogonal do vetor  $b$  no espaço coluna da matriz  $A$  encontramos um vetor  $\bar{b}$ .



A solução aproximada é obtida quando  $x$  atua sobre  $A$  para produzir  $\bar{b}$  no espaço coluna da matriz  $A$ .

Para este exemplo  $\bar{x} = [0.336 \ 0.168]$ .

Figura 4.15: Solução Aproximada,  $m = n > r$ .

$$\bar{\mathbf{b}} = Proj_{[3 \ 1]^T} \begin{bmatrix} 2 \\ 2 \end{bmatrix} = 0.8 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 2.4 \\ 0.8 \end{bmatrix}$$

Portanto quando o vetor  $\mathbf{b}$  estiver fora do espaço coluna da matriz  $\mathbf{A}$  podemos criar uma base ortogonal utilizando o espaço perpendicular ao espaço coluna, conforme mostramos na Figura 4.15.

Para determinar a melhor solução do sistema original deve-se resolver o sistema aproximado:

$$\begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2.4 \\ 0.8 \end{bmatrix}$$

portanto:

$$\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b} = \mathbf{A}^\dagger \bar{\mathbf{b}} = \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 2 \\ 2 \end{bmatrix} = \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix}^\dagger \begin{bmatrix} 2.4 \\ 0.8 \end{bmatrix} = \begin{bmatrix} 0.336 \\ 0.168 \end{bmatrix}$$

onde  $\bar{\mathbf{x}} = [x_1 \ x_2]^T = [0.336 \ 0.168]^T$  é a melhor solução aproximada para o sistema.

**Exemplo 4.9** Para o sistema:

$$\mathbf{Ax} = \mathbf{b} \Rightarrow \begin{bmatrix} 1 \\ 2 \end{bmatrix} \mathbf{x} = \begin{bmatrix} 10 \\ 4 \end{bmatrix}$$

podemos ter como solução  $\mathbf{x} = 10$  ou  $\mathbf{x} = 2$ , mas estamos interessados em uma única solução e que esta solução seja a melhor.

Se existe uma matriz  $\mathbf{A}^\dagger$  de tal forma que  $\mathbf{A}^\dagger \mathbf{A} = \mathbf{I}$  podemos encontrar uma solução ótima para o sistema fazendo:

$$\mathbf{A}^\dagger \mathbf{A} = \mathbf{I} \Rightarrow \mathbf{A}^\dagger \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \mathbf{I} \quad e \quad \mathbf{A}^\dagger = [\mathbf{a}_1 \ \mathbf{a}_2]$$

onde

$$\mathbf{a}_1 + 2\mathbf{a}_2 = 1$$

Pela Figura 4.16 que ilustra este exemplo vemos no lugar  $\mathbf{A}^\dagger$  todas as soluções para  $\mathbf{a}_1 + 2\mathbf{a}_2 = 1$ . Logo  $\mathbf{A}^\dagger$  para o exemplo dado não é única. Então buscamos a solução de norma mínima e a pseudo-inversa à esquerda,  $\mathbf{A}_e^\dagger$ , correspondente. Para esta solução tem-se:

$$\mathbf{A}_e^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$$



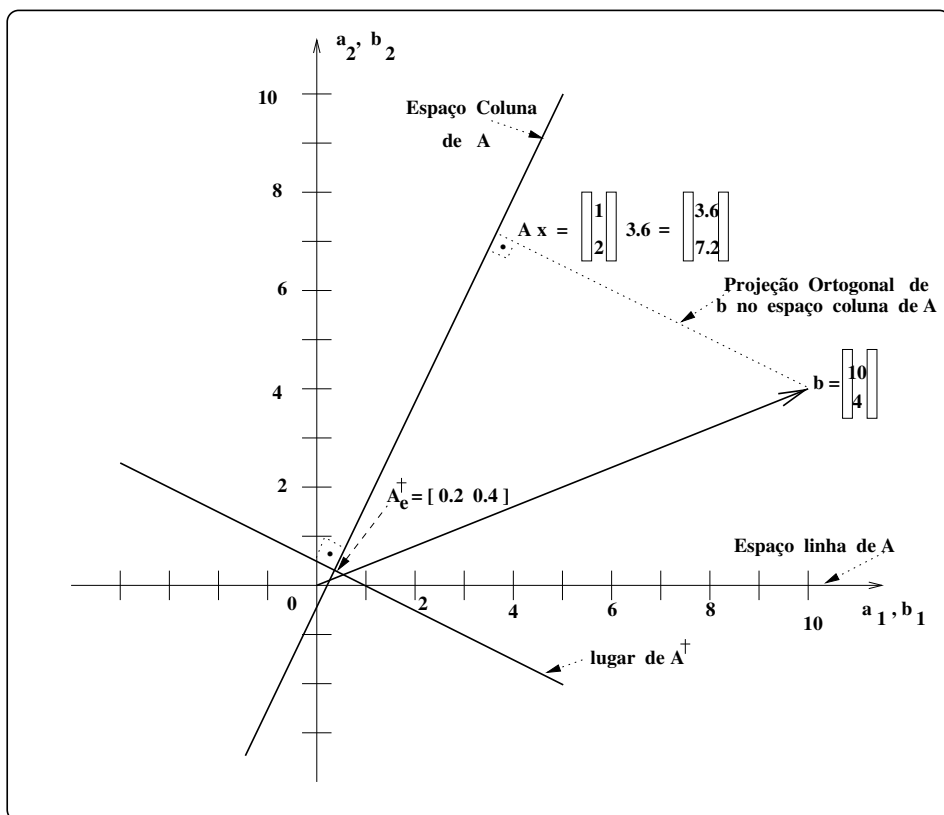


Figura 4.16: Representação geométrica: pseudo-inversa à esquerda.

Como

$$\mathbf{A}^T \mathbf{A} = (1)^2 + (2)^2 = 5$$

$$\mathbf{A}^\dagger = (5)^{-1} [1 \ 2] = [0.2 \ 0.4]$$

donde o melhor valor de  $\mathbf{x}$  é:

$$\bar{\mathbf{x}} = [0.2 \ 0.4] \begin{bmatrix} 10 \\ 4 \end{bmatrix} = 3.6$$

Portanto a melhor solução corresponde:

$$\bar{\mathbf{b}} = \begin{bmatrix} 1 \\ 2 \end{bmatrix} 3.6 = \begin{bmatrix} 3.6 \\ 7.2 \end{bmatrix}$$

Observe ainda pela Figura 4.16 que a solução encontrada é a projeção ortogonal de  $\mathbf{b}$  em  $\mathbf{A}$ .

## 4.7 Matrizes Projeção Ortogonal e Pseudo-Inversa à Direita

Consideremos a projeção ortogonal  $\bar{\mathbf{x}}$  do vetor  $\mathbf{b}$  sobre o espaço linha de  $\mathbf{A}_{m \times n}$  com  $r = m < n$ . A matriz  $\mathbf{A}$  tem  $m$  linhas:

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}$$

onde  $\mathbf{a}_i^T$  é o  $i$ -ésimo vetor linha  $n$ -dimensional de  $\mathbf{A}$ ,  $i = 1, 2, \dots, m$ . Como as  $m$  linhas de  $\mathbf{A}$  são linearmente independentes, o vetor  $\bar{\mathbf{x}} \in \Re(\mathbf{A}^T) \equiv \mathbf{R}^m$  e como  $\mathbf{x} = \bar{\mathbf{x}} + \mathbf{x}_h \in \mathbf{R}^n$  tem-se  $\mathbf{x}_h \in \mathcal{N}(\mathbf{A})$ , e então  $\mathbf{A}\mathbf{x}_h = 0$  donde:

$$\mathbf{b} = \mathbf{A}\mathbf{x} = \mathbf{A}(\bar{\mathbf{x}} + \mathbf{x}_h) = \mathbf{A}\bar{\mathbf{x}} \quad (4.35)$$

Para  $\bar{\mathbf{x}} \in \Re(\mathbf{A}^T)$  tem-se:

$$\bar{\mathbf{x}} = \sum_{i=1}^m \alpha_i \mathbf{c}_i = \mathbf{A}^T \alpha \quad (4.36)$$

onde  $\mathbf{c}_i$   $i = 1, 2, 3, \dots, m$  é um conjunto de vetores linearmente independentes no espaço linha de  $\mathbf{A}$ . Das equações (4.35) e (4.36):

$$\mathbf{A}\mathbf{A}^T \alpha = \mathbf{b} \in \mathbf{R}^m$$

donde

$$\alpha = (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$$

pois  $\mathbf{A}\mathbf{A}^T$  é regular; então:

$$\bar{\mathbf{x}} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{b}$$

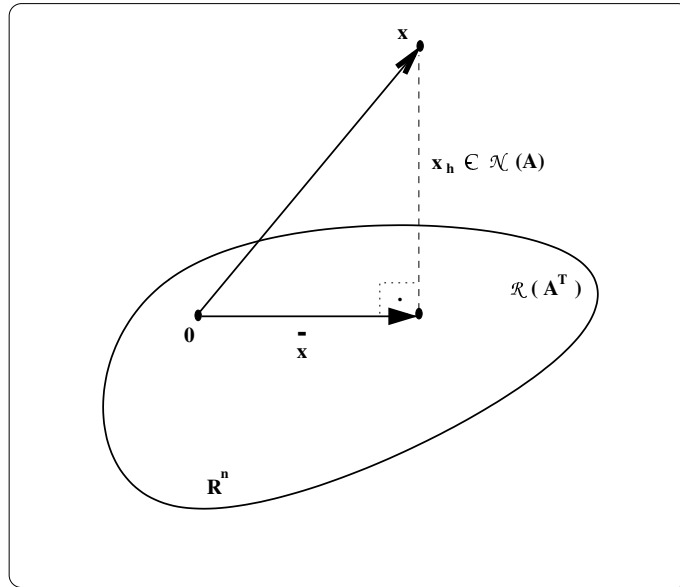
é a projeção ortogonal de  $\mathbf{b}$  no espaço linha de  $\mathbf{A}$ .

Portanto  $\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}$ , onde  $\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$  é a pseudo-inversa de  $\mathbf{A}$  à direita e como  $\mathbf{b} = \mathbf{A}\mathbf{x}$ :

$$\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{A}\mathbf{x} = P_l \mathbf{x} \quad (4.37)$$

$P_l \mathbf{x}$  é a projeção ortogonal de  $\mathbf{x}$  em  $\Re(\mathbf{A}^T)$ , onde a matriz de projeção ortogonal no espaço linha da matriz  $\mathbf{A}$  é dada por:

$$P_l = \mathbf{A}^\dagger \mathbf{A} = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1} \mathbf{A} \quad (4.38)$$

Figura 4.17: Projeção no espaço linha de  $\mathbf{A}$ .

e então  $\mathbf{I}_m - P_l$  é a projeção ortogonal em  $\mathcal{N}(\mathbf{A})$ , que é o subespaço ortogonal ao gerado por  $\mathbf{a}_i$ ,  $i = 1, 2, \dots, m$ . Esta projeção está ilustrada na Figura 4.17

Observemos que para o sistema subdeterminado podemos tecer os seguintes comentários:

- Se  $\mathbf{A}_{m \times n}$  tem posto  $= m$  então:

$$\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$$

e

$$\mathbf{A} \mathbf{A}^\dagger = \mathbf{I}_{m \times m}$$

Esta forma de pseudo-inversa, pseudo-inversa à direita, é esquematizada na Figura 4.18.

- A projeção ortogonal do vetor  $\mathbf{b}$  agora é sobre o espaço linha da matriz  $\mathbf{A}$  e é obtida através da pseudo-inversa à direita.
- 

$$r = \text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A} \mathbf{A}^T) = m$$

logo  $\mathbf{A} \mathbf{A}^T$  é regular.

A equação  $Ax = b$ ,  $m < n$  :

$$\begin{array}{c}
 \text{desconhecido} \\
 \begin{array}{ccc}
 & \text{n colunas} & \\
 & \boxed{A_{m \times n}} & \\
 \text{m linhas} & & 
 \end{array}
 \end{array}
 \begin{array}{c}
 \boxed{x_{n \times 1}} \\
 \\
 \end{array}
 =
 \begin{array}{c}
 \text{conhecido} \\
 \boxed{b_{m \times 1}}
 \end{array}$$

se a matriz  $A$  tem posto  $n$ , existe uma matriz  $A_{n \times m}^+$  tal que :

$$\begin{array}{ccc}
 \boxed{A_{m \times n}} & \boxed{A_{n \times m}^+} & = \boxed{I_{m \times m}}
 \end{array}$$

Deste modo podemos resolver a equação inicial como :

$$\begin{array}{ccc}
 \boxed{x_{n \times 1}} & = & \boxed{A_{n \times m}^+} \boxed{b_{m \times 1}}
 \end{array}$$

Onde  $A_{n \times m}^+$  é a pseudo-inversa à direita dada por :

$$\begin{array}{ccc}
 \boxed{A_{n \times m}^+} & = & \boxed{A_{n \times m}^T} \boxed{(AA^T)^{-1}_{m \times m}}
 \end{array}$$

Figura 4.18: Sistema subdeterminado: Pseudo-inversa à direita.

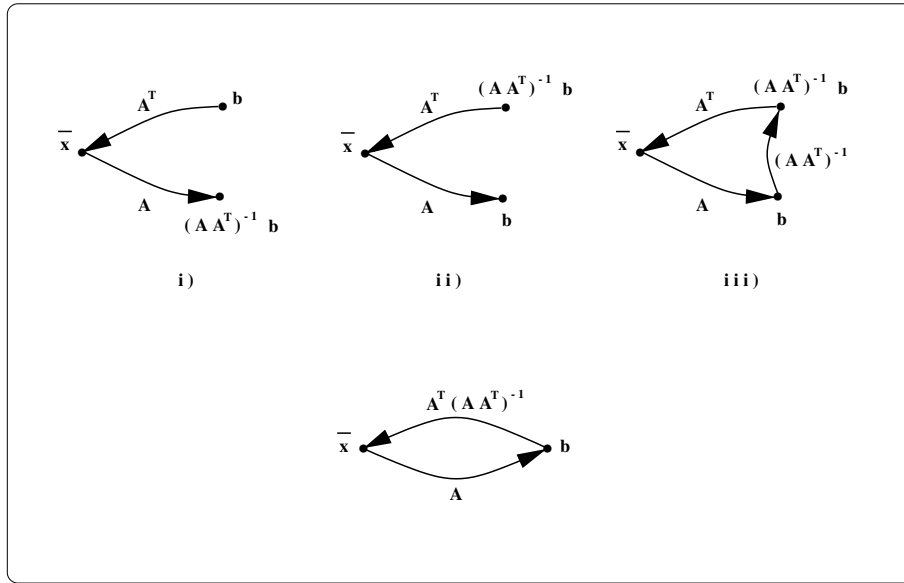


Figura 4.19: Pseudo-inversa à direita.

- $\mathbf{A}\mathbf{A}^T$  é regular e a pseudo-inversa à direita é:

$$\mathbf{A}_d^\dagger = \mathbf{A}^T (\mathbf{A}\mathbf{A}^T)^{-1}$$

A Figura 4.19 ilustra a pseudo-inversa à direita.

## 4.8 Solução de Mínimos Quadrados de Sistema Subdeterminado Através de Matriz Pseudo-Inversa

Na Figura 4.20 temos uma representação geométrica da pseudo-inversa na solução de  $\mathbf{A}\mathbf{x} = \mathbf{b}$ ,  $m < n$ , por mínimos quadrados. Este problema sem solução, dado que o vetor  $\mathbf{x}$  não está no espaço linha da matriz  $\mathbf{A}$  terá uma solução aproximada que será a solução ótima obtida pela projeção do vetor  $\mathbf{x}$  no espaço linha de  $\mathbf{A}$ :  $\bar{\mathbf{x}} = P_l \mathbf{x}$ , após o qual  $\mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}}$  tem solução  $\bar{\mathbf{x}} = \mathbf{A}_d^\dagger \bar{\mathbf{b}}$ .

**Teorema 4.3** *Seja  $\mathbf{A}_{m \times n}$ ,  $m < n$ , com posto completo  $r = m$ . Então:*

- $\mathbf{A}\mathbf{x} = \mathbf{b}$  tem solução para qualquer  $\mathbf{b} \in \mathbf{R}^m$ .
- O vetor  $\bar{\mathbf{x}} = \mathbf{A}_d^\dagger \mathbf{b}$  é a solução de norma mínima de  $\mathbf{A}\mathbf{x} = \mathbf{b}$ .

**Prova:**

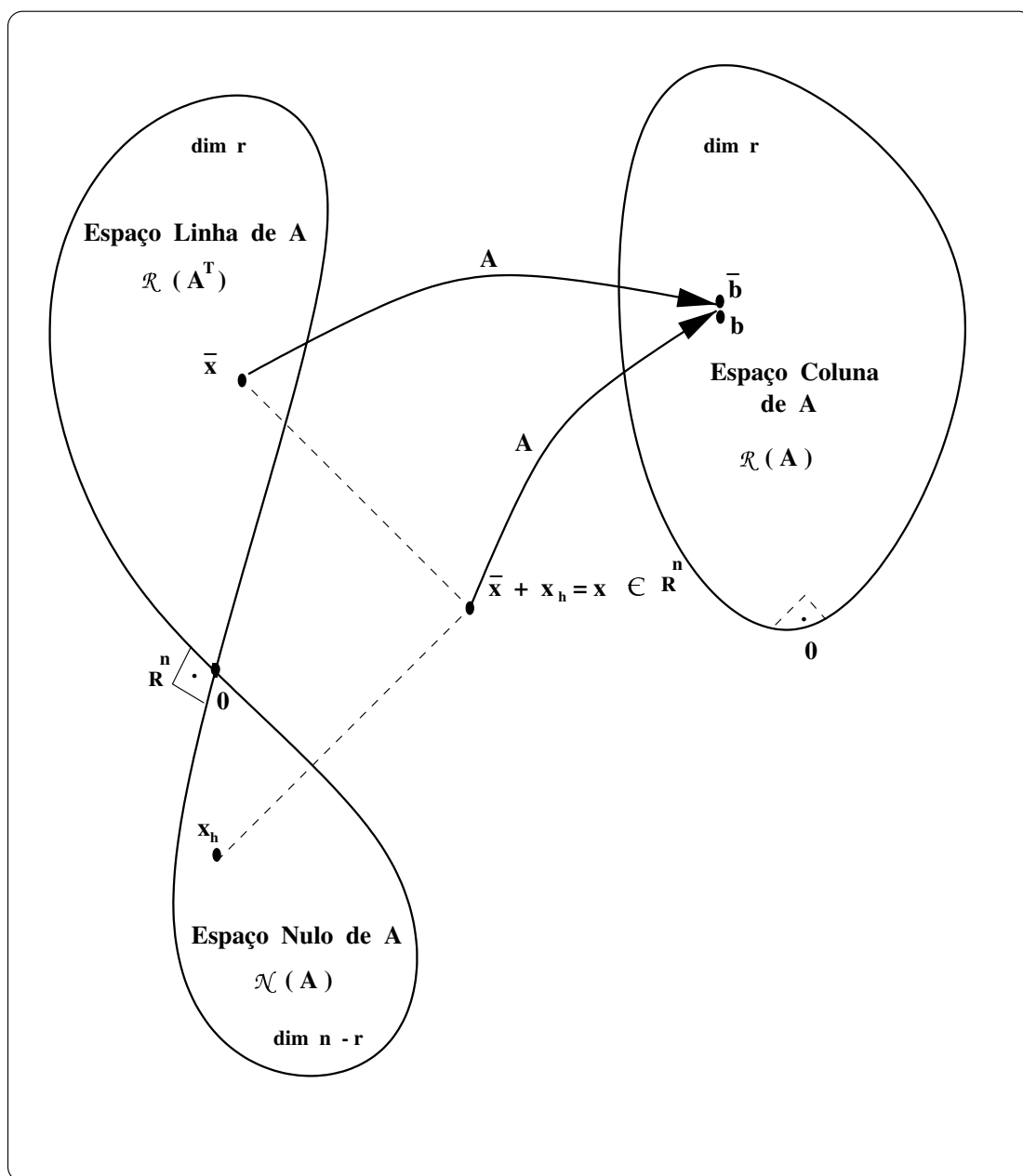


Figura 4.20: Solução de mínimos quadrados de  $\mathbf{Ax} = \mathbf{b}$ ,  $m < n$ .

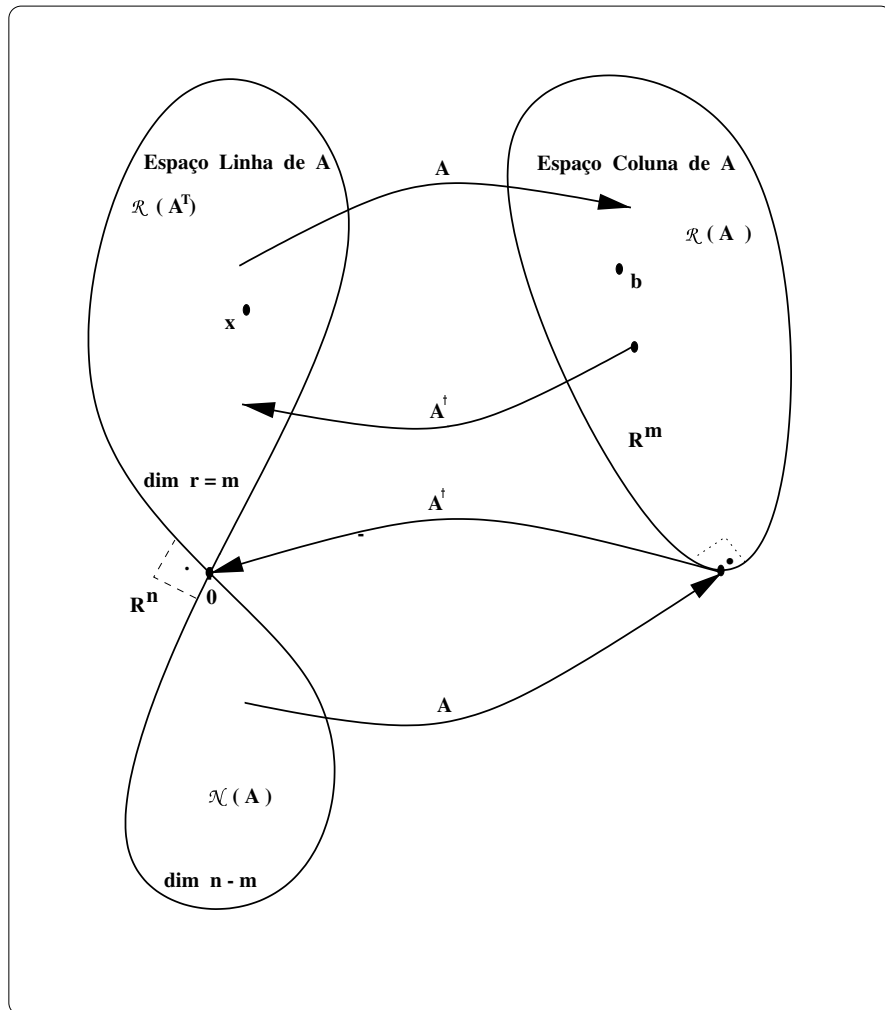


Figura 4.21: Pseudo-inversa à direita: estrutura da transformação.

- a) Óbvia da estrutura da transformação, Figura 4.21
- b)  $\bar{\mathbf{x}}$  é uma solução:

$$\mathbf{A}\bar{\mathbf{x}} = \mathbf{A}\mathbf{A}^\dagger \mathbf{b} = \mathbf{A}\mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b} = \mathbf{b}$$

$\bar{\mathbf{x}}$  é a solução de norma mínima: como visto, para qualquer  $\mathbf{x}$  dado,  $\mathbf{x} = \mathbf{x}_p + \mathbf{x}_h$  é solução, onde  $\mathbf{x}_h \in \mathcal{N}(\mathbf{A})$ , mas existe um único  $\mathbf{x}_p = \bar{\mathbf{x}} \in \mathfrak{R}(\mathbf{A}^T)$  que é dado por  $\bar{\mathbf{x}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}$  e pelo teorema de Pitágoras:  $\|\mathbf{x}\|^2 = \|\bar{\mathbf{x}}\|^2 + \|\mathbf{x}_h\|^2$ , logo  $\|\mathbf{x}\|^2 > \|\bar{\mathbf{x}}\|^2$  para qualquer  $\mathbf{x}_h$  não nulo.

Claramente  $\bar{\mathbf{x}}$  é a solução de norma mínima pois qualquer outra solução tem uma componente em  $\mathcal{N}(\mathbf{A})$  e isto iria aumentar sua norma. Portanto, a solução de norma mínima é:

$$\bar{\mathbf{x}} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{b}$$

C.Q.D.

**Exemplo 4.10** Para o sistema:

$$\mathbf{A}\mathbf{x} = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} = 10 = \mathbf{b}$$

Podemos ter um número infinito de soluções sobre o espaço linha de  $\mathbf{A}$ , conforme mostrado na Figura 4.22; se existe uma matriz  $\mathbf{A}^\dagger$  de tal forma que  $\mathbf{A}\mathbf{A}^\dagger = \mathbf{I}$  podemos encontrar uma solução ótima para o sistema fazendo:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \mathbf{A}^\dagger = \mathbf{I} \quad e \quad \mathbf{A}^\dagger = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \end{bmatrix}$$

donde  $\mathbf{a}_1 + 2\mathbf{a}_2 = 1$ . Logo  $\mathbf{A}^\dagger$  para o exemplo dado não é única. Então busquemos a solução de norma mínima e a pseudo-inversa á direita,  $\mathbf{A}_d^\dagger$ , correspondente. Para esta solução tem-se:

$$\mathbf{A}_d^\dagger = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$$

Como

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = 5 \quad \text{então} \quad (\mathbf{A}\mathbf{A}^T)^{-1} = \frac{1}{5}$$

e

$$\mathbf{A}_d^\dagger = \frac{1}{5} \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix}$$



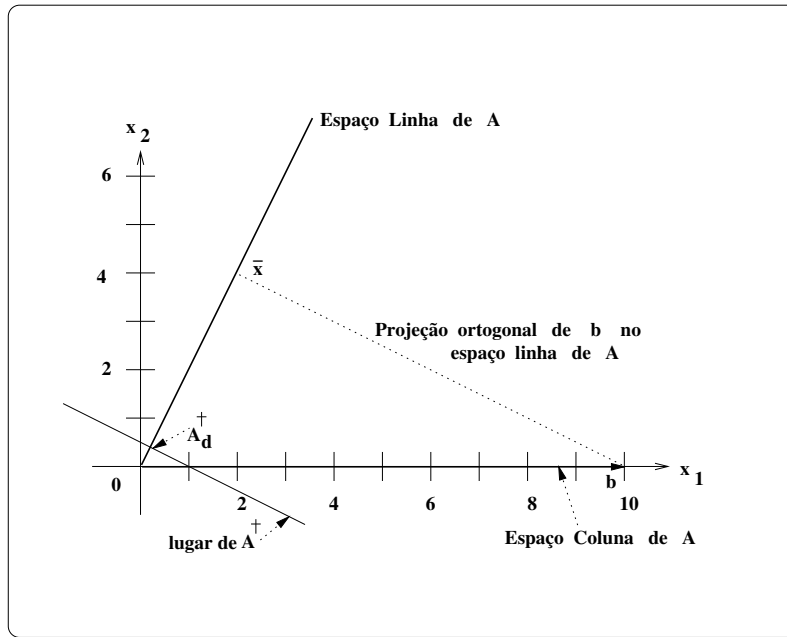


Figura 4.22: Representação geométrica: pseudo-inversa à direita.

Portanto o melhor valor de  $\mathbf{x}$  é:

$$\bar{\mathbf{x}} = \begin{bmatrix} 0.2 \\ 0.4 \end{bmatrix} 10 = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$$

então a melhor solução corresponde:

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 2 \\ 4 \end{bmatrix} = 10$$

**Exemplo 4.11** Para  $\mathbf{Ax} = \mathbf{b}$  com  $\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}$  e  $\mathbf{b} = 2$ , conforme ilustrado na Figura 4.23 temos:

$$\bar{\mathbf{x}} = \mathbf{A}^T(\mathbf{AA}^T)^{-1}\mathbf{b} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \left( \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \right)^{-1} 2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Como consequência dos resultados obtidos podemos tecer alguns comentários e enunciar os seguintes corolários:

**Corolário 4.1** Em termos dos subespaços fundamentais para  $\mathbf{Ax} = \mathbf{b}$ , o operador projeção ortogonal  $P$  decompõe  $\mathbf{b} \in \mathbb{R}^m$  de forma única em:

$$\mathbf{b} = P\mathbf{b} + (\mathbf{I}_m - P)\mathbf{b}$$

onde  $P\mathbf{b} \in \mathcal{R}(\mathbf{A})$  e  $(\mathbf{I}_m - P)\mathbf{b} \in \mathcal{N}(\mathbf{A}^T)$ .

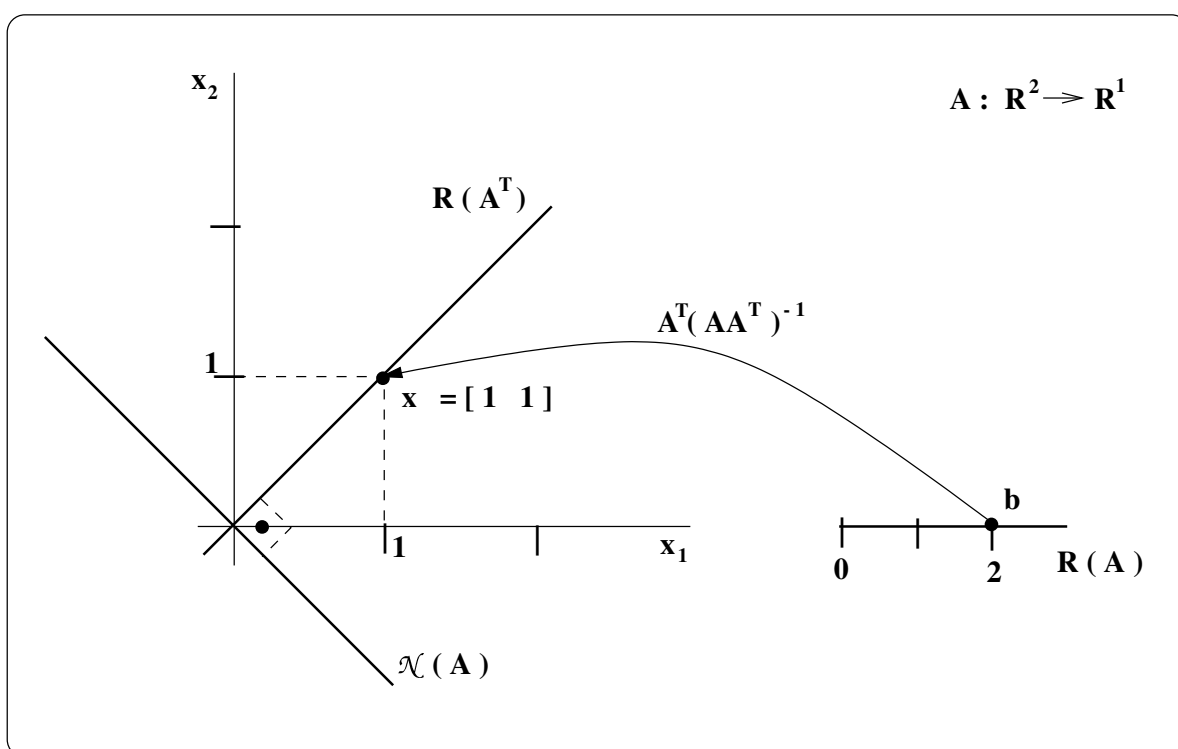


Figura 4.23: Representação geométrica: pseudo-inversa à direita.

**Corolário 4.2** *O operador projeção ortogonal  $P_l$  decompõe  $\mathbf{x} \in \mathbf{R}^n$  de forma única em:*

$$\mathbf{x} = P_l \mathbf{x} + (\mathbf{I}_n - P_l) \mathbf{x}$$

onde  $P_l \mathbf{x} \in \mathcal{R}(\mathbf{A}^T)$  e  $(\mathbf{I}_n - P_l) \mathbf{x} \in \mathcal{N}(\mathbf{A})$ .

**Corolário 4.3**  *$P = \mathbf{A} \mathbf{A}^\dagger$  é a projeção ortogonal de  $\mathbf{R}^m$  sobre  $\mathcal{R}(\mathbf{A})$  e  $P_l = \mathbf{A}^\dagger \mathbf{A}$  é a projeção ortogonal de  $\mathbf{R}^n$  sobre  $\mathcal{R}(\mathbf{A}^T)$ .*

**Corolário 4.4** *Se  $\mathbf{b} \in \mathcal{R}(\mathbf{A})$ ,  $\mathbf{A} \mathbf{x} = \mathbf{b}$  tem pelo menos uma solução se e somente se  $(\mathbf{I}_m - \mathbf{A} \mathbf{A}^\dagger) \mathbf{b} = 0$ . Quando isto ocorre, a solução geral é dada por:*

$$\mathbf{x} = \mathbf{A}^\dagger \mathbf{b} + (\mathbf{I}_n - \mathbf{A}^\dagger \mathbf{A}) \mathbf{z}$$

onde  $\mathbf{z}$  é um vetor arbitrário de dimensão  $n$ . Observe que  $(\mathbf{I}_n - \mathbf{A}^\dagger \mathbf{A})$  é a projeção ortogonal de  $\mathbf{R}^n$  sobre  $\mathcal{N}(\mathbf{A})$ ,  $P_{\mathcal{N}(\mathbf{A})}$ , e portanto  $P_{\mathcal{N}(\mathbf{A})} \mathbf{z}$  é um vetor arbitrário de  $\mathcal{N}(\mathbf{A})$ .

- Se  $\mathbf{z} = 0$  ou se  $\mathbf{z} \in \mathcal{R}(\mathbf{A}^T)$  tem-se solução de norma mínima  $\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}$ .

**Corolário 4.5** *Se  $\mathbf{b} \notin \mathcal{R}(\mathbf{A})$ ,  $\mathbf{A} \mathbf{x} = \mathbf{b}$  não tem solução. O vetor  $\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}$  de norma mínima minimiza o erro  $\mathbf{e} = \mathbf{A} \mathbf{x} - \mathbf{b}$  no sentido de mínimos quadrados conforme exposto neste capítulo.*

- Portanto neste caso não exigiremos que  $\mathbf{A} \mathbf{x}$  iguale  $\mathbf{b}$ , bem como dentre todas as soluções aproximadas escolheremos a de norma mínima.

Também podemos dizer o seguinte:

Dado  $\mathbf{A}$ ,  $\forall \mathbf{b} \in \mathbf{R}^m$  determine o vetor  $\bar{\mathbf{x}}$  de norma mínima que minimiza  $\|\mathbf{e} = \mathbf{A} \mathbf{x} - \mathbf{b}\|$ .  $\bar{\mathbf{x}}$  pode ser descrito da seguinte maneira:

Projete  $\mathbf{b}$  ortogonalmente sobre  $\mathcal{R}(\mathbf{A})$ , obtendo  $\bar{\mathbf{b}} \in \mathcal{R}(\mathbf{A})$ . Existe um e apenas um vetor em  $\mathcal{R}(\mathbf{A}^T)$ ,  $\bar{\mathbf{x}}$  tal que  $\mathbf{A} \bar{\mathbf{x}} = \bar{\mathbf{b}}$  e tal que:

$$\|\mathbf{A} \bar{\mathbf{x}} - \mathbf{b}\| = \min_{\mathbf{x} \in \mathbf{R}^n} \|\mathbf{A} \mathbf{x} - \mathbf{b}\|,$$

e

$$\bar{\mathbf{x}} = \mathbf{A}^\dagger \mathbf{b}.$$

- Os projetores ortogonais nos quatro subespaços fundamentais de  $\mathbf{A}$  são:

$$P \equiv P_{\mathcal{R}(\mathbf{A})} = \mathbf{A} \mathbf{A}^\dagger$$

$$P_{\mathcal{N}(\mathbf{A})} = \mathbf{I} - \mathbf{A}^\dagger \mathbf{A}$$

$$P_l \equiv P_{\mathcal{R}(\mathbf{A}^T)} = \mathbf{A}^\dagger \mathbf{A}$$

$$P_{\mathcal{N}(\mathbf{A}^T)} = \mathbf{I} - \mathbf{A} \mathbf{A}^\dagger$$

- A definição mais geral e formal de matriz pseudo-inversa: Pseudo-Inversa de Moore- Penrose é a que resulta do seguinte teorema enunciado e provado por Penrose [863]:

**Teorema 4.4** Para qualquer matriz  $\mathbf{A}_{m \times n}$ ,  $\mathbf{A}_{n \times m}^\dagger$  é a única solução do sistema de equações:

a)

$$\mathbf{A}^\dagger \mathbf{A} \mathbf{A}^\dagger = \mathbf{A}^\dagger$$

b)

$$\mathbf{A} \mathbf{A}^\dagger \mathbf{A} = \mathbf{A}$$

c)

$$(\mathbf{A}^\dagger \mathbf{A})^T = \mathbf{A}^\dagger \mathbf{A}$$

d)

$$(\mathbf{A} \mathbf{A}^\dagger)^T = \mathbf{A} \mathbf{A}^\dagger$$

- Para  $\mathbf{A}$  de posto completo:

a) Se  $m > n = r$ :  $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^{-1} \mathbf{A}^T$

b) Se  $r = m < n$ :  $\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^{-1}$

- Para  $\mathbf{A}$  de posto incompleto:

a) Se  $m > n > r$ :  $\mathbf{A}^\dagger = (\mathbf{A}^T \mathbf{A})^\dagger \mathbf{A}^T$

b) Se  $r < m < n$ :  $\mathbf{A}^\dagger = \mathbf{A}^T (\mathbf{A} \mathbf{A}^T)^\dagger$

- $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$

•

$(\mathbf{AB})^\dagger$  não necessariamente é igual a  $\mathbf{B}^\dagger \mathbf{A}^\dagger$

•

$\mathbf{AA}^\dagger$  não necessariamente é igual a  $\mathbf{A}^\dagger \mathbf{A}$

•

$(\mathbf{A}^k)^\dagger$  não necessariamente é igual a  $(\mathbf{A}^\dagger)^k$

•

$\text{posto} \mathbf{A} = \text{posto}(\mathbf{A}^\dagger) = \text{posto}(\mathbf{AA}^\dagger) = \text{posto}(\mathbf{A}^\dagger \mathbf{A})$

•

$(\mathbf{AA}^T)^\dagger \mathbf{AA}^T = \mathbf{AA}^T$

•

$(\mathbf{A}^T \mathbf{A})^\dagger \mathbf{A}^T \mathbf{A} = \mathbf{A}^T \mathbf{A}$

•

$\Re(\mathbf{A}^\dagger) = \Re(\mathbf{A}^T)$ .

## 4.9 Princípios para Decomposição em Valores Singulares

A partir dos conceitos de autovalor e autovetor pode-se obter a pseudo-inversa de uma matriz.

Seja uma matriz quadrada  $\mathbf{A}_{n \times n}$  e de forma intuitiva vejamos os conceitos de autovalor e autovetor, que supomos conhecidos:

- **Autovetor:** Quase todos os vetores quando multiplicados por uma matriz  $\mathbf{A}$  mudam de direção. Certamente um número excepcional de vetores  $\mathbf{x}$ , não nulos, tem a mesma direção que  $\mathbf{Ax}$ . Estes vetores são os autovetores de  $\mathbf{A}$ .
- **Autovalor:** A multiplicação de um autovetor  $\mathbf{x}$  pela matriz  $\mathbf{A}$  gera um vetor que será  $\lambda$  vezes o vetor original  $\mathbf{x}$ , podendo ser menor, igual ou maior que este. Daí podermos escrever que  $\mathbf{Ax} = \lambda \mathbf{x}$ , onde o escalar  $\lambda$  é um autovalor da matriz  $\mathbf{A}$ . Um autovalor poderá ser zero. O autovalor pode ser interpretado fisicamente como um ganho.

Para calcular autovetor e autovalor deve-se resolver o sistema:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

então,

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$$

Os autovetores  $\mathbf{x}$  formam o espaço nulo de  $\mathbf{A} - \lambda\mathbf{I}$ . Conhecendo um autovalor  $\lambda$ , podemos encontrar um autovetor  $\mathbf{x}$  para solucionar  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$ .

Para que  $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = 0$  tenha solução não nula,  $(\mathbf{A} - \lambda\mathbf{I})$  deverá ser singular.

Então o número  $\lambda$  é um autovalor de  $\mathbf{A}$  se e somente se

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$$

Esta é a equação característica que envolve somente  $\lambda$  e não  $\mathbf{x}$ . Cada autovalor,  $\lambda$ , da matriz  $\mathbf{A}$  conduzirá a pelo menos um autovetor,  $\mathbf{x}$ .

Os autovalores da matriz  $\mathbf{A}$  tem as seguintes propriedades:

- O produto dos  $n$  autovalores é igual ao determinante de  $\mathbf{A}$ .

$$\det(\mathbf{A}) = \prod_{i=1}^n \lambda_i$$

- A soma dos  $n$  autovalores é o traço da matriz  $\mathbf{A}$ , que é igual a soma dos elementos da diagonal da matriz  $\mathbf{A}$ :

$$\text{tr}(\mathbf{A}) = \sum_{i=1}^n \lambda_i = \sum_{i=1}^n a_{ii}$$

#### 4.9.1 Diagonalização de Uma Matriz Quadrada

Como é bem conhecido, a diagonalização de uma matriz quadrada  $\mathbf{A}_{n \times n}$  é facilmente obtida através de uma transformação de similaridade conforme mostra o seguinte teorema:

**Teorema 4.5** *Se uma matriz  $\mathbf{A}_{n \times n}$  tem autovetores linearmente independentes e  $\mathbf{T}$  é uma matriz tendo como colunas os autovetores de  $\mathbf{A}$ , então:*

a) *A matriz  $\mathbf{A}$  é diagonalizável e:*

$$\Lambda = \mathbf{T}^{-1}\mathbf{A}\mathbf{T} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \quad (4.39)$$

b) Se  $\mathbf{x}$  é um autovetor de  $\mathbf{A}$ ,  $\mathbf{T}^{-1}\mathbf{x}$  é um autovetor de  $\Lambda$ .

**Prova:**

a) Como os autovetores de  $\mathbf{A}$  são linearmente independentes, por hipótese, pode-se escrever:

$$\mathbf{AT} = \mathbf{A} \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} = \begin{bmatrix} \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \cdots & \lambda_n \mathbf{x}_n \end{bmatrix}$$

então

$$\begin{aligned} \mathbf{AT} &= \begin{bmatrix} \lambda_1 \mathbf{x}_1 & \lambda_2 \mathbf{x}_2 & \cdots & \lambda_n \mathbf{x}_n \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{x}_1 & \mathbf{x}_2 & \cdots & \mathbf{x}_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \mathbf{T}\Lambda \end{aligned}$$

Como  $\mathbf{T}$  é invertível, multiplicando à esquerda ambos os membros da equação  $\mathbf{AT} = \mathbf{T}\Lambda$  tem-se:

$$\Lambda = \mathbf{T}^{-1}\mathbf{AT} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

que constitui a forma diagonal de  $\mathbf{A}$ .

b) De a),  $\mathbf{T}^{-1}\mathbf{AT} = \Lambda$  donde  $\mathbf{A} = \mathbf{T}\Lambda\mathbf{T}^{-1}$ . Como  $\mathbf{Ax} = \lambda\mathbf{x}$  temos:

$$\mathbf{T}\Lambda\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{x}$$

então

$$\Lambda\mathbf{T}^{-1}\mathbf{x} = \lambda\mathbf{T}^{-1}\mathbf{x}$$

Portanto  $\lambda$  e  $\mathbf{T}^{-1}\mathbf{x}$  são respectivamente autovalor e autovetor de  $\Lambda$ .

**C.Q.D.**

Observe que:

$$\Lambda^2 = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}^2 = \begin{bmatrix} \lambda_1^2 & & \\ & \ddots & \\ & & \lambda_n^2 \end{bmatrix} \quad (4.40)$$

e

$$\Lambda^{-1} = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}^{-1} = \begin{bmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_n^{-1} \end{bmatrix} \quad (4.41)$$

### 4.9.2 Diagonalização de Uma Matriz Qualquer

A decomposição em valores singulares ( *Singular Value Decomposition* ou *SVD* ) de uma matriz  $\mathbf{A}_{m \times n}$  tem grande importância teórica e prática, por exemplo para a solução de problemas de mínimos quadrados. Ela dá uma forma diagonal à matriz  $\mathbf{A}$  equivalente a uma transformação ortogonal.

Para a fatoração da matriz quadrada  $\mathbf{A} = \mathbf{T}\mathbf{\Lambda}\mathbf{T}^{-1}$  escolhe-se a base dos autovetores. Quando a matriz  $\mathbf{A}$  é simétrica seus autovetores são ortonormais. Seja  $\mathbf{T} = \mathbf{Q}$  a matriz ortonormal com estes autovetores como colunas, então a matriz diagonal  $\mathbf{\Lambda} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \mathbf{Q}^T\mathbf{A}\mathbf{Q}$  e a fatoração de uma matriz simétrica vem a ser  $\mathbf{A} = \mathbf{Q}\mathbf{\Lambda}\mathbf{Q}^T$ .

Os quatro subespaços fundamentais existem para todo operador, donde para toda matriz que o representa, no entanto se suas bases não são ortonormais tal matriz não é necessariamente diagonal. Logo para diagonalizar uma matriz  $\mathbf{A}_{m \times n}$ , de posto  $r$ , precisamos determinar bases ortonormais  $\mathbf{v}_1 \dots \mathbf{v}_r$ , para o espaço linha e  $\mathbf{u}_1 \dots \mathbf{u}_r$  para o espaço coluna.

Sejam as matrizes  $\mathbf{V}$  e  $\mathbf{U}$  associadas às bases ortonormais dos quatro subespaços fundamentais:

- primeiras  $r$  colunas de  $\mathbf{V}$ : espaço linha de  $\mathbf{A}$
- últimas  $n - r$  colunas de  $\mathbf{V}$ : espaço nulo de  $\mathbf{A}$
- primeiras  $r$  colunas de  $\mathbf{U}$ : espaço coluna de  $\mathbf{A}$
- últimas  $m - r$  colunas de  $\mathbf{U}$ : espaço nulo de  $\mathbf{A}^T$

As matrizes ortogonais  $\mathbf{V}$  e  $\mathbf{U}$  têm respectivamente dimensões  $n \times n$  e  $m \times m$  diferentemente da matriz  $\mathbf{T}$  associada à diagonalização da matriz  $\mathbf{A}$  quadrada. Agora a base do espaço linha difere da base do espaço coluna. A matriz  $\mathbf{V}$  tem como primeiras colunas  $v_1 \dots v_r$  que constituem a base do espaço linha de  $\mathbf{A}$  e é completada com os vetores ortogonais  $v_{r+1} \dots v_n$  do espaço nulo de  $\mathbf{A}$ . Analogamente, a matriz  $\mathbf{U}$  tem como primeiras colunas  $u_1 \dots u_r$  que constituem a base do espaço coluna de  $\mathbf{A}$  e é completada com os vetores ortogonais  $u_{r+1} \dots u_m$  que constituem a base do espaço nulo de  $\mathbf{A}^T$ .

**Teorema 4.6** *Uma matriz  $\mathbf{A}_{m \times n}$ , de posto  $r$ , pode ser fatorada na forma  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  onde  $\mathbf{\Sigma}$  é a matriz diagonal:*

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & & & \\ & \ddots & & & \\ & & \sigma_r & & \\ & & & \ddots & \\ & 0 & & & 0 \end{bmatrix}_{m \times n} = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}_{m \times n} \quad e \quad \Sigma_1 = \begin{bmatrix} \sigma_1 & & \\ 0 & \ddots & \\ & & \sigma_r \end{bmatrix}_{r \times r}$$



com a diagonal principal tendo os valores singulares  $\sigma_i$  da matriz  $\mathbf{A}$ , que são números positivos com ordenação obedecendo o critério:  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq \sigma_{r+1} = \dots = \sigma_m = 0$ , e:

$$\mathbf{U} = \begin{bmatrix} u_1 & \dots & u_r & \dots & u_m \\ & & & & \end{bmatrix}_{m \times m}$$

e

$$\mathbf{V} = \begin{bmatrix} v_1 & \dots & v_r & \dots & v_n \\ & & & & \end{bmatrix}_{n \times n}^T$$

são matrizes ortonormais com  $\mathbf{v}_1 \dots \mathbf{v}_r$ , e  $\mathbf{u}_1 \dots \mathbf{u}_r$  formando respectivamente as bases dos espaços linha e coluna da matriz  $\mathbf{A}$  tal que:

$$\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i, \quad i = 1, 2, \dots, r$$

**Prova:**

Sejam  $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$  os autovalores não nulos de  $\mathbf{A}^T\mathbf{A}$  ou de  $\mathbf{A}\mathbf{A}^T$  simétricas:

a) então existe uma matriz ortogonal  $\mathbf{U}$  de ordem  $m$  tal que:

$$\mathbf{A}\mathbf{A}^T = \mathbf{U} \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & \\ & & & \ddots & \\ 0 & & & & 0 \end{bmatrix}_{m \times m} \mathbf{U}^T \quad (4.42)$$

Observe que as primeiras  $r$  colunas da matriz  $\mathbf{U}$  serão multiplicadas por  $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$ , e as restantes por autovalores nulos; então podemos fazer:

$$\mathbf{U} = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \dots \quad \mathbf{u}_m]_{m \times m} = [\mathbf{U}_1 \quad | \quad \mathbf{U}_2]_{m \times m} \quad (4.43)$$

onde  $\mathbf{U}_1$  tem dimensão  $m \times r$  e  $\mathbf{U}_2$  tem dimensão  $m \times (m - r)$  e tem-se:

$$\mathbf{A}\mathbf{A}^T\mathbf{U}_1 = \mathbf{U}_1\Sigma_1^2 \quad (4.44)$$

onde,

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & & & \\ & \ddots & & \\ & & \sigma_i & \\ & 0 & & \ddots \\ & & & & \sigma_r \end{bmatrix} \quad (4.45)$$

A partir da equação (4.44), premultiplicando por  $\mathbf{A}^T$  e posmultiplicando por  $\Sigma_1^{-1}$ :

$$\mathbf{A}^T \mathbf{A} \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} = \mathbf{A}^T \mathbf{U}_1 \Sigma_1^2 \Sigma_1^{-1}$$

Fazendo

$$\mathbf{V}_1 = \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} \quad (4.46)$$

temos

$$\mathbf{A}^T \mathbf{A} \mathbf{V}_1 = \mathbf{V}_1 \Sigma_1^2 \quad (4.47)$$

Observe que a matriz  $\mathbf{V}_1$  tem dimensão  $n \times r$  e satisfaz:

$$\mathbf{V}_1^T \mathbf{V}_1 = (\mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1})^T (\mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1}) = \Sigma_1^{-1} \mathbf{U}_1^T \mathbf{A} \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} = I_{r \times r} \quad (4.48)$$

devido a (4.44).

Como a matriz  $\mathbf{A}$  possui  $n$  colunas e somente  $r$  são linearmente independentes, podemos sempre encontrar uma matriz  $\mathbf{V}_2$  ortogonal de dimensão  $n \times (n - r)$  tal que:

$$\mathbf{A} \mathbf{V}_2 = 0, \quad \mathbf{V}_2^T \mathbf{V}_2 = \mathbf{I}_{(n-r) \times (n-r)} \quad (4.49)$$

Fazendo  $\mathbf{V} = [\mathbf{V}_1 \mid \mathbf{V}_2]$  e calculando  $\mathbf{V}^T \mathbf{V}$  obtemos:

$$\mathbf{V}^T \mathbf{V} = \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{V}_1^T \mathbf{V}_1 & \mathbf{V}_1^T \mathbf{V}_2 \\ \mathbf{V}_2^T \mathbf{V}_1 & \mathbf{V}_2^T \mathbf{V}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{r \times r} & 0 \\ 0 & \mathbf{I}_{(n-r) \times (n-r)} \end{bmatrix} = \mathbf{I}$$

pois

$$\mathbf{V}_1^T \mathbf{V}_2 = \Sigma_1^{-1} \mathbf{U}_1^T \mathbf{A} \mathbf{V}_2 = 0 \quad \text{e} \quad \mathbf{V}_2^T \mathbf{V}_1 = \mathbf{V}_2^T \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} = 0 \quad (4.50)$$

Portanto:

$$\mathbf{V} = \begin{bmatrix} \mathbf{V}_1 & | & \mathbf{V}_2 \end{bmatrix} \quad (4.51)$$

é uma matriz ortogonal de dimensão  $n \times n$ ; donde  $\mathbf{V}^T = \mathbf{V}^{-1}$ .

A partir da equação (4.44), pela posmultiplicação por  $\Sigma_1^{-1}$  e usando a equação (4.46), nós temos:

$$\mathbf{A}\mathbf{A}^T\mathbf{U}_1\Sigma_1^{-1} = \mathbf{U}_1\Sigma_1^2\Sigma_1^{-1} \Rightarrow \mathbf{A}\mathbf{A}^T\mathbf{U}_1\Sigma_1^{-1} = \mathbf{U}_1\Sigma_1$$

$$\mathbf{A}\mathbf{V}_1 = \mathbf{U}_1\Sigma_1 \quad (4.52)$$

Como a matriz  $\mathbf{V}$  é ortogonal:

$$\mathbf{V}\mathbf{V}^T = \mathbf{V}\mathbf{V}^{-1} = \mathbf{I} \quad (4.53)$$

e

$$\mathbf{V}\mathbf{V}^T = \begin{bmatrix} \mathbf{V}_1 & | & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \frac{\mathbf{V}_1^T}{\mathbf{V}_2^T} \end{bmatrix} = \mathbf{V}_1\mathbf{V}_1^T + \mathbf{V}_2\mathbf{V}_2^T = \mathbf{I} \quad (4.54)$$

então  $\mathbf{V}_1\mathbf{V}_1^T = \mathbf{I} - \mathbf{V}_2\mathbf{V}_2^T$ .

Posmultiplicando (4.52) por  $\mathbf{V}_1^T$  obtemos a forma reduzida:

$$\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T \quad (4.55)$$

que pode ser escrita como:

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^T \quad (4.56)$$

onde:

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}_{m \times n} \quad (4.57)$$

Evidentemente, as bases ortogonais associadas a  $\mathbf{U}$  e  $\mathbf{V}$  podem ser normalizadas, donde  $\mathbf{U}$  e  $\mathbf{V}$  feitas ortonormais.

Da equação (4.52) tem-se:

$$\mathbf{A}\mathbf{v}_i = \sigma_i\mathbf{u}_i, \quad i = 1, \dots, r \quad (4.58)$$

- b) Analogamente para  $\mathbf{A}^T \mathbf{A}$  existe uma matriz ortogonal  $\mathbf{V}$  de ordem  $n$  tal que:

$$\mathbf{A}^T \mathbf{A} = \mathbf{V} \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & 0 \\ & & \sigma_r^2 & \\ & 0 & & \ddots \\ & & & & 0 \end{bmatrix}_{n \times n} \mathbf{V}^T \quad (4.59)$$

Observe que as primeiras  $r$  colunas da matriz  $\mathbf{V}$  serão multiplicadas por  $\sigma_1^2, \sigma_2^2, \dots, \sigma_r^2$ , e as restantes por autovalores nulos, então podemos fazer:

$$\mathbf{V} = [\mathbf{v}_1 \ \mathbf{v}_2 \ \dots \ \mathbf{v}_n]_{n \times n} = [\mathbf{V}_1 \ | \ \mathbf{V}_2]_{n \times n} \quad (4.60)$$

onde  $\mathbf{V}_1$  tem dimensão  $n \times r$  e  $\mathbf{V}_2$  tem dimensão  $n \times (n - r)$ , tem-se:

$$\mathbf{A}^T \mathbf{A} = [\mathbf{V}_1 \ | \ \mathbf{V}_2] \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix}$$

$$\mathbf{A}^T \mathbf{A} = \mathbf{V}_1 \Sigma_1^2 \mathbf{V}_1^T$$

posmultiplicando ambos os termos por  $\mathbf{V}_1$ :

$$\mathbf{A}^T \mathbf{A} \mathbf{V}_1 = \mathbf{V}_1 \Sigma_1^2 \mathbf{V}_1^T \mathbf{V}_1$$

como  $\mathbf{V}_1^T \mathbf{V}_1 = \mathbf{I}_{r \times r}$

$$\mathbf{A}^T \mathbf{A} \mathbf{V}_1 = \mathbf{V}_1 \Sigma_1^2 \quad (4.61)$$

A partir da equação (4.61), premultiplicando por  $\mathbf{A}$  e posmultiplicando por  $\Sigma_1^{-1}$ :

$$\mathbf{A} \mathbf{A}^T \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} = \mathbf{A} \mathbf{V}_1 \Sigma_1^2 \Sigma_1^{-1}$$

Fazendo

$$\mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} \quad (4.62)$$

temos

$$\mathbf{A} \mathbf{A}^T \mathbf{U}_1 = \mathbf{U}_1 \Sigma_1^2 \quad (4.63)$$

Observe que a matriz  $\mathbf{U}_1$  tem dimensão  $m \times r$  e satisfaz:

$$\mathbf{U}_1^T \mathbf{U}_1 = (\mathbf{A} \mathbf{V}_1 \Sigma_1^{-1})^T (\mathbf{A} \mathbf{V}_1 \Sigma_1^{-1}) = \Sigma_1^{-1} \mathbf{V}_1^T \underbrace{\mathbf{A}^T \mathbf{A} \mathbf{V}_1}_{\mathbf{V}_1 \Sigma_1^2} \Sigma_1^{-1} = I_{r \times r} \quad (4.64)$$

devido a (4.61).

Como a matriz  $\mathbf{A}$  possui  $n$  colunas e somente  $r$  são linearmente independentes, nós podemos sempre encontrar uma matriz  $\mathbf{U}_2$  ortogonal de dimensão  $m \times (m - r)$  tal que:

$$\mathbf{A}^T \mathbf{U}_2 = 0, \quad \mathbf{U}_2^T \mathbf{U}_2 = \mathbf{I}_{(m-r) \times (m-r)} \quad (4.65)$$

fazendo  $\mathbf{U} = [\mathbf{U}_1 \mid \mathbf{U}_2]$  e calculando  $\mathbf{U}^T \mathbf{U}$  obtemos:

$$\mathbf{U}^T \mathbf{U} = \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{bmatrix} \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{U}_1^T \mathbf{U}_1 & \mathbf{U}_1^T \mathbf{U}_2 \\ \mathbf{U}_2^T \mathbf{U}_1 & \mathbf{U}_2^T \mathbf{U}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{I}_{r \times r} & 0 \\ 0 & \mathbf{I}_{(m-r) \times (m-r)} \end{bmatrix} = \mathbf{I}_{m \times m}$$

pois

$$\mathbf{U}_1^T \mathbf{U}_2 = \Sigma_1^{-1} \mathbf{V}_1^T \mathbf{A}^T \mathbf{U}_2 = 0 \quad \text{e} \quad \mathbf{U}_2^T \mathbf{U}_1 = \mathbf{U}_2^T \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} = 0 \quad (4.66)$$

Portanto:

$$\mathbf{U} = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \quad (4.67)$$

é uma matriz ortogonal de dimensão  $m \times m$ ; donde  $\mathbf{U}^T = \mathbf{U}^{-1}$ .

A partir da equação (4.61), pela posmultiplicação por  $\Sigma_1^{-1}$  e usando a equação (4.62), nós temos:

$$\mathbf{A}^T \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} = \mathbf{V}_1 \Sigma_1^2 \Sigma_1^{-1} \Rightarrow \mathbf{A}^T \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} = \mathbf{V}_1 \Sigma_1$$

$$\mathbf{A}^T \mathbf{U}_1 = \mathbf{V}_1 \Sigma_1 \quad (4.68)$$

Como a matriz  $\mathbf{U}$  é ortogonal:

$$\mathbf{U} \mathbf{U}^T = \mathbf{U} \mathbf{U}^{-1} = \mathbf{I} \quad (4.69)$$

e

$$\mathbf{U}\mathbf{U}^T = [\mathbf{U}_1 \mid \mathbf{U}_2] \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{bmatrix} = \mathbf{U}_1\mathbf{U}_1^T + \mathbf{U}_2\mathbf{U}_2^T = \mathbf{I}_{m \times m}$$

então  $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{I} - \mathbf{U}_2\mathbf{U}_2^T$

Posmultiplicando (4.68) por  $\mathbf{V}_1^T$  reobtemos a equação (4.55).

**C.Q.D.**

Como consequência dos resultados obtidos podemos tecer alguns comentários:

- Como  $\mathbf{V}$  e  $\mathbf{U}$  são ortonormais, tem-se:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^{-1}$$

donde  $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$ .

•

$$\mathbf{A}^T\mathbf{A} = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T) = \mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \mathbf{V}\mathbf{\Sigma}^T\mathbf{\Sigma}\mathbf{V}^T$$

onde:

$$\mathbf{\Sigma}^T\mathbf{\Sigma} = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix}_{n \times n}$$

$\Sigma_1^2$  é uma matriz diagonal, com  $\sigma_1^2 \cdots \sigma_r^2$  que são os autovalores não nulos da matriz  $\mathbf{A}^T\mathbf{A}$  simétrica:

$$\mathbf{A}^T\mathbf{A} = \mathbf{V} \begin{bmatrix} \sigma_1^2 & & & \\ & \ddots & & 0 \\ & & \sigma_r^2 & \\ & 0 & & \ddots \\ & & & & 0 \end{bmatrix} \mathbf{V}^T \quad (4.70)$$

Nas colunas de  $\mathbf{V}$  estão os autovetores de  $\mathbf{A}^T\mathbf{A}$ .

•

$$\mathbf{A}\mathbf{A}^T = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)(\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)^T = (\mathbf{U}\mathbf{\Sigma}\mathbf{V}^T)(\mathbf{V}\mathbf{\Sigma}^T\mathbf{U}^T) = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T$$

onde

$$\mathbf{\Sigma}^T\mathbf{\Sigma} = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix}_{m \times m}$$

$$\mathbf{A}\mathbf{A}^T = \mathbf{U} \begin{bmatrix} \sigma_1^2 & & & & \\ & \ddots & & & \\ & & \sigma_r^2 & & 0 \\ & 0 & & \ddots & \\ & & & & 0 \end{bmatrix} \mathbf{U}^T \quad (4.71)$$

$\Sigma_1^2$  é uma matriz diagonal, com  $\sigma_1^2 \cdots \sigma_n^2$  sendo autovalores não nulos da matriz simétrica  $\mathbf{A}\mathbf{A}^T$ .

Nas colunas de  $\mathbf{U}$  estão os autovetores de  $\mathbf{A}\mathbf{A}^T$ .

- Os autovalores valores não nulos de  $\mathbf{A}\mathbf{A}^T$  e  $\mathbf{A}^T\mathbf{A}$  são  $\sigma_i^2$ ,  $i = 1, 2, \dots, r$ , pois estas matrizes são simétricas.
- $\mathbf{A}^T\mathbf{A}$  e  $\mathbf{A}\mathbf{A}^T$  são chamadas matrizes de Gram e podemos observar que:

a) Se

$$\mathbf{A} = \left[ \begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{array} \right]$$

e

$$\mathbf{A}^T = \left[ \begin{array}{c} \hline \mathbf{a}_1 \\ \hline \mathbf{a}_2 \\ \hline \vdots \\ \hline \mathbf{a}_n \end{array} \right]$$

então:

$$\begin{aligned} \mathbf{A}^T\mathbf{A} &= \left[ \begin{array}{c} \hline \mathbf{a}_1 \\ \hline \mathbf{a}_2 \\ \hline \vdots \\ \hline \mathbf{a}_n \end{array} \right] \left[ \begin{array}{c|c|c|c} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{array} \right] = \\ &= \left[ \begin{array}{cccc} \langle \mathbf{a}_1, \mathbf{a}_1 \rangle & \langle \mathbf{a}_1, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{a}_1, \mathbf{a}_n \rangle \\ \vdots & & \dots & \vdots \\ \langle \mathbf{a}_n, \mathbf{a}_1 \rangle & \langle \mathbf{a}_n, \mathbf{a}_2 \rangle & \dots & \langle \mathbf{a}_n, \mathbf{a}_n \rangle \end{array} \right] \end{aligned}$$

que é a matriz dos produtos escalares das colunas de  $\mathbf{A}$ .

b)

$$\mathbf{A} = \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix}$$

e

$$\mathbf{A}^T = \left[ \begin{array}{c|c|c|c} \mathbf{a}_1^T & \mathbf{a}_2^T & \dots & \mathbf{a}_m^T \end{array} \right]$$

então:

$$\begin{aligned} \mathbf{A}\mathbf{A}^T &= \begin{bmatrix} \mathbf{a}_1^T \\ \mathbf{a}_2^T \\ \vdots \\ \mathbf{a}_m^T \end{bmatrix} \left[ \begin{array}{c|c|c|c} \mathbf{a}_1^T & \mathbf{a}_2^T & \dots & \mathbf{a}_m^T \end{array} \right] = \\ &= \begin{bmatrix} \langle \mathbf{a}_1^T, \mathbf{a}_1^T \rangle & \langle \mathbf{a}_2^T, \mathbf{a}_1^T \rangle & \dots & \langle \mathbf{a}_m^T, \mathbf{a}_1^T \rangle \\ \vdots & \vdots & \dots & \vdots \\ \langle \mathbf{a}_1^T, \mathbf{a}_m^T \rangle & \langle \mathbf{a}_2^T, \mathbf{a}_m^T \rangle & \dots & \langle \mathbf{a}_m^T, \mathbf{a}_m^T \rangle \end{bmatrix} \end{aligned}$$

que é a matriz dos produtos escalares das linhas de  $\mathbf{A}$ .

- Vê-se da equação (4.58) que com 2 bases ortonormais diferentes pode-se diagonalizar  $\mathbf{A}_{m \times n}$  fazendo com que cada vetor de base  $\mathbf{v}_i$  se transforme em cada vetor de base  $\mathbf{u}_i$ :

$$\mathbf{u}_i = \frac{\mathbf{A}}{\sigma_i} \mathbf{v}_i$$

- Das equações (4.55) e (4.58) tem-se:

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (4.72)$$



- Comparando com o caso de uma matriz quadrada, a matriz diagonal  $\Sigma$  é semelhante a  $\Lambda$ , onde  $\Sigma$  contém os valores singulares  $[\sigma_1, \sigma_2, \dots]$  e a matriz  $\Lambda$  contém os autovalores  $[\lambda_1, \lambda_2, \dots]$ .

Se  $\mathbf{V} = \mathbf{U} = \mathbf{T}$ , onde  $\mathbf{T}$  é uma matriz com autovetores de uma matriz quadrada, temos:

$$\mathbf{AT} = \mathbf{T}\Lambda$$

donde  $\mathbf{T}^{-1}\mathbf{AT} = \Lambda$ .

A matriz é diagonalizada, porém os autovetores não são ortonormais em geral.

**Exemplo 4.12** Dada a matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix}$$

obtenhamos sua decomposição em valores singulares.

Duas maneiras de se obter a decomposição em valores singulares são:

a) Através de  $\mathbf{A}^T\mathbf{A}$  temos:

$$\mathbf{A}^T\mathbf{A} = (\mathbf{V}\Sigma\mathbf{U}^T)(\mathbf{U}\Sigma^T\mathbf{V}^T) = \mathbf{V}\Sigma\Sigma^T\mathbf{V}^T$$

desta forma obteremos  $\mathbf{V}$  e posteriormente  $\mathbf{U}$ , da seguinte maneira:

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix} \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

tem a equação característica dada por:  $\det(\mathbf{A}^T\mathbf{A} - \lambda\mathbf{I}) = 0$ , seus autovalores são:  $\lambda_1 = 9$  e  $\lambda_2 = 1$  que correspondem aos valores singulares:  $\sigma_1^2$  e  $\sigma_2^2$  da matriz  $\mathbf{A}$ . Os autovetores  $\mathbf{v}_i$  associados à matriz  $\mathbf{A}^T\mathbf{A}$  são obtidos pela solução do sistema  $(\mathbf{A}^T\mathbf{A} - \lambda_i\mathbf{I})\mathbf{x} = 0$ :

$$\lambda_1 = 9 \Rightarrow \mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \quad e \quad \lambda_2 = 1 \Rightarrow \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Observe que os vetores  $\mathbf{v}_1$  e  $\mathbf{v}_2$  são unitários ( $\sqrt{\mathbf{v}_1^2 + \mathbf{v}_2^2} = 1$ ).

Os autovetores unitários  $\mathbf{u}_1$  e  $\mathbf{u}_2$  são obtidos de  $\mathbf{AV} = \mathbf{U}\Sigma$ :

$$\mathbf{A}[\mathbf{v}_1 \quad \mathbf{v}_2] = [\sigma_1\mathbf{u}_1 \quad \sigma_2\mathbf{u}_2] = [\mathbf{u}_1 \quad \mathbf{u}_2] \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

donde,

$$\lambda_1 = 9 \implies \mathbf{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \quad e \quad \lambda_2 = 2 \implies \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Portanto:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$$

b) Através de  $\mathbf{A}\mathbf{A}^T$  temos:

$$\mathbf{A}\mathbf{A}^T = (\mathbf{U}\mathbf{\Sigma}^T\mathbf{V}^T)(\mathbf{V}\mathbf{\Sigma}\mathbf{U}^T) = \mathbf{U}\mathbf{\Sigma}\mathbf{\Sigma}^T\mathbf{U}^T$$

desta forma obteremos  $\mathbf{U}$  e posteriormente  $\mathbf{V}$ , da seguinte maneira:

Como

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} \begin{bmatrix} 1 & -2 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} 5 & -4 \\ -4 & 5 \end{bmatrix}$$

tem a equação característica dada por:  $\det(\mathbf{A}\mathbf{A}^T - \lambda\mathbf{I}) = 0$ , seus autovalores são:  $\lambda_1 = 9$  e  $\lambda_2 = 1$  que correspondem aos valores singulares:  $\sigma_1^2$  e  $\sigma_2^2$  da matriz  $\mathbf{A}$ . Os autovetores  $\mathbf{u}_i$  associados à matriz  $\mathbf{A}^T\mathbf{A}$  são obtidos pela solução do sistema  $(\mathbf{A}\mathbf{A}^T - \lambda_i\mathbf{I})\mathbf{x} = 0$ :

$$\lambda_1 = 9 \implies \mathbf{u}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad e \quad \lambda_2 = 1 \implies \mathbf{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Observe que os vetores  $\mathbf{u}_1$  e  $\mathbf{u}_2$  são unitários ( $\sqrt{\mathbf{u}_1^2 + \mathbf{u}_2^2} = 1$ ).

Os autovetores unitários  $\mathbf{v}_1$  e  $\mathbf{v}_2$  são obtidos de  $\mathbf{A}\mathbf{V} = \mathbf{U}\mathbf{\Sigma}$ :

$$\mathbf{A}[\mathbf{v}_1 \quad \mathbf{v}_2] = [\sigma_1\mathbf{u}_1 \quad \sigma_2\mathbf{u}_2] = [\mathbf{u}_1 \quad \mathbf{u}_2] \begin{bmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{bmatrix}$$

donde,

$$\lambda_1 = 9 \implies \mathbf{v}_1 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad e \quad \lambda_2 = 1 \implies \mathbf{v}_2 = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}$$

Portanto:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \begin{bmatrix} 1 & 2 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}^T$$

**Teorema 4.7**

$$a) \quad \mathbf{U}_1 \mathbf{U}_1^T = \mathbf{A} \mathbf{A}^T \quad (4.73)$$

$$b) \quad \mathbf{V}_1 \mathbf{V}_1^T = \mathbf{A}^T \mathbf{A} \quad (4.74)$$

**Prova:**

a) Da equação (4.62) tem-se:

$$\mathbf{U}_1 = \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1}$$

fazendo:

$$\mathbf{U}_1 \mathbf{U}_1^T = \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} (\mathbf{A} \mathbf{V}_1 \Sigma_1^{-1})^T = \mathbf{A} \mathbf{V}_1 \Sigma_1^{-1} \Sigma_1^{-1} \mathbf{V}_1^T \mathbf{A}^T = \mathbf{A} \mathbf{A}^T$$

b) Da equação (4.46) tem-se:

$$\mathbf{V}_1 = \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1}$$

fazendo:

$$\mathbf{V}_1 \mathbf{V}_1^T = \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} (\mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1})^T = \mathbf{A}^T \mathbf{U}_1 \Sigma_1^{-1} \Sigma_1^{-1} \mathbf{U}_1^T \mathbf{A} = \mathbf{A}^T \mathbf{A}$$

**C.Q.D.**

**Teorema 4.8**

a) O projetor ortogonal sobre o espaço coluna da matriz  $\mathbf{A}$  é:

$$P_{\mathcal{R}(\mathbf{A})} = P = (\mathbf{A} \mathbf{A}^\dagger)^T = \mathbf{A} \mathbf{A}^\dagger = \mathbf{U}_1 \mathbf{U}_1^T \quad (4.75)$$

b) O projetor ortogonal sobre o espaço linha da matriz  $\mathbf{A}$  é:

$$P_{\mathcal{R}(\mathbf{A}^T)} = P_l = (\mathbf{A}^\dagger \mathbf{A})^T = \mathbf{A}^\dagger \mathbf{A} = \mathbf{V}_1 \mathbf{V}_1^T \quad (4.76)$$

c) O projetor ortogonal sobre o espaço nulo da matriz  $\mathbf{A}$  é:

$$P_{\mathcal{N}(\mathbf{A})} = \mathbf{I} - \mathbf{A}^\dagger \mathbf{A} = \mathbf{V}_2 \mathbf{V}_2^T \quad (4.77)$$

d) O projetor ortogonal sobre o espaço nulo da matriz  $\mathbf{A}^T$  é:

$$P_{\mathcal{N}(\mathbf{A}^T)} = \mathbf{I} - \mathbf{A}\mathbf{A}^\dagger = \mathbf{U}_2\mathbf{U}_2^T \quad (4.78)$$

**Prova:**

a) Das equações (4.18) e (4.55)

$$P = \mathbf{A}\mathbf{A}^\dagger = \mathbf{A}(\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$$

e

$$\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$$

e do teorema 4.7 - parte b)  $\mathbf{V}_1\mathbf{V}_1^T = \mathbf{A}^T\mathbf{A}$  tem-se:

$$P = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T(\mathbf{V}_1\mathbf{V}_1^T)^{-1}(\mathbf{U}_1\Sigma_1\mathbf{V}_1^T)^T = \mathbf{U}_1\mathbf{U}_1^T$$

b) Das equações (4.38) e (4.55)

$$P_l = \mathbf{A}_d^\dagger\mathbf{A} = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}\mathbf{A}$$

e

$$\mathbf{A} = \mathbf{U}_1\Sigma_1\mathbf{V}_1^T$$

e do teorema 4.7 - parte a)  $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{A}\mathbf{A}^T$  tem-se:

$$P_l = (\mathbf{U}_1\Sigma_1\mathbf{V}_1^T)^T(\mathbf{U}_1\mathbf{U}_1^T)^T\mathbf{U}_1\Sigma_1\mathbf{V}_1^T = \mathbf{V}_1\mathbf{V}_1^T$$

c) Como  $\mathbf{V}\mathbf{V}^T = \mathbf{I}$  teremos:

$$\mathbf{V}\mathbf{V}^T = \begin{bmatrix} \mathbf{V}_1 & | & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} = \mathbf{V}_1\mathbf{V}_1^T + \mathbf{V}_2\mathbf{V}_2^T = \mathbf{I}$$

então  $\mathbf{V}_2\mathbf{V}_2^T = \mathbf{I} - \mathbf{V}_1\mathbf{V}_1^T$ .

Da equação (4.76) sabemos que  $\mathbf{V}_1\mathbf{V}_1^T = \mathbf{A}^\dagger\mathbf{A}$ .

Logo por substituição obtemos  $\mathbf{V}_2\mathbf{V}_2^T = \mathbf{I} - \mathbf{A}^\dagger\mathbf{A}$ .

d) Como  $\mathbf{U}\mathbf{U}^T = \mathbf{I}$  teremos:

$$\mathbf{U}\mathbf{U}^T = \begin{bmatrix} \mathbf{U}_1 & | & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{bmatrix} = \mathbf{U}_1\mathbf{U}_1^T + \mathbf{U}_2\mathbf{U}_2^T = \mathbf{I}$$

então  $\mathbf{U}_2\mathbf{U}_2^T = \mathbf{I} - \mathbf{U}_1\mathbf{U}_1^T$  Da equação (4.75) sabemos que  $\mathbf{U}_1\mathbf{U}_1^T = \mathbf{A}\mathbf{A}^\dagger$ . Logo por substituição obtemos  $\mathbf{U}_2\mathbf{U}_2^T = \mathbf{I} - \mathbf{A}\mathbf{A}^\dagger$ .

C.Q.D.

---

Pelo que vimos neste capítulo é conveniente neste ponto fazer as seguintes observações:

Se  $P$  quadrada é idempotente, então:

a)  $P^T$  e  $\mathbf{I} - P$  são idempotentes.

b) Seus autovalores são 0 ou 1.

A multiplicidade do autovalor 1 é  $\text{traço}(P)$ .

c)  $\text{posto}(P) = \text{traço}(P)$ .

d)  $P(\mathbf{I} - P) = (\mathbf{I} - P)P = 0$ .

e)  $P\mathbf{x} = \mathbf{x}$  se e somente se  $\mathbf{x} \in \mathfrak{R}(P)$

f)  $\mathcal{N}(P) = \mathfrak{R}(\mathbf{I} - P)$

### 4.9.3 Decomposição em Valores Singulares e Pseudo-Inversa de Matriz

Neste ponto, um pouco de intuição e de raciocínio elementar é oportuno. A decomposição em valores singulares de  $\mathbf{A}_{m \times n}$  permite escrever:

$$\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i, \quad i = 1, \dots, r$$

donde  $\mathbf{A}\mathbf{V}_1 = \Sigma_1 \mathbf{U}_1$ , em termos dos vetores  $\mathbf{v}_i$  do espaço linha e  $\mathbf{u}_i$  do espaço coluna de  $\mathbf{A}$ .

Em termos dos quatro subespaços fundamentais de  $\mathbf{A}$  a operação inversa deve fazer o oposto. Se  $\mathbf{A}\mathbf{v}_i = \sigma_i \mathbf{u}_i$ , então:

$${}^{\text{''}}\mathbf{A}^{-1}\text{''}\mathbf{u}_i = \frac{\mathbf{v}_i}{\sigma_i}$$

Logo os valores singulares de  ${}^{\text{''}}\mathbf{A}^{-1}\text{''}$  são  $1/\sigma_i$ , da mesma forma que os autovalores de  $\mathbf{A}^{-1}$  são  $1/\lambda_i$ , e as bases ortonormais são trocadas: os vetores  $\mathbf{u}_i$  e  $\mathbf{v}_i$  geram respectivamente o espaço linha e o espaço coluna de  ${}^{\text{''}}\mathbf{A}^{-1}\text{''}$ ,  $i = 1, \dots, r$ .

Porém se  $\mathbf{A}^{-1}$  não existe, o que fazer ?

Como já exploramos suficientemente neste capítulo o conceito de pseudo-inversa, e como existe uma matriz que multiplicada por  $\mathbf{u}_i$  produz  $\mathbf{v}_i/\sigma_i$ , evidentemente esta é a matriz pseudo-inversa de  $\mathbf{A}$  :  $\mathbf{A}^\dagger$ , e podemos escrever:

$$\mathbf{A}^\dagger = \sum_{i=1}^r \sigma_i^{-1} \mathbf{v}_i \mathbf{u}_i^T \quad (4.79)$$

ou equivalentemente:

$$\begin{aligned} \mathbf{A}^\dagger &= [\mathbf{v}_1 \quad \dots \quad \mathbf{v}_r \quad \dots \quad \mathbf{v}_n] \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} [\mathbf{u}_1 \quad \dots \quad \mathbf{u}_r \quad \dots \quad \mathbf{u}_m]^T = \\ &= \mathbf{V} \Sigma^\dagger \mathbf{U}^T = \mathbf{V}_1 \Sigma_1^\dagger \mathbf{U}_1^T \end{aligned} \quad (4.80)$$

onde:

$$\Sigma^\dagger = \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & \ddots & \\ & & & & 0 \end{bmatrix} = \text{diag}(\sigma_i^{-1})$$

com:

$$\sigma_i^{-1} = \begin{cases} 1/\sigma_i & \text{para } i \leq r \\ 0 & \text{para } i > r \end{cases}$$

Portanto a matriz  $\mathbf{A}^\dagger$  além de única, pode ser obtida através da decomposição em valores singulares de  $\mathbf{A}$ .

Por outro lado pode-se relacionar a decomposição em valores singulares e a matriz pseudo-inversa de maneira formal através do seguinte procedimento relacionado à solução de mínimos quadrados de  $\mathbf{Ax} = \mathbf{b}$ :

Das equações:

$$(4.32): \quad \bar{\mathbf{b}} = \mathbf{Pb} \quad ,$$

$$(4.33): \quad \mathbf{A}\bar{\mathbf{x}} = \bar{\mathbf{b}} \quad ,$$

$$(4.34): \quad \bar{\mathbf{x}} = \mathbf{A}^\dagger \bar{\mathbf{b}} \quad ,$$

$$(4.55): \quad \mathbf{A} = \mathbf{U}_1 \Sigma_1 \mathbf{V}_1^T \quad \text{e}$$

$$(4.75): \quad \mathbf{P} = \mathbf{U}_1 \mathbf{U}_1^T \quad \text{tem-se:}$$

$$\bar{\mathbf{b}} = \mathbf{Pb} = \mathbf{U}_1 \mathbf{U}_1^T \mathbf{b}$$

donde

$$\mathbf{U}_1 \Sigma_1 \mathbf{V}_1^T \bar{\mathbf{x}} = \mathbf{U}_1 \mathbf{U}_1^T \mathbf{b}$$

então:

$$\Sigma_1 \mathbf{V}_1^T \bar{\mathbf{x}} = \mathbf{U}_1^T \mathbf{b}$$

portanto:

$$\mathbf{A}^\dagger = \mathbf{V}_1 \Sigma_1^{-1} \mathbf{U}_1^T = \sum_{i=1}^r \mathbf{v}_i \sigma_i^{-1} \mathbf{u}_i \quad (4.81)$$

**Exemplo 4.13** *A pseudo-inversa da matriz:*

$$\mathbf{A} = \begin{bmatrix} 2 & 2 \\ -1 & 1 \end{bmatrix}$$

*obtida através da sua decomposição em valores singulares é:*

$$\mathbf{A}^\dagger = \mathbf{V} \Sigma^\dagger \mathbf{U}^T = \begin{bmatrix} 0.25 & -0.5 \\ 0.25 & 0.5 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/2\sqrt{2} & 0 \\ 0 & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^T$$

Na Figura 4.24 temos uma representação geométrica da pseudo-inversa nos espaços fundamentais, onde o espaço linha da matriz  $\mathbf{A}^\dagger$  é formado pelos vetores  $\mathbf{u}_i$  e o espaço coluna pelos vetores  $\mathbf{v}_i$  da decomposição em valores singulares da matriz  $\mathbf{A}$ , e como já ressaltado as bases são trocadas.

Portanto  $\mathbf{A}$  tem uma pseudo-inversa  $\mathbf{A}^\dagger$  tal que:

$$\mathbf{A}^\dagger \mathbf{u}_i = \frac{\mathbf{v}_i}{\sigma_i} \quad \text{para } i \leq r \quad (4.82)$$

e

$$\mathbf{A}^\dagger \mathbf{u}_i = 0 \quad \text{para } i > r \quad (4.83)$$

#### 4.9.4 Interpretações Geométricas da Decomposição em Valores Singulares

Dos estudos anteriores podemos escrever:

$$\mathbf{A}_{m \times n} = \mathbf{U}_{m \times m} \Sigma_{m \times n} \mathbf{V}_{n \times n}^T = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \quad (4.84)$$

e

$$\mathbf{A}_{n \times m}^\dagger = \mathbf{V}_{n \times n} \Sigma_{n \times m}^{-1} \mathbf{U}_{m \times m}^T = \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix} \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{U}_1^T \\ \mathbf{U}_2^T \end{bmatrix} \quad (4.85)$$

onde:

- $\mathbf{U}_1$  tem dimensão  $m \times r$ .
- $\mathbf{U}_2$  tem dimensão  $m \times (m - r)$ .

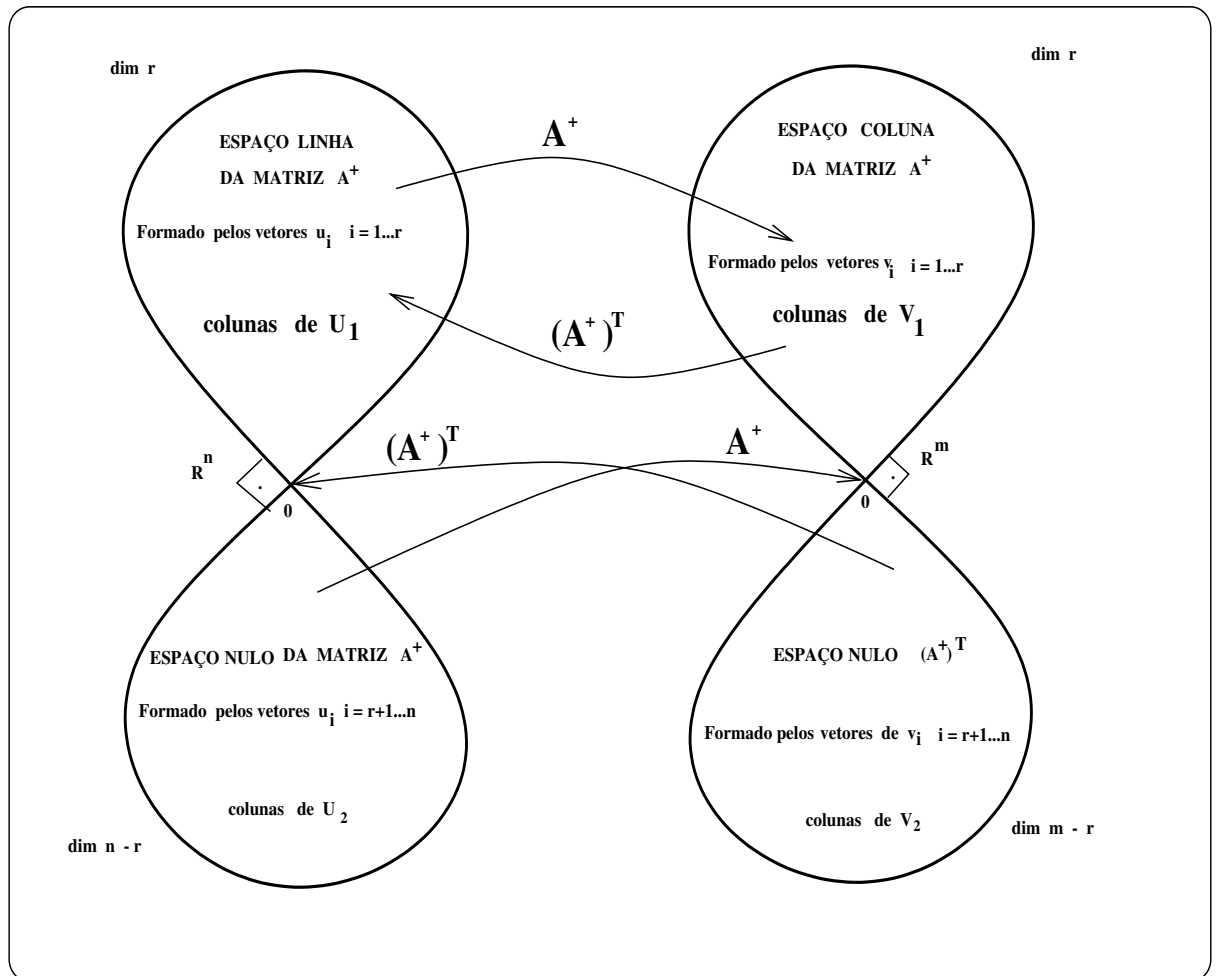


Figura 4.24: Representação Geométrica da Pseudo-Inversa nos Espaços Fundamentais.



- $\mathbf{V}_1$  tem dimensão  $n \times r$ .
- $\mathbf{V}_2$  tem dimensão  $n \times (n - r)$ .

As matrizes  $\mathbf{U}_1$ ,  $\mathbf{U}_2$ ,  $\mathbf{V}_1$  e  $\mathbf{V}_2$  fornecem bases ortogonais para os quatro subespaços fundamentais associados com a matriz  $\mathbf{A}$ ;  $\mathbf{U}_1$  e  $\mathbf{V}_1$  geram respectivamente o espaço coluna e o espaço linha da matriz  $\mathbf{A}$ .

A propriedade que faz a decomposição em valores singulares especial é o fato que  $\Sigma_{1r \times r}$  é uma matriz diagonal, invertível, de posto  $r$  que permite um desacoplamento tal que:

$$\mathbf{A} = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (4.86)$$

$$\mathbf{A}^\dagger = \sum_{i=1}^r \mathbf{v}_i \sigma_i^{-1} \mathbf{u}_i^T \quad (4.87)$$

$$\sigma_i \mathbf{u}_i = \mathbf{A} \mathbf{v}_i \quad (4.88)$$

e

$$\sigma_i \mathbf{v}_i = \mathbf{A}^\dagger \mathbf{u}_i \quad (4.89)$$

com  $\mathbf{u}_i$ , a  $i$ -ésima coluna da matriz  $\mathbf{U}_1$  e  $\mathbf{v}_i$ , a  $i$ -ésima coluna da matriz  $\mathbf{V}_1$ .

Isto cria a possibilidade de ordenar os vetores no espaço coluna e no espaço linha da matriz  $\mathbf{A}$ : a direção mais importante do espaço coluna é  $\mathbf{u}_1$ , com escala  $\sigma_1$ , e é atingida aplicando  $\mathbf{A}$  ao vetor  $\mathbf{v}_1$ ; a segunda direção mais importante será  $\mathbf{u}_2$  ... e assim sucessivamente.

Destas discussões podemos representar os subespaços fundamentais em termos das bases ortonormais  $\mathbf{u}_i$  e  $\mathbf{v}_i$  pela Figura 4.25.

Para outra interpretação da decomposição em valores singulares observemos que nela ocorrem rotações, reflexões e estiramento, como ilustrado na Figura 4.26. Note que  $\Sigma$  é uma matriz de estiramento. De fato os valores singulares,  $\sigma_i$ , são afetados pelas transformações e trocas de escalas, e isto é evidenciado pelo exemplo a seguir:

#### Exemplo 4.14

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} = \mathbf{U} \Sigma \mathbf{V}^T = \begin{bmatrix} 0.4046 & 0.9145 \\ 0.9145 & -0.4046 \end{bmatrix} \begin{bmatrix} 5.4650 & 0 \\ 0 & 0.3660 \end{bmatrix} \begin{bmatrix} 0.5760 & -0.8174 \\ 0.8174 & 0.5760 \end{bmatrix}^T$$

$$\mathbf{A} = \begin{bmatrix} 10 & 20 \\ 30 & 40 \end{bmatrix} = \mathbf{U} \Sigma \mathbf{V}^T = \begin{bmatrix} 0.4046 & 0.9145 \\ 0.9145 & -0.4046 \end{bmatrix} \begin{bmatrix} 54.6499 & 0 \\ 0 & 3.6579 \end{bmatrix} \begin{bmatrix} 0.5760 & -0.8174 \\ 0.8174 & 0.5760 \end{bmatrix}^T$$

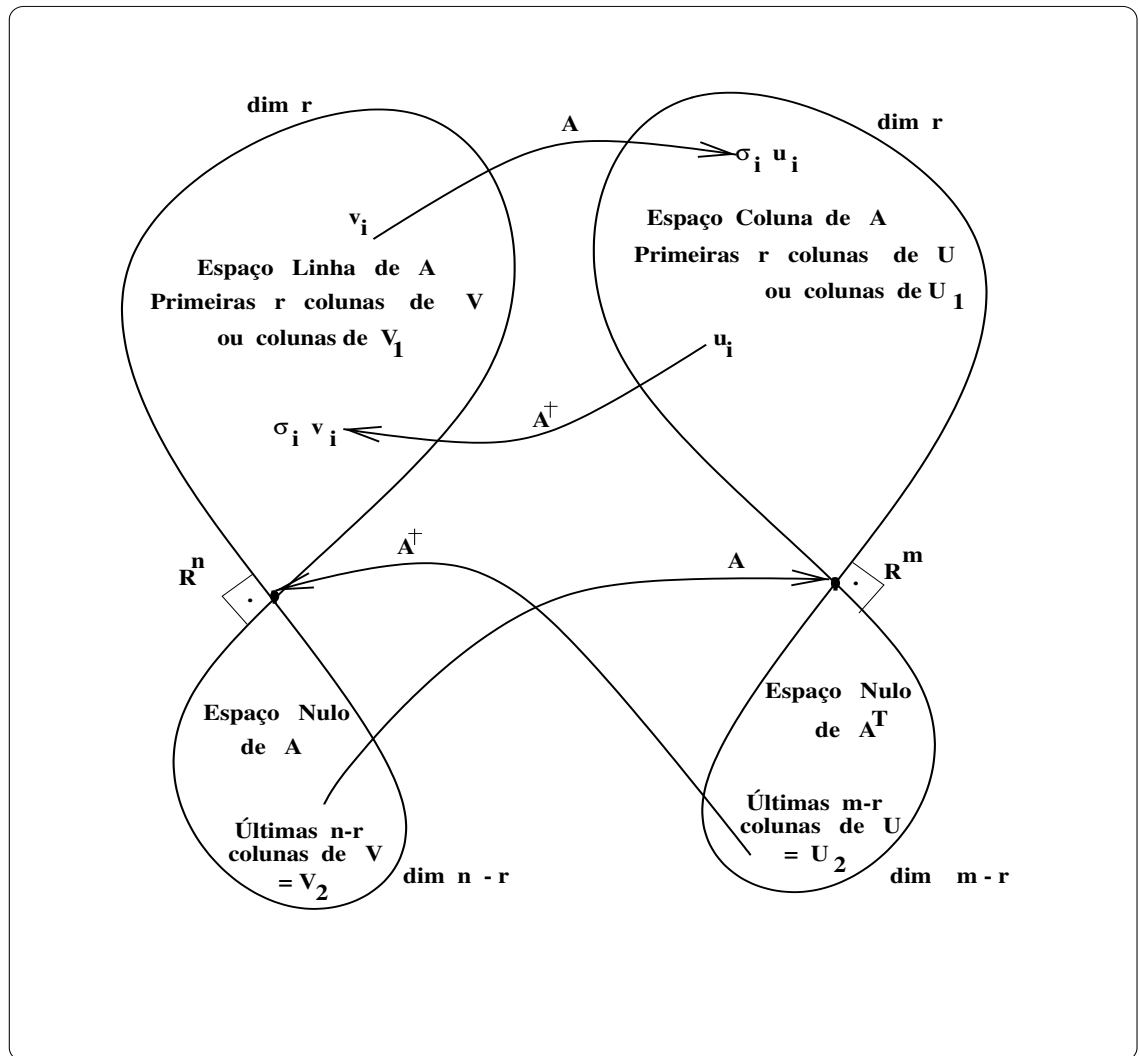


Figura 4.25: Representação Geométrica da Decomposição em Valores Singulares nos Espaços Fundamentais .

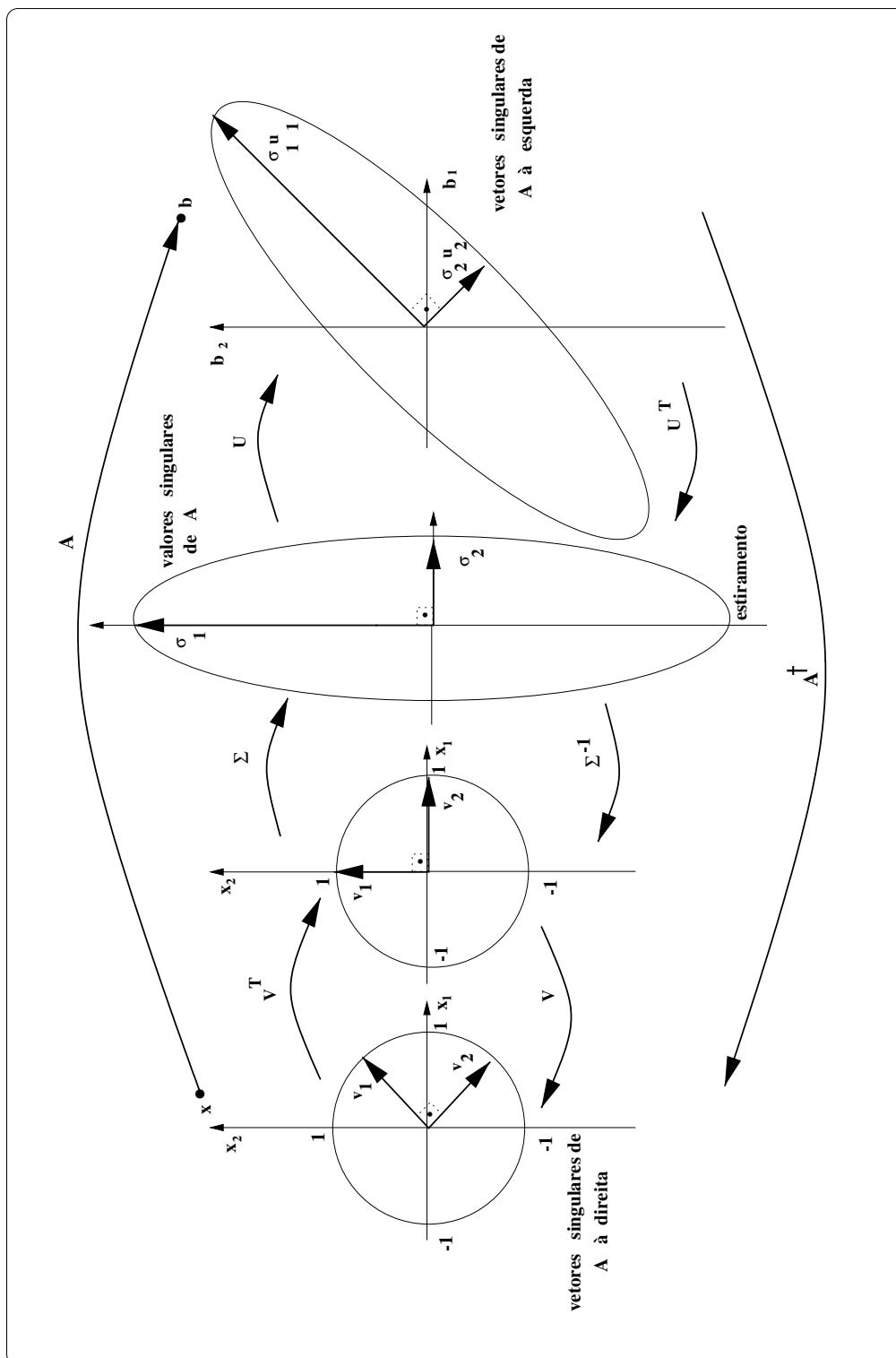


Figura 4.26: Rotações, Reflexões e Estiramento na Decomposição em Valores Singulares de  $A_{2 \times 2}$ .

A decomposição em valores singulares da matriz  $\mathbf{A}$  mostra o comportamento do mapeamento  $\mathbf{Ax} = \mathbf{b}$ :  $\mathbf{x}$  é rodado no espaço  $r$ -dimensional por  $\mathbf{V}^T$ , então escalado por  $\sigma_i$ ,  $i \leq r$  ( para  $i > r$  os vetores são projetados em zero) e finalmente rodado no espaço por  $\mathbf{U}$  resultando  $\mathbf{b}$ , Figura 4.26.

Observe que pela transformação  $\mathbf{A}_{m \times n}$  a imagem de uma hipersfera unitária é uma hiperelipse.

#### 4.9.5 Propriedades da Decomposição em Valores Singulares

A seguir relacionamos algumas propriedades da decomposição em valores singulares de uma matriz  $\mathbf{A}_{m \times n}$  de posto  $r$ :

•

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T, \quad \mathbf{A}^T = \mathbf{V}\mathbf{\Sigma}\mathbf{U}^T, \quad \mathbf{A}^\dagger = \mathbf{V}\mathbf{\Sigma}^{-1}\mathbf{U}^T$$

•

$$(\mathbf{A}^\dagger)^T = (\mathbf{A}^T)^\dagger = \mathbf{U}\mathbf{\Sigma}^{-1}\mathbf{V}^T$$

•

$$\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{\Sigma}^2\mathbf{V}^T, \quad \mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{\Sigma}^2\mathbf{U}^T$$

•

$$(\mathbf{A}^T\mathbf{A})^\dagger = \mathbf{A}^\dagger\mathbf{A}^{T\dagger} = \mathbf{V}\mathbf{\Sigma}^{-2}\mathbf{V}^T,$$

•

$$(\mathbf{A}\mathbf{A}^T)^\dagger = \mathbf{A}^{T\dagger}\mathbf{A}^\dagger = \mathbf{U}\mathbf{\Sigma}^{-2}\mathbf{U}^T,$$

•

$$\mathbf{P}_l = \mathbf{A}^\dagger\mathbf{A} = \mathbf{V}\mathbf{V}^T, \quad \mathbf{P} = \mathbf{A}\mathbf{A}^\dagger = \mathbf{U}\mathbf{U}^T$$

•

$$\|\mathbf{A}\| = \|\mathbf{A}^T\| = \|\mathbf{\Sigma}\| = \sigma_1$$

•

$$\|\mathbf{A}^\dagger\| = \|\mathbf{\Sigma}^{-1}\| = \sigma_r^{-1}$$

•

$$\|\mathbf{A}^T\mathbf{A}\| = \|\mathbf{A}\mathbf{A}^T\| = \|\mathbf{\Sigma}^2\| = \sigma_1^2$$

•

$$\| (\mathbf{A}^T \mathbf{A})^\dagger \| = \| (\mathbf{A} \mathbf{A}^T)^\dagger \| = \| \boldsymbol{\Sigma}^{-2} \| = \sigma_r^{-2}$$

onde à norma euclidiana de um vetor  $\mathbf{y} \in \mathbf{R}^m$  :

$$\| \mathbf{y} \| = \left( \sum_{i=1}^m y_i^2 \right)^{1/2} = \langle \mathbf{y}^T, \mathbf{y} \rangle^{1/2}$$

corresponde a seguinte norma de matriz:

**Definição 4.1** A norma espectral de  $\mathbf{A}_{m \times n}$  é o máximo valor singular de  $\mathbf{A}$ :

$$\| \mathbf{A} \| \stackrel{\text{def}}{=} \sigma_1$$

## 4.10 Projeção Oblíqua

Considere uma matriz  $\mathbf{G}_{m \times n}$  tal que:

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} & \dots & g_{1n} \\ g_{21} & g_{22} & \dots & g_{2n} \\ \vdots & & & \\ g_{m1} & g_{m2} & \dots & g_{mn} \end{bmatrix} = \begin{bmatrix} \mathbf{g}_1^T \\ \mathbf{g}_2^T \\ \vdots \\ \mathbf{g}_m^T \end{bmatrix}$$

onde

$$\mathbf{g}_i^T = [g_{i1} \ g_{i2} \ \dots \ g_{in}] , \quad i = 1, 2, \dots, m.$$

Um elemento  $\mathbf{g}$  que pertence ao espaço linha de  $\mathbf{G}$  é uma combinação linear dos vetores linha  $[\mathbf{g}_1, \mathbf{g}_2, \dots, \mathbf{g}_m]$  de  $\mathbf{G}$ :

$$\mathbf{g} = \alpha_1 \mathbf{g}_1 + \alpha_2 \mathbf{g}_2 + \dots + \alpha_m \mathbf{g}_m = [\mathbf{g}_1 \ \mathbf{g}_2 \ \dots \ \mathbf{g}_m] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix} = \sum_{i=1}^m \alpha_i \mathbf{g}_i = \mathbf{G}^T \boldsymbol{\alpha}$$

onde:

$$\boldsymbol{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_m \end{bmatrix}$$

Seja  $\mathbf{g} = \mathbf{x}|\mathbf{G}$  a projeção ortogonal de um vetor arbitrário  $\mathbf{x}$  no espaço linha de  $\mathbf{G}$  que pode ser escrita como  $\mathbf{g} = \mathbf{G}^T \alpha$ , para algum  $\alpha$ ; como estamos tratando de projeção ortogonal,  $\mathbf{g}$  também deve satisfazer  $(\mathbf{x} - \mathbf{g})^T \mathbf{g} = 0$ . Portanto:  $(\mathbf{x} - \mathbf{G}^T \alpha)^T \mathbf{G}^T \alpha = 0$  donde:

$$\mathbf{x}^T \mathbf{G}^T \alpha - \alpha^T \mathbf{G} \mathbf{G}^T \alpha = 0$$

$$(\mathbf{x}^T \mathbf{G}^T - \alpha^T \mathbf{G} \mathbf{G}^T) \alpha = 0$$

como  $\alpha$  é qualquer:

$$\mathbf{x}^T \mathbf{G}^T - \alpha^T \mathbf{G} \mathbf{G}^T = 0$$

transpondo tem-se:

$$\mathbf{G} \mathbf{x} - \mathbf{G} \mathbf{G}^T \alpha = 0$$

$$\mathbf{G} \mathbf{G}^T \alpha = \mathbf{G} \mathbf{x}$$

A solução  $\alpha$  depende de  $\mathbf{G} \mathbf{G}^T$ :

- se  $\mathbf{G} \mathbf{G}^T$  é não singular, então:

$$\alpha = (\mathbf{G} \mathbf{G}^T)^{-1} \mathbf{G} \mathbf{x}$$

- se  $\mathbf{G} \mathbf{G}^T$  é singular, utilizamos a pseudo-inversa de  $\mathbf{G} \mathbf{G}^T$ :

$$\alpha = (\mathbf{G} \mathbf{G}^T)^\dagger \mathbf{G} \mathbf{x}$$

Portanto a projeção  $\mathbf{g} = \mathbf{x}|\mathbf{G}$  é dada por:

$$\mathbf{g} = \mathbf{G}^T \alpha = \mathbf{G}^T (\mathbf{G} \mathbf{G}^T)^\dagger \mathbf{G} \mathbf{x} = \overline{\mathbf{P}}_{\mathbf{G}} \mathbf{x}, \quad \text{onde} \quad \overline{\mathbf{P}} \equiv \mathbf{P}_l$$

onde  $\mathbf{P}_l = \overline{\mathbf{P}}_{\mathbf{G}} = \mathbf{G}^T (\mathbf{G} \mathbf{G}^T)^\dagger \mathbf{G}$  é o operador projeção ortogonal no espaço linha de  $\mathbf{G}$ .

Seja  $\mathbf{a}_i$  um vetor linha de  $\mathbf{A}$ ; a projeção ortogonal de  $\mathbf{a}_i$  no espaço linha de  $\mathbf{G}$  é:

$$\mathbf{a}|\mathbf{G} = \overline{\mathbf{P}}_{\mathbf{G}} \mathbf{a}_i \quad \text{ou} \quad \mathbf{a}|\mathbf{G} = \mathbf{a}_i^T \overline{\mathbf{P}}_{\mathbf{G}} \quad (4.90)$$

então podemos descrever matematicamente a projeção ortogonal do espaço linha de  $\mathbf{A}$  sobre o espaço linha de  $\mathbf{G}$  por:

$$\mathbf{A}|\mathbf{G} = \mathbf{A} \overline{\mathbf{P}}_{\mathbf{G}} = \mathbf{A} \mathbf{G}^T (\mathbf{G} \mathbf{G}^T)^\dagger \mathbf{G} \quad (4.91)$$

Podemos decompor uma matriz  $\mathbf{A}$  como a combinação linear de 2 matrizes ortogonais:  $\mathbf{G}$  e  $\mathbf{G}^\perp$ ; Figura 4.27:

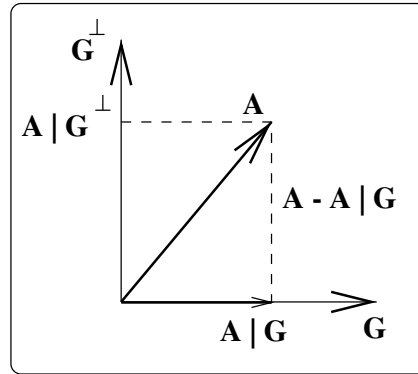


Figura 4.27: Projeção ortogonal de  $\mathbf{A}$  no espaço linha de  $\mathbf{G}$ .

$$\mathbf{A} = \mathbf{A}|_{\mathbf{G}} + \mathbf{A}|_{\mathbf{G}^\perp} \quad (4.92)$$

Das equações (4.91) e (4.92) pode-se escrever:

$$\mathbf{A} = \mathbf{A}|_{\mathbf{G}} + \mathbf{A}|_{\mathbf{G}^\perp} = \mathbf{A}\bar{\mathbf{P}}_{\mathbf{G}} + \mathbf{A}\bar{\mathbf{P}}_{\mathbf{G}^\perp}$$

donde

$$\bar{\mathbf{P}}_{\mathbf{G}^\perp} = \mathbf{I}_n - \bar{\mathbf{P}}_{\mathbf{G}} \quad (4.93)$$

Também podemos decompor  $\mathbf{A}$  através de suas projeções oblíquas, Figura 4.28. Se  $\mathbf{A}/_{\mathbf{G}}\mathbf{F}$  denota a projeção oblíqua de  $\mathbf{A}$  sobre  $\mathbf{F}$  ao longo de  $\mathbf{G}$  e  $\mathbf{A}/_{\mathbf{F}}\mathbf{G}$  denota a projeção oblíqua de  $\mathbf{A}$  sobre  $\mathbf{G}$  ao longo de  $\mathbf{F}$  tem-se:

$$\mathbf{A} = \mathbf{A}/_{\mathbf{F}}\mathbf{G} + \mathbf{A}/_{\mathbf{G}}\mathbf{F} \quad (4.94)$$

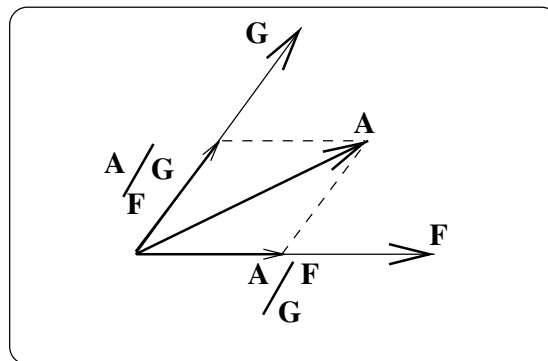


Figura 4.28: Projeção oblíqua de uma matriz  $\mathbf{A}$ .

Se  $P_{\mathbf{G},\mathbf{F}}$  denota o projetor oblíquo sobre  $\mathbf{G}$  ao longo de  $\mathbf{F}$  tem-se:

$$\mathbf{A}/_{\mathbf{F}}\mathbf{G} \equiv P_{\mathbf{G},\mathbf{F}}\mathbf{A} \quad (4.95)$$

Analogamente podemos decompor  $\mathbf{A}$  como combinação linear de duas matrizes  $\mathbf{G}$  e  $\mathbf{F}$  não ortogonais e de seus complementos ortogonais através das projeções ortogonais  $P_{[\mathbf{F},\mathbf{G}]}\mathbf{A}$  e  $P_{[\mathbf{F},\mathbf{G}]}^{\perp}\mathbf{A}$ :

$$\mathbf{A} = \mathbf{A}|[\mathbf{F} \ \mathbf{G}] + \mathbf{A}|[\mathbf{F} \ \mathbf{G}]^{\perp} \equiv P_{[\mathbf{F},\mathbf{G}]}\mathbf{A} + P_{[\mathbf{F},\mathbf{G}]}^{\perp}\mathbf{A} \quad (4.96)$$

que representamos na Figura 4.29.

Como

$$P_{[\mathbf{F},\mathbf{G}]}\mathbf{A} = \mathbf{A}|[\mathbf{F} \ \mathbf{G}] = \mathbf{A}/_{\mathbf{F}}\mathbf{G} + \mathbf{A}/_{\mathbf{G}}\mathbf{F} \equiv P_{\mathbf{G},\mathbf{F}}\mathbf{A} + P_{\mathbf{F},\mathbf{G}}\mathbf{A} \quad (4.97)$$

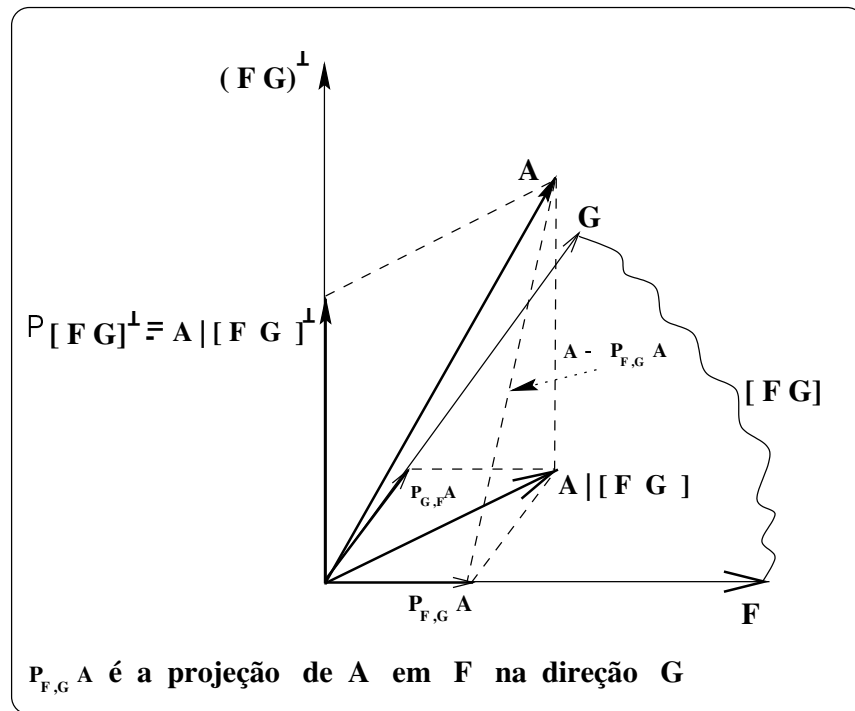


Figura 4.29: Interpretação da projeção oblíqua de uma matriz.

A matriz  $\mathbf{A}$  pode ser descrita por uma combinação das equações (4.96) e (4.97), Figura 4.29:

$$\mathbf{A} = P_{\mathbf{G},\mathbf{F}}\mathbf{A} + P_{\mathbf{F},\mathbf{G}}\mathbf{A} + P_{[\mathbf{F},\mathbf{G}]}^{\perp}\mathbf{A} \quad (4.98)$$

Da equação (4.98) tem-se:



$$\mathbf{I} = P_{\mathbf{G}, \mathbf{F}} + P_{\mathbf{F}, \mathbf{G}} + P_{[\mathbf{F} \ \mathbf{G}]^\perp} \quad (4.99)$$

Do que vimos neste capítulo a projeção ortogonal de  $\mathbf{A}_{m \times n}$  no espaço coluna de  $\mathbf{G}_{q \times n}$  é:

$$\mathbf{A}|\mathbf{G} = P_{\mathbf{G}} \mathbf{A} = \mathbf{G}(\mathbf{G}^T \mathbf{G})^\dagger \mathbf{G}^T \mathbf{A} = \mathbf{U}_{1\mathbf{G}} \mathbf{U}_{1\mathbf{G}}^T \mathbf{A}$$

onde  $\mathbf{G} = \mathbf{U}_{1\mathbf{G}} \boldsymbol{\Sigma}_{1\mathbf{G}} \mathbf{V}_{1\mathbf{G}}^T$ , e o projetor ortogonal no espaço coluna de  $\mathbf{G}$  é:

$$P_{\mathbf{G}} = \mathbf{G}(\mathbf{G}^T \mathbf{G})^\dagger \mathbf{G}^T = \mathbf{U}_{1\mathbf{G}} \mathbf{U}_{1\mathbf{G}}^T$$

para  $\mathbf{G}$  com posto coluna completo.

**Teorema 4.9** *A projeção oblíqua do espaço coluna de  $\mathbf{A}_{m \times n}$ , de posto  $r$ , no espaço coluna de  $\mathbf{F}_{m \times p}$  ao longo do espaço coluna de  $\mathbf{G}_{m \times q}$ , com  $m \geq p+q \geq n \geq r$  é:*

$$\begin{aligned} \mathbf{P}_{\mathbf{F}, \mathbf{G}} \mathbf{A} &= \mathbf{A} /_{\mathbf{F}} \mathbf{G} = \mathbf{F} [\mathbf{I}_r \ 0] \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{G} \\ \mathbf{G}^T \mathbf{F} & \mathbf{G}^T \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{A} = \\ &= [\mathbf{U}_{1\mathbf{F}} \ 0] \begin{bmatrix} \mathbf{U}_{1\mathbf{F}}^T \\ \mathbf{U}_{1\mathbf{G}}^T \end{bmatrix} = [\mathbf{U}_{1\mathbf{F}} \ 0] [\mathbf{U}_{1\mathbf{F}} \ \mathbf{U}_{1\mathbf{G}}]^{-1} \end{aligned} \quad (4.100)$$

onde

$$\mathbf{F} = \mathbf{U}_{1\mathbf{F}} \boldsymbol{\Sigma}_{1\mathbf{F}} \mathbf{V}_{1\mathbf{F}}^T \quad e \quad \mathbf{G} = \mathbf{U}_{1\mathbf{G}} \boldsymbol{\Sigma}_{1\mathbf{G}} \mathbf{V}_{1\mathbf{G}}^T$$

**Prova:** Primeiramente determinemos o projetor ortogonal no espaço coluna formado pela matriz  $[\mathbf{F} \ \mathbf{G}]$  composta pelas colunas de  $\mathbf{F}_{m \times p}$  e  $\mathbf{G}_{m \times q}$  supondo os espaços coluna de  $\mathbf{F}$  e  $\mathbf{G}$  disjuntos o que requer,  $r \leq n \leq p+q \leq m$  (tal que o posto desta matriz seja inferior a  $n$ ).

Para o caso de projeções oblíquas suponhamos que os espaços coluna das matrizes  $\mathbf{F}$  e  $\mathbf{G}$  são dois subespaços de  $\mathbf{A}$  de tal forma que  $\mathbf{F} \oplus \mathbf{G} = \mathbf{A}$ . Note que  $\mathbf{F}$  e  $\mathbf{G}$  não precisam ter colunas linearmente independentes. Então através de uma escolha adequada de base teremos a projeção sobre  $\mathbf{F}$  no sentido de  $\mathbf{G}$ .

O projetor ortogonal no espaço coluna de  $[\mathbf{F} \ \mathbf{G}]$  é:

$$\begin{aligned} P_{[\mathbf{F} \ \mathbf{G}]} &= [\mathbf{F} \ \mathbf{G}] [\mathbf{F} \ \mathbf{G}]^\dagger = [\mathbf{F} \ \mathbf{G}] \left( [\mathbf{F} \ \mathbf{G}]^T \ [\mathbf{F} \ \mathbf{G}] \right)^\dagger [\mathbf{F} \ \mathbf{G}]^T = \\ &= [\mathbf{F} \ \mathbf{G}] \left[ \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \ [\mathbf{F} \ \mathbf{G}] \right]^\dagger [\mathbf{F} \ \mathbf{G}]^T = \end{aligned}$$

$$= [\mathbf{F} \ \mathbf{G}] \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{G} \\ \mathbf{G}^T \mathbf{F} & \mathbf{G}^T \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \quad (4.101)$$

donde a projeção ortogonal de  $\mathbf{A}$  no espaço  $[\mathbf{F} \ \mathbf{G}]$  é:

$$P_{[\mathbf{F} \ \mathbf{G}]} \mathbf{A} = \mathbf{A} | [\mathbf{F} \ \mathbf{G}] = [\mathbf{F} \ \mathbf{G}] \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{G} \\ \mathbf{G}^T \mathbf{F} & \mathbf{G}^T \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{A}$$

O problema estaria solucionado caso  $\mathbf{G}$  e  $\mathbf{F}$  tivessem a mesma dimensão, o que nem sempre ocorre; portanto devemos escrever o espaço coluna gerado por  $\mathbf{G}$  e  $\mathbf{F}$  da seguinte forma:

$$[\mathbf{F} \ \mathbf{G}] = \mathbf{F} [\mathbf{I}_p \ 0] + \mathbf{G} [0 \ \mathbf{I}_q]$$

onde as matrizes identidade  $\mathbf{I}_p$  e  $\mathbf{I}_q$  têm dimensões adequadas para viabilizar a junção dos subespaços.

Logo:

$$\begin{aligned} \mathbf{A} | [\mathbf{F} \ \mathbf{G}] &= (\mathbf{F} [\mathbf{I}_p \ 0] + \mathbf{G} [0 \ \mathbf{I}_q]) \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{G} \\ \mathbf{G}^T \mathbf{F} & \mathbf{G}^T \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{A} = \\ &= \mathbf{U}_{1[\mathbf{F} \ \mathbf{G}]} \mathbf{U}_{1[\mathbf{F} \ \mathbf{G}]}^T \mathbf{A} \end{aligned}$$

onde  $[\mathbf{F} \ \mathbf{G}] = \mathbf{U}_{1[\mathbf{F} \ \mathbf{G}]} \boldsymbol{\Sigma}_{1[\mathbf{F} \ \mathbf{G}]} \mathbf{V}_{1[\mathbf{F} \ \mathbf{G}]}^T$ . Portanto da equação (4.97) podemos escrever:

$$\begin{aligned} P_{\mathbf{F}, \mathbf{G}} \mathbf{A} &= \mathbf{A} /_{\mathbf{F} \ \mathbf{G}} = \mathbf{F} [\mathbf{I}_r \ 0] \begin{bmatrix} \mathbf{F}^T \mathbf{F} & \mathbf{F}^T \mathbf{G} \\ \mathbf{G}^T \mathbf{F} & \mathbf{G}^T \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F}^T \\ \mathbf{G}^T \end{bmatrix} \mathbf{A} = \\ &= [\mathbf{U}_{1\mathbf{F}} \ 0] \mathbf{U}_{[\mathbf{F} \ \mathbf{G}]}^T \mathbf{A} = [\mathbf{U}_{1\mathbf{F}} \ 0] [\mathbf{U}_{1\mathbf{F}}^T \ \mathbf{U}_{1\mathbf{G}}]^{-1} \mathbf{A} \end{aligned}$$

C.Q.D.

**Teorema 4.10** *A projeção oblíqua do espaço linha de  $\mathbf{A}_{m \times n}$ , de posto  $r$ , no espaço linha de  $\mathbf{F}_{p \times n}$ , na direção do espaço linha de  $\mathbf{G}_{q \times n}$ , com  $r \leq m \leq p + q < n$  é:*

$$\mathbf{A} /_{\mathbf{G}} \mathbf{F} = \mathbf{A} [\mathbf{F}^T \ \mathbf{G}^T] \begin{bmatrix} \mathbf{F} \mathbf{F}^T & \mathbf{F} \mathbf{G}^T \\ \mathbf{G} \mathbf{F}^T & \mathbf{G} \mathbf{G}^T \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{I}_r \\ 0 \end{bmatrix} \mathbf{F} = \mathbf{A} [\mathbf{V}_{1\mathbf{F}} \ \mathbf{V}_{1\mathbf{G}}] [\mathbf{V}_{1\mathbf{F}}^T \ 0] \quad (4.102)$$

**Prova:** O projetor ortogonal no espaço linha de  $\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$  é:

$$\begin{aligned} \overline{P} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} &= \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}^T \left[ \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}^T \right]^\dagger \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \left[ \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \right]^\dagger \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \\ &= \mathbf{V}_1 \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \mathbf{V}_1^T \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \end{aligned} \tag{4.103}$$

onde:

$$\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{U}_1 \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \boldsymbol{\Sigma}_1 \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \mathbf{V}_1^T \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$$

Portanto:

$$\mathbf{A} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \overline{P} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$$

O problema estaria solucionado caso  $\mathbf{G}$  e  $\mathbf{F}$  tivessem a mesma dimensão, o que nem sempre ocorre; portanto devemos escrever o espaço linha gerado por  $\mathbf{G}$  e  $\mathbf{F}$  da seguinte forma:

$$\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{I}_p \\ 0 \end{bmatrix} \mathbf{F} + \begin{bmatrix} 0 \\ \mathbf{I}_q \end{bmatrix} \mathbf{G}$$

Observe que as matrizes identidade  $\mathbf{I}_p$  e  $\mathbf{I}_q$  terão dimensões adequadas para viabilizar a junção dos subespaços. Logo,

$$\begin{aligned} \mathbf{A} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} &= \mathbf{A} \overline{P} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^\dagger \left[ \begin{bmatrix} \mathbf{I}_p \\ 0 \end{bmatrix} \mathbf{F} + \begin{bmatrix} 0 \\ \mathbf{I}_q \end{bmatrix} \mathbf{G} \right] = \\ &= \mathbf{A} \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{I}_p \\ 0 \end{bmatrix} \mathbf{F} + \mathbf{A} \begin{bmatrix} \mathbf{F}^T & \mathbf{G}^T \end{bmatrix} \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^\dagger \begin{bmatrix} 0 \\ \mathbf{I}_q \end{bmatrix} \mathbf{G} \end{aligned} \tag{4.104}$$

Observe que a equação (4.104) representa a projeção ortogonal da matriz  $\mathbf{A}$  no espaço formado por:

$$\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$$

que é a soma de duas projeções oblíquas:

1 - a projeção do espaço linha de  $\mathbf{A}$  no espaço linha de  $\mathbf{F}$  na direção do espaço linha de  $\mathbf{G}$

2 - a projeção do espaço linha de  $\mathbf{A}$  no espaço linha de  $\mathbf{G}$  na direção do espaço linha de  $\mathbf{F}$

Portanto como estamos interessados no termo da primeira projeção temos:

A projeção oblíqua: projeção do espaço linha de  $\mathbf{A}_{m \times n}$ , no espaço linha de  $\mathbf{F}_{p \times n}$ , na direção do espaço linha de  $\mathbf{G}_{q \times n}$  é:

$$\begin{aligned} \mathbf{A}\overline{\mathbf{P}}_{\mathbf{F},\mathbf{G}} &= \mathbf{A}/_{\mathbf{G}}\mathbf{F} = \mathbf{A}[\mathbf{F}^T \quad \mathbf{G}^T] \begin{bmatrix} \mathbf{F}\mathbf{F}^T & \mathbf{F}\mathbf{G}^T \\ \mathbf{G}\mathbf{F}^T & \mathbf{G}\mathbf{G}^T \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{I}_r \\ 0 \end{bmatrix} \mathbf{F} = \\ &= \mathbf{A} \begin{bmatrix} \mathbf{V}_{1\mathbf{F}} & \mathbf{V}_{1\mathbf{G}} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{1\mathbf{F}}^T \\ 0 \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{V}_{1\mathbf{F}} & \mathbf{V}_{1\mathbf{G}} \end{bmatrix} \mathbf{V}_{1\mathbf{F}}^{-1} \end{aligned}$$

C.Q.D.

**Corolário 4.6** *A projeção oblíqua do espaço linha de  $\mathbf{A}_{m \times n}$ , de posto  $r$ , no espaço linha de  $\mathbf{F}_{p \times n}$ , na direção do espaço linha de  $\mathbf{G}_{q \times n}$ , com  $r \leq m \leq p + q < n$  é:*

$$\mathbf{A}/_{\mathbf{G}}\mathbf{F} = \mathbf{A}\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \left[ \left( \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \right)^{\dagger} \right]_{:,1:r} \mathbf{F} \quad (4.105)$$

onde o subescrito  $:,1:r$  corresponde a  $r$  primeiras colunas.

**Prova:** Da equação (4.103), o projetor ortogonal de  $\mathbf{A}$  no espaço linha de  $\begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$  é:

$$\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$$

donde

$$\mathbf{A}\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}^{\dagger} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix}$$

Esta projeção ortogonal poderá ser decomposta em duas projeções oblíquas:

$$\mathbf{A}\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A}/_{\mathbf{G}}\mathbf{F} + \mathbf{A}/_{\mathbf{F}}\mathbf{G}$$

de forma análoga esta decomposição é ilustrada para o espaço coluna na Figura 4.29; donde tem-se:

$$\mathbf{A}/_{\mathbf{G}}\mathbf{F} = \mathbf{A}\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = \mathbf{A} \left[ \left( \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} \right)^{\dagger} \right]_{:,1:r} \mathbf{F}$$

C.Q.D.

**Corolário 4.7** *A projeção oblíqua do espaço linha de  $\mathbf{A}_{m \times n}$ , de posto  $r$ , no espaço linha de  $\mathbf{F}_{p \times n}$ , na direção do espaço linha de  $\mathbf{G}_{q \times n}$ , com  $r \leq m \leq p + q < n$  é:*

$$\mathbf{A}/_{\mathbf{G}}\mathbf{F} = \mathbf{A}\overline{\mathbf{P}} \begin{bmatrix} \mathbf{F} \\ \mathbf{G} \end{bmatrix} = [\mathbf{A} \mid \mathbf{G}^{\perp}] [\mathbf{F} \mid \mathbf{G}^{\perp}]^{\dagger} \mathbf{F} \quad (4.106)$$

**Prova:** Pode-se escrever que:

$$\mathbf{A}/_{\mathbf{G}}\mathbf{F} \mid \mathbf{G}^{\perp} = \mathbf{A} \mid \mathbf{G}^{\perp} \quad (4.107)$$

do mesmo modo tem-se que:

$$\mathbf{A} \mid \mathbf{G}^{\perp} = \alpha [\mathbf{F} \mid \mathbf{G}^{\perp}] \quad (4.108)$$

donde tem-se:

$$\alpha = [\mathbf{A} \mid \mathbf{G}^{\perp}] [\mathbf{F} \mid \mathbf{G}^{\perp}]^{\dagger} \quad (4.109)$$

das equações (4.108) e (4.107) tem-se que:

$$\alpha [\mathbf{F} \mid \mathbf{G}^{\perp}] = \mathbf{A}/_{\mathbf{G}}\mathbf{F} \mid \mathbf{G}^{\perp}$$

logo

$$\alpha [\mathbf{F}] = \mathbf{A}/_{\mathbf{G}}\mathbf{F} \quad (4.110)$$

substituindo a equação (4.109) em (4.110) tem-se:

$$\mathbf{A}/_{\mathbf{G}}\mathbf{F} = [\mathbf{A} \mid \mathbf{G}^{\perp}] [\mathbf{F} \mid \mathbf{G}^{\perp}]^{\dagger} \mathbf{F}$$

Para ilustrar esta prova apresentamos a Figura 4.30.

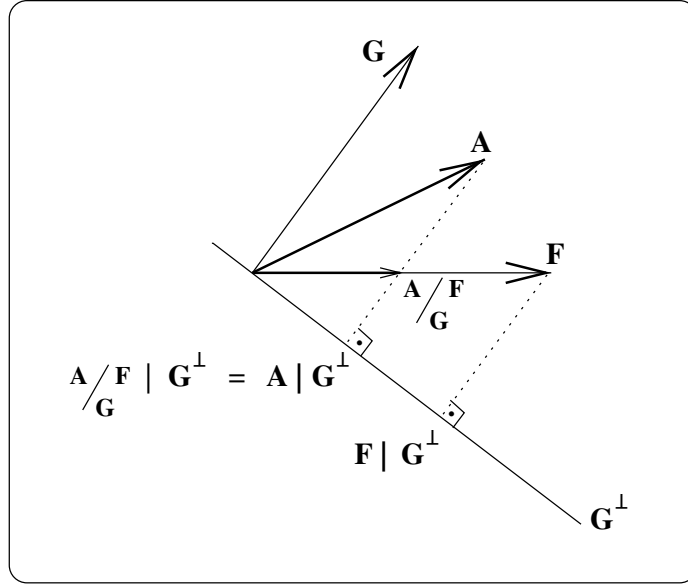


Figura 4.30: Representação Geométrica de Uma Projeção Oblíqua em Função de  $\mathbf{G}^\perp$ .

C.Q.D.

Note que quando  $\mathbf{G} = \mathbf{0}$  ou quando  $\mathbf{G}$  for ortogonal ao espaço linha de  $\mathbf{F}$  ( $\mathbf{G}\mathbf{F}^T = 0$ ), a projeção oblíqua se reduz a uma projeção ortogonal,  $\mathbf{A}/_{\mathbf{G}}\mathbf{F} = \mathbf{A}|\mathbf{F}$ . Para o caso especial em que  $\mathbf{G} = \mathbf{F}^\perp$  temos um projetor ortogonal:

$$\begin{aligned}\overline{P}_{\mathbf{F},\mathbf{F}^\perp} &= \begin{bmatrix} \mathbf{V}_{1\mathbf{F}} & \mathbf{V}_{1\mathbf{G}} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{1\mathbf{F}} & 0 \end{bmatrix}^{-1} = \\ &= \begin{bmatrix} \mathbf{V}_{1\mathbf{F}} & \mathbf{V}_{1\mathbf{G}} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{1\mathbf{F}}^T \\ 0 \end{bmatrix} = \\ &= \mathbf{V}_{1\mathbf{F}}\mathbf{V}_{1\mathbf{F}}^T = \mathbf{F}^\dagger\mathbf{F} = \overline{P}_{\mathbf{F}}\end{aligned}$$

onde  $\overline{P}_{\mathbf{F}}$  é a projeção ortogonal da matriz  $\mathbf{F}$  e é idêntica à projeção oblíqua de  $\mathbf{F}$  ao longo de  $\mathbf{F}^\perp$ .

## 4.11 Ângulo e Direções Principais Entre Subespaços

O cosseno do ângulo  $\theta$  entre dois vetores coluna  $\mathbf{x}$  e  $\mathbf{y}$  é dado por:

$$\cos \theta = \frac{\mathbf{x}^T \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{y}\|} \equiv \frac{\mathbf{x}^T \mathbf{y}}{(\mathbf{x}^T \mathbf{x})^{1/2} (\mathbf{y}^T \mathbf{y})^{1/2}} \equiv (\mathbf{x}^T \mathbf{x})^{-1/2} (\mathbf{x}^T \mathbf{y}) (\mathbf{y}^T \mathbf{y})^{-1/2} \quad (4.111)$$

portanto, o cosseno do ângulo  $\alpha$  entre dois vetores linha  $\mathbf{x}^T$  e  $\mathbf{y}^T$  será:

$$\cos \alpha = \frac{\mathbf{xy}^T}{\|\mathbf{x}\| \|\mathbf{y}\|} \equiv \frac{\mathbf{xy}^T}{(\mathbf{xx}^T)^{1/2} (\mathbf{yy}^T)^{1/2}} \equiv (\mathbf{xx}^T)^{-1/2} (\mathbf{xy}^T) (\mathbf{yy}^T)^{-1/2} \quad (4.112)$$

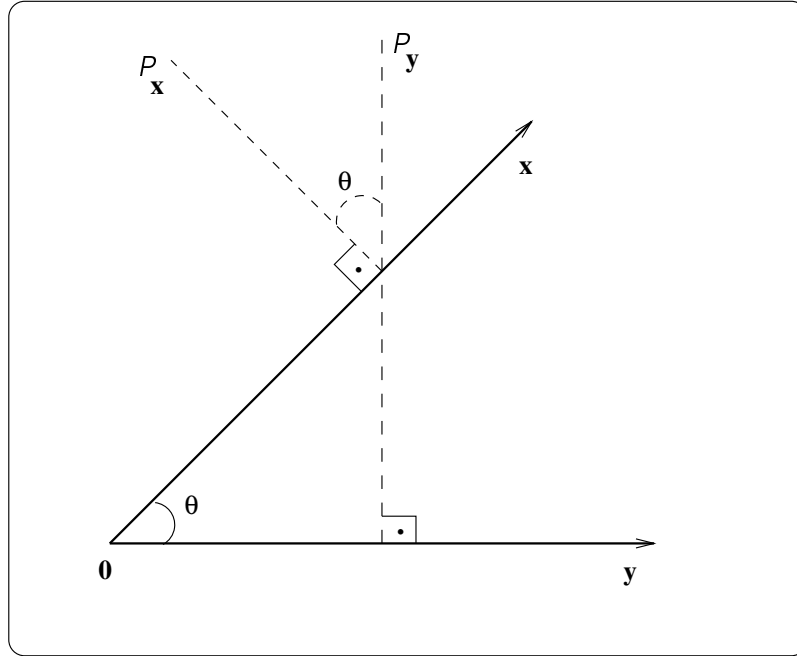


Figura 4.31: Interpretação de ângulo entre vetores.

**Definição 4.2** *Sejam as matrizes  $P_{\mathbf{x}} = \mathbf{x}(\mathbf{xx}^T)^{-1}\mathbf{x}^T$  e  $P_{\mathbf{y}} = \mathbf{y}(\mathbf{yy}^T)^{-1}\mathbf{y}^T$  os projetores ortogonais sobre os vetores coluna  $\mathbf{x}$  e  $\mathbf{y}$  respectivamente. O coeficiente de inclinação  $\mathcal{R}$  entre um par de vetores coluna  $\mathbf{x}$  e  $\mathbf{y}$  separados pelo ângulo  $\theta$ , Figura 4.31, é dado por:*

$$\begin{aligned} \mathcal{R} &= \cos^2 \alpha = (\mathbf{xx}^T)^{-1} (\mathbf{x}^T \mathbf{y})^2 (\mathbf{yy}^T)^{-1} \equiv \\ &\equiv \mathbf{x}(\mathbf{xx}^T)^{-1} \mathbf{x}^T \mathbf{y} (\mathbf{yy}^T)^{-1} \mathbf{y}^T \equiv P_{\mathbf{x}} P_{\mathbf{y}} \end{aligned} \quad (4.113)$$

Generalizando o conceito de ângulo entre dois vetores conceituamos ângulos e direções principais entre dois subespaços.

Sejam os espaços coluna de duas matrizes  $\mathbf{F}_{m \times p}$  e  $\mathbf{G}_{m \times q}$ . O primeiro ângulo principal  $\theta_1$  (o menor de todos) é obtido pela escolha de vetores unitários  $\mathbf{f}_1 \in \mathfrak{R}(\mathbf{F})$  e  $\mathbf{g}_1 \in \mathfrak{R}(\mathbf{G})$  e minimizando o ângulo entre eles. Este é o primeiro ângulo principal e os vetores unitários  $\mathbf{f}_1$  e  $\mathbf{g}_1$  são as primeiras direções principais. A seguir escolhemos um vetor unitário  $\mathbf{f}_2 \in \mathfrak{R}(\mathbf{F})$  ortogonal à  $\mathbf{f}_1$  e  $\mathbf{g}_2 \in \mathfrak{R}(\mathbf{G})$  ortogonal à  $\mathbf{g}_1$  e minimizamos o ângulo  $\theta_2$  entre eles. Estes são os segundos ângulos e direções principais. Continuamos desta forma até determinar  $\min(p, q)$  ângulos.

Portanto:

**Definição 4.3** *Os ângulos principais  $0 \leq \theta_1 \leq \theta_2 \leq \dots \leq \pi/2$  entre os espaços coluna de  $\mathbf{F}_{m \times p}$  e  $\mathbf{G}_{m \times q}$  e as direções principais correspondentes  $\mathbf{f}_i \in \mathfrak{R}(\mathbf{F})$  e  $\mathbf{g}_i \in \mathfrak{R}(\mathbf{G})$  são definidos recursivamente para  $i = 1, 2, \dots, \min(p, q)$  por:*

$$\theta_i = \cos^{-1} \left( \max_{\mathbf{f} \in \mathfrak{R}(\mathbf{F})} \max_{\mathbf{g} \in \mathfrak{R}(\mathbf{G})} \mathbf{f}^T \mathbf{g} \right) = \cos^{-1}(\mathbf{f}_i^T \mathbf{g}_i) \quad (4.114)$$

sujeito as restrições:

$$\begin{aligned} \|\mathbf{f}\| &= 1 \quad e \\ \|\mathbf{g}\| &= 1 \\ e \text{ para } i > 1: \\ \mathbf{f}_j^T \mathbf{f} &= 0, \\ \mathbf{g}_j^T \mathbf{g} &= 0, \quad j = 1, 2, \dots, i-1. \end{aligned}$$

Observe que  $\mathbf{f}_i$  e  $\mathbf{g}_i$  são  $\mathbf{f}$  e  $\mathbf{g}$  que atingem os máximos indicados na equação (4.114) e que para o  $i$ -ésimo ângulo principal eles são restritos à serem ortogonais à todos os  $\mathbf{f}_j$  e  $\mathbf{g}_j$  anteriores, respectivamente.

Sejam  $P_{\mathbf{F}} = \mathbf{F}(\mathbf{F}^T \mathbf{F})^{-1} \mathbf{F}^T, p \times p$ , e  $P_{\mathbf{G}} = \mathbf{G}(\mathbf{G}^T \mathbf{G})^{-1} \mathbf{G}^T, q \times q$ , os projetores ortogonais sobre os espaços coluna das matrizes  $\mathbf{F}$  e  $\mathbf{G}$ , subespaços do espaço coluna de  $\mathbf{A}$ .

Sejam  $p = \text{posto}(P_{\mathbf{F}}) = \text{traço}(P_{\mathbf{F}})$  e  $q = \text{posto}(P_{\mathbf{G}}) = \text{traço}(P_{\mathbf{G}})$ . Os números  $\rho = \text{posto}(P_{\mathbf{F}} P_{\mathbf{G}})$  e  $\mathcal{R} = \text{traço}(P_{\mathbf{F}} P_{\mathbf{G}})$  tais que  $\mathcal{R} \leq \rho$  definem a dimensão e o coeficiente de inclinação entre  $\mathbf{F}$  e  $\mathbf{G}$ :

- a)  $\rho = \mathcal{R}$  se e somente se  $\mathbf{F}$  e  $\mathbf{G}$  são ortogonalmente inclinados,
- b)  $\rho = \mathcal{R} = 0$  se e somente se  $\mathbf{F}$  e  $\mathbf{G}$  são ortogonais,
- c)  $\mathbf{F}$  é completamente inclinado para  $\mathbf{G}$  se  $\rho = p$  e
- d)  $\mathbf{F}$  é inclinado em  $\mathbf{G}$  se  $\mathcal{R} = p$ .



Do que discutimos anteriormente podemos enunciar os seguintes resultados:

**Teorema 4.11** *Os ângulos e direções principais entre os espaços coluna das matrizes  $\mathbf{F}_{m \times p}$  e  $\mathbf{G}_{m \times q}$  são dados pela decomposição em valores singulares da matriz  $P_{\mathbf{F}}P_{\mathbf{G}} = \mathbf{U}\Sigma\mathbf{V}^T$ , onde  $P_{\mathbf{F}} = \mathbf{F}(\mathbf{F}^T\mathbf{F})^{-1}\mathbf{F}^T$  e  $P_{\mathbf{G}} = \mathbf{G}(\mathbf{G}^T\mathbf{G})^{-1}\mathbf{G}^T$ .*

*As direções principais entre  $\Re(\mathbf{F})$  e  $\Re(\mathbf{G})$  são dadas pelas colunas de  $\mathbf{U}$  e pelas colunas de  $\mathbf{V}$ . Os cossenos dos ângulos principais entre  $\Re(\mathbf{F})$  e  $\Re(\mathbf{G})$  são os valores singulares de  $\Sigma$ .*

Em decorrência podemos escrever:

$$\text{a) } \Re(\mathbf{A}) \triangleleft \Re(\mathbf{B}) \stackrel{\text{def}}{=} \mathbf{U}$$

$$\text{b) } \Re(\mathbf{A}) \triangleleft \Re(\mathbf{B}) \stackrel{\text{def}}{=} \mathbf{V}$$

$$\text{c) } \Re(\mathbf{A}) \triangleleft \Re(\mathbf{B}) \stackrel{\text{def}}{=} \Sigma$$

**Corolário 4.8** *O  $i$ -ésimo ângulo principal entre  $\Re(\mathbf{F})$  e  $\Re(\mathbf{G})$  é dado por:*

$$\cos \theta_i = \mathbf{f}_i^T \mathbf{g}_i = \gamma_i \quad (4.115)$$

onde  $\gamma_i$  é o  $i$ -ésimo valor singular de  $\mathbf{U}_{1F}^T \mathbf{U}_{1G}$

Analogamente

**Definição 4.4** *Os ângulos principais  $0 \leq \alpha_1 \leq \alpha_2 \leq \dots \leq \pi/2$  entre os espaços linha de  $\mathbf{F}_{p \times n}$  e  $\mathbf{G}_{q \times n}$  e as direções principais correspondentes  $\mathbf{f}_i^T \in \Re(\mathbf{F}^T)$  e  $\mathbf{g}_i^T \in \Re(\mathbf{G}^T)$  são definidos recursivamente para  $i = 1, 2, \dots, \min(p, q)$  por:*

$$\alpha_i = \cos^{-1} \left( \max_{\mathbf{f}^T \in \Re(\mathbf{F}^T)} \max_{\mathbf{g}^T \in \Re(\mathbf{G}^T)} \mathbf{f}^T \mathbf{g} \right) = \cos^{-1}(\mathbf{f}_i^T \mathbf{g}_i) \quad (4.116)$$

sujeito às restrições  $\|\mathbf{f}\| = 1$  e  $\|\mathbf{g}\| = 1$  e para  $i > 1$ :  $\mathbf{f}_j^T \mathbf{f} = 0$ ,  $\mathbf{g}_j^T \mathbf{g} = 0$ ,  $j = 1, 2, \dots, i-1$ .

**Teorema 4.12** *Os ângulos e direções principais entre os espaços linha das matrizes  $\mathbf{F}_{p \times n}$  e  $\mathbf{G}_{q \times n}$  são dados pela decomposição em valores singulares da matriz  $\overline{P}_{\mathbf{F}}\overline{P}_{\mathbf{G}} = \mathbf{U}\Sigma\mathbf{V}^T$ , onde  $\overline{P}_{\mathbf{F}} = \mathbf{F}^T(\mathbf{F}\mathbf{F}^T)^{-1}\mathbf{F}$  e  $\overline{P}_{\mathbf{G}} = \mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^{-1}\mathbf{G}$ .*

*As direções principais entre  $\Re(\mathbf{F}^T)$  e  $\Re(\mathbf{G}^T)$  são dadas pelas linhas de  $\mathbf{U}^T$  e pelas linhas de  $\mathbf{V}^T$ . Os cossenos dos ângulos principais entre  $\Re(\mathbf{F}^T)$  e  $\Re(\mathbf{G}^T)$  são os valores singulares de  $\Sigma$ .*

Em decorrência podemos escrever:

$$a) \mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T) \stackrel{\text{def}}{=} \mathbf{U}^T$$

$$b) \mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T) \stackrel{\text{def}}{=} \mathbf{V}^T$$

$$c) \mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T) \stackrel{\text{def}}{=} \Sigma$$

**Corolário 4.9** *O  $i$ -ésimo ângulo principal entre  $\mathfrak{R}(\mathbf{F}^T)$  e  $\mathfrak{R}(\mathbf{G}^T)$  é dado por:*

$$\cos \alpha_i = \mathbf{f}_i^T (\mathbf{g}_i^T)^T = \beta_i$$

onde  $\beta_i$  é o  $i$ -ésimo valor singular de  $\mathbf{V}_{1F}^T \mathbf{V}_{1G}$ .

**Corolário 4.10**

$$\beta_i = \gamma_i \Rightarrow \alpha_i = \theta_i$$

Para formas alternativas de cálculo dos ângulos e direções principais podemos, com base nas discussões anteriores, enunciar os teoremas seguintes:

**Teorema 4.13** *Os ângulos e direções principais entre os espaços coluna de  $\mathbf{F}$  e  $\mathbf{G}$  são dados pela decomposição em valores singulares da matriz:*

$$(\mathbf{F}^T \mathbf{F})^{-1/2} (\mathbf{F}^T \mathbf{G}) (\mathbf{G}^T \mathbf{G})^{-1/2} = \mathbf{U} \Sigma \mathbf{V}^T \quad (4.117)$$

Como:

$$a) [\mathfrak{R}(\mathbf{A}) \triangleleft \mathfrak{R}(\mathbf{B})] = \mathbf{U}^T (\mathbf{F}^T \mathbf{F})^{-1/2} \mathbf{F}^T$$

$$b) [\mathfrak{R}(\mathbf{A}) \triangleleft \mathfrak{R}(\mathbf{B})] = \mathbf{V}^T (\mathbf{G}^T \mathbf{G})^{-1/2} \mathbf{G}^T$$

$$c) [\mathfrak{R}(\mathbf{A}) \triangleleft \mathfrak{R}(\mathbf{B})] = \Sigma$$

**Teorema 4.14** *Os ângulos e direções principais entre os espaços linha de  $\mathbf{F}$  e  $\mathbf{G}$  são dados pela decomposição em valores singulares da matriz:*

$$(\mathbf{F} \mathbf{F}^T)^{-1/2} (\mathbf{F} \mathbf{G}^T) (\mathbf{G} \mathbf{G}^T)^{-1/2} = \mathbf{U} \Sigma \mathbf{V}^T \quad (4.118)$$

Como:

$$a) [\mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T)] = \mathbf{U}^T (\mathbf{F} \mathbf{F}^T)^{-1/2} \mathbf{F}$$

$$b) [\mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T)] = \mathbf{V}^T (\mathbf{G} \mathbf{G}^T)^{-1/2} \mathbf{G}$$

$$c) [\mathfrak{R}(\mathbf{A}^T) \triangleleft \mathfrak{R}(\mathbf{B}^T)] = \Sigma$$

Os resultados apresentados nesta secção podem ser estendidos para projeção oblíqua.

**Teorema 4.15** *Sejam os valores singulares  $\underline{\sigma}_i$  do projetor oblíquo  $P_{\mathbf{F},\mathbf{G}} = \underline{\mathbf{U}} \underline{\Sigma}_i \underline{\mathbf{V}}^T$ . Se os ângulos principais entre os dois subespaços  $\Re(\mathbf{F})$   $\Re(\mathbf{G})$  são  $\theta_i = \cos^{-1} \gamma_i$ , então:*

$$\text{sen } \theta_i = \underline{\sigma}_i^{-1}, \quad i = 1, 2, \dots, \min(p, q) \quad (4.119)$$

**Prova:**

Como  $\gamma_i = \cos \theta_i$  é um valor singular de  $\mathbf{U}_{1\mathbf{F}}^T \mathbf{U}_{1\mathbf{G}}$ ,  $\gamma_i^2$  é um autovalor de  $\mathbf{U}_{1\mathbf{F}}^T \mathbf{U}_{1\mathbf{G}} \mathbf{U}_{1\mathbf{G}}^T \mathbf{U}_{1\mathbf{F}} = \mathbf{U}_{1\mathbf{F}}^T P_{\mathbf{G}} \mathbf{U}_{1\mathbf{F}}$  e então:

- a)  $\cos^2 \theta_i$  é um autovalor de  $\mathbf{U}_{1\mathbf{F}}^T P_{\mathbf{G}} \mathbf{U}_{1\mathbf{F}}$  donde,
- b)  $1 - \cos^2 \theta_i$  é um autovalor de  $\mathbf{I} - \mathbf{U}_{1\mathbf{F}}^T P_{\mathbf{G}} \mathbf{U}_{1\mathbf{F}}$  donde,
- c)  $\text{sen}^2 \theta_i$  é um autovalor de  $\mathbf{U}_{1\mathbf{F}}^T (\mathbf{I} - P_{\mathbf{G}}) \mathbf{U}_{1\mathbf{F}}$  donde,
- d)  $1/\text{sen}^2 \theta_i$  é um autovalor de  $(\mathbf{U}_{1\mathbf{F}}^T P_{\mathbf{G}}^\perp \mathbf{U}_{1\mathbf{F}})^{-1}$ .

Como os autovalores são invariantes para uma transformação ortogonal, tem-se também que  $\frac{1}{\text{sen}^2 \theta_i}$  é um autovalor de:

$$\mathbf{U}_{1\mathbf{F}} (\mathbf{U}_{1\mathbf{F}}^T P_{\mathbf{G}}^\perp \mathbf{U}_{1\mathbf{F}})^{-1} \mathbf{U}_{1\mathbf{F}}^T = P_{\mathbf{F},\mathbf{G}} P_{\mathbf{F},\mathbf{G}}^T \quad (4.120)$$

que pode ser obtido de:

$$P_{\mathbf{F},\mathbf{G}} = \mathbf{F} (\mathbf{F}^T P_{\mathbf{G}}^\perp \mathbf{F}^T)^{-1} P_{\mathbf{G}}^\perp$$

Portanto  $\frac{1}{\text{sen}^2 \theta_i}$  é um valor singular  $\underline{\sigma}_i$  de  $P_{\mathbf{F},\mathbf{G}}$  donde:

$$\text{sen } \theta_i = \frac{1}{\underline{\sigma}_i}, \quad i = 1, 2, \dots, \min(p, q).$$

**C.Q.D.**

**Teorema 4.16** *Sejam os valores singulares  $\bar{\sigma}_i$  do projetor oblíquo  $\bar{P}_{\mathbf{F},\mathbf{G}} = \bar{\mathbf{U}} \bar{\Sigma} \bar{\mathbf{V}}$ . Se os ângulos principais entre os dois subespaços  $\Re(\mathbf{F}^T)$  e  $\Re(\mathbf{G}^T)$  são  $\alpha_i = \cos^{-1} \beta_i$ , então:*

$$\text{sen } \alpha_i = \bar{\sigma}_i^{-1}, \quad i = 1, \dots, \min(p, q). \quad (4.121)$$

**Prova:**

Como  $\beta_i = \cos \alpha_i$  é um valor singular de  $\mathbf{V}_{1\mathbf{F}}^T \mathbf{V}_{1\mathbf{G}}$ ,  $\beta_i^2$  é um autovalor de:

$$\mathbf{V}_{1\mathbf{F}}^T \mathbf{V}_{1\mathbf{G}} \mathbf{V}_{1\mathbf{G}}^T \mathbf{V}_{1\mathbf{F}} = \mathbf{V}_{1\mathbf{F}}^T \bar{P}_{\mathbf{G}} \mathbf{V}_{1\mathbf{F}}$$

e então:

- a)  $\beta_i^2 = \cos^2 \alpha_i$  é um autovalor de  $\mathbf{V}_{1\mathbf{F}}^T \overline{\mathbf{P}}_{\mathbf{G}} \mathbf{V}_{1\mathbf{F}}$  donde,
- b)  $1 - \cos^2 \alpha_i$  é um autovalor de  $\mathbf{I} - \mathbf{V}_{1\mathbf{F}}^T \overline{\mathbf{P}}_{\mathbf{G}} \mathbf{V}_{1\mathbf{F}}$  donde,
- c)  $\sin^2 \alpha_i$  é um autovalor de  $\mathbf{V}_{1\mathbf{F}}^T (\mathbf{I} - \overline{\mathbf{P}}_{\mathbf{G}}) \mathbf{V}_{1\mathbf{F}}$  donde,
- d)  $1/\sin^2 \alpha_i$  é um autovalor de  $(\mathbf{V}_{1\mathbf{F}}^T \mathbf{P}_{\mathbf{G}}^\perp \mathbf{V}_{1\mathbf{F}})^{-1}$ .

Como os autovalores são invariantes à uma transformação ortogonal, então:  $\frac{1}{\sin^2 \alpha_i}$  também é um autovalor de:

$$\mathbf{V}_{1\mathbf{F}} (\mathbf{V}_{1\mathbf{F}}^T \mathbf{P}_{\mathbf{G}}^\perp \mathbf{V}_{1\mathbf{F}})^{-1} \mathbf{V}_{1\mathbf{F}}^T = \overline{\mathbf{P}}_{\mathbf{F},\mathbf{G}} \overline{\mathbf{P}}_{\mathbf{F},\mathbf{G}}^T \quad (4.122)$$

que pode ser deduzido de  $\overline{\mathbf{P}}_{\mathbf{F},\mathbf{G}} = \overline{\mathbf{P}}_{\mathbf{G}} \mathbf{F}^T (\mathbf{F} \overline{\mathbf{P}}_{\mathbf{G}}^{-1} \mathbf{F}^T)^{-1} \mathbf{F}$

Portanto  $\frac{1}{\sin^2 \alpha_i}$  é um valor singular  $\bar{\sigma}_i$  de  $\overline{\mathbf{P}}_{\mathbf{F},\mathbf{G}}$ , donde  $\sin \alpha_i = \frac{1}{\bar{\sigma}_i}$ ,  $i = 1, 2, \dots, r$ .

**C.Q.D.**

Como os ângulos principais  $\gamma_i$  vão de 0 a  $\frac{\pi}{2}$ , e os senos de  $\theta_i$  vão de 0 a 1, os valores singulares correspondentes de  $P_{\mathbf{F},\mathbf{G}}$  vão de 1 a  $\infty$ . Sendo de baixo posto, uma projeção oblíqua também tem alguns valores singulares nulos que não correspondem a qualquer ângulo principal. Portanto, também acabamos de mostrar que os valores singulares de uma projeção oblíqua podem ser 0, 1 ou qualquer valor superior a 1.

Pode-se provar que:

$$P_{\mathbf{F},\mathbf{G}} = \mathbf{F}(\mathbf{F}^T P_{\mathbf{G}}^\perp \mathbf{F})^{-1} \mathbf{F}^T P_{\mathbf{G}}^\perp = P_{\mathbf{F}}(\mathbf{I} - \mathbf{G}(\mathbf{G}^T P_{\mathbf{F}}^\perp \mathbf{G})^{-1} \mathbf{G}^T P_{\mathbf{F}}^\perp) \quad (4.123)$$

A primeira igualdade da equação (4.123) mostra que o operador  $P_{\mathbf{G}}^\perp$  é um operador que anula tudo no espaço de interferência  $\mathbf{G}$ .

Prova-se que  $P_{\mathbf{F},\mathbf{G}}$  é idempotente e tem como espaço coluna o de  $\mathbf{F}$  e um espaço nulo que inclui o espaço coluna de  $\mathbf{G}$ . Para completar o espaço nulo, seja o espaço:

$$\mathbf{B} = [\mathbf{F} \quad \mathbf{G}]^\perp,$$

então  $P_{\mathbf{G}}^\perp \mathbf{B} = \mathbf{B}$  e  $\mathbf{F}^T \mathbf{B} = 0$ , donde,

$$P_{\mathbf{F},\mathbf{G}} \mathbf{B} = \mathbf{F}(\mathbf{F}^T P_{\mathbf{G}}^\perp \mathbf{F})^{-1} \mathbf{F}^T P_{\mathbf{G}}^\perp \mathbf{B} = \mathbf{F}(\mathbf{F}^T P_{\mathbf{G}}^\perp \mathbf{F})^{-1} \mathbf{F}^T \mathbf{B} = 0 \quad (4.124)$$

e vemos que  $\Re(\mathbf{B})$  também está no espaço nulo de  $P_{\mathbf{F},\mathbf{G}}$  e então concluímos que

a)

$$\Re(P_{\mathbf{F},\mathbf{G}}) = \Re(\mathbf{F}) \quad \text{e} \quad \mathcal{N}(P_{\mathbf{F},\mathbf{G}}) = \Re([\mathbf{G} \ \mathbf{B}])$$

b)

$$P_{\mathbf{F},\mathbf{G}} = P_{\mathbf{F}}(\mathbf{I} - P_{\mathbf{G},\mathbf{F}}) = P_{\mathbf{F}}P_{\mathbf{F},\mathbf{G}}$$

c)

$$\mathbf{F}^\dagger P_{\mathbf{F},\mathbf{G}} = (P_{\mathbf{G}}^\perp \mathbf{F})^\dagger$$

$$P_{\mathbf{G}}^\perp P_{\mathbf{F},\mathbf{G}} P_{\mathbf{G}}^\perp = P_{\mathbf{G}}^\perp P_{P_{\mathbf{G}}^\perp, \mathbf{F}} P_{\mathbf{G}}^\perp = P_{P_{\mathbf{G}}^\perp, \mathbf{F}}$$

d) *traço* ( $P_{\mathbf{F},\mathbf{G}}$ ) =  $p$ .e) *traço* ( $P_{\mathbf{G},\mathbf{F}}$ ) =  $q$ .

f) Os valores singulares das projeções são 0 ou 1 como o são seus autovaleores.

## 4.12 Ferramentas Estatísticas e Ferramentas Geométricas

Em sua excelente tese de doutorado, *Subspace Identification - Theory - Implementation - Application*, Peter Van Overschee, [1091]; estabelece algumas relações entre ferramentas estatísticas e geométricas das quais nos beneficiaremos neste trabalho, e que apresentamos a seguir.

Em modelagem computacional de dados por métodos de subespaço é típico supor a disponibilidade de longas séries temporais de dados ( $j \rightarrow \infty$ ) e que os dados são ergódicos. Desta forma podemos substituir o operador esperança  $E$  (média sobre um número de experimentos infinito) pelo operador esperança  $E_j$  aplicado à soma de variáveis (média sobre um experimento infinitamente longo). Por exemplo, para a correlação entre as sequências de média zero:  $\mathbf{a}_k \in \mathbf{R}^{n_a}$  e  $\mathbf{e}_k \in \mathbf{R}^{n_e}$ ,  $k = 0, 1, 2, \dots, j$  faz-se:

$$E[\mathbf{a}_k \mathbf{e}_k^T] = \lim_{j \rightarrow \infty} \left[ \frac{1}{j} \sum_{i=0}^j \mathbf{a}_i \mathbf{e}_i^T \right] = E_j \left[ \sum_{i=0}^j \mathbf{a}_i \mathbf{e}_i^T \right]$$

com uma definição óbvia de  $E_j$ :

$$E_j[\bullet] \stackrel{\text{def}}{=} \lim_{j \rightarrow \infty} \frac{1}{j}[\bullet] \quad (4.125)$$

Em contexto estocástico definimos a covariância  $\Phi_{[\mathbf{F}, \mathbf{G}]}$  entre as matrizes  $\mathbf{F}_{p \times n}$  e  $\mathbf{G}_{q \times n}$  como:

$$\Phi_{[\mathbf{F}, \mathbf{G}]} \stackrel{\text{def}}{=} E_j[\mathbf{F}, \mathbf{G}^T] \quad (4.126)$$

Agora podemos estender as ferramentas geométricas até aqui introduzidas em contexto determinístico para o contexto estocástico, ou seja, para a facilidade de notação e para facilitar as deduções teóricas redefinimos as operações geométricas em um contexto estatístico. Esta redefinição consiste, por exemplo para o espaço linha, simplesmente em realizar a seguinte substituição em todas as definições:

$$\mathbf{F}\mathbf{G}^T \leftarrow \Phi_{[\mathbf{F}, \mathbf{G}]}$$

Portanto, em contexto estocástico obtemos para:

$$\mathbf{F}|\mathbf{G} = \mathbf{F}\overline{\mathbf{P}}_{\mathbf{G}} = \mathbf{F}\mathbf{G}^T(\mathbf{G}\mathbf{G}^T)^\dagger \mathbf{G}$$

e

$$\mathbf{F}|\mathbf{G}^\perp = \mathbf{F}(\mathbf{I} - \overline{\mathbf{P}}_{\mathbf{G}})$$

respectivamente:

$$\mathbf{F}|\mathbf{G} = \mathbf{F}\overline{\mathbf{P}}_{\mathbf{G}} = \Phi_{[\mathbf{F}, \mathbf{G}]} \Phi_{[\mathbf{G}, \mathbf{G}]}^\dagger \mathbf{G} \quad (4.127)$$

e

$$\mathbf{F}|\mathbf{G}^\perp = \mathbf{F} - \Phi_{[\mathbf{F}, \mathbf{G}]} \Phi_{[\mathbf{G}, \mathbf{G}]}^\dagger \mathbf{G} \quad (4.128)$$

Analogamente obtemos:

$$\mathbf{A}\overline{\mathbf{P}}_{\mathbf{G}, \mathbf{F}} = \mathbf{A}/_{\mathbf{F}}\mathbf{G} = [\Phi_{[\mathbf{A}, \mathbf{G}]} \quad \Phi_{[\mathbf{A}, \mathbf{F}]}] \begin{bmatrix} \Phi_{[\mathbf{G}, \mathbf{G}]} & \Phi_{[\mathbf{G}, \mathbf{F}]} \\ \Phi_{[\mathbf{F}, \mathbf{G}]} & \Phi_{[\mathbf{F}, \mathbf{F}]} \end{bmatrix}^\dagger \begin{bmatrix} \mathbf{I}_r \\ 0 \end{bmatrix} \mathbf{G} =$$

$$[\mathbf{A}\mathbf{P}_{\mathbf{F}^\perp}][\mathbf{G}\mathbf{P}_{\mathbf{F}^\perp}]^\dagger \mathbf{G} = [\mathbf{A}|\mathbf{F}^\perp][\mathbf{G}|\mathbf{F}^\perp]^\dagger \mathbf{G} \quad (4.129)$$

e os ângulos e direções principais a partir da decomposição em valores singulares de:

$$\Phi_{[\mathbf{F}, \mathbf{F}]}^{-1/2} \Phi_{[\mathbf{F}, \mathbf{G}]} \Phi_{[\mathbf{G}, \mathbf{G}]}^{-1/2} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (4.130)$$

Como:

a)

$$[\Re(\mathbf{F}^T) \triangleleft \Re(\mathbf{G}^T)] = \mathbf{U}^T \Phi_{[\mathbf{F}, \mathbf{F}]}^{-1/2} \mathbf{F} \quad (4.131)$$

b)

$$[ \Re(F^T) \triangleleft \Re(G^T) ] = \mathbf{V}^T \Phi_{[\mathbf{G}, \mathbf{G}]}^{-1/2} \mathbf{G} \quad (4.132)$$

c)

$$[ \Re(F^T) \triangleleft \Re(G^T) ] = \Sigma \quad (4.133)$$

Podemos usar a mesma notação para as operações geométricas determinística e estocástica porque, quando da implementação dos algoritmos o número de medidas sempre será finito ( $j \neq \infty$ ) e aproximaremos  $\Phi_{[\mathbf{F}, \mathbf{G}]}$  por:

$$\Phi_{[\mathbf{F}, \mathbf{G}]} \cong \frac{1}{j} \mathbf{F} \mathbf{G}^T \quad (4.134)$$

Portanto as duas definições levemente diferentes nos arcabouços determinístico e estocástico coincidem. Por exemplo, a projeção ortogonal:

$$\begin{aligned} \mathbf{F} \overline{P}_{\mathbf{G}} &= \mathbf{F} | \mathbf{G} = \Phi_{[\mathbf{F}, \mathbf{G}]} \Phi_{[\mathbf{G}, \mathbf{G}]}^{\dagger} \mathbf{G} \\ &= \left[ \frac{1}{j} \mathbf{F} \mathbf{G}^T \right] \left[ \frac{1}{j} \mathbf{G} \mathbf{G}^T \right]^{\dagger} \mathbf{G} = \mathbf{F} \mathbf{G}^T [\mathbf{G} \mathbf{G}^T]^{\dagger} \mathbf{G} \end{aligned} \quad (4.135)$$

que é idêntica a relação para o contexto determinístico. O contexto deverá esclarecer qual definição está subjacente, principalmente no Capítulo 5.

Contudo, para deduções teóricas em contexto estocástico sempre suporemos  $j \rightarrow \infty$ .





## Capítulo 5

# REALIZAÇÃO, OPERADOR DE HANKEL E FUNDAMENTOS PARA MODELAGEM DE DADOS NO ESPAÇO DE ESTADO

### 5.1 Introdução

Embora os sistemas dinâmicos sejam descritos ou caracterizados de muitas maneiras, um modo usual para especificar um sistema dinâmico é a sequência de resposta ao impulso unitário ou sequência de multiplicadores dinâmicos. Por exemplo, sabemos que a resposta de um sistema causal discreto linear estacionário monovariável: uma entrada - uma saída, para qualquer sequência de entrada a partir do estado inicial zero, resposta ao estado nulo, é descrita por:

$$\bar{\mathbf{y}}(z) = \bar{\mathbf{G}}(z)\bar{\mathbf{u}}(z) \quad (5.1)$$

onde a função de transferência discreta  $\bar{\mathbf{G}}(z) = \sum_{i=0}^{\infty} \mathbf{G}_i z^{-i}$  é a transformada  $z$  da resposta ao impulso,  $\bar{\mathbf{y}}(z) = \sum_{k=0}^{\infty} \mathbf{y}_k z^{-k}$  e  $\bar{\mathbf{u}}(z) = \sum_{k=0}^{\infty} \mathbf{u}_k z^{-k}$ .

Outra formulação para tal sistema é a descrita por:

$$\mathbf{y} = \mathbf{T}\mathbf{u} \quad (5.2)$$

onde  $\mathbf{T}$  é o operador de transferência do sistema, que mapeia um sinal de entrada  $\mathbf{u}$ , um vetor que representa uma sequência de entradas:

$$\mathbf{u} = \left[ \dots \quad \mathbf{u}_{-1} \quad \mathbf{u}_0 \quad \mathbf{u}_1 \quad \dots \right]^T$$

em um sinal de saída  $\mathbf{y}$ , um vetor que representa uma seqüência de saídas:

$$\mathbf{y} = \left[ \dots \quad \mathbf{y}_{-1} \quad \mathbf{y}_0 \quad \mathbf{y}_1 \quad \dots \right]^T$$

Ainda outra formulação muito importante para descrever um sistema discreto linear com dimensão finita invariante no tempo é a chamada representação de estado do seu modelo matemático:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix}, \quad \mathbf{x}_0; \quad k = 0, 1, 2 \dots \quad (5.3)$$

onde  $\mathbf{x}_k$  é o vetor de estado, de dimensão  $n$ ; supondo que o sistema tem  $m$  entradas e  $l$  saídas, então tem-se as matrizes do sistema com as seguintes dimensões:  $\mathbf{A} \in \Re^{n \times n}$ ,  $\mathbf{B} \in \Re^{n \times m}$ ,  $\mathbf{C} \in \Re^{l \times n}$ ,  $\mathbf{D} \in \Re^{l \times m}$  e estado inicial  $\mathbf{x}_0$ . O número inteiro  $n$  é a ordem do sistema. Na Figura 5.1 mostramos uma representação no espaço de estado para um instante particular de tempo  $k$ , onde ocorre a operação de mapeamento linear da entrada atual  $\mathbf{u}_k$  e do estado atual  $\mathbf{x}_k$  para a saída  $\mathbf{y}_k$  e o próximo estado  $\mathbf{x}_{k+1}$ .

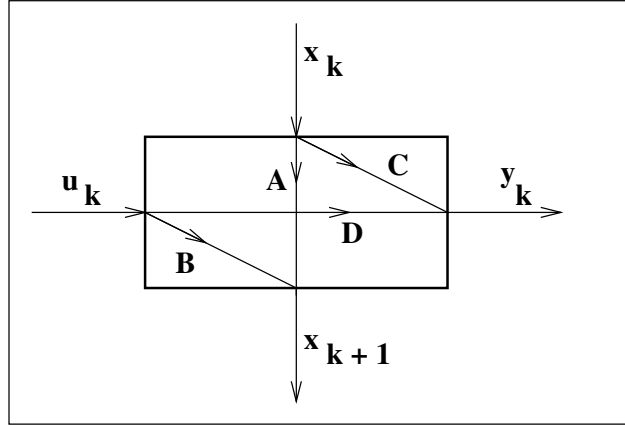


Figura 5.1: Representação no Espaço de Estado.

Para um sistema discreto linear estacionário multivariável, para qualquer seqüência vetorial de entrada,  $\mathbf{u}(k) \equiv \mathbf{u}_k$ , a resposta ao estado nulo pode ser descrita por:

$$\mathbf{y}(k) = \sum_{i=0}^{\infty} \mathbf{G}(i) \mathbf{u}(k-i) \equiv \sum_{i=0}^{\infty} \mathbf{G}_i \mathbf{u}_{k-i} \quad (5.4)$$

sendo  $\mathbf{y}(k) \equiv \mathbf{y}_k$  o vetor de saída no  $k$ -ésimo instante e  $\mathbf{G}(i) \equiv \mathbf{G}_i \in \Re^{l \times m}$  os coeficientes da função de transferência discreta matricial:

$$\overline{\mathbf{G}}(z) = \sum_{i=0}^{\infty} \mathbf{G}(i) \mathbf{z}^{-i}. \quad (5.5)$$

$\mathbf{G}(i)$  é a matriz resposta à sequência impulso unitário discreto ou meramente a resposta ao impulso discreto do sistema multivariável que rediscutiremos a seguir. Usaremos indiferentemente as notações  $\mathbf{y}_k$ ,  $\mathbf{y}(k)$  e  $\mathbf{y}(t_k)$  para o valor de  $\mathbf{y}$  no instante  $t_k$ ,  $k$  inteiro.

### 5.1.1 Resposta ao Impulso de Sistema Monovariável

Se temos um sistema discreto monovariável representado no espaço de estado e na sua entrada aplicamos um impulso discreto unitário

$$\mathbf{u}_k = \delta_k = \begin{cases} 1, & k = 0 \\ 0, & k \neq 0 \end{cases}$$

com o estado inicial sendo nulo,  $\mathbf{x}_0 = 0$ , a partir do instante em que ocorre o impulso tem-se:

- Instante 0 :

$$\begin{cases} \mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{u}_0 \\ \mathbf{y}_0 = \mathbf{C}\mathbf{x}_0 + \mathbf{D}\mathbf{u}_0 \end{cases} \text{ como } \mathbf{x}_0 = 0 \text{ e } \mathbf{u}_0 = 1 \implies \mathbf{y}_0 = \mathbf{D}\mathbf{u}_0 = \mathbf{D} = \mathbf{G}_0$$

- Instante 1

$$\begin{cases} \mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{u}_1 \\ \mathbf{y}_1 = \mathbf{C}\mathbf{x}_1 + \mathbf{D}\mathbf{u}_1 \end{cases} \text{ como } \mathbf{x}_1 = \mathbf{B} \text{ e } \mathbf{u}_1 = 0 \implies \mathbf{y}_1 = \mathbf{C}\mathbf{B}\mathbf{u}_0 = \mathbf{CB} = \mathbf{G}_1$$

- Instante 2

$$\begin{cases} \mathbf{x}_3 = \mathbf{A}\mathbf{x}_2 + \mathbf{B}\mathbf{u}_2 \\ \mathbf{y}_2 = \mathbf{C}\mathbf{x}_2 + \mathbf{D}\mathbf{u}_2 \end{cases} \text{ como } \mathbf{x}_2 = \mathbf{AB} \text{ e } \mathbf{u}_2 = 0 \implies \mathbf{y}_2 = \mathbf{CAB}\mathbf{u}_0 = \mathbf{CAB} = \mathbf{G}_2$$

- Instante 3

$$\begin{cases} \mathbf{x}_4 = \mathbf{A}\mathbf{x}_3 + \mathbf{B}\mathbf{u}_3 \\ \mathbf{y}_3 = \mathbf{C}\mathbf{x}_3 + \mathbf{D}\mathbf{u}_3 \end{cases} \text{ como } \mathbf{x}_3 = \mathbf{A}^2\mathbf{B} \text{ e } \mathbf{u}_3 = 0 \implies \mathbf{y}_3 = \mathbf{CA}^2\mathbf{B}\mathbf{u}_0 = \mathbf{CA}^2\mathbf{B} = \mathbf{G}_3$$

e assim sucessivamente.

Portanto a resposta ao impulso de tal sistema tem a forma:

$$\mathbf{G}_k = \begin{cases} \mathbf{D}, & k = 0 \\ \mathbf{CA}^{k-1}\mathbf{B}, & k \geq 1 \end{cases} \quad (5.6)$$

As matrizes  $\mathbf{CA}^{k-1}\mathbf{B}$  da expressão (5.6) constituem os parâmetros de Markov e expressam os efeitos das entradas passadas na saída atual  $\mathbf{y}_k$ .

### 5.1.2 Resposta ao Impulso de Sistema Multivariável

Para um sistema discreto multivariável com  $m$  entradas pode-se pensar a entrada  $\mathbf{u}_k$  como constituída de uma combinação linear de impulsos  $\delta_i = [\dots 0 \ 0 \ 1 \ 0 \ 0 \dots]^T$  com 1 apenas na  $i$ -ésima componente, um para cada  $i$ -ésima componente da entrada, ponderado por:

$$\mathbf{u}_{ki} = \begin{cases} 1 & \text{para } k = i \\ 0 & \text{para } k \neq i \end{cases} \quad (5.7)$$

$$\mathbf{u}_k = \sum_{i=1}^m \mathbf{u}_{ki} \delta_i = \mathbf{u}_{k1} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \mathbf{u}_{k2} \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \mathbf{u}_{k3} \begin{bmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{bmatrix} + \dots + \mathbf{u}_{km} \begin{bmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{bmatrix} \quad (5.8)$$

A resposta ao impulso do sistema multivariável é então a combinação linear das respostas aos impulsos  $\delta_i$  ponderadas por  $\mathbf{u}_{ki}$ .

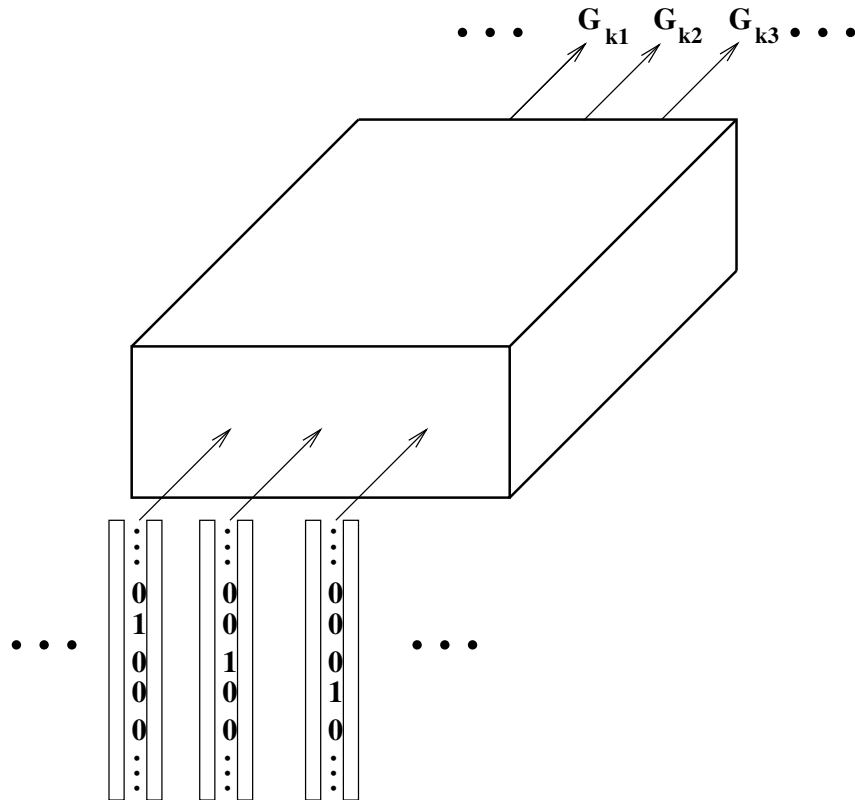


Figura 5.2: Um Hipotético Impulso Multivariável Aplicado a Um Sistema.

Na Figura 5.2 representamos a obtenção da resposta ao impulso de um sistema multivariável e podemos fazer uma interpretação em que a resposta ao impulso de um sistema multivariável é obtida fazendo-se a aplicação em paralelo de impulsos unitários  $\delta_i$ ,  $i = 1, \dots, m$  em cada  $i$ -ésima entrada, e determinando-se a respectiva resposta ao impulso,  $\mathbf{G}_i$ . A matriz resposta ao impulso unitário no instante  $k$  será:

$$\mathbf{G}(k) \stackrel{\text{def}}{=} [\mathbf{G}_{k1} \ \mathbf{G}_{k2} \ \dots \ \mathbf{G}_{ki} \ \dots \ \mathbf{G}_{km}] \quad (5.9)$$

onde  $\mathbf{G}_{ki}$  é a resposta do sistema inicialmente em repouso (relaxado) à entrada impulsiva  $\delta_i$ , onde  $\delta_i$  é um vetor  $m \times 1$ , com 1 na  $i$ -ésima coordenada e zeros nas demais, como mostrado na Figura 5.2.

Para exprimir  $\mathbf{G}(k)$  em termos das matrizes  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  e  $\mathbf{D}$  do sistema, considere a resposta do sistema à entrada  $\mathbf{u}_k$ , dada pela equação (5.8), supondo  $\mathbf{x}(0) = 0$ . Como o sistema é causal - não responde antes de ser excitado - por hipótese, segue-se que a resposta para  $k < 0$  é nula; para  $k = 0$ , a resposta é:

$$\mathbf{G}_{0i} = \mathbf{D}\mathbf{u}_{0i}\delta_i.$$

Para  $k = 1$  teremos:

$$\mathbf{x}_1 = \mathbf{A}\mathbf{x}_0 + \mathbf{B}\mathbf{u}_{1i}\delta_i = \mathbf{B}\mathbf{u}_{1i}\delta_i = \mathbf{B}\delta_i$$

implicando que a resposta é  $\mathbf{G}_{1i} = \mathbf{C}\mathbf{B}\delta_i$ . Para  $k = 2$ ,  $\mathbf{x}_2 = \mathbf{A}\mathbf{x}_1 + \mathbf{B}\mathbf{u}_{2i}\delta_i = \mathbf{A}\mathbf{B}\delta_i$  e desta forma  $\mathbf{G}_{2i} = \mathbf{C}\mathbf{A}\mathbf{B}\delta_i$ . De forma geral:

$$\mathbf{G}_{ki} = \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}\delta_i \quad \text{para } k \geq 1$$

Portanto a matriz resposta ao impulso tem a estrutura dada por:

$$\mathbf{G}(k) = \begin{cases} 0, & k < 0 \\ \mathbf{D}, & k = 0 \\ \mathbf{C}\mathbf{A}^{k-1}\mathbf{B}, & k \geq 1 \end{cases} \quad (5.10)$$

Por exemplo, considere um sistema multivariado com 3 saídas e 2 entradas. Para a saída:  $\mathbf{y}_k = [\mathbf{y}_{k1} \ \mathbf{y}_{k2} \ \mathbf{y}_{k3}]^T$  e a entrada:  $\mathbf{u}_k = [\mathbf{u}_{k1} \ \mathbf{u}_{k2}]^T$ , no momento inicial,  $k = 0$  tem-se:

$$\mathbf{y}_0 = \begin{bmatrix} \mathbf{y}_{01} \\ \mathbf{y}_{02} \\ \mathbf{y}_{03} \end{bmatrix} = \mathbf{D} \begin{bmatrix} \mathbf{u}_{01} \\ \mathbf{u}_{02} \end{bmatrix}$$

$$\begin{bmatrix} \mathbf{y}_{01} \\ \mathbf{y}_{02} \\ \mathbf{y}_{03} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{11} & \mathbf{D}_{21} \\ \mathbf{D}_{12} & \mathbf{D}_{22} \\ \mathbf{D}_{13} & \mathbf{D}_{23} \end{bmatrix} \begin{bmatrix} \mathbf{u}_{01} \\ \mathbf{u}_{02} \end{bmatrix}$$

supondo que o impulso aplicado seja:  $\mathbf{u}_k = [\mathbf{u}_{k1} \ \mathbf{u}_{k2}]^T = [1 \ 0]^T$

$$\begin{bmatrix} \mathbf{y}_{01} \\ \mathbf{y}_{02} \\ \mathbf{y}_{03} \end{bmatrix} = \begin{bmatrix} \mathbf{D}_{11} \\ \mathbf{D}_{12} \\ \mathbf{D}_{13} \end{bmatrix}$$

note que o impulso  $[1 \ 0]$  gera a primeira coluna da matriz  $\mathbf{D}$  no instante  $k = 0$ .

## 5.2 Realização Mínima da Resposta ao Impulso

As equações de saída do sistema linear, discreto, apresentadas nas equações (5.10), considerando o estado inicial  $\mathbf{x}_0 = 0$ , podem ser reescritas como:

$$\mathbf{y}_k = \sum_{j=k_0}^k \mathbf{G}(k, j) \mathbf{u}(j), \quad \text{para } k \geq k_0 \quad (5.11)$$

sendo a matriz  $\mathbf{G}$  dada por:

$$\mathbf{G}(k, j) \equiv \mathbf{G}(k - j) = \begin{cases} \mathbf{D} & , \ k = j \\ \mathbf{C} \mathbf{A}^{(k-j-1)} \mathbf{B}, & k \geq j + 1 \end{cases} \quad (5.12)$$

Quando existe uma equação de estado correspondente a uma dada  $\mathbf{G}(k, j)$ , existem muitas, pois uma transformação de variável de estado não altera a resposta ao impulso. Também existem equações de estado de dimensões diferentes dando uma resposta ao impulso especificada. Por exemplo, se as equações (5.3) com  $\mathbf{D} = 0$  correspondem ao comportamento entrada - saída em (5.12), então uma equação de estado da forma:

$$\begin{bmatrix} \mathbf{x}(k+1) \\ \mathbf{z}(k+1) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & 0 \\ 0 & \mathbf{F} \end{bmatrix} \begin{bmatrix} \mathbf{x}(k) \\ \mathbf{z}(k) \end{bmatrix} + \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix} \mathbf{u}(k)$$

$$\mathbf{y}(k) = \begin{bmatrix} \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}(k) \\ \mathbf{z}(k) \end{bmatrix}$$

produz o mesmo comportamento entrada - saída conforme a Figura 5.3 evidencia e / ou porque a matriz de transição,  $\Phi(k, j+1) = \mathbf{A}^{(k-j-1)}$ , para tal sistema é bloco diagonal, pois:

$$\begin{bmatrix} \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \Phi_A(k, j+1) & 0 \\ 0 & \Phi_F(k, j+1) \end{bmatrix} \begin{bmatrix} \mathbf{B} \\ 0 \end{bmatrix} = \mathbf{C} \Phi_A(k, j+1) \mathbf{B}, \quad k \geq j+1$$

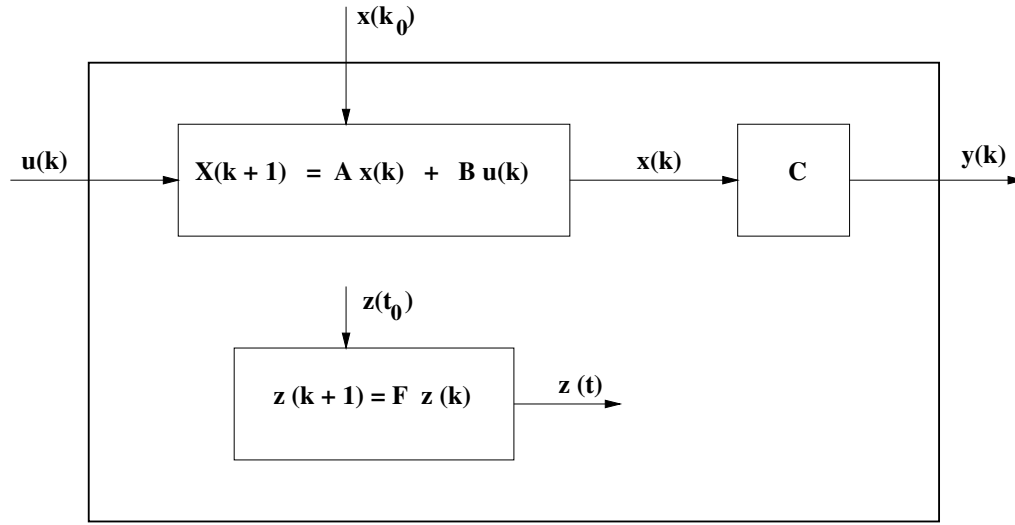


Figura 5.3: Múltiplas Realizações de Um Sistema.

O exemplo se refere ao fato de que novas variáveis de estado que são desconectadas da entrada, da saída ou de ambas, podem ser adicionadas a uma equação de estado sem modificar o comportamento entrada - saída associado. Este exemplo mostra que se uma equação linear de estado de dimensão  $n$  tem comportamento entrada - saída especificado por  $\mathbf{G}(k, j)$ , então para qualquer inteiro positivo  $q$  existem equações de estado de dimensão  $n + q$  que têm o mesmo comportamento entrada - saída: o mesmo operador de transferência, a mesma resposta ao impulso, a mesma matriz de transferência, os mesmos parâmetros de Markov, para o caso estacionário. Portanto nosso principal interesse teórico é considerar equações lineares de estado de dimensão mínima, que realizam uma  $\mathbf{G}(k, j)$  especificada.

Assim uma equação linear de estado de dimensão  $n$ :

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & 0 \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix} \quad (5.13)$$

é chamada uma realização da matriz resposta ao impulso unitário,  $\mathbf{G}(k, j)$  se  $\mathbf{G}(k, j) = \mathbf{C} \Phi_A(k, j+1) \mathbf{B}$ , para todo  $k, j$  tal que  $k \geq j+1$ .

Se uma realização (5.13) existe, então a resposta ao impulso é dita realizável, e (5.13) é chamada uma realização mínima se nenhuma realização de  $\mathbf{G}(k, j)$  com dimensão inferior a  $n$  existir.

### 5.3 Realização e Operador de Hankel

Portanto podemos dizer que o problema de realização é encontrar uma representação no espaço de estado, tal que para um dado operador  $\mathbf{T}$  de transferência tenhamos a seguinte relação:

$$\begin{bmatrix} \mathbf{x}_{k+1} \\ \mathbf{y}_k \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k \\ \mathbf{u}_k \end{bmatrix} \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{T}\mathbf{u} \quad (5.14)$$

onde  $\mathbf{y}$  e  $\mathbf{u}$  representam respectivamente seqüências de vetores de saída,  $\mathbf{y}_k$ , e de entrada,  $\mathbf{u}_k$ , tal que as representações destas equações sejam equivalentes. Existe um número infinito de realizações no espaço de estado para um dado sistema. O vetor de estado,  $\mathbf{x}_k$  poderá conter alguns estados que não são observados na saída ou que nunca são excitados pela entrada. Portanto concentraremos nossa atenção para as realizações mínimas no espaço de estado, isto é, modelos para os quais a dimensão do vetor de estado é mínima. Voltaremos a este ponto.

Para desenvolver expressões gerais que possibilitem relacionar entradas e saídas de um sistema qualquer a partir de um instante inicial,  $k$ , para sistemas discretos representados no espaço de estado por:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \end{cases} \quad (5.15)$$

observa-se que no instante seguinte tem-se:

$$\begin{cases} \mathbf{x}_{k+2} = \mathbf{A}\mathbf{x}_{k+1} + \mathbf{B}\mathbf{u}_{k+1} \\ \mathbf{y}_{k+1} = \mathbf{C}\mathbf{x}_{k+1} + \mathbf{D}\mathbf{u}_{k+1} \end{cases} \quad (5.16)$$

Substituindo a equação (5.15) em (5.16) tem-se:

$$\begin{cases} \mathbf{x}_{k+2} = \mathbf{A}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + \mathbf{B}\mathbf{u}_{k+1} \\ \mathbf{y}_{k+1} = \mathbf{C}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + \mathbf{D}\mathbf{u}_{k+1} \end{cases} \quad \Rightarrow \quad \begin{cases} \mathbf{x}_{k+2} = \mathbf{A}^2\mathbf{x}_k + \mathbf{A}\mathbf{B}\mathbf{u}_k + \mathbf{B}\mathbf{u}_{k+1} \\ \mathbf{y}_{k+1} = \mathbf{C}\mathbf{A}\mathbf{x}_k + \mathbf{C}\mathbf{B}\mathbf{u}_k + \mathbf{D}\mathbf{u}_{k+1} \end{cases} \quad (5.17)$$

No instante seguinte tem-se:

$$\begin{cases} \mathbf{x}_{k+3} = \mathbf{A}\mathbf{x}_{k+2} + \mathbf{B}\mathbf{u}_{k+2} \\ \mathbf{y}_{k+2} = \mathbf{C}\mathbf{x}_{k+2} + \mathbf{D}\mathbf{u}_{k+2} \end{cases} \quad (5.18)$$

que referenciando ao instante inicial resulta em:

$$\begin{cases} \mathbf{x}_{k+3} = \mathbf{A}^3\mathbf{x}_k + \mathbf{A}^2\mathbf{B}\mathbf{u}_k + \mathbf{A}\mathbf{B}\mathbf{u}_{k+1} + \mathbf{B}\mathbf{u}_{k+2} \\ \mathbf{y}_{k+2} = \mathbf{C}\mathbf{A}^2\mathbf{x}_k + \mathbf{C}\mathbf{A}\mathbf{B}\mathbf{u}_k + \mathbf{C}\mathbf{B}\mathbf{u}_{k+1} + \mathbf{D}\mathbf{u}_{k+2} \end{cases}$$

E assim sucessivamente.



A solução completa da equação (5.14) para um dado instante de tempo  $k \geq 0$  pode ser escrita como:

$$\mathbf{y}_k = \begin{cases} \mathbf{C}\mathbf{x}_0 + \mathbf{D}\mathbf{u}_0, & k = 0 \\ \mathbf{C}\mathbf{A}^k\mathbf{x}_0 + \sum_{j=0}^{k-1} \mathbf{C}\mathbf{A}^{k-j-1}\mathbf{B}\mathbf{u}_j + \mathbf{D}\mathbf{u}_k, & k \geq 1 \end{cases} \quad (5.19)$$

A equação de saída do sistema linear discreto, apresentada na equação (5.19), nos permite observar de outra maneira o problema de realização ou de identificação: a partir das equações (5.19) obtenha as equações (5.15).

Então para uma coleção de saídas,  $\mathbf{y}$ , relacionadas com uma coleção de entradas,  $\mathbf{u}$ , obtidas a partir do instante 0 até o instante  $k$  tem-se:

$$\mathbf{y} = \begin{bmatrix} \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \mathbf{y}_3 \\ \vdots \\ \mathbf{y}_k \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \mathbf{C}\mathbf{A}^3 \\ \vdots \\ \mathbf{C}\mathbf{A}^k \\ \vdots \end{bmatrix} \mathbf{x}_0 + \begin{bmatrix} \mathbf{D} & 0 & 0 & \cdots \\ \mathbf{C}\mathbf{B} & \mathbf{D} & 0 & \cdots \\ \mathbf{C}\mathbf{A}\mathbf{B} & \mathbf{C}\mathbf{B} & \mathbf{D} & \cdots \\ \mathbf{C}\mathbf{A}^2\mathbf{B} & \mathbf{C}\mathbf{A}\mathbf{B} & \mathbf{C}\mathbf{B} & \cdots \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{C}\mathbf{A}^k\mathbf{B} & \mathbf{C}\mathbf{A}^{k-1}\mathbf{B} & \cdots & \mathbf{D} \\ \vdots & \vdots & \ddots & \vdots \end{bmatrix} \mathbf{u} \quad (5.20)$$

onde  $\mathbf{u} = [\mathbf{u}_0 \ \mathbf{u}_1 \ \mathbf{u}_2 \ \mathbf{u}_3 \ \dots \ \mathbf{u}_k \ \dots]^T$ .

Pode-se associar à equação (5.20) um impulso unitário  $\mathbf{u}$  e considerando o estado inicial  $\mathbf{x}_0 = 0$ , teremos uma resposta ao impulso da seguinte forma:

$$\mathbf{G} = [\dots \ 0 \ 0 \ \mathbf{G}_0 \ \mathbf{G}_1 \ \mathbf{G}_2 \ \dots \ \mathbf{G}_i \ \dots]^T = \mathbf{T} [\dots \ 0 \ 0 \ 1 \ 0 \ 0 \ \dots]^T \quad (5.21)$$

sendo  $\mathbf{G}_i = \mathbf{C}\mathbf{A}^{i-1}\mathbf{B}$  e  $\mathbf{G}_0 = \mathbf{D}$ , com o operador  $\mathbf{T}$  dado por:

$$\mathbf{T} = \begin{bmatrix} \ddots & & \vdots & & \ddots \\ \cdots & \mathbf{G}_0 & 0 & 0 & \cdots \\ & \mathbf{G}_1 & \mathbf{G}_0 & 0 & \\ & \mathbf{G}_2 & \mathbf{G}_1 & \mathbf{G}_0 & \\ & \ddots & \vdots & & \ddots \end{bmatrix} \quad (5.22)$$

Na representação no espaço de estado  $\mathbf{G}$  é:

$$\mathbf{G} = [\dots \ 0 \ \mathbf{D} \ \mathbf{C}\mathbf{B} \ \mathbf{C}\mathbf{A}\mathbf{B} \ \mathbf{C}\mathbf{A}^2\mathbf{B} \ \dots]^T \quad (5.23)$$

A  $k$ -ésima coluna de  $\mathbf{T}$  contém a resposta ao impulso devido a um impulso no instante  $k$ ; e a sequência  $\mathbf{G}$  contém a informação sobre a resposta dinâmica do modelo. Tal relação baseia-se na linearidade do sistema.

Portanto, para um sistema monovariável tem-se:

$$\mathbf{y} = \begin{bmatrix} \mathbf{D} \\ \mathbf{CB} \\ \mathbf{CAB} \\ \mathbf{CA}^2\mathbf{B} \\ \vdots \\ \mathbf{CA}^k\mathbf{B} \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{D} & 0 & 0 & 0 & \dots \\ \mathbf{CB} & \mathbf{D} & 0 & & \dots \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & & \dots \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \dots \\ \vdots & \vdots & & & \\ \mathbf{CA}^k\mathbf{B} & \mathbf{CA}^{k-1}\mathbf{B} & \dots & \mathbf{D} & \dots \\ \vdots & \vdots & & \ddots & \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ \vdots \end{bmatrix}$$

donde

$$\mathbf{y} = \mathcal{T}[1 \ 0 \ 0 \ \dots]^T \quad (5.24)$$

onde  $\mathcal{T}$  é uma matriz Toeplitz.

Voltando ao problema inicial de realização: dado um operador de transferência  $\mathbf{T}$  como pode um modelo no espaço de estado que realize este operador de transferência ser determinado? A solução para o problema de realização no contexto da metodologia de subespaço exige um operador de Hankel o qual nos definiremos a seguir.

Suponhamos que aplicamos entradas somente até o instante  $k = -1$  (que chamamos de passado com respeito ao instante presente,  $k = 0$ ) e medimos as saídas para os instantes  $k \geq 0$  (que chamamos de futuro). Na Figura 5.4 apresentamos uma representação onde entradas são aplicadas até  $k = -1$  (passado) e saídas são medidas a partir de  $k = 0$  (futuro); produzindo informações sobre o estado a partir do instante  $k = 0$ . Observe que o estado  $\mathbf{x}_0$  acopla às medidas de entradas passadas às medidas de saídas futuras. De posse dos dados de entradas passadas e saídas futuras a representação no espaço de estado pode ser obtida.

Escrevendo  $\mathbf{y} = \mathbf{T}\mathbf{u}$  tem-se:

$$\begin{bmatrix} \vdots \\ \times \\ \times \\ \mathbf{y}_0 \\ \mathbf{y}_1 \\ \mathbf{y}_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \ddots & & & \vdots & & \ddots \\ \cdots & \times & \times & 0 & 0 & \cdots \\ & \times & \times & \times & 0 & \\ & & \mathbf{G}_3 & \mathbf{G}_2 & \mathbf{G}_1 & \times \\ \cdots & & & \mathbf{G}_3 & \mathbf{G}_2 & \cdots \\ & & & \mathbf{G}_4 & \mathbf{G}_3 & \\ & & & \vdots & & \ddots \end{bmatrix} \begin{bmatrix} \vdots \\ \mathbf{u}_{-2} \\ \mathbf{u}_{-1} \\ 0 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \quad (5.25)$$

ou seja

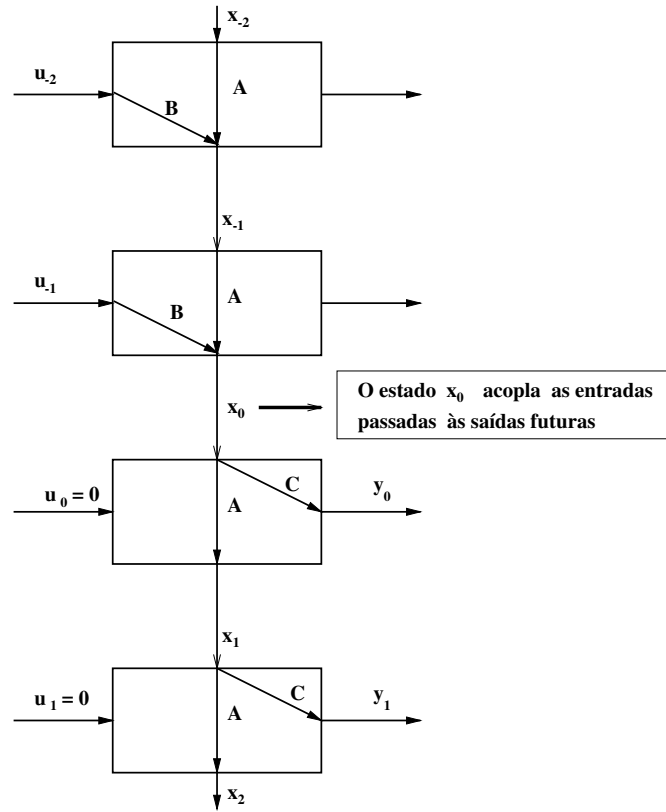


Figura 5.4: Representação de dinâmica discreta no espaço de estado.

$$\begin{cases} y_0 = \cdots + \mathbf{G}_3 \mathbf{u}_{-3} + \mathbf{G}_2 \mathbf{u}_{-2} + \mathbf{G}_1 \mathbf{u}_{-1} + 0 + 0 + \cdots \\ y_1 = \cdots + \mathbf{G}_4 \mathbf{u}_{-3} + \mathbf{G}_3 \mathbf{u}_{-2} + \mathbf{G}_2 \mathbf{u}_{-1} + 0 + 0 + \cdots \\ y_2 = \cdots + \mathbf{G}_5 \mathbf{u}_{-3} + \mathbf{G}_4 \mathbf{u}_{-2} + \mathbf{G}_3 \mathbf{u}_{-1} + 0 + 0 + \cdots \\ \vdots \end{cases} \quad (5.26)$$

e assim sucessivamente; reordenando a série de equações de (5.26) e desprezando os termos nulos tem-se:

$$\begin{cases} y_0 = \mathbf{G}_1 \mathbf{u}_{-1} + \mathbf{G}_2 \mathbf{u}_{-2} + \mathbf{G}_3 \mathbf{u}_{-3} + \cdots \\ y_1 = \mathbf{G}_2 \mathbf{u}_{-1} + \mathbf{G}_3 \mathbf{u}_{-2} + \mathbf{G}_4 \mathbf{u}_{-3} + \cdots \\ y_2 = \mathbf{G}_3 \mathbf{u}_{-1} + \mathbf{G}_4 \mathbf{u}_{-2} + \mathbf{G}_5 \mathbf{u}_{-3} + \cdots \\ \vdots \end{cases} \quad (5.27)$$

donde

$$\mathbf{y}_k = \sum_{i=0}^{\infty} \mathbf{G}_i \mathbf{u}_{k-i} \quad (5.28)$$

Da equação (5.25) e da série de equações (5.27) concluímos que podemos nos restringir à utilização de apenas uma parte de  $\mathbf{T}$ ; esta parte mapeia uma sequência finita e podemos através das definições de sequências de entradas passadas e de saídas futuras, respectivamente:

$$\mathbf{u}^- \triangleq [\mathbf{u}_{-1} \quad \mathbf{u}_{-2} \quad \mathbf{u}_{-3} \quad \dots]^T$$

e

$$\mathbf{y}^+ \triangleq [\mathbf{y}_0 \quad \mathbf{y}_1 \quad \mathbf{y}_2 \quad \dots]^T$$

reescrever tal parte da matriz  $\mathbf{T}$ ; desta maneira podemos reescrever a equação (5.25) como  $\mathbf{y}^+ = \mathbf{H}\mathbf{u}^-$  com  $\mathbf{H}$  definido como:

$$\mathbf{H} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \mathbf{G}_3 & \dots \\ \mathbf{G}_2 & \mathbf{G}_3 & \mathbf{G}_4 & \\ \mathbf{G}_3 & \mathbf{G}_4 & \ddots & \\ \vdots & & & \end{bmatrix} \quad (5.29)$$

$\mathbf{H}$  tem suas antidiagonais constantes e recebe o nome de operador de Hankel.

Para a submatriz bloco  $L \times J$  superior à esquerda da matriz de Hankel, resultante de um truncamento finito da matriz infinita de Hankel tem-se:

$$\mathbf{H}_{J,L} = \begin{bmatrix} \mathbf{G}_1 & \mathbf{G}_2 & \dots & \mathbf{G}_L \\ \mathbf{G}_2 & \mathbf{G}_3 & \dots & \mathbf{G}_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ & \dots & \mathbf{G}_i & \dots \\ \vdots & \vdots & & \ddots & \vdots \\ \mathbf{G}_J & \mathbf{G}_{J+1} & \dots & \mathbf{G}_{J+L-1} \end{bmatrix} \quad (5.30)$$

## 5.4 Gênese de Operador de Hankel

O operador de Hankel,  $\mathbf{H}$ , é muito importante e útil em vários aspectos de análise e de projeto de sistemas. Sua conexão com problemas baseados em dados amostrados é bem conhecida e estabelecida; recentemente ele tem sido estudado e usado para novas aplicações. Nestas estão incluídos: projetos de filtros digitais, realizações markovianas, modelamento de séries temporais multivariadas, redução de modelos, realizações balanceadas entre outras. Em cada um destes casos, a obtenção de função de transferência, impedância ou

realização segundo algum critério de otimalidade é reduzida à determinação de alguns valores e vetores singulares de uma matriz de Hankel semi-infinita.

A seguir mostraremos, por simplicidade para sistemas monovariáveis, que esta estrutura de Hankel pode ser obtida a partir das equações (5.15), (5.16) e (5.18). Desta seqüência de equações, as saídas são dadas por:

$$\begin{cases} \mathbf{y}_{k+1} &= \mathbf{C}\mathbf{x}_{k+1} + \mathbf{D}\mathbf{u}_{k+1} \\ \mathbf{y}_{k+2} &= \mathbf{C}\mathbf{x}_{k+2} + \mathbf{D}\mathbf{u}_{k+2} \\ \mathbf{y}_{k+3} &= \mathbf{C}\mathbf{x}_{k+3} + \mathbf{D}\mathbf{u}_{k+3} \\ \vdots & \end{cases} \quad (5.31)$$

como apenas a entrada  $\mathbf{u}_0 = 1$  e as demais são nulas, as séries de equações apresentadas em (5.31) podem ser reescritas como:

$$\begin{cases} \mathbf{y}_{k+1} &= \mathbf{C}\mathbf{x}_{k+1} \\ \mathbf{y}_{k+2} &= \mathbf{C}\mathbf{x}_{k+2} \\ \mathbf{y}_{k+3} &= \mathbf{C}\mathbf{x}_{k+3} \\ \vdots & \end{cases} \quad (5.32)$$

Retornando à série de equações (5.15), (5.16) e (5.18) podemos substituir os estados  $\mathbf{x}_{k+2}, \mathbf{x}_{k+3} \dots$  nas equações (5.32):

$$\begin{cases} \mathbf{y}_{k+1} &= \mathbf{C}\mathbf{x}_{k+1} \\ \mathbf{y}_{k+2} &= \mathbf{C}\mathbf{x}_{k+2} = \mathbf{C}(\mathbf{A}\mathbf{x}_{k+1} + \mathbf{B}\mathbf{u}_{k+1}) \\ \mathbf{y}_{k+3} &= \mathbf{C}\mathbf{x}_{k+3} = \mathbf{C}(\mathbf{A}\mathbf{x}_{k+2}) = \mathbf{C}\mathbf{A}(\mathbf{A}\mathbf{x}_{k+1} + \mathbf{B}\mathbf{u}_{k+1}) \\ \vdots & \end{cases} \quad (5.33)$$

como tais entradas são nulas, a expressão (5.33) é simplificada para:

$$\begin{cases} \mathbf{y}_{k+1} &= \mathbf{C}\mathbf{x}_{k+1} \\ \mathbf{y}_{k+2} &= \mathbf{C}\mathbf{A}\mathbf{x}_{k+1} \\ \mathbf{y}_{k+3} &= \mathbf{C}\mathbf{A}^2\mathbf{x}_{k+1} \\ \vdots & \end{cases} \quad (5.34)$$

então

$$\begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \mathbf{y}_{k+3} \\ \vdots \end{bmatrix} = \mathcal{O}\mathbf{x}_{k+1} \quad (5.35)$$

sendo,  $\mathcal{O}$  é a matriz de observabilidade estendida dada por:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \vdots \end{bmatrix} \quad (5.36)$$

A partir da equação (5.15) podemos reescrever o vetor de estado  $\mathbf{x}_{k+1}$  como:

$$\begin{aligned}
 \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\
 \mathbf{x}_{k+1} &= \mathbf{A}(\mathbf{A}\mathbf{x}_{k-1} + \mathbf{B}\mathbf{u}_{k-1}) + \mathbf{B}\mathbf{u}_k \\
 \mathbf{x}_{k+1} &= \mathbf{A}^2\mathbf{x}_{k-1} + \mathbf{A}\mathbf{B}\mathbf{u}_{k-1} + \mathbf{B}\mathbf{u}_k \\
 \mathbf{x}_{k+1} &= \mathbf{A}^2(\mathbf{A}\mathbf{x}_{k-2} + \mathbf{B}\mathbf{u}_{k-2}) + \mathbf{A}\mathbf{B}\mathbf{u}_{k-1} + \mathbf{B}\mathbf{u}_k \\
 \mathbf{x}_{k+1} &= \mathbf{A}^3\mathbf{x}_{k-2} + \mathbf{A}^2\mathbf{B}\mathbf{u}_{k-2} + \mathbf{A}\mathbf{B}\mathbf{u}_{k-1} + \mathbf{B}\mathbf{u}_k \\
 &\vdots
 \end{aligned} \tag{5.37}$$

logo

$$\mathbf{x}_{k+1} = \mathbf{A}^{(\dots)}\mathbf{x}_0 + \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \mathbf{u}_{k-2} \\ \vdots \end{bmatrix} \tag{5.38}$$

O vetor de estado  $\mathbf{x}_{k+1}$  pode ser escrito em termos da matriz de atingibilidade estendida  $\mathcal{C}$ :

$$\mathcal{C} = \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots \end{bmatrix} \tag{5.39}$$

Considerando o vetor de estado inicial  $\mathbf{x}_0 = 0$  e substituindo a equação (5.38) na equação (5.35) vemos que as observações futuras são expressas em função das entradas atuais e passadas através de uma matriz  $\mathbf{H}$  definida pelo produto das matrizes  $\mathcal{C}$  e  $\mathcal{O}$ :

$$\begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \mathbf{y}_{k+3} \\ \vdots \end{bmatrix} = \mathcal{O}\mathcal{C} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \mathbf{u}_{k-2} \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{C}\mathbf{B} & \mathbf{C}\mathbf{A}\mathbf{B} & \mathbf{C}\mathbf{A}^2\mathbf{B} & \dots \\ \mathbf{C}\mathbf{A}\mathbf{B} & \mathbf{C}\mathbf{A}^2\mathbf{B} & \mathbf{C}\mathbf{A}^3\mathbf{B} & \dots \\ \mathbf{C}\mathbf{A}^2\mathbf{B} & \mathbf{C}\mathbf{A}^3\mathbf{B} & \mathbf{C}\mathbf{A}^4\mathbf{B} & \dots \\ \vdots & & & \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \mathbf{u}_{k-2} \\ \vdots \end{bmatrix} \tag{5.40}$$

$$\begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \mathbf{y}_{k+3} \\ \vdots \end{bmatrix} = \mathbf{H} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k-1} \\ \mathbf{u}_{k-2} \\ \vdots \end{bmatrix} \tag{5.41}$$

A matriz de Hankel expressa todas as saídas futuras em termos das entradas atual e passadas; desta maneira podemos reescrever (5.41) como:

$$\mathbf{y}_{k+1}^+ = \mathbf{H}\mathbf{u}_k^- \tag{5.42}$$

Os elementos da matriz  $\mathbf{H}$  são os parâmetros de Markov. O operador de Hankel tem importantes propriedades para modelagem de dados no espaço de estado.

## 5.5 Propriedades da Matriz de Hankel

**Teorema 5.1** *A matriz de Hankel admite as seguintes fatorações:*

$$\begin{aligned} \mathbf{H} = \mathcal{O}\mathbf{C} &= \begin{bmatrix} \mathbf{C} \\ \mathbf{C}\mathbf{A} \\ \mathbf{C}\mathbf{A}^2 \\ \vdots \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots \end{bmatrix} \\ &= \hat{\mathcal{O}}\hat{\mathcal{C}} = \begin{bmatrix} \hat{\mathbf{C}} \\ \hat{\mathbf{C}}\hat{\mathbf{A}} \\ \hat{\mathbf{C}}\hat{\mathbf{A}}^2 \\ \vdots \end{bmatrix} \begin{bmatrix} \hat{\mathbf{B}} & \hat{\mathbf{A}}\hat{\mathbf{B}} & \hat{\mathbf{A}}^2\hat{\mathbf{B}} & \dots \end{bmatrix} \end{aligned} \quad (5.43)$$

sendo  $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$  e  $\{\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}\}$  duas realizações mínimas para a mesma função de transferência.

A transformação de similaridade  $\mathbf{S}$ , tal que:

$$\hat{\mathbf{A}} = \mathbf{S}^{-1}\mathbf{A}\mathbf{S}, \quad \hat{\mathbf{B}} = \mathbf{S}^{-1}\mathbf{B}, \quad \hat{\mathbf{C}} = \mathbf{C}\mathbf{S}, \quad \text{e} \quad \hat{\mathbf{D}} = \mathbf{D}$$

é única e pode ser descrita por:

$$\mathbf{S} = \mathcal{C}\hat{\mathcal{C}}^T(\hat{\mathcal{C}}\hat{\mathcal{C}}^T)^{-1} = \mathcal{C}\hat{\mathcal{C}}^\dagger \quad (5.44)$$

ou por:

$$\mathbf{S}^{-1} = (\hat{\mathcal{O}}^T\hat{\mathcal{O}})^{-1}\hat{\mathcal{O}}^T\mathcal{O} = \hat{\mathcal{O}}^\dagger\mathcal{O} \quad (5.45)$$

sendo  $\mathcal{O}$ ,  $\mathcal{C}$  e  $\hat{\mathcal{O}}$ ,  $\hat{\mathcal{C}}$  respectivamente as matrizes de observabilidade e atingibilidade das referidas realizações e são dadas por:

$$\hat{\mathcal{C}} = \hat{\mathcal{O}}^\dagger\mathcal{O}\mathcal{C} = \hat{\mathcal{O}}^\dagger\mathbf{H} \quad (5.46)$$

$$\hat{\mathcal{O}} = \mathcal{O}\mathcal{C}\hat{\mathcal{C}}^\dagger = \mathbf{H}\hat{\mathcal{C}}^\dagger \quad (5.47)$$

**Prova:**

Como as realizações são mínimas, elas são observáveis e atingíveis e desta forma:

$$\rho(\mathcal{O}) = \rho(\hat{\mathcal{O}}) = \rho(\mathcal{C}) = \rho(\hat{\mathcal{C}}) = n$$

e  $\mathcal{O}^T\mathcal{O}$ ,  $\hat{\mathcal{O}}^T\hat{\mathcal{O}}$ ,  $\mathcal{C}\mathcal{C}^T$  e  $\hat{\mathcal{C}}\hat{\mathcal{C}}^T$  são não singulares. Para demonstrar (5.43) basta observar que:

$$\begin{aligned}\hat{\mathcal{O}}\hat{\mathcal{C}} &= \begin{bmatrix} \mathbf{CS} \\ \mathbf{CSS}^{-1}\mathbf{AS} \\ \mathbf{CS}(\mathbf{S}^{-1}\mathbf{AS})^2 \\ \vdots \end{bmatrix} \begin{bmatrix} \mathbf{S}^{-1}\mathbf{B} & \mathbf{S}^{-1}\mathbf{ASS}^{-1}\mathbf{B} & (\mathbf{S}^{-1}\mathbf{AS})^2\mathbf{B} & \dots \end{bmatrix} = \\ &= \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \end{bmatrix} \begin{bmatrix} \mathbf{B} & \mathbf{AB} & \mathbf{A}^2\mathbf{B} & \dots \end{bmatrix} = \mathcal{O}\mathcal{C} = \mathbf{H}\end{aligned}$$

Da igualdade acima e porque  $\hat{\mathcal{O}}^T\hat{\mathcal{O}}$  é não singular podemos escrever  $\hat{\mathcal{C}}$  utilizando a pseudo inversa de  $\hat{\mathcal{O}}$  como:

$$\hat{\mathcal{C}} = (\hat{\mathcal{O}}^T\hat{\mathcal{O}})^{-1}\hat{\mathcal{O}}^T\mathcal{O}\mathcal{C} = \hat{\mathcal{O}}^\dagger\mathcal{O}\mathcal{C} \quad (5.48)$$

Analogamente tem-se:

$$\hat{\mathcal{O}} = \mathcal{O}\mathcal{C}\hat{\mathcal{C}}^T(\hat{\mathcal{C}}\hat{\mathcal{C}}^T)^{-1} = \mathcal{O}\mathcal{C}\hat{\mathcal{C}}^\dagger \quad (5.49)$$

Como  $\mathcal{C}$  e  $\hat{\mathcal{C}}$  têm posto completo,  $\mathbf{S}$  é regular e como:

$$\hat{\mathcal{C}} = \begin{bmatrix} \hat{\mathbf{B}} & \hat{\mathbf{A}}\hat{\mathbf{B}} & \hat{\mathbf{A}}^2\hat{\mathbf{B}} & \dots \end{bmatrix} = \begin{bmatrix} \mathbf{S}^{-1}\mathbf{B} & \mathbf{S}^{-1}\mathbf{ASS}^{-1}\mathbf{B} & (\mathbf{S}^{-1}\mathbf{AS})^2\mathbf{B} & \dots \end{bmatrix} = \mathbf{S}^{-1}\mathcal{C} \quad (5.50)$$

e

$$\hat{\mathcal{O}} = \begin{bmatrix} \hat{\mathbf{C}} \\ \hat{\mathbf{C}}\hat{\mathbf{A}} \\ \hat{\mathbf{C}}\hat{\mathbf{A}}^2 \\ \vdots \end{bmatrix} = \begin{bmatrix} \mathbf{CS} \\ \mathbf{CSS}^{-1}\mathbf{AS} \\ \mathbf{CS}(\mathbf{S}^{-1}\mathbf{AS})^2 \\ \vdots \end{bmatrix} = \mathcal{O}\mathbf{S} \quad (5.51)$$

das equações (5.46) e (5.49) pode-se escrever:

$$\hat{\mathcal{O}}\hat{\mathcal{C}} = \mathcal{O}\mathcal{C}\hat{\mathcal{C}}^\dagger\hat{\mathcal{O}}^\dagger\mathcal{O}\mathcal{C} = \mathcal{O}\mathcal{C} \quad (5.52)$$

e das equações (5.47) e (5.50) tem-se:

$$\hat{\mathcal{O}}\hat{\mathcal{C}} = \mathcal{O}\mathbf{S}\mathbf{S}^{-1}\mathcal{C} = \mathcal{O}\mathcal{C} \quad (5.53)$$

donde

$$\mathbf{S} = \mathcal{C}\hat{\mathcal{C}}^\dagger$$

e

$$\mathbf{S}^{-1} = \hat{\mathcal{O}}^\dagger\mathcal{O}$$



Suponhamos distintas  $\mathbf{S}$  e  $\hat{\mathbf{S}}$  que levassem  $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$  à  $\{\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}\}$  então teríamos:

$$\hat{\mathcal{O}} = \mathcal{O}\mathbf{S}$$

e

$$\hat{\mathcal{O}} = \mathcal{O}\hat{\mathbf{S}}$$

donde

$$\mathcal{O}(\mathbf{S} - \hat{\mathbf{S}}) = 0$$

e portanto

$$\mathbf{S} = \hat{\mathbf{S}}$$

donde  $\mathbf{S}$  é única.

**C.Q.D.**

**Corolário 5.1** *Os parâmetros de Markov das realizações mínimas  $\{\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}\}$  e  $\{\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}, \hat{\mathbf{D}}\}$  de uma função de transferência são idênticos.*

**Prova:** É imediata.

**C.Q.D.**

**Corolário 5.2** *A matriz  $\mathbf{H}$  tem posto igual a ordem mínima,  $n$  do sistema.*

**Prova:**

Das equações (5.36) , (5.39) e (5.42) tem-se:

$$\begin{aligned} \mathbf{x}_0 &= \mathcal{C}\mathbf{u}_k^- \\ \mathbf{y}_{k+1}^+ &= \mathcal{O}\mathbf{x}_0 \end{aligned} \quad (5.54)$$

Como a matriz  $\mathbf{H} = \mathcal{O}\mathcal{C}$  e para a realização mínima as matrizes de observabilidade e atingibilidade têm posto completo igual à dimensão do sistema,  $n$ ,  $\mathbf{H}$ : produto de duas matrizes de posto  $n$  também terá posto  $n$ .

**C.Q.D.**

**Corolário 5.3** *A matriz  $\mathbf{H}$  tem estrutura invariante aos deslocamentos para cima e para à esquerda abaixo definidos.*

**Prova:**

A estrutura de uma matriz qualquer é dita ser invariante aos deslocamentos à esquerda se o espaço coluna do operador deslocado estiver contido no espaço coluna do operador original. Definindo a matriz de Hankel deslocada  $\vec{\mathbf{H}}$  como:

$$\vec{\mathbf{H}} = \begin{bmatrix} \mathbf{G}_2 & \mathbf{G}_3 & \mathbf{G}_4 & \\ \mathbf{G}_3 & \mathbf{G}_4 & \mathbf{G}_5 & \\ \mathbf{G}_4 & \mathbf{G}_5 & \ddots & \dots \\ \vdots & & & \end{bmatrix} \quad (5.55)$$

vê-se imediatamente que ela satisfaz às relações:

$$\vec{\mathbf{H}} = \vec{\mathcal{O}} \mathcal{C} = \mathcal{O} \mathbf{A} \mathcal{C} \quad (5.56)$$

e

$$\overleftarrow{\mathbf{H}} = \mathcal{O} \overleftarrow{\mathcal{C}} = \mathcal{O} \mathbf{A} \mathcal{C} \quad (5.57)$$

onde os símbolos  $\rightarrow$  e  $\leftarrow$  representam que a primeira linha (deslocamento para cima) e a primeira coluna (deslocamento à esquerda), respectivamente, do operador original  $\mathbf{H}$ , por exemplo, foram eliminadas.

Para as matrizes  $\mathcal{O}$  e  $\mathcal{C}$  tem-se que os deslocamentos para cima ou para a esquerda são equivalentes a multiplicações pela matriz  $\mathbf{A}$  no centro da fatoração respectiva.

A matriz  $\mathbf{H}$  tem esta propriedade porque seu espaço coluna contém todas as saídas do sistema a partir do instante  $k = 0$  devidas às entradas até  $k = -1$ ; o espaço coluna de  $\vec{\mathbf{H}}$  contém todas as saídas possíveis dos sistema a partir de  $k = 1$ , devido as entradas até  $k = -1$ .

**C.Q.D.**

- A fatoração de  $\mathbf{H}$  pode ser obtida através da decomposição em valores singulares:

$$\mathbf{H} = \mathbf{U} \mathbf{\Sigma} \mathbf{V}^T = \begin{bmatrix} \mathbf{U}_1 & \mathbf{U}_2 \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{11} & 0 \\ 0 & \mathbf{\Sigma}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_2 \end{bmatrix}^T \quad (5.58)$$

donde  $\mathcal{O} = (\mathbf{U}_1 \mathbf{\Sigma}_{11}^{1/2}) \mathbf{S}^{-1}$  e  $\mathcal{C} = \mathbf{S} (\mathbf{\Sigma}_{11}^{1/2} \mathbf{V}_1^T)$  para alguma transformação  $\mathbf{S}$ .

- Como as matrizes  $\mathcal{O}$  e  $\mathcal{C}$  têm posto completo o sistema de matrizes  $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$  e  $\mathbf{D}$  pode ser determinado a partir de  $\vec{\mathbf{H}}$ , onde das equações (5.40) e (5.55) tem-se:

$$\vec{\mathbf{H}} = \begin{bmatrix} \mathbf{CAB} & \mathbf{CA}^2\mathbf{B} & \mathbf{CA}^3\mathbf{B} & \dots \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CA}^3\mathbf{B} & \mathbf{CA}^4\mathbf{B} & \dots \\ \mathbf{CA}^3\mathbf{B} & \mathbf{CA}^4\mathbf{B} & \mathbf{CA}^5\mathbf{B} & \dots \\ \vdots & & & \end{bmatrix} \quad (5.59)$$

o que permite escrever as expressões abaixo:

$$\mathbf{A} = \mathcal{O}^\dagger \vec{\mathbf{H}} \mathcal{C}^\dagger \quad (5.60)$$

$$\mathbf{B} = \mathcal{O}^\dagger \vec{\mathbf{H}}^{\rightarrow 1} \quad (5.61)$$

$$\mathbf{C} = \vec{\mathbf{H}}^{\rightarrow 1} \mathcal{C}^\dagger \quad (5.62)$$

onde  $\vec{\mathbf{H}}^{\rightarrow 1}$  e  $\vec{\mathbf{H}}^{\rightarrow 1}$  representam os primeiros blocos coluna e linha respectivamente da matriz  $\vec{\mathbf{H}}$ . Observe que as matrizes  $\mathbf{B}$  e  $\mathbf{C}$  também podem ser obtidas diretamente do primeiro bloco coluna de  $\mathcal{C}$  e do primeiro bloco linha de  $\mathcal{O}$  respectivamente. Além disso, o procedimento pode ser executado com uma matriz bloco de Hankel finita  $\mathbf{H}_{J \times J}$  para algum  $J > n$ , onde  $n$  é a ordem do sistema.

- Outra alternativa para obter a realização no espaço de estado é através do seguinte procedimento:
  - a) Obtenha um operador de Hankel a partir dos dados de entrada e saída. Determine o posto,  $n$ , deste operador e fatorize  $\mathbf{H} = \mathcal{O}\mathcal{C}$  onde as matrizes  $\mathcal{O}$  e  $\mathcal{C}$  terão postos,  $n$ , completos.
  - b) Conhecendo as matrizes  $\mathcal{O}$  e  $\mathcal{C}$  e utilizando a propriedade de invariância ao deslocamento obtenha:  $\vec{\mathcal{O}} = \mathcal{O}\mathbf{A}$ , o que implica  $\mathbf{A} = \mathcal{O}^\dagger \vec{\mathcal{O}}$ , onde  $\mathcal{O}^\dagger$  é a matriz pseudo inversa de  $\mathcal{O}$ . Como  $\mathcal{O}$  tem posto linha  $n$  completo, tem-se:  $\mathcal{O}^\dagger = (\mathcal{O}^T \mathcal{O})^{-1} \mathcal{O}^T$ . Então determine a matriz  $\mathbf{A}$ . As matrizes seguintes serão:  $\mathbf{B} = \mathcal{C}_{:,1}$ ;  $\mathbf{C} = \mathcal{O}_{:,1}$  e finalmente  $\mathbf{D}$  é determinado a partir da equação (5.21), como o primeiro elemento da resposta ao impulso  $\mathbf{D} = \mathbf{G}_0$  onde o subscrito  $(:, 1)$  representa a primeira coluna associada a matriz e  $(1, :)$  representa a primeira linha.

Estas relações são úteis para estudar as propriedades de uma realização no espaço de estado a partir das seqüências de entrada e de saída e várias abordagens surgiram nos últimos anos para viabilizar esquemas de modelagem de dados baseados em teoria de realização.

## 5.6 Modelos Estendidos no Espaço de Estado

**Teorema 5.2** *Para o sistema discreto linear estacionário multivariável modelado pela equação (5.3) pode-se escrever os seguintes modelos estendidos no espaço de estado (MEEE):*

$$a) \quad \mathbf{Y}_H(k) = \mathcal{O}_i \mathbf{X}(k) + \mathcal{T} \mathbf{U}_H(k) \quad (MEEE1) \quad (5.63)$$

$$b) \quad \mathbf{Y}_H(k) = \mathcal{O}_i \mathcal{C}_{j-1} \mathbf{U}_T(k) + \mathcal{T} \mathbf{U}_H(k) \quad (MEEE2) \quad (5.64)$$

$$c) \quad \mathbf{Y}_H(k+i) = \mathcal{A}_i \mathbf{Y}_H(k) + \mathcal{B}_i \mathbf{U}_{2i}(k) \quad (MEEE3) \quad (5.65)$$

onde a matriz  $\mathbf{Y}_H(k)$  com  $li$  linhas e  $j$  colunas terá consecutivos vetores de saída (de dimensão  $l \times 1$ , onde  $l$  é o número de saídas), ordenados da seguinte maneira:

$$\mathbf{Y}_H(k) = \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \cdots & \mathbf{y}_{k+j-1} \\ \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \cdots & \mathbf{y}_{k+j} \\ \mathbf{y}_{k+2} & \mathbf{y}_{k+3} & \cdots & \mathbf{y}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{k+i-1} & \mathbf{y}_{k+i} & \cdots & \mathbf{y}_{k+j+i-2} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_k^+ & \mathbf{y}_{k+1}^+ & \cdots & \mathbf{y}_{k+j-1}^+ \end{bmatrix} \quad (5.66)$$

com

$$\mathbf{y}_k^+ = \begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \vdots \\ \mathbf{y}_{k+i-1} \end{bmatrix}$$

A matriz  $\mathbf{U}_H(k)$  com a dimensão  $(mi) \times j$  contém consecutivos vetores de entrada (de dimensão  $m \times 1$ , onde  $m$  é o número de entradas), ordenados da seguinte maneira:

$$\mathbf{U}_H(k) = \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \mathbf{u}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \dots & \mathbf{u}_{k+j} \\ \mathbf{u}_{k+2} & \mathbf{u}_{k+3} & \dots & \mathbf{u}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \dots & \mathbf{u}_{k+j+i-2} \end{bmatrix} = [\mathbf{u}_k^+ \quad \mathbf{u}_{k+1}^+ \quad \dots \quad \mathbf{u}_{k+j-1}^+] \quad (5.67)$$

com

$$\mathbf{u}_k^+ = \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \vdots \\ \mathbf{u}_{k+i-1} \end{bmatrix}$$

e  $\mathbf{U}_{2i}(k)$  com a dimensão  $(2mi) \times i$ .  $\mathbf{U}_T(k)$  é uma matriz Toeplitz bloco superior de entradas com dimensão  $((j-1)m \times (j-1))$  dada por:

$$\mathbf{U}_T(k) = \begin{bmatrix} 0 & \mathbf{u}_k & \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \dots & \dots & \mathbf{u}_{k+j-2} \\ 0 & 0 & \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \dots & \mathbf{u}_{k+j-3} \\ 0 & 0 & 0 & \mathbf{u}_k & \dots & \dots & \mathbf{u}_{k+j-4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \mathbf{u}_k \end{bmatrix} \quad (5.68)$$

$\mathbf{X}(k)$  é uma matriz  $n \times j$  que contém a sequência dos vetores de estado:

$$\mathbf{X}(k) = [\mathbf{x}_k \quad \mathbf{x}_{k+1} \quad \mathbf{x}_{k+2} \quad \dots \quad \mathbf{x}_{k+j-1}] \quad (5.69)$$

$\mathcal{T}$  é uma matriz Toeplitz triangular inferior  $(li) \times (mi)$  contendo os parâmetros de Markov do sistema:  $\mathbf{D}$  na diagonal principal e  $\mathbf{CA}^{i-1}\mathbf{B}$  abaixo desta diagonal:

$$\mathcal{T} = \begin{bmatrix} \mathbf{D} & 0 & 0 & \dots & 0 \\ \mathbf{CB} & \mathbf{D} & 0 & \dots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \dots & 0 \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \dots & \mathbf{D} \end{bmatrix} \quad (5.70)$$

já obtida pela equação (5.24).  $\mathcal{O}_i$  é uma matriz  $(li) \times n$ , obtida da equação (5.36):

$$\mathcal{O}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-2} \\ \mathbf{CA}^{i-1} \end{bmatrix}$$

$\mathcal{C}_j$  e  $\mathcal{C}_j^{\leftrightarrow}$  são matrizes  $n \times (mj)$ , obtidas da equação (5.39):

$$\begin{aligned} \mathcal{C}_j &= [\mathbf{B} \quad \mathbf{AB} \quad \mathbf{A}^2\mathbf{B} \quad \dots \quad \mathbf{A}^{j-1}\mathbf{B}] \\ \mathcal{C}_j^{\leftrightarrow} &= [\mathbf{A}^{j-1}\mathbf{B} \quad \mathbf{A}^{j-2}\mathbf{B} \quad \dots \quad \mathbf{AB} \quad \mathbf{B}] \end{aligned}$$

$\mathcal{C}_j^{\leftrightarrow}$  é a matriz de atingibilidade reordenada e:

$$\mathcal{A}_i = \mathcal{O}_i \mathbf{A}^i \mathcal{O}_i^\dagger \quad (5.71)$$

$$\mathcal{B}_i = [\mathcal{O}_i \mathcal{C}_j^{\leftrightarrow} \quad \mathcal{T}] - \mathcal{A}_i [\mathcal{T} \quad 0_{li \times mj}] \quad (5.72)$$

**Prova:**

- a) Para o primeiro modelo estendido (MEEE1) a prova dá-se imediatamente pela substituição na equação (5.63) das matrizes (5.36), (5.66), (5.67), (5.69) e (5.70):

$$\begin{aligned} & \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \dots & \mathbf{y}_{k+j-1} \\ \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \dots & \mathbf{y}_{k+j} \\ \mathbf{y}_{k+2} & \mathbf{y}_{k+3} & \dots & \mathbf{y}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{k+i-1} & \mathbf{y}_{k+i} & \dots & \mathbf{y}_{k+j+i-2} \end{bmatrix} = \\ & = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} [\mathbf{x}_k \quad \mathbf{x}_{k+1} \quad \mathbf{x}_{k+2} \quad \dots \quad \mathbf{x}_{k+j-1}] + \\ & \begin{bmatrix} \mathbf{D} & 0 & 0 & \dots & 0 \\ \mathbf{CB} & \mathbf{D} & 0 & \dots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \dots & 0 \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \dots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \mathbf{u}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \dots & \mathbf{u}_{k+j} \\ \mathbf{u}_{k+2} & \mathbf{u}_{k+3} & \dots & \mathbf{u}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \dots & \mathbf{u}_{k+j+i-2} \end{bmatrix} \square. \end{aligned}$$

- b) Para provar o segundo modelo estendido (MEEE2) existe a necessidade de escrever a seqüência de equações de estado, equação (5.69), em função das entradas e da matriz de atingibilidade, deste modo tem-se:

$$\mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k ;$$

$$\begin{aligned} \mathbf{x}_{k+2} &= \mathbf{A}\mathbf{x}_{k+1} + \mathbf{B}\mathbf{u}_{k+1} = \mathbf{A}(\mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k) + \mathbf{B}\mathbf{u}_{k+1} = \\ &= \mathbf{A}^2\mathbf{x}_k + \mathbf{A}\mathbf{B}\mathbf{u}_k + \mathbf{B}\mathbf{u}_{k+1} , \end{aligned}$$

$$\begin{aligned} \mathbf{x}_{k+3} &= \mathbf{A}\mathbf{x}_{k+2} + \mathbf{B}\mathbf{u}_{k+2} = \mathbf{A}(\mathbf{A}^2\mathbf{x}_k + \mathbf{A}\mathbf{B}\mathbf{u}_k + \mathbf{B}\mathbf{u}_{k+1}) + \mathbf{B}\mathbf{u}_{k+2} = \\ &= \mathbf{A}^3\mathbf{x}_k + \mathbf{A}^2\mathbf{B}\mathbf{u}_k + \mathbf{A}\mathbf{B}\mathbf{u}_{k+1} + \mathbf{B}\mathbf{u}_{k+2} , \end{aligned}$$

generalizando tem-se:

$$\mathbf{x}_{k+j} = \mathbf{A}^j\mathbf{x}_k + \mathbf{A}^{j-1}\mathbf{B}\mathbf{u}_k + \mathbf{A}^{j-2}\mathbf{B}\mathbf{u}_{k+1} + \dots + \mathbf{A}\mathbf{B}\mathbf{u}_{k+j-2} + \mathbf{B}\mathbf{u}_{k+j-1} \quad (5.73)$$

então a seqüência de estados é dada por:

$$\begin{aligned} \mathbf{X}_k &= \begin{bmatrix} \mathbf{x}_k & \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \dots & \mathbf{x}_{k+j-1} \end{bmatrix} = \begin{bmatrix} \mathbf{x}_k & \mathbf{A}\mathbf{x}_k & \mathbf{A}^2\mathbf{x}_k & \dots & \mathbf{A}^{j-1}\mathbf{x}_k \end{bmatrix} + \\ &+ \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{B} & \mathbf{A}^2\mathbf{B} & \dots & \mathbf{A}^{j-2}\mathbf{B} \end{bmatrix} \begin{bmatrix} 0 & \mathbf{u}_k & \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \dots & \dots & \mathbf{u}_{k+j-2} \\ 0 & 0 & \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \dots & \mathbf{u}_{k+j-3} \\ 0 & 0 & 0 & \mathbf{u}_k & \dots & \dots & \mathbf{u}_{k+j-4} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & \dots & \mathbf{u}_k \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{x}_k & \mathbf{A}\mathbf{x}_k & \mathbf{A}^2\mathbf{x}_k & \dots & \mathbf{A}^{j-2}\mathbf{x}_k \end{bmatrix} + \mathcal{C}_{j-1}\mathbf{U}_{\mathcal{T}}(k) \end{aligned} \quad (5.74)$$

Porém os efeitos de um estado inicial,  $\mathbf{x}_k$ , não nulo, podem ser tratados como sendo causados por uma entrada adicional; para que isto ocorra, o sistema original ( $\mathbf{A}$ ,  $\mathbf{B}$ ,  $\mathbf{C}$ ,  $\mathbf{D}$ ) será redefinido como ( $\mathbf{A}$ ,  $\hat{\mathbf{B}}$ ,  $\mathbf{C}$ ,  $\hat{\mathbf{D}}$ ) com uma entrada adicional. Este novo sistema tem o estado inicial nulo,  $\hat{\mathbf{x}}_k = 0$ ,  $\hat{\mathbf{B}} = [\mathbf{B} \ \mathbf{A}\mathbf{x}_k]$  e  $\hat{\mathbf{D}} = [\mathbf{D} \ \mathbf{C}\mathbf{x}_k]$ ; substituindo na equação (5.73) tem-se:

$$\begin{aligned} \mathbf{x}_{k+j} &= \mathbf{A}^{j-1} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{x}_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ 1 \end{bmatrix} + \mathbf{A}^{j-2} \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{x}_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_{k-1} \\ 0 \end{bmatrix} + \dots + \\ &+ \begin{bmatrix} \mathbf{B} & \mathbf{A}\mathbf{x}_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_{k+j-1} \\ 0 \end{bmatrix} \end{aligned} \quad (5.75)$$

Observe que as equações (5.75) e (5.73) são iguais, no entanto na equação (5.75) o estado inicial original não nulo foi substituído por um vetor de entrada onde se incorpora uma entrada impulso unitário, o que permite afirmar que um estado inicial não nulo pode ser tratado pela inclusão de uma entrada adicional impulsiva e de uma coluna extra,  $\mathbf{Ax}_k$ , na matriz de entrada.

Neste caso, para as saídas do sistema tem-se:

$$\mathbf{y}_k = \mathbf{Cx}_k + \mathbf{Du}_k = \begin{bmatrix} \mathbf{D} & \mathbf{Cx}_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ 1 \end{bmatrix} \quad (5.76)$$

$$\mathbf{y}_{k+j-1} = \mathbf{Cx}_{k+j-1} + \mathbf{Du}_{k+j-1} = \mathbf{Cx}_{k+j-1} + \begin{bmatrix} \mathbf{D} & \mathbf{Cx}_k \end{bmatrix} \begin{bmatrix} \mathbf{u}_{k+j-1} \\ 0 \end{bmatrix} \quad (5.77)$$

Portanto são equivalentes os sistemas i) e ii) a seguir:

i)

$$\begin{cases} \mathbf{x}_{k+j} &= \mathbf{Ax}_{k+j-1} + \mathbf{Bu}_{k+j-1} & \text{para } \mathbf{x}_k \neq 0 \\ \mathbf{y}_{k+j-1} &= \mathbf{Cx}_{k+j-1} + \mathbf{Du}_{k+j-1} \end{cases}$$

ii)

$$\begin{cases} \hat{\mathbf{x}}_{k+j} &= \mathbf{A}\hat{\mathbf{x}}_{k+j-1} + \hat{\mathbf{B}}\hat{\mathbf{u}}_{k+j-1} & \text{para } \hat{\mathbf{x}}_k = 0 \\ \hat{\mathbf{y}}_{k+j-1} &= \mathbf{C}\hat{\mathbf{x}}_{k+j-1} + \hat{\mathbf{D}}\hat{\mathbf{u}}_{k+j-1} \end{cases}$$

sendo

$$\hat{\mathbf{u}}_{k+j-1} = \begin{bmatrix} \mathbf{u}_{k+j-1} \\ \delta_k \end{bmatrix}$$

e para evitar desnecessária complicação notacional, a partir deste ponto supõe-se estado inicial  $\mathbf{x}_k = 0$  e retorna-se à notação anterior, porém com esta hipótese, que o resultado acima justifica.

A prova do segundo modelo estendido (MEEE2) é obtida considerando o estado inicial  $\mathbf{x}_k = 0$  e a equação (5.74):

$$\mathbf{X}(k) = \mathcal{C}_{j-1} \mathbf{U}_{\mathcal{T}}(k) \quad (5.78)$$

Uma característica do MEEE2 é a possibilidade de concatenar as matrizes com os parâmetros de Markov bem como as entradas:



$$\mathbf{Y}_H(k) = \begin{bmatrix} \mathcal{O}_i \mathcal{C}_{j-1} & \mathcal{T} \end{bmatrix} \begin{bmatrix} \mathbf{U}_{\mathcal{T}}(k) \\ \mathbf{U}_H(k) \end{bmatrix} \quad (5.79)$$

A estrutura matricial Toeplitz das matrizes  $\mathcal{O}_i \mathcal{C}_{j-1}$  e  $\mathbf{U}_{\mathcal{T}}(k)$  possibilita reescrever a equação (5.79) como:

$$\mathbf{Y}_H(k) = \begin{bmatrix} \mathbf{G}_{j-1} & \mathbf{G}_{j-2} & \dots & \mathbf{G}_2 & \mathbf{G}_1 & \mathbf{G}_0 & 0 & \dots & 0 \\ \mathbf{G}_j & \mathbf{G}_{j-1} & \dots & \mathbf{G}_3 & \mathbf{G}_2 & \mathbf{G}_1 & \mathbf{G}_0 & \dots & 0 \\ \mathbf{G}_{j+1} & \mathbf{G}_j & \dots & \mathbf{G}_4 & \mathbf{G}_3 & \mathbf{G}_2 & \mathbf{G}_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots & & \vdots & & \dots & 0 \\ \mathbf{G}_{i+j-2} & \mathbf{G}_{i+j-3} & \dots & \mathbf{G}_{i-3} & \mathbf{G}_{i-2} & \mathbf{G}_{i-1} & & \dots & \mathbf{G}_0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & \mathbf{u}_k \\ 0 & 0 & \dots & 0 & \mathbf{u}_k & \mathbf{u}_{k+1} \\ 0 & 0 & \dots & \mathbf{u}_k & \mathbf{u}_{k+1} & \mathbf{u}_{k+2} \\ \vdots & \vdots & \vdots & \ddots & & \vdots \\ 0 & \mathbf{u}_k & \dots & \mathbf{u}_{k+j-3} & \mathbf{u}_{k+j-2} \\ \mathbf{u}_k & \mathbf{u}_{k+1} & \dots & \mathbf{u}_{k+j-2} & \mathbf{u}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \dots & \mathbf{u}_{k+j-1} & \mathbf{u}_{k+j} \\ \mathbf{u}_{k+2} & \mathbf{u}_{k+3} & \dots & \mathbf{u}_{k+j} & \mathbf{u}_{k+j+1} \\ \vdots & \vdots & \vdots & & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \dots & \mathbf{u}_{k+j+i-1} & \mathbf{u}_{k+j+i-2} \end{bmatrix} \quad (5.80)$$

Na equação (5.80) tem-se o produto da matriz bloco,  $\mathbf{M}_g$ , Toeplitz com os parâmetros de Markov, pela matriz bloco,  $\mathbf{M}_u$ , de Hankel com os vetores de entrada:

$$\mathbf{Y}_H(k) = \mathbf{M}_g \mathbf{M}_u \quad (5.81)$$

- c) Para provar o terceiro modelo estendido (MEEE3) considere o modelo estendido (MEEE1), equação (5.63), donde para uma seqüência vetorial,  $\mathbf{y}_k^+$ , tem-se:

$$\mathbf{y}_k^+ = \mathcal{O}_i \mathbf{x}_k + \mathcal{T} \mathbf{u}_k^+ \quad (5.82)$$

considerando que esta seqüência ocorre  $i$  passos à frente tem-se:

$$\mathbf{y}_{k+i}^+ = \mathcal{O}_i \mathbf{x}_{k+i} + \mathcal{T} \mathbf{u}_{k+i}^+ \quad (5.83)$$

A seqüência de vetores de estado  $\mathbf{x}(k+i)$  pode ser escrita em relação à seqüência de estado presente  $\mathbf{x}(k)$  como:

$$\begin{aligned}
\mathbf{x}(k+i) &= \mathbf{A}^i \mathbf{x}(k) + \begin{bmatrix} \mathbf{A}^{i-1} \mathbf{B} & \mathbf{A}^{i-2} \mathbf{B} & \dots & \mathbf{A} \mathbf{B} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \vdots \\ \mathbf{u}_{k+i-2} \\ \mathbf{u}_{k+i-1} \end{bmatrix} \\
&= \mathbf{A}^i \mathbf{x}(k) + \begin{bmatrix} \mathbf{A}^{i-1} \mathbf{B} & \mathbf{A}^{i-2} \mathbf{B} & \dots & \mathbf{A} \mathbf{B} & \mathbf{B} \end{bmatrix} \mathbf{u}_k^+
\end{aligned} \tag{5.84}$$

observe que para escrever a equação (5.84) utiliza-se da equação (5.74) com reordenações das entradas, para obter  $\mathbf{u}_k^+$ , e da matriz de atingibilidade; desta forma a matriz reordenada de atingibilidade é:

$$\mathcal{C}_i^{\epsilon p} = \begin{bmatrix} \mathbf{A}^{i-1} \mathbf{B} & \mathbf{A}^{i-2} \mathbf{B} & \dots & \mathbf{A} \mathbf{B} & \mathbf{B} \end{bmatrix} \tag{5.85}$$

substituindo a equação (5.84) em (5.83) tem-se:

$$\mathbf{y}_{k+i}^+ = \mathcal{O}_i \mathbf{A}^i \mathbf{x}_k + \mathcal{O}_i \mathcal{C}_i^{\epsilon p} \mathbf{u}_k^+ + \mathcal{T} \mathbf{u}_{k+i}^+ \tag{5.86}$$

A seqüência de estados atual é obtida da equação (5.82):

$$\mathbf{x}_k = \mathcal{O}_i^\dagger (\mathbf{y}_k^+ - \mathcal{T} \mathbf{u}_k^+) \tag{5.87}$$

finalmente substituindo  $\mathbf{x}_k$ , definido pela equação (5.87), na equação (5.86) tem-se:

$$\mathbf{y}_{k+i}^+ = \mathcal{O}_i \mathbf{A}^i \mathcal{O}_i^\dagger (\mathbf{y}_k^+ - \mathcal{T} \mathbf{u}_k^+) + \mathcal{O}_i \mathcal{C}_i^{\epsilon p} \mathbf{u}_k^+ + \mathcal{T} \mathbf{u}_{k+i}^+ \tag{5.88}$$

fazendo

$$\mathcal{A}_i = \mathcal{O}_i \mathbf{A}^i \mathcal{O}_i^\dagger \tag{5.89}$$

tem-se

$$\begin{aligned}
\mathbf{y}_{k+i}^+ &= \mathcal{A}_i (\mathbf{y}_k^+ - \mathcal{T} \mathbf{u}_k^+) + \mathcal{O}_i \mathcal{C}_i^{\epsilon p} \mathbf{u}_k^+ + \mathcal{T} \mathbf{u}_{k+i}^+ \\
\mathbf{y}_{k+i}^+ &= \mathcal{A}_i \mathbf{y}_k^+ + (\mathcal{O}_i \mathcal{C}_i^{\epsilon p} - \mathcal{A}_i \mathcal{T}) \mathbf{u}_k^+ + \mathcal{T} \mathbf{u}_{k+i}^+ \\
\mathbf{y}_{k+i}^+ &= \mathcal{A}_i \mathbf{y}_k^+ + \begin{bmatrix} \mathcal{O}_i \mathcal{C}_i^{\epsilon p} & \mathcal{T} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k^+ \\ \mathbf{u}_{k+i}^+ \end{bmatrix} - \begin{bmatrix} \mathcal{A}_i \mathcal{T} & 0_{il \times jm} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k^+ \\ \mathbf{u}_{k+i}^+ \end{bmatrix}
\end{aligned} \tag{5.90}$$

$$\mathbf{y}_{k+i}^+ = \mathcal{A}_i \mathbf{y}_k^+ + \left[ \begin{bmatrix} \mathcal{O}_i \mathcal{C}_i^{\leftarrow p} & \mathcal{T} \end{bmatrix} - \begin{bmatrix} \mathcal{A}_i \mathcal{T} & 0_{il \times jm} \end{bmatrix} \right] \begin{bmatrix} \mathbf{u}_k^+ \\ \mathbf{u}_{k+i}^+ \end{bmatrix}$$

fazendo

$$\mathcal{B}_i = \left[ \begin{bmatrix} \mathcal{O}_i \mathcal{C}_i^{\leftarrow p} & \mathcal{T} \end{bmatrix} - \begin{bmatrix} \mathcal{A}_i \mathcal{T} & 0_{li \times mj} \end{bmatrix} \right]$$

tem-se:

$$\mathbf{y}_{k+i}^+ = \mathcal{A}_i \mathbf{y}_k^+ + \mathcal{B}_i \begin{bmatrix} \mathbf{u}_k^+ \\ \mathbf{u}_{k+i}^+ \end{bmatrix} \quad (5.91)$$

como

$$\mathbf{u}_{k+i}^+ = \begin{bmatrix} \mathbf{u}_{k+i} \\ \mathbf{u}_{k+i+1} \\ \mathbf{u}_{k+i+2} \\ \vdots \\ \mathbf{u}_{k+i+j-1} \end{bmatrix}$$

para  $i = j$  pode-se escrever:

$$\mathbf{y}_{k+i}^+ = \mathcal{A}_i \mathbf{y}_k^+ + \mathcal{B}_i \mathbf{u}_{2i}^+ \quad (5.92)$$

donde tem-se:

$$\mathbf{Y}_H(k+i) = \mathcal{A}_i \mathbf{Y}_H(k) + \mathcal{B}_i \mathbf{U}_{2i}(k)$$

o que corresponde ao MEEE3; em particular para  $i = 1$ :

$$\mathbf{Y}_H(k+1) = \mathcal{A}_1 \mathbf{Y}_H(k) + \mathcal{B}_1 \mathbf{U}_2(k) \quad (5.93)$$

**C.Q.D.**

## 5.7 Métodos de Subespaço para Modelagem de Dados: Idéias Básicas.

Utilizando os modelos estendidos apresentados no teorema 5.2 pode-se desenvolver métodos de subespaços para modelagem de dados no espaço de estado explorando tanto projeções como interseções.

### 5.7.1 Projeções Ortogonais para um Modelo de Estado Estendido

Uma idéia básica para modelagem de dados em subespaços é recuperar o termo  $\mathcal{O}_i \mathbf{X}(k)$  da equação (5.63); para isto podemos fazer a projeção de cada termo desta equação no complemento ortogonal de  $\mathbf{U}_H$ ,  $\mathbf{U}_H^\perp$ , Figura 5.5, donde:

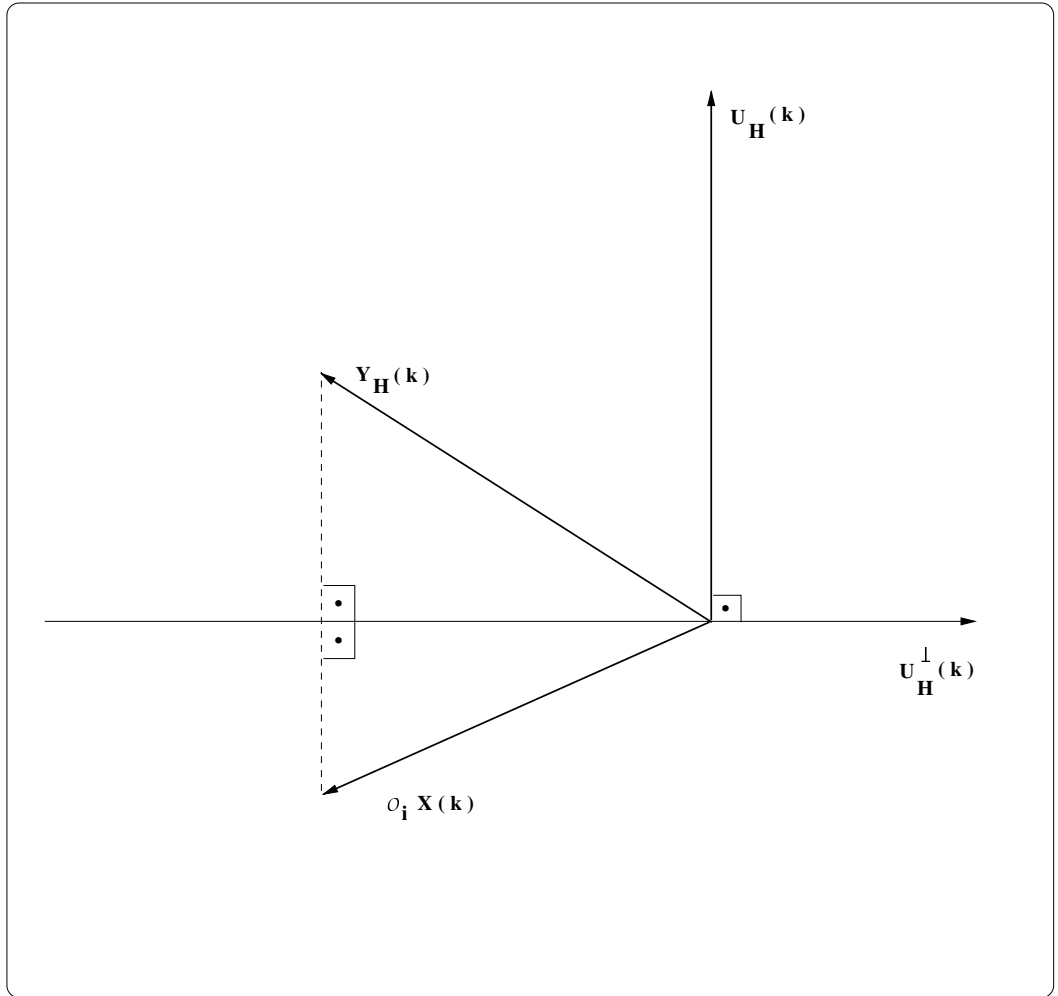


Figura 5.5: Projeções de  $\mathbf{Y}$  e  $\mathbf{X}$  em  $\mathbf{U}_H^\perp$ .

$$\mathbf{Y}_H(k)|\mathbf{U}_H(k)^\perp = \mathcal{O}_i \mathbf{X}(k)|\mathbf{U}_H(k)^\perp + \mathcal{T} \mathbf{U}_H(k)|\mathbf{U}_H(k)^\perp \quad (5.94)$$

logo

$$\mathbf{Y}_H(k)|\mathbf{U}_H(k)^\perp = \mathcal{O}_i \mathbf{X}(k)|\mathbf{U}_H(k)^\perp \quad (5.95)$$

possibilita obter a matriz de seqüência de estados,  $\mathbf{X}(k)$ , e ou a matriz de observabilidade,  $\mathcal{O}_i$ , através de técnicas de decomposição matricial; por exemplo utilizando decomposições em valores singulares ou  $\mathbf{RQ}$  tem-se:

$$\mathbf{Y}_H(k)|\mathbf{U}_H(k)^\perp = [\mathbf{U}_1 \quad \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} = \mathbf{U}_1 \Sigma_1^{1/2} \Sigma_1^{1/2} \mathbf{V}_1^T = \mathbf{RQ}$$

desta forma  $\mathcal{O}_i$  e  $\mathbf{X}(k)$  podem ser obtidos como:

$$\mathcal{O}_i = \mathbf{U}_1 \Sigma_1^{1/2} = \mathbf{R}$$

$$\mathbf{X}(k)|\mathbf{U}_H(k)^\perp = \Sigma_1^{1/2} \mathbf{V}_1^T = \mathbf{Q}$$

A exploração desta idéia para os diversos modelos estendidos permite-nos propor uma variedade de métodos de subespaço para modelagem de dados no espaço de estado por constituir a base dos mesmos. Nas Figuras 5.6, 5.7, 5.8, 5.9, 5.10, 5.11, 5.12 e 5.13 estão representações geométricas dos espaços fundamentais das matrizes da equação (5.95).

Para resolver o problema de modelagem de dados precisa-se estimar a ordem do sistema e identificar as matrizes do modelo no espaço de estado. Para tanto o sistema deve estar persistentemente excitado:

**Definição 5.1** *Para o sistema representado pela equação (5.3) a entrada é persistentemente excitante de ordem  $mi$  quando posto( $\mathbf{U}_H(k)$ ) =  $mi$ .*

Considerando o teorema 5.2, enuncia-se e prova-se o seguinte corolário:

**Corolário 5.4** *Se  $n$  é a ordem da realização mínima do sistema e se:*

- a)  $i \geq n$
- b)  $\mathbf{U}_H(k)$  é uma matriz retangular com mais colunas que linhas:  $j > mi + n$ .<sup>1</sup>
- c)  $\text{posto}(\mathbf{X}(k)) = n$ , ( todos os modos são suficientemente excitados).
- d) A intersecção do espaço linha de  $\mathbf{X}(k)$  com o espaço linha de  $\mathbf{U}_H(k)$  é vazia,  $\Re(\mathbf{X}^T(k)) \cap \Re(\mathbf{U}_H^T(k)) = \emptyset$ .

---

<sup>1</sup>Como  $\text{posto}(\mathbf{U}_H(k)) \neq j$ , pois se  $\text{posto}(\mathbf{U}_H(k)) = j$  não teremos complemento ortogonal do espaço linha de  $\mathbf{U}_H(k)$ , logo  $j > \text{posto}(\mathbf{U}_H(k))$ . Para que todas as entradas no instante  $k$  sejam independentes:  $\text{posto}(\mathbf{U}_H(k)) = mi < j$ .

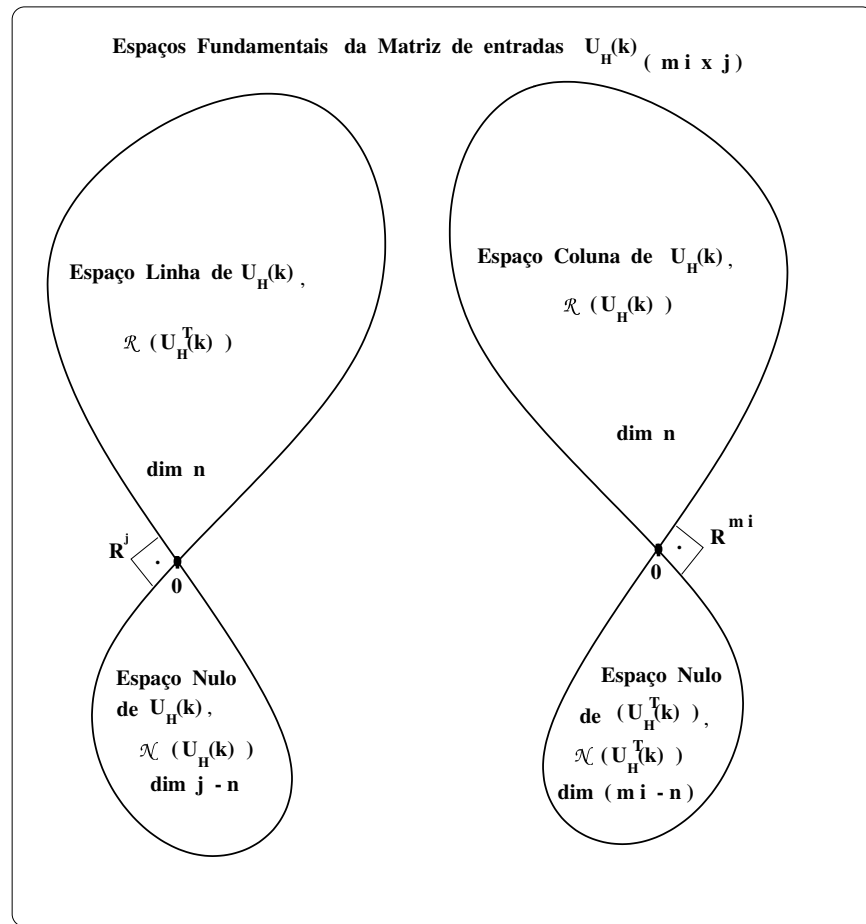


Figura 5.6: Representação Geométrica dos Espaços Fundamentais da matriz  $U_H(k)$ .

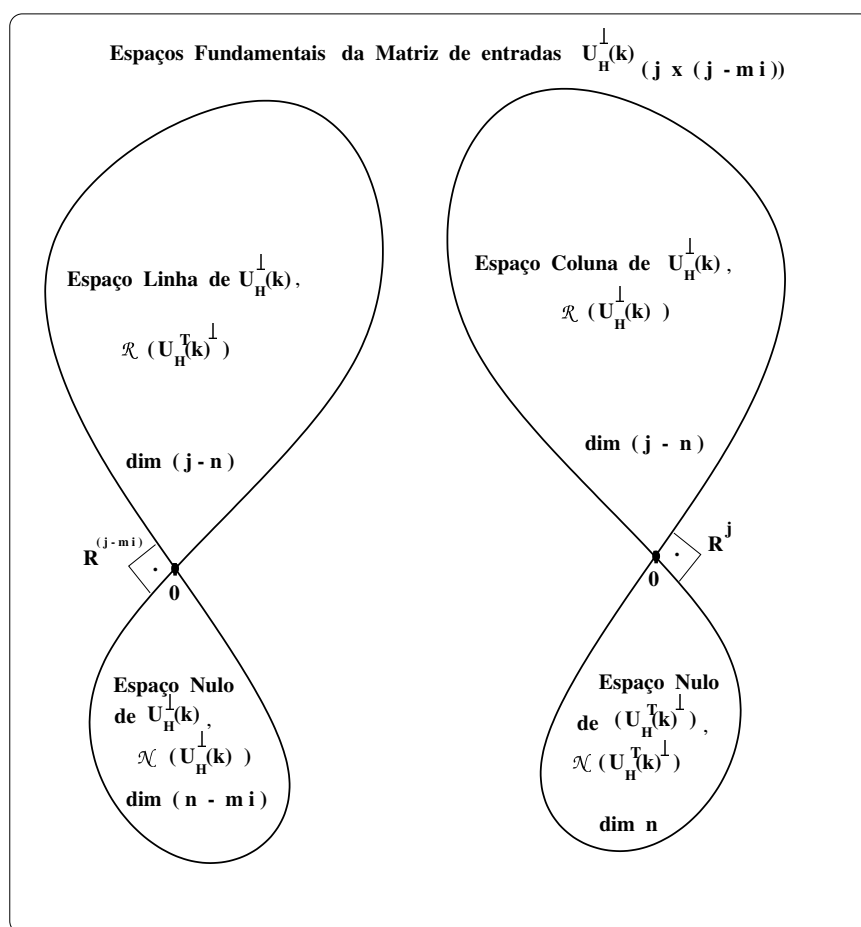


Figura 5.7: Representação Geométrica dos Espaços Fundamentais da matriz  $U_H(k)^\perp$ .

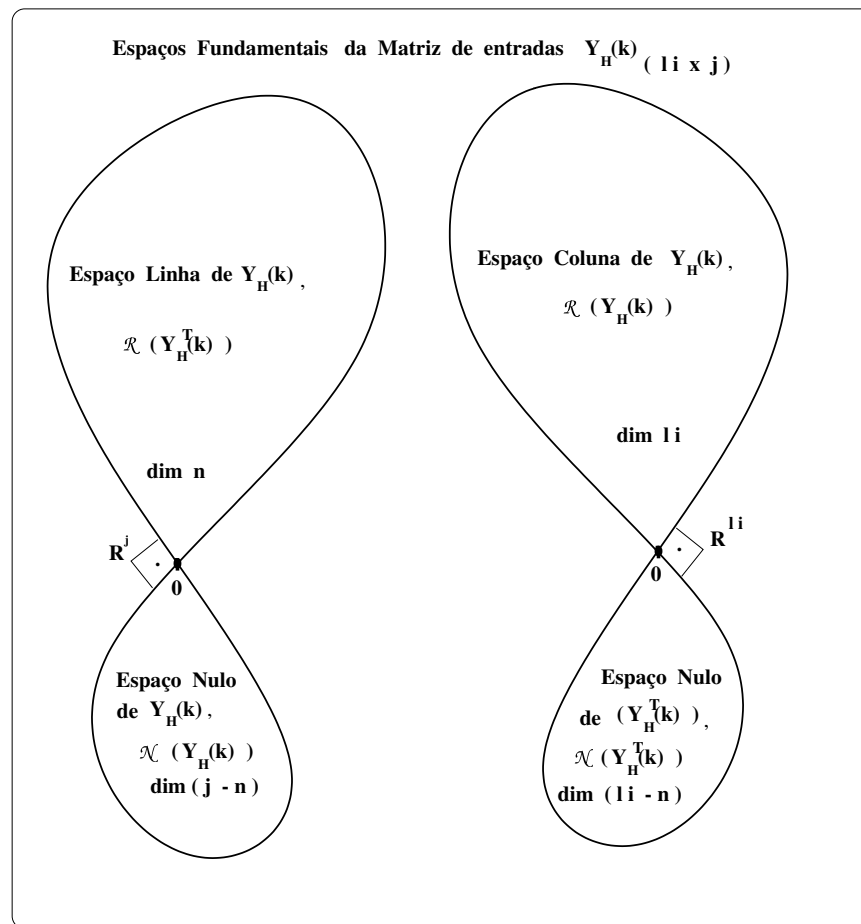


Figura 5.8: Representação Geométrica dos Espaços Fundamentais da matriz  $\mathbf{Y}_H(k)$ .



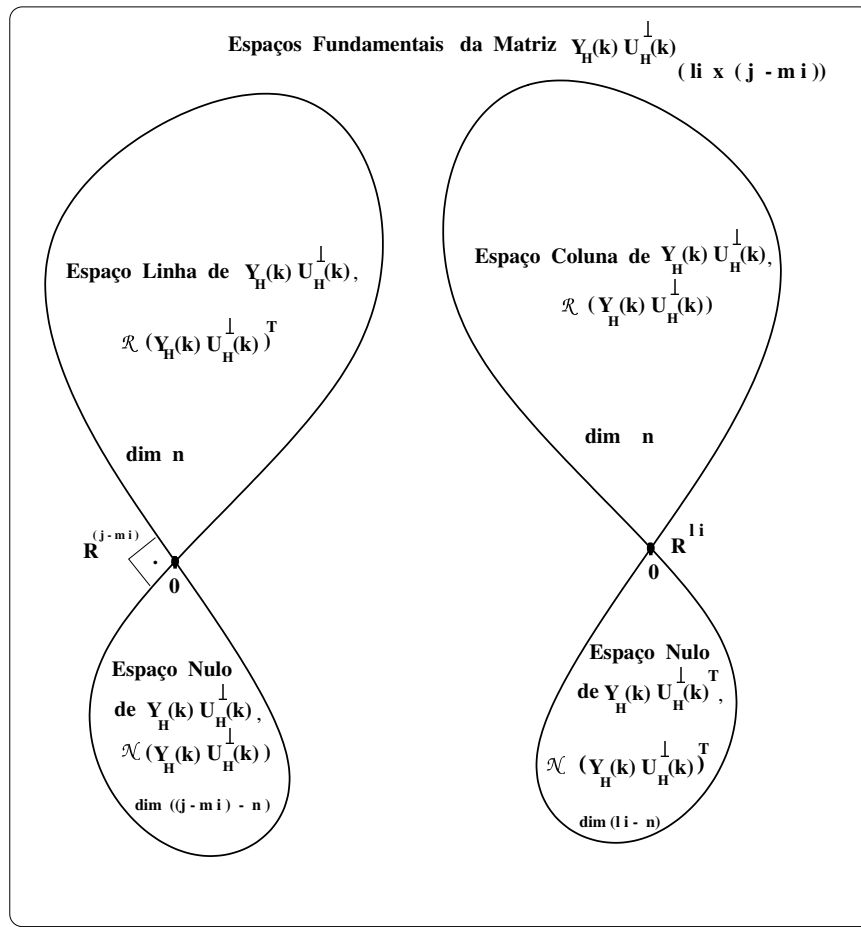


Figura 5.9: Representação Geométrica dos Espaços Fundamentais da matriz  $\mathbf{Y}_H(k) \mathbf{U}_H(k)^\perp$ .

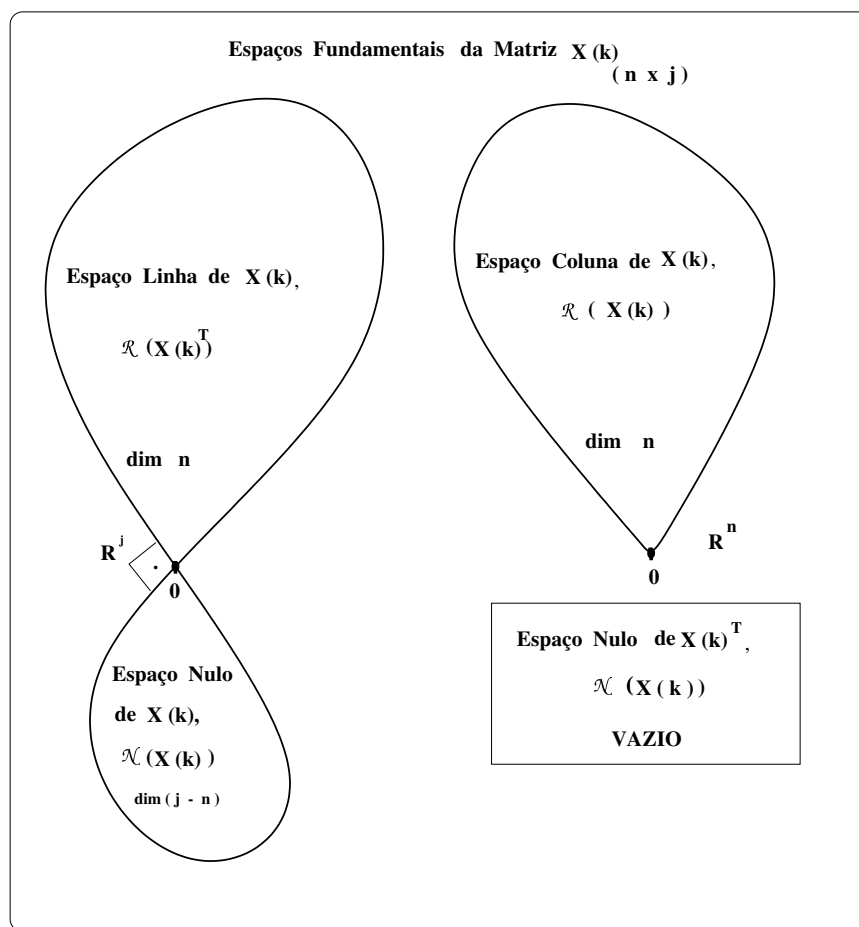


Figura 5.10: Representação Geométrica dos Espaços Fundamentais da matriz  $\mathbf{X}(k)$ .

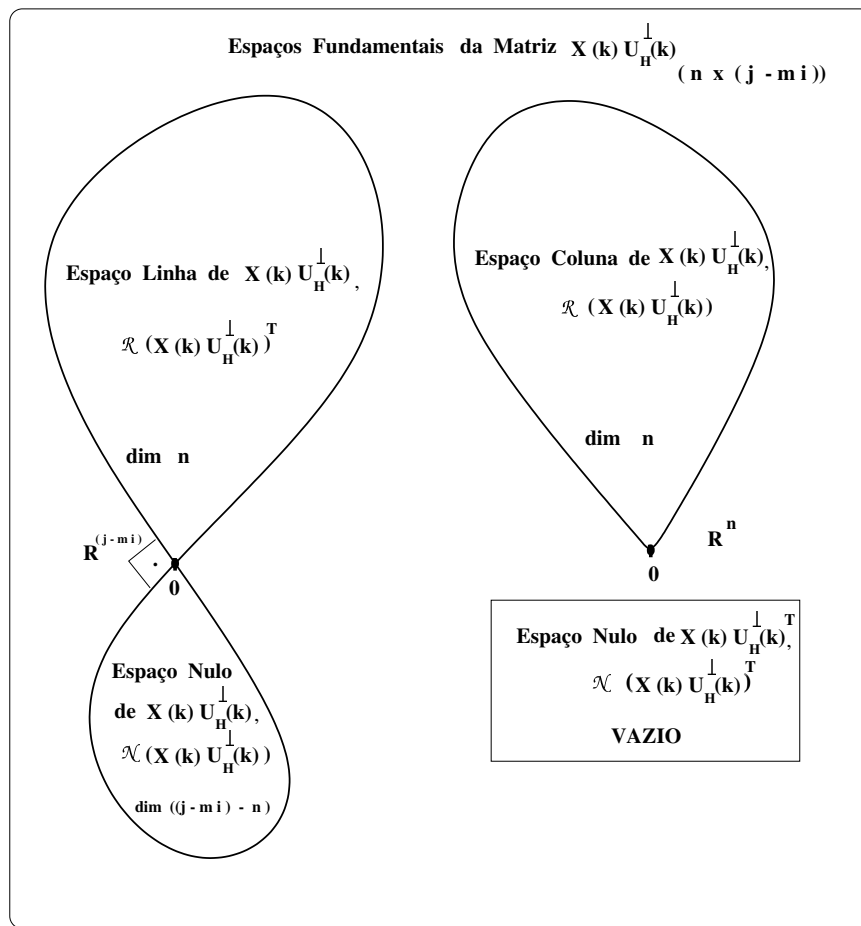


Figura 5.11: Representação Geométrica dos Espaços Fundamentais da matriz  $\mathbf{X}(\mathbf{k}) \mathbf{U}_H(\mathbf{k})^\perp$ .

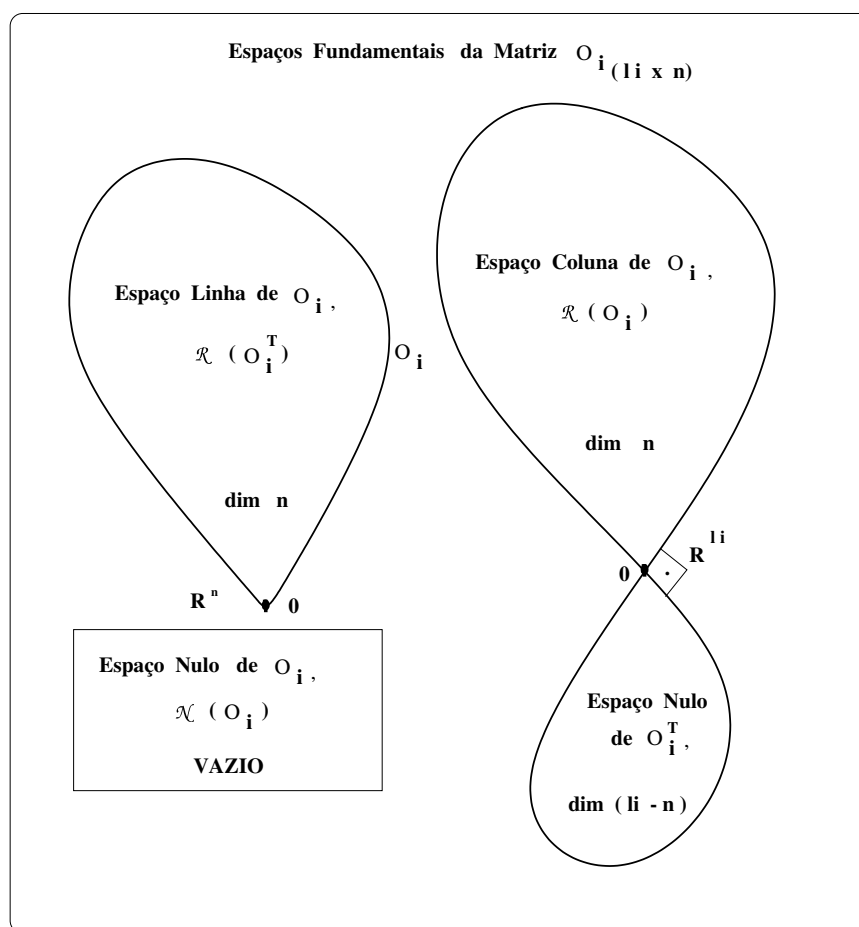


Figura 5.12: Representação Geométrica dos Espaços Fundamentais da matriz  $O_i$ .

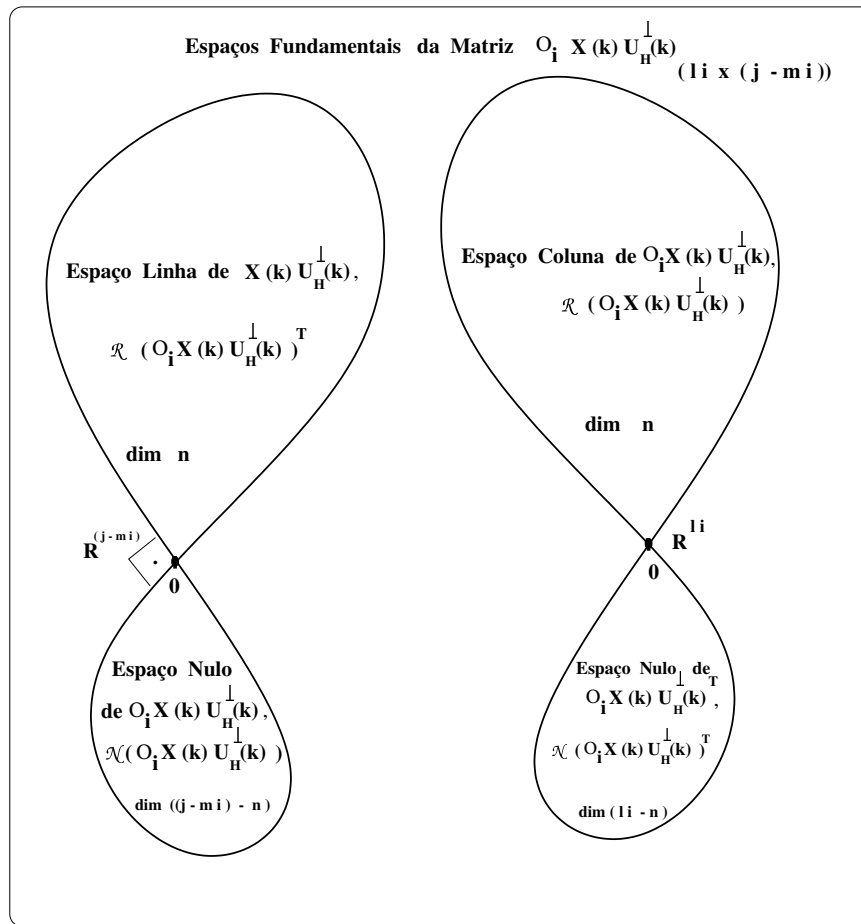


Figura 5.13: Representação Geométrica dos Espaços Fundamentais da matriz  $O_i X(k) U_H(k)^\perp$ .

e)  $\mathbf{W}$  é uma matriz concatenada:

$$\mathbf{W} = \begin{bmatrix} \mathbf{Y}_H(k) \\ \mathbf{U}_H(k) \end{bmatrix} \quad (5.96)$$

então:

$$i) \quad \mathbf{X}(k) = \mathcal{O}_i^\dagger [\mathbf{I}_{li} - \mathcal{T}] \mathbf{W} \quad (5.97)$$

$$ii) \quad \text{posto}(\mathbf{Y}_H(k) \mathbf{U}_H(k)^\perp) = n \quad (5.98)$$

$$iii) \quad \text{posto}(\mathbf{W}) = \text{posto}(\mathbf{U}_H(k)) + n = mi + n \quad (5.99)$$

O  $\text{posto}(\mathbf{U}_H(k)) = mi$ , onde  $mi$  = número de linhas linearmente independentes de  $\mathbf{U}_H(k)$ .

**Prova:**

i) Da equação (5.63) e da matriz (5.96) é imediata.

ii) Da equação (5.95) tem-se:

$$\mathbf{Y}_H(k) \mathbf{U}_H(k)^\perp = \mathcal{O}_i \mathbf{X}(k) \mathbf{U}_H(k)^\perp$$

portanto

$$\text{posto}(\mathbf{Y}_H(k) \mathbf{U}_H(k)^\perp) = \text{posto}(\mathcal{O}_i \mathbf{X}(k) \mathbf{U}_H(k)^\perp)$$

onde o espaço coluna de  $\mathbf{U}_H(k)^\perp$  é o espaço nulo de  $\mathbf{U}_H(k)$ . Considerando que  $\mathcal{O}_i$  tem posto coluna completo ( condição de observabilidade) e fazendo uso do Teorema de Sylvester <sup>2</sup> e das Figuras 5.6, 5.7, 5.10, 5.11, 5.12 e 5.13 tem-se:

---

<sup>2</sup>Sendo as matrizes  $\mathbf{C} = \mathbf{AB}$ , onde  $\mathbf{A}_{m \times n}$ ,  $\mathbf{B}_{n \times p}$  e  $\mathbf{C}_{m \times p}$  de posto  $r(\mathbf{A})$ ,  $r(\mathbf{B})$  e  $r(\mathbf{C})$  respectivamente. Então:

$$r(\mathbf{C}) = \min(r(\mathbf{A}), r(\mathbf{B})) - \dim[\Re(\mathbf{A}^T) \cap \Re^\perp(\mathbf{B})]$$

**Prova** Gantmacher, [368]; donde temos:

$$\dim[\Re(\mathbf{A}) \cap \Re(\mathbf{B})] = r(\mathbf{A}) + r(\mathbf{B}) - r(\mathbf{AB})$$

$$\dim[\Re(\mathbf{A}^T) \cap \Re(\mathbf{C}^T)] = r(\mathbf{A}) + r(\mathbf{C}) - r(\mathbf{AC})$$

$$\begin{aligned}
& posto(\mathcal{O}_i \mathbf{X}(k) \mathbf{U}_H(k)^\perp) = \\
& = posto(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) - \dim[\Re(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) \cap (\Re(\mathcal{O}_i)^T)^\perp] = \\
& = posto(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) - \dim(\Re(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) \cap \emptyset) = posto(\mathbf{X}(k) \mathbf{U}_H(k)^\perp)
\end{aligned}$$

Fazendo o uso da condição d)

$$\begin{aligned}
& posto(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) = posto(\mathbf{X}(k)) - \dim[\Re(\mathbf{X}(k))^T \cap (\Re(\mathbf{U}_H(k)^\perp)^\perp)] = \\
& posto(\mathbf{X}(k)) - \dim[\Re(\mathbf{X}(k))^T \cap (\Re(\mathbf{U}_H^T(k)))] = posto(\mathbf{X}(k))
\end{aligned}$$

Portanto  $posto(\mathbf{X}(k) \mathbf{U}_H(k)^\perp) = n$  o que prova ii).

Uma interpretação de  $\mathcal{O}_i \mathbf{X}(k) \mathbf{U}_H(k)^\perp$  é apresentada na Figura 5.14.

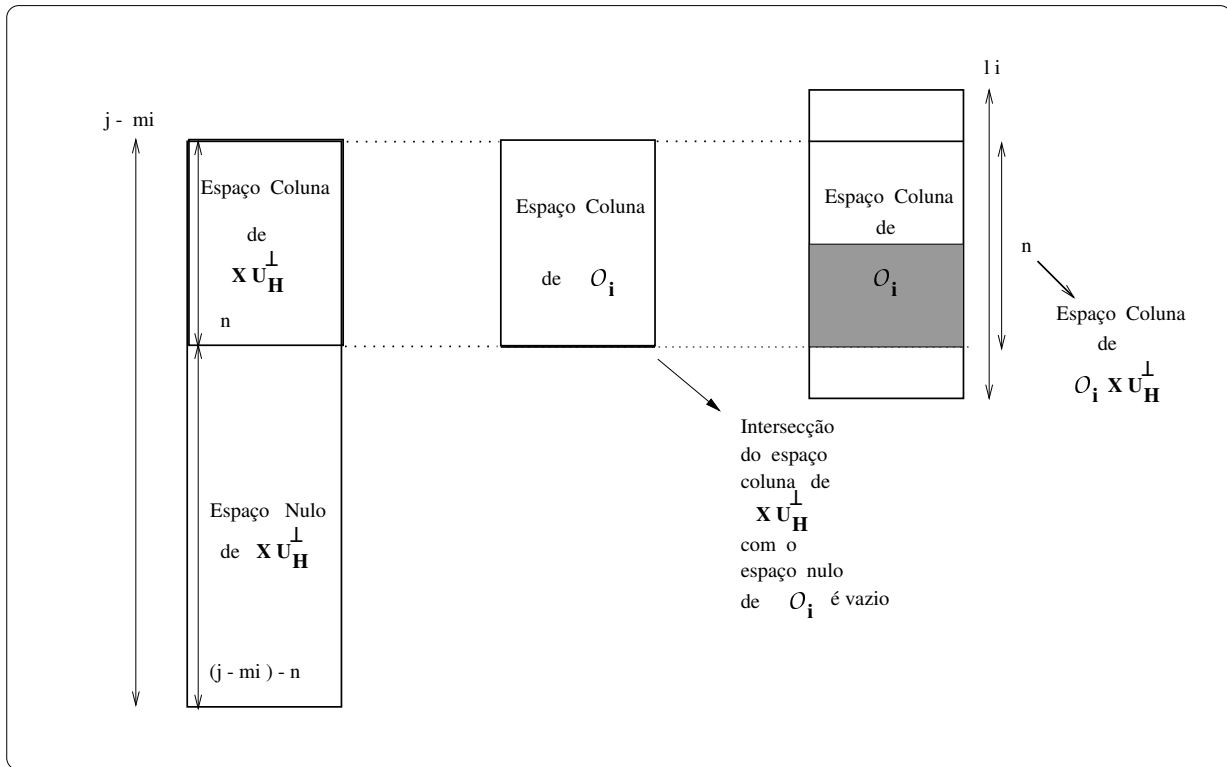


Figura 5.14: Interpretação Geométrica Para Determinação da Dimensão do Estado do Sistema.

iii) Como

$$posto(\mathbf{W}) = posto(\mathbf{W} \mathbf{U}_H^\perp(k)) + posto(\mathbf{W} \mathbf{U}_H(k)) \quad (5.100)$$

$$\text{posto}(\mathbf{W}\mathbf{U}_H^\perp(k)) = \text{posto}(\mathbf{Y}_H\mathbf{U}_H^\perp(k)) \quad (5.101)$$

$$\text{posto}(\mathbf{W}\mathbf{U}_H(k)) = \text{posto}(\mathbf{U}_H(k)) \quad (5.102)$$

Das equações (5.100), (5.101), (5.102) e (5.98) tem-se:

$$\text{posto}(\mathbf{W}) = n + mi$$

Portanto o espaço linha de  $\mathbf{Y}_H(k)$  adiciona  $n$  dimensões ao espaço linha de  $\mathbf{U}_H(k)$ .

**C.Q.D.**

### 5.7.2 Algumas Propriedades Geométricas de Um Modelo de Estado Estendido

Do MEEE1 tem-se:

$$\mathbf{Y}_f = \mathcal{O}_i\mathbf{X}_f + \mathcal{T}\mathbf{U}_f \quad (5.103)$$

e de forma análoga:

$$\mathbf{Y}_p = \mathcal{O}_i\mathbf{X}_p + \mathcal{T}\mathbf{U}_p \quad (5.104)$$

onde os subíndices  $p$  e  $f$  denotam coleções de dados passados e futuros respectivamente; onde:

$$\mathbf{Y}_{2H} = \begin{bmatrix} \mathbf{Y}_p \\ \mathbf{Y}_f \end{bmatrix} = \begin{bmatrix} \mathbf{y}_0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{j-1} \\ \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_j \\ \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{i-1} & \mathbf{y}_i & \cdots & \mathbf{y}_{j+i-2} \\ \mathbf{y}_i & \mathbf{y}_{i+1} & \cdots & \mathbf{y}_{i+j-1} \\ \mathbf{y}_{i+1} & \mathbf{y}_{i+2} & \cdots & \mathbf{y}_{i+j} \\ \mathbf{y}_{i+2} & \mathbf{y}_{i+3} & \cdots & \mathbf{y}_{i+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{2i-1} & \mathbf{y}_{2i} & \cdots & \mathbf{y}_{2i+j-2} \end{bmatrix}$$



$$\mathbf{U}_{2H} = \left[ \begin{array}{c} \mathbf{U}_p \\ \mathbf{U}_f \end{array} \right] = \left[ \begin{array}{cccc} \mathbf{u}_0 & \mathbf{u}_1 & \dots & \mathbf{u}_{j-1} \\ \mathbf{u}_1 & \mathbf{u}_2 & \dots & \mathbf{u}_j \\ \mathbf{u}_2 & \mathbf{u}_3 & \dots & \mathbf{u}_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{i-1} & \mathbf{u}_i & \dots & \mathbf{u}_{j+i-2} \\ \mathbf{u}_i & \mathbf{u}_{i+1} & \dots & \mathbf{u}_{i+j-1} \\ \mathbf{u}_{i+1} & \mathbf{u}_{i+2} & \dots & \mathbf{u}_{i+j} \\ \mathbf{u}_{i+2} & \mathbf{u}_{i+3} & \dots & \mathbf{u}_{i+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{2i-1} & \mathbf{u}_{2i} & \dots & \mathbf{u}_{2i+j-2} \end{array} \right]$$

e para os estados

$$\mathbf{X}_p = \left[ \begin{array}{cccc} \mathbf{x}_0 & \mathbf{x}_1 & \mathbf{x}_2 & \dots & \mathbf{x}_{j-1} \end{array} \right] \quad (5.105)$$

$$\mathbf{X}_f = \left[ \begin{array}{cccc} \mathbf{x}_i & \mathbf{x}_{i+1} & \mathbf{x}_{i+2} & \dots & \mathbf{x}_{i+j-1} \end{array} \right] \quad (5.106)$$

logo tem-se:

$$\mathbf{X}_f = [\mathcal{O}_i^\dagger \quad - \quad \mathcal{O}_i^\dagger \mathcal{T}] \left[ \begin{array}{c} \mathbf{Y}_f \\ \mathbf{U}_f \end{array} \right] \quad (5.107)$$

e

$$\mathbf{X}_p = [\mathcal{O}_i^\dagger \quad - \quad \mathcal{O}_i^\dagger \mathcal{T}] \left[ \begin{array}{c} \mathbf{Y}_p \\ \mathbf{U}_p \end{array} \right] \quad (5.108)$$

Utilizando a equação (5.84) e a matriz de atingibilidade reordenada,  $\mathcal{C}_j^{\leftarrow p}$ , tem-se:

$$\mathbf{X}_f = \mathbf{A}^i \mathbf{X}_p + \mathcal{C}_j^{\leftarrow p} \mathbf{U}_p \quad (5.109)$$

### Intersecção de Subespaços

Algoritmos para modelagem de dados utilizando técnicas de intersecção foram primeiramente apresentados por De Moor e co-autores [259, 763, 1091], e a discussão a seguir fundamenta-se nas suas contribuições. Nestes algoritmos o espaço linha de uma seqüência futura de estados,  $\mathbf{X}_f$ , é calculado como a intersecção do espaço linha de entradas e saídas passadas,  $\mathbf{W}_p$ , com o espaço linha de entradas e saídas futuras,  $\mathbf{W}_f$ :

$$\Re(\mathbf{X}_f^T) = \Re(\mathbf{W}_p^T) \cap \Re(\mathbf{W}_f^T), \quad (5.110)$$

onde

$$\mathbf{W}_p = \begin{bmatrix} \mathbf{U}_p \\ \mathbf{Y}_p \end{bmatrix} \quad \text{e} \quad \mathbf{W}_f = \begin{bmatrix} \mathbf{U}_f \\ \mathbf{Y}_f \end{bmatrix}$$

Portanto a intersecção dos espaços linha de matrizes bloco de Hankel construídas com vetores de dados de entrada-saída passados e futuros respectivamente nos dá a sequência de vetores de estado futuros.

A equação (5.110) pode ser demonstrada a partir das equações (5.103) e (5.104). Da equação (5.99) tem-se:

$$\text{posto}(\mathbf{W}_p) = \text{posto}(\mathbf{W}_f) = mi + n$$

logo

$$\dim(\mathfrak{R}(\mathbf{W}_p^T)) = \dim(\mathfrak{R}(\mathbf{W}_f^T)) = mi + n$$

Da mesma forma para a concatenação:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_p \\ \mathbf{W}_f \end{bmatrix}$$

tem-se:

$$\dim(\mathbf{W}_p + \mathbf{W}_f) = \dim(\mathbf{W}) = 2mi + n$$

Pelo teorema de dimensão de Grassmann:

$$\begin{aligned} \dim(\mathbf{W}_p \cap \mathbf{W}_f) &= \dim(\mathbf{W}_p) + \dim(\mathbf{W}_f) - \dim(\mathbf{W}_p + \mathbf{W}_f) = \\ &= mi + n + mi + n - 2mi + n = n \end{aligned}$$

Da equação (5.107) tem-se que  $\mathbf{X}_f = [\mathcal{O}_i^\dagger \quad -\mathcal{O}_i^\dagger \mathcal{T}] \mathbf{W}_f$ , donde o espaço linha de  $\mathbf{X}_f$  está contido no espaço linha de  $\mathbf{W}_f$ . Analogamente o espaço linha de  $\mathbf{X}_p$  está contido no espaço linha de  $\mathbf{W}_p$ . Da equação (5.109) tem-se que  $\mathbf{X}_p$  e  $\mathbf{U}_p$  determinam completamente  $\mathbf{X}_f$ ; e como o espaço linha de  $\mathbf{X}_p$  está contido no espaço linha de  $\mathbf{W}_p$ , o mesmo ocorre para o espaço linha de  $\mathbf{X}_f$ .

Na prática, devido a perturbações nos dados de entrada e saída é possível que não ocorra intersecção entre os espaços linha, sendo necessário encontrar uma intersecção aproximada, utilizando os  $n$  vetores principais da matriz. O procedimento para obter a intersecção se utiliza de uma decomposição em valores singulares da matriz  $\mathbf{W}$ :

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_f \\ \mathbf{W}_p \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \boldsymbol{\Sigma}_{11} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V}^T$$

Pode-se mostrar que  $\mathbf{U}_{21}^T \mathbf{W}_p = -\mathbf{U}_{22}^T \mathbf{W}_f$  e que o espaço linha de  $\mathbf{U}_{21}^T \mathbf{W}_p$  é igual a intersecção dos espaços linhas de  $\mathbf{W}_p$  e  $\mathbf{W}_f$ ; pois as dimensões das matrizes são:

$$\mathbf{U}_{11} = (mi + li) \times (2mi + n)$$

$$\mathbf{U}_{12} = (mi + li) \times (2li - n)$$

$$\mathbf{U}_{21} = (mi + li) \times (2mi + n)$$

$$\mathbf{U}_{22} = (mi + li) \times (2li - n)$$

$$\mathbf{\Sigma}_{11} = (2mi + n) \times (2mi + n)$$

Da propriedade de ortogonalidade da matriz  $\mathbf{U}$  tem-se:

$$\begin{bmatrix} \mathbf{U}_{11}^T & \mathbf{U}_{12}^T \\ \mathbf{U}_{21}^T & \mathbf{U}_{22}^T \end{bmatrix} \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{11}^T \mathbf{U}_{11} + \mathbf{U}_{21}^T \mathbf{U}_{21} & \mathbf{U}_{11}^T \mathbf{U}_{12} + \mathbf{U}_{21}^T \mathbf{U}_{22} \\ \mathbf{U}_{12}^T \mathbf{U}_{11} + \mathbf{U}_{22}^T \mathbf{U}_{21} & \mathbf{U}_{12}^T \mathbf{U}_{12} + \mathbf{U}_{22}^T \mathbf{U}_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

desta forma podemos escrever que:

$$\mathbf{U}_{12}^T \mathbf{U}_{11} + \mathbf{U}_{22}^T \mathbf{U}_{21} = 0$$

o que nos permite obter a partir das equações anteriores:

$$\begin{bmatrix} \mathbf{U}_{12}^T & \mathbf{U}_{22}^T \end{bmatrix} \begin{bmatrix} \mathbf{W}_p \\ \mathbf{W}_f \end{bmatrix} = 0$$

ou  $\mathbf{U}_{12}^T \mathbf{W}_p = -\mathbf{U}_{22}^T \mathbf{W}_f$ .

Então o vetor com a seqüência de estados é dado por:

$$\mathbf{X}_f = \mathbf{U}_q^T \mathbf{U}_{12}^T \mathbf{W}_p$$

sendo que a matriz  $\mathbf{U}_q$  é utilizada para a redução de  $(2li - n)$  vetores linha mutuamente dependentes; a matriz  $\mathbf{U}_q$  é obtida através de:

$$\mathbf{U}_{12}^T \mathbf{U}_{11} \mathbf{\Sigma}_{11} = \begin{bmatrix} \mathbf{U}_q \\ \mathbf{U}_q^\perp \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_q & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_q^T \\ \mathbf{V}_q^{\perp T} \end{bmatrix} \mathbf{V}^{\perp T}$$

### Um Procedimento de Modelagem de Dados no Espaço de Estado

Um algoritmo matricial explorando intersecção para a modelagem de dados no espaço de estado para sistemas multivariáveis lineares, invariantes no tempo, a partir de medidas de entrada e saída pode ser descrito e entendido nas seguintes etapas: Primeiramente um vetor com uma seqüência de vetores de estado é determinado como a intersecção dos espaços linha de duas matrizes bloco de Hankel construídas com dados de entrada - saída. A seguir as matrizes do sistema são obtidas através da solução de um problema de mínimos quadrados.

Estendendo a representação da equação (5.14) como:

$$\begin{bmatrix} \mathbf{X}_f(k+1) \\ \mathbf{Y}_f(k) \end{bmatrix} = \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{X}_f(k) \\ \mathbf{U}_f(k) \end{bmatrix} \quad (5.111)$$

como  $\mathbf{X}_f(k)$  é conhecido do item anterior, tal sistema pode ser realizado pela solução de um conjunto linear de equações.

Redefinimos as matrizes  $\mathbf{W}_p$  e  $\mathbf{W}_f$  como:

$$\mathbf{W}_p = \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \cdots & \mathbf{u}_{k+j-1} \\ \mathbf{y}_k & \mathbf{y}_{k+1} & \cdots & \mathbf{y}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \cdots & \mathbf{u}_{k+j} \\ \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \cdots & \mathbf{y}_{k+j} \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \cdots & \mathbf{u}_{k+j+i-2} \\ \mathbf{y}_{k+i-1} & \mathbf{y}_{k+i} & \cdots & \mathbf{y}_{k+j+i-2} \end{bmatrix} \quad (5.112)$$

$$\mathbf{W}_f = \begin{bmatrix} \mathbf{u}_{k+i} & \mathbf{u}_{k+i+1} & \cdots & \mathbf{u}_{k+i+j-1} \\ \mathbf{y}_{k+i} & \mathbf{y}_{k+i+1} & \cdots & \mathbf{y}_{k+i+j-1} \\ \mathbf{u}_{k+i+1} & \mathbf{u}_{k+i+2} & \cdots & \mathbf{u}_{k+i+j} \\ \mathbf{y}_{k+i+1} & \mathbf{y}_{k+i+2} & \cdots & \mathbf{y}_{k+i+j} \\ \vdots & \vdots & & \vdots \\ \mathbf{u}_{k+2i-1} & \mathbf{u}_{k+2i} & \cdots & \mathbf{u}_{k+j+2i-2} \\ \mathbf{y}_{k+2i-1} & \mathbf{y}_{k+2i} & \cdots & \mathbf{y}_{k+j+2i-2} \end{bmatrix} \quad (5.113)$$

Realizando a decomposição em valores singulares de  $\mathbf{W}$  tem-se:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_p \\ \mathbf{W}_f \end{bmatrix} = \begin{bmatrix} \mathbf{U}_{11} & \mathbf{U}_{12} \\ \mathbf{U}_{21} & \mathbf{U}_{22} \end{bmatrix} \begin{bmatrix} \mathbf{\Sigma}_{11} & 0 \\ 0 & 0 \end{bmatrix} \mathbf{V}^T$$

Definindo  $\mathbf{U}_q$  como anteriormente podemos escrever:

$$\begin{aligned} & \begin{bmatrix} \mathbf{U}_q^T \mathbf{U}_{12}^T \mathbf{U}_{(m+l+1:(i+1)(m+l),:)} \mathbf{\Sigma} \\ \mathbf{U}_{(mi+li+m+1:(m+l)(i+1),:)} \mathbf{\Sigma} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{U}_q^T \mathbf{U}_{12}^T \mathbf{U}_{(1:mi+li,:)} \mathbf{\Sigma} \\ \mathbf{U}_{(mi+li+1:mi+li+m,:)} \mathbf{\Sigma} \end{bmatrix} \end{aligned}$$

### Projeção Oblíqua de Dados em Subespaço

**Teorema 5.3** *A projeção oblíqua das saídas futuras,  $\mathbf{Y}_f$  no subespaço das saídas e entradas passadas,  $\mathbf{W}_p$ , na direção das entradas futuras,  $\mathbf{U}_f$ :*

$$\mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p = \mathcal{O}_i \mathbf{X}_f \quad (5.114)$$

*sob as seguintes condições:*

- a) A entrada  $\mathbf{U}_H$  é persistentemente excitante de ordem  $2i$  .
- b) A intersecção do espaço linha das entradas futuras,  $\mathbf{U}_f(k)$ , com o espaço linha dos estados passados,  $\mathbf{X}_p(k)$ , é vazia.

**Prova:**

Substituindo na equação (5.109) a equação (5.108) tem -se:

$$\begin{aligned}\mathbf{X}_f(k) &= \mathbf{A}^i \mathbf{X}_p(k) + \mathcal{C}_j^{\leftarrow p} \mathbf{U}_p(k) \\ &= \mathbf{A}^i [\mathcal{O}_i^\dagger \mathbf{Y}_p(k) - \mathcal{O}_i^\dagger \mathcal{T} \mathbf{U}_p(k)] + \mathcal{C}_j^{\leftarrow p} \mathbf{U}_p(k) \\ &= [\mathcal{C}_j^{\leftarrow p} - \mathbf{A}^i \mathcal{O}_i^\dagger \mathcal{T}] \mathbf{U}_p(k) + [\mathbf{A}^i \mathcal{O}_i^\dagger] \mathbf{Y}_p(k)\end{aligned}$$

Portanto,

$$\mathbf{X}_f(k) = \mathcal{L}_p \mathbf{W}_p \quad (5.115)$$

com

$$\mathcal{L}_p = [\mathbf{A}^i \mathcal{O}_i^\dagger \mid \mathcal{C}_j^{\leftarrow p} - \mathbf{A}^i \mathcal{O}_i^\dagger \mathcal{T}]$$

Utilizando a equação (5.115) pode-se reescrever (5.103) como:

$$\mathbf{Y}_f(k) = \mathcal{O}_i \mathcal{L}_p \mathbf{W}_p + \mathcal{T} \mathbf{U}_f(k)$$

e fazendo a projecção no complemento ortogonal de  $\mathbf{U}_f(k)$  tem-se:

$$\mathbf{Y}_f(k) | \mathbf{U}_f^\perp(k) = \mathcal{O}_i \mathcal{L}_p \mathbf{W}_p | \mathbf{U}_f^\perp(k) + \mathcal{T} \mathbf{U}_f(k) | \mathbf{U}_f^\perp(k)$$

donde,

$$\mathbf{Y}_f(k) | \mathbf{U}_f^\perp(k) = \mathcal{O}_i \mathcal{L}_p \mathbf{W}_p | \mathbf{U}_f^\perp(k)$$

multiplicando ambos os membros por  $[\mathbf{W}_p | \mathbf{U}_f^\perp(k)]^\dagger \mathbf{W}_p$  tem-se:

$$[\mathbf{Y}_f(k) | \mathbf{U}_f^\perp(k)] [\mathbf{W}_p | \mathbf{U}_f^\perp(k)]^\dagger \mathbf{W}_p = \mathcal{O}_i \mathcal{L}_p \mathbf{W}_p = \mathcal{O}_i \mathbf{X}_f(k) \quad (5.116)$$

da equações (5.114) e (5.116) tem -se:

$$[\mathbf{Y}_f(k) | \mathbf{U}_f^\perp(k)] [\mathbf{W}_p | \mathbf{U}_f^\perp(k)]^\dagger \mathbf{W}_p = \mathbf{Y}_{f/\mathbf{U}_f} \mathbf{W}_p = \mathcal{O}_i \mathbf{X}_f \quad (5.117)$$

Observe que a expressão,  $[\mathbf{Y}_f(k) | \mathbf{U}_f^\perp(k)] [\mathbf{W}_p | \mathbf{U}_f^\perp(k)]^\dagger \mathbf{W}_p$ , da equação (5.117) corresponde a uma projecção oblíqua, equação (4.106).

**C.Q.D.**

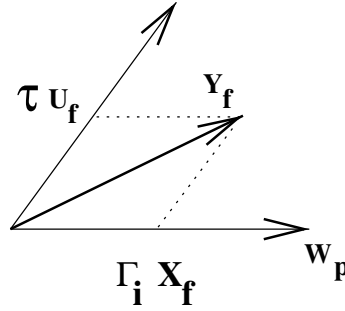


Figura 5.15: Interpretação de projeção oblíqua em espaço linha.

---

Do Teorema 5.3 pode-se concluir que:

$$\text{posto}(\mathbf{Y}_f / \mathbf{U}_f \mathbf{W}_p) = n$$

$$\Re((\mathbf{Y}_f / \mathbf{U}_f \mathbf{W}_p)^T) = \Re(\mathbf{X}_f^T(k))$$

$$\Re(\mathbf{Y}_f / \mathbf{U}_f \mathbf{W}_p) = \Re(\mathcal{O}_i)$$

Na Figura 5.15 apresentamos uma interpretação geométrica da Eq.(5.114), Van Overschee & De Moor, (1996), [1094].

## 5.8 Realização Mínima Estocástica de Sequência de Covariâncias

Considere um processo estocástico regular <sup>3</sup>, estacionário,  $l$ -dimensional com saída  $\mathbf{y}_k$  representado pelo mínimo sistema no espaço de estado  $n$  dimensional dado por:

$$\begin{aligned} \mathbf{x}_{k+1} &= \mathbf{A}\mathbf{x}_k + \mathbf{K}\mathbf{e}_k \\ \mathbf{y}_k &= \mathbf{C}\mathbf{x}_k + \mathbf{e}_k \end{aligned} \tag{5.118}$$

onde  $\mathbf{e}_k$  é uma sequência ruído branco com dimensão  $l$  e matriz de covariância dada por  $\mathbf{\Delta} = E(\mathbf{e}_k \mathbf{e}_k^T)$ . Este modelo é chamado inovativo, Aoki, [40]. A matriz de covariância para  $\mathbf{y}_k$  pode ser expressa em termos da matriz resposta ao impulso e das matrizes do modelo inovativo:

---

<sup>3</sup>Ver referências sobre processos estocásticos: Caines, [171], ; Åström & Wittenmark, [50]; Hannan & Deistler, [453].

$$\Lambda_0 = E(\mathbf{y}_k \mathbf{y}_k^T) = \sum_{k=0}^{\infty} (\mathbf{C} \mathbf{A}^{k-1} \mathbf{K}) \Delta (\mathbf{C} \mathbf{A}^{k-1} \mathbf{K})^T \quad (5.119)$$

e

$$\Lambda_i = E(\mathbf{y}_{k+1} \mathbf{y}_k^T) = \sum_{k=0}^{\infty} (\mathbf{C} \mathbf{A}^{k+i-1} \mathbf{K}) \Delta (\mathbf{C} \mathbf{A}^{k+i-1} \mathbf{K})^T, \quad (5.120)$$

para  $i = \pm 1, \pm 2, \dots$

Supondo que um conjunto de dados de saída está disponível:

$$\mathbf{y}_k, \mathbf{y}_{k-1}, \mathbf{y}_{k-2}, \dots, \mathbf{y}_{k-L+1}$$

a matriz de covariância entre as pilhas de vetores com os dados passados e com os dados futuros da série temporal tem a estrutura de uma matriz de Hankel

$$\mathbf{H}_{J,L} = E \left( \begin{bmatrix} \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \mathbf{y}_{k+3} \\ \vdots \\ \mathbf{y}_{k+J} \end{bmatrix} \begin{bmatrix} \mathbf{y}_k^T & \mathbf{y}_{k-1}^T & \dots & \mathbf{y}_{k-L+1}^T \end{bmatrix} \right) = \begin{bmatrix} \Lambda_1 & \Lambda_2 & \dots & \Lambda_L \\ \Lambda_2 & \Lambda_3 & \dots & \Lambda_{L+1} \\ \vdots & \vdots & \ddots & \vdots \\ \Lambda_J & \Lambda_{J+1} & \dots & \Lambda_{J+L-1} \end{bmatrix} \quad (5.121)$$

onde  $J$  se refere ao horizonte para o futuro.

Se o problema for obter uma representação no espaço de estado do processo estocástico este será referido como um problema de realização estocástica. Este tema será brevemente tratado nesta tese, para maior aprofundamento sugerimos: Akaike , [8]; Aoki , [40]; Caines , [171]; Faure , [338]; Lindquist & Picci , [665] Anderson , [28] e Cáceres , [169].

Para um processo estocástico estacionário, a sequência de estados  $\mathbf{x}_k$  implica em que a matriz de covariância do estado  $\mathbf{\Pi} = E(\mathbf{x}_k \mathbf{x}_k^T)$  não dependa do instante de tempo  $k$ . A partir das equações (5.118) tem-se que  $\mathbf{\Pi}$  satisfaz as relações:

$$\mathbf{\Pi} = \mathbf{A} \mathbf{\Pi} \mathbf{A}^T + \mathbf{K} \Delta \mathbf{K}^T \quad (5.122)$$

$$\mathbf{M} = \mathbf{A} \mathbf{\Pi} \mathbf{C}^T + \mathbf{K} \Delta \quad (5.123)$$

$$\Lambda_0 = \mathbf{C} \mathbf{\Pi} \mathbf{C}^T + \Delta \quad (5.124)$$

A sequência de covariâncias,  $(\Lambda_i)_{i=0}^{\infty}$ , é determinada unicamente por  $(\Lambda_0, \mathbf{A}, \mathbf{M}, \mathbf{C})$ , dado que  $\Lambda_i = \mathbf{C} \mathbf{A}^{i-1} \mathbf{M}$  para  $i \geq 1$ , com  $\mathbf{M} = E(\mathbf{x}_{k+1} \mathbf{y}_k^T)$ .

Supondo  $\Delta > 0$ ,  $\Delta$  e  $\mathbf{K}$  podem ser expressos como:

$$\Delta = \Lambda_0 - \mathbf{C}\Pi\mathbf{C}^T \quad (5.125)$$

$$\mathbf{K} = (\mathbf{M} - \mathbf{A}\Pi\mathbf{C}^T)\Delta^{-1}, \quad (5.126)$$

das equações (5.122), (5.123) e (5.124) tem-se:

$$\Pi = \mathbf{A}\Pi\mathbf{A}^T + (\mathbf{M} - \mathbf{A}\Pi\mathbf{C}^T)(\Lambda_0 - \mathbf{C}\Pi\mathbf{C}^T)^{-1}(\mathbf{M} - \mathbf{A}\Pi\mathbf{C}^T)^T, \quad (5.127)$$

Esta é a equação algébrica de Riccati para o problema de realização estocástica. Observe que  $\mathbf{K}$  e  $\Delta$  são únicas para uma dada matriz de covariância de estado  $\Pi$ , e que existe uma correspondência biunívoca entre o conjunto de soluções da equação (5.127) e o de todas as possíveis realizações  $\mathbf{A}$ ,  $\mathbf{K}$  e  $\mathbf{C}$  no espaço de estado para um sistema fixo de coordenadas que é determinado pelas matrizes  $\mathbf{A}$  e  $\mathbf{C}$ . O problema de realização estocástica pode ser decomposto nas seguintes etapas:

- a) Determinar as matrizes  $\Lambda_0$ ,  $\mathbf{A}$ ,  $\mathbf{M}$  e  $\mathbf{C}$  que representam um modelo para uma seqüência de covariâncias  $\Lambda_i$  de um conjunto de saídas  $\mathbf{y}_k$ .

Para calcular  $\Lambda_0$ ,  $\mathbf{A}$ ,  $\mathbf{M}$  e  $\mathbf{C}$  pode-se utilizar um procedimento semelhante ao utilizado para a realização determinística agora supondo que a matriz de Hankel de covariâncias possa ser fatorada como:  $\mathbf{H} = \mathcal{O}\Omega$  com  $\Omega = (\mathbf{M}, \mathbf{A}\mathbf{M}, \dots)$ .

- b) Resolver a equação de Riccati, (5.127), para obter a covariância do estado  $\Pi$ . Destacamos neste ponto um algoritmo numérico paralelo e distribuído desenvolvido por Tamariz, [1030], durante a elaboração desta tese de doutorado.
- c) Calcular  $\mathbf{K}$  e  $\Delta$  a partir de  $\Lambda_0$ ,  $\mathbf{A}$ ,  $\mathbf{M}$ ,  $\mathbf{C}$  e  $\Pi$ , pela solução das equações (5.125) e (5.126).

## 5.9 Sistemas com Entradas Exógenas: Os Subsistemas Determinístico e Estocástico

Seja o sistema estocástico linear discreto da forma geral:



$$\left\{ \begin{array}{l} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k + \mathbf{v}_k \\ \text{com} \\ \mathbf{E} \left[ \begin{pmatrix} \mathbf{w}_p \\ \mathbf{v}_p \end{pmatrix} \begin{pmatrix} \mathbf{w}_q \\ \mathbf{v}_q \end{pmatrix}^T \right] = \begin{pmatrix} \mathbf{Q} & \mathbf{S} \\ \mathbf{S}^T & \mathbf{R} \end{pmatrix} \delta_{p,q} \geq 0 . \end{array} \right. \quad (5.128)$$

onde  $\mathbf{E}$  representa o operador esperança e  $\delta_{p,q}$  é o delta de Kronecker.

Os vetores  $\mathbf{u}_k \in \Re^{m \times 1}$  e  $\mathbf{y}_k \in \Re^{l \times 1}$  são respectivamente as medidas no instante  $k$ -ésimo das  $m$  entradas e  $l$  saídas do processo.

O vetor  $\mathbf{x}_k$  representa o estado do processo no instante de tempo  $k$ ;  $\mathbf{v}_k \in \Re^{l \times 1}$  e  $\mathbf{w}_k \in \Re^{n \times 1}$  são vetores de sinais não observáveis,  $\mathbf{v}_k$  é o vetor de ruído nas medidas e  $\mathbf{w}_k$  é o vetor de ruído no processo. Supõe-se estes vetores ruídos brancos, estacionários, com média zero, e não correlatos com as entradas observáveis,  $\mathbf{u}_k$ .

$\mathbf{A} \in \Re^{n \times n}$ ;  $\mathbf{B} \in \Re^{n \times m}$ ;  $\mathbf{C} \in \Re^{l \times n}$ ;  $\mathbf{D} \in \Re^{l \times m}$ ;  $\mathbf{Q} \in \Re^{n \times n}$ ;  $\mathbf{S} \in \Re^{n \times l}$  e  $\mathbf{R} \in \Re^{l \times l}$  são matrizes de covariância da seqüências de ruídos  $\mathbf{w}_k$  e  $\mathbf{v}_k$ . Na identificação por subespaços supõe-se tipicamente que o número de dados de medidas disponíveis tende a infinito e que os dados são ergódicos.

Podemos decompor o sistema (5.128) na soma de um sistema determinístico, superescrito  $d$ , com um sistema estocástico, superescrito  $e$ ; por superposição, os estados do sistema global são dados pela soma dos estados das partes determinística e estocástica (idem para as saídas):

$$\left\{ \begin{array}{l} \mathbf{x}_k = \mathbf{x}_k^d + \mathbf{x}_k^e \\ \mathbf{y}_k = \mathbf{y}_k^d + \mathbf{y}_k^e \end{array} \right. \quad (5.129)$$

onde:

a) o subsistema determinístico:

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1}^d = \mathbf{A}\mathbf{x}_k^d + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k^d = \mathbf{C}\mathbf{x}_k^d + \mathbf{D}\mathbf{u}_k \end{array} \right. \quad (5.130)$$

descreve a influência da entrada determinística  $\mathbf{u}_k$  sobre a saída determinística  $\mathbf{y}_k^d$ .

b) o subsistema estocástico:

$$\left\{ \begin{array}{l} \mathbf{x}_{k+1}^e = \mathbf{A}\mathbf{x}_k^e + \mathbf{w}_k \\ \mathbf{y}_k^e = \mathbf{C}\mathbf{x}_k^e + \mathbf{v}_k \end{array} \right. \quad (5.131)$$

descreve a influência das seqüências de ruído  $\mathbf{w}_k$  e  $\mathbf{v}_k$  sobre a saída estocástica  $\mathbf{y}_k^e$ .

Outra importante representação de um sistema linear dinâmico estocástico: modelo de estado inovativo com entradas exógenas, é dada por:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k + \mathbf{K}\mathbf{e}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{e}_k \end{cases} \quad (5.132)$$

onde  $\mathbf{K} \in \Re^{n \times l}$  e a inovação  $\mathbf{e}_k$  é um vetor ruído branco  $l$ -dimensional não correlato com a entrada observável  $\mathbf{u}_k$ .

Um sistema com entradas exógenas, Figura 5.16 a), pode ser representado como na Figura 5.16 b), onde  $\mathbf{G}(k, j)$  e  $\mathbf{L}(k, j)$  são as respectivas matrizes resposta ao impulso unitário para entradas  $\mathbf{u}_k$  e  $\mathbf{e}_k$  e de modo semelhante à equação (5.11) podemos expressar  $\mathbf{y}_k$  por:

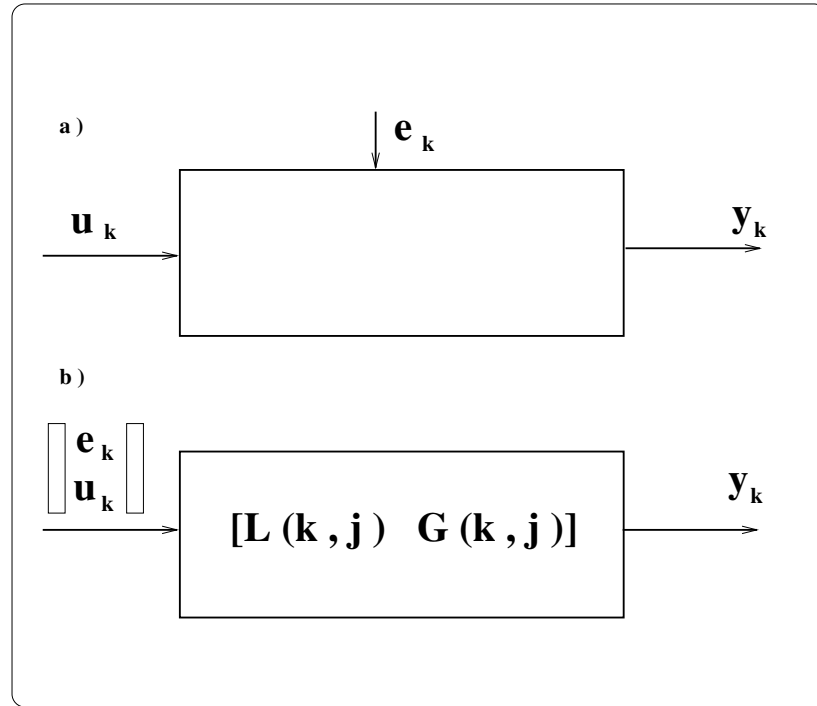


Figura 5.16: Representações de Sistemas com Entradas Exógenas.

$$\mathbf{y}_k = \sum_{j=k_0}^k \mathbf{L}(k, j)\mathbf{e}(j) + \sum_{j=k_0}^k \mathbf{G}(k, j)\mathbf{u}(j) = \sum_{j=k_0}^k \begin{bmatrix} \mathbf{L}(k, j) & \mathbf{G}(k, j) \end{bmatrix} \begin{bmatrix} \mathbf{e}(j) \\ \mathbf{u}(j) \end{bmatrix} \quad (5.133)$$

para  $k \geq k_0$ .

Reescrevendo as equações (5.132) como:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \begin{bmatrix} \mathbf{K} & \mathbf{B} \end{bmatrix} \begin{bmatrix} \mathbf{e}_k \\ \mathbf{u}_k \end{bmatrix} \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \begin{bmatrix} \mathbf{I}_s & 0 \end{bmatrix} \begin{bmatrix} \mathbf{e}_k \\ \mathbf{u}_k \end{bmatrix} \end{cases} \quad (5.134)$$

tem-se para a matriz resposta ao impulso  $[\mathbf{L}(k, j) \quad \mathbf{G}(k, j)]$  os respectivos parâmetros de Markov:

$$[\mathbf{L}(k, j) \quad \mathbf{G}(k, j)] = \begin{cases} [\mathbf{I}_s \quad 0] & k = j \\ \mathbf{C}\mathbf{A}^{k-j-1}[\mathbf{K} \quad \mathbf{B}] & k \geq j + 1 \end{cases}$$

Como consequência dos desenvolvimentos deste capítulo podemos apresentar os seguintes resultados para sistemas multivariáveis com entradas exógenas:

- Para  $\mathbf{x}_k = 0$ :

$$\mathbf{y}_k^+ = \begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \vdots \\ \mathbf{y}_{k+n-1} \end{bmatrix} = \begin{bmatrix} [\mathbf{I}_s \quad 0] & 0 & 0 & \dots & 0 \\ \mathbf{C}[\mathbf{K} \quad \mathbf{B}] & [\mathbf{I}_s \quad 0] & 0 & & \vdots \\ \mathbf{C}\mathbf{A}[\mathbf{K} \quad \mathbf{B}] & \mathbf{C}[\mathbf{K} \quad \mathbf{B}] & [\mathbf{I}_s \quad 0] & & \vdots \\ \vdots & \vdots & & \ddots & \\ \mathbf{C}\mathbf{A}^{n-2}[\mathbf{K} \quad \mathbf{B}] & \mathbf{C}\mathbf{A}^{n-3}[\mathbf{K} \quad \mathbf{B}] & & \dots & [\mathbf{I}_s \quad 0] \end{bmatrix} \begin{bmatrix} \begin{bmatrix} \mathbf{e}_k \\ \mathbf{u}_k \end{bmatrix} \\ \begin{bmatrix} \mathbf{e}_{k+1} \\ \mathbf{u}_{k+1} \end{bmatrix} \\ \begin{bmatrix} \mathbf{e}_{k+2} \\ \mathbf{u}_{k+2} \end{bmatrix} \\ \vdots \\ \begin{bmatrix} \mathbf{e}_{k+n-1} \\ \mathbf{u}_{k+n-1} \end{bmatrix} \end{bmatrix} \quad (5.135)$$

- **Teorema 5.4** *O sistema (5.132) é mínimo se e somente se  $(\mathbf{A}, \mathbf{C})$  é observável e  $(\mathbf{A}, [\mathbf{K} \ \mathbf{B}])$  é atingível.*

A prova deste teorema é apresentada em Hannan, E.J. & Deistler, [452].

- **Definição 5.2** *Uma matriz infinita  $\mathbf{H}$  bloco de Hankel de uma matriz resposta ao impulso  $[\mathbf{L} \ \mathbf{G}]$  é:*

$$\mathbf{H} = \begin{bmatrix} [\mathbf{L}_1 \ \mathbf{G}_1] & [\mathbf{L}_2 \ \mathbf{G}_2] & [\mathbf{L}_3 \ \mathbf{G}_3] & \cdots \\ [\mathbf{L}_2 \ \mathbf{G}_2] & [\mathbf{L}_3 \ \mathbf{G}_3] & [\mathbf{L}_4 \ \mathbf{G}_4] & \cdots \\ [\mathbf{L}_3 \ \mathbf{G}_3] & [\mathbf{L}_4 \ \mathbf{G}_4] & \cdots & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \quad (5.136)$$

- Analogamente definimos as matrizes Toeplitz respectivamente para  $\mathbf{u}_k$  e  $\mathbf{e}_k$  como:

$$\mathcal{T}_u = \begin{bmatrix} 0 & 0 & 0 & \cdots \\ \mathbf{G}_1 & 0 & 0 & \cdots \\ \mathbf{G}_2 & \mathbf{G}_1 & 0 & \cdots \\ \vdots & \vdots & \cdots & \ddots \end{bmatrix}, \quad \mathcal{T}_e = \begin{bmatrix} \mathbf{I}_s & 0 & 0 & \cdots \\ \mathbf{L}_1 & \mathbf{I}_s & 0 & \cdots \\ \mathbf{L}_2 & \mathbf{L}_1 & \mathbf{I}_s & \cdots \\ \vdots & \vdots & \cdots & \ddots \end{bmatrix}, \quad (5.137)$$

e podemos escrever:

$$\begin{bmatrix} \mathbf{y}_k \\ \mathbf{y}_{k+1} \\ \mathbf{y}_{k+2} \\ \vdots \end{bmatrix} = \mathbf{H} \begin{bmatrix} \mathbf{e}_{k-1} \\ \mathbf{u}_{k-1} \\ \mathbf{e}_{k-2} \\ \vdots \end{bmatrix} + \mathcal{T}_u \begin{bmatrix} \mathbf{u}_k \\ \mathbf{u}_{k+1} \\ \mathbf{u}_{k+2} \\ \vdots \end{bmatrix} + \mathcal{T}_e \begin{bmatrix} \mathbf{e}_k \\ \mathbf{e}_{k+1} \\ \mathbf{e}_{k+2} \\ \vdots \end{bmatrix}$$

donde

$$\mathbf{y}_k^+ = \mathbf{H} \begin{bmatrix} \mathbf{e}_{k-1}^- \\ \mathbf{u}_{k-1}^- \end{bmatrix} + [\mathcal{T}_e \ \mathcal{T}_u] \begin{bmatrix} \mathbf{e}_k^+ \\ \mathbf{u}_k^+ \end{bmatrix} \quad (5.138)$$

- Para  $\mathbf{u}_k = 0$  os modelos inovativos são tais que o ruído  $\mathbf{e}_k$  é igual a  $\mathbf{y}_k - \mathbf{y}_{k|k-1}$  onde  $\mathbf{y}_{k|k-1}$  é a projecção ortogonal de  $\mathbf{y}_k$  no subespaço formado por  $\mathbf{y}_{k-1}, \mathbf{y}_{k-2}, \dots, \mathbf{y}_1$ . Eles são completamente especificados somente quando as matrizes  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{K}$  e a matriz de covariância das inovações,  $\Delta$ , são determinadas.

- Para identificação determinística e ou estocástica, de sistemas multivariáveis no espaço de estado, pode-se utilizar técnicas de fatoração na estrutura (5.136).
- Para séries temporais os modelos de estado podem ser determinados de pelo menos duas formas:
  - i) Usar a estrutura (5.136) para obter as matrizes do sistema.
  - ii) Usar a estrutura (5.121) de modo similar para o mesmo propósito.

A semelhança entre os itens i) e ii) pode ser observada notando que:

$$\mathbf{G}_i = \mathbf{C}\mathbf{A}^{i-1}\mathbf{B}$$

e

$$\mathbf{\Lambda}_i = \mathbf{C}\mathbf{A}^{i-1}\mathbf{M}$$

O item i) usa informações das seqüências de matrizes resposta ao impulso (estimadas) e o item ii) informações das matrizes de covariâncias das amostras.

- Para a seqüência de estados:

$$\mathbf{X}_i \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{x}_i & \mathbf{x}_{i+1} & \mathbf{x}_{i+2} & \dots & \mathbf{x}_{i+j-1} \end{bmatrix}$$

definimos respectivamente as seqüências de estado determinística  $\mathbf{X}_i^d$  e estocásticas  $\mathbf{X}_i^e$  como:

$$\mathbf{X}_i^d \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{x}_i^d & \mathbf{x}_{i+1}^d & \mathbf{x}_{i+2}^d & \dots & \mathbf{x}_{i+j-1}^d \end{bmatrix}$$

$$\mathbf{X}_i^e \stackrel{\text{def}}{=} \begin{bmatrix} \mathbf{x}_i^e & \mathbf{x}_{i+1}^e & \mathbf{x}_{i+2}^e & \dots & \mathbf{x}_{i+j-1}^e \end{bmatrix}$$

e de forma análoga à que fizemos para as equações (5.105) e (5.106) para as seqüências passada e futura de estados determinísticos e estocásticos definimos:

$$\begin{aligned} \mathbf{X}_p^d &= \mathbf{X}_0^d & , & & \mathbf{X}_f^d &= \mathbf{X}_i^d \\ \mathbf{X}_p^e &= \mathbf{X}_0^e & , & & \mathbf{X}_f^e &= \mathbf{X}_i^e \end{aligned} \quad (5.139)$$

- Para o sistema com entradas exógenas, que também chamaremos de sistema combinado determinístico-estocástico, podemos escrever modelos estendidos; em particular, para o sistema combinado representado pela equação (5.132) o modelo estendido MEEE1 pode ser escrito como:

$$\mathbf{Y}_f = \mathcal{O}_i \mathbf{X}_f^d + [\mathcal{T}_u \ 0] \mathbf{U}_f + [\mathcal{T}_e \ 0] \mathbf{E}_{fe}(k) \quad (5.140)$$

e de forma análoga:

$$\mathbf{Y}_p = \mathcal{O}_i \mathbf{X}_p^d + [\mathcal{T}_u \ 0] \mathbf{U}_p + [\mathcal{T}_e \ 0] \mathbf{E}_{pe}(k) \quad (5.141)$$

onde os subíndices  $p$  e  $f$  denotam passado e futuro respectivamente e  $\mathbf{E}_{pe}$  e  $\mathbf{E}_{fe}$  denotam matrizes com vetores inovação:

$$\mathbf{E}_{2H} = \left[ \frac{\mathbf{E}_{pe}}{\mathbf{E}_{fe}} \right] = \begin{bmatrix} \mathbf{e}_0 & \mathbf{e}_1 & \dots & \mathbf{e}_{j-1} \\ \mathbf{e}_1 & \mathbf{e}_2 & \dots & \mathbf{e}_j \\ \mathbf{e}_2 & \mathbf{e}_3 & \dots & \mathbf{e}_{j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_{i-1} & \mathbf{e}_i & \dots & \mathbf{e}_{j+i-2} \\ \hline \mathbf{e}_i & \mathbf{e}_{i+1} & \dots & \mathbf{e}_{i+j-1} \\ \mathbf{e}_{i+1} & \mathbf{e}_{i+2} & \dots & \mathbf{e}_{i+j} \\ \mathbf{e}_{i+2} & \mathbf{e}_{i+3} & \dots & \mathbf{e}_{i+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{e}_{2i-1} & \mathbf{e}_{2i} & \dots & \mathbf{e}_{2i+j-2} \end{bmatrix} \quad (5.142)$$

- De forma análoga à equação (5.109) tem-se:

$$\mathbf{X}_f^d = \mathbf{A}^i \mathbf{X}_p^d + \mathcal{C}_j^{\leftarrow p d} \mathbf{U}_p \quad (5.143)$$

onde

$$\mathcal{C}_j^{\leftarrow p d} = \begin{bmatrix} \mathbf{A}^{i-1} \mathbf{B} & \mathbf{A}^{i-2} \mathbf{B} & \dots & \mathbf{A} \mathbf{B} & \mathbf{B} \end{bmatrix} \quad (5.144)$$

- O termo  $[\mathcal{T}_u \ 0] \mathbf{U}_f$  pode ser removido da equação (5.140) utilizando uma projeção  $P$  de tal forma que:  $\mathbf{U}_f P = 0$ . Desta maneira,  $P$  é um projetor ortogonal no espaço nulo de  $\mathbf{U}_f$ .
- Para eliminar o termo desconhecido  $\mathbf{E}_{fe}$  da equação (5.140) pode-se recorrer à uma variável instrumental dado que o ruído é não correlato com a entrada e nem com a saída:

$$\lim_{k \rightarrow \infty} \frac{1}{k} \mathbf{E}_{fe} \mathbf{W}_f^T = 0 \quad (5.145)$$

A variável instrumental  $\mathbf{W}_f$  é dada por:

$$\mathbf{W}_f = \begin{bmatrix} \mathbf{Y}_f \\ \mathbf{U}_f \end{bmatrix} \quad (5.146)$$

- Supondo o par  $\mathbf{A}$ ,  $\mathbf{Q}^{1/2}$  atingível definimos as seguintes matrizes Toeplitz de covariâncias de saídas:

a)

$$\mathbf{C}_i = \Phi_{[\mathbf{Y}_f, \mathbf{Y}_p]} \quad (5.147)$$

b)

$$\mathbf{L}_i = \Phi_{[\mathbf{Y}_f, \mathbf{Y}_f]} = \Phi_{[\mathbf{Y}_p, \mathbf{Y}_p]} \quad (5.148)$$

- O procedimento descrito a partir da equação (5.140) é frequentemente usado nos métodos de identificação de sistemas em subespaços, dos quais destacamos:
  - *CVA - Canonical Variate Analysis*, Larimore, [641].
  - *N4SID - Numerical algorithms for subspace state space system identification*, Van Overschee & De Moor, [1090] e [1091].
  - *MOESP - MIMO Output-Error State Space model identification*, Verhaegen & Dewilde (1992), [1113].
  - *ERA - Eigensystem Realization Algorithm*, Juang & Pappa [549].
  - *DSR - Combined Deterministic and Stochastic System Identification and Realization*, Ruscio, [295].

- **Estimação de Estados Através de Dados Entrada-Saída de Sistema Combinado Determinístico e Estocástico**

- a) **Estados e Filtro de Kalman**

Na obtenção de algoritmos para modelagem de dados em subespaços para sistema combinado determinístico-estocástico, o filtro de Kalman pode ter papel fundamental. Nesta seção é apresentada uma forma de equação para o regime transitório onde uma sequência de estados estimados para o sistema combinado é gerada através de um banco de filtros de Kalman. Denotamos por  $\hat{\mathbf{x}}_k$  a estimativa do estado  $\mathbf{x}_k$ .

**Teorema 5.5** *Sejam:*

- a) *A estimativa do estado inicial:  $\hat{\mathbf{x}}_0$*
- b) *A matriz de covariância para a estimativa do estado inicial:*  
 $\mathbf{\Pi}_0 = E[\hat{\mathbf{x}}_0 \hat{\mathbf{x}}_0^T]$

c) *As medidas de entradas e saídas:  $\mathbf{y}_0, \mathbf{u}_0, \mathbf{y}_1, \mathbf{u}_1, \dots, \mathbf{y}_{k-1}, \mathbf{u}_{k-1}$  então a estimativa do estado  $\hat{\mathbf{x}}_k$  é dada pelo filtro de Kalman recursivo:*

$$\hat{\mathbf{x}}_k = \mathbf{A}\hat{\mathbf{x}}_{k-1} + \mathbf{B}\mathbf{u}_{k-1} + \mathbf{K}_{k-1}(\mathbf{y}_{k-1} - \mathbf{C}\hat{\mathbf{x}}_{k-1} - \mathbf{D}\mathbf{u}_{k-1}) \quad (5.149)$$

$$\mathbf{K}_{k-1} = (\mathbf{M} - \mathbf{A}\mathbf{\Pi}_{k-1}\mathbf{C}^T)(\mathbf{\Lambda}_0 - \mathbf{C}\mathbf{\Pi}_{k-1}\mathbf{C}^T)^{-1} \quad (5.150)$$

$$\mathbf{\Pi}_k = \mathbf{A}\mathbf{\Pi}_{k-1}\mathbf{A}^T + (\mathbf{M} - \mathbf{A}\mathbf{\Pi}_{k-1}\mathbf{C}^T)(\mathbf{\Lambda}_0 - \mathbf{C}\mathbf{\Pi}_{k-1}\mathbf{C}^T)^{-1}(\mathbf{M} - \mathbf{A}\mathbf{\Pi}_{k-1}\mathbf{C}^T)^T \quad (5.151)$$

*que pode ser escrito explicitamente como:*

$$\hat{\mathbf{x}}_k = (\mathbf{A}^k - \Omega_k \mathcal{O}_k \mid \mathcal{C}_k^{\epsilon pe} - \Omega_k \mathcal{T}_u \mid \Omega_k) \begin{bmatrix} \hat{\mathbf{x}}_0 \\ \mathbf{u}_0 \\ \vdots \\ \mathbf{u}_{k-1} \\ \mathbf{y}_0 \\ \vdots \\ \mathbf{y}_{k-1} \end{bmatrix} \quad (5.152)$$

*onde:*

$$\Omega_k = (\mathcal{C}_k^{\epsilon pe} - \mathbf{A}^k \mathbf{\Pi}_0 \mathcal{O}_k^T)(\mathbf{H}_k - \mathcal{O}_k \mathbf{\Pi}_0 \mathcal{O}_k^T)^{-1} \quad (5.153)$$

*e:*

$$\mathcal{C}_k^{\epsilon pe} = \begin{bmatrix} \mathbf{A}^{k-1}\mathbf{M} & \mathbf{A}^{k-2}\mathbf{M} & \dots & \mathbf{A}\mathbf{M} & \mathbf{M} \end{bmatrix} \quad (5.154)$$

*é a matriz de atingibilidade reordenada para o sistema estocástico e  $\mathbf{H}_k$  é uma matriz de covariâncias da saída:*

$$\mathbf{H}_k = \begin{bmatrix} \mathbf{\Lambda}_0 & \mathbf{\Lambda}_{-1} & \dots & \mathbf{\Lambda}_{1-k} \\ \mathbf{\Lambda}_1 & \mathbf{\Lambda}_0 & \dots & \mathbf{\Lambda}_{2-k} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{\Lambda}_{k-1} & \mathbf{\Lambda}_{k-2} & \dots & \mathbf{\Lambda}_0 \end{bmatrix}$$

*com  $\mathbf{\Lambda}_i = E(\mathbf{y}_{k+i}\mathbf{y}_k^T)$ ,  $\mathbf{M} = E(\mathbf{x}_k\mathbf{y}_k^T)$  e a matriz de covariância do estado  $\mathbf{\Pi}_k = E(\mathbf{x}_k\mathbf{x}_k^T)$ .*

*A solução explícita da matriz  $\mathbf{\Pi}_k$  é igual a:*



$$\mathbf{\Pi}_k = \mathbf{A}_k \mathbf{\Pi}_0 (\mathbf{A}^T)^k + (\mathcal{C}_k^{\epsilon p e} - \mathbf{A}^k \mathbf{\Pi}_0 \mathcal{O}_k^T) (\mathbf{L}_k - \mathcal{O}_k \mathbf{\Pi}_0 \mathcal{O}_k^T)^{-1} (\mathcal{C}_k^{\epsilon p e} - \mathbf{A}^k \mathbf{\Pi}_0 \mathcal{O}_k^T)^T \quad (5.155)$$

onde  $\mathbf{L}_k$  é uma matriz de covariâncias de saídas, equação (5.148).

**Prova:** A prova é obtida por indução e pode ser encontrada em Van Overschee & De Moor, [1094].

**C.Q.D.**

Este teorema indica que a estimativa dos estado através do filtro de Kalman,  $\hat{\mathbf{x}}_k$ , pode ser escrita como uma combinação de medidas de entradas-saídas passadas  $\mathbf{y}_0, \mathbf{u}_0, \mathbf{y}_1, \mathbf{u}_1, \dots, \mathbf{y}_{k-1}, \mathbf{u}_{k-1}$  e de uma estimativa do estado inicial,  $\hat{\mathbf{x}}_0$ .

Esta observação permite ver que a seqüência de estados pode ser estimada pela combinação de algoritmos de modelagem de dados determinístico e estocástico em subespaços pois:

$$\begin{aligned} \hat{\mathbf{X}}_f &= (\hat{\mathbf{x}}_i \ \hat{\mathbf{x}}_{i+1} \ \dots \ \hat{\mathbf{x}}_{i+j+1}) \\ &= (\mathbf{A}^i - \Omega_i \mathcal{O}_i \mid \mathcal{C}_i^{\epsilon p e} - \Omega_i \mathcal{T}_u \mid \Omega_i) \begin{bmatrix} \hat{\mathbf{X}}_p \\ \mathbf{U}_p \\ \mathbf{Y}_p \end{bmatrix} \\ &= (\mathbf{A}^i - \Omega_i \mathcal{O}_i \mid \mathcal{C}_i^{\epsilon p e} - \Omega_i \mathcal{T}_u \mid \Omega_i) \begin{bmatrix} \hat{\mathbf{X}}_p \\ \mathbf{W}_p \end{bmatrix} \end{aligned} \quad (5.156)$$

Esta seqüência de estados é gerada por um banco de filtros de Kalman não estacionários, trabalhando em paralelo cada uma das colunas de uma matriz bloco de Hankel com entradas e saídas passadas,  $\mathbf{W}_p$ .

Na Figura 5.17 temos uma representação da estimação da seqüência de estados  $\hat{\mathbf{X}}_f$  a partir das medidas de entrada-saída e da seqüência de estados  $\hat{\mathbf{X}}_p$  utilizando filtros de Kalman. Quando as matrizes,  $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}, \mathbf{Q}, \mathbf{R}$  e  $\mathbf{S}$  do sistema são conhecidas, o estado  $\hat{\mathbf{x}}_{i+q}$  pode ser determinado através de um filtro de Kalman não estacionário baseado em  $i$  medidas de  $\mathbf{u}_k$  e  $\mathbf{y}_k$ , no estado inicial  $\hat{\mathbf{x}}_q$  e na matriz de covariância do erro inicial:

$$\tilde{\mathbf{\Pi}}_0 = \Sigma^e - \mathbf{\Pi}_0, \quad \Sigma^e \stackrel{\text{def}}{=} E[\hat{\mathbf{x}}_k \hat{\mathbf{x}}_k^T] \quad (5.157)$$

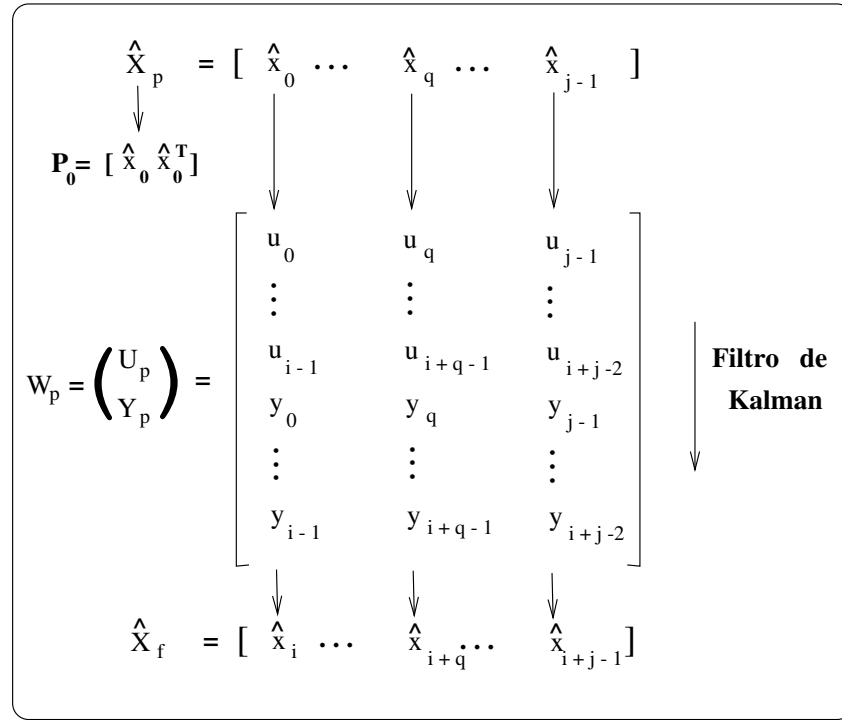


Figura 5.17: Representação da Estimação da Seqüência de Estados Através de um Banco de Filtros de Kalman.

Iterando o  $q$ -ésimo filtro de Kalman recursivo sobre  $i$  passos obter-se-ia a estimativa do estado  $\hat{x}_{i+q}$ . O mesmo procedimento pode ser feito para todas as  $j$  colunas e então dizemos ter um banco de  $j$  filtros de Kalman em paralelo.

Contudo, para algoritmos de subespaços é importante notar que as matrizes  $A$ ,  $B$ ,  $C$ ,  $D$ ,  $Q$ ,  $R$  e  $S$  não precisam ser conhecidas para determinar a seqüência de estados,  $\hat{X}_f$ . Esta seqüência de estados pode ser determinada diretamente através dos dados de entrada-saída, Teorema 5.7.

A modelagem de dados no espaço de estado é realizada a partir de  $i$  medidas de entrada-saída e de um estado inicial conhecido. A Figura 5.18 mostra uma representação de métodos tradicionais e métodos de subespaços. Observe que rodando no sentido anti-horário a partir das medidas de entrada-saída podemos obter a seqüência de estados,  $\hat{X}_f$ , utilizando projeções oblíquas ou ortogonais e as matrizes do sistema podem ser obtidas através de técnicas de mínimos quadrados. Rodando no sentido horário primeiramente as matrizes do sistema são obtidas para posterior-

mente obtermos os estados.

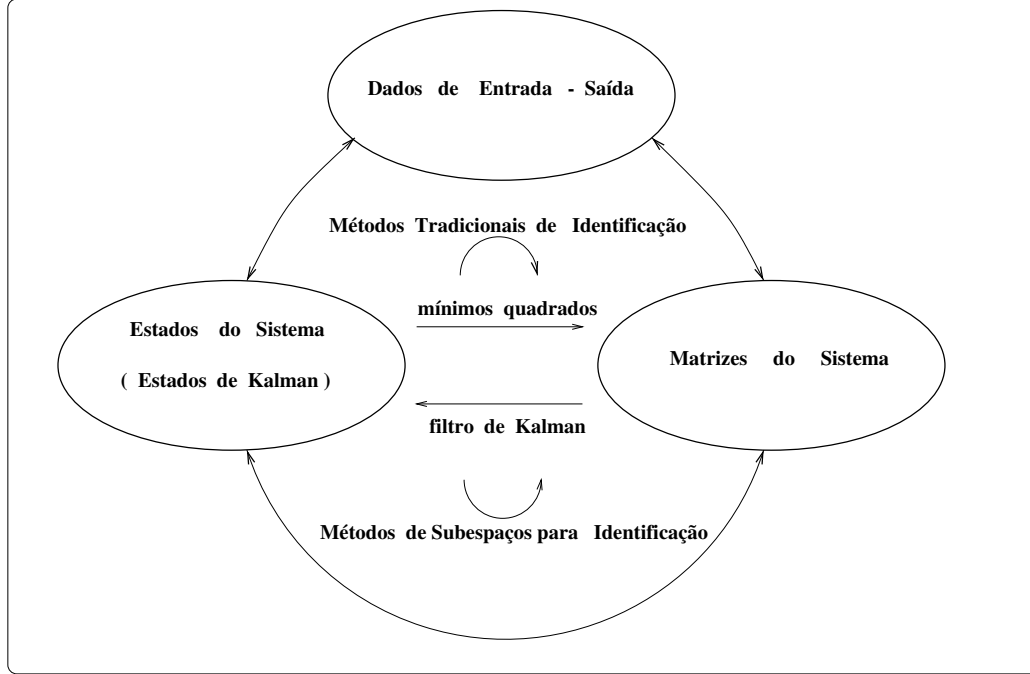


Figura 5.18: Comparação entre Métodos para Modelagem de Dados.

#### b) **Projeção Ortogonal**

Através de uma projeção ortogonal é possível relacionar os estados do filtro de Kalman com os dados de entrada-saída.

**Teorema 5.6** *A projeção ortogonal de saídas futuras sobre entradas-saídas passadas e entradas futuras é:*

$$\mathbf{Y}_f | \begin{bmatrix} \mathbf{W}_p \\ \mathbf{U}_f \end{bmatrix} = \mathcal{O}_i \hat{\mathbf{X}}_i + \mathcal{T} \mathbf{U}_f \quad (5.158)$$

onde

$$\hat{\mathbf{X}}_0 = \Phi^{xu} (\Phi^{uu})^{-1} \begin{bmatrix} \mathbf{U}_p \\ \mathbf{U}_f \end{bmatrix} \quad (5.159)$$

$$\Pi_0 = -[\Phi^{xx} - \Phi^{xu} (\Phi^{uu})^{-1} (\Phi^{xu})^T] \quad (5.160)$$

com

$$\Phi^{uu} \stackrel{\text{def}}{=} E[\mathbf{U}_{2H} \mathbf{U}_{2H}^T]$$

$$\Phi^{xu} \stackrel{\text{def}}{=} E[\mathbf{X}_p^d \mathbf{U}_{2H}]$$

$$\Phi^{xx} \stackrel{\text{def}}{=} E[\mathbf{X}_p^d \mathbf{X}_p^d]$$

sob as seguintes condições:

- a) A entrada  $\mathbf{u}_k$  é não correlata com os ruídos  $\mathbf{w}_k$  e  $\mathbf{v}_k$ .
- b) A entrada determinística  $\mathbf{u}_k$  é persistentemente excitante de ordem  $2i$ .
- c) Um número infinito de medidas está disponível,  $j \rightarrow \infty$ .
- d) O ruído do processo  $\mathbf{w}_k$  e o ruído de medida  $\mathbf{v}_k$  são não nulos.

**Prova:** A prova pode ser encontrada no Apêndice 6 do livro de Van Overschee & De Moor, [1094].

**C.Q.D.**

A importância do Teorema 5.6 é revelar uma forma pela qual a seqüência de estados do filtro de Kalman,  $\hat{\mathbf{X}}_f$ , está relacionada com os dados de entrada-saída. A projeção ortogonal pode ser computada dos dados sem o conhecimento das matrizes do sistema. Através das equações (5.159) e (5.160) tem-se:

$$\hat{\mathbf{X}}_0 = \mathbf{X}_p \left| \begin{bmatrix} \mathbf{U}_p \\ \mathbf{U}_f \end{bmatrix} \right. \quad (5.161)$$

$$\Pi_0 = -\Phi \left[ \mathbf{X}_p \left| \begin{bmatrix} \mathbf{U}_p \\ \mathbf{U}_f \end{bmatrix} \right|^\perp, \mathbf{X}_p \left| \begin{bmatrix} \mathbf{U}_p \\ \mathbf{U}_f \end{bmatrix} \right|^\perp \right] \quad (5.162)$$

O estado inicial “ real ” seria  $\mathbf{X}_p^d + \mathbf{X}_p^e$ ; no entanto a parte estocástica do estado inicial,  $\mathbf{X}_p^e$ , impossível de ser estimada, é zerada. Isto explica o fato de  $\mathbf{X}_0^e$  não aparecer na equação (5.161). Note que o espaço linha de  $\hat{\mathbf{X}}_0$ , equação (5.161), deve estar na combinação dos espaços linha de  $\mathbf{U}_p$  e  $\mathbf{U}_f$ , e é a melhor estimativa de  $\mathbf{X}_p^d$ , encontrada no espaço linha das entradas passadas e futuras.

### c) Projeção Oblíqua

O Teorema 5.7 permite o cálculo do espaço linha da seqüência de estado de um filtro de Kalman e do espaço coluna de uma matriz de observabilidade estendida,  $\mathcal{O}_i$  diretamente dos dados de entrada-saída sem necessidade do conhecimento das matrizes do

sistema. As matrizes do sistema podem ser obtidas posteriormente através da sequência de estado,  $\hat{\mathbf{X}}_f$  ou de  $\mathcal{O}_i$ . Um resumo dos Teoremas 5.6 e 5.7 é apresentado na Figura 5.19

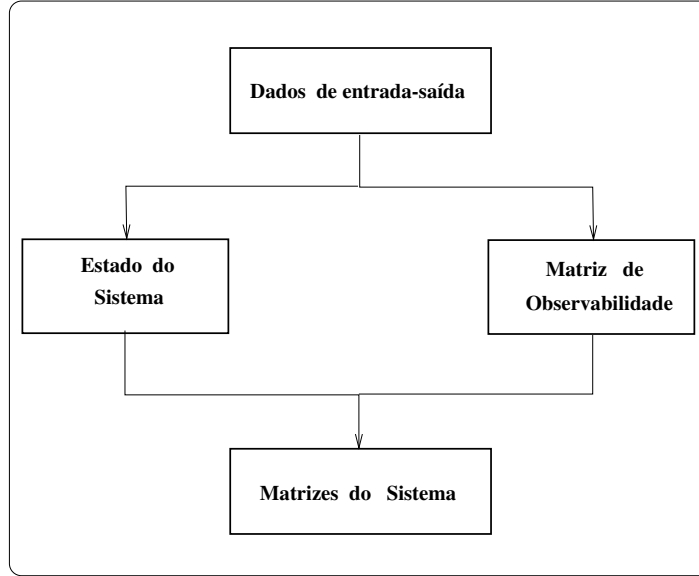


Figura 5.19: Representação dos Teoremas 5.6 e 5.7.

**Teorema 5.7** *A projeção oblíqua das saídas futuras,  $\mathbf{Y}_f$  no subespaço das saídas e entradas passadas,  $\mathbf{W}_p$ , na direção das entradas futuras,  $\mathbf{U}_f$ :*

$$\mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p = \mathcal{O}_i \hat{\mathbf{X}}_f = [\mathbf{U}_1 \mid \mathbf{U}_2] \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \mathbf{V}_1^T \\ \mathbf{V}_2^T \end{bmatrix} \quad (5.163)$$

onde

$$\hat{\mathbf{X}}_0 = \mathbf{X}_p /_{\mathbf{U}_f} \mathbf{U}_p \quad (5.164)$$

$$\mathbf{\Pi}_0 = -[\mathbf{\Phi}^{xx} - \mathbf{\Phi}^{xu}(\mathbf{\Phi}^{uu})^{-1}(\mathbf{\Phi}^{xu})^T] \quad (5.165)$$

sob as seguintes condições:

- A entrada  $\mathbf{u}_k$  é não correlata com os ruídos  $\mathbf{w}_k$  e  $\mathbf{v}_k$ .
- A entrada determinística  $\mathbf{u}_k$  é persistentemente excitante de ordem  $2i$ .
- Um número infinito de medidas está disponível,  $j \rightarrow \infty$ .

d) O ruído do processo  $\mathbf{w}_k$  e o ruído de medida  $\mathbf{v}_k$  são não nulos.  
tem-se:

a) A ordem do sistema é igual ao número de valores singulares não nulos da equação (5.163).

b) A matriz de observabilidade estendida  $\mathcal{O}_i$  é dada por:

$$\mathcal{O}_i = \mathbf{U}_1 \Sigma_1^{1/2} \quad (5.166)$$

c) A seqüência de estados,  $\hat{\mathbf{X}}_f$ , é:

$$\hat{\mathbf{X}}_f = \mathcal{O}_i^\dagger \mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p \quad (5.167)$$

**Prova:** A prova pode ser encontrada no Apêndice 7 do livro de Van Overschee & De Moor, [1094].

**C.Q.D.**

Do Teorema 5.7 pode-se concluir que:

$$\text{posto}(\mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p) = n$$

$$\Re((\mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p)^T) = \Re(\mathbf{X}_f^T)$$

$$\Re(\mathbf{Y}_f /_{\mathbf{U}_f} \mathbf{W}_p) = \Re(\mathcal{O}_i)$$

A seqüência de estados,  $\hat{\mathbf{X}}_f$  obtida através do Teorema 5.7 é diferente da seqüência de estados,  $\hat{\mathbf{X}}_f$  obtida através do Teorema 5.6. As duas seqüências são diferentes pois o estado inicial  $\hat{\mathbf{X}}_0$  utilizado em cada Teorema é diferente.

## Capítulo 6

# IDENTIFICAÇÃO MULTIVARIADA PARALELA E DISTRIBUÍDA VIA TRATAMENTO MOESP

### 6.1 Introdução

A representação de um sistema em variáveis no espaço de estado é muito conveniente em engenharia pois processos industriais multivariados podem ser descritos com exatidão por esta classe de modelos, além disso, hoje, as ferramentas de projeto de sistemas de controle que estão disponíveis estão baseadas neste tipo de tratamento. Por esta razão a classe de modelos no espaço de estado vem a ser muito interessante.

O algoritmo aqui implementado é classificado como um esquema de identificação de modelo em subespaço do espaço de estado a partir de dados de entrada - saída. Esta estrutura é baseada no emprego da teoria da realização. Características comuns na organização destes algoritmos são a execução de uma fatoração, por exemplo **QR**, em seguida uma decomposição em valores singulares, seguida pela solução de um conjunto superdeterminado de equações. Um tratamento computacional de alto desempenho para modelagem de dados baseado no algoritmo *MOESP - MIMO Output-Error State Space model identification* de Verhaegen & Dewilde, [1113], é realizado neste capítulo.

O esquema apresentado pressupõe que dispomos de sequências de medidas de entrada - saída.

Na Figura 6.1 temos a representação esquemática do sistema dinâmico que pretendemos modelar. As perturbações podem ser consideradas prove-

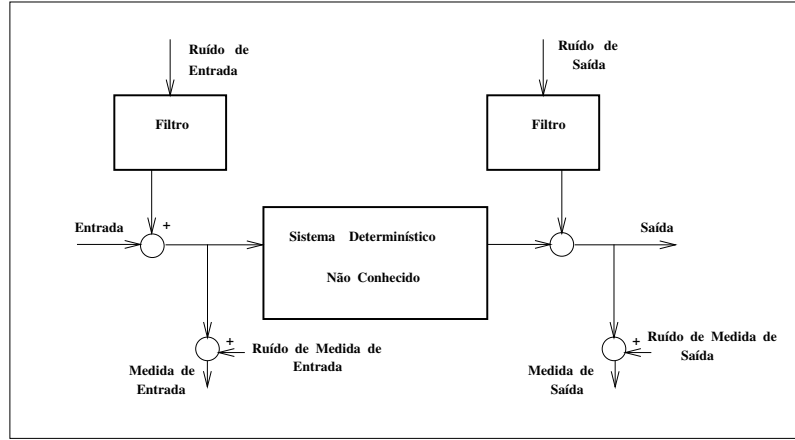


Figura 6.1: Diagrama Esquemático Geral de Sistema a Ser Modelado.

nientes de dois tipos de fonte:

- 1) Interno: devido a efeitos parasitas ou a entradas não medidas no sistema, representado na Figura 6.1 pelo ruído de saída.
- 2) Externo: seja devido aos equipamentos de medida e transdutores, representado pelos ruídos de medida de entrada e de saída, seja devido a ruídos de entrada.

Matematicamente o sistema dinâmico determinístico é descrito pelo seguinte conjunto de equações:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{B}\mathbf{u}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{D}\mathbf{u}_k \end{cases} \quad (6.1)$$

onde os vetores  $\mathbf{u}_k \in \Re^m$  e  $\mathbf{y}_k \in \Re^l$  são as medidas no  $k$ -ésimo instante de respectivamente,  $m$  entradas e  $l$  saídas do processo. O vetor  $\mathbf{x}_k \in \Re^n$  é o vetor de estado do processo. As matrizes  $\mathbf{A} \in \Re^{n \times n}$ ,  $\mathbf{B} \in \Re^{n \times m}$ ,  $\mathbf{C} \in \Re^{l \times n}$  e  $\mathbf{D} \in \Re^{l \times m}$  terão dimensões adequadas com a ordem do sistema.

Para o modelamento neste Capítulo, se todos os erros internos e externos são supostos nulos, a menos do ruído de medida de saída,  $\mathbf{v}_k$ , então para a medida de saída tem-se:

$$\mathbf{z}_k = \mathbf{y}_k + \mathbf{v}_k \quad (6.2)$$

donde um modelo de espaço de estado com erro de saída, como visto no Capítulo 1. Poderíamos supor  $\mathbf{v}_k$  um processo estocástico arbitrário de média zero, contudo, restringiremos nossa apresentação neste capítulo supondo  $\mathbf{v}_k = 0$ .



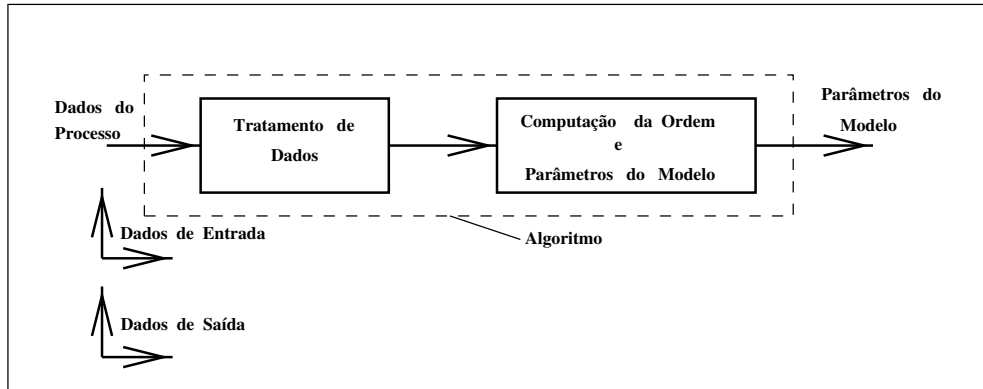


Figura 6.2: Diagrama Esquemático de um Procedimento para Computação de Modelo no Espaço de Estado.

O problema investigado neste capítulo é o de identificar o modelo no espaço de estado, incluindo a ordem do sistema e a quádrupla de matrizes, a partir de um algoritmo que explora as propriedades de subespaços. Para obter a ordem do sistema e a quádrupla de matrizes utiliza-se a aproximação de um subespaço definido pelo espaço gerado por colunas de matrizes determinadas por dados de entrada-saída do sistema, Figura 6.2.

## 6.2 Método de Subespaços

O método MOESP tratado neste Capítulo para obter a ordem do sistema e a quádrupla de matrizes utiliza a aproximação de um subespaço, definido pelo espaço gerado por colunas de matrizes determinadas por dados de entrada-saída do sistema e maiores detalhes são encontrados em Verhaegen & Dewilde, [1113, 1114, 1115]. Os aspectos relevantes deste algoritmo são apresentados a seguir.

Representações úteis no estudo das propriedades de realização no espaço de estado, baseadas na seqüência de entrada-saída, foram estudadas no Capítulo 5 e aqui enfatizamos o modelo MEEE1 para o caso determinístico:

$$\mathbf{Y}_H = \mathcal{O}_i \mathbf{X}_f^d(k) + \mathcal{T} \mathbf{U}_H(k) \quad (6.3)$$

A matriz  $\mathbf{Y}_H$  com  $l.i$  linhas e  $j$  colunas tem consecutivos vetores de saída  $\mathbf{y}_k$ , (de dimensão  $l \times 1$ ), ordenados da seguinte maneira:

$$\mathbf{Y}_H = \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \cdots & \mathbf{y}_{k+j-1} \\ \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \cdots & \mathbf{y}_{k+j} \\ \mathbf{y}_{k+2} & \mathbf{y}_{k+3} & \cdots & \mathbf{y}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{k+i-1} & \mathbf{y}_{k+i} & \cdots & \mathbf{y}_{k+j+i-2} \end{bmatrix} \quad (6.4)$$

A matriz  $\mathbf{U}_H$  com a mesma estrutura da matriz  $\mathbf{Y}_H$  contém consecutivos vetores de entrada  $\mathbf{u}_k$ , (de dimensão  $m \times 1$ ), ordenados da seguinte maneira:

$$\mathbf{U}_H = \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \cdots & \mathbf{u}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \cdots & \mathbf{u}_{k+j} \\ \mathbf{u}_{k+2} & \mathbf{u}_{k+3} & \cdots & \mathbf{u}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \cdots & \mathbf{u}_{k+j+i-2} \end{bmatrix} \quad (6.5)$$

$\mathbf{X}_f^d(k)$  contém a seqüência de vetores de estado:

$$\mathbf{X}_f^d = [\mathbf{x}_k \quad \mathbf{x}_{k+1} \quad \mathbf{x}_{k+2} \quad \cdots \quad \mathbf{x}_{k+j-1}] \quad (6.6)$$

$\mathcal{O}_i$  é uma matriz com estrutura semelhante à da matriz de observabilidade, matriz de observabilidade estendida, dada por:

$$\mathcal{O}_i = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \quad (6.7)$$

e finalmente,  $\mathcal{T}$  é uma matriz triangular inferior dada por:

$$\mathcal{T} = \begin{bmatrix} \mathbf{D} & 0 & 0 & \cdots & 0 \\ \mathbf{CB} & \mathbf{D} & 0 & \cdots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \cdots & 0 \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \cdots & \mathbf{D} \end{bmatrix} \quad (6.8)$$

Observa-se que  $\mathbf{D}$  está na diagonal principal da matriz  $\mathcal{T}$ .

A equação (6.3) pode ser obtida através da substituição direta na equação (6.1), das matrizes acima apresentadas, propiciando uma representação muito importante para estudar as propriedades de uma realização no espaço de estado a partir das seqüências multivariadas de entrada e de saída, como

visto no Capítulo 5, bem como para o desenvolvimento de algoritmos de identificação multivariável no espaço de estado, como o aqui implementado.

$$\begin{aligned}
 & \begin{bmatrix} \mathbf{y}_k & \mathbf{y}_{k+1} & \cdots & \mathbf{y}_{k+j-1} \\ \mathbf{y}_{k+1} & \mathbf{y}_{k+2} & \cdots & \mathbf{y}_{k+j} \\ \mathbf{y}_{k+2} & \mathbf{y}_{k+3} & \cdots & \mathbf{y}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_{k+i-1} & \mathbf{y}_{k+i} & \cdots & \mathbf{y}_{k+j+i-2} \end{bmatrix} = \\
 & = \begin{bmatrix} \mathbf{C} \\ \mathbf{CA} \\ \mathbf{CA}^2 \\ \vdots \\ \mathbf{CA}^{i-1} \end{bmatrix} \begin{bmatrix} \mathbf{x}_k & \mathbf{x}_{k+1} & \mathbf{x}_{k+2} & \cdots & \mathbf{x}_{k+j-1} \end{bmatrix} + \\
 & \begin{bmatrix} \mathbf{D} & 0 & 0 & \cdots & 0 \\ \mathbf{CB} & \mathbf{D} & 0 & \cdots & 0 \\ \mathbf{CAB} & \mathbf{CB} & \mathbf{D} & \cdots & 0 \\ \mathbf{CA}^2\mathbf{B} & \mathbf{CAB} & \mathbf{CB} & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{CA}^{i-2}\mathbf{B} & \mathbf{CA}^{i-3}\mathbf{B} & \mathbf{CA}^{i-4}\mathbf{B} & \cdots & \mathbf{D} \end{bmatrix} \begin{bmatrix} \mathbf{u}_k & \mathbf{u}_{k+1} & \cdots & \mathbf{u}_{k+j-1} \\ \mathbf{u}_{k+1} & \mathbf{u}_{k+2} & \cdots & \mathbf{u}_{k+j} \\ \mathbf{u}_{k+2} & \mathbf{u}_{k+3} & \cdots & \mathbf{u}_{k+j+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{u}_{k+i-1} & \mathbf{u}_{k+i} & \cdots & \mathbf{u}_{k+j+i-2} \end{bmatrix} \\
 & \tag{6.9}
 \end{aligned}$$

### 6.3 Implementação de Alto Desempenho De Método de Subespaço

O algoritmo MOESP ordinário, classificado como um procedimento de identificação de sistemas baseado em subespaços, requer o processamento de uma matriz associada com subespaços gerados por dados de entrada-saída construída numa forma particular: de Hankel. Aplicando operações da Álgebra Linear Numérica tais como fatoração  $\mathbf{RQ}$  e decomposição em valores singulares da matriz de Hankel, obtemos um sistema de equações matriciais lineares super-determinado cuja solução fornece uma realização para o suposto sistema dinâmico que gera as seqüências de dados de entrada-saída, que desejamos identificar.

#### 6.3.1 Cálculo das Matrizes $\mathbf{A}$ e $\mathbf{C}$

Suponha que aplicamos ao sistema uma seqüência de  $N$  entradas independentes  $\mathbf{u}_k, k = 0, 1, 2, \dots, N-1$ , mas que temos apenas um segmento finito de

uma seqüência de saída de  $k = 0, 1, 2 \dots L$ , com  $d \leq L \ll N$ . Podemos então através de uma escolha adequada de seqüências de entradas e saídas medidas relacionar as entradas conhecidas com as saídas medidas de tal forma que:

$$\mathbf{W} = \begin{bmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \dots & \mathbf{u}_{N-1} \\ \mathbf{y}_0 & \mathbf{y}_1 & \dots & \mathbf{y}_{N-1} \end{bmatrix} \mathbf{Q} = \begin{bmatrix} \mathbf{u}'_0 & \mathbf{u}'_1 & \dots & \mathbf{u}'_L & 0 & \dots & 0 \\ \mathbf{y}'_0 & \mathbf{y}'_1 & \dots & \mathbf{y}'_L & \mathbf{y}'_{L+1} & \dots & \mathbf{y}'_{N-1} \end{bmatrix} \quad (6.10)$$

Através de transformações apropriadas, por exemplo fatoração  $\mathbf{QR}$ , podemos reescrever Eq.(6.10) como:

$$\mathbf{W} = \begin{bmatrix} \mathbf{U}_H \\ \mathbf{Y}_H \end{bmatrix} = \left[ \begin{array}{c|c} \mathbf{R}_{11} & 0 \\ \mathbf{R}_{21} & \mathbf{R}_{22} \end{array} \right] \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \end{bmatrix} \quad (6.11)$$

onde  $\mathbf{R}_{11}$  e  $\mathbf{R}_{22}$  são matrizes triangulares inferiores de dimensão  $(L+1) \times (L+1)$  e  $[(\mathbf{Q}_1)^T (\mathbf{Q}_2)^T]$  são as primeiras  $2(L+1)$  colunas de uma matriz unitária  $\mathbf{Q}$  tendo dimensão  $N \times N$ . Do Corolário 5.4, *posto*  $\mathbf{W} = mi+n$ , onde  $mi$  é o número de linhas linearmente independentes de  $\mathbf{U}_H$  e  $n$  é a ordem da realização mínima.

Logo,

$$[\mathbf{y}'_0 \ \mathbf{y}'_1 \ \dots \ \mathbf{y}'_{N-1}] = [\mathbf{R}_{22} \ 0] \quad (6.12)$$

Devido à propriedade de invariância ao deslocamento da matriz  $\mathcal{O}_i$ , a partir da decomposição em valores singulares de  $\mathbf{R}_{22}$ , obtemos:

$$\mathbf{R}_{22} = [\mathbf{U}_n \ |\mathbf{U}_n^\perp] \begin{bmatrix} \mathbf{S}_n & 0 \\ 0 & \mathbf{S}_2 \end{bmatrix} \begin{bmatrix} \mathbf{V}_n^T \\ (\mathbf{V}_n^\perp)^T \end{bmatrix} \quad (6.13)$$

Então as matrizes  $\mathbf{A}$  e  $\mathbf{C}$  serão dadas por:

$$\mathbf{U}_n^{(1)} \mathbf{A} = \mathbf{U}_n^{(2)} \quad (6.14)$$

$$\mathbf{C} = \mathbf{U}_n(1:l,:) \quad (6.15)$$

onde  $\mathbf{U}_n^{(1)}$  é uma submatriz composta pelas primeiras  $(i-1)$  colunas da matriz  $\mathbf{U}_n$  e  $\mathbf{U}_n^{(2)}$  é uma submatriz composta pelas últimas  $(i-1)$  colunas da matriz  $\mathbf{U}_n$ , a menos de uma transformação de similaridade. Observe que  $\mathbf{U}_n(1:l,:)$  refere-se a notação utilizada em *MATLAB*.

### 6.3.2 Cálculo das Matrizes $\mathbf{B}$ e $\mathbf{D}$

O desenvolvimento algébrico para o cálculo das matrizes  $\mathbf{B}$  e  $\mathbf{D}$  também é realizado a partir da fatoração  $\mathbf{QR}$  da matriz de dados de entrada e saída da equação (6.11), da qual:

$$\mathbf{U}_H = \mathbf{R}_{11} \mathbf{Q}_1 \quad (6.16)$$

Retomando a Equação (6.3):  $\mathbf{Y}_H = \mathcal{O}_i \mathbf{X}_f^d + \mathcal{T} \mathbf{U}_H$ , podemos substituir  $\mathbf{Y}_H$  obtido da equação (6.11); deste modo encontramos:

$$\mathbf{Y}_H = \mathcal{O}_i \mathbf{X}_f^d + \mathcal{T} \mathbf{U}_H = \mathbf{R}_{21} \mathbf{Q}_1 + \mathbf{R}_{22} \mathbf{Q}_2 \quad (6.17)$$

Como o espaço coluna de  $\mathbf{R}_{22}$  é igual ao espaço coluna de  $\mathcal{O}_i$ , se a decomposição em valores singulares de  $\mathbf{R}_{22} = \mathbf{U}_n \mathbf{S}_n \mathbf{V}_n^T$ , então haverá uma matriz de transformação,  $\mathbf{T}$ , tal que  $\mathcal{O}_i \mathbf{T} = \mathbf{U}_n$  e podemos escrever:

$$\mathbf{U}_n \mathbf{T}^{-1} \mathbf{X}_f^d + \mathcal{T} \mathbf{R}_{11} \mathbf{Q}_1 = \mathbf{R}_{21} \mathbf{Q}_1 + \mathbf{U}_n \mathbf{S}_n \mathbf{V}_n^T \mathbf{Q}_2 \quad (6.18)$$

Como  $\mathbf{U}_n^\perp \mathbf{U}_n = 0$ , obtemos da Eq. (6.18):

$$[(\mathbf{U}_n^\perp)^T \mathcal{T} \mathbf{R}_{11} - (\mathbf{U}_n^\perp)^T \mathbf{R}_{21}] \mathbf{Q}_1 = 0 \quad (6.19)$$

ou

$$(\mathbf{U}_n^\perp)^T \mathcal{T} \mathbf{R}_{11} - (\mathbf{U}_n^\perp)^T \mathbf{R}_{21} = 0 \quad (6.20)$$

Como  $\mathbf{R}_{11}$  tem inversa:

$$(\mathbf{U}_n^\perp)^T \mathcal{T} - (\mathbf{U}_n^\perp)^T \mathbf{R}_{21} \mathbf{R}_{11}^{-1} = 0 \quad (6.21)$$

como  $\mathcal{T} = \text{func}[\mathbf{B} \mid \mathbf{D}]$  temos a quádrupla de matrizes determinada.

### 6.3.3 Proposta de Implementação Multiprocessada

Na Figura 6.3 apresentamos uma proposta para paralelização do algoritmo, onde é possível notar que existe uma paralelização natural do mesmo, no cálculo das matrizes  $\mathbf{Y}_H$  e  $\mathbf{U}_H$ , partes 6 e 7, por exemplo. No entanto, na implementação efetuada paralelizamos as fatorações  $\mathbf{QR}$  que devido ao grande número de operações algébricas envolvidas possuem elevada complexidade computacional, da ordem de  $2mn^2$  para matriz  $m \times n$ .

## 6.4 Ambiente de Programação

Neste trabalho a obtenção da solução paralela para determinação de uma realização no espaço de estado é feita em um ambiente que utiliza como plataforma de programação paralela o *MPI: Message Passing Interface*, com memória distribuída, comunicação síncrona com semântica bloqueante das funções para enviar e receber mensagens, com modo de programação *SPMD*:

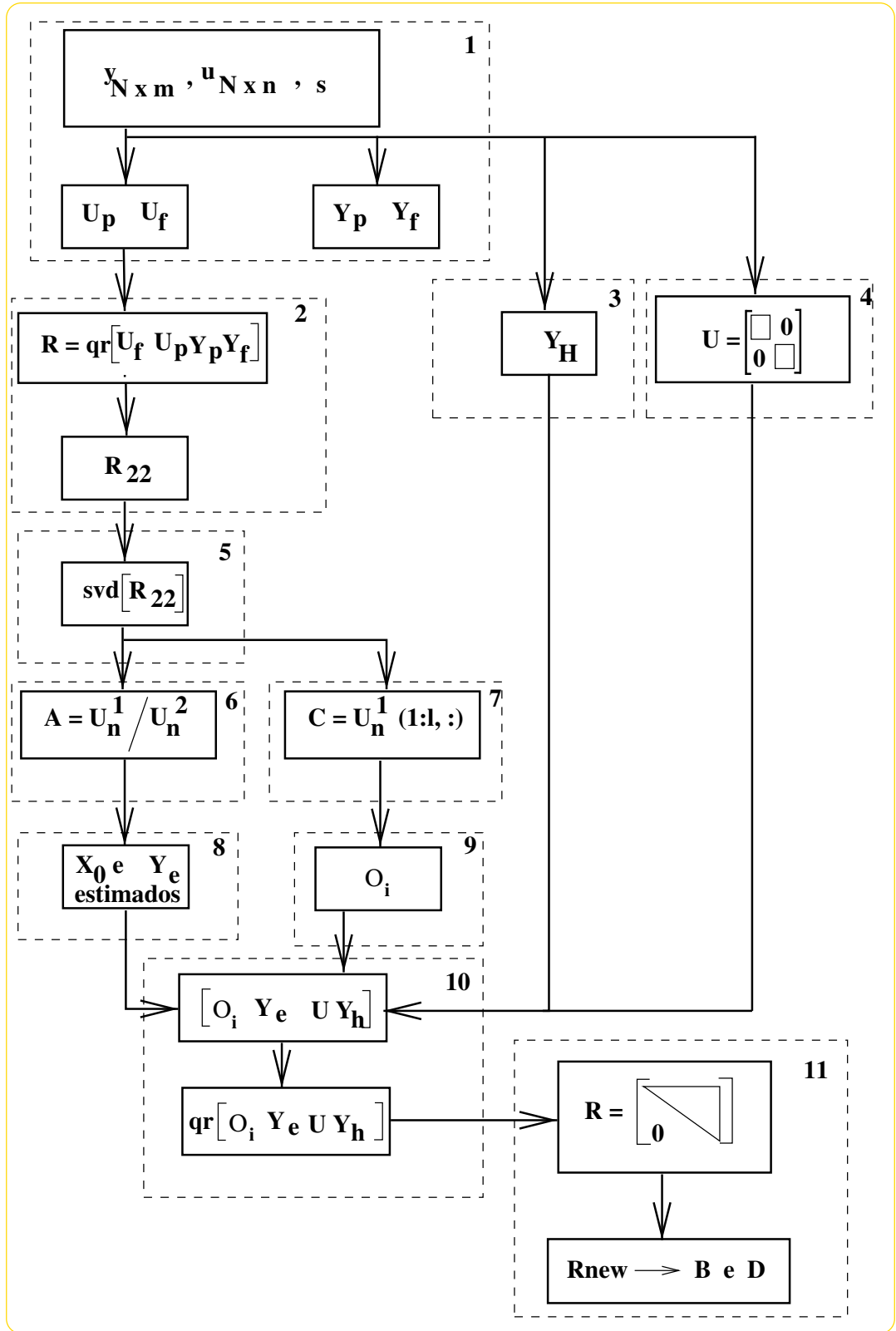


Figura 6.3: Proposta de Paralelização do Algoritmo MOESP.

*Single Program, Multiple Data*, onde todos os processadores executam o mesmo programa, podendo porém executar tarefas diferentes sobre dados diferentes. A solução foi implementada utilizando as subrotinas da biblioteca pública *Scalapack*.

A *Scalapack* é uma biblioteca de alto desempenho com rotinas de Álgebra Linear para computadores *MIMD*<sup>1</sup> com paradigma de troca de mensagens e memória distribuída em redes de estações de trabalho suportando *PVM* e ou *MPI* e foi apresentada no Capítulo 3. Ela utiliza como parte de suas rotinas, a biblioteca sequencial *BLAS* - *Basic Linear Algebra Subroutines* e a paralela *PBLAS* - *Parallel BLAS*, os subprogramas de comunicação da álgebra linear básica *BLACS* - *Basic Linear Algebra Communication Subroutines* e as rotinas da *LAPACK* - *Linear Algebra PACKage*.

Similarmente ao *LAPACK*, as rotinas da *Scalapack* são baseadas em algoritmos com particionamento em blocos, o que minimiza a frequência de movimento de dados entre os diferentes níveis da memória hierárquica. Os algoritmos são apresentados como processos e a comunicação pode ser ponto-a-ponto ou em grupo quando necessário. De forma geral, a *Scalapack* foi desenvolvida e testada para o caso de um processo por processador.

Dos vários algoritmos paralelos que têm sido propostos, um dos que receberam maior atenção recentemente é o baseado em multiplicação de matrizes, pois a multiplicação de matrizes de alta ordem é forte candidata à paralelização, Henry & Watkins & Dongarra, [485].

### 6.4.1 Aspectos Paralelos Para Processamento Computacional

Em todas as rotinas os elementos locais de uma matriz são armazenados em arranjos uni-dimensionais. Por exemplo, os elementos locais de uma matriz  $A$  ( $m \times n$ ) podem ser armazenados num arranjo vetorial real  $A(*)$ . É mais conveniente considerar  $A$  como um arranjo de duas dimensões ( $LDA, *$ ), onde  $LDA$  deve ser maior ou igual ao número de linhas de  $A$  armazenadas na linha específica da grade de processadores e o arranjo  $A$  deve ter um número de colunas maior ou igual ao número de colunas de  $A$  armazenadas na coluna específica da grade dos processadores. Maiores informações sobre a distribuição da matriz sobre os processadores participantes no processo são encapsuladas num arranjo denominado "*descriptor de arranjo*". Tal descriptor é associado a cada uma das matrizes distribuídas. As coordenadas do des-

---

<sup>1</sup>Esta arquitetura combina paralelismo nos fluxos de dados e de instruções; é composta de unidades de processamento, cada qual capaz de realizar operações lógicas e aritméticas padrões. As unidades operam assincronamente sob o controle de fluxo de instruções individuais.

critor unicamente determinam o mapa das coordenadas da matriz sobre a memória dos processadores locais. Por exemplo, na fatoração  $\mathbf{QR}$  e rotinas associadas, o descritor, DESCAs, tem dimensão (9), e armazena as seguintes informações:

DESCA(1) - Tipo do descritor ( para a distribuição de bloco cíclica de duas dimensões deve ser igual a 1).

DESCA(2) - Encadeamento para a biblioteca ( variável retornada pela rotina de inicialização para identificar a matriz  $\mathbf{A}$ ).

DESCA(3) -  $m$ , número de linhas de  $\mathbf{A}$ .

DESCA(4) -  $n$ , número de colunas de  $\mathbf{A}$ .

DESCA(5) -  $M_b$ , fator bloqueante usado para distribuir as linhas de  $\mathbf{A}$ , isto é, o número de linhas armazenadas num bloco.

DESCA(6) -  $N_b$ , fator bloqueante usado para distribuir as colunas de  $\mathbf{A}$ , isto é, o número de colunas armazenadas num bloco.

DESCA(7) - Índice do processador linha sobre o qual a primeira linha de  $\mathbf{A}$  é armazenada.

DESCA(8) - Índice do processador coluna sobre o qual a primeira coluna de  $\mathbf{A}$  é armazenada.

DESCA(9) - Dimensão ( $LDA$ ) do arranjo local  $\mathbf{A}$  armazenando os blocos locais de  $\mathbf{A}$ .

Em geral, o descritor de arranjo não muda através do ciclo de vida da matriz para a qual ele esta associado.

### 6.4.2 Estrutura da rede de estações de trabalho

As simulações foram realizadas num computador paralelo virtual formado por computadores conectados numa rede local com máquinas Sun do tipo SPARC-20, SPARC-4, SPARC-Classic e SPARC-2. Esta rede pertence ao DMCSI: *Departamento de Máquinas, Componentes e Sistemas Inteligentes*, e sua configuração está apresentada na Figura 6.4. Esta rede é constituída de duas subredes que possuem como ponto de ligação (*Gateway*) entre elas a máquina *Papagaio*. As simulações foram realizadas na sub rede LCSi: *Laboratório de Controle e Sistemas Inteligentes*, utilizando a máquina *Águia*, que é a servidora desta rede e é o ponto de ligação com a rede da FEEC: *Faculdade de Engenharia Elétrica e Computação*.



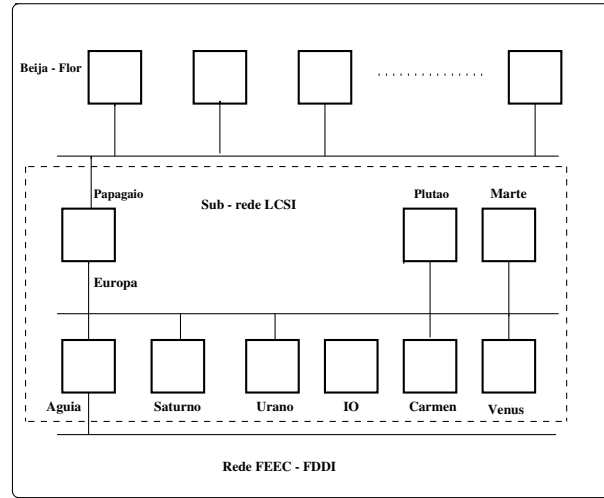


Figura 6.4: Estrutura da Rede de Estações de Trabalho.

### 6.4.3 Subrotina para Fatoração QR

Para o processamento usamos o algoritmo **QR**, aplicado em uma matriz de Hankel, **H**, construída com os dados de entrada - saída do sistema. A fatoração **QR** é usada para comprimir os dados. O algoritmo para fatoração **QR** explora a estrutura da matriz de Hankel. O método implementado é numericamente estável e eficiente.

#### Descrição

A subrotina para fatoração **QR** de uma matriz real,  $\mathbf{A}_s, m \times n$ , calcula  $\mathbf{A}_s = \mathbf{Q}\mathbf{R}$ , onde  $\mathbf{A}_s$  é uma submatriz de tamanho  $m_{\mathbf{A}} \times n_{\mathbf{A}}$  da matriz **A**, isto é,

$$\mathbf{A}_s(1 : m, 1 : n) = \mathbf{A}(i_{\mathbf{A}} : i_{\mathbf{A}} + m - 1, j_{\mathbf{A}} : j_{\mathbf{A}} + n - 1)$$

Se  $i_{\mathbf{A}} = j_{\mathbf{A}} = 1, m = m_{\mathbf{A}}, n = n_{\mathbf{A}}$  então  $\mathbf{A}_s = \mathbf{A}$ .

A matriz ortogonal **Q** não é formada explicitamente, mas é representada como o produto de refletores elementares:

$$\mathbf{Q} = \mathbf{H}_1 \mathbf{H}_2 \dots \mathbf{H}_k, \text{ onde } k = \min(m, n) \quad (6.22)$$

Cada refletor elementar  $\mathbf{H}_l$  tem a forma:

$$\mathbf{H}_l = \mathbf{I} - \frac{2\mathbf{w}\mathbf{w}^T}{\mathbf{w}^T\mathbf{w}}, \quad (6.23)$$

conforme apresentado no Capítulo 2.

### 6.4.4 Subrotina para Decomposição em Valores Singulares

A subrotina calcula a decomposição em valores singulares (*SVD*) de uma matriz real cujas colunas são distribuídas numa grade de duas dimensões.

A *SVD* de uma matriz  $A$  real retangular  $m \times n$ , onde  $m \geq n$ , é definida como

$$A = \begin{bmatrix} U & \tilde{U} \end{bmatrix} \begin{bmatrix} \Sigma \\ 0 \end{bmatrix} V^T = U \Sigma V^T$$

onde  $U$  é uma matriz  $m \times n$  de vetores singulares esquerdos,  $\tilde{U}$  é uma matriz  $m \times (m - n)$ ,  $V$  é uma matriz ortogonal  $n \times n$  de vetores singulares direitos e  $\Sigma$  é uma matriz diagonal  $n \times n$  de valores singulares. Os valores singulares  $\sigma_1, \sigma_2, \dots, \sigma_n$  são não negativos e decrescentes em ordem de magnitude. A matriz  $\begin{bmatrix} U & \tilde{U} \end{bmatrix}$  é ortogonal.

Para o caso  $m < n$  a *SVD* pode ser definida na forma

$$A = U \begin{bmatrix} \Sigma & 0 \end{bmatrix} \begin{bmatrix} V & \tilde{V} \end{bmatrix}^T = U \Sigma V^T$$

onde  $U$  é uma matriz  $m \times m$  ortogonal,  $\tilde{V}$  é uma matriz  $n \times (n - m)$ ,  $V$  é uma matriz ortogonal  $n \times m$  de vetores singulares direitos e  $\Sigma$  é uma matriz diagonal  $m \times m$  de valores singulares. A matriz  $\begin{bmatrix} V & \tilde{V} \end{bmatrix}$  é ortogonal.

Em geral, as matrizes  $\tilde{U}$  e  $\tilde{V}$  não são únicas.

### 6.4.5 Especificações para Subrotinas em FORTRAN

#### Especificação para Subrotina QR

```
SUBROUTINE PSGEQRF(M, N, A, IA, JA, DESCA, TAU, WORK, LWORK, INFO)
REAL          A(*), TAU(*), WORK(*)
INTEGER       M, N, IA, JA, DESCA(*), LWORK, INFO
```

A matriz  $A$  deve ser particionada em  $M_b^A \times N_b^A$  blocos retangulares os quais são armazenados num arranjo  $A$  com uma distribuição de bloco cíclica de duas dimensões.

#### Especificação para Subrotina SVD

```
SUBROUTINE PSGESVD( JOBU, JOBVT, M, N, A, IA, JA, DESCA, S, U, IU,
$                  JU, DESCU, VT, IVT, JVT, DESCVT, WORK, LWORK,
$                  INFO )
```

Processador $\{0, 0\}$	Processador $\{0, 1\}$
$N_x = 3$	$N_x = 3$
colunas (1 : 3)	colunas (4 : 6)
Processador $\{1, 0\}$	Processador $\{1, 1\}$
$N_x = 3$	$N_x = 2$
colunas (7 : 9)	colunas (10 : 11)

Tabela 6.1: Partionamento de Matriz em 4 Processadores.

```

INTEGER          DESCA( * ), DESCU( * ), DESCVT( * )
REAL             A( * ), S( * ), U( * ), VT( * ), WORK( * )
CHARACTER        JOBU, JOBVT
INTEGER          IA, INFO, IU, IVT, JA, JU, JVT, LWORK, M, N

```

As colunas da matriz  $A$  são alocadas a processadores lógicos da grade, linha por linha, isto é, na linha de maior ordem da grade, começando no processador lógico  $\{0, 0\}$ . Cada processador lógico que contém colunas da matriz, contém  $NB = \lceil N/P \rceil$  colunas, exceto para o último processador que contém dados e o número de colunas deve ser menor que  $NB$ . Este processador deverá conter  $\text{mod}(N, NB)$  colunas se  $\text{mod}(N, NB) \neq 0$  e poderia conter  $N_b$  colunas. Alguns processadores lógicos podem não conter qualquer coluna da matriz se  $N$  não é relativamente maior que  $p$ , mas se  $N > (p - 1)^2$  então todos os processadores podem certamente conter colunas da matriz, na seção 7.2.2 ilustramos este particionamento.

O número de processadores lógicos que contém colunas da matriz é dado por  $p_d = \lceil N/NB \rceil$ .

O seguinte exemplo ilustra o caso onde o último processador com dados é o último processador da grade. Além do mais, o número de colunas no último processador com dados não é igual ao número de colunas nos demais processadores.

Se  $m_p = 2$ ,  $n_p = 2$  então  $p = m_p \times n_p = 4$ . Se  $n = 11$  então  $NB = \lceil N/p \rceil = \lceil 2.75 \rceil = 3$ ,  $\text{mod}(N, NB) = 2 \neq 0$  e  $p_d = \lceil N/NB \rceil = \lceil 3.6667 \rceil = 4$ .

## 6.5 Resultados e Comentários

Nesta seção apresentamos alguns dos resultados obtidos com uma versão paralela do método MOESP utilizando a biblioteca *ScaLAPACK*.

Todos os experimentos foram realizados utilizando Fortran 77 e aritmética com dupla precisão, padrão IEEE.

Uma sequência de entrada-saída, com vetores de entrada multivariada de 3 elementos e vetores de saída multivariada de 2 elementos, é obtida a

partir de uma realização “benchmark” no espaço de estado com as seguintes matrizes:  $\mathbf{A}_{4 \times 4}$ ;  $\mathbf{B}_{4 \times 3}$ ;  $\mathbf{C}_{2 \times 4}$  e  $\mathbf{D}_{2 \times 3}$ , representantes do conjunto de dados original e apresentadas a seguir:

$$\mathbf{A} = \begin{bmatrix} 0.2128 & 0.1360 & 0.1979 & -0.0836 \\ 0.1808 & 0.4420 & -0.3279 & 0.2344 \\ -0.5182 & 0.1728 & -0.5448 & -0.3083 \\ 0.2252 & -0.0541 & -0.4679 & 0.8290 \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} -0.0101 & 0.0317 & -0.9347 \\ -0.0600 & 0.5621 & 0.1657 \\ -0.3310 & -0.3712 & -0.5846 \\ -0.2655 & 0.4255 & 0.2204 \end{bmatrix},$$

$$\mathbf{C} = \begin{bmatrix} 0.6557 & -0.2502 & -0.5188 & -0.1229 \\ 0.6532 & -0.1583 & -0.0550 & -0.2497 \end{bmatrix} \quad \text{e} \quad \mathbf{D} = \begin{bmatrix} -0.4326 & 0.1253 & -1.1465 \\ -1.6656 & 0.2877 & 1.1909 \end{bmatrix}$$

Através da modelagem computacional aqui desenvolvida para um conjunto com 490 amostras de valores entrada - saída, obtivemos o seguinte modelo no espaço de estado:

$$\mathbf{A}_c = \begin{bmatrix} 0.3386 & 0.4402 & -0.5891 & -0.0241 \\ -0.3428 & -0.6781 & -0.5043 & -0.3215 \\ -0.0519 & 0.0062 & 0.7988 & -0.5805 \\ 0.0003 & 0.0043 & 0.0103 & 0.4796 \end{bmatrix}, \quad \mathbf{B}_c = \begin{bmatrix} -0.0940 & 0.2708 & 1.0339 \\ 0.2965 & 0.2712 & 0.3916 \\ 0.1043 & -0.1980 & 0.1669 \\ 0.0088 & 0.0115 & -0.0037 \end{bmatrix},$$

$$\mathbf{C}_c = \begin{bmatrix} -0.5725 & 0.5875 & -0.1078 & -0.4545 \\ -0.6617 & 0.0679 & -0.0000 & 0.5835 \end{bmatrix} \quad \text{e} \quad \mathbf{D}_c = \begin{bmatrix} -0.4326 & 0.1253 & -1.1465 \\ -1.6656 & 0.2877 & 1.1909 \end{bmatrix}$$

As matrizes do sistema “benchmark” e do sistema calculado são diferentes, no entanto o sistema calculado pode representar adequadamente o conjunto de dados entrada-saída do sistema desde que os parâmetros de Markov do sistema “benchmark” e do sistema calculado sejam idênticos; resultado obtido através do Corolário 5.1. Desta forma tem-se:

a) Os parâmetros de Markov para o sistema “benchmark”:

i)

$$\mathbf{D} = \begin{bmatrix} -0.4326 & 0.1253 & -1.1465 \\ -1.6656 & 0.2877 & 1.1909 \end{bmatrix}$$

ii)

$$\mathbf{CB} = \begin{bmatrix} 0.2127 & 0.0204 & -0.3781 \\ 0.0874 & -0.1541 & -0.6597 \end{bmatrix}$$

iii)

$$\mathbf{CAB} = \begin{bmatrix} -0.1651 & -0.2765 & -0.6657 \\ -0.0360 & -0.2262 & -0.3273 \end{bmatrix}$$

iv)

$$\mathbf{CA}^2\mathbf{B} = \begin{bmatrix} 0.1256 & -0.0024 & 0.2887 \\ 0.0984 & -0.0833 & 0.1614 \end{bmatrix}$$

e assim sucessivamente.

b) Os parâmetros de Markov para o modelo obtido:

i)

$$\mathbf{D}_c = \begin{bmatrix} -0.4326 & 0.1253 & -1.1465 \\ -1.6656 & 0.2877 & 1.1909 \end{bmatrix}$$

ii)

$$\mathbf{C}_c\mathbf{B}_c = \begin{bmatrix} 0.2128 & 0.0204 & -0.3782 \\ 0.0875 & -0.1541 & -0.6597 \end{bmatrix}$$

iii)

$$\mathbf{C}_c\mathbf{A}_c\mathbf{B}_c = \begin{bmatrix} -0.1651 & -0.2766 & -0.6658 \\ -0.0359 & -0.2262 & -0.3273 \end{bmatrix}$$

iv)

$$\mathbf{C}_c\mathbf{A}_c^2\mathbf{B}_c = \begin{bmatrix} 0.1256 & -0.0024 & 0.2888 \\ 0.0984 & -0.0832 & 0.1615 \end{bmatrix}$$

e assim sucessivamente.

Como os parâmetros de Markov do sistema “benchmark”:

$$[\mathbf{D}, \mathbf{CB}, \mathbf{CAB}, \mathbf{CA}^2\mathbf{B}, \dots]$$

coincidem com os parâmetros de Markov do modelo obtido:

$$[\mathbf{D}_c, \mathbf{C}_c\mathbf{B}_c, \mathbf{C}_c\mathbf{A}_c\mathbf{B}_c, \mathbf{C}_c\mathbf{A}_c^2\mathbf{B}_c, \dots]$$

concluímos que a verificação da qualidade da realização é adequada.

Os tempos de processamento do algoritmo MOESP correspondentes a dois experimentos são apresentados na Tabela 6.2.

Para comentar a Tabela 6.2 faremos referência a um conceito muito importante em paralelismo que é a granularidade. A granularidade consiste no “grau de divisão” dos dados de um problema entre os processadores que

Tarefa do Algoritmo	Tempo de Execução, (s)	
	1 processador	2 processadores
Leitura de Dados	0.0741	0.0657
Achando <b>QR</b>	0.1554	0.3366
Formando $\mathbf{R}_{32}$	0.0231	0.0152
Achando SVD	0.0478	0.1960
Achando $\mathbf{U}_1, \mathbf{U}_2$	0.0008	0.0008
Achando SVD de $\mathbf{U}_1$	0.0170	0.1486
Achando $inv(\mathbf{S}_1)$	0.0009	0.0005
Formando $\mathbf{Y}_e$	0.0929	0.0709
Formando $\mathcal{O}_i$	0.0844	0.0709
Formando $[\mathcal{O}_i \mathbf{Y}_e \mathbf{U} \mathbf{Y}_h]$	0.5720	0.4777
Achando <b>QR</b>	0.1123	0.4738
Formando <b>Rnew</b>	0.7976	0.5932
Achando SVD de <b>Rnew</b>	0.0501	0.2436
Achando $inv(\mathbf{S})$	0.0397	0.0295
Tempo Total	2.0681	2.7230

Tabela 6.2: Tempos de Processamento do Algoritmo MOESP.

Tarefa do Algoritmo	Ganho ( “ <i>speedup</i> ” )	Eficiência ( % )	Função Custo	Eficácia
Leitura de Dados	1.1279	56.3950	1.7732	0.6361
Achando <b>QR</b>	0.4617	23.0850	4.3318	0.1066
Formando $\mathbf{R}_{32}$	1.5197	75.9850	1.3160	1.1547
Achando SVD	0.2439	12.1950	8.2001	0.0297
Achando $\mathbf{U}_1, \mathbf{U}_2$	1	50	2	0.5000
Achando SVD de $\mathbf{U}_1$	0.1144	5.7200	17.4825	0.0065
Achando $\text{inv}(\mathbf{S}_1)$	1.3103	65.5150	1.5264	0.8584
Formando $\mathbf{Y}_e$	1.3103	65.5150	1.5264	0.8584
Formando $\mathcal{O}_i$	1.1904	59.5200	1.6801	0.7085
Formando $[\mathcal{O}_i \mathbf{Y}_e \mathbf{U} \mathbf{Y}_h]$	1.1974	59.8700	1.6703	0.7169
Achando <b>QR</b>	0.2370	11.8500	8.4388	0.0281
Formando <b>Rnew</b>	1.3446	67.2300	1.4874	0.9040
Achando SVD de <b>Rnew</b>	0.2057	10.2850	9.7229	0.0212
Achando $\text{inv}(\mathbf{S})$	1.3458	67.2900	1.4861	0.9056
Tempo Total	0.7595	37.9750	2.6333	0.2884

Tabela 6.3: Ganhos no Processamento do Algoritmo MOESP.

executam a tarefa. Assim, a granularidade está diretamente relacionada com o tempo de processamento total (quanto maior ela for, menor será o tempo de processamento) e com o tempo de comunicação entre os processadores (uma vez que quanto mais processadores participarem da tarefa, maior será o tempo de comunicação). Por esta razão tarefas como a decomposição em valores singulares e fatoração **QR** tiveram um pior desempenho quando da execução em paralelo. Como uma das conclusões temos que a utilização de mais processadores não leva, nestas tarefas, para este exemplo, a um aumento na velocidade de processamento. Existe um número finito de processadores que produz um desempenho máximo, no entanto este número é dependente do tamanho das matrizes executadas e do tipo de processadores envolvidos.

A coluna de ganho, relação (3.1), da Tabela 6.3 ressalta as conclusões sobre a granularidade das etapas referentes a decomposição em valores singulares e a fatoração **QR**, comentadas quando da apresentação da Tabela 6.2. Os valores de ganho menores que 1 mostram as atividades em que o aumento do número de processadores provocou um aumento no tempo de processamento.

Ressaltamos ainda na Tabela 6.3, os valores de eficiência, da função custo e da eficácia para as diferentes etapas do algoritmo:

- a) os valores de eficiência, relação (3.5): observe que a eficiência de 100% ocorre para nosso exemplo. Quando a adição do segundo processador acontece, provoca-se a queda do tempo de processamento à metade para uma determinada tarefa.
- b) os valores da função custo, relação (3.6): observe que para valores abaixo de 2, (igual ao número de processadores), a utilização de processamento paralelo levou a um ganho de velocidade.
- c) os valores da eficácia, relação (3.7), mostram a relação entre o aumento da razão de aceleração e o aumento da eficiência.

Dos resultados obtidos podemos observar que:

- Uma rede de estações de trabalho pode ter o seu desempenho computacional superado por uma única máquina de alto desempenho; no entanto o processamento distribuído e paralelo possui qualidades não disponíveis no processamento sequencial.
- Considerar apenas o tempo de execução de um algoritmo como critério para avaliar a qualidade do processamento computacional distribuído e paralelo é uma visão muito particular e polarizada; outros critérios são importantes, por exemplo a taxa de utilização de cada processador.
- O processamento computacional distribuído e paralelo é eficaz.

Como trabalho futuro sugerimos que o desempenho do algoritmo deva ser avaliado através de variações nos parâmetros que especificam a disposição de dados nos processadores. Em máquinas de memória compartilhada, esta avaliação pode ser obtida controlando o tamanho de cada bloco, enquanto que para o processamento em máquinas de memória distribuída pode-se controlar tanto o tamanho de cada bloco como a configuração de distribuição do algoritmo. O objetivo deste experimento é o projeto de algoritmos lineares a serem utilizados em processamento de alto desempenho, tal que a frequência com que os dados são transferidos entre os diferentes níveis hierárquicos de memória seja minimizada; embora a quantidade total de dados movidos seja inalterada, o custo associado com a transmissão é reduzido através de um menor número de mensagens necessárias para mover os dados.



## Capítulo 7

# MODELAGEM PARALELA E DISTRIBUÍDA DE SÉRIES TEMPORAIS EM SUBESPAÇOS DO ESPAÇO DE ESTADO

### 7.1 Introdução

A teoria de sistemas lineares, especialmente no século XX sofreu grandes desenvolvimentos. A partir do final da década de 1950, a abordagem via espaço de estados por Kalman foi de grande impacto, Kailath [566], Caines [171]. A modelagem de séries temporais, a identificação e a análise de sistemas foram e são as grandes beneficiárias desta teoria, mas em diferentes épocas e abordagens. A abordagem via espaço de estados para modelagem de séries temporais foi beneficiada pelos trabalhos de Akaike [8] e Faurre [337] na década de 1970, e através do trabalho de Aoki [40] na década de 1980. Na década de 1990, o livro de análise de séries temporais de Hamilton [447], por exemplo, tem grande parte dedicada à modelagem de séries temporais no espaço de estados e este tratamento têm se destacado por exemplo em economia. Computação de alto desempenho é mais recente, e modelagem no espaço de estado e computação de séries temporais multivariadas têm aplicações em diversas áreas: economia, engenharia, ecologia, medicina.

Neste Capítulo apresentamos implementação distribuída e paralela do algoritmo de Aoki [40], para modelagem de séries temporais multivariáveis. Um sistema dinâmico linear, invariante no tempo e de dimensão finita, pode representar uma família de séries temporais vetoriais, Willems [1162]. Caso

esta família de vetores seja disposta em uma forma matricial conveniente podemos ter uma matriz de Hankel. O tratamento da matriz de Hankel pelo algoritmo de Aoki através de técnicas numéricas adequadas, nos leva a obter uma das realizações do sistema dinâmico em estudo.

## 7.2 Modelagem Paralela de Séries Temporais no Espaço de Estado

A modelagem de séries temporais multivariadas no espaço de estado requer múltiplos experimentos com sequências de dados de entradas. Tal modelo pode ser representado por um sistema discreto estocástico linear multivariável e invariante no tempo, com ruído branco na entrada:

$$\begin{cases} \mathbf{x}_{k+1} = \mathbf{A}\mathbf{x}_k + \mathbf{K}\mathbf{e}_k \\ \mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{e}_k \end{cases} \quad (7.1)$$

onde  $\mathbf{x}_k \in \mathbb{R}^n$  é o vetor de estado estacionário no sentido fraco;  $\mathbf{e}_k \in \mathbb{R}^l$  é o vetor de ruído com média zero, serialmente não correlato, estacionário no sentido fraco;  $\mathbf{y}_k \in \mathbb{R}^l$  é o vetor de observação (saída);  $\mathbf{A} \in \mathbb{R}^{n \times n}$  é a matriz dinâmica do sistema;  $\mathbf{K} \in \mathbb{R}^{n \times l}$  é a matriz de entrada do sistema;  $\mathbf{C} \in \mathbb{R}^{l \times n}$  é a matriz de saída.

Para o modelo descrito na equação (7.1), definimos uma matriz de observabilidade estendida dada por:

$$\mathcal{O} = \begin{bmatrix} \mathbf{C}^T & \mathbf{A}^T \mathbf{C}^T & (\mathbf{A}^T)^2 \mathbf{C}^T & \dots & (\mathbf{A}^T)^{n-1} \mathbf{C}^T & \dots \end{bmatrix} \quad (7.2)$$

e uma matriz de atingibilidade estendida dada por:

$$\mathbf{\Omega} = \begin{bmatrix} \mathbf{M} & \mathbf{A}\mathbf{M} & (\mathbf{A})^2 \mathbf{M} & \dots & (\mathbf{A}^{n-1})\mathbf{M} & \dots \end{bmatrix} \quad (7.3)$$

sendo a matriz  $\mathbf{M} = E(\mathbf{x}_{k+1}\mathbf{y}_k^T)$ , uma matriz de covariância dada por:

$$\mathbf{M} = \mathbf{A}\mathbf{\Pi}\mathbf{C}^T + \mathbf{K}\mathbf{\Delta}. \quad (7.4)$$

com:

$$\mathbf{\Delta} = E(\mathbf{e}_k \mathbf{e}_k^T), \quad \mathbf{\Delta} > 0 \quad \text{condição de regularidade} \quad (7.5)$$

e  $\mathbf{\Pi}$  é a matriz de covariância do vetor de estado:

$$\mathbf{\Pi} = E(\mathbf{x}_k \mathbf{x}_k^T) \quad (7.6)$$

O produto das matrizes de observabilidade e atingibilidade gera uma matriz de Hankel :

$$\mathbf{H} = [\mathbf{h}_{i,j}]_{i,j=0}^{\infty} \quad (7.7)$$

Matrizes desta forma têm a propriedade de serem armazenadoras de dados de processos e têm sido estudadas e aplicadas ao modelamento de sistemas dinâmicos. A matriz de Hankel é uma matriz Toeplitz com as linhas em ordem reversa. Desta forma a matriz de Hankel possui os mesmos elementos nas anti-diagonais. A matriz de Hankel também pode ser formada a partir dos dados de saída do sistema, gerando duas matrizes,  $\mathbf{Y}_-$  e  $\mathbf{Y}_+$  dadas por:

$$\mathbf{Y}_- = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{N-1} \\ 0 & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{N-2} \\ 0 & 0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{N-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{y}_{N-K} & \mathbf{y}_{N-K+1} \\ 0 & 0 & \cdots & \mathbf{y}_{N-K-1} & \mathbf{y}_{N-K} \end{bmatrix} \quad \text{onde } \mathbf{Y}_- \in \Re^{(Kl) \times (N-1)} \quad (7.8)$$

$$\mathbf{Y}_+ = \begin{bmatrix} \mathbf{y}_2 & \mathbf{y}_3 & \mathbf{y}_4 & \cdots & \mathbf{y}_{N-1} & \mathbf{y}_N \\ \mathbf{y}_3 & \mathbf{y}_4 & \mathbf{y}_5 & \cdots & \mathbf{y}_N & 0 \\ \mathbf{y}_4 & \mathbf{y}_5 & \mathbf{y}_6 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_J & \mathbf{y}_{J+1} & \mathbf{y}_{J+2} & \cdots & 0 & 0 \\ \mathbf{y}_{J+1} & \mathbf{y}_{J+2} & \mathbf{y}_{J+3} & \cdots & 0 & 0 \end{bmatrix} \quad \text{onde } \mathbf{Y}_+ \in \Re^{(Jl) \times (N-1)} \quad (7.9)$$

O produto das matrizes  $\mathbf{Y}_+ \mathbf{Y}_-^T / N$  gera uma matriz de Hankel,  $\mathbf{H}$ , associada aos dados de saída de um sistema ergódico, sem calcular explicitamente as covariâncias  $\mathbf{\Lambda}_k = E(\mathbf{y}_k \mathbf{y}_k^T)$ .

A decomposição em valores singulares da matriz de Hankel nos permite a partir de um procedimento algébrico, Masanao Aoki [40], a obtenção da tripla  $(\mathbf{A}, \mathbf{K}$  e  $\mathbf{C})$  do modelo no espaço de estado de um sistema dinâmico; a obtenção desta realização fornece a ordem do sistema. Considera-se que o par de matrizes  $\{\mathbf{A}, \mathbf{C}\}$  é observável e que o par de matrizes  $\{\mathbf{A}, \mathbf{K}\}$  é atingível e ainda que  $\{\mathbf{A}, \mathbf{K}, \mathbf{C}\}$  é uma realização mínima balanceada para o sistema dado na equação (7.1), isto é:

$$\mathbf{G}_c = \mathbf{G}_o \quad (7.10)$$

onde  $\mathbf{G}_o$  é o gramiano de observabilidade e  $\mathbf{G}_c$  é o gramiano de atingibilidade,

Lee [656]:

$$\mathbf{G}_o = \sum_{k=0}^{\infty} (\mathbf{A}^T)^k \mathbf{C}^T \mathbf{C} \mathbf{A}^k \quad \text{e} \quad \mathbf{G}_c = \sum_{k=0}^{\infty} \mathbf{A}^k \mathbf{K} \mathbf{K}^T (\mathbf{A}^T)^k \quad (7.11)$$

O algoritmo de Aoki é executado em 7 etapas de paralelização:

- Etapa 1 - Geração simultânea das matrizes  $\mathbf{H}^A$ ,  $\mathbf{H}^M$ ,  $\mathbf{H}^C$ ,  $\mathbf{H}$ ,  $\mathbf{Y}_-$  e  $\mathbf{Y}_+$ .

A partir dos dados de saída do sistema podemos gerar duas matrizes,  $\mathbf{Y}_-$  e  $\mathbf{Y}_+$  dadas por:

$$\mathbf{Y}_- = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & \cdots & \mathbf{y}_{N-1} \\ 0 & \mathbf{y}_1 & \mathbf{y}_2 & \cdots & \mathbf{y}_{N-2} \\ 0 & 0 & \mathbf{y}_1 & \cdots & \mathbf{y}_{N-3} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathbf{y}_{N-K} & \mathbf{y}_{N-K+1} \\ 0 & 0 & \cdots & \mathbf{y}_{N-K-1} & \mathbf{y}_{N-K} \end{bmatrix} \quad \text{onde} \quad \mathbf{Y}_- \in \Re^{(Kl) \times (N-1)} \quad (7.12)$$

$$\mathbf{Y}_+ = \begin{bmatrix} \mathbf{y}_2 & \mathbf{y}_3 & \mathbf{y}_4 & \cdots & \mathbf{y}_{N-1} & \mathbf{y}_N \\ \mathbf{y}_3 & \mathbf{y}_4 & \mathbf{y}_5 & \cdots & \mathbf{y}_N & 0 \\ \mathbf{y}_4 & \mathbf{y}_5 & \mathbf{y}_6 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \mathbf{y}_J & \mathbf{y}_{J+1} & \mathbf{y}_{J+2} & \cdots & 0 & 0 \\ \mathbf{y}_{J+1} & \mathbf{y}_{J+2} & \mathbf{y}_{J+3} & \cdots & 0 & 0 \end{bmatrix} \quad \text{onde} \quad \mathbf{Y}_+ \in \Re^{(Jl) \times (N-1)} \quad (7.13)$$

O produto das matrizes  $\mathbf{Y}_+$  e  $\mathbf{Y}_-^T$  permite gerar uma matriz de Hankel,  $\mathbf{H}$ , sem calcular explicitamente as covariâncias  $\mathbf{\Lambda}_k$ . Assim,

$$\mathbf{H}_{J,K} = \frac{\mathbf{Y}_+ \mathbf{Y}_-^T}{N}, \quad (7.14)$$

então nós obtemos uma matriz de Hankel tal que:

$$\mathbf{H}_{J,K} = \begin{bmatrix} \mathbf{C}\mathbf{M} & \mathbf{C}\mathbf{A}\mathbf{M} & \cdots & \mathbf{C}\mathbf{A}^{K-1}\mathbf{M} \\ \mathbf{C}\mathbf{A}\mathbf{M} & \mathbf{C}\mathbf{A}^2\mathbf{M} & \cdots & \mathbf{C}\mathbf{A}^K\mathbf{M} \\ \mathbf{C}\mathbf{A}^2\mathbf{M} & \mathbf{C}\mathbf{A}^3\mathbf{M} & \cdots & \mathbf{C}\mathbf{A}^{K+1}\mathbf{M} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{C}\mathbf{A}^{J-1}\mathbf{M} & \mathbf{C}\mathbf{A}^J\mathbf{M} & \cdots & \mathbf{C}\mathbf{A}^{J+K-1}\mathbf{M} \end{bmatrix} = \begin{bmatrix} \mathbf{\Lambda}_1 & \mathbf{\Lambda}_2 & \cdots & \mathbf{\Lambda}_K \\ \mathbf{\Lambda}_2 & \mathbf{\Lambda}_3 & \cdots & \mathbf{\Lambda}_{K+1} \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{\Lambda}_J & \mathbf{\Lambda}_{J+1} & \cdots & \mathbf{\Lambda}_{J+K} \end{bmatrix} \quad (7.15)$$

Ainda, dos dados de saída do sistema podemos gerar três novas matrizes:

$$\mathbf{H}^A = \begin{bmatrix} \mathbf{\Lambda}_2 & \mathbf{\Lambda}_3 & \dots & \mathbf{\Lambda}_{K+1} \\ \mathbf{\Lambda}_3 & \mathbf{\Lambda}_4 & \dots & \mathbf{\Lambda}_K \\ \mathbf{\Lambda}_4 & \mathbf{\Lambda}_5 & \dots & \mathbf{\Lambda}_K \\ \vdots & \vdots & \vdots & \vdots \\ \mathbf{\Lambda}_{J+1} & \mathbf{\Lambda}_{J+2} & \dots & \mathbf{\Lambda}_{J+K+1} \end{bmatrix}, \quad \mathbf{H}^M = \begin{bmatrix} \mathbf{\Lambda}_1 \\ \mathbf{\Lambda}_2 \\ \mathbf{\Lambda}_3 \\ \vdots \\ \mathbf{\Lambda}_J \end{bmatrix} \text{ e } (7.16)$$

$$\mathbf{H}^C = [\mathbf{C}\mathbf{M} \quad \mathbf{C}\mathbf{A}\mathbf{M} \quad \dots \quad \mathbf{C}\mathbf{A}^{K-1}\mathbf{M}] = [\mathbf{\Lambda}_1 \quad \mathbf{\Lambda}_2 \quad \dots \quad \mathbf{\Lambda}_K]. \quad (7.17)$$

bem como

$$\mathbf{\Lambda}_0 = E(\mathbf{y}_k \mathbf{y}_k^T) = \mathbf{C}\mathbf{\Pi}\mathbf{C}^T + \mathbf{\Delta}, \quad (7.18)$$

e

$$\mathbf{\Lambda}_k = E(\mathbf{y}_{i+k} \mathbf{y}_i) = \mathbf{C}\mathbf{A}^{k-1}\mathbf{M}, \quad k > 1 \quad (7.19)$$

onde  $\mathbf{\Lambda}_0$  é a autocovariância do vetor de saída, e onde considera-se que o ruído de entrada tenha uma matriz de covariância dada por:

$$E(\mathbf{e}_k \mathbf{e}_s^T) = \mathbf{\Delta} \delta_{t,s}, \quad \mathbf{\Delta} > 0 \quad (\text{condição de regularidade}) \quad (7.20)$$

e que o estado possua uma matriz de covariância dada por:

$$E(\mathbf{x}_k \mathbf{x}_k^T) = \mathbf{\Pi}, \quad (7.21)$$

com a matriz  $\mathbf{\Pi}$  constante.

- Etapa 2 - A decomposição em valores singulares da matriz  $\mathbf{H}_{J,K}$  é obtida :

$$\mathbf{H}_{J,K} = \mathbf{U}\mathbf{\Sigma}^{1/2}\mathbf{\Sigma}^{1/2}\mathbf{V}^T. \quad (7.22)$$

Um problema importante na área da computação matricial é a determinação da decomposição em valores singulares da matriz  $\mathbf{H}_{J,K}$ . Por sua importância nesta tese, este problema recebeu um tratamento teórico aprofundado nos Capítulos 4 e 5 e sua solução computacional

que também apresenta dificuldades consideráveis, é tratada adequadamente nos Capítulos 3, 6 e 7.

Dada uma matriz  $\Sigma \in \mathbb{R}^{m \times n}$ , sempre existe a decomposição

$$\mathbf{U}^T \Sigma \mathbf{V} = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_r) \in \mathbb{R}^{m \times n} \quad r = \min\{m, n\} \quad (7.23)$$

com  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$  e  $\mathbf{U}$  e  $\mathbf{V}$  matrizes unitárias.

Os números  $\sigma_i$  são valores singulares de  $\Sigma$ , as colunas de  $\mathbf{U}$  são os vetores singulares de  $\Sigma$  à esquerda e formam uma base para o espaço coluna de  $\Sigma \Sigma^T$ ; as colunas de  $\mathbf{V}$  são os vetores singulares de  $\Sigma$  à direita e formam uma base para o espaço coluna de  $\Sigma^T \Sigma$ .

Por inspeção do conjunto de valores singulares de uma matriz, pode-se determinar seu posto numérico, visto que seu menor valor singular é uma medida da distância da matriz ao espaço vetorial das matrizes reais  $m \times n$  com posto incompleto, Golub *et al.* [399].

- Etapa 3 - Neste nível são calculadas em paralelo as matrizes do sistema (7.1):

$$\mathbf{A} = \Sigma^{-1/2} \mathbf{U}^T \mathbf{H}^A \mathbf{V} \Sigma^{-1/2}, \quad (7.24)$$

$$\mathbf{C} = \mathbf{H}^C \mathbf{V} \Sigma^{-1/2}. \quad (7.25)$$

e a matriz  $\mathbf{M}$ , é dada por:

$$\mathbf{M} = \Sigma^{-1/2} \mathbf{U}^T \mathbf{H}^M, \quad (7.26)$$

onde:

$$\mathbf{M} = \mathbf{A} \Pi \mathbf{C}^T + \mathbf{K} \Delta. \quad (7.27)$$

- Etapa 4 - Construimos a matriz simplética:

$$\Phi = \begin{bmatrix} \Psi - \hat{\mathbf{Q}} \Psi^{-T} \mathbf{D} & \hat{\mathbf{Q}} \Psi^{-T} \\ -\Psi^{-T} \mathbf{D} & \Psi^{-T} \end{bmatrix}, \quad (7.28)$$

com  $\Psi^{-T} \equiv (\Psi^{-1})^{-T}$  e onde as matrizes auxiliares são dadas por:

$$\Psi = \mathbf{A}^T - \mathbf{C}^T \Lambda_0^{-1} \mathbf{M}^T, \quad \hat{\mathbf{Q}} = \mathbf{C}^T \Lambda_0^{-1} \mathbf{C}, \quad \text{e } \mathbf{D} = \mathbf{M} \Lambda_0^{-1} \mathbf{M}^T. \quad (7.29)$$

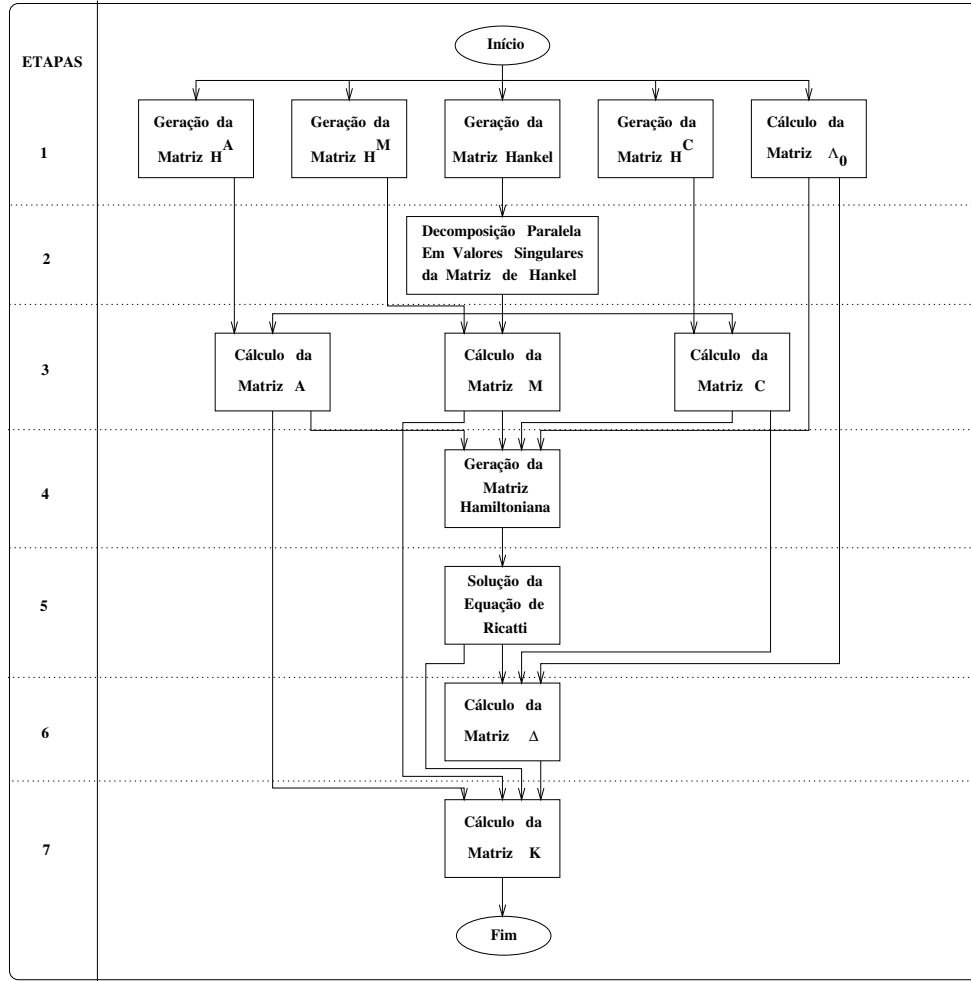


Figura 7.1: Implementação Paralela do Algoritmo de Aoki.

- Etapa 5 - A matriz  $\Pi$ , solução da equação de Ricatti, é determinada a partir da equação:

$$\Pi = A\Pi A^T + \mathcal{F}(\Pi) \quad (7.30)$$

onde

$$\mathcal{F}(\Pi) = (M - A\Pi C^T)(\Lambda_0 - (C\Pi C^T)^{-1}(M - A\Pi C^T)^T. \quad (7.31)$$

A solução  $\Pi$  pode ser obtida a partir da decomposição de Schur real da matriz  $\Phi$ :

$$\Phi = \mathbf{W}\mathbf{T}\mathbf{W}^T. \quad (7.32)$$

Esta decomposição é obtida a partir do método **QR**.

Particionando a matriz unitária  $\mathbf{W}$  conforme:

$$\mathbf{W} = \begin{bmatrix} \mathbf{W}_{11} & \mathbf{W}_{12} \\ \mathbf{W}_{21} & \mathbf{W}_{22} \end{bmatrix}, \quad (7.33)$$

enunciamos a seguinte proposição, Vaughan [1099]:

**Proposição :** *A matriz  $\Pi = \mathbf{W}_{21}\mathbf{W}_{11}^{-1}$  é a solução da Equação de Riccati*

Uma nova proposta para Solução Computacional da Equação Algébrica de Riccati em formas Sequencial e Paralela é apresentada por Tamariz, [1030], e foi utilizada neste Capítulo.

- Etapa 6 - Cálculo da matriz  $\Delta$  , onde  $\Delta$  é dada por:

$$\Delta = \Lambda_0 - \mathbf{C}\Pi\mathbf{C}^T \quad (7.34)$$

- Etapa 7 - Obtenção da matriz  $\mathbf{K}$ :

$$\mathbf{K} = (\mathbf{M} - \mathbf{A}\Pi\mathbf{C}^T)\Delta^{-1} \quad (7.35)$$

As etapas de paralelização estão representadas na Figura 7.1.

Na etapa inicial do processamento são construídas matrizes de covariância a partir do mesmo conjunto de dados, visto que estas matrizes são independentes e podem ser distribuídas entre os vários processadores e executadas simultaneamente, Stevens, [999]. Os resultados parciais são então combinados por um processador gerente que armazena e envia as matrizes obtidas mais recentemente aos processadores que executarão as próximas tarefas. O cálculo final das matrizes do sistema depende da finalização do cálculo dos valores singulares da matriz de Hankel. A decomposição em valores singulares da matriz de Hankel é executada em paralelo num ambiente multi processado, pois esta tarefa é a que exige maior desempenho computacional, Golub *et al.*, [400].

As etapas da decomposição em valores singulares da matriz e da solução paralela da equação de Riccati por serem de maior demanda de processamento são destacadas nas próximas seções.



### 7.2.1 Solução Paralela da Decomposição em Valores Singulares

A *SVD* é escrita como  $\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$  onde  $\mathbf{\Sigma}$  é uma matriz  $m \times n$  cujos elementos são iguais a zero exceto os seus  $\min(m, n)$  elementos de diagonal,  $\mathbf{U}$  é uma matriz ortogonal  $m \times m$ , e  $\mathbf{V}$  é uma matriz ortogonal  $n \times n$ . Os elementos da diagonal de  $\mathbf{\Sigma}$  são os valores singulares de  $\mathbf{A}$  e as colunas de  $\mathbf{U}$  e  $\mathbf{V}$  são os correspondentes vetores singulares, esquerdos e direitos, respectivamente, como já vimos nos Capítulos anteriores.

Os valores singulares são retornados em  $\mathbf{\Sigma}$  em ordem decrescente e só as primeiras  $\min(m, n)$  colunas de  $\mathbf{U}$  e linhas de  $\mathbf{V}^T$  são calculadas. Se  $m \gg n$ , um melhor desempenho computacional é obtido por meio da fatoração  $\mathbf{QR}$  de  $\mathbf{A}$ ; então a decomposição em valores singulares é realizada em  $\mathbf{R}$  que é uma matriz  $n \times n$ ; desta forma, se  $\mathbf{A} = \mathbf{QR}$  e  $\mathbf{R} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$ , então a *SVD* de  $\mathbf{A}$  é tal que  $\mathbf{A} = (\mathbf{QU})\mathbf{\Sigma}\mathbf{V}^T$ . Observe que nesta transformação os valores singulares da matriz  $\mathbf{A}$  são preservados.

Para realizar a implementação paralela do algoritmo, as linhas ou colunas da matriz são distribuídas em uma grade de processadores de duas dimensões, pois as subrotinas da *ScaLAPACK* utilizadas, Blackford, [120], são baseadas em algoritmos com particionamento em blocos, o que minimiza a frequência de movimento de dados entre os diferentes níveis de memória hierárquica. Ou seja, em cada passo do algoritmo um bloco de linhas ou colunas distribuídas na grade é atualizado, e a maioria da computação é executada por operações matriz-matriz nestes blocos.

### 7.2.2 Solução Paralela da Equação Algébrica de Ricatti

A obtenção da solução paralela da equação algébrica de Riccati foi dividida em duas partes: a primeira corresponde à formação da matriz em forma de Hessenberg via TSEE,<sup>1</sup> utilizando como plataforma de programação paralela o MPI - Message Passing Interface, com memória distribuída, comunicação síncrona, e semântica bloqueante das funções para enviar e receber mensagens com modo de programação SPMD; a segunda parte corresponde à obtenção da matriz em Forma Real de Schur que foi implementada utilizando as subrotinas da *ScaLAPACK*.

A biblioteca de alto desempenho com rotinas de Álgebra Linear, *ScaLAPACK*, para computadores MIMD<sup>2</sup> com paradigma de troca de mensagens e memória

<sup>1</sup>TSEE: transformações de similaridade elementares estabilizadas.

<sup>2</sup>Esta arquitetura combina paralelismo nos fluxos de dados e de instruções; é composta de unidades de processamento, cada qual capaz de realizar operações lógicas e aritméticas

*distribuída em redes de estações de trabalho suporta PVM - Parallel Virtual Machine ou MPI. Utiliza também como parte de suas rotinas a biblioteca sequencial BLAS - Basic Linear Algebra Subroutines, a paralela PBLAS - Parallel BLAS, os subprogramas de comunicação da álgebra linear básica BLACS - Basic Linear Algebra Communication Subroutines e as rotinas da LAPACK - Linear Algebra PACKage .*

*Similarmente ao LAPACK, as rotinas da ScaLAPACK são baseadas em algoritmos com particionamento em blocos, o que minimiza a frequência de movimento de dados entre os diferentes níveis da memória hierárquica. Os algoritmos são apresentados como processos e a comunicação pode ser ponto-a-ponto ou em grupo quando necessário. De forma geral, a ScaLAPACK foi desenvolvida e testada para o caso de um processo por processador.*

*Interessa-nos a implementação paralela do método de Schur para o caso de matrizes reais não simétricas utilizando TSEE: método de Schur-Modificado paralelo.*

*Consideramos a paralelização do algoritmo de Schur-Modificado para matrizes densas, não simétricas sobre um sistema de memória distribuída. O modo de programação paralela utilizado é SPMD, ou seja todos os processadores vão executar o mesmo programa, porém podem executar tarefas diferentes sobre dados diferentes. A estratégia de paralelização aplicada segue um esquema assíncrono. Descreve-se as principais subrotinas utilizadas na solução do problema apresentado.*

*A implementação paralela da primeira parte da solução da equação algébrica de Ricatti que corresponde a formação da matriz de Hessenberg via TSEE é baseada em operações do tipo produto matriz-matriz, que representam maior custo computacional; essas operações são efetuadas acessando os elementos da matriz por linhas; desta maneira formamos blocos que são distribuídos entre os diferentes processadores participantes do processamento, Figura 7.2. Neste caso é possível paralelizar o produto tal que cada processador calcule os valores dos elementos diferentes do vetor resultante. O método para acessar os elementos da matriz é a Distribuição por linhas: os  $P$  processadores que participam do processamento paralelo podem ser apresentados como um arranjo linear uni-dimensional, onde os elementos são distribuídos em blocos. Assim ao distribuir uma matriz  $(6 \times 6)$  em 4 processadores, numeramos os processadores de 0 até  $P - 1$  e as linhas da matriz de 1 até  $N$  ( $N = 6$ ). A Tabela 7.1 apresenta uma distribuição de blocos por linha, onde cada linha ou submatriz formada é nomeada por um número de processador. Neste caso, cada processador armazena no máximo  $(NB)$  número de linhas, onde  $NB = \lceil N/P \rceil$  e a linha  $k$  é armazenada no processador  $\lfloor k/NB \rfloor$ . A*

---

padrões. As unidades operam assincronamente sob o controle de fluxo de instruções individuais.

distribuição atribui blocos de linhas de tamanho  $(NB)$  aos processadores sucessivos. Se o valor de  $P$  é divisível pelo valor de  $N$ , então cada processador vai receber blocos do mesmo tamanho, de forma a garantir um melhor balanceamento de carga.

0
1
2
3

Tabela 7.1: Armazenagens por linha

Com o modo de programação SPMD, cada processador participante realiza uma operação de multiplicação ( $C = A * B$ ) com os dados enviados pelo processador principal, ou seja, os dados são divididos e cada processador calcula no máximo  $(NB)$  blocos da matriz resultante na multiplicação matriz-matriz que realiza. O processador principal, encarregado de distribuir os dados e realizar seus cálculos envia no máximo  $(NB)$  blocos da matriz  $A$   $(NB, N)$  e a matriz  $B$   $(N, N)$  para cada processador participante do processo e recebe uma matriz de dimensão  $(NB, N)$  com o produto final de cada processador. A Figura 7.2, ilustra o procedimento descrito para um sistema de ordem  $(6 \times 6)$  onde os processadores recebem as linhas seguintes:

- $P_0$  = linha  $A_{1n}$  da matriz (processador principal)
- $P_1$  = linhas  $A_{2n}$  e  $A_{3n}$  da matriz
- $P_2$  = linhas  $A_{4n}$  e  $A_{5n}$  da matriz
- $P_3$  = linha  $A_{6n}$  da matriz

A implementação paralela da segunda parte do algoritmo proposto para solução da equação algébrica de Ricatti, a correspondente a achar a Forma Real de Schur, é baseada na utilização das subrotinas definidas na biblioteca ScaLAPACK.

Na implementação estão presentes os mecanismos básicos *ScaLAPACK*, tais como criação da topologia malha, comunicação entre tarefas e entrada-saída de dados. A ScaLAPACK utiliza uma distribuição de dados com blocos cíclicos de dimensão-2, na qual a matriz, dividida em blocos de linhas e colunas  $(MB \times NB)$ , é atribuída ao mesmo processador. A distribuição de dados cíclica é parametrizada por quatro valores:  $P$ ,  $Q$ ,  $r$  e  $c$ , onde  $(P \times Q)$  é a definição da topologia malha utilizada e  $(r \times c)$  é a dimensão do bloco. Supõe-se que os processadores estão arranjados logicamente como uma malha de  $P$

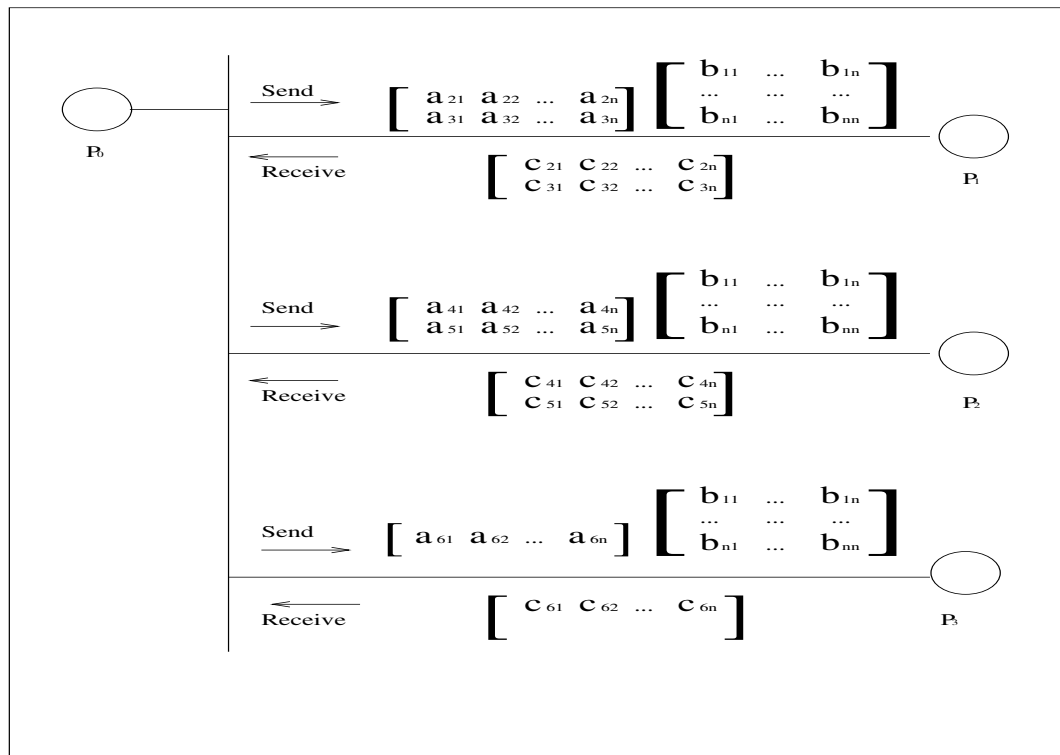


Figura 7.2: Processo Paralelo de Distribuição de matrizes.

linhas e  $Q$  colunas; assim o número de processadores utilizados corresponde ao produto  $R = P \times Q$ .

Após o particionamento da matriz ( $N \times N$ ) em blocos ( $MB \times NB$ ), o primeiro bloco é associado ao processador com coordenadas ( $RSRC, CSRC$ ), onde  $RSRC$  e  $CSRC$  correspondem respectivamente aos números dos processos linha e coluna aos quais o primeiro bloco da matriz é atribuído. Seja a entrada ( $I, J$ ); cada elemento da matriz global está localizado no processador definido pelas coordenadas ( $p_r, p_c$ ), do bloco local ( $l, m$ ), na posição ( $x, y$ ), onde:

$$\begin{aligned} (l, m) &= (\lfloor (I - 1) / (P * MB) \rfloor, \lfloor (J - 1) / (Q * NB) \rfloor), \\ (p_r, p_c) &= ((RSRC + \lfloor (I - 1) / MB \rfloor) \bmod P, (CSRC + \lfloor (J - 1) / NB \rfloor) \bmod Q), \\ (x, y) &= (\bmod(I - 1, MB) + 1, \bmod(J - 1, NB) + 1). \end{aligned}$$

Apresentamos na Figura 7.3 o mapeamento de uma matriz  $A$  ( $6 \times 6$ ) particionada em blocos ( $2 \times 2$ ) e na Figura 7.4 o ponto de vista do processador para a distribuição da matriz ( $6 \times 6$ ) em uma topologia malha ( $2 \times 2$ ) processadores, onde podemos observar que a entrada local  $A(2, 3)$ , definida no processador de coordenadas  $(0, 0)$ , contém o valor da entrada da matriz global  $a_{25}$ .

$a_{11}$	$a_{12}$	$a_{13}$	$a_{14}$	$a_{15}$	$a_{16}$
$a_{21}$	$a_{22}$	$a_{23}$	$a_{24}$	$a_{25}$	$a_{26}$
$a_{31}$	$a_{32}$	$a_{33}$	$a_{34}$	$a_{35}$	$a_{36}$
$a_{41}$	$a_{42}$	$a_{43}$	$a_{44}$	$a_{45}$	$a_{46}$
$a_{51}$	$a_{52}$	$a_{53}$	$a_{54}$	$a_{55}$	$a_{56}$
$a_{61}$	$a_{62}$	$a_{63}$	$a_{64}$	$a_{65}$	$a_{66}$

Figura 7.3: Matriz  $6 \times 6$  particionada em blocos  $2 \times 2$ .

Existe uma subrotina da *ScaLAPACK* denominada *BLACS-GRIDINIT* que realiza o mapeamento dos elementos da matriz global em cada processador por linha, ou seja vai distribuindo os elementos ordenados por linha, veja a definição na Tabela 7.2, onde quatro processos são distribuídos em uma malha ( $2 \times 2$ ) com uma distribuição por linha:

Para um exemplo de ordem ( $20 \times 20$ ), os parâmetros da subrotina que inicializam a topologia malha vão estar definidos como:  $MB=NB=6$ ,  $P=Q=2$  e  $LLD=12$ , onde  $LLD$  corresponde aos números máximos de linhas e ou colunas em cada bloco da submatriz; para o exemplo que estamos apresentando os parâmetros estão definidos como:  $MB = NB = 6$ ,  $P = Q = 2$  e  $LLD = 6$ .

	0				1	
0	$a_{11}$	$a_{12}$	$a_{15}$	$a_{16}$	$a_{13}$	$a_{14}$
	$a_{21}$	$a_{22}$	$a_{25}$	$a_{26}$	$a_{23}$	$a_{24}$
	$a_{51}$	$a_{52}$	$a_{55}$	$a_{56}$	$a_{53}$	$a_{54}$
	$a_{61}$	$a_{62}$	$a_{65}$	$a_{66}$	$a_{63}$	$a_{64}$
1	$a_{31}$	$a_{32}$	$a_{35}$	$a_{36}$	$a_{33}$	$a_{34}$
	$a_{41}$	$a_{42}$	$a_{45}$	$a_{46}$	$a_{43}$	$a_{44}$

Figura 7.4: Matriz  $6 \times 6$  em Topologia Malha  $2 \times 2$ .

0	1
2	3

Tabela 7.2: Quatro Processos Mapeados em Malha  $2 \times 2$ .

## 7.3 Resultados Experimentais

### 7.3.1 Caso A

Considerando o seguinte sistema “*benchmark*” no espaço de estado:

$$\mathbf{A} = \begin{bmatrix} -0.735 & -0.363 \\ 0.333 & -0.565 \end{bmatrix}, \quad \mathbf{K} = \begin{bmatrix} -1.502 & -0.949 \\ -0.945 & -0.088 \end{bmatrix} \quad \text{e} \quad \mathbf{C} = \begin{bmatrix} -1.438 & -0.680 \\ 1.067 & -0.531 \end{bmatrix}$$

obtém-se uma seqüência de saída multivariada da seguinte forma:

$$\mathbf{y}(k) = \begin{bmatrix} \mathbf{y}_1(k) \\ \mathbf{y}_2(k) \end{bmatrix}, \quad k = 1, 2, 3, \dots, 60. \quad (7.36)$$

A matriz de Hankel foi gerada com as covariâncias  $\mathbf{\Lambda}_1$ ,  $\mathbf{\Lambda}_2$  e  $\mathbf{\Lambda}_3$  e em sua decomposição em valores singulares, os valores singulares  $\sigma_i$ ,

$$\frac{\sigma_{max}}{\sigma_i} > Z \quad (7.37)$$

foram desconsiderados, onde  $Z$  é uma constante determinada experimentalmente. Bons resultados foram obtidos com  $10 \leq Z \leq 100$ , para este experimento. Na Figura 7.5 o eixo das abcissas corresponde ao número de observações indexadas no tempo e o eixo das ordenadas ao vetor de saída da realização original e ou da realização calculada.

Para avaliação dos resultados obtidos, comparamos a seqüência calculada  $\{\mathbf{y}_c(k)\}$ , (Fig. 7.5 C, D -  $\mathbf{y}_{c1}(k)$  e  $\mathbf{y}_{c2}(k)$ ), gerada com as matrizes  $\mathbf{A}_c$ ,

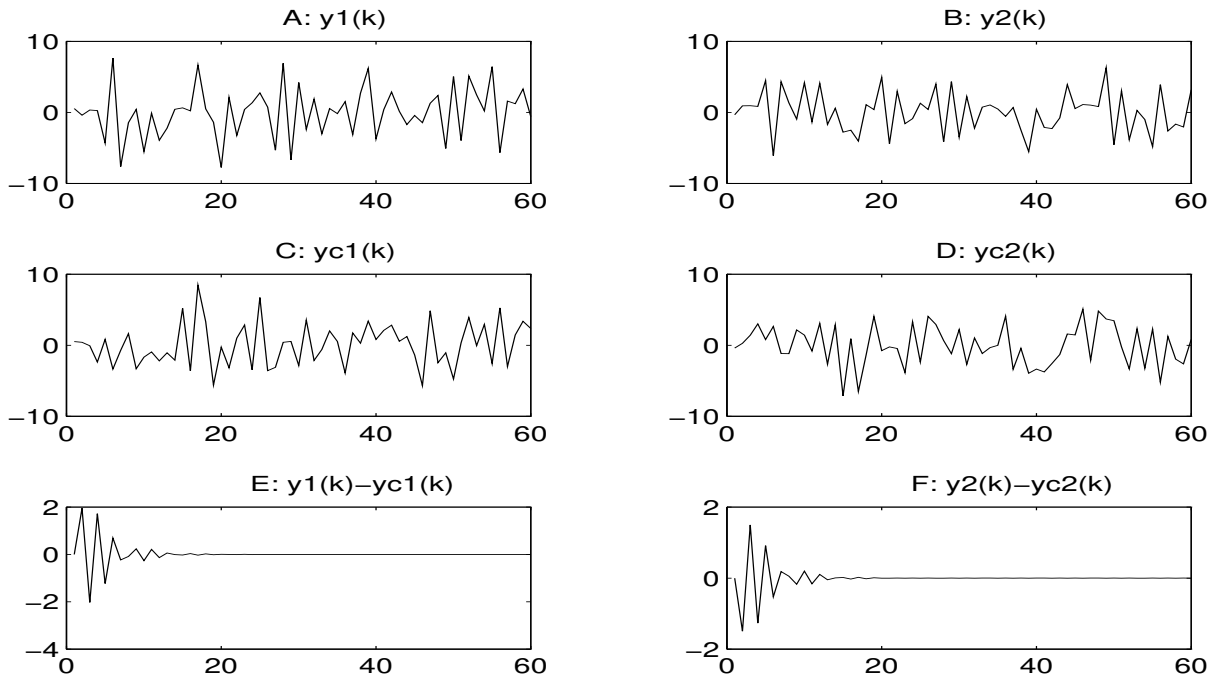


Figura 7.5: Comparações de Resultados do Algoritmo Paralelo.

$\mathbf{K}_c$  e  $\mathbf{C}_c$  calculadas, com a sequência vetorial  $\{\mathbf{y}(k)\}$ , (Fig. 7.5 A, B, -  $\mathbf{y}_1(k)$  e  $\mathbf{y}_2(k)$ ), ambas obtidas em resposta à mesma excitação de ruído branco. Consideramos, também, a diferença entre as respostas impulsivas ( $\mathbf{e}_1(k) = \delta$ ,  $\mathbf{e}_2(k) = 0$ ) para as duas saídas, (Fig. 7.5 E, F -  $\mathbf{y}_1(k) - \mathbf{y}_{c1}(k)$  e  $\mathbf{y}_2(k) - \mathbf{y}_{c2}(k)$ ).

Para o processamento do algoritmo considera-se uma rede paralela e distribuída com 5 computadores para a qual obtém-se a razão de aceleração 1.67.

### 7.3.2 Caso B

Considerando o seguinte sistema “*benchmark*” no espaço de estado:

$$\mathbf{A} = \begin{bmatrix} -0.5465 & 0.6630 & -0.1199 \\ -0.8468 & -0.8542 & -0.0653 \\ -0.2463 & -1.2013 & 0.4853 \end{bmatrix}; \quad \mathbf{K} = \begin{bmatrix} -0.5955 & -0.0793 \\ -0.1497 & 1.5352 \\ -0.4348 & -0.6065 \end{bmatrix}$$

e

$$\mathbf{C} = \begin{bmatrix} -1.3474 & -0.9036 & -0.6275 \\ 0.4694 & 0.0359 & 0.5354 \end{bmatrix}$$

geramos a série temporal multivariada,  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$

O algoritmo distribuído e paralelo proposto foi aplicado na modelagem desta série temporal multivariada,  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$ ,  $\mathbf{y}_k \in \mathbb{R}^2$ , gerando a seguinte tripla de matrizes:

$$\mathbf{A}_N = \begin{bmatrix} -0.2497 & -0.6753 & -0.0392 \\ 0.7490 & 0.3505 & -0.0535 \\ 0.0045 & 0.0649 & -0.4679 \end{bmatrix}; \quad \mathbf{K}_N = \begin{bmatrix} -0.2225 & 0.2349 \\ 0.3285 & 0.6982 \\ 0.0597 & 0.1756 \end{bmatrix}$$

e

$$\mathbf{C}_N = \begin{bmatrix} -2.6917 & 1.6029 & -0.1129 \\ -0.2622 & -1.0398 & -0.0745 \end{bmatrix}$$

Da sequência  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$  tomamos os  $q$  primeiros vetores  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q\}$  com  $q \ll N$  e calculamos a tripla de matrizes do modelo pelo algoritmo descrito. As matrizes assim obtidas são chamadas de  $\mathbf{A}_q$ ,  $\mathbf{K}_q$  e  $\mathbf{C}_q$ . O mesmo procedimento é repetido para  $2q, 3q, \dots, N$  e assim obtém-se uma sequência de triplas formadas com os dados disponíveis desde  $k = 1$  até  $k = rq$ , sendo  $r = N/q$ .

A seguir apresentamos as três primeiras triplas de matrizes obtidas por este procedimento:

$$\mathbf{A}_{1q} = \begin{bmatrix} -0.1795 & -0.7008 & -0.0684 \\ 0.7542 & 0.4025 & -0.1320 \\ 0.0529 & -0.0127 & -0.5074 \end{bmatrix}; \quad \mathbf{K}_{1q} = \begin{bmatrix} -0.2243 & 0.1940 \\ 0.2599 & 0.6136 \\ 0.2930 & 0.3757 \end{bmatrix}$$

e

$$\mathbf{C}_{1q} = \begin{bmatrix} -2.8609 & 1.6482 & -0.3140 \\ -0.1056 & -0.9674 & -0.3234 \end{bmatrix};$$

$$\mathbf{A}_{2q} = \begin{bmatrix} -0.2322 & -0.6571 & -0.1596 \\ 0.7278 & 0.3926 & -0.1402 \\ 0.1428 & -0.1056 & 0.5753 \end{bmatrix}; \quad \mathbf{K}_{2q} = \begin{bmatrix} -0.2187 & 0.2190 \\ 0.2821 & 0.6441 \\ 0.2501 & 0.3228 \end{bmatrix}$$

e

$$\mathbf{C}_{2q} = \begin{bmatrix} -2.8214 & 1.6826 & -0.2861 \\ -0.1910 & -1.0127 & -0.1833 \end{bmatrix};$$

$$\mathbf{A}_{3q} = \begin{bmatrix} -0.2388 & -0.6636 & -0.1326 \\ 0.7331 & 0.3790 & -0.1149 \\ 0.1240 & -0.0921 & 0.6399 \end{bmatrix}; \quad \mathbf{K}_{3q} = \begin{bmatrix} -0.2086 & 0.2220 \\ 0.2911 & 0.6670 \\ 0.2039 & 0.2647 \end{bmatrix}$$

e

$$\mathbf{C}_{3q} = \begin{bmatrix} -2.8339 & 1.6876 & -0.2486 \\ -0.2097 & -1.0485 & -0.1816 \end{bmatrix}$$

sendo  $N = 8192$  e  $q = 512$ .



Na seção 7.3.3 são avaliados os tempos computacionais correspondentes à solução da equação algébrica de Ricatti de formas sequencial e paralela. Na seção 7.4 os modelos obtidos com dados intermediários são usados no projeto de preditores de um passo a frente por filtro de Kalman, e as covariâncias dos erros de predição em cada intervalo são avaliadas pelo seu traço conforme a Figura 7.7.

### 7.3.3 Resultados Computacionais: Equação Algébrica de Ricatti

Os resultados que apresentamos foram obtidos com diferentes valores dos parâmetros. Por exemplo, para uma matriz de ordem  $(6 \times 6)$  que definimos a seguir, os parâmetros são os seguintes:  $N=6$ ,  $P=3$ ,  $NB=2$ . Com este tamanho de matriz, obtemos tempos de execução paralela de cada uma das partes do programa maiores do que os tempos respectivos obtidos na implementação sequencial, veja Tabela 7.3; mas isso é diferente quando utilizamos  $P=2$ .

Atividade	SEQUENCIAL	PARALELO	
	1-proc	2-proc	3-proc
Leitura dos dados	0.1256	0.0296	0.0323
Matriz Hessenberg	0.1699	0.1060	0.2049
Forma Real de Schur	0.2228	0.1412	0.3623
Tempo Total de Execução	0.5183	0.2768	0.5995

Tabela 7.3: Tempos de Execução das partes do Algoritmo em uma matriz  $(6 \times 6)$

Com nossa implementação obtemos uma solução simétrica, definida não negativa da equação algébrica de Riccati com *TSEE*.

Quando testamos com matrizes de ordens maiores, por exemplo  $(50 \times 50)$  obtemos, Tabela 7.4, tempos de execução paralela de cada uma das partes do programa bem menores do que os tempos respectivos obtidos na implementação sequencial; assim podemos concluir que a paralelização surte melhores resultados quando o problema envolve cálculos com matrizes de ordens bem maiores.

Contrariamente, temos um exemplo teste com uma matriz de ordem  $(100 \times 100)$ , na qual os resultados não são favoráveis, ou seja obtemos tempos de execução paralela de cada uma das partes do programa bem maiores do que os tempos respectivos obtidos na implementação sequencial, veja Tabela 7.5. Acharmos que a causa deste problema esteja na definição do tamanho

Atividade	SEQUENCIAL	PARALELO	
	1-proc	2-proc	3-proc
Leitura dos dados	0.3092	0.2642	0.2869
Matriz Hessenberg	59.0045	8.3596	10.3933
Forma Real de Schur	1.8914	0.9707	1.3932
Tempo Total de Execução	61.2051	9.5945	12.0734

Tabela 7.4: Tempos de Execução das partes do Algoritmo em uma matriz ( $50 \times 50$ )

dos pacotes no sistema operacional ou na grande comunicação entre os processadores para obter o resultado final do problema.

Atividade	SEQUENCIAL	PARALELO	
	1-proc	2-proc	3-proc
Leitura dos dados	0.5088	0.2939	0.3293
Matriz Hessenberg	73.2661	121.6329	132.3660
Forma Real de Schur	1.2397	2.1775	3.2429
Tempo Total de Execução	75.6775	125.1755	137.1508

Tabela 7.5: Tempos de Execução das partes do Algoritmo em uma matriz ( $100 \times 100$ )

## 7.4 Avaliação da Modelagem de Séries Temporais Multivariadas no Espaço de Estado

A modelagem de dados nos leva à obtenção de modelos no espaço de estado que são somente uma aproximação do fenômeno dinâmico em estudo. Tradicionalmente a obtenção da função de transferência a partir do modelo estimado no espaço de estado nos fornece informação sobre a qualidade da estimação obtida, no entanto este resultado não é facilmente obtido e a validação do modelo é mais complicada do que comparar meramente números. Com esta preocupação apresentamos um procedimento de álgebra linear numérica que apresenta um valor da qualidade do modelo obtido a partir dos dados de entrada - saída do sistema original.

As triplas de matrizes obtidas são avaliadas por meio de um procedimento de comparação entre a sequência vetorial predita pelo modelo e a sequência

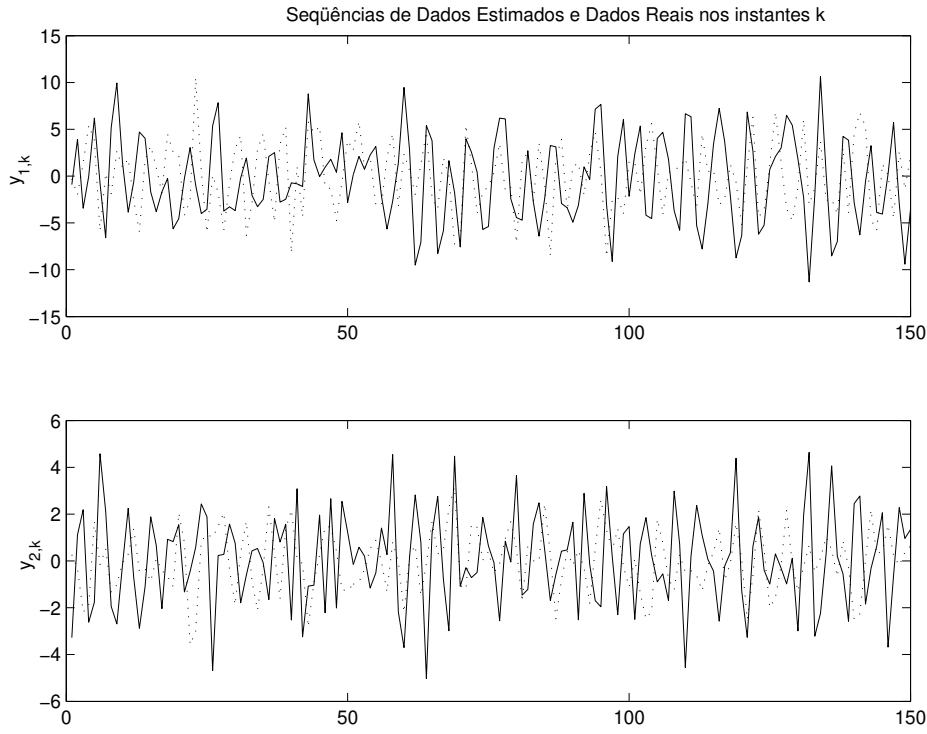


Figura 7.6: Valores medidos e valores estimados em função do instante  $k$ . A linha sólida representa os dados reais e a pontilhada os preditos.

vetorial do “*benchmark*”. A Figura 7.6 mostra o traçado de observações originais da série e de observações estimadas das últimas 150 amostras da sequência vetorial  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_N\}$ . A série temporal predita e o sinal original não são coincidentes em sua totalidade pois o sinal original é um processo puramente estocástico.

O teste de predição para avaliação de modelos obtidos é descrito a seguir. Dada a sequência  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q\}$  calcula-se a tripla  $\mathbf{A}_q$ ,  $\mathbf{K}_q$  e  $\mathbf{C}_q$ . Um estimador de um passo à frente é construído e usado para projetar o vetor de saída no subespaço gerado pelas saídas anteriores, dadas por  $\{\mathbf{y}_{2|1}, \mathbf{y}_{3|2}, \dots, \mathbf{y}_{q|q-1}\}$  com:

$$\mathbf{x}_{k+1|k} = (\mathbf{A} - \mathbf{BC})\mathbf{x}_{k|k-1} + \mathbf{B}\mathbf{y}_k$$

e

$$\mathbf{y}_{k+1|k} = \mathbf{C}\mathbf{x}_{k+1|k} = \mathbf{C}\mathbf{B}\mathbf{y}_k + \mathbf{C}(\mathbf{A} - \mathbf{BC})\mathbf{x}_{k|k-1}$$

Se o estimador for consistente, o vetor erro de predição  $\mathbf{e}_{k|k-1} = \mathbf{y}_{k|k-1} - \mathbf{y}_k$  terá uma matriz de covariância que se aproxima da covariância da inovação, conforme a quantidade  $q \rightarrow \infty$ . Nesta avaliação do modelo usamos os primeiros  $q$  vetores de observações para calcular as matrizes do modelo  $\mathbf{A}_q$ ,

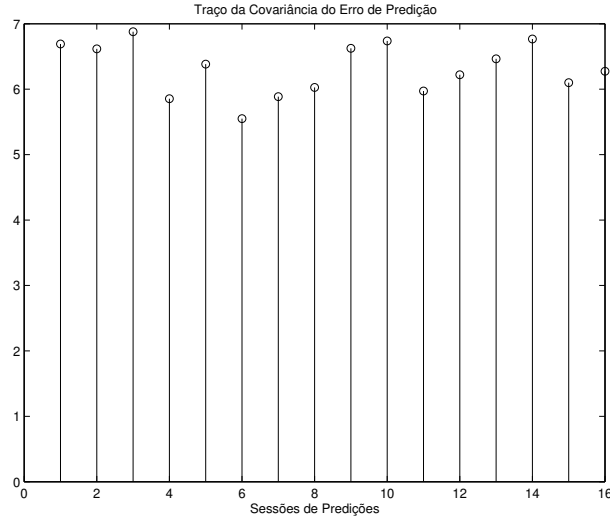


Figura 7.7: Traço da Covariância do Erro de Predição.

$\mathbf{K}_q$  e  $\mathbf{C}_q$ . A predição é feita sobre os vetores  $\{\mathbf{y}_1, \mathbf{y}_2, \dots, \mathbf{y}_q\}$ . Isto é feito para verificar se o modelo obtido pelas matrizes  $\mathbf{A}_q$ ,  $\mathbf{K}_q$  e  $\mathbf{C}_q$  representa a sequência de vetores originais  $\mathbf{y}_k$ ,  $k = 1, 2, \dots, q$ . Este procedimento de predição é repetido nos intervalos  $q + 1$  a  $2q$ ,  $2q + 1$  a  $3q$  e assim sucessivamente até  $(r - 1)q + 1$  a  $rq$ , sendo  $rq = N$ . A Figura 7.7 mostra os traços das matrizes de covariâncias do erro de predição conforme o intervalo de  $q = 512$  vetores aproxima-se da  $N$ -ésima observação. Vemos que o estimador formado com as matrizes calculadas mantém o traço da covariância do erro de estimação dentro de um intervalo fixo de 5.5 a 7.0, portanto próximo do traço da matriz de inovação,  $\Delta = E(\mathbf{e}_k \mathbf{e}_k^T)$ , calculada:

$$\Delta = \begin{bmatrix} 3.8912 & -1.8080 \\ -1.8080 & 2.1277 \end{bmatrix}$$

definida positiva, cujo valor do traço é 6.0188.

Para os parâmetros de Markov do modelo resultante

$$[\mathbf{C}_N \mathbf{M}_N, \mathbf{C}_N \mathbf{A}_N \mathbf{M}_N, \mathbf{C}_N \mathbf{A}_N^2 \mathbf{M}_N, \dots]$$

com

$$\mathbf{M}_N = \begin{bmatrix} -2.2081 & 1.8132 \\ -1.5392 & 0.3711 \\ 0.0694 & 0.2565 \end{bmatrix}$$

tem-se para as três primeiras realizações:

$$\begin{aligned}
\mathbf{H}_{1,3} &= \begin{bmatrix} \Lambda_1 & \Lambda_2 & \Lambda_3 \end{bmatrix} = \\
&= \begin{bmatrix} 3.4596 & -4.3192 & -7.7935 & 4.2835 & -2.2730 & 2.1491 \\ 2.1655 & -0.8811 & 1.8741 & -1.3480 & -0.7408 & 0.2078 \end{bmatrix} \\
\text{e} \quad & \begin{bmatrix} \mathbf{C}_N \mathbf{M}_N & \mathbf{C}_N \mathbf{A}_N \mathbf{M}_N & \mathbf{C}_N \mathbf{A}_N^2 \mathbf{M}_N \end{bmatrix} = \\
&= \begin{bmatrix} 3.4684 & -4.3147 & -7.7804 & 4.2937 & -2.2490 & 2.1558 \\ 2.1744 & -0.8805 & 1.8786 & -1.3395 & -0.7256 & 0.2170 \end{bmatrix}
\end{aligned}$$

Os parâmetros de Markov da matriz de Hankel montada com as covariâncias amostrais são próximos dos parâmetros de Markov calculados com a realização produzida pelo método implementado, respeitando a relação:

$$\Lambda_k = E(\mathbf{y}_{i+k}\mathbf{y}_i) = \mathbf{C}\mathbf{A}^{k-1}\mathbf{M}, \quad k > 1$$

## 7.5 Conclusões

Dos resultados obtidos podemos observar que:

- Em relação à comparação em tempo de execução dos algoritmos implementados em formas sequencial e paralela, podemos dizer que o processamento paralelo nem sempre foi melhor que o processamento sequencial quanto ao tempo de execução, pois no processamento paralelo existe uma alta comunicação de dados entre os diferentes processadores, o que implica que a solução possa ser mais demorada; contudo o processamento paralelo propiciou uma menor carga computacional para os processadores envolvidos.
- Em nossa implementação devido a utilização de uma rede de estações de trabalho, e a forma de implementar o algoritmo paralelo, o tempo de execução obtido para alguns casos foi bem maior do que o tempo de execução obtido no algoritmo sequencial, por causa da comunicação entre os diferentes processadores participantes do procedimento. Em muitos algoritmos e sistemas paralelos e ou distribuídos o tempo consumido pela comunicação entre processadores é uma fração elevada do tempo total necessário para resolver o problema.
- O particionamento inerente ao processamento paralelo e distribuído cria um problema de comunicação por exigir maior comunicação entre os processadores. A alteração do número de canais de comunicação

entre processadores começa a ser estudada através de estratégias de eletrônica reconfigurável, por exemplo com a utilização de máquinas computacionais baseadas em *FPGA* - *Field Programmable Gate Array*, onde é possível reconfigurar a topologia da máquina durante o processamento, bem como a criação e destruição de canais de comunicação de acordo com a demanda, o que viabiliza o processamento paralelo e distribuído de tal forma que o problema tratado crie eletronicamente canais de comunicação.

- Além da redução no tempo de execução do algoritmo pelo uso de processamento paralelo e distribuído, notamos que o uso de filtros de Kalman construídos com matrizes sistêmicas obtidas na modelagem produz estimativas próximas dos vetores originais da série, conforme Figura 7.6. Uma avaliação das covariâncias dos erros de predição em cada intervalo também é mostrada na Figura 7.7. Vemos que os traços das covariâncias dos erros de predição permanecem dentro de um intervalo limitado indicando estacionariedade do processo de estimação e que o estimador é consistente.

## Capítulo 8

# CONCLUSÕES E COMENTÁRIOS

Neste trabalho apresentamos metodologias para modelagem computacional distribuída e paralela no espaço de estado de dados de sistemas multivariáveis e de séries temporais multivariadas.

O problema de identificação multivariável no espaço de estado foi abordado com base no método *MOESP - MIMO Output-Error State sPace model identification* de Verhaegen & Dewilde.

O problema de modelagem de séries temporais multivariadas no espaço de estado foi abordado com base no método proposto por Masanao Aoki em seu livro “*State Space Modeling of Time Series*”.

Para desenvolver e implementar tais metodologias, estudos teóricos sobre fundamentos de modelagem de dados com base na teoria da realização no espaço de estado para sistemas determinísticos, desenvolvida por Kalman, foram realizados de forma personalizada e com contribuições originais, seja para uma visão bem estruturada e bem fundamentada da mesma, seja com relação à formulação, reformulação e/ou demonstração de alguns resultados básicos e indispensáveis ao bom entendimento e à boa aplicação desta teoria, nesta tese.

Da mesma forma, para a concretização do trabalho sobre fundamentos realizado no Capítulo 5, na maneira por nós desejada, foi indispensável a elaboração de estudos personalizados e com contribuições originais, com a devida profundidade teórica, desenvolvidos nos Capítulos 2 e 4.

Em especial acreditamos que os Capítulos 4 e 5 constituem contribuições importantes desta tese, não apenas para o seu objetivo central de modelagem de dados, como para o estudo teórico de sistemas lineares de forma mais avançada e atual, visando suas múltiplas aplicações e implicações em nosso meio.

No Capítulo 3 uma abordagem adequada aos objetivos desta tese foi apresentada sobre computação de alto desempenho, e alguns pontos de vista, bem como alguns desenvolvimentos originais também foram apresentados.

Os algoritmos distribuídos e paralelos propostos nos Capítulos 6 e 7 são contribuições muito relevantes desta tese e as implementações computacionais mostram a praticabilidade destes algoritmos. Algumas conclusões e resultados também são apresentadas nestes Capítulos.

Para continuação deste trabalho sugerimos: testar os algoritmos aqui propostos em máquina paralela; realizar implementações diferentes das que fizemos para os nossos algoritmos, procurando reduzir a comunicação em favor da computação; aplicar nossos resultados nas mais diversas áreas, como exemplo a aeroespacial e a econômica, tanto para modelagem puramente, como para controle multivariável baseado em modelo e para predição baseada em modelo multivariável de séries temporais.



# Referências Bibliográficas

- [1] Aasnaes, H. and T. Kailath, T.. *An innovations approach to least-squares estimation, Pt. VII: Some applications of vector autoregressive moving average models.* IEEE Trans. Automat. Control, AC-18(6):601-607, Dec. 1973.
- [2] Abdelmalek, N.N.. *On the Solution of The Linear Least Squares Problems and Pseudo-Inverses.* Computing, Vol.13, Fasc 3-4, 215-228, December, 1974.
- [3] Abed-Meraim, Karim and Hua, Yingbo. *Weighted Minimum Noise Subspace Method for Blind System Identification.* Signal Processing, Vol.69, Issue 1, 15-27, August, 1998.
- [4] Adamjan, V.M.; Arov, D.Z. and Krĕin, M.G.. *Analytic Properties of Schimidt Pairs for a Hankel Operator and The Generalized Schur-Takagi Problem.* Mathematics of the URSS - SBORNIK, Vol.15, No.1, 31-73, 1971.
- [5] Adeyemi, S.O.; Wu, S.M. and Berthouex. *Modeling and Control of a Phosphorus Removal Process By Multivariate Time Series Method.* Automatica, Pergamon Press, 105-112, 1979.
- [6] Afriat, S.N.. *Orthogonal and Oblique Projectors and The Characteristics of Pairs of Vector Spaces.* Proceedings of the Cambridge Philosophical Society - Mathematical and Physical Sciences - Volume 53, Part 4 - 800-816, October, 1957.
- [7] Ainsleigh, Phillip L.. *Observations on Oblique Projectors and Pseudoinverses.* IEEE Transactions on Signal Processing, Vol.45, No.7, 1886-1889, July, 1997.
- [8] Akaike, Hirotugu. *Stochastic Theory of Minimal Realization.* IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 667-674, December, 1974.
- [9] Akaike, Hirotugu. *A New Look At The Statistical Model Identification..* IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 716-723, December, 1974.
- [10] Akaike, Hirotugu. *Markovian Representation of Stochastic Processes By Canonical Variables.* SIAM Journal Control and Optimization, Vol. 13, number 1, pp.162-173, 1975.
- [11] Akaike, Hirotugu and Nakagawa, Toichiro. *Statistical Analysis and Control of Dynamic Systems.* KTK Scientific Publishers and Kluwer Academic Publishers, 1988.
- [12] Akçay, Hüseyin; Gu, Guoxiang and Khargonekar, Pramod P.. *A Class Algorithms for Identification in  $H_\infty$  : Continuos-Time Case.* IEEE Transactions on Automatic Control, Vol.38, No.2, 289-294, February, 1993.
- [13] Akke, Magnus and Lampe, Karl-Heinz. *Interaction of DC cables and Ship Steering Autopilots.* Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4086-4091, December, 1994.
- [14] Akman, Ibrahim and De Gooijer, Jan G. *Component extraction analysis of multivariate time series .* Computational Statistics & Data Analysis Volume: 21, Issue: 5, pp. 487-499, May, 1996.

- [15] Al-Saggaf, Ubaid M. and Franklin, Gene F.. *Model Reduction Via Balancead Realizations: An Extension and Frequency Weighting Techniques*. IEEE Transactions on Automatic Control, Vol.33, No.7, 687-692, July, 1988.
- [16] Albert, Arthur Lemuel. *Regression and The Moore-Penrose Pseudoinverse*. Academic Press, 1972.
- [17] Albert, Arthur E. and Gardner Jr., Leland A.. *Stochastic Approximation and Nonlinear Regression*. MIT Press, 1967.
- [18] Alengin, G.; Bucy, R.S.; Moura, J.M.F.; Pagés, J.; and Ribeiro, M.I.. *ARMA Identification*. Journal of Optimization Theory and Applications, Vol.55, No.3, 345-357, December, 1987.
- [19] Allison, Bruce J.; Isaksson, Alf J. and Karlström, Anders. *Grey-Box Identification of a TMP Refiner*. Internal Report IR-S3-REG-9513 -Department of Signals, Sensors and Systems, Automatic Control, Stockholm, Sweden, 1995.
- [20] Alvarado, Fernando L.. *Parallel Solution of Transient Problems by Trapezoidal Integration*. IEEE Transactions on Power Apparatus and Systems, Vol.PAS-98, No.3, 1080-1090, May/June, 1979.
- [21] Alvarruiz, F.; Henández, V.; Ruiz, P.A. and Vidal, A.M.. *Computing Minimal Realizations of Control Linear Systems. An Approach Based on Distributed Memory Algorithms*. Proceedings of CONTROLO' 98 - 3<sup>rd</sup> Portuguese Conference on Automatic Control, 727-732, Coimbra, Portugal, 9-11 september, 1998.
- [22] Alvin, Kenneth F. and Peterson, Lee D.. *Determination of Modal Residues and Residual Flexibility For Time-Domain System Realization*. Journal of Guidance, Control and Dynamics, Vol. 21, No. 4, pp. 588-594, 1998.
- [23] Amaral, Wagner Caradori do. *Identificação de Sistemas Discretos por Métodos Sequenciais*. Tese de Mestrado, Faculdade de Engenharia de Campinas - UNICAMP, 1976.
- [24] Amaral, Wagner Caradori do. *Identificação Recursiva de Sistemas Multivariáveis*. Tese de Doutorado, Faculdade de Engenharia de Campinas - UNICAMP, 1980.
- [25] Amicucci, G.L. and Monaco S.. *An Observer for A Class of Nonlinear Systems*. Proceedings of 3rd European Control Conference, Rome, Italy, 3423-3428, September, 1995.
- [26] Amorim, C. L., Barbosa, V. C. e Fernandes, E.S.T.. *Uma Introdução à Computação Paralela e Distribuída*. VI Escola de Computação, UNICAMP, IMECC, Campinas, 1988.
- [27] Anan'ev, B.I.. *A Guaranteed Filtering Scheme for Hereditary Differential Systems With No Information On Initial States*. Proceedings of 3rd European Control Conference, Rome, Italy, 966-971, September, 1995.
- [28] Andersen, Anders H.; Gash, Don M. and Avison, Malcolm J.. *Principal component analysis of the dynamic response measured by fMRI: a generalized linear systems framework*. Magnetic Resonance Imaging Volume: 17, Issue: 6, pp. 795-815, July, 1999.
- [29] Anderson, Brian D.O.. *An Algebraic Solution To the Spectral Factorization Problem*. IEEE Trans. Automatic Control, Volume AC-12, pp.410-414, 1967.
- [30] Anderson, Brian D.O.. *A System Theory Criterion For Positive Real Matrices*. Journal Siam Control, Volume 5, No.2, pp.171-182, 1967.
- [31] Anderson, Brian D.O.. *The Inverse Problem of Stationary Covariance Generation*. Journal of Statistical Physics, Volume 1, No.1, pp.133-147, 1969.
- [32] Anderson, Brian D.O.. *An Approach to Multivariable System Identification*. Automatica, Volume 13, No.4, 401-408, July, 1977.
- [33] Anderson, Brian D.O.; Deistler, M. and Scherrer, W.. *Solution Set Properties for Static Errors-in-variables Problems*. Automatica, Volume: 32, Issue: 7, pp. 1031-1035, July, 1996.

- [34] Ängeby, Jakob and Söderström, Torsten. *Determination Of Thermal Parameters In Houses*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1650-1655, July, 1991.
- [35] Antoulas, A.C. and Anderson B.D.O.. *On Optimal Inputs For Identification : A summary*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2886-2887, December, 1994.
- [36] Aoki, Masanao and Staley, M.. *On Input Signal Synthesis in Parameter Identification*. Automatica, Volume 6, No.6, 431-440, May, 1970.
- [37] Aoki, Masanao and Li, Mu T.. *Partial Reconstruction of State Vectors in Decentralized Dynamic Systems*. IEEE Transactions on Automatic Control, Vol.AC-18, No.3, 289-292, June, 1973.
- [38] Aoki, Masanao. *On Subspaces Associated With Partial Reconstruction of State Vectors, the Structure Algorithm, and The Predictable Directions of Riccati Equations*. IEEE Transactions on Automatic Control, Vol.AC-18, No.4, 399-400, August, 1973.
- [39] Aoki, Masanao and Havenner, Arthur. *A method For Approximate Representation of Vector-Valued Time Series : An Alternative to VAR Modeling*. Presented em NFS-NBER seminar on Time Series- University of California - pag. 1-29 - November, 1984.
- [40] Aoki, Masanao. *State Space Modeling of Time Series*. - Springer-Verlag, 1987.
- [41] Aoki, Masanao. *On Alternative State Space Representations Of Time Series Models*. Journal of Economic Dynamics and Control , Vol.- 12, No.2/3, 595-607, June/September, 1988.
- [42] Aoki, Masanao. *Optimization of Stochastic Systems*. Academic Press, Inc., 1989.
- [43] Aoki, Masanao and Havenner, Arthur. *State Space Modeling of multiple Time Series*. Econometric Reviews, Vol.10, n.1, 1-59, 1991.
- [44] Aoki, Masanao. *Modeling Economic Time Series By Forward and Backward State Space Innovation Models and IV Estimators*. European Journal of Operational Research, Volume 73, 265-278, Number 2, March 19, 1994.
- [45] Aoki, Masanao. *Two Complementary Representations of Multiple Time Series in State-space Innovation Forms*. Journal Forecasting, Volume 13, 69-90, Number 13, 1994.
- [46] Aoki, Masanao and Havenner, Arthur M. (Editors). *Applications Of Computer Aided Time Series Modeling*. Lecture Notes in Statistics , No.119, 1996.
- [47] Arena , P.; Caponetto, R.; Fortuna, L. and Muscatto, G.. *Non-linear System Observability As An Optimisation Problem : Solution Via Evolutionary Computing*. Proceedings of 3rd European Control Conference, Rome , Italy, 3412-3416, September, 1995.
- [48] Ariaans, L.J.J.M. ; Weiland, S. and Damen, A.A.H.. *On Optimal Deterministic Identification With Uncertain Data Treated As Exact Data*. Proceedings of 3rd European Control Conference, Rome , Italy, 1-6, September, 1995.
- [49] Arruda, Lúcia Valéria R. de. *Um Supervisor Baseado Em Conhecimento Para Modelagem de Processos*. Tese de Mestrado, Faculdade de Engenharia Elétrica - UNICAMP, 1988.
- [50] Åström, Karl J. and Wittenmark, Björn. *Computer Controlled Systems*. Prentice Hall, 1984.
- [51] Åström, Karl J. and Eykhoff, P.. *System Identification - A Survey*. Automatica, Volume 7, No.2, 123-162, March, 1971.
- [52] Åström, Karl J. and Wittenmark, Björn. *Problems of Identification and Control*. Journal of Mathematical Analysis and Applications, Volume 34, No.2, 90-113, May, 1971.
- [53] Astfalk, Greg. *Parallel Inverse Iteration for Eigenvalue and Singular Value Decompositions*. SIAM NEWS, Vol.27, No.5, 14-16, May/June, 1994.

- [54] Athans, Michael. *The Role and Use of Stochastic Linear-Quadratic-Gaussian Problem In Control System Design*. IEEE Transactions On Automatic Control, Volume AC-16, No.6, 529-552, December, 1971.
- [55] Atherton, D.P.; Hussain, D.M.A. and Gul, E.. *Target Tracking Using Transputers As Parallel Processors*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 510-515, July, 1991.
- [56] Atkinson, P. and Nixon, M.. *Monitoring a Deposition Process Using an Extended Adaptive Kalman Filter*. IEE Proceedings, Vol.131, Pt.D, No. 1, 1-7, January, 1984.
- [57] Awad, Adnan M.. *Properties of the akaike information criterion*. Microelectronics and Reliability Volume: 36, Issue: 4, pp. 457-464, April, 1996.
- [58] Azimi-Sadjadi, Mahmood R., Lu, Tongxin, and Nebot, Eduardo M.. *Parallel and Sequential Block Kalman Filtering and Their Implementations Using Systolic Arrays*. IEEE Transactions on Signal Processing, Vol.39, No.1, 137-147, January, 1991.
- [59] Baase, Sara. *Computer Algorithms*. Addison-Wesley, 1988.
- [60] Backx, Antonius Cornelius Petrus Maria. ( O Dr. Backx assina também como Ton Backx). *Identification Of An Industrial Process : A Markov Parameter Approach*. PhD. Thesis, Dept. El.Eng. Eindhoven University Of Tech., 1987.
- [61] Bagchi, Arunabha and Brummelhuis, Paul Ten. *Parameter Identification in Tidal Models With Uncertain Boundaries*. Automatica, Volume 30, No.5, 745-759, May, 1994.
- [62] Baheti, Radhakisan S., O'Hallaron, David R. and Itzkowitz, Howard R.. *Mapping Extended Kalman Filters Onto Linear Arrays*. IEEE Transactions on Automatic Control, Vol.35, No.12, 1310-1319, December, 1990.
- [63] Bai, Er-Wei; Tempo, Roberto and Cho, Hyonyong. *Membership Set Estimators : Size, Optimal Inputs and Relations With Least Squares*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 255-262, December, 1994.
- [64] Bai, Er-Wei and Nagpal, Krishan M.. *Least Squares Type Algorithms for Identification in the Presence of Modeling Uncertainty*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3602-3607, December, 1994.
- [65] Bai, Er-Wei and Tempo, Roberto. *On the Worst Case Identification an Its Applications*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4148-4149, December, 1994.
- [66] Bai, Zhaojun. *A Parallel Algorithm For Computing The Generalized Singular Value Decomposition*. Journal of Parallel and Distributed Computing, Vol.20, No.3, 280-288, March, 1994.
- [67] Baksalary, Jerzy K. and Baksalary, Oskar Maria *Idempotency of linear combinations of two idempotent matrices*. Linear Algebra and its Applications Volume: 321, Issue: 1-3, pp. 3-7, December 15, 2000.
- [68] Baksalary, J.K. and Kala, R.. *Two Relations Between Oblique and  $\Lambda$ -Orthogonal Projectors*. Linear Algebra and Its Applications, Vol.24, 99-103, April, 1979.
- [69] Baratchart, L. ; Leblond J.; Torkhani, N. and Partington, J.R.. *Robust Identification From Partial Frequency Data*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3900-3905, December, 1994.
- [70] Baratchart, L. ; Leblond J. and Torkhani, N.. *Best Bounded  $H^2$  Extension of Partial Frequency Data*. Proceedings of 3rd European Control Conference, Rome , Italy, 1330-1335, September, 1995.
- [71] Barrett, R.; Berry, M.; Chan, T. F.; Demmel, J.; Donato, J.; Dongarra, J.; Eijkhout, V.; Pozo; R.;Romine; C. and Van der Vorst, C.H.. *Templates for the Solution of Linear Systems: Building Blocks for Iterative Methods*. SIAM, 2nd Edition, 1994.

- [72] Barreto, G.; Bottura, C.P. e Costa Filho, J.T.. *Proposta Para Computação Paralela e Distribuída De Algoritmo de Mínimos Quadrados Para Identificação Recursiva de Sistemas*. XVI - Encontro Nacional de Automática, São Luís - MA, pg. 1-13, 1995.
- [73] Barreto, G.; Bottura, C.P. e Costa Filho, J.T.. *Computação Paralela e Distribuída De Modelo no Espaço de Estado Aplicada Em Identificação Multivariável de Sistemas*. 47a. - Reunião Anual da SBPC, São Luís - MA, pg. 34, 1995.
- [74] Barros, Péricles Barros and Lucena Jr., Vicente F... *The Use of Exponential Weighting in The Excitation Dependent Least Squares Estimator*. Proceedings of the 32th Conference on Decision and Control, San Antonio, Texas, 2056-2057, December, 1993.
- [75] Bartoszewicz, Andrzej. *A comment on 'A time-varying sliding surface for fast and robust tracking control of second-order uncertain systems'*. Automatica, Volume 31, No.12, 1893-1895, December, 1995.
- [76] Barwell, Vitor and George, Alan.. *A Comparison of Algorithms for Solving Symmetric Indefinite Systems of Linear Equations*. ACM Transactions On Mathematical Software, Vol.2, No.3, 242-251, September, 1976.
- [77] Bastogne, T.; Noura, H.; Sibille, P.; and Richard, A.. *Multivariable Identification of A Winding Process By Subspace Methods For Tension Control*. Control Engineering Practice, Vol.6 , 1077-1088, 1998.
- [78] Baur, U. and Isermann, R.. *On-Line Identification Of a Heat Exchanger With a Process Computer - A Case Study*. Automatica, Volume 13, No.5, 487-496, September, 1977.
- [79] Bauer, D.; Deistler, M. and Scherrer, W.. *The analysis of The Asymptotic Variance Of Subspace Algorithms*. 11th IFAC Symposium on System Identification, Vol.3, 1087-1091, July, 1997.
- [80] Bauer, D.; Deistler, M. and Scherrer, W.. *User Choices In Subspace Algorithms*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, WM07-7, December, 1998.
- [81] Bauer, Dietmar. *Some Asymptotic Theory For The Estimation of Linear Systems Using Maximum Likelihood Methods Or Subspace Algorithms*. Ph.D. Thesis, TU Wien, supervisor Prof. Manfred Deistler, 1998.
- [82] Bauer, Dietmar and Deistler, Manfred. *Balanced Canonical Forms for System Identification*. IEEE Transactions on Automatic Control, Vol. 44, No. 6, 1118-1131, June, 1998.
- [83] Bauer, Dietmar; Deistler, Manfred and Scherrer, W.. *Consistency and asymptotic normality of some subspace algorithms for systems without observed inputs*. Automatica, Volume: 35, Issue: 7, pp. 1243-1254, July, 1999.
- [84] Bauer, Dietmar . *On Data Preprocessing for Subspace Methods*. submitted to CDC'2000 conference, Sydney, Australia, 2000.
- [85] Bauer, Dietmar . *Asymptotic efficiency of the CCA subspace method in the case of no exogenous inputs*. submitted to Journal of Time Series Analysis, 2000.
- [86] Bauer, Dietmar and Jansson, M.. *Analysis of the asymptotic properties of the MOESP type of subspace algorithms*. Automatica, Volume: 36, Issue: 4, pp. 497-509, April, 2000.
- [87] Bauer, Dietmar . *Order estimation for subspace methods*. submitted to Automatica, 2000.
- [88] Baz, Didier El. *Parallel Iterative Algorithms For The Solution Of Markov Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2524-2527, December, 1994.
- [89] Becerra, V.M. and Roberts, P.D.. *Discrete Dynamic Integrated System Optimisation and Parameter Estimation and Its Application To Optimal Batch Process Control*. Proceedings of 3rd European Control Conference, Rome , Italy, 1255-1260, September, 1995.

- [90] Beck, V. James and Arnold, Kenneth J.. *Parameter Estimation in Engineering and Science*. John Wiley & Sons, 1977.
- [91] Beck, M.B.. *System Identification and Control*. Dynamic Modeling and Expert Systems in Wastewater Engineering - Gilles G. Patry and David Chapman Lewis Publishers, Inc., 261-323, 1989.
- [92] Beer, R. de. *c59 Quantitative In Vivo NMR* . University of Technology Delft, Departament of Applied Physics - Internet Report - August 3, 1994.
- [93] Behrens, Richard T. and Scharf, Louis L.. *Signal Processing Applications of Oblique Projection Operators*. IEEE Transactions on Signal Processing, Vol.42, No.6, 1413-1424, June, 1994.
- [94] Bellman, Richard and Roth, Robert, S.. *Systems Identification With Partial Information*. Journal of Mathematical Analysis and Applications, Vol.68, No.2, 321-333, April, 1979.
- [95] Ben-Ari, M.. *Principles of Concurrent and Distributed Programming*. Prentice Hall, 1990.
- [96] Ben-Israel, A. and Charnes, A.. *Contributions To The Theory of Generalized Inverses* JSIAM - Journal of the Society for Industrial and Applied Mathematics, Volume 11, No.3, 667-699, September, 1963.
- [97] Benner, Peter; Mehrmann, Volker; Sima, Vasile; Van Huffel, Sabine and Varga, Andras. *SLICOT - A Subroutine Library in Systems and Control Theory*. NICONET Report 97-3, June, 1997.
- [98] Benner, Peter; Laub, Alan J. and Mehrmann, Volker. *A Collection of Benchmark Examples for The Numerical Solution of Algebraic Riccati Equations II : Discrete-Time Case*. Obtido em correspondencia pessoal com o Prof. Laub em 1997 e-mail : laub@ece.ucsb.edu .
- [99] Bergeon, B., Ermine, J.L., Khaddad, A. and Monsion, M.. *Generic Expert System for Identification of Industrial Processes*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 275-280, July, 1991.
- [100] Bernard, Olivier and Gouzé, Jean-Luc. *Robust Validation of Uncertain Models*. Proceedings of 3rd European Control Conference, Rome, Italy, 1261-1266, September, 1995.
- [101] Bertsekas, D. and Tsitsiklis, J.. *Parallel and Distributed Computations*. Prentice Hall, 1989.
- [102] Bertram, J.E. and Sarachik, P.E.. *On Optimal Computer Control*. IFAC, International and Remote Control, Vol.1, pp. 419-422, 1960.
- [103] Betta, A. and Linkens, D.A.. *Intelligent Knowledge-Based System for Dynamic System Identification*. IEE Proceedings, Vol.137, Pt.D, No. 1, 1-12, January, 1990.
- [104] Bezanson, L.W. and Harris, S.L.. *Identification and Control of an Extruder Using Multivariable Algorithms*. IEE Proceedings, Vol.133, Pt.D, No. 4, 145-152, July, 1986.
- [105] Bhattacharyya, S.P.. *On Calculating Maximal (A,B) Invariant Subspaces*. IEEE Transactions on Automatic Control, Vol.AC-20, No.2, 264-265, April, 1975.
- [106] Bierman, Gerald J.. *Factorization Methods for Discrete Sequential Estimation*. Academic Press, Inc., 1977.
- [107] Bingulac, S. and Krtolica, R.V.. *An Algorithm For Simultaneous Order and Parameter Identification in Multivariable Systems*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1020-1025, August, 1988.
- [108] Birkhoff, Garrett e MacLane, Saunders. *Álgebra Moderna Básica*. Guanabara Dois, 1980.
- [109] Bischof, Christian and Loan, Charles Van. *The WY Representation for Products of Householder Matrices*. SIAM J.SCI.STAT.COMPUT. - Volume 8, No. 1, s2-s13, 1987.

- [110] Bittanti, Sergio; Bolzern, Paolo and Campi, Marco. *Recursive Least-Squares Identification Algorithms With Incomplete Excitation: Convergence Analysis and Application to Adaptive Control*. IEEE Transactions on Automatic Control, Vol.35, No.12, 1371-1373, December, 1990.
- [111] Bittanti, Sergio, Bolzern, Paolo; Colaneri, Patrizio and Delrio, Paolo. *Identification of a Helicopter Dynamic Model For Active Control of Vibrations*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 668-673, July, 1991.
- [112] Bittanti, Sergio and Campi, Marco. *Least-Squares Identification of Autoregressive Models With Time-Varying Parameters*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3610-3611, December, 1994.
- [113] Bittanti, Sergio and Picci, Giorgio. *Identification, Adaptation, Learning : the Science of Learning Models from Data*. NATO ASI Series; Springer, 1996.
- [114] Björck, Åke. *Least Squares Methods*. Handbook of Numerical Analysis - editors P.G. Ciarlet and J.L. Lions - Elsevier Science Publishing Company Inc., 1992.
- [115] Björck, Åke. *Numerical Methods for Least Squares Problems*. Siam - Society for Industrial and Applied Mathematics, 1996.
- [116] Björck, Åke and Elfving, Tommy. *Accelerated Projection Methods For Computing Pseudoinverse Solutions Of Systems of Linear Equations*. BIT, Vol.19, 145-163, 1979.
- [117] Björck, Åke and Golub, Gene H.. *Numerical Methods for Computing Angles Between Linear Subspaces*. Mathematics of Computation, Volume 27, Number 123, 579-594, July, 1973.
- [118] Blackford, L. S.; J. Choi, A. Cleary, E. D'Ázevedo, J. Demmel, I. Dhillon, J. Dongarra, S. Hammarling, G. Henry, A. Petitet, K. Stanley, D. Walker, and R. C. Whaley, 1997. *ScaLAPACK Users' Guide*, SIAM Publications, Philadelphia, (1997).
- [119] Blackford, L. S.; Cleary, A.; Petitet, Antoine; Whaley, R. C. ; Demmel, James; Dhillon, I.; Ren, H.; Stanley, K.; Dongarra, Jack and Hammarling, Sven. *Practical Experience in the Numerical Dangers of Heterogeneous Computing*. ACM Transactions on Mathematical Software (TOMS) - 23(2): 133-147 ,1997
- [120] Blackford, L. S.; Choi, J.; Cleary, S. A.; Demmel, J.; Dhillon, I.; Dongarra, J.; Hammarling, G.; Henry, G.; Petitet, A.; Stanley, K.; Walker, D. and Whaley, R. C.. *SCaLAPACK: A Portable Linear Algebra Library for Distributed Memory Computers - Design Issues and Performance*. Proceedings of the Supercomputer 96 meeting, November 1996, IEEE Computer Society Press, 1996.
- [121] Blessing, Peter. *Parameter Estimation Of State-Space Models For Multivariable Systems With Correlation Analysis and Method Of Least Squares*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 385-394, July, 1977.
- [122] Bobrow, James E. and Murray, Walter. *An Algorithm for RLS Identification of Parameters that Vary Quickly with Time*. IEEE Transactions on Control, Vol.38, NO. 2, 351-354, February, 1993.
- [123] Bodin, Per and Wahlberg. *Thresholding in High Order Transfer Function Estimation*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3400-3405, December, 1994.
- [124] Bodin, Per. *On the Selection of Best Orthonormal Basis in System Identification and Signal Processing*. PhD. Thesis School of Electrical Engineering, Royal Institute of Technology, Stockholm, Sweden, 1997.
- [125] Bokor, József. *Approximate identification for robust control*. Annual Reviews in Control Volume: 22, pp. 187-198 , 1998.

- [126] Boom, Ton van den; Klompstra, Martin and Damen, Ad. *System Identification for  $H_\infty$ - Robust Control Design*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1431-1436, July, 1991.
- [127] Borrie, John A. *Stochastic Systems for Engineers : modelling, estimation and control*. Prentice Hall, 1992.
- [128] Bosgra, Okko H. and Van Der Weiden, Antonius J.J.. *Input-Output Invariants for Linear Multi-variable Systems*. IEEE Transactions on Automatic Control, Vol. AC-25, No.1, 20-36, February, 1980.
- [129] Bottura, Celso Pascoli.  
*Princípios de Controle e Servomecanismo*.  
Editora Guanabara Dois, 1982.
- [130] Bottura, Celso Pascoli.  
*Análise Linear de Sistemas* .  
Editora Guanabara Dois, 1982.
- [131] Bottura, C.P.; Barreto, G.; Bordon, M.J. e Costa Filho, J.T.. *Proposta de Procedimento para Avaliação de Desempenho Numérico de Algoritmos Matriciais*. XVI CILAMCE - Congresso Íbero Latino- Americano de Métodos Computacionais de Engenharia, Curitiba, 1995.
- [132] Bottura, C.P.; Barreto, G.; Bordon, M.J. e Costa Filho, J.T.. *Computação de Alto Desempenho dos Valores Singulares de Uma Matriz de Hankel Aplicada ao Modelamento de Séries Temporais No Espaço De Estados*. 49a. Reunião Anual da Sociedade Brasileira Para o Progresso da Ciência, Belo Horizonte, pg.48, 1997.
- [133] Bottura, C.P.; Barreto, G.; Bordon, M.J. e Costa Filho, J.T.. *Uma Abordagem Computacional Para Cálculo Da Decomposição Em Valores Singulares De Matriz De Hankel*. XVIII CILAMCE - Congresso Íbero Latino- Americano de Métodos Computacionais de Engenharia, Brasília, 1997.
- [134] Bottura, C.P.; Barreto, G.; Bordon, M.J. e Costa Filho, J.T.. *Séries Temporais No Espaço de Estado: Modelamento e Computação de Alto Desempenho*. XIV - Congresso Brasileiro de Engenharia Mecânica, Bauru, 1997.
- [135] Bottura, C.P.; Barreto, G.; Bordon, M.J. e Costa Filho, J.T.. *State Space Modelling and High Performance Computing of Multivariate Time Series*. Eighth SIAM Conference on Parallel Processing for Scientific Computing, Minneapolis, Minnesota, USA, March, 1997.
- [136] Bottura, C.P.; Barreto, G. and Bordon, M.J.. *Multivariate Time Series State Space Modelling With High Performance SVD Computation of an Hankel Matrix* . Sixth SIAM Conference on Applied Linear Algebra, Snowbird - Utah, USA, March, 1997.
- [137] Bottura, Celso Pascoli; Barreto, Gilmar; Bordon, José Maurício and Tamariz, Annabell Del Real. *A High Computational Performance Approach for a Subspace Identification Method*. Ninth SIAM Conference on Parallel Processing for Scientific Computing, San Antonio, Texas, March 22-24, 1999.
- [138] Bottura, Celso P. ; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R.. *Tratamento Computacional de Alto Desempenho em Método de Subespaços para Modelagem de Dados*. XV Congresso Brasileiro de Engenharia Mecânica - Cobem 99, Águas de Lindóia, São Paulo, Brasil.
- [139] Bottura, Celso P. ; Tamariz, Annabell D.R.; Barreto, Gilmar and Fonseca Neto, Joao Viana da. *A modified-Schur method for solving algebraic Riccati equations*. 38th IEEE Conference on Decision and Control, Phoenix, Arizona, USA, 1999.
- [140] Bottura, Celso P.; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R.. *Parallel and distributed state-space modeling for computation of time series using realization theory*. 3rd MATHMOD - Third International Symposium on Mathematical Modelling, Vienna, Austria, 2000.



- [141] Bottura, Celso P. ; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R.. *Parallel and distributed computational data modelling via Verhaegen & Dewilde's subspace method*. 2000 American Control Conference, Chicago, USA, 2000.
- [142] Bottura, Celso P. ; Barreto, Gilmar and Cáceres, Angel F. Torrico. *River Water Quality Modeling Via Verhaegen and Dewilde's Subspace Method*. IFAC - Agricontrol 2000 Conference, Wageningen, The Netherlands, 2000.
- [143] Bottura, Celso P. ; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R.. *An Approach to State Space Computational Modeling and Prediction of Time Series in Parallel and Distributed Computers*. 16th Brazilian Congress Of Mechanical Engineering, november, 26-30, Uberlândia - Minas Gerais, Brasil, 2001.
- [144] Bottura, Celso P. ; Barreto, Gilmar; Tamariz, Annabell D.R. and João Viana da Fonseca Neto. *Parallel and Distributed Multivariable Identification Via the MOESP Approach*. 16th Brazilian Congress Of Mechanical Engineering, november, 26-30, Uberlândia - Minas Gerais, Brasil, 2001.
- [145] Bottura, Celso P. ; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R. *Parallel and Distributed Computational Multivariate Time Series Modeling In State Space*. 2002 American Control Conference, Anchorage, Alaska, USA, May 8-10, 2002.
- [146] Bottura, Celso P. ; Barreto, Gilmar; Bordon, Mauricio José and Tamariz, Annabell D.R. *Parallel and Distribution Computational Multivariate Time Series Modeling In State Space*. Submitted.
- [147] Bottura, Celso P. ; Tamariz, Annabell D.R.; Barreto, Gilmar and Cáceres, Angel F. Torrico. *Parallel and Distributed MOESP Computational System's Modelling*. 10th Mediterranean Conference On Control and Automation - MED 2002, Lisboa, Portugal, July 9-12, 2002.
- [148] Bottura, Celso Pascoli. *Notas para um Livro de Controle Multivariável*. em fase de edição.
- [149] Bounit, H. and Hammouri, H.. *Observers For Distributed Bilinear Control Systems*. Proceedings of 3rd European Control Conference, Rome , Italy, 3417-3422, September, 1995.
- [150] Box, George E. P. and Jenkins, Gwilym M.. *Time Series Analysis : Forecasting and Control*. Holden-Day Series in Time Series Analysis and Digital Processing, Holden-Day, 1976.
- [151] Bowerman, Bruce L.; O'Connell, Richard T.. *Time Series Forecasting*. Duxbury Press, 1979.
- [152] Boyd, Stephen; Tang, Y.S. and Chua, Leon O.. *Measuring Volterra Kernels*. IEEE Transactions on Circuits and Systems, Volume CAS-30, Number 8, pp.571-577, 1983.
- [153] Breiman, Leo and Friedman, Jerome H.. *Estimating Optimal Transformations for Multiple Regression and Correlation*. JASA - Journal of the American Statistical Association, Volume 50, Number 391, 580-619, September, 1985.
- [154] Brent, Richard P.. *The Parallel Evaluation of General Arithmetic Expressions*. Journal of the Association for Computing Machinery, Vol.21, No.2, 201-206, April, 1974.
- [155] Brent, Richard P. and Luk, Franklin T.. *The Solution of Singular-Value and Symmetric Eigenvalue Problems on Multiprocessor Arrays*. SIAM Journal on Scientific and Statistical Computing, Vol.6, No.1, 69-84, January, 1985.
- [156] Bretthauer, G. *Identifiability of Subsystems in Interconnected Systems*. Proceedings of the IFAC/IFORS Symposium - Large Scale Systems: Theory and Applications 1983 - Warsaw, Poland, 417-422 , July, 1983.
- [157] Brockett, Roger W.. *Gramians, Generalized Inverses, and the Least-Squares Approximation of Optical Flow*. Journal of Visual Communication and Image Representation, Vol.1, No. 1, pp. 3-11, September, 1990.
- [158] Brown, D.W. and Gaston, F.M.F. *Hierachical Signal Flow Graph Representation Of The Square-Root Covariance Kalman Filter*. Algorithms and Parallel VLSI Architectures III - M. Moonen and F. Catthoor (Editors), pp. 37 - 48, Elsivier Science, 1995.

- [159] Bruder, S., Johnson, R. and Farooq, M.. *A Multi-Sensor Tracking and Identification Algorithm Using The Square-Root Information Filter*. Proceedings of the 32th Conference on Decision and Control, San Antonio, Texas, 940-941, December, 1993.
- [160] Bruyne, F. De and Gevers, M.. *Identification For Control : Can The Optimal Restricted Complexity Model Always Be Identified*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3912-3917, December, 1994.
- [161] Bubnicki, Z. *Optimization Problems in Large-Scale Systems Modelling and Identification*. Proceedings of the IFAC/IFORS Symposium - Large Scale Systems: Theory and Applications 1983 - Warsaw, Poland, 411-416, July, 1983.
- [162] Budde, C.Donald La. *Two New Classes of Algorithms for Finding the Eigenvalues and Eigenvectors of Real Symmetric Matrices*. Journal of the Association for Computing Machinery, Volume 11, Number 1, 53-58, January, 1964.
- [163] Budin, Michael A.. *Minimal Realization of Discrete Linear Systems from Input-Output Observations*. IEEE Transactions on Automatic Control, Vol.AC-16, No.5, 395-401, October, 1971.
- [164] Bull, J.M.; Riley, G.D.; Rasbash, J. and Goldstein, H., *Parallel implementation of a multilevel modelling package*. Computational Statistics & Data Analysis, Volume: 31, Issue: 4, pp. 457-474, October 28, 1999.
- [165] Bultheel, A. and Moor, B. De *Rational approximation in linear systems and control*. Journal of Computational and Applied Mathematics Volume: 121, Issue: 1-2, September 1, pp. 355-378, 2000.
- [166] Bunse-Gerstner, Angelika and Faßbender, Heike. *A Jacobi-Like Method for Solving Algebraic Riccati Equations on Parallel Computers*. IEEE Transactions on Automatic Control, Vol.AC-42, No.8, 1071-1084, August, 1997.
- [167] Businger, P. and Golub, G.H.. *Linear Least Squares Solutions by Householder Transformations*. Numerische Mathematik 7, 269-276, 1965.
- [168] Byrnes, Christopher I.; Lindquist, Anders; Gusev, Sergei V. and Matveev, Alexei S.. *The Geometry of Positive Real Functions With Applications To The Rational Covariance Extension Problem*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3883-3888, December, 1994.
- [169] Cáceres, Angel Fernando Torrico. *Identificação de Sistemas Multivariáveis no Espaço de Estado e Controle de Sistemas da Qualidade da Água de Rios*. Tese de Doutorado a ser apresentada na Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas em 2002.
- [170] Cadzow, James A. and Solomon Jr., Otis M.. *Algebraic Approach to System Identification*. IEEE - ASSP - Acoustics, Speech and Signal Processing, Volume ASSP-34, Number 3, 462-469, June 1986.
- [171] Caines, P.E.. *Linear Stochastic Systems*. Wiley, 1988.
- [172] Callafon, R.A. de and Van den Hof, P.M.J.. *Filtering and Parametization Issues in Feedback Relevant Identification Based on Fractional Model Representations*. Proceedings of 3rd European Control Conference, Rome , Italy, 441-446, September, 1995.
- [173] Calzarossa, Maria and Serazzi, Giuseppe. *Workload characterization: a survey*. Proceedings of the IEEE, Vol. 81, Number 8, 1136-1150, August, 1993.
- [174] Candy, J. V.; Bullock, T.E. and Warren, M.E.. *Invariant System Description of the Stochastic Realization*. Automatica, Vol.15, Number 4, pp.493-495 , July, 1979.
- [175] Cao, Liangyue; Mees, Alistair and Judd, Kevin . *Dynamics from multivariate time series*. Physica D Volume: 121, Issue: 1-2, pp. 75-88, October 1, 1998.

- [176] Carey, G. F.. *Parallel Supercomputing: Methods, Algorithms and Applications*. Wiley Series in Parallel Computing, 1989.
- [177] Carrete, P. ; Genin, Y.; Bastin, G. and Gevers, M.. *Discarding Data May Improve The Parameter Estimation Accuracy In System Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 827-832, September, 1995.
- [178] Castaldi, P.; Guidorzi, R.P. ; Soverini, U. and Beghelli, S.. *On The Application Of The Frisch Scheme Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 851-855, September, 1995.
- [179] Castro, Maria Cristina Felippetto de. *Predição Não-Linear de Séries Temporais Usando redes Neurais RBF por Decomposição em Componentes Principais*. Tese de Doutorado, Faculdade de Engenharia Elétrica e Computação - UNICAMP, 2001.
- [180] Castrucci, Plínio B.L. e Latre, Luiz G. *Identificação e Controle Estocástico*. 4º Congresso Brasileiro de Automática - Cursos, 34-72, 1982.
- [181] Cavanaugh, Joseph E. and Shumway, Robert H.. *An Akaike information criterion for model selection in the presence of incomplete data*. Journal of Statistical Planning and Inference Volume: 67, Issue: 1, pp. 45-65 , March 16, 1998.
- [182] Cavanaugh, Joseph E.. *Unifying the derivations for the Akaike and corrected Akaike information criteria*. Statistics & Probability Letters Volume: 33, Issue: 2, pp. 201-208 , April 30, 1997.
- [183] Chai, Tian You.. *A Multivariable Robust Self-Tuning Controller*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 310-315, July, 1991.
- [184] Chakraborty, Kanad; Mehrotra, Kishan; Mohan, Chilukuri K. and Ranka, Sanjay. *Forecasting the Behavior of Multivariate Time Series Using Neural Networks*. Neural Networks, Vol.5, 961-970, 1992.
- [185] Chan, David. *Implementation and Evaluation of a Parallel Singular Value Decomposition Algorithm for Direction of Arrival Analysis*. Master's thesis, Worcester Polytechnic Institute, MA, 1995.
- [186] Chan, H. and Özgüner, Ü.. *Control of Interconnected Systems Over A Communications Network With Queues*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4104-4109, December, 1994.
- [187] Chan, Tony F.. *An Improved Algorithm for Computing the Singular Value Decomposition*. ACM Transactions on Mathematical Software, Volume 8, No.1, 72-83, March, 1982.
- [188] Chang, Xiao-Wen and Paige, Christopher.. *A Perturbation Analysis for  $\mathbf{R}$  in the  $\mathbf{QR}$  Factorization*. obtido via chirs@cs.mcgill.ca, Technical Report No. SOCS-95.7, 1-20, November, 1995.
- [189] Chen, Chi-Tsong. *Linear Systems Theory and Design*. Holt, Rinehart and Winston, New york, 1984.
- [190] Chen, Chi-Tsong. *Analog and Digital Control Systems Design: Transfer-Function, State-Space and Algebraic Methods*. Sanders College Publishing, 1995.
- [191] Chen, Hai-Wen. *Modeling and Identification of Parallel and Feedback Nonlinear Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2267-2272, December, 1994.
- [192] Chen, Jie; Farrel, Jay; Nett, Carl N. and Zhou, Kemin. *Realization of Positive Linear Systems Using Polyhedral Cones*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4158-4163, December, 1994.

- [193] Chen, Shyh-Ching and Kuck, David J.. *Time and Parallel Processor Bounds For Linear Recurrence Systems*. IEEE Transactions on Computers, Vol.C-24, No.7, 701-717, July, 1975.
- [194] Chernousko, F.L.; Ovssevich, A.I. and Shmatkov, A.M.. *Optimal Two - Sided Ellipsoidal State Estimation In Dynamical Systems*. Proceedings of 3rd European Control Conference, Rome , Italy, 961-965, September, 1995.
- [195] Chin, Toshio M.; Karl, William C. and Willsky, Alan S.. *A Distributed and Iterative Method for Square Root Filtering in Space-Time Estimation*. Automatica, Volume 31, Number 1, 67-82, January, 1995.
- [196] Chisci, Luigi and Zappa, Giovanni. *High Throughput Parallel Implementation Of RLS*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 812-817, August, 1988.
- [197] Chisci, Luigi, Garulli, A. and Zappa, Giovanni. *Properties Of Optimal Bounding Parallelotope Algorithms*. Proceedings of 3rd European Control Conference, Rome , Italy, 7-12, September, 1995.
- [198] Cho, Young Man. *Fast Subspace Based System Identification: Theory and Practice*. Stanford University - A dissertation submitted for the degree of Doctor of Philosophy, August, 1993.
- [199] Cho, Young Man; Xu, Guanghan and Kailath, Thomas. *Fast Identification of State-Space Models Via Exploration of Displacement Structure*. IEEE Transactions on Automatic Control, Vol.39, No.10, 2004-2017, October, 1994.
- [200] Cho, Young Man; Xu, Guanghan and Kailath, Thomas. *Fast Recursive Identification of State-Space Models Via Exploration of Displacement Structure*. Automatica, Vol.30, No.1, 45-59, January, 1994.
- [201] Choi, J.; Dongarra, J.; Pozo, R. and Walker, D.W.. *Parallel Matrix Transpose Algorithms on Distributed Memory Concurrent Computers*. Technical Report UT-CS-93-215 - LAPACK Working Note 65, University of Tennessee, 1993.
- [202] Choi, J.; Demmel, J.; Dhillon, I.; Dongarra, J.; Ostrouchov, S.; Petitet, A.; Stanley, K.; Walker, D. and Whaley, R. C.. *LAPACK Working Note 95 : ScaLAPACK: A Portable Linear Algebra Library for Distributed Memory Computers - Design Issues and Performance*. LAPACK Working Note 95, March 1995.
- [203] Chongrid, Krit and Hara, Shinji. *Hankel Norm of Sampled-Data Systems*. IEEE Transactions on Automatic Control, Vol.40, No.11, 1939-1942, November, 1995.
- [204] Chou, Chun Tung. *Geometry of Linear Systems and Identification*. Cambridge University - A dissertation submitted for the degree of Doctor of Philosophy, March, 1994.
- [205] Chou, C.T. and Hanzon, Bernard. *A Continuation method approach to the  $L_2$  model reduction problem*. Proceedings of 3rd European Control Conference, Rome , Italy, 1336-1341, September, 1995.
- [206] Chou, C.T. and Verhaegen, Michel. *Subspace Algorithms for the Identification of Multivariable Dynamic Errors-in-Variables Models*. Automatica, Volume 33, No.10 , 1857-1869, October, 1997.
- [207] Christov, N.D. ; Petkov, P. Hr. and Konstantinov, M.M.. *An Efficient QR-Decomposition Algorithm For Block-Hankel Matrices*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1105-1107, August, 1988.
- [208] Chu, L.. *A Variable-Structure Model For Policy Formulation of Mineral Resources Systems*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1551-1554, August, 1988.
- [209] Chui, N. L. C. and Maciejowski, J.M.. *Realization of Stable Models with Subspace Methods*. Automatica, Volume 32, No.11, 1587-1595, November, 1996.

- [210] Chun, J., Kailath, T. and Lev-Ari, H.. *Fast Parallel Algorithms For QR and Triangular Factorizations*. SIAM - Journal on Scientific and Statistical Computing, Volume 8, No.6, 899-913, November, 1987.
- [211] Cimmino, Gianfranco. *Calcolo Approssimato per le Soluzioni dei Sistemi Di Equazioni Lineari*. La Ricerca Scientifica, Vol.2 , n.9, pp.325-338, 1938.
- [212] Cipra, Barry . *Engineers Look to Kalman Filtering for Guidance*. SIAM News, Vol. 26, No. 5, August, 1993.
- [213] Coleman, T.F. and Loan, C.V.. *Handbook for Matrix Computations*. SIAM - Society for Industrial and Applied Mathematics, 1988.
- [214] Connell, Edwin H.. *Elements of Abstract and Linear Algebra*. Department of Mathematics, University of Miami, <http://www.math.miami.edu/~ec/book/>, 1999.
- [215] Cooper D.M. and Thompson, R.. *A Note on Estimation of Parameters of The Autoregressive-Moving Average Process*. Biometrika, Volume 64, No.3 , pag. 625-628 - December, 1977.
- [216] Corrêa, G.O. and Glover, K.. *Pseudo-canonical Forms, Identifiable Parametrizations and Simple Parameter Estimation for Linear Multivariable Systems : Input-Output Models*. Automatica, Volume 20, No.4, 429-442, July, 1984.
- [217] Corrêa, G.O. and Glover, K.. *Pseudo-canonical Forms, Identifiable Parametrizations and Simple Parameter Estimation for Linear Multivariable Systems : Parameter Estimation*. Automatica, Volume 20, No.4, 443-452, July, 1984.
- [218] Cosnard, Michel. *Distributed Algorithms: Their Nature & The Problems Encountered*. edited by Cosnard, Michel; Robert, Yves; Quinton, Patrice and Raynal, Michel. Parallel & Distributed Algorithms. Elsevier Science Publishers B.V. - 1989.
- [219] Costa Filho, J. Tarcísio. *Processamento Paralelo de Algoritmos de Controle Hierárquico*. Tese de Mestrado, Faculdade de Engenharia Elétrica - UNICAMP, 1988.
- [220] Costa Filho, J. Tarcísio. *Proposta para Computação Assíncrona Paralela e Distribuída de Estruturas Especiais de Jogos Dinâmicos*. Tese de Doutorado, Faculdade de Engenharia Elétrica - UNICAMP, 1992.
- [221] Costa Filho, José Tarcísio; Bottura, Celso Pascoli. *Parallel and Distributed Dynamic Games Computation On A Network of Multiple Workstations*. IFAC Workshop on Distributed Computer Control Systems Beijing, China, 23-25 /8 /1992.
- [222] Costa Filho, José Tarcísio; Bottura, Celso Pascoli e Barreto, Gilmar. *Um Método de Identificação Recursiva Paralela*. V Simpósio Brasileiro de Arquitetura de Computadores e Processamento de Alto Desempenho, 426-431 - Florianópolis, 1993.
- [223] Costa Filho, José Tarcísio; Bottura, Celso Pascoli; Barreto, Gilmar e Souza, Wagner A.M.. *Um Algoritmo Paralelo de Mínimos Quadrados Recursivo Baseado em Transformações Ortogonais*. XIX Congresso Nacional de Matemática Aplicada e Computacional, 365-366 - Goiânia, 1996.
- [224] Crabbe, Leland and Young, Brian B.. *Modeling Multiple Time Series Via The Internally Balanced State Space Representation*. pag. 1-31 - July 1986 - correspondência pessoal do Prof. Dr. Celso Pascoli Bottura.
- [225] Crosta, Giovanni F.; Santoni, Francesca. *Continuous Flows Which Identify Distributed Parameters*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2265-2266, December, 1994.
- [226] Crow, M.L. and Ilic', M. . *The Parallel Implementation of The Waveform Relaxation Method for Transient Stability Simulations*. IEEE Transactions on Power Systems, Vol.PWRS-5, No.3, 922-932, August, 1990.

- [227] Cryer, Jonathan D. and Ledolter, Johannes. *Small-sample Properties of The Maximum Likelihood estimator in the First-order Moving Average Model*. Biometrika, Volume 68, No.3 , pag. 691-694 - December, 1981.
- [228] Csanky, L.. *Fast Parallel Matrix Inversion Algorithms*. SIAM J.Computing , Vol.5, No.4 , 618-623, December, 1976.
- [229] Cunha, Rudnei Dias da e Hopkins, Tim. *The Parallel Solution of Triangular Systems of Linear Equations*. Symposium in High Performance Computing, Montpellier, France , pag. 1-11, 7-9 October, 1991.
- [230] Cunha, Rudnei Dias da e Hopkins, Tim. *The Parallel Iterative Methods (PIM) Package for the Solution of Systems of Linear Equations on Parallel Computers*. SuperComp 94 - Seminário de Computação Aplicada, Universidade Federal do Rio Grande do Sul - Porto Alegre - 287-290 - Setembro de 1994.
- [231] Dahlén, Anders. *What is Wrong With Subspace Identification Of Time Series And What Can Be Done Instead ?* Thesis Licentiate - Division of Optimization and Systems Theory, Royal Institute of Technology - Stockholm - Sweden - 1998.
- [232] Dahlén, Anders, Lindquist, Anders and Mari, Jorge. *Experimental Evidence Showing That Stochastic Subspace Identification Methods May Fail*. Systems & Control Letters, V.34, pag. 303-312, 1998.
- [233] Dasgupta, Soura and Anderson, Brian D.O.. *Physically Based Parametrizations For Designing Adaptive Algorithms*. Automatica, Volume 23, No.4, 469-477, July 1987.
- [234] Dastych, J. and Unbehauen, H.. *Some Aspects on The Identification of Dead-Times*. Proceedings of 3rd European Control Conference, Rome , Italy, 3697-3701, September, 1995.
- [235] Datta, Biswa Nath. *Parallel and Large-Scale Matrix Computations in Control : Some Ideas*. Linear Algebra and Applications, Vol.121, 243-264, 1989.
- [236] Date, Paresh. *Identification for Control: Deterministic Algorithms and Error Bounds*. Ph.D. Thesis Control Group - Department of Engineering - University of Cambridge - June, 2000.
- [237] Davis, R.A. and Glover, K.. *An application of recent model validation techniques to flight test data*. Proceedings of 3rd European Control Conference, Rome , Italy, 1249-1254, September, 1995.
- [238] DeCarlo, Raymond A.. *Linear Systems - A State Variable Approach With Numerical Implementation*. Prentice Hall, 1990.
- [239] De Cock, Katrien and De Moor, Bart. *Identification of the first order parameters of a circulant modulated Poisson process*. Proc. of the International Conference on Telecommunications (ITC'98), Porto Carras, pp. 420-424 (Vol. II); Greece, Jun. 1998.
- [240] De Cock, Katrien and De Moor, Bart. *Subspace angles and distances between ARMA models*. Internal Report 00-39, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), and in Proceedings of the Fourteenth International Symposium of Mathematical Theory of Networks and Systems (MTNS 2000), Perpignan, France, June 2000.
- [241] De Cock, Katrien and De Moor, Bart. *Subspace angles between AR models*. Internal Report 00-44, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2000.
- [242] De Cock, Katrien and De Moor, Bart. *Subspace angles between linear stochastic models*. Internal Report 99-101a, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2000.
- [243] De Cuyper, Joris *Subspace identification os Multi Axis Durability Test Rigs using SLICOT*. Second NICONET WORKSHOP: Numerical Control Software: SLICOT, a useful tool in industry; INRIA Rocquencourt, France, December, 1999.

- [244] Dehaene, Jeroen. *Continuous-Time Matrix Algorithms, Systolic Algorithms and Adaptive Neural Networks*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, Oktober, 1995.
- [245] De Hoog, Thomas J.; Heuberger, Peter S.C. and Van den Hof, Paul M.J.. *Partial realization in generalized bases: algorithm and example*. Proceedings of the 12th IFAC Symposium on System Identification - SYSID, 2000.
- [246] Deistler, Manfred and Gevers, Michel. *Properties of The Parametrization Of Monic ARMA Systems*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1341-1348, August, 1988.
- [247] Deistler, M. ; Peternell, K. and Scherrer, W.. *Consistency and relative efficiency of subspace methods*. Automatica, Volume 31, No.12, 1865-1875, December, 1995.
- [248] Deistler, Manfred. *Identification of Linear Systems: A General Framework for the Theory*. Proceedings of the 38th Conference on Decision & Control, Phoenix, Arizona, USA, 2822-2826, December, 1999.
- [249] De Jong, Lieuwe Sytse. *Numerical Aspects of Realization Algorithms in Linear Systems Theory*. PhD. Dissertation, Eindhoven University Technology, Eindhoven, The Netherlands, 1975.
- [250] Delgado, Catarina J.M.; Santos, P.Lopes dos and Carvalho, J.L. Martins. *A Recursive Subspace Identification Algorithm*. Proceedings of CONTROLO' 98 - 3<sup>rd</sup> Portuguese Conference on Automatic Control, 465-470, Coimbra, Portugal, 9-11 september, 1998.
- [251] Delgado, Catarina J.M.; Santos, P.Lopes dos and Carvalho, J.L. Martins. *On-Line Subspace Identification Algorithm*. VIII Congresso Latinoamericano de Control Automático - Viña del Mar, Chile, 673-678, 9 - 13 de Noviembre, 1998.
- [252] Delgado, Catarina J.M.; Santos, P.Lopes dos and Carvalho, J.L. Martins. *Numerical Algorithm for Recursive Subspace Identification*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, TA08-5, December, 1998.
- [253] Delmaire, G.; Cassar, J.P. and Staroswiecki, M.. *Identification and Parity Space Techniques For Failure Detection in SISO Systems Including Modelling Errors*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2279-2285, December, 1994.
- [254] De Lathauwer, Lieven and De Moor, Bart. *From Matrix to Tensor : Multilinear Algebra and Signal Processing*. Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, ESAT-SISTA/TR 1996-93, December, 1996.
- [255] De Lathauwer, Lieven. *Signal Processing Based On Multilinear Algebra*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, September, 1997.
- [256] Demmel, James W.. *LAPACK: A Portable Linear Algebra Library for Supercomputers* IEEE Control Systems Society, Workshop on Computer-Aided Control System Design (CACSD), 1-6, December, 1989.
- [257] Demmel, James W.; Heath, Michael T. and Vorst, Henk A. van der.. *Parallel Numerical Linear Algebra* Acta Numerica, Cambridge University Press, October, 1992.
- [258] Demmel, James W.; Heath, Michael T. and Vorst, Henk A. van der.. *Numerical Linear Algebra* E-mail : demmel@cs.berkeley.edu , October, 1994.
- [259] De Moor, Bart. *Mathematical Concepts and Techniques for Modelling of Static and Dynamic Systems*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, June, 1988.
- [260] De Moor, Bart; Vandewalle, Joos; Moonen, Marc; Vanderberghe, Lieven and Miegheem, Piet Van. *A Geometrical Strategy for the Identification of State Space Models Of Linear Multivariable System With Singular Value Decomposition*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 700-704, August, 1988.

- [261] De Moor, Bart and Vandewalle, Joos. *The Uncertainty Principle of Mathematical Modelling*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 2017- 2022, August, 1988.
- [262] De Moor, Bart L.R. and Golub, Gene H.. *Generalized Singular Value Decompositions: A Proposal for a Standardized Nomenclature*. Numerical Analysis Project - Manuscript NA-89-04, Stanford University - 1-12, April, 1989.
- [263] De Moor, Bart L.R.. *The Restricted Singular Value Decompositions: Properties and Applications*. Numerical Analysis Project - Manuscript NA-89-03, Stanford University - 1-69, April, 1989.
- [264] De Moor, Bart L.R.. *On The Structure and Geometry of The Product Singular Value Decomposition..* Numerical Analysis Project - Manuscript NA-89-06, Stanford University - 1-52, May, 1989.
- [265] De Moor, Bart L.R. and Boyd, Stephen. *Analytic Properties Of Singular Values and Vectors*. ESAT-SISTA report 1989-28, 1-16, 1989.
- [266] De Moor, B. and Zha, Hongyuan. *A Tree Of Generalizations of The Ordinary Singular Value Decomposition*. Linear Algebra and Its Applications, Volume 147, 469-500, March, 1991.
- [267] De Moor, B. and Golub, Gene. *The restricted singular value decomposition: properties and applications*. SIAM - Journal on Matrix Analysis and Applications, Volume 12, No.3, 401-425, July, 1991.
- [268] De Moor, Bart; Van Overschee, Peter and Suykens, Johan. *Subspace Algorithms for System Identification and Stochastic Realization*. Proc. of the International Symposium on Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing (MTNS'91), Kobe, Japan, pp. 589-595; Jun. 1991.
- [269] De Moor, Bart L.R.. *The Singular Value Decomposition and Long and Short Spaces of Noisy Matrices*. ESAT-SISTA report 139, 1-39, 1992.
- [270] De Moor, B. and Paul Van Dooren. *Generalizations of the singular value and QR decomposition*. SIAM - Journal on Matrix Analysis and Applications, Volume 13, No.4, 993-1014, October, 1992.
- [271] De Moor, Bart. *Numerical Algorithms for State Space Subspace System Identification*. Published in Academia Analecta, Klasse der Wetenschappen, Koninklijke Akademie voor Wetenschappen, Jaargang, 55, nr.5, 1993
- [272] De Moor, Bart. *Structured Total Least Squares and  $L_2$  approximation problems*. Special Issue on Numerical Linear Algebra Methods in Control, Signal and Systems, Linear Algebra and Its Applications, 1993.
- [273] De Moor, Bart. *On the structure of generalized singular value and QR decompositions*. SIAM - Journal on Matrix Analysis and Applications, Volume 15, No.1, 347-358, January, 1994.
- [274] De Moor, Bart; Gevers, Michel and Goodwin, Graham C..  *$L_2$ -overbiased,  $L_2$ -underbiased and  $L_2$ -unbiased Estimation of Transfer Functions*. Automatica, Volume 30, No.5, 893-898, May, 1994.
- [275] De Moor, Bart and Roorda, Berent.  *$L_2$ -Optimal Linear System Identification Structured Total Least Squares For SISO Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2874-2879, December, 1994.
- [276] De Moor, Bart and Overschee, Peter Van. *Graphical User Interface Software for System Identification* ESAT/SISTA TR 1994-06 - Department of Electrical Engineering - Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, June, 1994.
- [277] De Moor, Bart and Overschee, Peter Van. *Numerical Algorithms for State Space Subspace System Identification*. Editado por Alberto Isidori em Trends in Control - A European Perspective, 385-422, Springer-Verlag, 1995.



- [278] De Moor, Bart ; Gersem, Peter De; Schutter, Bart De and Favoreel, Wouter. *DAISY : A Database for Identification of Systems* . Journal A, Vol.38, no.3, pp 4-5, September, 1997.
- [279] De Moor B.; Van Overschee P. and Favoreel W.. *Algorithms for Subspace State Space System Identification - An Overview*. Birkhauser Book Series on Applied and Computational Control, Signals and Circuits (Biswa Datta Editor), Chapter 6 - pp. 271 - 336, 1999.
- [280] De Schepper S.; Van Overschee P. and De Moor B., *Transfer function Optimization based on subspace identification models*. International Model Analysis Conference (IMAC), Kissimmee, Florida, Feb. 1999.
- [281] De Schutter, Bart. *Max-Algebraic Systems Theory for Discrete Event Systems*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, February, 1996.
- [282] De Schutter, Bart *Minimal state-space realization in linear system theory: an overview* . Journal of Computational and Applied Mathematics Volume: 121, Issue: 1-2, pp. 331-354, September 1, 2000.
- [283] Denham, Michael J.. *Canonical Forms for the Identification of Multivariable Linear Systems*. IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 646-656, December, 1974.
- [284] Deprettere, E.F.(ed.), *SVD and Signal Processing: Algorithms, Applications and Architectures*. North-Holland, Elsevier Science Pub. Co., 1988.
- [285] Deprettere, E.F. and Van der Veen, Alle-Jan (ed.), *Algorithms and Parallel VLSI Architectures. Volume A: Tutorials* Elsevier Science Pub., 1991.
- [286] Desai, Uday and Pal, Debajyoti. *A transformation Approach to Stochastic Model Reduction*. IEEE Transactions on Automatic Control, Volume AC-29, Number 12, pp.1097-1100 , December, 1984.
- [287] Desoer, C.A. and Whalen, B.H.. *A Note On Pseudoinverses*. JSIAM - Journal of the Society for Industrial and Applied Mathematics, Volume 11, No.2, 442-447, June, 1963.
- [288] Dewilde, Patrick and Van der Venn, Alle-Jan. *On the Hankel- Norm Approximation of Upper-Triangular Operators and Matrices*. Integral Equations and Operator Theory, vol. 17, no. 1, pp. 1-45, 1993.
- [289] Dewilde, Patrick and Van der Venn, Alle-Jan. *Time-Varying Systems and Computations*. Kluwer Academic Publishers, 1998.
- [290] Dey, Subhrakanti; Krishnamurthy, Vikram and Salmon-Legagneur, Thierry. *Maximum Likelihood Estimation of Time-Series With Markov Regime*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2856-2857, December, 1994.
- [291] Di Ruscio, David. *Topics in model based control with applications to the thermo mechanical pulping process*. Phd thesis, Department of Engineering Cybernetics, The Norwegian Institute of Technology, 1993.
- [292] Di Ruscio, David. *A method for the identification of state space models from input and output measurements*. Modeling, Identification and Control, vol. 16, no. 3, p.129-143, 1995.
- [293] Di Ruscio, David. *Subspace System Identification : Theory and Applications*. Lecture Notes Telemark Institute of Technology, January , 1995.
- [294] Di Ruscio, David . *A Method for Identification of Combined Deterministic Stochastic Systems*. Proceedings of 3rd European Control Conference, Rome , Italy, 429-434, September, 1995.
- [295] Di Ruscio, David . *The DSR Software : A Description*. Comunicação pessoal, 1-3, April, 1997.
- [296] Di Ruscio, David and Foss, Bjarne. *On state space model based predictive control*. Proceedings of the IFAC DYCOPS5 conference, 1998.

- [297] Di Ruscio, David . *Partial Least Squares (PLS) : Direct derivation via the Cayley Hamilton Theorem*. Lecture at : PROST, NTNU, October, 27, 1999.
- [298] Di Ruscio, David . *A weighted view on the partial least-squares algorithm*. Automatica, Volume: 36, Issue: 6, pp. 831-850 , June, 2000.
- [299] Dickinson, Bradley W.; Morf, Martin and Kailath, Thomas. *A Minimal Realization Algorithm For Matrix Sequences*. IEEE Transactions on Automatic Control, Vol.AC-19, No.1, 31-38, February, 1974.
- [300] Dickinson, B.W.; Kailath, T. and Morf, M.. *Canonical Matrix Fraction and State-Space Descriptions for Deterministic and Stochastic Linear Systems*. IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 656-674, December, 1974.
- [301] Diebold, Francis X.. *State Space Modeling Of Time Series - A Review Essay*. Journal of Economic Dynamics and Control , Vol.- 13, No.4, 597-612, October, 1989.
- [302] Dietrich, G.. *A New Formulation of The Hypermatrix Householder-QR Decomposition* Computer Methods In Applied Mechanics and Engineering, Volume 9, No.3, 273-280, 1976.
- [303] Ding, Zhi and Kennedy, Rodney A.. *An Algebraic Self-Training Algorithm For Multi-Channel System Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3612-3613, December, 1994.
- [304] Dongarra, Jack J., Kaufman, Linda and Hammarling, Sven.. *Squeezing the Most out of Eigenvalue Solvers on High-Performance Computers*. Linear Algebra and Its Applications, Vol.77, 113-136, May, 1977.
- [305] Dongarra, Jack and Eijkhoutd, Victor. *Standardized Numerical Linear Algebra Software* Encyclopedia of Computer Science and Technology - Volume 41, Marcel Dekker Publisher, pp 207-233, Edited by Allen Kent and James Williams, August, 1999.
- [306] Dongarra, Jack and Sullivan, Francis. *Guest Editors' Introduction: The Top 10 Algorithms*. Computing in Science & Engineering, Vol. 2, No. 1, p. 22-23, January/February, 2000.
- [307] Dou, H.-S. and Thien, Phan-. *On The Scalability of Parallel Computations On A Network Of Workstations*. Computational Mechanics, Vol.22, 344-354, 1998.
- [308] Dunkel, W.. *Identification Of A Nonlinear Model for State Estimation in an Airplane*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 661-667, July, 1991.
- [309] Duncan, D.B. and Horn, S.D.. *Linear Dynamic Recursive Estimation from the Viewpoint of Regression Analysis*. JASA - Journal of the American Statistical Association, Volume 67, No. 67, 815-821, December, 1972.
- [310] Ekongolo, Simon; Loverini, Corinne and Ragot, José. *Identification of LTI Systems Using Laguerre Models*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1728-1730, December, 1994.
- [311] El-Rewini, Hesham and Lewis, Ted G.. *Distributed and Parallel Computing*. Manning Publications Co. , 1997.
- [312] El-Sherief, H. and Sinha, N.K.. *Stochastic Approximation Algorithm for the Identification of Linear Multivariable Systems*. IEEE Transactions on Automatic Control, Vol.AC-24, No.2, 331-333, April, 1979.
- [313] El-Sherief, H. and Sinha, N.K.. *Bootstrap Estimation of Parameters and States of Linear Multivariable Systems*. IEEE Transactions on Automatic Control, Vol.AC-24, No.2, 340-343, April, 1979.
- [314] El-Sherief, H. and Sinha, N.K.. *Choice of Models for the Identification of Linear Multivariable Discrete-Time Systems* . Proceedings of IEE, Vol.126, No.12, 1326-1330, December, 1979.

- [315] El-Sherief, H.. *Parametric Identification of a State-Space Model of Multivariable Systems Using the Extended Least-Square Method*. IEEE Transactions on Systems, Man, and Cybernetics, Vol.SMC-11, No.3, 223-227, March, 1981.
- [316] El-Sherief, H.. *Multivariable Systems Structure and Parametric Identification Using the Correlation Method*. Automatica, Vol.17, No.3, 541-544, May, 1981.
- [317] El-Sherief, H. and Maud, M.A. *Dedicated Microprocessors for Realtime Identification for of Multivariable Systems*. Automatica, Volume 20, No.1, 129-131, January, 1984.
- [318] Ern, Alexandre and Giovangigli, Vincent. *Projected Iterative Algorithms With Application to Multicomponent Transport*. Linear Algebra and its Applications Volume: 250, pp. 289-315, January 1, 1997.
- [319] Eskinat, Esref. *System Identification Using Constrained Estimation*. Proceedings of 3rd European Control Conference, Rome , Italy, 856-861, September, 1995.
- [320] Eykhoff, Pieter. *System Identification - Parameter and State Estimation*. John Wiley & Sons, 1974.
- [321] Eykhoff, Pieter. *Trends and Progress in System Identification*. Automatica, Volume 18, No. 5, 635-636, September 1982.
- [322] Faddeeva, V. N.. *Computational methods of linear algebra*. Dover Publications, Inc., 1959.
- [323] Falcão, Djalma M.. *Parallel and Distributed Processing Applications in Power System Simulation and Control*. Controle Automação, Volume 5, No. único, 125-143, Out/Nov., 1994.
- [324] Fang, Wen-Hsien and Yagle, Andrew E.. *Two Methods for Toeplitz-plus-Hankel Approximation to a Data Covariance Matrix*. IEEE Transactions on Signal Processing, Vol.40, No.6, 1490-1498, June, 1992.
- [325] Farebrother, R.W.. *Linear Least Squares Computations*. Marcel Dekker, Inc., 1988.
- [326] Farwing, H.Meier zu and Unbehauen, H.. *Knowledge-Based System Identification*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 20-28, July, 1991.
- [327] Favoreel, W.. *Benchmark For Subspace System Identification Algorithms*. Katholieke Universiteit Leuven - Departement Elektrotechniek, 1-31, ESAT-SISTA/ TR - 1998.
- [328] Favoreel, W.; De Moor, Bart; and Van Overschee, Peter. *Subspace State Space System Identification For Industrial Processes*. Katholieke Universiteit Leuven - Departement Elektrotechniek, 1-31, ESAT-SISTA/ TR , 1-9, 1998.
- [329] Favoreel, Wouter; De Moor, Bart; Gevers, Michel and Van Overschee, Peter. *Model-free subspace-based LQG-design*. Internal Report 98-34, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1998.
- [330] Favoreel, Wouter and De Moor, Bart. *SPC: Subspace predictive control*. Internal Report 98-49, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1998.
- [331] Favoreel, Wouter; De Moor, Bart; Gevers, Michel and Van Overschee, Peter. *Closed-loop model-free subspace-based LQG-design*. Proc. of the IEEE Mediterranean Conference on Control and Automation, Haifa, Israel, pp. TP5-4, June 1999.
- [332] Favoreel, Wouter; De Moor, Bart; Gevers, Michel and Van Overschee, Peter. *Model-free subspace-based LQG-design*. Proc. of The America Control Conference, ACC99, San Diego, CA, USA, June 2-4, 1999.
- [333] Favoreel, Wouter; Sima, Vasile; Van Huffel, Sabine; Verhaegen, Michel and De Moor, Bart. *Subspace Model Identification of Linear Systems in SLICOT*. SLICOT Working Note 1998-6, p.1-28, October 1998.

- [334] Favoreel, Wouter; Van Huffel, Sabine; De Moor, Bart; Sima, Vasile and Verhaegen, Michel. *Comparative study between three different subspace identification methods*. Proceedings of the European Control Conference 1999, p. 1-6, Karlsruhe, Germany, August 31 - September 3, 1999.
- [335] Favoreel, Wouter; De Moor, Bart and Van Overschee, Peter. *Subspace Identification of Bilinear Systems Subject to White Inputs*. IEEE Transactions On Automatic Control, Vol. 44, No.6, pp. 1157-1165, June, 1999.
- [336] Favoreel, W.. *Subspace Methods For Identification and Control of Linear and Bilinear Systems*. Katholieke Universiteit Leuven - Departement Elektrotechniek PhD. Thesis, 1999.
- [337] Faurre, P.L.. *Realisations Markoviennes de Processus Stationnaires*. IRIA - Rapport de Recherche, No.13, March, 1973.
- [338] Faurre, P.L.. *Stochastic Realization Algorithms*. in *System Identification : Advances and Case Studies*, ed. Mehra, Raman K. and Lainiotis, Dimitri G.; Academic Press, 1976.
- [339] Fernando, K.V. and Nicholson, H.. *Identification of Linear Systems With Input and Output Noise: the Koopmans-Levin Method*. IEE Proceedings, Vol.132, Pt.D, No.1, 30-36, January, 1985.
- [340] Ferreres, G. and M'Saad, M.. *Output Error Methods For Set-Membership Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 25-30, September, 1995.
- [341] Ferreres, G.; Fromion, V. and M'Saad, M.. *Set-Membership Identification for Adaptive Control*. Proceedings of 3rd European Control Conference, Rome , Italy, 972-977, September, 1995.
- [342] Ferretti, G.; Maffezzoni C. and Scattolini, R.. *On The identification Of Time Delay With Least Squares Methods*. Proceedings of 3rd European Control Conference, Rome , Italy, 3691-3696, September, 1995.
- [343] Filaretov, V.F.; Dyda, A.A. and Vukobratovic, M.. *Sliding Regimes in Adaptive Robot Control*. Proceedings of 3rd European Control Conference, Rome , Italy, 2363-2367, September, 1995.
- [344] Filipovic, V.Z.. *Identification in  $H_\infty$  in the Presence of Deterministic and Stochastic Noise*. Proceedings of 3rd European Control Conference, Rome , Italy, 31-35, September, 1995.
- [345] Fischer, Sergio. *Séries Univariantes de Tempo - Metodologia de Box & Jenkins*. Fundação de Economia e Estatística - Porto Alegre, RS - Agosto, 1982.
- [346] Fitzsimons, Philip M.. *Reducing the computation required to solve a standard minimax problem*. Automatica, Volume 31, No.12, 1885-1887, December, 1995.
- [347] Flynn, Michael J.. *Some Computer Organization and Their Effectiveness*. IEEE Transactions on Computers, Vol.C-21, No.9, 948-960, September, 1972.
- [348] Forsman, Krister. *Hybrid Control Systems and Comprehensive Gröbner Bases*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4092-4097, December, 1994.
- [349] Forsythe, George E.. and Moler, Cleve B.. *Computer Solution of Linear Algebraic Systems*. Prentice-Hall, Inc., 1967.
- [350] Forsythe, George E.; Malcolm, Michael A. and Moler, Cleve B.. *Computer Methods for Mathematical Computations*. Prentice-Hall Inc., 1977.
- [351] Fosdick, Lloyd and Jessup, Elizabeth. *An Overview of Scientific Computing*. High Performance Scientific Computing - University of Colorado At Boulder HPSC Course Notes, p. 1-29, September, 1995.
- [352] Fox, Geoffrey and Messina, Paul C.. *Advanced Computer Architectures*. Scientific American - Volume 257, 45-52, October, 1985.

- [353] Fraleigh, J.B. and Beauregard, R.A.. *Linear Algebra*. Addison-Wesley Publishing Company, 1990.
- [354] Francis, J.G.F.. *The QR Transformation A Unitary Analogue to the LR Transformation - Part 1*. The Computer Journal, Vol.4, No.3, 265-271, October, 1961.
- [355] Francis, J.G.F.. *The QR Transformation - Part 2*. The Computer Journal, Vol.4, No.4, 332-345, January, 1962.
- [356] Franklin, Gene F. and Powell, J. David. *Digital Control of Dynamic Systems*. Addison-Wesley Publishing Company, 1981.
- [357] Frost, P. and Kailath, T.. *An innovations approach to least-squares estimation, Pt. III: Nonlinear estimation in white Gaussian noise*. IEEE Trans. Automat. Control, AC-16(3), 217-226, June 1971.
- [358] Fuchs, Jean-Jacques. *Multiscale identification of real sinusoids in noise*. Automatica, Vol. 30, No.1, 147-155, January, 1994.
- [359] Funkquist, Jonas. *On Modeling and Control of a Continuous Pulp Digester*. Teknologie Licentiate Thesis - Royal Institute of Technology, Department of Signals, Sensors and Systems, Automatic Control, Stockholm, Sweden, 1993.
- [360] Funkquist, Jonas. *A Dynamic Model of the Continuous Digester suitable for System Identification*. Proc. 36th SIMS Simulation Conference, Stockholm, Sweden, Aug.17-19, 1994.
- [361] Funkquist, Jonas. *On Modeling and Identification of a Continuous Pulp Digester*. Internal Report IR-S3-REG-9401 -Department of Signals, Sensors and Systems, Automatic Control, Stockholm, Sweden, 1994.
- [362] Furuta, Katsuhisa and Wongsaisuwan, Manop. *Discrete-time LQG Dynamic Controller Design Using Plant Markov Parameters*. Automatica, Vol.31, No. 9, 1317-1324, 1995.
- [363] Gajski, Daniel D.. *An Algorithm for Solving Linear Recurrence Systems On Parallel and Pipelined Machines*. IEEE Transactions on Computers, Vol.C-30, No.3, 190-206, March, 1981.
- [364] Gajski, Daniel D. and Peir, Jih-Kwon.. *Essential Issues in Multiprocessor Systems*. Computers, Volume 18, Number 6, 9-27, June, 1985.
- [365] Galiana, Francisco D. ; Handschin, Edmund and Fiechter, Albert R.. *Identification of Stochastic Electric Load Models From Physical Data*. IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 887-893, December, 1974.
- [366] Gallivan, K.A.; Heath, Michael T.; Ng, Esmond; Ortega, James M.; Peyton, Barry W.; Plemmons, R.J.; Romine, Charles H.; Sameh, A.H. and Voigt, Robert G.. *Parallel Algorithms for Matrix Computations*. SIAM - Society for Industrial and Applied Mathematics, 1990.
- [367] Gallivan, K.A.; Plemmons, R.J. and Sameh, A.H.. *Parallel Algorithms for Dense Linear Algebra Computations*. SIAM - Review, Vol.32, 54-135, 1990.
- [368] Gantmacher, F.R.. *The Theory of Matrices*. Chelsea Publishing Company - Vol. 1 e 2, 1977.
- [369] Gao, A.J.. *Weighted Least Squares. Algorithm For Continuous-Time Model* Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 263-264, December, 1994.
- [370] Gardner G.; Harvey, A. C. and Phillips, G.D.A.. *An Algorithm for Exact Maximum Likelihood Estimation of Autoregressive-Moving Average Models for Means of Kalman Filtering*. Applied Statistics, Vol.29, No. 3, 311-322, 1980.
- [371] Garnier, Hugues; Sibille, Patrick and Nguyen, Hoang L.. *A New Bias-Compensating Least-Squares Method For Continuous-Time MIMO Systems Identification Applied To a Laboratory-Scale Process*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 1711-1716, August, 1994.

- [372] Gautier, M. and Vandanjon, P.O.. *Reduction of Robot Base Parameters*. Proceedings of 3rd European Control Conference, Rome , Italy, 2368-2373, September, 1995.
- [373] Gautier, M.; Restrepo, P.P. and Khalil, W.. *Identification Of An Industrial Robot Using Filtered Dynamic Model*. Proceedings of 3rd European Control Conference, Rome , Italy, 2380-2385, September, 1995.
- [374] Geesey, R. and Kailath, T. . *An innovations approach to least-squares estimation Pt. V: Innovations representations and recursive estimation in colored noise*. IEEE Trans. Automat. Control, AC-18(5), 435-453, Oct. 1973.
- [375] Geist, Al; Beguelin, Adam; Dongarra, Jack; Jiang, Weicheng and Sunderam, Vaidy. *PVM 3 User's Guide and Reference Manual*. Oak Ridge National Laboratory - Oak Ridge, Tennessee - May, 1994.
- [376] Geist, Al; Beguellini, Adam; Dongarra, Jack; Jiang, Weicheng; Manchek, Robert and Sunderam, Vaidy. *PVM: Parallel Virtual Machine A Users' Guide and Tutorial for Networked Parallel Computing*. The MIT Press, 1994.
- [377] Gentleman, W. Morven. *Least Squares Computations by Givens Transformations Without Square Roots*. J.Inst.Maths.Applics - Volume 12, 329-336, 1973.
- [378] George, Alan; Liu, Joseph and Ng; Esmond. *A Data Structure for Sparse QR and LU Factorizations*. SIAM - Scientific and Statistical Computing, Vol.9 , No.1, 100-121, January, 1988.
- [379] George, Edwin Bruce and Youngblod, John N. *Parameter Estimating State Reconstruction*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 593-607, July, 1977.
- [380] Gersch, Will and Stone, David. *A Circular Lattice-Smoothness Priors Multichannel Time Varying Autoregressive Model*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 107-111, July, 1991.
- [381] Gevers, Michel and Kailath, Thomas. *An innovations approach to least-squares estimation, Pt. VI: Discrete-time innovations representations and recursive estimation*. IEEE Trans. Automat. Control, AC-18(6), 588-600, Dec. 1973.
- [382] Gevers, Michel and Wertz, Vincent. *Uniquely Identifiable State-space and ARMA Parametrizations For Multivariable Linear Systems*. Automatica, Volume 20, No.3, 333-347, May, 1984.
- [383] Gevers, Michel. *Connecting Identification and Robust Control: A new Challenge*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1-10, July, 1991.
- [384] Ghosal, Dipak; Serazzi, Giuseppe and Tripathi, Satish. *The Processor Working Set and Its Use in Scheduling Multiprocessor Systems*. IEEE Transactions On Software Engineering, Vol. 17, No.5, 443-453, May, 1991.
- [385] Giarre, Laura and Milanese, Mario. *Model Class Optimality In Set Membership Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1741-1745, December, 1994.
- [386] Giarre, L.; Kaciewicz, B.Z. and Milanese, M.. *Model Quality Evaluation In Set Membership Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 937-942, September, 1995.
- [387] Gilbert, Paul D.. *State Space and ARMA Models : An Overview of the Equivalence*. Bank of Canada, March, 1993.
- [388] Gill, Philip E.; Murray, Walter and Wright, Margaret H.. *Numerical Linear Algebra and Optimization*. Addison-Wesley, 1991.

- [389] Givens, Wallace. *Computation of Plane Unitary Rotations Transforming a General Matrix to Triangular Form*. J.Soc.Indust.App.Math - Volume 6, No.1, 26-50, 1958.
- [390] Glover, Keith. *Structural Aspects of System Identification*. Ph.D. Thesis - Massachusetts Institute of Technology - MIT, 1973.
- [391] Glover, Keith. *All Optimal Hankel-norm Approximations of Linear Multivariable Systems and Their  $L^\infty$  - error bounds*. International Journal of Control - Volume 39, No.6, 1115-1193, June , 1984.
- [392] Glover, Keith. *An Approximate Realization Algorithm Which Directly Identifies The System Poles*. IFAC - Helsinki - 1789-1796, 1978.
- [393] Goldberg, Jack L.. *Matrix Theory with Applications*. McGraw-Hill, Inc., 1992.
- [394] Golub, G.H. and Kahan, W.. *Calculating The Singular Values and Pseudo-Inverse of a Matrix*. J.SIAM Numer. Anal., Vol.2, 205-224, 1965.
- [395] Golub, G.H.. *Numerical Methods for Solving Linear Least Squares Problems*. Numerische Mathematik 7, 206-216, 1965.
- [396] Golub, G.H.. *Matrix Decompositions and Statistical Calculations*. in Statistical Computation - ed.Milton, Roy C. and Nelder, John A. , 365-397, Academic Press, Inc., 1969.
- [397] Golub, G.H. and Reinsch, C.. *Singular Value Decomposition and Least Squares Solutions*. Numerische Mathematik 14, 403-420, 1970.
- [398] Golub, Gene H.. and Loan, Charles F. Van. *An Analysis of The Total Least Squares Problem*. SIAM J.Numer.Anal., Vol.17, No.6 , 883-893, December, 1980.
- [399] Golub, Gene H. and Loan, Charles F. Van. *Matrix Computations*. The Johns Hopkins University Press, 1989.
- [400] Golub, Gene H. and Ortega, James M.. *Scientific Computing: An Introduction with Parallel Computing*. Academic Press, 1993.
- [401] Gonzaga, Clovis C.; Lara, Hugo J.. *A Note on Properties of Condition Numbers*. Linear Algebra and its Applications Volume: 261, pp. 269-273, August, 1997.
- [402] Goodrich, Robert L. and Caines, Peter E.. *Linear System Identification from Nonstationary Cross-Sectional Data*. IEEE Transactions on Automatic Control, Vol. AC-24, No.3, 403-411, June, 1979.
- [403] Goodwin, G. C. and Payne, R. L.. *Dynamic System Identification*. Academic Press, 1977.
- [404] Gopalaratnam, Giriya; Parameswaran, V.; Raol, Jr. and Balakrishna, S.. *Estimation Of Aircraft Performance & Stability Characteristics by Dynamic Manoeuvres Using Maximum Likelihood Estimation Techniques*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 649-654, July, 1991.
- [405] Gopinath, B. *On The Identification of Linear Time-Invariant Systems From Input-Output Data*. The Bell System - Technical Journal - Vol.48, No.5, 1101-1113, May-June, 1969.
- [406] Göransson; Jansson, Magnus and Ottersten, Björn. *Spatial and Temporal Frequency Estimation of Uncorrelated Signals Using Subspace Fitting*, Proc. of 8th IEEE Signal Processing Workshop on Statistical Signal and Array Processing, Corfu, Greece, pp. 94-96, June, 1996.
- [407] Goto, Satoru and Uosaki, Katsuji. *An On-Line Identification Procedure For Linear Stochastic Systems*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1260-1265, July, 1991.
- [408] Gourieroux, Christian and Monfort, Alain. *Statistics and Econometric Models. Volume 1 - General Concepts, Estimation, Prediction and Algorithms* Cambridge University Press, 1995.

- [409] Gouriéroux, Christian and Monfort, Alain. *Statistics and Econometric Models. Volume 2 - Testing, Confidence Regions, Model Selection and Asymptotic Theory*. Cambridge University Press, 1995.
- [410] Gouriéroux, Christian and Monfort, Alain. *Time Series and Dynamic Models*. Cambridge University Press, 1997.
- [411] Graupe, D., Krause, D.J. and Moore, J.B. *Identification of Autoregressive Moving-Average Parameters of Time Series*. IEEE Transactions on Automatic Control, Vol.AC-20, No.1, 104-107, February, 1975.
- [412] Gray, R.M. and Davisson, L.D.. *Introduction to Statistical Signal Processing*. obtido em : <http://ee-www.stanford.edu/gray/sp.html> ; may, 2000.
- [413] Grepper, Pierre O. *Direct Identification of Dynamical Systems With Application To Air-Vehicles*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 583-592, July, 1977.
- [414] Greville, T. N. E. *Some Applications of The Pseudoinverse of The a Matrix*. The SIAM Review, Volume 2, No.1, 15-22, January, 1960.
- [415] Greville, T. N. E. *Solutions of The Matrix Equation  $\mathbf{XAX} = \mathbf{X}$ , and Relations Between Oblique and Orthogonal Projectors*. The SIAM Journal on Applied Mathematics, Volume 26, No.4, 828-832, June, 1974.
- [416] Groß, Jürgen. *On the product of orthogonal projectors*. Linear Algebra and its Applications Volume: 289, Issue: 1-3, pp. 141-150, March 1, 1999.
- [417] Groß, Jürgen; Trenkler, Götz and Liski, Erkki P. Razzaghi. *Necessary and sufficient conditions for superiority of misspecified restricted least squares regression estimator*. Journal of Statistical Planning and Inference Volume: 71, Issue: 1-2, pp. 109-116, August 1, 1998.
- [418] Grübel, G.; Varga, A.; Boom, A. van den and Geurts, A. J.. *Towards a Coordinated Development of Numerical CACSD Software: the RASP/SLICOT Compatibility Concept*. Proc. CACSD'94 Symposium, Tucson, Arizona, pp. 499-504, 1994.
- [419] Guidorzi, R.. *Canonical Structures In The Identification Of Multivariable Systems*. Automatica, Volume 11, No.4, 361-374, July, 1975.
- [420] Guidorzi, Roberto P.. *Invariants and Canonical Forms For Systems Structural and Parametric Identification*. Automatica, Volume 17, No.1, 117-133, January, 1981.
- [421] Guidorzi, Roberto P.. *Identification of The Maximal Number Of Linear Relations From Noisy Data*. Systems & Control Letters, Volume 24, Number 3 - pag.159-165 - February, 1995.
- [422] Gulliksson, Mårten and Wedin, Per-Åke. *Modifying The QR-Decomposition To Constrained and Weighted Linear Least Squares*. SIAM - Journal on Matrix Analysis and Applications, Volume 13, No.4, 1298-1313, October, 1992.
- [423] Gunnarsson, Svante and Hjalmarsson, Håkan. *Some aspects of iterative identification and control design schemes*. obtido através de : <http://joakim.isy.liu.se>, 1-20, artigo escrito em : June, 1994.
- [424] Guo, Lei; and Ljung, Lennart. *Performance Analysis of General Tracking Algorithms*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2851-2855, December, 1994.
- [425] Guo, Lei; and Ljung, Lennart. *The Role of Model Validation for Assessing The Size of The Unmodelled Dynamics*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3894-3899, December, 1994.
- [426] Gupta, Narendra K.. *Efficient Computation of Gradient and Hessian of Likelihood Function in Linear Dynamic Systems*. IEEE Transactions on Automatic Control, Vol. AC-21, No.5, 781-783, October, 1976.



- [427] Gupta, R.D. and Fairman, F.W.. *Parameter Estimation For Multivariable Systems*. IEEE Transactions on Automatic Control, Vol. AC-19, No.5, 546-549, October, 1974.
- [428] Gustafson, John L.. *Reevaluating Amdahl's Law*. Communications of the ACM, Vol. 31, No.5, 532-533, May, 1988.
- [429] Gustafsson, F. and Ninness, B.M. *Asymptotic Power and The Benefit of Under-Modeling in Change Detection*. Proceedings of 3rd European Control Conference, Rome, Italy, 1237-1242, September, 1995.
- [430] Gustafsson, Fredrik; Gunnarsson and Ljung, Lennart. *On Time-Frequency Resolution Of Signal Properties Using Parametric Techniques*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2259-2264, December, 1994.
- [431] Gustafsson, T.K. and Mäkilä, P.M.. *Black-Box Methods For Identification Of Set-Valued Models For Control*. Proceedings of 3rd European Control Conference, Rome, Italy, 943-948, September, 1995.
- [432] Gustafsson, Tony. *On The State Estimation Problem for Subspace Identification*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, FM07-2, December, 1998.
- [433] Gustafsson, Tony; Lovera, Marco and Verhaegen, Michel. *A Novel Algorithm for Recursive Instrumental Variable Based Subspace Identification*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, FM07-3, December, 1998.
- [434] Gustafsson, Tony. *Subspace Identification Using Instrumental Variable Techniques*. Automatica, Vol.37, N.12, pp. 2005-2010, 2001.
- [435] Gustavsson, I.. *Survey Of Applications Of Identification in Chemical And Physical Processes*. Automatica, Volume 11, No.1, 3-24, January, 1975.
- [436] Gustavsson, I., Ljung L. and Söderström, T.. *Identification Of Process in Closed Loop - Identifiability and Accuracy Aspects*. Automatica, Volume 13, No.1, 59-75, January, 1977.
- [437] Haas, Violet B.. *Reduced Order State Estimation For Linear Systems With Exact Measurements*. Automatica, Volume 20, No.2, 225-229, March, 1984.
- [438] Hager, William W.. *Applied Numerical Linear Algebra*. Prentice Hall, 1988.
- [439] Hahn, H.; Leimbach, K.-D. and Piepenbrink. *Inertia Parameter Identification of Rigid Bodies Using A Multi-Axis Test Facility*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 1735-1737, August, 1994.
- [440] Hajdasiński, A.K.. *Linear Multivariate Systems : Preliminary Problems in Mathematical, Description, Modelling and Identification*. Eindhoven University of Technology, TH-Report 80-E-106, April, 1980.
- [441] Hakvoort, Richard G. and Van den Hof, Paul M.J.. *An Instrumental Variable Procedure for the Identification of Probabilistic Frequency Response Uncertainty Regions*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3596-3601, December, 1994.
- [442] Hakvoort, Richard G. and Van den Hof, Paul M.J.. *A Practically Feasible Procedure for Identification of  $H_\infty$  Error Bounds*. Proceedings of 3rd European Control Conference, Rome, Italy, 925-930, September, 1995.
- [443] Hale, Jack K. and La Salle, Joseph P.. *Differential Equations and Dynamical Systems*. Academic Press, 1967.
- [444] Halmos, Paul Richard. *Measure Theory*. Van Nostrand Reinhold, 1950.
- [445] Halmos, Paul Richard. *Finite-Dimensional Vector Spaces*. Princeton, Van Nostrand, 1958.

- [446] Halmos, Paul Richard. *Teoria Ingênua dos Conjuntos*. Editora Polígono S.A., 1970.
- [447] Hamilton, J.D. *Time Series Analysis*, Princeton U.P., 1994.
- [448] Hammer, Jacob. *On Some Control Problems In Molecular Biology*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4098-4103, December, 1994.
- [449] Han, Guang-Wen. *Impulse Transfer Function Equivalent Models For Multivariable Linear Stochastic System and Its Realization With Finite Markov Parameters*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1996-2001, August, 1988.
- [450] Hannan, E.J.. *Time Series Analysis*. Science Paperbacks and Methuen & Co. Ltd., 1960.
- [451] Hannan, E.J. and Kavalieris, L.. *Multivariate Linear Time Series Models*. Advances In Applied Probability, Volume 16, Number 3, 492-561, September, 1984.
- [452] Hannan, E.J.; Krishnaiah, P.R. and Rao M.M. (Ed.). *Handbook of Statistics 5 : Time Series in the Time Domain*. North-Holland, 1985.
- [453] Hannan, E.J. and Deistler, M.. *The Statistical Theory Of Linear Systems*. John Wiley & Sons, 1988.
- [454] Hanselman, Duane e Littlefield, Bruce. *MATLAB 5 - Versão do Estudante*. Makron Books, 1999.
- [455] Hansen, Per Brinch. *Householder Reduction of Linear Equations*. ACM - Computing Surveys - Volume 24, No. 2, 185-194, June, 1992.
- [456] Hanson, Richard J. and Lawson, Charles L.. *Extensions and Applications of the Householder Algorithm for Solving Linear Least Squares Problems*. Mathematics of Computation - Volume 23, No. 108, 787-812, October, 1969.
- [457] Hanson, Richard J. and Dyer, P. *A Computacional Algorithm For Sequential Estimation*. The Computer Journal, Vol.14, No. 3, 285-290, August, 1971.
- [458] Hanson, Richard J. and Lueche, Glenn R.. *Parallel Inverse Iteration for Eigenvalue and Singular Value Decompositions*. SIAM NEWS - Volume 27, No. 5, 14-16, May/June, 1994.
- [459] Hanzon B.. *A Geometric Approach To System Identification Using Model-Reduction Techniques*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 983-988, August, 1988.
- [460] Hari; K. V. S. and Gummadavelli, U.. *Effect of spatial smoothing on the performance of subspace methods in the presence of array model errors*. Automatica, Vol. 30, No.1, 11-26, January, 1994.
- [461] Harvey, Andrew C.. *Time Series Models*. Harvester Wheatsheaf, 1993.
- [462] Harvey, Andrew C.. *Forecasting, structural time series models and the Kalman filter*. Cambridge University Press, 1992.
- [463] Harshavardhana, P. ; Jonckheere, E.A. and Silverman, L.M.. *Stochastic Balancing and Approximation-Stability and Minimality*. IEEE Transactions on Automatic Control, Vol. AC-29, No.8, 744-746, August, 1984.
- [464] Hartnett, M.K.; Lightbody, G. and Irwin, G.W.. *Identification of state models using principal components analysis*. Chemometrics and Intelligent Laboratory Systems, Vol. 46, No. 2, pp. 181-196, 1999.
- [465] Hatanaka, Toshiharu and Uosaki, Katsuji. *Optimal Input Design For Discrimination of Linear Stochastic Models Based On Kullback-Leibler Discrimination Information Measure*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 841-845, August, 1988.

- [466] Hatanaka, Toshiharu; Uosaki, Katsuji and Yamada, Yasuhiro. *Evolutionary Approach To System Identification*. 11th IFAC Symposium on System Identification, Vol.3, 1383-1388, July, 1977.
- [467] Hautus, M.L.J. and Silverman, L.M.. *System Structure and Singular Control*. Linear Algebra and Its Applications, Vol.50, 369-402, April, 1983.
- [468] Havenner, Arthur and Leng, Zhiqiang. *Improved Estimates of The Parameters of State Space Time Series Models*. Journal of Economic Dynamics and Control, Vol. 20, No.5, 767-789, May, 1996.
- [469] Haverkamp, B.R.J.; Chou, C.T. and Verhaegen, M.. *Subspace Identification Of Continuous-Time Wiener Models*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, TA08-4, December, 1998.
- [470] Haverkamp, Bert. *Efficient implementation of subspace method identification algorithms*. NICO-NET Report 1999-3; p.1-12; March 1999.
- [471] Hayes, Monson H.. *Statistical Digital Signal Processing and Modeling*. John Wiley & Sons, 1996.
- [472] Haykin, Simon. *Adaptive Filter Theory*. Prentice Hall, 1991.
- [473] Heath, M.T. and Sorensen, D.C.. *A Pipelined Givens Method for Computing The QR Factorization of a Sparse Matrix*. Linear Algebra and Its Applications, Volume 77, 189-203, May, 1986.
- [474] Heath, M.T.; Laub, A.J.; Paige, C. C. and Ward, R.C.. *Computing the Singular Value Decomposition of the Product of Two Matrices*. SIAM - Journal on Scientific and Statistical Computing, Volume 7, No.4, 1147-1159, October, 1986.
- [475] Hecker, O., Knapp, T. and Isermann, R.. *Robust Adaptive Control of a Time Varying Process Using Parallel Recursive Estimators*. Automatica, Volume 30, No.4, 599-604, April, 1994.
- [476] Heemink, A.W.; Bolding, K. and Verlaan, M.. *Storm Surge Forecasting Using Kalman Filtering (a review)*. Report 95-99, Faculty of Technical Mathematical and Informatics, 1999.
- [477] Hefes, H.. *The Effect of Erroneous Models on The Kalman Filter Response*. IEEE Transactions On Automatic Control, Volume AC-11, No.3, 541-542, July, 1966.
- [478] Heij, C.. *Deterministic Identification of Dynamical Systems*. Lecture Notes in Control and Information Sciences - Springer-Verlag, Volume 127, 1989.
- [479] Heij, C. and Scherrer, W.. *On The Consistency of Identification by Dynamic Factor Models*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2880-2885, December, 1994.
- [480] Heinig, G. and Jankowski, P. *Parallel and Superfast Algorithms for Hankel Systems of Equations*. Numerische Mathematik, Volume 58, Number 1, 109-127, 1990.
- [481] Hekstra, Gerben J.; Deprettere, Ed F.; Monari, Monica and Heusdens, Richard. *Recursive Approximate Realization Of Image Transforms With Orthonormal Rotations*. correspondência pessoal gerben@cas.et.tudelft.nl, 1999.
- [482] Heller, Don. *A Survey of Parallel Algorithms in Numerical Linear Algebra*. SIAM Review - Volume 20, No. 4, 740-777, October, 1978.
- [483] Helstrom, Carl W.. *Probability and Stochastic Processes for Engineers*. Macmillan Publishing Company, 2nd ed., 1991.
- [484] Henderson, H. V. and Searle, S.R.. *On Deriving The Inverse of A Sum of Matrices*. SIAM Review - Volume 23, No. 1, 53-60, January, 1981.
- [485] Henry, Greg; Watkins, David and Dongarra, Jack. *A Parallel Implementation of the Nonsymmetric QR Algorithm for Distributed Memory Architectures*. University of Tennessee Computer Science Technical Report, CS-97-352, March 1997 and LAPACK Working Note 121.

- [486] Hernández, V. and Vidal, A.M.. *Parallel Algorithms For The Singular Value Decomposition and Their Applications in Signal and Control Systems*. Proceedings of the First European Control Conference - Grenoble, France, 596-601, July, 1991.
- [487] Hill, David R. and Moler, Cleve B.. *Experiments in Computational Matrix Algebra*. Random House, 1988.
- [488] Hirashima, Yoichi and Iiguni, Youji. *An Identification System Design Using A CMAC With a Learning Algorithm Based on The Kalman Filter*. 11th IFAC Symposium on System Identification, Vol.3, 1389-1394, July, 1977.
- [489] Hjalmarsson, Håkan ; Gevers, Michel; Bruyne, Franky De and Leblond, Juliette. *Identification for Control : Closing The Loop Gives More Accurate Controllers*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4150-4155, December, 1994.
- [490] Ho, B. L. and Kalman, R.E.. *Effective Construction of Linear State-Variable Models from Input/Output Data*. Proceedings, 3rd Annual Allerton Conference on Circuit and system Theory, pp. 449-459, 1965.
- [491] Ho, B.L. and Kalman, R.E. *Efficient construction of linear state variable models from input/output functions*. Regelungstechnik, V.14, pp. 545-548, 1966.
- [492] Hockney, R.W. and Jesshope, C.R.. *Parallell Computers - Architecture, Programming and Algorithms*. IOP Publishing Ltd, 1988.
- [493] Hodges, S.D. and Moore P.G.. *Data Uncertainties and Least-Squares Regression*. Applied Statistics, Volume 21, No.2, 185-195, 1972.
- [494] Hodžić, Migdat I. *Iterative Identification Of Large Scale Systems*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 672-677, August, 1988.
- [495] Hoffman, M.R. and Reynolds, L.K.. *The Role of Control Systems in The Management of a Modern Water Utility*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 1479-1485, August, 1994.
- [496] Holcomb, Tyler R.; Hjalmarsson, Håkan; Morari, Manfred and Tyler, Matthew L.. *Significance regression: a statistical approach to partial least squares*. Journal of Chemometrics, Volume 11, Issue 4, 1997.
- [497] Holzhüter, T.. *Order Selection for ARX-Models via Hankel Approximation of Hight Order Estimates*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 936-941, July, 1991.
- [498] Homssi, L., Tilli, A. and Despujols, A.. *Continuous-Time Process Identification: Comparison of Eight Methods and Particular Aspects*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1634-1642, July, 1991.
- [499] Hongyue, Zhang and Xiangui, Yu. *New Rescue Methods for Preventing the Divergence of the Fast Least-Squares Algorithm*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1108-1113, August, 1988.
- [500] Horn, Roger A. and Johnson , Charles R.. *Matrix Analysis*. Cambridge University Press, 1985.
- [501] Höskuldsson, Agnar. *Dimension of Linear Models*. Chemometrics and Intelligent Laboratory Systems, Vol. 32, No. 1, pp. 37-55, 1996.
- [502] Hotelling, Harold. *Relations Between Two Sets Of Variates*. Biometrika, Volume 28, 321-377, 1936.
- [503] Hou, Daqing and Hsu, Chin Shung. *State-space Model Identification With Data Correlation*. International Journal of Control, Vol.53, No.1, 181-192, January, 1991.

- [504] Householder, Alston S.. *A Survey of Some Closed Methods for Inverting Matrices*. Journal of the Society for Industrial and Applied Mathematics, Vol.5, No.3, 155-169, September, 1957.
- [505] Householder, Alston S.. *A Class of Methods for Inverting Matrices*. Journal of the Society for Industrial and Applied Mathematics, Vol.6, No.2, 189-195, June, 1958.
- [506] Householder, Alston S.. *Generated Error in Rotational Tridiagonalization*. Journal of the Association for Computing Machinery, Vol.5, 335-338, 1958.
- [507] Householder, Alston S.. *Unitary Triangularization of a Nonsymmetric Matrix*. Journal of the Association for Computing Machinery - JACM, Vol.5, 339-342, 1958.
- [508] Householder, Alston S.. *The Theory of Matrices in Numerical Analysis*. Dover Publications, Inc., 1964.
- [509] Huffel, Sabine Van and Vandeswalle, Joos.. *The Total Least Squares Problem Computational Aspects and Analysis*. SIAM - Society for Industrial and Applied Mathematics, 1991.
- [510] Hukl, G. ; Belanský, J.; Belavý, C. ; Antoniová, M. ; Szuda, J. and Végh, P.. *Identification and Modelling of Lumped Input and Distributed Output Systems at The Accuracy Requested*. Proceedings of 3rd European Control Conference, Rome , Italy, 447-452, September, 1995.
- [511] Hung, George; Stark, Lawrence and Eykhoff, Pieter. *On Interpretation of Kernels. I. Computer Simulation of Responses to Impulse Pairs*. Annals of Biomedical Engineering, Volume 5, Number 2, pp. 130-143, June, 1977.
- [512] Hwang, Kai and Briggs, Fayé A.. *Computer Architecture and Parallel Processing*. McGraw Hill, 1984.
- [513] Hwang, Kai.. *Multiprocessors Supercomputers for Scientific/Engineering Applications*. Computers, Volume 18, Number 6, 57-73, June, 1985.
- [514] Hwang, Kai and Xu, Zhiwei. *Scalable Parallel Computing*. WCB/McGraw-Hill, 1998.
- [515] IBM *RS/6000 SP Overview, Planning and Installation - Instructor Guide 1*. IBM Education and Training, 1996.
- [516] Ilic'-Spong, M. ; Crow, M.L. and Pai, M.A.. *Transient Stability Simulation by Waveform Relaxation Methods*. IEEE Transactions on Power Systems, Vol.PWRS-2, No.4, 943-952, November, 1987.
- [517] Ionescu, Vlad; Oară and Weiss, Martin. *General Matrix Pencil Techniques for The Solution of Algebraic Riccati Equations : A Unified Approach*. IEEE Transactions on Automatic Control, Vol.42, No.8, 1085-1097, August, 1997.
- [518] Ipsen, Ilse C.F.; Saad, Youcef and Schultz, Martin H.. *Complexity of Dense-Linear-System Solution on a Multiprocessor Ring*. Linear Algebra and Its Applications, Volume 77, 205-239, May, 1986.
- [519] Isermann, R.; Baur, U.; Bamberger, P.; Kneppo, P. and Siebert, H.. *Comparison of Six On-Line Identification and Parameter Estimation Methods*. Automatica, Volume 10, No.1, 81-103, January, 1974.
- [520] Isermann, R.. *Practical Aspects of Process Identification*. Automatica, Volume 16, No.5, 575-587, September, 1980.
- [521] Ito, Y.. *Adaptive Control of Hierarchical Econometric Model With Decentralized Decision-Makers*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1871-1875, August, 1988.
- [522] Jainandunsing K. and Deprettere, E.F.. *A New Class Of Parallel Algorithms For Solving Systems Of Linear Equations*. SIAM J.Sci.Stat.Comput. Vol.10, No.5, 880-912, 1989.

- [523] Jakoby, W. and Pandit, M.. *A Prediction-error-method for Recursive Identification of Nonlinear Systems*. Automatica, Volume 23, No.4, 491-496, July, 1987.
- [524] Janakiraman, P.A. and Renganathan, S.. *Recursive Computation of Pseudo-Inverse of Matrices*. Automatica, Volume 18, No. 5, 631-633, September 1982.
- [525] Jansson, Magnus and Wahlberg, Bo. *On Weighting in State-Space Subspace System Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 435-440, September, 1995.
- [526] Jansson, Magnus. *On Performance Analysis of Subspace Methods in System Identification and Sensor Array Processing* Licentiate Thesis - Royal Institute of Technology - Stockholm, 1995.
- [527] Jansson, Magnus and Wahlberg, Bo. *A Linear Regression Approach To State-Space Subspace System Identification*. Signal Processing, Vol.52, 103-129, 1996.
- [528] Jansson, Magnus; Göransson, Bo and Ottersten, Björn. *Frequency Estimation of Uncorrelated Signals Using Subspace Fitting*. Royal Institute of Technology - Stockholm, IR-S3-REG-9616, 1996.
- [529] Jansson, Magnus. *On Subspace Methods in System Identification and Sensor Array Signal Processing*. Ph.D. Thesis - Royal Institute of Technology - Stockholm, 1997.
- [530] Jansson, Magnus and Wahlberg, Bo. *On Consistency of Subspace System Identification Methods*. Automatica, vol. 34, no. 12, pp. 1507-1519, Dec. 1998.
- [531] Jansson, Magnus, Swindlehurst, A. Lee and Ottersten, Björn. *Robust Weighted Subspace Fitting in the Presence of Array Model Errors* . IEEE International Conference on Acoustics, Speech and Signal Processing, 1998.
- [532] Jansson, Magnus; Göransson, Bo and Ottersten, Björn. *A Subspace Method for Direction of Arrival Estimation of Uncorrelated Emitter Signals*. IEEE Trans. on Signal Processing, vol. 47, no. 4, pp. 945-956, April, 1999.
- [533] Jansson, Magnus and Stoica, Petre. *Forward-only and forward-backward sample covariances - A comparative study*. Signal Processing, Volume: 77, Issue: 3, pp. 235-245, September, 1999.
- [534] Jean-Marie, Alain; Lefebvre-Barbaroux, Sophie and Liu, Zhen. *An analytical approach to the performance evaluation of master-slave computational models*. Parallel Computing Volume: 24, Issue: 5-6, June, pp. 841-862, 1998.
- [535] Jennings, Alan and McKeown, J.J.. *Matrix Computation*. John Wiley, 1977.
- [536] Jennings, Alan. *Matrix Computation For Engineers and Scientists*. John Wiley & Sons, 1977.
- [537] Jepsen, O.N. and Rake, H.. *The Adjusted Least Squares Method For Identification Of Dynamical Systems*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1290-1293, July, 1991.
- [538] Johansen, Tor A. and Foss, Bjarne A.. *Identification of Non-linear Systems Structure and Parameters Using Regime Decomposition*. Automatica, Volume 31, Number 2, 321-326, February, 1995.
- [539] Johansen, Tor A. and Weyer, Erik. *Structure Identification Using Separate Validation Data - Asymptotic Properties*. Proceedings of 3rd European Control Conference, Rome , Italy, 833-838, September, 1995.
- [540] Johansen, Tor Arne. *Operating Regime Based Process Modeling and Identification* . Dr. Ing. Thesis, Department of Engineering Cybernetics - The Norwegian Institute of Technology - University of Trondheim, 1994.

- [541] Johansson, Rolf. *Multivariable System Identification Via Continued-Fraction Approximation*. IEEE Transactions on Automatic Control, Vol. 40, NO.3, 507-512, March, 1995. and Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2288-2293, December, 1994.
- [542] Johansson, Rolf; Holm, Half; Olsson, Bertil S. and Brandt, Johan. *System Identification of Atrial Activation During Chronic Atrial Fibrillation in Man*. IEEE Transactions on Automatic Control, Vol.AC-43, No.6, 790-799, June, 1998.
- [543] Johansson, R.; Verhaegen, M.; Chou, C.T. and Robertsson, A. *Behavioral Model Identification*. Proceedings of the 37th IEEE Conference on Decision and Control, Volume: 1, Page(s): 126 -131, vol.1, 1998
- [544] Johansson, Rolf. *System Modeling and Identification*. Prentice Hall, Englewood Cliffs, 2001.
- [545] Johnson, M.A. and Daniels, M.W.. *Identification of Essential States For Reduced-Order Models Using a Modal Analysis*. IEE Proceedings, Vol.132, Pt.D, No. 3, 111-118, May, 1985.
- [546] Johnson, Richard A. and Wichern, Dean W.. *Applied Multivariate Statistical Analysis*. Prentice Hall, 1998.
- [547] Jones, A.. *Spiral - A New Algorithm for Non-Linear Parameter Estimation Using a Least Squares*. The Computer Journal, Vol.13, No. 3, 301-308, August, 1970.
- [548] Jover, J.M. and Kailath,T.. *A Parallel Architecture for Kalman Filter Measurement Update and Parameter Estimation*. Automatica, Volume 22, No.1, 43-57, January 1995.
- [549] Juang, Jer-Nan and Pappa, Richard S.. *An Eigensystem Realization Algorithm for Modal Parameters Identification and Model Reduction*. Journal Guidance, Vol.8, n.5, pp. 620-627 , September-October, 1985.
- [550] Juang, Jer-Nan and Pappa, Richard S.. *Effects of Noise on Modal Parameters Identified By the Eigensystem Realization Algorithm*. Journal Guidance, Vol.9, n.3, pp. 294-303 , May-June 1986.
- [551] Juang, Jer-Nan. *Applied System Identification*. Prentice Hall, 1994.
- [552] Juang, Jer-Nan. *State-Space System Realization With Input- and Output-Data Correlation*. NASA Technical Paper-3622, pp. 41, April 1997.
- [553] Juang, Jer-Nan and Longman, Richard W.. *Optimized System Identification*. NASA Technical Paper-209711, pp. 31, October 1999.
- [554] Juang, Jer-Nan and Chen, Ya-Chin. *Signal Prediction With Input Identification*. NASA Technical Paper-209705, pp. 28, October 1999.
- [555] Juditsky, Anatoli; Hjalmarsson, H åkan; Benveniste,Alberto; Delyon, Bernard; Ljung, Lennart; Sjöberg, Jonas and Zhang, Qinghua. *Nonlinear black-box models in system identification: mathematical foundations*. Automatica, Volume 31, No.12, 1725-1750, December, 1995.
- [556] Jutan, Arthur; Tremblay, J.P.; MaGregor, J.F. and Wright, J.D.. *Multivariable Computer Control of a Butane Hydrogenolysis Reactor : Part I - State Space Reactor Modeling , Part II - Data Collection, Parameter Estimation and Stochastic Disturbance Identification, Part III - On-line Quadratic Stochastic Control Studies*. Aiche Journal - Chemical Engineering Research and Development, Volume 23, No.5, 732-758, September, 1977.
- [557] Kadlec, Jiří. *A Recursive Modified Gram-Schmidt Identification With Directional Tracking of Parameters*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1707-1711, July, 1991.
- [558] Kahan, Robert E.. *Networks For Advanced Computing*. Scientific American - Volume 257, 128-176, October, 1985.

- [559] Kahan, W. *Numerical Linear Algebra*. Canadian Mathematical Bulletin, Vol.9, no.6, 757-801, 1966.
- [560] Kahaner, David; Moler, Cleve and Nash, Stephen. *Numerical Methods and Software*. Prentice-Hall, 1989.
- [561] Kailath, Thomas. *An innovations approach to least-squares estimation, Pt. I: Linear filtering in additive noise*. IEEE Trans. Automat. Control, AC-13(6), 646-655, Dec. 1968.
- [562] Kailath, T. and Frost P.. *An innovations approach to least-squares estimation, part ii: Linear smoothing in additive white noise*. IEEE Trans. Automat. Control, AC-13(6), 655-660, Dec. 1968.
- [563] Kailath, T. and Geesey, R.. *An innovations approach to least-squares estimation, Part IV: Recursive estimation given the covariance functions*. IEEE Trans. Automat. Control, AC-16(6), 720-727, Dec. 1971.
- [564] Kailath, Thomas. *Some New Algorithms for Recursive Estimation in Constant Linear Systems*. IEEE Transactions on Information Theory, Volume IT-19, No.6, 750-760, November, 1973.
- [565] Kailath, Thomas; Kung, Sun-Yuan and Morf, Martin.. *Displacement Ranks of Matrices and Linear Equations*. Journal of Mathematical Analysis and Applications, Vol.68, No.2, 395-407, April, 1979.
- [566] Kailath, Thomas. *Linear Systems*. Prentice-Hall, Inc., 1980.
- [567] Kailath, Thomas and Chun, J. *Generalized Displacement Structure for Block-Toeplitz, Toeplitz-Block and Toeplitz-Derived Matrices*. SIAM - Journal on Matrix Analysis and Applications, Volume 15, No.1, 114-128, January, 1994.
- [568] Kailath, Thomas and Sayed, Ali H.. *Displacement Structure: Theory and Applications*. SIAM - Review, Volume 37, No.3, 297-386, September, 1995.
- [569] Kailath, Thomas. *Array Algorithms for Structured Matrices*. Notas do seminário proferido no Information Systems Laboratory, 2000.
- [570] Kagiwada, Harriet H.. *System Identification - Methods and Applications*. Addison-Wesley, 1974.
- [571] Kalivas, John H.. *Basis sets for multivariate regression*. Analytica Chimica Acta Volume: 428, Issue: 1, pp. 31-40, February 1, 2001.
- [572] Kalman, R.E. and Bertram, J.E.. *A Unified Approach To The Theory of Sampling Systems*. Journal of The Franklin Institute, Vol.267, pp. 405-436, 1959.
- [573] Kalman, R.E.. *On the General Theory of Control Systems*. IFAC, International and Remote Control, Vol.1, pp. 481-492, 1960.
- [574] Kalman, R.E.. *A New Approach to Linear Filtering and Prediction Problems*. Trans. ASME, J.Basic Eng. Series 82D, pp. 35-45, March, 1960.
- [575] Kalman, R.E. and Bertram, J. E.. *Control Systems Analysis and Design Via the "Second Method" of Lyapunov*. Transactions of the ASME - Journal Basic Engineering, pp. 371-393, June, 1960.
- [576] Kalman, R.E.. *Mathematical Description of Linear Dynamical Systems*. Journal SIAM Control, Ser. A, Vol.1, N.2, pp. 152-192, 1963.
- [577] Kalman, R.E.. *Irreducible Realizations and the Degree of a Rational Matrix*. Journal Of The Society For Industrial and Applied Mathematics, Volume 13, Number 2, pp.520-544, June, 1965.
- [578] Kalman, R.E.. *Algebraic Structure Of Linear Dynamical Systems, I. The Module of  $\Sigma$* . Proceedings of The National Academy of Sciences Of The United States of America, Vol.54, Number 6, pp. 1503-1508, December, 1965.



- [579] Kalman, R.E.. *Algebraic Aspects of the Theory of Dynamical Systems*. Differential Equations and Dynamical Systems ( Proceedings of an International Symposium held at the University of Puerto Rico, Mayaguez, Puerto Rico, December 27-30, 1965. ) Edited by Jack k. Hale and Joseph P. LaSalle, pp.133-146, Academic Press, 1967.
- [580] Kalman, R.E.; Falb, P.L. and Arbib, M. A.. *Topics in Mathematical System Theory*. McGraw-Hill Book Company, 1969.
- [581] Kalman, R.E.. *On Minimal Partial Realizations of a Linear Input/Output Map*. (Note : A collection of papers in honor of E.A. Guillemin) Aspects of Network and System Theory, edited by R.E. Kalman and N. DeClaris, New York, Holt, Rinehart and Winston, pp. 385-407, 1971.
- [582] Kalman, R.E.. *Realization Theory of Linear Dynamical Systems*. (Note : A collection of papers in honor of E.A. Guillemin) International Seminar Course Trieste 1974 Proceedings, Control Theory and Topics in Functional Analysis; Vienna International Atomic Energy Agency, pag : 235 - 256, 1976.
- [583] Kalman, R.E.. *System Identification From Noisy Data*. International Symposium on Dynamical Systems, Florida , 135-164, February, 1981.
- [584] Kalman, R.E.. *Identification of Noisy Systems*. Russian Mathematical Surveys., Vol.40, No.4, 25-41, 1985.
- [585] Kamel, M.S.; Ma K.S. and Enright W.H.. *ODEXPRT - An Expert System to Select Numerical Solvers for Initial Value ODE Systems*. ACM - Transactions on Mathematical Software, Volume 19, No.1, 44-62, March, 1993.
- [586] Katayama, Tohru and Picci, Giorgio. *An Approach to Realization of Stochastic Systems With Exogenous Input*. 11th IFAC Symposium on System Identification, Vol.3, 1107-1112, July, 1977.
- [587] Katayama, Tohru and Sugimoto, Sueo (Editors). *Statistical Methods In Control And Signal Processing*. Marcel Dekker, Inc., August, 1997.
- [588] Katayama, Tohru; Omori, Shogo and Picci, Giorgio. *on A Comparison of Some Subspace Identification Methods* . Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, TA08-6, December, 1998.
- [589] Katayama, Tohru and Picci, Giorgio. *Realization of Stochastic Systems With Exogenous Inputs and Subspace Identification Methods*. Automatica, Vol.35, 1635-1652, 1999.
- [590] Kaufman, Linda. *Application of Dense Householder Transformations to a Sparse Matrix*. ACM Transactions On Mathematical Software, Vol. 5, No.4, 442-450, December, 1979.
- [591] Kaufman, Linda. *The Generalized Householder Transformation and Sparse Matrix*. Linear Algebra and Its Applications, Vol. 90, 221-234, May, 1987.
- [592] Kaufman, Linda. *An Algorithm for The Banded Symmetric Generalized Matrix Eigenvalue Problem*. SIAM - Journal on Matrix Analysis and Applications, Volume 14, No.2, 372-389, April, 1993.
- [593] Kaufman, Linda; Sylvester, Garret S. and Wright, Margaret H.. *Structured Linear Least-Squares Problems In System Identification and Separable Nonlinear Data Fitting*. SIAM J.Optimization, Vol.4, No.4, 847-871, November, 1994.
- [594] Kayalar Selahattin and Weinert, Howard L.. *Oblique Projections : Formulas, Algorithms, and Error Bounds*. Mathematics of Control, Signals, and Systems, Vol.2, 33-45, 1989.
- [595] Kemma, Andreas H.; Larimore, Wallace E.; Seborg, D.E. and Mellichamp, Duncan A.. *On-line Multivariable Identification and Control of Chemical Processes Using Canonical Variate Analysis* . Proceedings of The American Control Conference - Baltimore - Maryland, June 1994.
- [596] Kempf, James. *Numerical Software Tools in C*. Prentice-Hall, 1987.

- [597] Kempthorne, Oscar. *The Design and Analysis of Experiments*. Wiley & Sons, 1975.
- [598] Keuchel, U. and Schmid, Chr.. *Identification and Control With CADACS-PC*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 679-684, July, 1991.
- [599] Kirk, Donald E.. *Optimal Control Theory*. Prentice-Hall Inc., 1970.
- [600] Kitagawa, Genshiro and Gersch, Will. *A Smoothness Priors-State Space Modeling of Time Series With Trend and Seasonality*. JASA - Journal of the American Statistical Association, Vol. 79, No.386, 378-389, June, 1984.
- [601] Kitagawa, Genshiro. *Non-Gaussian State Space Modeling Of Time Series*. Proceedings of the 26th IEEE Conference On Decision and Control, Los Angeles, California, pp.1700-1705, December, 1987.
- [602] Klema, Virginia C. and Laub, Alan J.. *The Singular Value Decomposition: Its Computation and Some Applications*. IEEE Transactions on Automatic Control, Vol. AC-25, NO.2, 164-176, April, 1980.
- [603] Knockaret, Luc. *A recursive Algorithm for Linear System Identification*. IEEE - ASSP - Acoustics, Speech and Signal Processing, Volume ASSP-34, Number 3, 492-498, June, 1986.
- [604] Knudsen, Torben. *Consistency analysis of subspace identification methods based on a linear regression approach*. Automatica Volume: 37, Issue: 1, pp. 81-89 , January, 2001.
- [605] Knudsen, Torben. *Subspace Identification Methods*. Notas de Aula obtidas em "http://www.control.auc.dk/ tk/ ", 1998.
- [606] Köppen, B., Seliger, R., Frank, P.M.. *Self-Tuning Multivariable Tracking Control For An Induction Motor*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 697-702, July, 1991.
- [607] Kosebalaban, Figen and Cinar, Ali. *Integration of multivariate SPM and FDD by parity space technique for a food pasteurization process*. Computers and Chemical Engineering Volume: 25, Issue: 2-3, March 15, 2001, pp. 473 - 491, 2001.
- [608] Krause, James M. and Khargonekar, Pramod P.. *Parameter Identification in Presence of Non-parametric Dynamic Uncertainty*. Automatica, Vol.26, No.1, 113-123, January, 1990.
- [609] Kreyszig, Erwin. *Advanced Engineering Mathematics*. John Wiley & Sons, Inc, Seventh Edition, 1993.
- [610] Krim, H. and Proakis, J.G.. *Smoothed Eigenspace-Based Parameter Estimation*. Automatica, Volume 30, Number 1, 27-38, January, 1994.
- [611] Krishnamurthy, Vikram. *Estimation Algorithm For Quantized Dynamic Shock Error Models in Multiplicative Noise*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2294-2299, December, 1994.
- [612] Krishnan , S.; Barton, G.W. and Perkins, J.D.. *Robust Parameter Estimation in On-Line Optimization - Part 1. Methodology and Simulated Case Study*. Computers & Chemical Engineering, Volume 16, Number 6, 545-562, June, 1992.
- [613] Krishnan , S.; Barton, G.W. and Perkins, J.D.. *Robust Parameter Estimation in On-Line Optimization - Part 2. Application To an Industrial Process*. Computers & Chemical Engineering, Volume 17, Number 7, 663-669, July, 1993.
- [614] Krzyżak, A.; Sasiadek, J.Z. and Zhu, P.. *Identification of Dynamic Nonlinear Systems Using The Hermite Series Approach*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1731-1732, December, 1994.

- [615] Kshirsagar, Anant M.. *Multivariate analysis*. Markel Dekker Inc., New York, 1972.
- [616] Kuck, David and Muraoka, Yoichi. *Bounds on The Parallel Evaluation of Arithemic Expressions Using Associativity and Commutativity*. Acta Informatica, Vol.3, Fasc.3, 203-216, 1974.
- [617] Kudva, P. and Narendra, K.S.. *An Identification Procedure For Discrete Multivariable Systems*. IEEE Transactions on Automatic Control, Vol. AC-19, No.5, 549-552, October, 1974.
- [618] Kulhavý, R.. *Recursive Nonlinear Estimation Through Global Approximation Of Model*. Proceedings of 3rd European Control Conference, Rome, Italy, 1273-1278, September, 1995.
- [619] Kullback, Solomon. *An Application of Characteristic Functions To The Distribution Problem of Statistics*. The Annals of Mathematical Statistics, Vol.V, No.4, 263-307 - December - 1934.
- [620] Kullback, Solomon. *On Samples From A Multivariate Normal Population*. The Annals of Mathematical Statistics, Vol.VI, No.4, 202-213 - December - 1935.
- [621] Kullback, Solomon. *An Application of Information Theory To Multivariate Analysis*. The Annals of Mathematical Statistics, Vol.23, No.1, 88-146 - March - 1952.
- [622] Kung, S.. *A New Identification and Model Reduction Algorithm Via Singular Value Decompositions*. Proc. 12th. Asilomar Conference on Circuits, Systems and Computers, Pacific Grove - California - 705-714, November, 1978.
- [623] Kung, Sun-Yuan and Lin, David W.. *Optimal Hankel-Norm Model Reductions : Multivariable Systems*. IEEE Transactions on Automatic Control, Vol. AC-26, No.4, 832-852, August, 1981.
- [624] Kung, Sun-Yuan and Lin, David W.. *A State-Space Formulation for Optimal Hankel-Norm Approximations*. IEEE Transactions on Automatic Control, Vol. AC-26, No.4, 942-946, August, 1981.
- [625] Kung, Sun-Yuan; Arun, K.S. and Rao, D.V. Bhaskar. *State-Space and Singular-value Decomposition-Based Approximation Methods For The Harmonic Retrieval Problem*. Journal of the Optical Society of America, Vol.73, No.12, 1799-1811, December, 1983.
- [626] Kung, Sun-Yuan. *Multivariable and Multidimensional Systems: Analysis and Design*. Ph.D. Thesis - Stanford University - Departament of Electrical Engineering, June, 1977.
- [627] Kurdila, Andrew J.; Narcowich, Francis J. and Ward, Joseph D.. *Persistency of Excitation, Identification, and Radial Basis Functions*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2273-2278, December, 1994.
- [628] Kurpel', Nikolai Stepanovich. *Projection-iterative methods for solutions of operator equations*. Translations of Mathematical Monographs, Vol. 46, American Mathematical Society, 1976.
- [629] Kurzhanski, A.B.. *State Estimation - A Common Framework For The Set - Membership (Bounding) and the H-inf Techniques*. Proceedings of 3rd European Control Conference, Rome, Italy, 955-960, September, 1995.
- [630] Kwakernaak, H.. *Maximum Likelihood Parameter Estimation for Linear Systems with Singular Observations*. IEEE Transactions on Automatic Control, Vol. AC-24, No.3, 496-498, June, 1979.
- [631] Kwan, C.M.. *Some News Results in Population Control*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4110-4111, December, 1994.
- [632] Laban, M. and Mulder, J.A.. *On-Line Identification Of Aircraft Aerodynamic Model Parameters*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 556-561, July, 1991.
- [633] Lages, N. A. de Castilho e Nogueira, J.M.Silva. *Introdução Aos Sistemas Distribuídos*. Editora da UNICAMP, 1986.

- [634] Lai, D. and Chen, G.. *Dynamical Systems Identification from Time-Series Data: A Hankel Matrix Approach* . Mathematical and Computer Modelling Volume: 24, Issue: 3, pp. 1-10, August, 1996.
- [635] Lancaster, H.O.. *Kolmogorov's Remark on the Hotelling Canonical Correlations*. Biometrika, Volume 53, number 3 and 4, pp.585-588 , 1966.
- [636] Landau, Ioan Doré. *System Identification and Control Design*. Prentice-Hall Inc., 1990.
- [637] Landau, Ioan Doré and Rolland, François. *An Approach For Closed Loop System Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4164-4169, December, 1994.
- [638] Larimore, Wallace E.. *System Identification Of Feedback and "Causality"Structure Using Canonical Variate Analysis*. 11th IFAC Symposium on System Identification, Vol.3, 1101-1106, July, 1977.
- [639] Larimore, Wallace E. and Mehra, Raman K.. *The Problem of Overfitting Data*. Byte , 167-180 , October , 1985.
- [640] Larimore, Wallace E.. *Identification of Nonlinear Systems Using Canonical Variate Analysis*. Proceedings of the 26th IEEE Conference On Decision and Control, Los Angeles, California, pp.1694-1699, December, 1987.
- [641] Larimore, Wallace E.. *Canonical Variate Analysis in Identification, Filtering and Adaptive Control*. Proceedings of the 29th Conference on Decision and Control - Hawaii - 596-604, December, 1990.
- [642] Larimore, Wallace E.. *Order-Recursive Factorization of The Pseudoinverse of a Covariance Matrix*. IEEE Transactions on Automatic Control, Vol. 35, No 12, pp. 1299-1303, December, 1990.
- [643] Larimore, Wallace E.. *Identification and Filtering of Nonlinear Systems Using Canonical Variate Analysis*. Nonlinear Modeling and Forecasting , SFI Studies in The Sciences of Complexity, Proc.Vol. XII, Eds M. Casdagli & S. Eubank, Addison-Wesley, pp.283-303, 1992.
- [644] Larimore, Wallace E.. *Accuracy Confidence Bands Including The Bias Of Model Under-Fitting*. Proceedings of The American Control Conference - California ,pp. 1995-1999, June , 1993.
- [645] Larimore, Wallace E.. *Automated and Optimal System Identification by Canonical Variables*. Proceedings of the American Control Conference - Baltimore , Maryland - June 1994.
- [646] Larimore, Wallace E.. *The Optimality of Canonical Variate Identification By Example*. Proceedings 10th IFAC Symposium on System Identification - Copenhagen, July , 1994.
- [647] Larimore, Wallace E.. *A Unified View of Reduced Rank Multivariate Prediction Using a Generalized Singular Value Decomposition*. Report - contract F33615-82-C-3600 Air Force Wright Aeronautical Laboratories 22 pages. (submitted for publication).
- [648] Larimore, Wallace E.. *Automated multivariable system identification and industrial applications* . Proceedings of the American Control Conference, pp. 1148 -1162 vol.2, 1999.
- [649] Larminat, Ph.de and Doncarli, C. *A Real-Time Generalized Least Squares Estimation Method for Identification of Linear System*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 379-384, July, 1977.
- [650] Lastman, G.J.. *Identification of Linear Multivariable Time-Invariant Systems From Samples of Input-Output Data*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1191-1195, July, 1991.
- [651] Lathauwer, Lieven De. *Signal Processing Based on Multilinear Algebra*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, September, 1997.
- [652] Laub, Alan J.. *Computation of "Balancing " Transformations*. Proceedings of Joint Control Conference - JACC - V.1 , paper FA8-E , 1980.

- [653] Laub, Alan J.. *A Schur Method for Solving Algebraic Riccati Equations*. IEEE Transactions on Automatic Control - Vol. AC-24, No.6, pp. 913 - 921, December, 1979.
- [654] Lawson, Charles L. and Hanson, Richard J.. *Solving Least Squares Problems*. Prentice-Hall - Series in Automatic Computation, 1974.
- [655] Lee, Jang G. and Lee, Dal H.. *Comparison of The Extended Kalman Filter and Maximum Likelihood Methods for Parameter Identification of an Autonomously Piloted Vehicle*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 674-678, July, 1991.
- [656] Lee, Lawton Hubert. *Identification and Robust Control of Linear Parameter-Varying Systems*. Ph.D. thesis Engineering-Mechanical Engineering - University of California at Berkeley , 1997.
- [657] Lemmerling, Philippe. *Structured Total Least Squares: Analysis, Algorithms and Applications*. PhD. Thesis - Katholieke Universiteit Leuven - May 1999.
- [658] Lemmerling, Philippe; Mastronardi, Nicola and Van Huffel, Sabine. *Fast Algorithm for Solving The Hankel/ Toeplitz Structured Total Least Squares Problem*. Numerical Algorithms, 23, pp. 371-392, 2000.
- [659] Lemmerling, Philippe and Van Huffel, Sabine. *A projection interpretation of the structured total least squares approach*. Internal Report 98-123, pp. 100-112, ESAT-SISTA, K.U.Leuven (Leuven, Belgium 1998.
- [660] Leonides, C. T.. *Control And Dynamic Systems - Volume 55 - Digital and Numeric Techniques and Their Applications In Control Systems* . Academic Press, Inc., 1993.
- [661] Leonhardt, Steffen. *A Parallel Algorithm For Process Identification*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 521-526, July, 1991.
- [662] Levy, Bernard C.; Benveniste, Albert and Nikoukhah, Ramine. *High-level Primitives For Linear Estimation*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3906-3911, December, 1994.
- [663] Lideng Pan and Huaijing, Du. *Structure and Parameters Identification of Multivariable CARMA Models Using SSM-Technique*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1003-1007, August, 1988.
- [664] Lindquist, Anders and Pavon, Michele. *Markovian Representation of Discrete-Time Stationary Stochastic Vector Processes*. Proceedings of the 20th IEEE Conference on Decision and Control, CDC, pp.1345-1356, 1981.
- [665] Lindquist, Anders and Picci, Giorgio. *Realization Theory For Multivariate Stationary Gaussian Processes*. SIAM Journal on Control and Optimization, V.23, pp. 809 - 857, 1985 .
- [666] Lindquist, Anders and Picci, Giorgio. *On noncausal estimation, stochastic realization, and the Riccati inequality*. Proceedings of the 28th IEEE Conference on Decision and Control, pp. 1207 - 1209, 1989 .
- [667] Lindquist, Anders and Picci, Giorgio. *Canonical Correlation Analysis, Approximate Covariance Extension and Identification of Stationary Time Series*. Automatica, Volume 32, No.5, 709-733, May, 1996.
- [668] Ling, Fuyun. *Order-Recursive Least-Squares Adaptive Algorithms - A Unified Framework*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1409-1414, August, 1988.
- [669] Linkens, D.A., Asbury, A.J., Rimmer, S.J. and Menad, M.. *Identification and Control of Muscle-Relaxant Anaesthesia*. IEE Proceedings, Vol.129, Pt.D, No. 4, 136-141, July, 1982.

- [670] Liu, D.Y.. *The Perfect Recursive Least-Squares Filtering Algorithms For Linear Systems*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1299-1304, July, 1991.
- [671] Liu, Ketao. *Identification of Multi-Input and Multi-Output Systems by Observability Range Space Extraction*. Proceedings of the 31st Conference on Decision and Control Arizona, 915-919, December, 1992.
- [672] Liu, Ketao and Skelton, Robert E.. *Q-Markov Covariance Equivalent Realization and its Application to Flexible Structure Identification*. Journal of Guidance, Control and Dynamics, Vol.16, No.2, 308-319, March-April, 1993.
- [673] Liu, K.J.Ray and Hsieh, S.F.. *Fast Orthogonalization Algorithm and Parallel Architecture for Ar Spectral Estimation Based on Forward-Backward Linear Prediction*. IEEE Transactions on Signal Processing, Vol. 41, No.3, 1453-1458, March, 1993.
- [674] Liu, R. and Suen, L.-C.. *Minimal Dimension Realization and Identifiability of Input-Output Sequences*. IEEE Transactions on Automatic Control, Vol. AC-22, No.2, 227-232, April, 1977.
- [675] Ljung, Lennart; Söderström, Torsten and Gustavsson, Ivar. *Counterexamples to General Convergence of A Commonly Used Recursive Identification Method*. IEEE Transactions on Automatic Control, Vol. AC-20, No.5, 643-652, October, 1975.
- [676] Ljung, Lennart. *Consistency of the Least-Squares Identification Method*. IEEE Transactions on Automatic Control, Vol. AC-21, No.5, 779-781, October, 1976.
- [677] Ljung, Lennart. *Fast Algorithms for Recursive Estimation and Identification*. Numerical Techniques for Stochastic Systems edited by Francesco Archetti and Marco Cugiani - 47-57 - North-Holland Publishing Company - 1980.
- [678] Ljung, L. and Söderström, T.. *Theory and Practice of Recursive Identification*. The MIT Press, 1983.
- [679] Ljung, L.. *System Identification : Theory for the User* Prentice Hall, 1987.
- [680] Ljung, Lennart. *Issues in System Identification*. IEEE Control Systems Magazine, Vol.11, No.1, 25-29, January, 1991.
- [681] Ljung, Lennart. *Identifying State-Space Models With Mathwork's System Identification Toolbox*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1582-1584, July, 1991.
- [682] Ljung, Lennart and Hjalmarsson, Håkan. *System Identification Through The Eyes Of Model Validation*. Proceedings of 3rd European Control Conference, Rome , Italy, 949-954, September, 1995.
- [683] Ljung, L., Sjöberg and McKelvey, T.. *On The Use Of Regularization In System Identification*. Seminário interno - Pages 1-15, August , 10, 1992. (obtido através : [ljung@isy.liu.se](mailto:ljung@isy.liu.se))
- [684] Ljung, L., Sjöberg and McKelvey, T.. *Subspace Identification from Closed Loop Data*. Signal Processing, Vol. 52, pp. 209-215, 1996.
- [685] Ljung, L., Sjöberg and McKelvey, T.. *A Least Squares Interpretation of Subspace Methods for System Identification*. Proc. 35th IEEE Conference on Decision and Control, Kobe, Japan, 335-342, December, 1996.
- [686] Ljung, L., Sjöberg and McKelvey, T.. *Interpretation of Subspace Methods: Consistency Analysis*. IFAC Symposium on System Identification, Fukuoka, Japan, 1997.
- [687] Loan, Charles F.Van. *Generalizing The Singular Value Decomposition*. SIAM Journal On Numerical Analysis, Vol. 13, No.1, 76-83, March, 1976.

- [688] Loomba, Jaime Terceiro. *Estimation of DynEconometric Models With Erros in Variables*. Lecture Notes in Economics and Mathematical Systems, Vol. 339, Springer-Verlag, 1990.
- [689] Lopez, J.F.; Oliveira, G.H.C.; Amaral, W.C. and Latre, L.G.. *Multivariable Constrained Predictive Control Methods Applied To The Shell Benckmark Problem : A Comparison*. Proceedings of 3rd European Control Conference, Rome , Italy, 3259-3264, September, 1995.
- [690] Lopes-Toledo, Alejandro A. and Athans, Michael. *Optimal Policies For Identification of Stochastic Linear Systems*. IEEE Transactions on Automatic Control, Vol. AC-20, No.6, 754-765, December, 1975.
- [691] Lorch, Edgar R.. *On A Calculus Of Operators In Reflexive Vector Spaces*. Transactions of the American Mathematical Society, Vol.45, pp.217-234, January to June, 1939.
- [692] Lovera, Marco. *Subspace Identification Methods : Theory and Applications*. Ph.D. Thesis - Dipartimento di Elettronica e Informazione - Politecnico Di Milano, 1998.
- [693] Lovera, Marco and Previdi, Fabio. *Identification of linear models for the dynamics of a photode-tector*. Control Engineering Practice, Volume: 8, pp. 1149-1158 , Issue: 10, October, 2000.
- [694] Ludlage, Jobert; Oudbier, Roel and Backx, Ton. *Use of A Priori Knowledge of Static Gains in MIMO Process Identification*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1196-1201, July, 1991.
- [695] Luenberger, D.G.. *Observers for Multivariable Systems*. IEEE Transactions on Automatic Control, Vol. AC-11, No.2, 190-197, April, 1966.
- [696] Luenberger, David G.. *Canonical Forms for Linear Multivariable Systems*. IEEE Transactions on Automatic Control, Vol. AC-12, No.3, 290-293, June, 1967.
- [697] Luenberger, David G.. *Optimization By Vector Space Methods*. John Wiley & Sons, 1969.
- [698] Luk, Franklin T. and Qiao, Sanzheng. *A new matrix decomposition for signal processing*. Automatica, Vol. 30, No.1, 39-43, January, 1994.
- [699] Lund, Eivind J.; Balchen, Jens G. and Foss, Bjarne A.. *Multiple Model Estimation With Inter-Residual Distance Feedback*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1512-1517, July, 1991.
- [700] Lundström, Eva. *Singular Value Computations For Toeplitz Matrices and Subspace Tracking*. Ph.D. Thesis; Department of Mathematics; Linköpings Universitet, Sweden, 1998.
- [701] Lus, Hilmi; Betti, Raimondo and Longman, Richard W.. *Identification of Linear Structural Systems Using Earthquake Induced Vibration Data*. M.ASCE Department of Civil Engineering and Engineering Mechanics, Columbia University, 1998. (hilmi@civil.columbia.edu),
- [702] Lütkepohl, Helmut. *Introduction to Multiple Time Series Analysis*. Spinger-Verlag, 1993.
- [703] Lynch, Edmond B. and Ramirez, W. Fred.. *Real-Time Time-Optimal Control of a Stirred-Tank Reactor Using Kalman Filtering for State Estimation*. AIChE Journal - Volume 21, Number 4 - 799-804, July, 1975.
- [704] Maciejowski, J.M. and Ober R.J.. *Balanced Parametrizations and Canonical Forms For Systems Indentification*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 989-996, August, 1988.
- [705] Maciejowski, J.M.. *Obtaining Initial Models from Real Data: Aproximate Realization and Subspace Methods*. NATO Advanced Study Institute - From Identification to Learning, Como, Italy, 15 pages - 22 August - 2 September 1994.
- [706] Maciejowski, J.M.. *Application to Input-Output Modelling and Time Series Analysis*. NATO Advanced Study Institute - From Identification to Learning, Como, Italy, 8 pages - 22 August - 2 September 1994.

- [707] Maciejowski, J.M.. *MIMO Parameter Estimation Using Balanced Realization*. NATO Advanced Study Institute - From Identification to Learning, Como, Italy, 14 pages - 22 August - 2 September 1994.
- [708] Maciejowski, J.M. and Chou, C.T.. *Applications of Estimation Using Balanced Parametrisations*. European Control Conference - ECC 95 - pag.1348-1353 - Roma - Italy - September, 1995.
- [709] Maciejowski, J.M.. *Guaranteed Stability With Subspace Methods*. Systems & Control Letters, Volume 26, Number 2 - pag.153-156 - September, 1995.
- [710] MacInnes, Craig S. and Vaccaro, Richard J.. *Tracking directions-of-arrival with invariant subspace updating*. Signal Processing Volume: 50, Issue: 1-2, pp. 137-150, April, 1996.
- [711] Maguire, L.P. and Irwin, G.W.. *Transputer Implementation of Kalman Filters*. IEE Proceedings, Vol.138, Pt.D, No.4, 355-362, July, 1991.
- [712] Maindonald, J.H.. *Statistical Computation*. John Wiley & Sons, 1984.
- [713] Maine, Richard E. and Iliff, Kenneth W.. *Formulation and Implementation of A Practical Algorithm For Parameter Estimation With Process and Measurement Noise*. SIAM Journal Applied Mathematics, Volume 41, Number 3, 558-579, December, 1981.
- [714] Maine, Richard E. and Iliff, Kenneth W.. *Identification of Dynamic Systems. Theory and Formulation*. NASA Reference Publication 1138, February, 1985.
- [715] Majdoul, A., Najim, M., Ouadou, M., Youlal, H., Zyoute, M. and Bennis, A.. *Identification Of An Industrial Fluidized Bed Phosphate Calcination Reactor*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1497-1502, August, 1988.
- [716] Mäkilä, P. M. ; Partington, J. R. and Gustafsson, T. K.. *Worst-case control-relevant identification*. Automatica, Volume 31, No.12, 1799-1819, December, 1995.
- [717] Makridakis , Spyros and Whellwright, Steven C.. *Interactive Forecasting*. John Wiley & Sons, 1984.
- [718] Makridakis , Spyros and Whellwright, Steven C.. *Forecasting - Methods and Applications*. Holden-Day, Inc. , 1978.
- [719] Mandel, John. *Use of the Singular Value Decomposition in Regression Analysis*. The American Statistician, Vol. 36, No.1, 15-24, February, 1982.
- [720] Markusson, Ola and Bohin, Torsten. *Identification of a Nonlinear EEG Generating Model*. IEEE Workshop On Nonlinear Signal and Image Processing, June, 1997.
- [721] Maryak; John L., Spall, James C. and Silberman, Geoffrey. *Uncertainties for recursive estimators in nonlinear state-space models, with applications to epidemiology*. Automatica, Volume 31, No.12, 1889-1892, December, 1995.
- [722] Maryak, John L.; Hunter, Lawrence W. and Favin, Stanley. *Automated System Monitoring and Diagnosis via Singular Value Decomposition*. Automatica, Volume 33, No.11, 2059-2063, November, 1997.
- [723] Masalsky, Ya.S., Koutcheriaev, B.V., Blinkov, A.E. and Genkin, A.L.. *The Criterial Identifier Of the Human-Operator's Control Strategy*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 420-424, July, 1991.
- [724] Mastronardi N.; Lemmerling P. and Van Huffel S.. *Fast structured Total Least Squares algorithm for solving the basic deconvolution problem*. Internal Report 98-89, pp. 1-23, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1998. ( <http://www.esat.kuleuven.ac.be/sista/members/lemmerli/> )



- [725] Mastronardi N.; Lemmerling P. and Van Huffel S.. *The Structured Total Least Squares problem*. Internal Report 99-83, pp. 1-18, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1999. (<http://www.esat.kuleuven.ac.be/sista/members/lemmerli/> )
- [726] Mastronardi N.; Kressner D.; Sima V.; Van Dooren P. and Van Huffel S.. *A fast algorithm for subspace state-space system identification via exploitation of the displacement structure*. Internal Report 00-08, pp. 1-15, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2000.
- [727] Mathelin, Michel de and Bodson, Marc. *Canonical vs Pseudo-Canonical Forms For the Structural and Parametric Identification of Multivariable Systems*. Proceedings of the First European Control Conference - Grenoble, France, 1283-1287, July, 1991.
- [728] Mathieu, B.; Oustaloup, A. and Levron, F.. *Transfer Function Parameter Estimation By Interpolation in The Frequency Domain*. Proceedings of 3rd European Control Conference, Rome , Italy, 2131-2136, September, 1995.
- [729] Matko, Drago and Zupančič, Borut. *Interactive Graphics In CACSD Packages*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 437-441, July, 1991.
- [730] Maybeck, Peter S.. *Stochastic models, estimation, and control*. Academic Press, 1982.
- [731] McElveen, J.Keith; Lee, Koucherng, Roger Lee and Bennett, John E.. *Identification of Multivariable Linear Systems from Input/Output Measurements*. IEEE Transactions on Industrial Electronics , Vol.39, No.3, 189-193, 1992.
- [732] McGinnie, Bernard Paul. *A Balanced View of System Identification*. University Engineering Department - Cambridge, PhD Thesis. e-mail do orientador (Dr. Jan Maciejowski): [jmm@eng.cam.ac.uk](mailto:jmm@eng.cam.ac.uk) April - 18, 1994.
- [733] McKelvey, T.. *System Identification Using Overparametrized State- Space Models*. Dept of EE Linköping University - Linköping, Sweden - Report LiTH-ISY-R-1454, 1993 .
- [734] McKelvey, T.. *On State-Space Models in System Identification*. Licentiate Thesis , Number 447; - Department of Electrical Engineering - Linköping University, S-581 83, Linköping , Sweden, May, 1994.
- [735] McKelvey, T.. *A Combined State-Space Identification Algorithm Applied To Data From A Modal Analysis Experiment On A Separation System*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2286-2287, December, 1994.
- [736] McKelvey, T. and Akçay, Hüseyin. *An Efficient Frequency Domain State-Space Identification Algorithm : Robustness and Stochastic Analysis*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3348-3353, December, 1994.
- [737] McKelvey, T.. *SSID - A MATLAB Toolbox for Multivariable State- Space Model Identification* . obtido através de : <http://rt.isy.liu.se>, 1-6, December - 1994.
- [738] McKelvey, T.. *Model Validation by Geometric Arguments*. Dept of EE Linköping University - Linköping, Sweden - Report LiTH-ISY-R-1773, 1995 .
- [739] McKelvey, T.. *Identification of State-Space Models From Time and Frequency Data*. PhD Thesis - Department of Electrical Engineering - Linköping University, S-581 83, Linköping , Sweden, 1995.
- [740] McKelvey, T.. *Model Validation by Geometric Arguments*. Proceedings of 3rd European Control Conference, Rome , Italy, 1324-1329, September, 1995.
- [741] McKelvey, T.; Ljung, L. and Akçay, H.. *Identification of Finite Dimensional Systems From Frequency Response Data*. Proceedings of 3rd European Control Conference, Rome, Italy, 2106-2111, September, 1995.

- [742] McKelvey, T. and Akçay, H.. *System Identification With Periodic Excitation Signals: A Subspace Based Algorithm*. Proceedings of 3rd European Control Conference, Rome , Italy, 423-428, September, 1995.
- [743] Meditch, J.S.. *Stochastic Optimal Linear Estimation and Control*. McGraw-Hill Book Company, 1969.
- [744] Mehra, Raman K.. *On-Line Identification of Linear Dynamic Systems With Applications to Kalman Filtering*. IEEE Transactions on Automatic Control, Vol. AC-16, No.1, 12-21, February, 1971.
- [745] Mehra, Raman K.. *Optimal Inputs for Linear System Identification*. IEEE Transactions on Automatic Control, Vol. AC-19, No.3, 192-200, June, 1974.
- [746] Mehra, Raman K. and Lainiotis, Dimitri G.. *System Identification : Advances and Case Studies*. Academic Press, 1976.
- [747] Mehra, Raman K.. *Kalman Filters and Theirs Applications to Forecasting*. Tims Studies In The Management Sciences, V.12, pages. 75-94, 1979.
- [748] Mehra, Raman K.. *Identification of State Space Models Using A Stochastic Realization Algorithm With Application To Modal Identification of Flexible Aerospace Structures*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 550-555, July, 1991.
- [749] Mehta, Amish; Kaufman, Howard and Ravi, R.. *Turbine System Identification : Experimental Results*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3593-3595, December, 1994.
- [750] Mendel, Jerry M.. *1- Multistage Least-Squares Parameter Estimators. 2 - On the Identification of State-Derivate-Coupled Systems*. IEEE Transactions on Automatic Control, Vol. AC-20, No.6, 775-785, December, 1975.
- [751] Meyer, Carl D.. *Matrix Analysis and Applied Linear Algebra*. SIAM Books, June, 2000.
- [752] Michaletzky, György and Ferrante, Augusto. *Splitting Subspaces and Acausal Spectral Factors*. Journal of Mathematical Systems, Estimation, and Control, Vol 5, No. 3, pp. 1-26, 1995.
- [753] Milanese M. and Elia, N.. *Linear Optimal Algorithms in Set Membership Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1746-1750, December, 1994.
- [754] Milanese M.. *Properties of Least Squares Estimates in Set Membership Identification*. Automatica, Volume 31, Number 2, 327-332, February, 1995.
- [755] Militký, Jiří . *Some graphical aids For Univariate Exploratory Data Analysis* . Analytica Chimica Acta - Volume 277, No.2, 215-221, May, 1993.
- [756] Milne, R.D.. *An Oblique Matrix Pseudoinverse*. Siam Journal on Applied Mathematics, Vol.16, No.5, 931-944, September, 1968.
- [757] Milton, Roy C. and Nelder, John A.. *Statistical Computation*. Academic Press, Inc., 1969.
- [758] Miyano, T.; Kimoto, S.; Shibuta, H.; Nakashima, K.; Ikenaga, Y.; and Aihara, K. 108 *Time series analysis and prediction on complex dynamical behavior observed in a blast furnace*. Physica D Volume: 135, Issue: 3-4, pp. 305-330, January 15, 2000.
- [759] Modi, Jagdish J.. *Parallel Algorithms and Matrix Computation*. Clarendon Press - Oxford , 1988.
- [760] Moitra, Abha and Iyengar, Sitharama. *A Maximally Parallel Balancing Algorithm for Obtaining Complete Balanced Binary Trees*. IEEE Transactions on Computers, Vol.C-34, No.6, 563-565, June, 1985.

- [761] Moler, Cleve; Liule, John ; Bangert, Steve and Kleiman, Steve. *PC-MATLAB for MS-DOS Personal Computers*. The MathWorks, Inc. - 1986
- [762] Morf, Martin and Kailath, Thomas. *Square-Root Algorithms For Least-Squares Estimation*. IEEE Transactions on Automatic Control, Vol. AC-20, No.4, 487- 497, August, 1975.
- [763] Moonen, Marc; Moor, Bart De; Vandenberghe, Lieven and Vandewalle, Joos. *On-and Off-Line Identification Of Linear State-Space Models*. INT.J. Control, Vol.49, No.1, 219-232, January 1989.
- [764] Moonen, Marc and Vandewalle, Joos. *QSVD Approach To On- and Off-Line State-Space Identification*. INT.J. Control, Vol.51, No.5, 1133-1146, May 1990.
- [765] Moonen, Marc. *Jacobi-type updating algorithms for signal processing, systems identifications and control*. Doctoral Dissertation, Department of Electrical Engineering, Katholieke Universiteit Leuven, Belgium, 1990.
- [766] Moonen, Marc and Moor, Bart De. *Comments on State-space Model Identification With Data Correlation*. International Journal of Control, Vol.55, No.1, 257-259, January, 1992.
- [767] Moonen, Marc; Moor, Bart De; and Ramos, José. *A Subspace Identification Algorithm For Descriptor Systems*. Systems & Control Letters, Volume 19, No.1, 47-52, July, 1992.
- [768] Moonen, Marc; Dooren, Paul Van and Vandewalle, Joos. *A Singular Value Decomposition Updating Algorithm For Subspace Tracking*. SIAM - Journal on Matrix Analysis and Applications, Volume 13, No.4, 1015-1038, October, 1992.
- [769] Moonen M.; Van Overschee P. and De Moor B.. *A subspace algorithm for combined stochastic-deterministic state space identification*. Internal Report 92-10, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1992.
- [770] Moonen, Marc and Ramos, José. *A Subspace Algorithm For Balanced State Space System Identification*. IEEE Transactions on Automatic Control, Vol. 38 , pp 1727-1729, 1993.
- [771] Moonen, Marc; Van Dooren and Vandewalle, J.. *An SVD updating algorithm for subspace tracking*. SIAM Journal on Matrix Analysis and Applications, Vol. 13, No. 4, pp 1015-1038, 1992.
- [772] Moonen, Marc and De Moor, B.. *Comments on 'State-space model identification with data correlation'*. International Journal of Control, Vol 55, No. 1, pp 257-259, 1992.
- [773] Moonen, Marc; De Moor, B.; and J. Ramos and S. Tan.. *A subspace identification algorithm for descriptor systems*. Systems & Control Letters, Vol. 19, pp 47-52, 1992.
- [774] Moonen, Marc; Van Overschee, Peter and De Moor, Bart. *A Subspace Algorithm for Combined Stochastic-Deterministic State Space Identification*. ESAT - SISTA REPORT 1992-10, p. 1-13, 1992.
- [775] Moonen, Marc. *Algorithms and Architectures for Recursive Total Least Squares Estimation*. Algorithms and Parallel VLSI Architectures II -P. Quinton and Y. Robert (Editors), pp. 39 - 46, Elsevier Science, 1992.
- [776] Moonen, Marc and McWhirter, John G.. *A Systolic Array For Recursive Least Squares By Inverse Updating*. Electronics Letters, Vol. 29, No. 13, pp 1217-1218, April-1993.
- [777] Moonen, Marc; Dooren, Paul Van and Vandewalle, Joos. *A Systolic Array For SVD updating*. ESAT - SISTA REPORT 1991-18 e-mail: moonen@esat.kuleuven.ac.be ou SIAM - Journal on Matrix Analysis and Applications, Volume 14, No.2, 353-371, April, 1993.
- [778] Moonen, Marc and Vandewalle, Joos. *A Systolic Array For Recursive Least Squares Computations*. IEEE Transactions on Signal Processing, Vol. 41 , No. 2, pp 906-912, 1993.

- [779] Moonen, Marc. *Systolic Algorithms For Recursive Total Least Squares Parameter Estimation and Mixed RLS/RTLS Problems*. International Journal of High Speed Electronics, special issue on 'Massively Parallel Computing', Vol. 4, No. 1, pp 55-68, 1993.
- [780] Moonen, Marc. *Implementing The Square-Root Information Kalman Filter On A Jacobi-Type Systolic Array*. Journal of VLSI Signal Processing, Vol. 8 (1994), No. 3, pp 1-9, 1994.
- [781] Moonen, Marc. *A Systolic Array For Recursive Least Squares Computations - Part II : Mapping directionally weighted RLS on an SVD updating array*. IEEE Transactions on Signal Processing, Vol. 44, No. 8, pp 2117-2121, August, 1996.
- [782] Moonen, Marc. *Introduction to Adaptive Signal Processing*. obtida em <ftp://wgs.esat.kuleuven.ac.be/pub/sista/moonen/reports/> pag. 1-176, edição no prelo, 1998.
- [783] Moore, Bruce C.. *Principal Component Analysis in Linear Systems : Controllability, Observability, and Model Reduction*. IEEE Transactions on Automatic Control, Vol.AC-26, No.1, 17-32, February, 1981.
- [784] Moore, John B.. *Application of Riccati Equations in Systems Engineering*. Electrical Engineering Transactions, 29-34, March, 1969.
- [785] Moose, Richard L.. *An Adaptive State Estimation Solution To The Maneuvering Target Problem*. IEEE Transactions on Automatic Control, Vol. AC-20, No.3, 359-362, June, 1975.
- [786] Morari, Manfred and Lee, Jay H. *Model predictive control: past, present and future*. Computers & Chemical Engineering Volume: 23, Issue: 4-5, May 1, pp. 667 - 682, 1999.
- [787] Morelli, Eugene A.. *Low Order Equivalent System Identification For The Tu-144LL Supersonic Aircraft*. NASA - AIAA - American Institute of Aeronautics and Astronautics, 2000-3902, 2000
- [788] Morettin, Pedro A. e Toloi, Clélia M.. *Séries Temporais*. Atual Editora, 1986.
- [789] Morimoto, Tetsuo; Fukuyama, Toshio; Yi, Youlin and Hashimoto, Yasushi. *Identification of Physiological Dynamics In Hydroponics*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1736-1741, August, 1988.
- [790] Morse, A.S.. *Estructural Invariants Of Linear Multivariable Systems*. Siam Journal Control, Vol. 11, No.3, 446-465, August, 1973.
- [791] Moustafa, Kamal A.F., Gaafar, M.Y., Yazid, T.Abu-El. *Modelling and Identification of a Stochastic Water-Quality System Using Actual Data*. IEE Proceedings, Vol.133, Pt.D, No. 4, 159-164, July, 1986.
- [792] Muller, David E. and Preparata, Franco P.. *Restructuring of Arithmetic Expressions For Parallel Evaluation*. Journal of the Association for Computing Machinery - JACM, Vol.23, No. 3, 534-543, July, 1976.
- [793] Muller, Jean-Michel. *On-Line Computing : A Survey and Some New Results*. Algorithms and Parallel VLSI Architectures II -P. Quinton and Y. Robert (Editors), pp. 261 - 272, Elsevier Science, 1992.
- [794] Muller, Keith E.. *Understanding Canonical Correlation Through The General Linear Model and Principal Components*. The American Statistician, Vol.36, No. 4, 342-367, November, 1982.
- [795] Müller, J., Conrad, T., Martinov, M. and Tesic, M.. *Computer Aided Control For Solar Drying*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 480-485, July, 1991.
- [796] Murakami, Kunio; Kakuta, Takeo; Onai, Rikio and Ito, Noriyoshi. *Research on Parallel machine Architecture for Fifth-Generation Computer Systems*. Computer, Vol.18, No.6, 76-92, June, 1985.

- [797] Murari, Carlos Alberto Favarin. *Estimação de Estado e Análise de Observabilidade em Redes de Energia Elétrica Utilizando Transformação Ortogonal*. Tese de Doutorado, Faculdade de Engenharia de Campinas, UNICAMP, 1986.
- [798] Murray, F.J.. *On Complementary Manifolds and Projections in Spaces  $L_p$  and  $l_p^\dagger$* . Transactions of the American Mathematical Society, Vol.41, pp.138-152, January to June, 1937.
- [799] Nagy, Peter A.J. and Ljung, Lennart. *An Intelligent Tool for System Identification* IEEE Control Systems Society, Workshop on Computer-Aided Control System Design (CACSD), 58-63, December, 1989.
- [800] Nagy, Peter A.J. and Ljung, Lennart. *Computer-Aided Model Structure Selection*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 918-923, July, 1991.
- [801] Nakajima, S.; Lindberg, C. F. and Carlsson, B.. *On-line Estimation of the Respiration Rate and the Oxygen Transfer Function Using an Extended Kalman Filter*. Signals, Sensors and Systems, KTH, Stockholm, Technical Reports in Reglerteknik -IR-S3-REG-9613, 1996.
- [802] Nassiri-Toussi, Karim and Ren, Wei. *Indirect Adaptive Pole-Placement Control of MIMO Stochastic Systems : Self-Tuning Results*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 265-270, December, 1994.
- [803] Negiz, Antoine and Çinar, Ali. *PLS, balanced, and canonical variate realization techniques for identifying VARMA models in state space*. Chemometrics and Intelligent Laboratory Systems, Vol. 38, No.2, pp. 209-221, 1997.
- [804] Nehari, Zeev. *On Bounded Bilinear Forms*. Annals of Mathematics, Vol.65, No.01, pp.153-162, January 1957.
- [805] Nerode, A.. *Linear Automaton Transformations*. Proceedings of the American Mathematical Society, Volume 9, Number 4, 541-544, August, 1958.
- [806] Neumann, D., Isermann, R., Nold, S.. *Comparison Of Some Parameter Estimation Methods For Continuous-Time Models*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1171-1176, August, 1988.
- [807] Neyran, B. Thomasset D. and Dufour, J.. *Direct Identification of a Variable Linear System Using a State Affine Model*. IEE Proceedings, Vol.134, Pt.D, No. 2, 89-94, March, 1987.
- [808] Ng, T.S.. *On-Line Model Order Determination and Parameter Estimation*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1049-1053, August, 1988.
- [809] Ng, T.S.; Goodwin, G.C. and Anderson, B.D.O.. *Identifiability of MIMO Linear Dynamic Systems Operating in Closed Loop*. Automatica, Vol.13, No.5, 477-485, September, 1977.
- [810] Ng, C.N. and Young, P.C. and Wang, C.. *Recursive Identification, Estimation and Forecasting of Multivariate Time-Series*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 860-865, August, 1988.
- [811] Niedźwiecki, Maciej. *Identification of Time-varying Systems With Abrupt Parameter Changes*. Automatica, Volume 30, No.3, 447-459, March, 1994.
- [812] Nilsson, Johan. *Analysis and Design of Real-Time Systems With Random Delays*. Licentiate Thesis, Department of Automatic Control, Lund Institute of Technology, 1996.
- [813] Ninness, Brett and Gustafsson, Fredrik. *A Unifying Construction of Orthogonal Bases for System Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3388-3393, December, 1994.
- [814] Ninness, Brett and Gustafsson, Fredrik. *Orthornormal Bases for System Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 13-18, September, 1995.

- [815] Ninness, Brett. *Integral Constraints On The Accuracy of Least Squares Estimation*. Proceedings of 3rd European Control Conference, Rome , Italy, 839-844, September, 1995.
- [816] Ninness, Brett and Goodwin, Graham C.. *Estimation of model quality*. Automatica, Volume 31, No.12, 1771-1797, December, 1995.
- [817] Nishimura, T.. *On the a Priori Information in Sequential Estimation Problems*. IEEE Transactions on Automatic Control, Vol. AC-11, No.2, 197-204, April, 1966.
- [818] Niu, Steve S. and Fisher, D. Grant.. *Monitoring Parameter Identifiability During On-Line Identification*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 1717-1722, August, 1994.
- [819] Niu, Steve S.. *Monitoring Parameter Identifiability With AUDI*. Proceedings of the Third IEEE Conference on Control Applications obtido através de : <http://joakim.isy.liu.se>, 1-16, June 24, 1994.
- [820] Niu, Steve S.; Ljung, Lennart and Björck Åke. *Decomposition Methods for Least-Squares Parameter Estimation*. obtido através de : <http://joakim.isy.liu.se>, 1-19, artigo escrito em :21-12-1994.
- [821] Niu, Steve S.and Ljung, Lennart. *Multiple Models Parameter Estimation*. obtido através de : <http://joakim.isy.liu.se>, 1-22, artigo escrito em :21-12-1994.
- [822] Niu, Steve S.; Fisher, D. Grant; Ljung, Lennart and Shah, Sirish L.. *A Tutorial On Multiple Models Least-Squares and Augmented UD Identification*. obtido através de : <http://joakim.isy.liu.se>, 1-62, artigo escrito em :21-12-1994.
- [823] Noble, Ben and Daniel, James W.. *Applied Linear Algebra*. Prentice-Hall, Inc., 1988.
- [824] Nordström, Kjell; Karlsson, Erlendur and Malmgren, Anders. *An Optimal Control and Nonparametric Modelling Design Procedure For Linear Systems With a Finite Data Set*. Proceedings of 3rd European Control Conference, Rome , Italy, 845-850, September, 1995.
- [825] Norton, J.P.. *Identification and Application of Bounded-Parameter Models*. Automatica, Volume 23, No.4, 497-507, July, 1987.
- [826] Oară, Cristian and Ionescu, Vlad. *Strong Stabilizing Solutions To Continuous ad Discrete-Time Riccati Equations*. Proceedings of CONTROLO' 98 - 3<sup>rd</sup> Portuguese Conference on Automatic Control, 757-761, Coimbra, Portugal, 9-11 september, 1998.
- [827] Obermeier, Klaus K.. *Side by Side*. BYTE Vol.13, No.12, 275-283, November, 1988.
- [828] Ogata, Katsuhiko. *Discrete-Time Control Systems*. Prentice-Hall, Inc., 1987.
- [829] Ohsumi, Akira; Takashima, Satoschi and Kameyama, Kentaro. *Subspace-based Identification Of Stochastic Systems Using Innovation Models*. 11th IFAC Symposium on System Identification, Vol.3, 1113-1118, July, 1977.
- [830] Olsen, Thor O. and Henriksen, Rolf. *Comparison Of Different Parameter Estimation Methods In Flotation Processes*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 463-468, July, 1977.
- [831] Olsson, Gustaf. *Practical Experiences of Identification and Modeling from Experiments*. Dynamic Modeling and Expert Systems in Wastewater Engineering - Gilles G. Patry and David Chapman Lewis Publishers, Inc., 325-344, 1989.
- [832] Orfanidis, Sophocles J.. *Optimum Signal Processing - An Introduction*. McGraw-Hill Book Company, 1988.
- [833] Ortega, James M. *Introduction to Parallel and Vector Solution of Linear Systems*. Plenum Press, 1988.

- [834] Östermark, Ralf. *Vector forecasting and dynamic portfolio selection : Empirical efficiency of recursive multiperiod strategies*. European Journal of Operational Research, Volume 55, Number 1, pp. 46-56, November 6, 1991.
- [835] Östermark, Ralf and Aoki, Masanao. *State Space Modelling and Spectral Analysis of Cointegrated Vector Processes ( Evidence from the U.S. and Scandinavian Economies)*. International Journal of Systems Science - Volume 26, No. 12, 2327-2353, December, 1995.
- [836] Jr. Ostrom, Charles W.. *Time Series Analysis - Regression Techniques*. Series : Quantitative Applications in the Social Sciences, Sage University Paper, N.9, 1990.
- [837] Ottersten, B.; Viberg, M. and Kailath, T.. *Analysis Of Subspace Fitting and ML Techniques for Parameter Estimation From Sensor Array Data*. Transactions on Signal Processing - pages 1-30, December, 4, 1990.
- [838] Ottersten, Björn.; Asztély, David; Kristensson, Martin and Parkvall, Stefan. *A Statistical Approach To Subspace Based Estimation With Applications in Telecommunications*. 2nd International Workshop on TLS and Errors-in-Variables Modeling, pp. 285-294, 1997.
- [839] Ouladsine, M.; Kobi, A. and Ragot, J.. *Robust Identification of Dynamic Systems With Respect to Abnormal Value*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4170-4171, December, 1994.
- [840] Pacheco, Peter S.. *Parallel Programming with MPI*. Morgan Kaufmann Publishers, 1997.
- [841] Paige, Chris C.. *Properties of Numerical Algorithms Related to Computing Controllability*. IEEE Transactions on Automatic Control, Vol. AC-26, No.1, 130-138, February, 1981.
- [842] Paige, C. C. and Saunders, M.A.. *Towards A Generalized Singular Value Decomposition*. SIAM - Journal on Numerical Analysis, Volume 18, No.3, 398-405, June, 1981.
- [843] Paige, Chris C.. *Computing the Generalized Singular Value Decomposition*. SIAM - Journal on Scientific and Statistical Computing, Volume 7, No.4, 1126-1146, October, 1986.
- [844] Paige, Chris C.. *Some Aspects of Generalized QR Factorizations*. Reliable Numerical Computation - edited M.G.Cox and S. Hammarling, 73-91, - Based on papers from a conference in honour of James Hardy Wilkinson, 1990.
- [845] Paknys, Robert. *Evaluation of Hankel Functions with Complex Argument and Complex Order*. IEEE Transactions on Antennas and Propagation, Vol. 40, No.5, 569-578, May, 1992.
- [846] Pan, Victor. *How Can We Speed Up Matrix Multiplication*. SIAM Review, 1988, Vol.26, No. 3, 393-415, July, 1984.
- [847] Panzieri, Stefano and Ulivi, Giovanni. *Disturbance Rejection of Iterative Learning Control Applied To Trajectory Tracking for A Flexible Manipulator*. Proceedings of 3rd European Control Conference, Rome , Italy, 2374-2379, September, 1995.
- [848] Pappas, T.; Laub, A.J. and Sandell Jr., N.R. *On The Numerical Solution of The Discrete Time Algebraic Riccati Equation*. IEEE Transactions on Automatic Control, AC-25, pp. 631-641, 1980.
- [849] Parlett; Beresford N.. *The QR Algorithm*. Computing in Science & Engineering, Vol. 2, p. 38-42 No. 1, January/February, 2000.
- [850] Partington, Jonathan R.. *An Introduction to Hankel Operators*. Cambridge Univerty Press , 1988.
- [851] Pati, Y.C. and Krishnaprasad. *Rational Wavelets in Model Reduction and System Identification*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3394-3399, December, 1994.
- [852] Patry, Gilles G. and Chapman, David. *Dynamic Modeling and Expert Systems in Wastewater Engineering*. Lewis Publishers, Inc., 1989.

- [853] Patton, Peter C.. *Multiprocessors: Architecture and Applications*. Computers, Volume 18, Number 6, 29-40, June 1985.
- [854] Pawlak, M. and Greblicki, W.. *On Nonparametric Identification of Cascade Nonlinear Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2866-2867, December, 1994.
- [855] Peeters, B. and De Roeck, G.. *The Performance Of Time Domain System Identification Methods Applied To Operation Data*. Proceedings of DAMAS 97, Structural Damage Assessment Using Advanced Signal Processing Procedures, pp. 377-386, University of Sheffield, UK, 30 June - 2 July, 1997.
- [856] Peeters, Bart and De Roeck, Guido. *Reference-Based Stochastic Subspace Identification For Output-Only Modal Analysis*. Mechanical Systems and Signal Processing, V.13, n.6, 855-878, 1999.
- [857] Peeters, Bart and De Roeck, Guido. *Reference-Based Stochastic Subspace Identification in Civil Engineering*. Inverse Problems in Engineering, V.00, 1-28, 1999.
- [858] Peeters, Ralf. *System Identification Based on Riemannian Geometry Theory and Algorithms*. Ph.D. Thesis, Free University, Amsterdam, 1993.
- [859] Peled, Abraham. *The Next Computer Revolution*. Scientific American - Volume 257, 35-42, October, 1985.
- [860] Peloubet Jr., R.P ; Haller, R.L. and Bolding, R.M.. *On-Line Adaptive Control of Unstable Aircraft Wing Flutter*. Proceedings of the 29th Conference on Decision and Control - Hawaii - pp. 643-651, December, 1990
- [861] Peña, Ricardo S. Sánchez e Galarza, Cecilia G.. *Practical Issues in Robust Identification*. IEEE Transactions on Control Systems Technology, Vol.2, No.1, 54-56, February, 1994.
- [862] Penny, John and Lindfield, George. *Numerical Methods Using Matlab*. Prentice Hall, 2000.
- [863] Penrose, R. *A Generalized Inverse For Matrices*. Proceedings of the Cambridge Philosophical Society, Vol.51, Part 3, 406-413, July, 1955.
- [864] Pernebo, Lars and Silverman, Leonard M.. *Model Reduction Via Balanced State Space Representations*. IEEE Transactions on Automatic Control, Vol. AC-27, No.2, 382-387, April, 1982.
- [865] Peternell, Klaus. *The Identification of Linear Dynamic Systems by Subspace and Realization-Based Algorithms*. PhD Thesis - Technischen Universität Wien - March, 1995.
- [866] Peternell, Klaus; Scherrer W. and Deistler, M.. *Statistical Analysis of Subspace Identification Methods*. European Control Conference - ECC 95 - pag.1342-1347 - Roma - Italy - September, 1995.
- [867] Peternell, Klaus; Scherrer W. and Deistler, M.. *Statistical Analysis of Novel Subspace Identification Methods*. Signal Processing, Vol.52, 161-177, 1996.
- [868] Peters, G. and Wilkinson, J.H.. *The Least Square Problem and Pseudo-Inverses*. The Computer Journal, Vol.13, No. 3, 309-316, August, 1970.
- [869] Petitet, A.; Casanova, H.; Dongarra, J.; Robert, Y. and Whaley, R.C.. *Parallel and Distributed Scientific Computing - A Numerical Linear Algebra Problem Solving Environment Designer's Perspective* Handbook on Parallel and Distributed Processing, 417-456, 1999.
- [870] Pfeiffer, Carl. *On The Sequential Estimation Technique* IEEE Transactions On Automatic Control, Vol.A.C-11, 132-133, January, 1966.
- [871] Pfeiffer, Carl. *Comment On "The Sequential Estimation Technique"* IEEE Transactions On Automatic Control, Vol.A.C-11, 767, October, 1966.



- [872] Phatak, Alope and De Jong, Sijmen. *The geometry of partial least squares*. Journal of Chemometrics, Volume 11, Issue 4, 311-338, 1997.
- [873] Phatak, Alope . *Evaluation of Some Multivariate Methods and Their Applications in Chemical Engineering*. Ph.D. Thesis in Chemical Engineering, University of Waterloo, 1993.
- [874] Picci, Giorgio and Pinzoni, Stefano. *On Joint Identification and Latent Variable Estimation In Factor Analysis Models*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3877-3882, December, 1994.
- [875] Picci, Giorgio. *Statistical Properties of Certain Subspace Identification Methods*. 11th IFAC Symposium on System Identification, Vol.3, 1093-1099, July, 1977.
- [876] Picci, Giorgio and Katayama, Tohru *Stochastic realization with exogenous inputs and 'subspace-methods' identification*. Signal Processing, Volume 52, Issue 2, Pages 145-160, July, 1996.
- [877] Picci, Giorgio and Katayama, Tohru *Stochastic realization with exogenous inputs and subspace-methods identification*. Automatica, Volume 35, Issue 10, Pages 1635-1652, October, 1999.
- [878] Pijnappel, W.W.F.; De Beer, R. and Ormondt, D. Van. *State Space Modelling of Medical NMR Signals*. Proceedings of the Workshop on Circuits, Systems and Signal Processing - Houthalen, 225-230, April 8-9, 1992.
- [879] Plackett, R.L.. *Some Theorems in Least Squares*. Biometrika, 37, 149-157, August, 1950.
- [880] Plackett, R.L.. *Studies in History of Probability and Statistics. XXIX - The discovery of the method of least squares*. Biometrika, V.59, pp. 239-251, 1972.
- [881] Plett, Gregory and Kollár, István. *Specification for Common IEEE Styles*. IEEE Transactions on Instrumentation and Measurement, Vol. 46, No.6, 1217-1223, December, 1997.
- [882] Poon, Martin W.Y.; Khan, Rafaat H. and Le-Ngoc, Son. *A Singular Value Decomposition (SVD) Based Method for Suppressing Ocean Clutter in High Frequency Radar*. IEEE Transactions on Signal Processing, Vol. 41, No.3, 1421-1425, March, 1993.
- [883] Popescu, Th. D.. *IDPACK/PC - Software Package For Identification and System Parameter Estimation*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 442-447, July, 1991.
- [884] Popović D.. *Parallel Processing in Real-Time Control : A State-of-Art Report*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 799-804, August, 1994.
- [885] Prada, C. de; Moreno, R. Poch, M. and Robusté, J.. *Recursive Estimation of OUR in Activated Sludge Process*. Proceedings of the First European Control Conference - Grenoble, France, 1648-1653, July, 1991.
- [886] Press, William H. ; Teukolsky, Saul A.; Vetterling, William T. and Flannery, Brian P.. *Numerical Recipes in C - The Art of Scientific Computing* . Cambridge University Press, 1992.
- [887] Previdi, Fabio; Lovera, Marco and Mambretti, Stefano. *Identification of the rainfall-runoff relationship in urban drainage networks*. Control Engineering Practice, Volume: 7, pp. 1489-1504 , Issue: 12, December, 1999.
- [888] Proakis, John G.; Rader, Charles M.; Ling Fuyun and Nikias, Chrysostomos L.. *Advanced Digital Signal Processing*. Macmillan Publishing Company, 1992.
- [889] Pronzato, Luc and Walter, Eric. *Robustness To Outliers of Bounded-Error Estimators, Consequences on Experiment Design*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 821-826, July, 1991.

- [890] Protsuto, V.S., Voronov, V.A., Braun, V.I. and Trushin, A.A.. *Identification Of Control Objects In Mineral Dressing Processes*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1532-1539, August, 1988.
- [891] Puthenpura, S. and Sinha, N.K.. *Robust Instrumental Variables Method for System Identification*. Control - Theory and Advanced Technology , Vol.1, No.3, 175-188, November, 1985.
- [892] Pyle, L. Duane. *Generalized Inverse Computations Using The Gradient Projection Method*. Journal of The Association for Computing Machinery - JACM, Vol.11, No. 4, 422-428, October, 1964.
- [893] Quach, T. and Farooq, M.. *Maximum Likelihood Track Formation With The Viterbi Algorithm*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 271-276, December, 1994.
- [894] Qin, S. Joe. *Partial Least Squares Regression For Recursive System Identification*. Proceedings of the 32th Conference on Decision and Control, San Antonio, Texas, 2617-2622, December, 1993.
- [895] Qingquan, Li and Haiyi, Liu. *Identification And Self-Tuning Control Of A Paper Machine*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1509-1514, August, 1988.
- [896] Ramirez, W. Fred and Maciejowski, Jan M.. *Balanced Realization for State-Space Identification and Optimal Output Regulation*. AIChE Journal - , Volume 41, No.5, 1217-1228, May, 1995.
- [897] Rangan, Sundeeep and Ren, Wei. *Stochastic  $H^\infty$  Identification : An Iteratively Weighted Least Squares Algorithms*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3374-3379, December, 1994.
- [898] Rao, Bhaskar D. and Arun, K. S.. *Model Based Processing of Signals: A State Space Approach*. Proceedings of IEEE, Vol.80, No.2, 283-309, February, 1992.
- [899] Rao, B.S.Y. and Durrant-Whyte, H.F.. *A Decentralized Bayesian Algorithm For Identification*. Proceedings of the 29th Conference on Decision and Control, Honolulu, Hawaii, December, 1990.
- [900] Rao, Sailesh K. and Kailath, Thomas. *Regular Iterative Algorithms and Theirs Implementation on Processor Arrays*. Proceedings of IEEE, Vol.76, No.3, 259-269, March, 1988.
- [901] Rapisarda, Paolo and Willems, J.C.. *State Maps for Linear Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2868-2873, December, 1994.
- [902] Rapisarda, Paolo. *Linear Differential Systems*. PhD. Thesis, Rijksuniversiteit Groningen, Netherlands, June, 1998.
- [903] Raynal, Michel. *Distributed Algorithms: Their Nature and The Problems Encountered*. Proceedings of the Parallel and Distributed Algorithms, Chateau de Bonas, Gers, France, Part 3, 179-185, October, 1988.
- [904] Regalia, P.A.. *Subspace Methods in System Identification and Source Localization*. Algorithms and Parallel VLSI Architectures III - M. Moonen and F. Catthoor (Editors), pp. 13 - 23, Elsevier Science, 1995.
- [905] Reinke, R.D.; Prätzel-Wolters, D. and Owens, D.H.. *Learning and Replication of Periodic Signals in Dynamic Networks*. Proceedings of 3rd European Control Conference, Rome , Italy, 2118-2124, September, 1995.
- [906] Rice, John R.,. *A Theory Of Condition*. SIAM J.Numerical Analysis , Vol.3, No.2 , 287-310, June, 1966.
- [907] Richalet, J.; Ata-Doss, S. Abu el and Coic, A.. *Global Identification and Optimal Input Design*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 809-814, July, 1991.

- [908] Rissanen, J.. *Recursive Identification of Linear Systems*. SIAM J. Control, Vol.9, No.3 , 420-430, August, 1971.
- [909] Rissanen, J.. *Complexity of Models*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1730-1732, July, 1991.
- [910] Rhodes, Ian B.. *A Tutorial Introduction to Estimation and Filtering*. IEEE Transactions on Automatic Control, Vol.AC-16, No.6, 688-706, December, 1971.
- [911] Rocha, Paula and Willems, Jan C.. *Controllability For Delay-Differential Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2894-2897, December, 1994.
- [912] Rodriguez-Canabal J.. *The Geometry of the Riccati Equation*. Stochastics, Volume 1, No.2, 129-149, 1973.
- [913] Rodríguez-Vásquez, K.; Fonseca, C.M. and Fleming, P.J.. *An Evolutionary Approach To Non-Linear Polynomial System Identification*. 11th IFAC Symposium on System Identification, Vol.3, 1395-1400, July, 1977.
- [914] Rogers, R.D. and Vemuri, V.. *Time Series and the Forecasting Problem*. Artificial Neural Networks - Forecasting Time Series - Edt. V. Rao Vemuri and Robert D. Rogers - IEEE Computer Society Press - 1-6, 1994.
- [915] Rolain, Y.,; Pintelon R.; Xu, K.Q. and Vold, H.. *On The Use of Orthogonal Polynomials in High Order Frequency Domain System Identification and its Application to Modal Parameter Estimation*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3365-3373, December, 1994.
- [916] Roorda, Berend and Heij, Christiaan. *Global Total Least Squares Modeling of Multivariable Time Series*. IEEE Transactions on Automatic Control, Vol.40, No.1, 50-63, January, 1995.
- [917] Roorda, Berend . *Algorithms for Global Total Least Squares Modelling of Finite Multivariable Time Series*. Automatica, Volume 31, Number 3, 391-404, March, 1995.
- [918] Roy III, Richard H.. *ESPRIT - Estimation of Signal Parameters Via Rotational Invariance Techniques*. PhD thesis, Stanford University, Department of Electrical Engineering, Stanford, CA, June, 1987.
- [919] Rugh, Wilson J.. *Linear System Theory*. Prentice Hall, 1996.
- [920] Rust, B.; Burrus, W.R. and Schneeburger, C.. *A Simple Algorithm for Computing The Generalized Inverse Of The Matrix*. Communications of The Association for Computing Machinery - CACM, Vol.9, No. 5, 381-387, May, 1966.
- [921] Rutan, Sarah C.. *Fast On-line Digital Filtering*. Chemometrics and Intelligent Laboratory Systems, Vol. 6, No. 3, pp. 191-201, 1989.
- [922] Sá, E. Marques de and Zhang, Yu-Lin. *The number of Kronecker indices of square pencils of a special kind*. Linear Algebra and its Applications, Volume: 305, Issue: 1-3, pp. 15-21, January 15, 2000.
- [923] Saad, Youcef and Schulttz, Martin H.. *Parallel Direct Methods for Solving Banded Linear Systems*. Linear Algebra and Its Applications, Volume 88/89, 623-650, April, 1987.
- [924] Salazar, James Wilson Aliaga. *Identificação de Sistemas de Estruturas Mecânicas e Aplicações*. Tese de Mestrado, Engenharia Mecânica da PUC/RJ, 2000.
- [925] Sameh, Ahmed H.. *On Jacobi and Jacobi-Like Algorithms For A Parallel Computer*. Mathematics of Computation, Volume 25, No.115, 579-590, July, 1971.
- [926] Sameh, Ahmed H. and Kuck, David J.. *A Parallel QR Algorithm for Symmetric Tridiagonal Matrices*. IEEE Transactions on Computers, Vol.C-26, No.2, 147-153, February, 1977.

- [927] Sameh, Ahmed H. and Brent, Richard P.. *Solving Triangular Systems On A Parallel Computer*. J.SIAM Numer. Anal., Vol.14, No.6, 1101-1113, December, 1977.
- [928] Sameh, Ahmed H. and Kuck, David J.. *On Stable Parallel Linear System Solvers*. Journal of The Association for Computing Machinery - CACM, Vol.25, No. 1, 81-91, January, 1978.
- [929] Sandoz, D.J.. *CAD for the Design and Evaluation of Industrial Control Systems*. IEE Proceedings, Vol.131, Pt.D, No. 4, 125-139, July, 1984.
- [930] Sano, A. and Kikucki, B.. *Adaptive Control of Arterial Oxygen Pressure of Newborn Infants Under Incubator Oxygen Treatments*. IEE Proceedings, Vol.132, Pt.D, No. 5, 205-211, September, 1985.
- [931] Santos, P. Lopes dos e Carvalho, J.L. Martins de, *Transfer Function Parameter Estimation From Bode Plots*. Proceedings of 3rd European Control Conference, Rome , Italy, 2137-2143, September, 1995.
- [932] Santos, Maria Inês dos e Groehs, Ademar Gilberto. *Vetorização de um Sistema de Elementos Finitos no Super Computador Cray*. SuperComp 94 - Seminário de Computação Aplicada, Universidade Federal do Rio Grande do Sul - Porto Alegre - 81-84 - Setembro de 1994.
- [933] Saridis, George N.. *Stochastic Approximation Methods For Identification and Control - A Survey*. IEEE Transactions on Automatic Control, Vol.AC-19, No.6, 798-809, December, 1974.
- [934] Saridis, George N. and Stein, G.. *Stochastic Approximation Algorithms For Linear Discrete-Time System Identification*. IEEE Transactions on Automatic Control, Vol.AC-13, No.5, 515-523, October, 1968.
- [935] Savage, John E.. *Models of Computation - Exploring the Power of Computing*. Addison-Wesley, 1998.
- [936] Saxén, Henrik and Östermark, Ralf. *State realization with exogenous variables - A test on blast furnace data*. European Journal of Operational Research Volume: 89, Issue: 1-2, pp.34-52, February 22, 1996.
- [937] Saxena, N. and Clark, J.J.. *Parallel Implementation Of The Double Brackett Matrix Flow For Eigenvalue-Eigenvector Computation and Sorting*. Algorithms and Parallel VLSI Architectures III - M. Moonen and F. Catthoor (Editors), pp. 85 - 96, Elsevier Science, 1995.
- [938] Sayed, A.H.; Nascimento, V.H. and Chandrasekaran, S.. *Estimation and control with bounded data uncertainties*. Linear Algebra and its Applications Volume: 284, Issue: 1-3, pp. 259-306, November 15, 1998.
- [939] Sayed, Ali H.; Kailath, Thomas and Lev-Ari, Hanoch. *Generalized Chandrasekhar Recursions from the Generalized Schur Algorithm*. IEEE Transactions on Automatic Control, Vol.39, No.11, 2265-2269, November, 1994.
- [940] Schachermayer, W. and Deistler, M.. *The set of observationally equivalent errors-in-variables models*. Systems & Control Letters; Volume: 34, Issue: 1-2, pp. 101-104, May 25, 1998.
- [941] Schaper, C.D. ; Larimore, Wallace E.; Seborg, D.E. and Mellichamp, D.A.. *Identification of Chemical Processes Using Canonical Variate Analysis* . Computers Chemical Engng. , Vol.18, No. 1, pp. 55-69, 1994.
- [942] Schaper, Charles D.; Molehi, Mehrdad M.; Saraswat, Krishna C. and Kailath, Thomas. *Modeling, Identification, and Control of Rapid Thermal Processing Systems*. Journal Electrochemical Society , Vol.141, No. 11, pp. 3200-3209, November, 1994.
- [943] Scharf, Louis L.. *Statistical Signal Processing - Detection, Estimation and Time Series Analysis*. Addison Wesley, 1991.
- [944] Schelfhout, Geert. *Model Reduction for Control Design*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, Januari, 1996.

- [945] Schendel, U. *Introduction To Numerical Methods For Parallel Computers*. Ellis Horwood Limited, 1984.
- [946] Schendel, U. *Sparse Matrices: Numerical Aspects with Applications for Scientists and Engineers*. Ellis Horwood Limited, 1989.
- [947] Scherpen, Jacqueliën M.A. *Balancing for Nonlinear Systems*. PhD thesis, Departament of Applied Mathematics, University of Twente, The Netherlands, September, 1994.
- [948] Schinkel, S.; Klauw, A.C. van der; Boom, T.J.J. van den and Hof, P.M.J. Van den. *Input Design for Frequency-Function Variance Reduction in Prediction Error Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 2112-2117, September, 1995.
- [949] Schipp, F.; Bokor, J. and Gianone, L.. *Basic Selection For  $H_\infty$  Identification in the Disc Algebra*. Proceedings of 3rd European Control Conference, Rome , Italy, 19-24, September, 1995.
- [950] Schoukens, J.; Rolain, Y.; Gustafsson, F. and Pintelon, R.. *Fast Calculation of Least-Squares Estimatives for System Identification*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, FA08-1, December, 1998.
- [951] Schreiber, Robert and Loan, Charles Van. *A Storage - Efficient WY Representation for Products of Householder Transformations*. SIAM J.SCI.STAT.COMPUT. - Volume 10, No.1, 53-57, January, 1989.
- [952] Schubert, Wolfgang and Stengel, Robert F.. *Parallel Synthesis of Robust Control Systems*. IEEE Transactions On Control Systems Technology, Volume 6, No.6, 701-706, November, 1998.
- [953] Schumann, Andreas. *INID - A Computer-Software for Experimental Modeling*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 431-436, July, 1991.
- [954] Schweppe, Fred C.. *Evaluation of Likelihood Functions for Gaussian Signals*. IEEE Transactions on Information Theory , Vol.IT-11, No.1, 61-70, January, 1965.
- [955] Seal; Hilary L.. *Studies in History of Probability and Statistics. XV - The historical development of the Gauss linear model*. Biometrika, V.54, pp. 1-24, 1967.
- [956] Secchi, A.R.; Morari, M. and Biscaia Jr., E.C.. *The Waveform Relaxation Method in The Concurrent Dynamic Process Simulation*. SuperComp 94 - Seminário de Computação Aplicada, Universidade Federal do Rio Grande do Sul - Porto Alegre - 126-129 - Setembro de 1994. Computers Chemical Engineering , Vol.17, No.7, 683-704, July, 1993.
- [957] Secchi, Argimiro R.. *Supercomputação Aplicada à Minimização de Resíduos Industriais*. SuperComp 94 - Seminário de Computação Aplicada, Universidade Federal do Rio Grande do Sul - Porto Alegre - 126-129 - Setembro de 1994.
- [958] Seliger, Ralf; Köppen, Birgit and Frank, Paul M.. *Self-Tuning Adaptive Control For Field-Oriented Controlled Induction Motors - A Simulation Study*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 703-708, July, 1991.
- [959] Sen, A. and Sinha, N.K.. *On-Line System Identification Algorithm Combining Stochastic Approximation and Pseudoinverse*. Automatica, Volume 11, No.4, 425-429, July, 1975.
- [960] Seong, Seung H. and Peterka, Jon A.. *Computer simulation of non-Gaussian multiple wind pressure time series*. Journal of Wind Engineering and Industrial Aerodynamics Volume: 72, Issue: 1-3, pp. 95-105, November 12, 1997.
- [961] Seywald, Hans. *Identification of The Optimal Switching Structure For Range-Optimal Atmospheric Flight Trajectories*. Proceedings of 3rd European Control Conference, Rome , Italy, 3154-3158, September, 1995.

- [962] Shanmugam, Sam K. and Breinphol, Arthur M.. *Random Signals : Detection, Estimation and Data Analysis*. John Wiley & Sons, 1988.
- [963] Shedler, G.S.. *Parallel Numerical Methods For The Solution Of Equations*. Communications of The Association for Computing Machinery - CACM, Vol.10, No. 5, 286-291, May, 1967.
- [964] Shi, Zhaoyun and Aoyama, Hisayuki. *Identification Of The Self-Organizing Exponential Autoregressive Models*. 11th IFAC Symposium on System Identification, Vol.3, 1377-1381, July, 1977.
- [965] Shibata, H.; Fujinaka, T.; Fukudome, Y. and Furumoto, N.. *Some Aspects On The Accuracy Of System Identification*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1294-1298, July, 1991.
- [966] Shokoohi, Shahriar and Silverman, Leonard M.. *Identification and Reduction of Time-Varying Discrete-Time Systems*. Automatica, Volume 23, No.4, 509-521, July, 1987.
- [967] Silverman, Leonard M.. *Realization of Linear Dynamical Systems*. IEEE Transactions on Automatic Control, Vol.AC-16, No.6, 554-567, December, 1971.
- [968] Silverman, Leonard M. and Bettayeb, Maamar. *Optimal Approximation of Linear Systems*. Proceedings Joint Automatic Control Conference - JACC - FA8-A - , 1980.
- [969] Sima, Vasile and Varga, A.. *RASP-IDENT : Subspace Model Identification Programs*. DLR - Oberpfaffenhofen , TR R176-94, October, 1996.
- [970] Sima, Vasile. *Algorithms and LAPACK-Based Software for Subspace Identification*. Proc. of the IEEE International Symposium on Computer Aided Control System, Design - CACSD'96 Symposium, Dearborn, MI, pp. 182-187, 1996.
- [971] Sima, Vasile; Varga, A. and Verhaegen, M.. *LAPACK-Based Software for Subspace Identification*. WGS Newsletter, No. 10, pp. 9-14, 1996.
- [972] Sima, Vasile. *Subspace-based Algorithms for Multivariable System Identification*. Studies in Informatics and Control, Vol.5, No. 4, pp. 335-344, December, 1996.
- [973] Sima, Vasile. *High-Performance Numerical Software for Control Systems, and Subspace-Based System Identification*. WGS Report 97-2,p. 1 - 24; March, 1997.
- [974] Sima, Vasile. *SLICOT Linear Systems Identification Toolbox*. SLICOT Working Note 2000-4, p. 1 - 28; July, 2000.
- [975] Sims, Craig S. and Park, Y.W.. *Ordered Sequential Filtering For Finite Dimensional Systems With Distributed Observations*. IEEE Transactions on Automatic Control, Vol. AC-20, No.3, 351-354, June, 1975.
- [976] Singh, Anita. *Multivariate Decision and Detection Limits* . Analytica Chimica Acta - Volume 277, No.2, 205-214, May, 1993.
- [977] Singh, M.G. and Titli, R.. *Systems : Decomposition, Optimization and Control*. Pergamon Press, 1978.
- [978] Sinha, N.K. and Pille, W.. *Online Parameter Estimation Using Matrix Pseudoinverse*. Proceedings IEE, Vol.118, No.8, 1041-1046, August, 1971.
- [979] Sinha, N.K. and Kwong, Y.H.. *Recursive Identification of the Parameters of Multivariable System*. Proceedings of the Fourth IFAC International Symposium - Multivariable Technological Systems - Fredericton, Canada, 323-327, July, 1977.
- [980] Sinha, N.K. and Kwong, Y.H.. *Recursive Estimation Of The Parameters of Linear Multivariable Systems*. Automatica, Vol. 15, No.4, 471-475, July, 1979.
- [981] Sinha, N.K. and Kuszta, B.. *Modeling and Identification of Dynamic Systems*. Van Nostrand Reinhold Company, 1983.

- [982] Sinha, N.K. and Puthenpura, S.. *Choice of the Sampling Interval for the Identification of Continuous-Time Systems From Samples of Input/Output Data*. IEE Proceedings, Vol.132, Pt.D, No.6, 263-267, November, 1985.
- [983] Sjöberg, Jonas; Zhang, Qinghua; Ljung, Lennart; Benveniste, Albert; Delyon, Bernard; Glorennec, Pierre-Yves ; Hjalmarssonand, H åkan and Juditsky , Anatoli. *Nonlinear black-box modeling in system identification: a unified overview*. Automatica, Volume 31, No.12, 1691-1724, December, 1995.
- [984] Sloman, Morris and Kramer, Jeff. *Distributed Systems and Computer Networks*. Prentice Hall, 1987.
- [985] Smith, Roy S. and Doyle, John C.. *Model Validation : A Connection Between Robust Control and Identification*. IEEE Transactions on Automatic Control, Vol.37, No.7, 942-951, July, 1992.
- [986] Snir, Marc; Otto, Steve; Huss-Lederman, Steven; Walker, David and Dongarra, Jack. *MPI: The Complete Reference*. The MIT Press, Cambridge, Massachusetts, 1996.
- [987] Speyer, Jason L. and Gustafson, Donald E.. *An Approximation Method for Estimation in Linear Systems With Parameter Uncertainty*. IEEE Transactions on Automatic Control, Vol. AC-20, No.3, 354-359, June, 1975.
- [988] Söderström, T. and Stewart, G. W.. *On The Numerical Properties of An iterative Method For Computing The Moore-Penrose Generalized Inverse*. SIAM - Journal on Numerical Analysis, Volume 11, No.11, 61-74, March, 1974.
- [989] Söderström, T. and Stoica, P.. *System Identification*. Prentice Hall, 1989.
- [990] Söderström, T. *Discrete-Time Stochastic Systems : Estimation and Control* . Prentice Hall International Series in Systems and Control Engineering, 1994.
- [991] Söderström, T. and Åstrom, K. J.. *Special issue on trends in system identification (Editorial)*. Automatica, Volume 31, No.12, 1689-1690, December, 1995.
- [992] Solo, Victor. *Transfer Function Estimation as an Inverse Problem*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3354-3358, December, 1994.
- [993] Sorensen, H.W.. *Least-Squares Estimation: From Gauss to Kalman*. IEEE - Spectrum, Volume 7, No.7, 63-68, July, 1970.
- [994] Sorensen, Harold W.. *Parameter Estimation: Principles and Problems* Marcel Dekker, Inc., 1980.
- [995] Sørli, James A.. *A Matlab Interface to IdKit* . Royal Institute of Technology,(KTH), Stockholm, Sweden, Technical Reports in Reglerteknik: IR-S3-REG-9302, 1993.
- [996] Stark, Henry and Yang, Yongyi. *Vector Space Projections*. John Wiley & Sons, Inc. , 1998.
- [997] Steedly, William M.; Ying, Ching-Hui J. and Moses, Randolph L.. *Statistical analysis of TLS-based prony techniques*. Automatica, Volume 30, No.1, 115-129, January, 1994.
- [998] Steinhardt, Allan O.. *Householder Transforms in Signal Processing*. IEEE - ASSP Magazine, 4-12, July, 1988.
- [999] Stevens, W. Richard. *UNIX Network Programming*. Prentice Hall Software Series, 1990.
- [1000] Stewart, G.W.. *Introduction to Matrix Computations*. Academic Press, 1973.
- [1001] Stewart, G.W.. *Error and Perturbation Bounds For Subspaces Associated With Certain Eigenvalue Problems*. SIAM Review, Vol.15, No. 4, 727-764, October, 1973.
- [1002] Stewart, G.W.. *Perturbation Bounds For The QR Factorization of a Matrix*. J.SIAM Numer. Anal., Vol.14, 509-518, June, 1977.

- [1003] Stewart, G.W.. *On The Perturbation Of Pseudo-Inverses, Projections and Linear Least Squares Problems*. J.SIAM Review, Vol.19, 634-662, October, 1977.
- [1004] Stewart, G.W. and Sun, Ji-guang. *Matrix Perturbation Theory*. Academic Press, 1990.
- [1005] Stewart, G.W.. *On the Early History Of The Singular Value Decomposition*. SIAM Review, Volume 35, No.4, 551-566, December, 1993.
- [1006] Stewart, G.W.. *The Decompositional Approach to Matrix Computation*. Computing in Science & Engineering, Vol. 2, p. 50-59 No. 1, January/February, 2000.
- [1007] Stigler, Stephen. *Gauss and The Invention of Least Squares*. The Annals of Statistics, V.9, No.3; pp. 465-474, 1981.
- [1008] Stoica, Petre; Nehorai, Arye and Söderström, Torsten. *Optimal Decentralized Methods For Source Location*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1404-1409, July, 1991.
- [1009] Stoica, Petre; Händel, Peter and Söderström, Torsten. *Approximate maximum likelihood frequency estimation*. Automatica, Volume 30, No.1, 131-145, January, 1994.
- [1010] Stoica, Petre; Mari, Jorge and McKelvey, Tomas. *Vector ARMA Estimation: an enhanced subspace approach*. Proceedings of the 38th Conference on Decision & Control, Phoenix, Arizona, USA, 3665-3670, December, 1999.
- [1011] Stoica, Petre and Jansson, Magnus. *A Transfer Function Approach to MIMO System Identification*. Proceedings of the 38th Conference on Decision & Control, Phoenix, Arizona, USA, December, 1999.
- [1012] Stoica, Petre and Jansson, Magnus. *MIMO System Identification: State-space versus Transfer Function*. IEEE Trans. On Signal Processing, Volume 48, No.11, 3087-3099, November, 2000.
- [1013] Stone, Harold S.. *An Efficient Parallel Algorithm for The Solution Of A Tridiagonal Linear System of Equations*. Journal of The Association for Computing Machinery - CACM, Vol.20, No. 1, 27-38, January, 1973.
- [1014] Stone, Harold S.. *Parallel Tridiagonal Equation Solvers*. ACM Transactions on Mathematical Software, Vol.1, No. 4, 289-307, December, 1975.
- [1015] Stoorvogel, A.A. and van Schuppen, J.H.. *System identification with information theoretic criteria*. Report National Research Institute for Mathematics and Computer Science in the Netherlands, BS-R9513, ISSN 0924-0659, 47 pages, 1995.
- [1016] Strang, Gilbert. *Linear Algebra and Its Applications*. Academic Press, 1976.
- [1017] Strang, Gilbert. *Introduction To Linear Algebra*. Wellesley-Cambridge Press, 1993.
- [1018] Strassen, Volker. *Gaussian Elimination Is Not Optimal*. Numerische Mathematik, Band 13 - Heft 4, 354-356, 1969.
- [1019] Strejc, V.. *Least Squares Parameter Estimation*. Automatica, Volume 16, No. 5, 535-550, September 1980.
- [1020] Strejc, Vladimir. *Trends in Identification*. Automatica, Volume 17, No. 1, 7-21, January 1981.
- [1021] Sugie, Toshiharu and Okada, Masafumi. *Iterative Controller Design Method Based on Closed-Loop Identification*. Proceedings of 3rd European Control Conference, Rome, Italy, 1243-1248, September, 1995.
- [1022] Sugisaka, M.. *A New Identification Method Using Neurocomputer*. 11th IFAC Symposium on System Identification, Vol.3, 1367-1370, July, 1977.



- [1023] Suneja, B. P. and Datta, T. K.. *Active control of ALP with improved performance function*. Ocean Engineering Volume: 25, Issue: 10, pp. 817-835, November, 1998,
- [1024] Swidenbank, E. and Hogg B.W.. *Application of System Identification Techniques to Modelling a Turbogenerator*. IEE Proceedings, Vol.136, Pt.D, No.3, 113-121, May, 1989.
- [1025] Swindlehurst, Arnold Lee. *Applications of Subspace Fitting To Estimation and Identification*. PhD thesis, Stanford University, Department of Electrical Engineering, Stanford, CA, June, 1991.
- [1026] Swindlehurst, Arnold Lee; Roy, R.; Ottersten, B. and Kailath, T.. *A Subspace Fitting Method for Identification of Linear State-Space Models*. IEEE Transactions on Automatic Control, Vol. 40, No.2, 311-316, February, 1995.
- [1027] Tadjine, M. and M'saad, M.. *On Identification and Control Design for Robust Performance*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3380-3387, December, 1994.
- [1028] Takane, Yoshio and Yanai, Haruo. *On oblique projectors*. Linear Algebra and its Applications Volume: 289, Issue: 1-3, pp. 297-310, March 1, 1999.
- [1029] Takeuchi, Kei; Yanai, Haruo and Mukherjee, Bishwa Nath. *The Foundations of Multivariate Analysis : A Unified Approach by Means of Projection Onto Linear Subspaces*. Wiley Eastern Limited, 1982.
- [1030] Tamariz, Annabell Del Real. *Uma nova proposta para Solução Computacional da Equação Algébrica de Riccati em formas Sequencial e Paralela*. Tese de Mestrado, LC-SI/DMCSI/FEEC/UNICAMP, (1999).
- [1031] Tamariz, Annabell Del Real; Bottura, Celso P.; Fonseca Neto, J.V. e Barreto, Gilmar. *Formas Sequencial e Paralela para Solução da Equação Algébrica de Riccati por um Algoritmo de Schur-Modificado*. XV Congresso Brasileiro de Engenharia Mecânica - Cobem 99, Águas de Lindóia, São Paulo, Brasil.
- [1032] Tanabe, Kunio. *Projection Method For Solving a Singular System Of Linear Equations and Its Applications*. Numerische Mathematik 17, 203-214, 1971.
- [1033] Tanaka, M.. *Approximate expression of Membership Set in Unknown But Bounded Error Approach by Neural Network*. Proceedings of 3rd European Control Conference, Rome , Italy, 931-936, September, 1995.
- [1034] Tao, Gang. *A Non-quadratic Gradient Algorithm*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3608-3609, December, 1994.
- [1035] Temeltas, H. and Asher, G.M.. *Real Time Identification of Robot Dynamic Parameters Using Parallel Processing*. IEE Proceedings Control Theory and Applications, Vol. 142, No.4, 359-377, July, 1995.
- [1036] Tempo R.; Barmish, B.R. and Trujillo J.. *Robust Estimation and Prediction Using ARMA Models : A Nonstatistical Approach*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1136-1139, August, 1988.
- [1037] Teo, K.L.; Zang, Z. and Yan, W.. *Learning Stabilization of Linear Unknown Unstable Systems*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4156-4157, december, 1994.
- [1038] Terada, Routo. *Desenvolvimento de Algoritmos e Complexidade de Computação*. Terceira Escola de Computação - PUC - Rio de Janeiro, 1982.
- [1039] Tewarson, R.P.. *A Computational Method for Evaluating Generalized Inverses*. The Computer Journal, Vol.10, No.4, 411-413, February, 1968.
- [1040] Therapos, C.P. *New Table to Implement Routh's Stability Test and Matrix Continued Fraction Expansion and Inversion*. IEE Proceedings, Vol.134, Pt.D, No. 5, 327-332, September, 1987.

- [1041] Tiao G.C. and Box G.E.P. *Modeling Multiple Time Series With Applications*. JASA - Journal of the American Statistical Association, Vol.76 , No.376, 802-816, December, 1981.
- [1042] Tong Lang and Perreau, Sylvie. *Multichannel Blind Identification: From Subspace to Maximum Likelihood Methods*. Proceedings of The IEEE, Vol.86, No.10, 1951-1968, October, 1998.
- [1043] Tou, J.T.. *On The Inverse of Vandermonde Matrix* IEEE Transactions On Automatic Control, Vol.A.C-11, 769, October, 1966.
- [1044] Tsai, Mi-Ching and Shih, Yen-Ping. *Balanced Minimal Realization Via Singular Value Decomposition of Sarason Operator*. Automatica, Volume 24, No.5, 701-705, September, 1988.
- [1045] Tsatsanis, M.K. and Giannakis , G.B.. *A Basic Expansion Approach for Detecting Transient Plant Disturbances and Jumps*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3412-3417, December, 1994.
- [1046] Tse, Edison and Weinert, Howard L.. *Structure Determination and Parameter Identification For Multivariable Stochastic Linear Systems*. IEEE Transactions on Automatic Control, Vol. AC-20, No.5, 603-613, October, 1975.
- [1047] Tse, Johnson; Bentsman, Joseph and Miller, Norman. *Minimax Long Range Parameter Estimation*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 277-282, December, 1994.
- [1048] Tu, Te-Ming; Shyu, Hsuen-Chyun; Lee, Ching-Hai and Chang, Chein-I. *An oblique subspace projection approach for mixed pixel classification in hyperspectral images*. Pattern Recognition Volume: 32, Issue: 8, pp. 1399-1408, August, 1999.
- [1049] Tuel, W.G.. *Correction to and Extension of "On the a priori Information in Sequential Estimation Problems"*. IEEE Transactions on Automatic Control, Vol. AC-12, No.1, 123, February, 1967.
- [1050] Tugnait, Jitendra K.; Ye, Yisong. *Stochastic System Identification With Noisy Input-Output Measurements Using Polyspectra*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2845-2850, December, 1994.
- [1051] Uciński, D. and Korbicz, J.. *An Algorithm for the Computation of Optimal Measurement Trajectories in Distributed Parameter Systems Identification*. Proceedings of 3rd European Control Conference, Rome , Italy, 1267-1272, September, 1995.
- [1052] Ulanicki, Bogumil and Kennedy, Paul Ritchie. *An Optimization Technique For Water Network Operations and Design*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4114-4115, December, 1994.
- [1053] Unbehauen, Heinz and Rao, Ganti Prasada. *Identification of Continuous Systems*. North-Holland, 1987.
- [1054] Unbehauen, Heinz and Rao, Ganti Prasada. *A review of identification in continuous-time systems*. Annual Reviews in Control Volume: 22, pp. 145-171, 1998.
- [1055] Ursella, Enrico; Soatto, Stefano and Perona, Pietro. *Motion and Attitude Estimation for Articulated Bodies Via Dynamic Vision*. Proceedings of 3rd European Control Conference, Rome , Italy, 2357-2362, September, 1995.
- [1056] Vaccaro, Richard J.. *Digital control : a state-space approach*. McGraw-Hill Series in Electrical and Computer Engineering, 1995.
- [1057] Vámos, Tibor and Bokor, József. *Bird's Eye View on Control Theory - motion, spaces, transformations*. Annual Reviews in Control Volume: 21, pp. 1-11, 1997.
- [1058] Vandaele, Piet and Moonen, Marc *Two Deterministic Blind Channel Estimation Algorithms Based on Oblique Projections*. Signal Processing Volume: 80, Issue: 3, pp. 481-495, March, 2000

- [1059] Van De Velde, Eric F.. *Concurrent Scientific Computing* . Springer-Verlag, 1994.
- [1060] Van den Hof, Paul J.M.. *Model Sets and parametrizations for Identification of Multivariable Equation Error Models*. Automatica, Volume 30, No.3, 433-446, March, 1994.
- [1061] Van den Hof, J.M. and van Schuppen, J.H.. *Realization of Positive Linear Systems Using Polyhedral Cones*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3889-3893, December, 1994.
- [1062] Van den Hof, J.M.. *Structural Identifiability of Linear Mamillary Compartmental Systems*. European Control Conference - ECC 95 - pag.1318-1323 - Roma - Italy - September, 1995.
- [1063] Van den Hof; Jacqueline . *Systems Theory and Systems Identification Compartmental Systems*. PhD. Thesis Rijksuniversiteit Groningen, Netherlands,, Netherlands, November , 1996.
- [1064] Van den Hof; J.M.. *Structural Identifiability of Linear Compartmental Systems*. IEEE Transactions on Automatic Control, Vol.AC-43, No.6, 800-818, June, 1998.
- [1065] Van den Hof, Paul J.M. and Schrama, Ruud J. P.. *Identification and control - closed-loop issues*. Automatica, Volume 31, No.12, 1751-1770, December, 1995.
- [1066] Van den Hof, Paul J.M. ; Heuberger, Peter S. C. and Bokor, József. *System identification with generalized orthonormal basis functions*. Automatica, Volume 31, No.12, 1821-1834, December, 1995.
- [1067] Van den Hof, Paul J.M. *Closed-loop issues in system identification*. Annual Reviews in Control Volume: 22, pp. 173-186 , 1998.
- [1068] Van den Hof, Paul J.M. *System Identification*. Lectures Notes tn 3111, Delft University of Technology, January, 2001.
- [1069] Van Der Veen, Alle-Jan and Dewilde, Patrick. *Time-Varying Systems Theory For Computational Networks*. Algorithms and Parallel VLSI Architectures II -P. Quinton and Y. Robert (Editors), pp. 103 - 127, Elsivier Science, 1992.
- [1070] Van Der Veen, Alle-Jan; Deprettere, Ed F. and Swindlehurst, Lee A.. *Subspace-Based Signal Analysis Using Singular Value Decomposition*. Proceedings of the IEEE, Vol. 81, No.9, 1275-1308, September, 1993.
- [1071] Van Der Veen, Alle-Jan. *Computation of Inner-Outer Factorization For Time-Varying Systems*. Technical Report - Delft University of Technology, 1-18 - July, 1993.
- [1072] Van Der Veen, Alle-Jan and Dewilde,P.M.. *Connections of Time-Varying Systems and Computational Linear Algebra*. In H. Dedieu (ed.), Elsevier, Circuit Theory and Design: Proc. 11-th ECCTD, pp. 167-172, Davos, Switzerland, August, 1993.
- [1073] Van Der Veen, Alle-Jan and Dewilde, P.M.. *Modeling Computational Networks by Time-varying Systems*. Integration, the VLSI Journal, vol. 16, no. 3, pp. 267-291, December, 1993.
- [1074] Van Der Veen, Alle-Jan and Verhaegen, Michel. *On Spectral Factorization and Riccati Equations For Time-Varying Systems in Discrete Time*. Technical Report - Delft University of Technology, 1-24 - December, 1993.
- [1075] Van Der Veen, Alle-Jan. *Algebraic Methods For Deterministic Blind Beamforming*. Proceedings of The IEEE, Vol.86, No.10, pp. 1987-2008, October, 1998.
- [1076] Van Dooren, Paul and Verhaegen Michel. *On The Use of Unitary State-Space Transformations*. Conference on Linear Algebra and Its Role in Systems Theory, Maine, 1984; publicado por Datta, Biswa; Johnson, Charles and Plemmons, Robert em Contemporary Mathematics, Volume 47, pp. 447-463, 1985.

- [1077] Van Dooren, Paul and Verhaegen Michel. *Condensed Forms For Efficient Time-Invariant Kalman Filtering*. SIAM Journal Scientific and Statistical Computing, Volume 9, Number 3, pp. 516-530, May 1988.
- [1078] Van Gestel, T.; Suykens, J.; De Moor, Bart and Van Overschee, P.. *Identification of Stable and Positive Real Models using Matrix Inequalities in Subspace Methods*. Internal Report 99-30, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1999.
- [1079] Van Gestel, Tony; Suykens, Johan A.K.; Van Dooren; Paul and De Moor, Bart. *Identification of stable models in subspace identification by using regularization*. Internal Report 99-81, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1999. Accepted for publication in IEEE Transactions on Automatic Control.
- [1080] Van Gestel, Tony; Suykens, Johan A.K.; Baesens, Bart Baesens; Viaene, Stijn; Vanthienen, Jan; Dedene, Guido; De Moor, Bart and Vandewalle, Joos. *Benchmarking Least Squares Support Vector Machine Classifiers*. Internal Report 00-37, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2000.
- [1081] Van Gestel, Tony; Suykens, Johan A.K.; Dooren, Paul Van; De Moor, Bart and Van Overschee P.. *Imposing stability in subspace identification by regularization*. Internal Report 00-30, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 2000. Accepted for publication in 39th IEEE Conference on Decision and Control, CDC 2000, Sydney, Australia.
- [1082] Van Gestel, Tony; Suykens, Johan A.K.; De Moor, Bart and Baestaens, Dirk-Emma. *Volatility Tube Support Vector Machines*. Neural Network World, Special Issue on PASE 2000, vol. 10, no. 1-2, pp. 287-297, Jan. 2000.
- [1083] Van Loan, Charles F.. *Generalizing The Singular Value Decomposition*. SIAM - Journal on Numerical Analysis, Volume 13, No.1, 76-83, March, 1976.
- [1084] Van Loan, Charles F.. *A General Matrix Eigenvalue Algorithm*. SIAM - Journal on Numerical Analysis, Volume 12, No.6, 819-834, December, 1975.
- [1085] Van Overschee, P.; De Moor, B. and Suykens J.. *Subspace Algorithms For System Identification and Stochastic Realization*. Proc. of the International Symposium on Recent Advances in Mathematical Theory of Systems, Control, Networks and Signal Processing (MTNS'91), Kobe, Japan, pp. 589-595, Jun. 1991
- [1086] Van Overschee P. and De Moor, B.. *Two Subspace Algorithms For the Identification of Combined Deterministic Stochastic Systems*. Proc. of the 31st IEEE Conference on Decision and Control (CDC'92), Tucson, Arizona, Dec. 1992, pp. 511-516, 1992.
- [1087] Van Overschee P. and De Moor, B.. *An exact subspace algorithm for the identification of combined deterministic-stochastic systems*. Internal Report 92-11, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1992.
- [1088] Van Overschee P. and De Moor, B.. *A Subspace Algorithm for Combined Stochastic-Deterministic State Space Identification*. Internal Report 92-10, ESAT-SISTA, K.U.Leuven (Leuven, Belgium), 1992.
- [1089] Van Overschee P. and De Moor, B.. *Subspace Algorithms for the Stochastic Identification Problem*. Automatica, Volume 29, Number 3, pp. 649-669, 1993.
- [1090] Van Overschee, Peter and De Moor, Bart. *N4SID: subspace algorithms for the identification of combined deterministic-stochastic systems*. Automatica, Volume 30, No.1, 75-93, January, 1994.
- [1091] Van Overschee, Peter. *Subspace Identification - Theory - Implementation - Application*. Phd Thesis - Katholieke Universiteit Leuven - Februari - 1995.
- [1092] Van Overschee, Peter and De Moor, Bart. *A Unifying Theorem for Three Subspace System Identification Algorithms*. Automatica, Volume 31, No.12, 1853-1864, December, 1995.
- [1093] Van Overschee, Peter and De Moor, Bart. *Choice of state-space basis in combined deterministic-stochastic subspace identification*. Automatica, Volume 31, No.12, 1877-1883, December, 1995.

- [1094] Van Overschee, Peter and De Moor, Bart. *Subspace Identification For Linear Systems*. Kluwer Academic Publishers - 1996.
- [1095] Van Overschee, Peter; De Moor, Bart; Dehandschutter, Wouter and Swevers, Jan. *A Subspace Algorithm for the Identification of Discrete Time Frequency Domain Power Spectra*. Automatica, Volume 33, No.12, 2147-2157, December, 1997.
- [1096] Van Overschee P.; De Moor B. and Favoreel W.. *Numerical algorithms for subspace state space system identification (N4SID)*. Proc. of the ASME Design Engineering Technical Conference (ASME DETC'97), Sacramento, 9 p., California, September, 1997.
- [1097] Vanpouche, Filiep. *Algorithms and Architectures For Adaptive Array Signal Processing*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, Februari, 1995.
- [1098] Van Schuppen, Jan H.. *Stochastic realization of  $\sigma$ -Algebras*. Proc. 14th International Symposium MTNS2000, Perpignan, June 2000, published on CD-ROM only, Universite de Perpignan, Perpignan, France, 2000.
- [1099] Vaughan, David R.. *A Nonrecursive Algebraic Solution for the Discrete Riccati Equation*. IEEE Transactions on Automatic Control, Vol. AC-15, No.5, 597-599, October, 1970.
- [1100] Verdult, V.; Verhaegen, M.; Chou, C.T. and Lovera, M.. *Efficient and Systematic Identification of MIMO Bilinear State Space Models*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, WP07-2, December, 1998.
- [1101] Verdult, V. and Verhaegen, M.. *Identification of Multivariable Linear Parameter Varying Systems Based on Subspace Techniques*. Conference on Decision & Control, CDC 2000, Sydney, Australia, December, 2000.
- [1102] Verdult, V. and Verhaegen, M.. *Bilinear State Space Systems for Nonlinear Dynamical Modelling*. Theory in Biosciences, Vol. 119, number 1, p. 1-9, 2000.
- [1103] Veres, Sándor M.. *Relations of Information Criteria for Structure Selection of Dynamic Systems*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 401-406, July, 1991.
- [1104] Veres, Sándor M.. *Relations Between Information Criteria for Model-Structure Selection. Part 1. The Role of Bayesian Model Order Estimation. Part2. Modelling By Shortest Data Description*. International Journal Of Control, Vol.52, No.2, 389-421, August, 1990.
- [1105] Verhaegen, Michel H.. *A New Class of Algorithms in Linear System Theory With Application to Real-Time Aircraft Model Identification*. PhD thesis, Katholieke Universiteit Leuven - Kard. Mercierlaan 94, 3001 - Leuven (Heverlee), Belgium, November, 1985.
- [1106] Verhaegen, Michel H. and Van Dooren, Paul. *Numerical Aspects of Different Kalman Filter Implementations*. IEEE Transactions on Automatic Control, Vol.AC-31, no.10, 907-917, October, 1986.
- [1107] Verhaegen, Michel H.. *The Use of The QR Factorization in The Partial Realization Problem*. NASA Technical Memorandum 100018, Ames Research Center, Moffett Field, California, september, 1987.
- [1108] Verhaegen, Michel H.. *The Minimal Residual QR Factorization Algorithm For Reliably Solving Subset Regression Problems*. NASA Technical Memorandum 100021, Ames Research Center, Moffett Field, California, September, 1987.
- [1109] Verhaegen, Michel H.. *Improved Understanding of The Loss-of-Symmetry Phenomenon in the Conventional Kalman Filter*. IEEE Transactions on Automatic Control, Vol.34, no.3, 331-333, March, 1989.

- [1110] Verhaegen, Michel H.; Kraan, R.M.C.; Rens, C.R. de; van Rhiju, A.P.J. and Verbruggen, H.B.. *Practical System Identification In The Process Industry*. Control and Computers - Volume 19, No.1, 16-20, July, 1991.
- [1111] Verhaegen, Michel. *A Novel Non-Iterative MIMO State Space Model Identification Technique*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1453-1458, July, 1991.
- [1112] Verhaegen, Michel and Deprettere, Ed. *A Fast, Recursive MIMO State Space Model Identification Algorithm*. Proceedings of the 30th Conference on Decision and Control - Brighton, England, 1349-1354, December, 1991.
- [1113] Verhaegen, Michel and Dewilde, Patrick. *Subspace Model Identification - Part 1 : The Output-Error State-Space Model Identification Class of Algorithms*. Internacional Journal of Control, Volume 56, Number 5, 1187-1210, November, 1992.
- [1114] Verhaegen, Michel and Dewilde, Patrick. *Subspace Model Identification - Part 2 : Analysis of The Elementary Output-Error State-Space Model Identification Algorithms*. Internacional Journal of Control, Volume 56, Number 5, 1211-1241, November, 1992.
- [1115] Verhaegen, Michel. *Subspace Model Identification - Part 3 : Analysis of The Ordinary Output-Error State-Space Model Identification Algorithms*. Internacional Journal of Control, Volume 58, Number 3, 555-586, September, 1993.
- [1116] Verhaegen, Michel. *Identification of The Deterministic Part of MIMO State-Space Models Given in Innovations Form From Input-Output Data*. Automatica, Volume 30, Number 1, 61-74, January, 1994.
- [1117] Verhaegen, Michel; Varga, Andras and Grübel, George. *Some Experience With The MOESP Class of Subspace Model Identification Methods in Identifying The BO105 Helicopter*. DLR-IB 515-94/16, p. 1-27, DLR Deutsche Forschungsanstalt für Luft- und Raumfahrt, D-82234 Oberpfaffenhofen, 1994.
- [1118] Verhaegen, Michel and Yu, Xiaode. *An Class of Subspace Model Identification Algorithms to Identify Periodically and Arbitrarily Time-varying Systems*. Automatica, Volume 31, Number 2, 201-216, February, 1995.
- [1119] Verhaegen, Michel and Westwick, David. *A Subspace Model Identification Algorithms to Identify MIMO Wiener Type Of Non-Linear Systems*. Proceedings of 3rd European Control Conference, Rome , Italy, 3429-3434, September, 1995.
- [1120] Verhaegen, Michel H.; Scherpen, J. and Haverkamp. *Modelvorming Identificatie en Simulatie (ET 11-31) Oefenopgaves*. Faculty of Information Technology and Systems, Department of Electrical Engineering Systems and Control Engineering Group, Delft University of Technology, notas de aula, February, 1998.
- [1121] Verhaegen, Michel H.. *Symbolic and Computational Pre-Processing in Physical Parameter Estimation of Multi-Body mechanical systems* . SLICOT Working Note 1999-20, p. 1-6; November 1999.
- [1122] Verhaegen, Michel H.. *Model Based Signal Processing for Control*. Course in Systems and Control Engineering Division University of Twente; Dec. 1999 - Feb. 2000.
- [1123] Viberg, Mats. *Effects of Unknown Noise Covariance On Parametric Array Processing Algorithms*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1392-1397, July, 1991.
- [1124] Viberg, Mats. *Subspace-based Methods for the Identification of Linear Time-invariant Systems*. Automatica, Volume 31, No.12, 1835-1851, December, 1995.
- [1125] Viberg, Mat; Wahlberg, Bo and Ottersten, Björn. *Analysis of State Space System Identification Methods Based on Instrumental Variables and Subspace Fitting*. Automatica, Volume 33, No.9, 1603-1616, September, 1997.

- [1126] Vidal, Maria Gloria Martínez. *Algoritmos Sistólicos para la Resolución de Ecuaciones Matriciales Lineales en Sistemas de Control*. Tesis Doctoral - Departamento de Sistemas Informáticos y Computación - Universidad Politécnica de Valencia, 1999.
- [1127] Vuchkov, I.N., and Boyadjieva, L.N.. *A New Algorithm Of the Recursive Generalized Least-Squares Method*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1114-1119, August, 1988.
- [1128] Xiujing, Wu; Zhang, Zhongjun, Zhang and Yuzhong Wang. *A New On-Line Structure Selection Algorithm For MIMO State Space Models*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 997-1002, August, 1988.
- [1129] Xu, Hong-guo. *Properties of a Quadratic Matrix Equation and the Solution of The Continuous-Time Algebraic Riccati Equation*. Linear Algebra and Its Applications, 222, 127-145, 1995.
- [1130] Xu, Guanghan. *Fast Subspace Decomposition and Applications*. Ph.D. Thesis Department of Electrical Engineering of Stanford University, September, 1991.
- [1131] Wah, Benjamin W.; Li, Guo-jie and Yu, Chee Fen.. *Multiprocessing of Combinatorial Search Problems*. Computers, Volume 18, No.6, 93-108, June, 1985.
- [1132] Wahlberg, Bo; and Jansson, Magnus. *4SID Linear Regression*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2858-2863, December, 1994.
- [1133] Walker, Homer F.. *Implementation of the GMRES Method Using Householder Transformations*. SIAM - Scientific and Statistical Computing, Vol.9 , No.1, 152-163, January, 1988.
- [1134] Wampler, Roy H.. *Solutions to Weighted Least Squares Problems by Modified Gram-Schmidt With Iterative Refinement*. ACM Transactions On Mathematical Software, Vol.5 , No.4, 457-465, December, 1979.
- [1135] Wang, Jinchun; Wada, Kiyoshi and Sun, Youxian. *Adaptive Genetic Algorithms For FFN-Based Dynamic Identification*. 11th IFAC Symposium on System Identification, Vol.3, 1371-1375, July, 1977.
- [1136] Wang, Le Yi and Lin, Lin. *Identification and Adaptive Stabilization of Slowly Varying Systems : An  $H^\infty$  Approach*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4172-4177, December, 1994.
- [1137] Wang, Liang; Libert, Gaëtane and Manneback, Pierre. *Kalman Filter Algorithm Based On Singular Value Decomposition*. Proceedings of the 31st Conference on Decision and Control, Tucson, Arizona, 1224-1229, December, 1992.
- [1138] Wang; Li-Xin and Mendel, Jerry M.. *Parallel Structured Networks For Solving a Wide Variety of Matrix Algebra Problems*. Journal of Parallel and Distributed Computing, Vol.14, No.3, 236-247, March, 1992.
- [1139] Wang, Xiaochun G.Ljung, Dumont Guy A., and Davies, Michael S.. *Modelling and Identification of Basic Weight Variations in Paper Machines*. IEEE Transactions on Control Systems Technology, Vol.1, NO.4, 230-237, December, 1993.
- [1140] Ward, Robert C.. *The QR Algorithm and Hyman's Method on Vector Computer*. Mathematics of Computation, Volume 30, No.133, 132-142, January, 1976.
- [1141] Wasan , M. T.. *Stochastic Approximation*. Cambridge At The University Press, 1969.
- [1142] Wassyng, A.. *Solving  $Ax = b$  : A Method With Reduced Storage Requirements*. SIAM Journal On Numerical Analysis, Vol.19, No.1, 197-204, February, 1982.
- [1143] Watkins, David S.. *Understanding The QR Algorithm*. SIAM Review, Vol.24, No.4, 427-440, October, 1982.

- [1144] Watkins, David S.. *Fundamentals of Matrix Computations*. John Wiley & Sons, 1991.
- [1145] Watkins, David S. *Shifting Strategies for The Parallel QR Algorithm*. Siam - Scientific Computing, Vol.15, N.4 953-958, July 1994
- [1146] Waugaman, William A. and Schrader, Cheryl B.. *Optimal Current Model From Surface Electrodes*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 4112-4113, December, 1994.
- [1147] Weiland, Siep. *A Method of Optimal System Identification With Applications in Control*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2888-2893, December, 1994.
- [1148] Weiss, Leonard. *Controllability, Realization and Stability of Discrete-Time Systems*. SIAM Journal on Control, Vol. 10, No.4, 230-264, November, 1972.
- [1149] Welch, Greg and Bishop, Gary. *An Introduction to the Kalman Filter*. TR 95-041, Department of Computer Science, University of North Carolina at Chapel Hill, p.16, 1995.
- [1150] Wentzell, Peter D.; Andrews, Darren T.; Hamilton, David C.; Faber, Klaas and Kowalski, Bruce R.. *Maximum likelihood principal component analysis*. Journal of Chemometrics, Volume 11, Issue 4, 339-366, 1997.
- [1151] Wernstedt, J., Otto, P., Puhlmann, R. and Rob, F.. *DIOPRAN- Expert - A Consulting/Expert Systems for Experimental Process Analysis*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 281-286, July, 1991.
- [1152] Wertz, Vicent; Gevers, Michel and Hannan, Edward J.. *The Determination of Optimum Structures for the State Space Representation of Multivariate Stochastic Processes*. IEEE Transactions on Automatic Control, Vol.C-27, No.6, 1200-1211, December, 1982.
- [1153] Westwick, David Thomas. *Methods for the Identification of Multiple-Input Nonlinear Systems*. PhD thesis, Department of Electrical and Biomedical Engineering Department , McGill University, Montreal, June, 1995.
- [1154] White, B.A. and El-Mahy, M.K.. *Parallel Implementations of Kalman Filters for Tracking Applications Onto Transputers Networks*. Proceedings of the Third IEEE Conference on Control Applications - Glasgow, Scotland, UK, 1073-1078, August, 1994.
- [1155] Wiener, Nobert. *Extrapolation, interpolation, and smoothing of stationary time series, with engineering applications*. New York, Wiley, 1949.
- [1156] Wieringa, Jakob Edo. *Statistical Process Control For Serially Correlated Data*. PhD. Thesis Rijksuniversiteit Groningen, Netherlands,, Netherlands, February , 1999.
- [1157] Wijshoff, Harry A.G. and Leeuwen, Jan van.. *The Structure of Periodic Storage Schemes For Parallel Memories*. IEEE Transactions on Computers, Vol.C-34, No.6, 501-505, June, 1985.
- [1158] Wilfert, H.-H. and Bretthauer, G. *Identification Aspects In Large Systems*. Proceedings of the IFAC/IFORS Symposium - Large Scale Systems: Theory and Applications 1983 - Warsaw, Poland, 405-410, July, 1983.
- [1159] Wilkinson, J. H.. *The Algebraic Eigenvalue Problem*. Oxford Science Publications, 1965.
- [1160] Willems, Jan C.. *Least Square Stationary Optimal Control and The Algebraic Riccati Equation*. IEEE Transactions on Automatic Control, Vol.AC-16, No.6, 621-634, December, 1971.
- [1161] Willems, Jan C.. *Input-Output and State-Space Representations of Finite-Dimensional Linear Time-Invariant Systems*. Linear Algebra and Its Applications, Vol.50, 581-608, April, 1983.
- [1162] Willems, Jan C.. *From Time Series to Linear System - Part I. Finite Dimensional Linear Time Invariant Systems*. Automatica, Volume 22, No.5, 561-580, September, 1986.



- [1163] Willems, Jan C.. *From Time Series to Linear System - Part II. Exact Modelling*. Automatica, Volume 22, No.6, 675-694, November, 1986.
- [1164] Willems, Jan C.. *From Time Series to Linear System - Part III. Approximate Modelling*. Automatica, Volume 23, No.1, 87-115, January, 1987.
- [1165] Willems, Jan C. (Ed.). *From Data to Model*. Springer-Verlag, 1989.
- [1166] Williams, D.; Yousefpour, P. and Swanick, B.H.. *Online Adaptive Control of a Fermentation Process*. IEE Proceedings, Vol.131, Pt.D, No. 4, 117-124, July, 1984.
- [1167] Wimmer, Harald K.. *Normal Forms of Symplectic Pencils and the Discrete-Time Algebraic Riccati Equation*. Linear Algebra and Its Applications, Volume 147, 411-440, March, 1991.
- [1168] Wolf, Sjoerd de and Hof, Paul M.J. Van den.. *Identification of a Pilot Crystallization Process With Output Error Methods*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1208-1213, July, 1991.
- [1169] Wolovich, W.A.. *The Determination of State-Space Representations for Linear Multivariable Systems*. Automatica, Volume 9, No.1, 97-106, January, 1973.
- [1170] Wong, Kwan Y. and Polak, Elijah. *Identification of Linear Discrete Time Systems Using The Instrumental Variable Method*. IEEE Transactions on Automatic Control, Vol.AC-12, No.6, 707-718, December, 1967.
- [1171] Wonham, W.M.. *Geometric State-Space Theory in Linear Multivariable Control : A Status Report*. Automatica, Volume 15, No.1, 5-13, January, 1979.
- [1172] Wonham, W.M.. *Linear Multivariable Control : A Geometric Approach*. Springer-Verlag, Third Edition, 1985.
- [1173] Wu, Wen-Teng, Chu, Y.T. and Chen, Kuo-Chieh. *Weighted Moving Identification and Control for Multivariable Systems*. IEE Proceedings, Vol.135, Pt.D, No. 3, 196-204, May, 1988.
- [1174] Wen, Wu; Rutan, Sarah C.; Baldovin, Antonietta and Massart, Désiré-Luc. *Feature selection using the Kalman filter for classification of multivariate data*. Analytica Chimica Acta Volume: 335, Issue: 1-2, pp. 11-22, December 10, 1996.
- [1175] Yamanoi, M. ; Naganawa, H. ; Sobajima, T. and Suzuki, T.. *Relaxation of the SPR-condition in the CGT method for the Input Tracking of Superposition of Sinusoids*. Proceedings of 3rd European Control Conference, Rome , Italy, 2125-2130, September, 1995.
- [1176] Yang, Ciann-Dong and Yeh, Fang-Bo. *An Efficient Algorithm on Optimal Hankel-Norm Approximation for Multivariable Systems*. IEEE Transactions on Automatic Control, Vol.37, No.6, 815-820, June, 1992.
- [1177] Yang, Hua and Hua Yingbo. *Statistical analysis of an eigendecomposition based method for 2-D frequency estimation*. Automatica, Volume 30, No.1, 157-168, January, 1994.
- [1178] Yeh, Fang-Bo and Yang, Ciann-Dong. *Singular Value Decomposition Of An Infinite Block-Hankel Matrix and Its Applications*. Numerical Functional Analysis and Optimization, Volume 9, No.7 & 8, 881-916, 1987.
- [1179] Yohai, V. J. and Ben, M.S. Garcia. *Canonical Variables As Optimal Predictors*. The Annals of Statistics, Volume 8, No.4, 865-869, July, 1980.
- [1180] York, Derek. *Least-Squares Fitting of A Straight Line*. Canadian Journal of Physics, Volume 44, No.5, 1079-1086, May, 1966.
- [1181] Youmin, Zhang, Quan, Pan, Zhang, Hongcai and Guanzhong, Dai. *A Parallel Decoupled Kalman Filtering Algorithm and Systolic Architecture*. Proceedings of the 32th Conference on Decision and Control, San Antonio, Texas, 3590-3595, December, 1993.

- [1182] Young, David M.. *Iterative Solution Of Large Linear Systems*. Academic Press, 1971.
- [1183] Young, N.J.. *The Singular-Value Decomposition of an Infinite Hankel Matrix*. Linear Algebra and Its Applications, Vol.50, 639-657, April, 1983.
- [1184] Young, Peter. *Parameter Estimation for Continuous-Time Models - A Survey*. Automatica, Volume 17, No. 1, 23-39, January 1981.
- [1185] Young, Peter. *Recursive Estimation and Time-Series Analysis*. Springer-Verlag, 1984.
- [1186] Young, Peter C.. *Simplified Refined Instrumental Variable (SRIV) Estimation and True Digital Control (TDC)*. Proceedings of the First European Control Conference - Grenoble, France, 1295-1306, July, 1991.
- [1187] Yu, Banmei; Li, Menglong, Liu, Aiping ; Li, Zhiliang; Shi, Leming and Pan, Zhongxiao. *Some Novel Methods Based on Recursive Optimal Estimation. Applications to Analytical Chemistry* . Analytica Chimica Acta - Volume 277, No.2, 199-204, May, 1993.
- [1188] Yu, Xiao-Hu and He, Zhen-ya. *Efficient Recursive Instrumental Variable Algorithms For Parameter Identification*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1100-1104, August, 1988.
- [1189] Yu, Xiaode and Verhaegen, Michel. *An Algorithm of Subspace Model Identification for Descriptor Systems*. Proceedings of 3rd European Control Conference, Rome , Italy, 417-422, September, 1995.
- [1190] Yufei, Xiong and Yuanzhang, Lu. *Identification of Multivariable Hydraulic Servo System and Noise Model*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1546-1550, August, 1988.
- [1191] Zadeh, Lotfi A. and Desoer, Charles A.. *Linear System Theory*. McGraw-Hill Book Company, 1963.
- [1192] Zakharov, Vasilli. *Parallelism and Array Processing*. IEEE Transactions on Computers, Vol.C-33, No.1, 45-78, January 1984.
- [1193] Zargham, Mehdi R.. *Computer Architecture - Single and Parallel Systems*. Prentice-Hall, 1996.
- [1194] Zarowski, Christopher J.. *A Schur Algorithm and Linearly Connected Processor Array for Toeplitz-plus-Hankel Matrices*. IEEE Transactions on Signal Processing, Vol.40, No.8, 2065-2078, August, 1992.
- [1195] Zeiger, H. Paul and McEwen, A.Julia. *Approximate Linear Realizations of Given Dimension via Ho's Algorithm*. IEEE Transactions on Automatic Control, Vol.19, 153, April, 1974.
- [1196] Zhang, Qinghu; Basseville, M. and Benveniste, A.. *Early warning of slight changes in systems*. Automatica, Volume 30, No.1, 95-113, January, 1994.
- [1197] Zhang, Youmin; Qingguo, Li ; Dai, Guanzhong and Zhang, Hongcai. *A New Recursive Least-Squares Identification Algorithm Based On Singular Value Decomposition*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 1733-1740, December, 1994.
- [1198] Zhang, Qinghua. *Using Wavelets in Nonparametric Estimation*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 3406-3411, December, 1994.
- [1199] Zhao, Zhikui; Ma, Chuan; Wang, Liqi; Han, Chongzhao and Wan, Baiwu. *CMAC Variable Structure Control of Three-Axis Simulation Turn-Table*. 11th IFAC Symposium on System Identification, Vol.2, 757-761, July, 1977.
- [1200] Zheng, W.X.. *Unbiased Identification of Multivariable Systems Subject To Colored Noise*. Proceedings of the 33rd Conference on Decision and Control, Lake Buena Vista, Florida, 2864-2865, December, 1994.

- [1201] Zheng, Wei Xing. *Noisy Input-Output System Identification Using The Least-Squares Based Algorithms*. Proceedings of the 37th IEEE Conference on Decision & Control, Tampa, Florida, USA, WM07-6, December, 1998.
- [1202] Zhou, Tong and Kimura Hidenori. *Simultaneous Identification of Nominal Model, Parametric Uncertainty and Unstructured Uncertainty For Robust Control*. Automatica, Volume 30, No.3, 391-402, March, 1994.
- [1203] Zhou, Wei-wu. *Linear and NonLinear Recursive Prediction Error Methods In State Space Models*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1092-1099, August, 1988.
- [1204] Zhu, Hai-An. *A Novel Method of Model Reduction Based On System Parameter Identification*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 910-915, August, 1988.
- [1205] Zhu, Y.C. and Backx, A.C.P.M.. *MIMO Process Identification for Controller Design : Test Signals, Nominal Model and Error Bounds*. Proceedings of the 9th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1991 - Budapest, Hungary, 1202-1207, July, 1991.
- [1206] Zhu, Yue-Xin and Wan, Bai-Wu. *The Parameter Identification of a Population Model of China*. Automatica, Volume 20, No.4, 415-428, July, 1984.
- [1207] Zhu, Guoming; Skelton, Robert E. and Li, Pingkang. *Q-Markov Cover Identification Using Pseudo Random Binary Signals*. ACC, 1-24, 1994 - (artigo obtido através do e-mail : skelton@ecn.purdue.edu ).
- [1208] Zomaya, A.Y.. *Extraction and Computation of Identifiable Parameters in Robot Dynamic Models : Theory and Application*. IEE Proceedings Control Theory and Applications, Vol. 141, No.1, 48-56, January, 1994.
- [1209] Zu-Shu, Li, Qi-Jian, Zhou and Ming, Xu. *Characteristic Identification, Characteristic Memory and Intelligent Controller*. Proceedings of the 8th IFAC/IFORS Symposium - Identification and System Parameter Estimation 1988 - Beijing, China, 1686-1690, August, 1988.