



UNIVERSIDADE ESTADUAL DE CAMPINAS
Faculdade de Engenharia Elétrica e de Computação

Henrique Evangelista de Oliveira

Identificação dos Parâmetros da Integral de Choquet via uma Abordagem baseada em Processamento de Sinais Esparsos

Campinas

2020

Henrique Evangelista de Oliveira

Identificação dos Parâmetros da Integral de Choquet via uma Abordagem baseada em Processamento de Sinais Esparsos

Tese apresentada à Faculdade de Engenharia Elétrica e de Computação da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Doutor em Engenharia Elétrica, na Área de Telecomunicações e Telemática.

Orientador: Prof. Dr. João Marcos Travassos Romano

Coorientador Prof. Dr. Leonardo Tomazeli Duarte

Este trabalho corresponde à versão final da tese defendida pelo aluno Henrique Evangelista de Oliveira, e orientada pelo Prof. Dr. João Marcos Travassos Romano.

Campinas

2020

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca da Área de Engenharia e Arquitetura
Rose Meire da Silva - CRB 8/5974

OL4i Oliveira, Henrique Evangelista de, 1987-
Identificação dos parâmetros da integral de Choquet via uma abordagem baseada em processamento de sinais esparsos / Henrique Evangelista de Oliveira. – Campinas, SP : [s.n.], 2020.

Orientador: João Marcos Travassos Romano.
Coorientador: Leonardo Tomazeli Duarte.
Tese (doutorado) – Universidade Estadual de Campinas, Faculdade de Engenharia Elétrica e de Computação.

1. Estimativa de parâmetro. 2. Esparsidade. I. Romano, João Marcos Travassos, 1960-. II. Duarte, Leonardo Tomazeli, 1982-. III. Universidade Estadual de Campinas. Faculdade de Engenharia Elétrica e de Computação. IV. Título.

Informações para Biblioteca Digital

Título em outro idioma: Identification of the Choquet integral parameters by means of sparse modeling

Palavras-chave em inglês:

Parameter estimation

Sparsity

Área de concentração: Telecomunicações e Telemática

Titulação: Doutor em Engenharia Elétrica

Banca examinadora:

João Marcos Travassos Romano [Orientador]

Cristiano Torezzan

Rosângela Ballini

Marcello Luiz Rodrigues de Campos

Ricardo Suyama

Data de defesa: 18-09-2020

Programa de Pós-Graduação: Engenharia Elétrica

Identificação e informações acadêmicas do(a) aluno(a)

- ORCID do autor: <https://orcid.org/0000-0002-9916-1562>

- Currículo Lattes do autor: <http://lattes.cnpq.br/3731642614007224>

COMISSÃO JULGADORA - TESE DE DOUTORADO

Candidato(a): Henrique Evangelista de Oliveira RA: 061419

Data de defesa: 18 de setembro de 2020

Titulo da Tese: "Identificação dos Parâmetros da Integral de Choquet via uma Abordagem baseada em Processamento de Sinais Esparsos"

Prof. Dr. João Marcos Travassos Romano (Presidente)

Prof. Dr. Cristiano Torezzan

Profa. Dra. Rosângela Ballini

Prof. Dr. Marcello Luiz Rodrigues de Campos

Prof. Dr. Ricardo Suyama

A Ata de Defesa, com as respectivas assinaturas dos membros da Comissão Julgadora, encontra-se no SIGA (Sistema de Fluxo de Dissertação/Tese) e na Secretaria de Pós-Graduação da Faculdade de Engenharia Elétrica e de Computação.

Dedico esta tese àquela que faz meus dias mais felizes, Suzana.

Agradecimentos

À Deus, principio e criador de tudo, que me fortalece e me ajuda a seguir em frente.

À minha amada esposa, Suzana, que não poupa esforços para cuidar de nossa família.

Aos meus queridos filhos: Antonio Álvaro, Tomás Henrique, José Fernando e Isadora Maria. A alegria de uma casa cheia compensa todas as dificuldades que se apresentam.

Aos meus pais e irmão, Jorge, Isabel e Lucas, por sempre me apoiarem e ajudarem.

Ao meu amigo e orientador, João Marcos Travassos Romano, pelos valiosos conselhos e exemplos passados ao longo desse trabalho.

Ao meu amigo e coorientador Leonardo Tomazeli Duarte, pelas horas dedicadas no enriquecimento desse trabalho.

Aos professores do doutorado em Engenharia Elétrica, pelos ensinamentos essenciais para o desenvolvimento desse trabalho.

Aos professores da banca, Cristiano Torezzan, Rosângela Ballini, Marcello Luiz Rodrigues de Campos e Ricardo Suyama por se disporem a colaborar imensamente nesse trabalho.

Aos amigos da UNICAMP.

Aos que compareceram virtualmente à defesa.

À toda minha família e aos demais amigos, que torcem por mim.

Ao amigo e compadre Toninho (in memoriam), que sempre me incentivou a terminar esse trabalho e, certamente, nesse momento intercede do céu para que eu o consiga.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

"Amanhã! Algumas vezes é prudência; muitas vezes, é o advérbio dos vencidos." (Josemaría Escrivá de Balaguer)

Resumo

A integral de Choquet tem sido usada como um operador de agregação na área de tomada de decisão multicritério. Devido à sua natureza não linear, ela pode modelar interações entre diferentes critérios, como sinergia e redundância. No entanto, a identificação dos parâmetros da integral de Choquet é um problema desafiador devido à sua natureza mal-posta. Em trabalhos recentes, esse problema foi abordado considerando termos de regularização baseados em esparsidade. Neste trabalho, essa abordagem também é considerada. Porém, diferentemente de estudos anteriores, nos quais a integral de Choquet é parametrizada por meio de uma medida *fuzzy*, exploramos aqui a esparsidade em um domínio transformado conhecido como representação de índice de interação. Um conjunto de experimentos numéricos é apresentado para comprovar que a utilização da hipótese de esparsidade se mostra mais adequada quando o domínio do índice de interação é considerado, sendo essa a primeira e principal contribuição da tese. A segunda contribuição dessa tese é a análise de unicidade, que mostra sob quais condições se terá uma solução única para o problema de identificação das capacidades de Choquet. Também é apresentado um conjunto de experimentos numéricos para verificar a unicidade em diferentes situações.

Palavras-chaves: Integral de Choquet; Identificação de Parâmetros; Medida *Fuzzy*; Esparsidade; *Compressive Sensing*.

Abstract

The Choquet integral has been used as an aggregation operator in the field of multiple criteria decision aiding. Due to its nonlinear nature, the Choquet integral can model interactions between different criteria, such as synergy and redundancy. However, the identification of the Choquet integral parameters is a challenging problem due to its ill-posed nature. In recent works, this problem has been addressed by considering regularization terms based on sparsity. In this work, this approach is also considered. However, differently from previous studies, in which the Choquet integral is parametrized by means of a fuzzy measure, we here exploit sparsity in a transformed domain known as interaction index representation. A set of numerical experiments is presented to attest that sparsity hypothesis is more appropriate when considered in the interaction index domain. This is the first and main contribution of the thesis. The second contribution of this thesis is the identifiability analysis, that shows the conditions for what the identification problem has unique solution. Also, a set of numerical experiments is presented to verify uniqueness in different situations.

Keywords: Choquet Integral; Parameter Identification, Fuzzy Measures, Sparsity, Compressive Sensing.

Lista de ilustrações

Figura 2.1 – Sistema mecânico.	20
Figura 2.2 – Conjunto com 5 alternativas.	21
Figura 2.3 – Problemática de escolha.	21
Figura 2.4 – Problemática de classificação.	22
Figura 2.5 – Problemática de ordenamento.	22
Figura 2.6 – Processo de agregação.	23
Figura 2.7 – Interpretação geométrica para a agregação linear e para a agregação feita utilizando a integral de Choquet.	32
Figura 3.1 – Vetor esparso.	41
Figura 3.2 – Indicador para o cálculo da norma l_0	42
Figura 3.3 – Lugares geométricos de raio unitário no \mathbb{R}^2 para as normas l_p com $p = 1, 2, \infty$, e para a quasi-norma l_p com $p = \frac{1}{2}$. Extraído de (ELDAR; KUTYNIOK, 2012a).	43
Figura 3.4 – Solução de um problema inverso através da minimização das normas l_p com $p = 1, 2, \infty$, e da quasi-norma l_p com $p = \frac{1}{2}$. Adaptado de (ELDAR; KUTYNIOK, 2012a).	44
Figura 4.1 – Valores das capacidades de Choquet para o caso aditivo.	48
Figura 4.2 – Valores dos índices de interação para o caso aditivo.	51
Figura 4.3 – Esquema representativo dos experimentos realizados.	53
Figura 4.4 – Caso <i>1-aditivo</i> : Dispersão bi-dimensional entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	56
Figura 4.5 – Caso <i>1-aditivo</i> : Valores das capacidades de Choquet identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	57
Figura 4.6 – Caso <i>1-aditivo</i> : Valores dos índices de interação generalizados identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	58
Figura 4.7 – Caso <i>2-aditivo</i> : Dispersão bi-dimensional entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	60
Figura 4.8 – Caso <i>2-aditivo</i> : Valores das capacidades de Choquet identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	61

Figura 4.9 – Caso <i>2-aditivo</i> : Valores dos índices de interação generalizados identifi- cados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização. .	62
Figura 5.1 – Número de vezes que o critério do <i>posto</i> é satisfeito (considerando um cenário com 1000 simulações), no cenário 4, para diferentes números de alternativas e para: (a) 3 critérios e (b) 4 critérios.	70
Figura 5.2 – Número de vezes que o critério do <i>spark</i> é satisfeito, no cenário 4, para diferentes números de alternativas e para: (a) 3 critérios e (b) 4 critérios.	72

Lista de tabelas

Tabela 2.1 – Ordenação dos estudantes pela média ponderada.	26
Tabela 2.2 – Ordenação dos estudantes por meio da integral de Choquet.	34
Tabela 4.1 – Caso 1-aditivo: MSE e MAE entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	55
Tabela 4.2 – Caso 1-aditivo: MSE e MAE entre os valores teóricos das capacidades de Choquet e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	56
Tabela 4.3 – Caso 1-aditivo: MSE e MAE entre os valores teóricos dos índices de interação generalizados e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	57
Tabela 4.4 – Caso 2-aditivo: MSE e MAE entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	59
Tabela 4.5 – Caso 2-aditivo: MSE e MAE entre os valores teóricos das capacidades de Choquet e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	60
Tabela 4.6 – Caso 2-aditivo: MSE e MAE entre os valores teóricos dos índices de interação generalizados e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.	61
Tabela 5.1 – Valores do <i>posto</i> para os quatro cenários para o caso com 3 critérios e 8 alternativas.	69
Tabela 5.2 – Valores do <i>spark</i> para os quatro cenários para o caso com 3 critérios e 6 alternativas.	71

Sumário

1	INTRODUÇÃO	15
2	MÉTODOS DE APOIO À DECISÃO MULTICRITÉRIO BASEADOS NA INTEGRAL DE CHOQUET	19
2.1	Métodos de apoio à decisão multicritério	19
2.2	O processo de agregação	23
2.3	A integral de Choquet	26
3	IDENTIFICAÇÃO DAS CAPACIDADES DE CHOQUET	35
3.1	Estimação dos parâmetros da integral de Choquet	35
3.1.1	Aprendizado supervisionado da integral de Choquet	36
3.1.1.1	Problemas inversos	39
3.2	Regularização do problema mal-posto	40
3.2.1	Soluções Esparsas	41
3.2.2	Otimização do problema com um termo de regularização	44
4	IDENTIFICAÇÃO NO DOMÍNIO DO ÍNDICE DE INTERAÇÃO	47
4.1	Motivação para trabalhar no domínio do índice de interação	47
4.2	Índice de interação generalizado	48
4.3	Regularização esparsa no domínio do índice de interação	51
4.4	Problema de identificação: resultados numéricos	53
4.4.1	Problema de identificação: resultados para o caso <i>1-aditivo</i>	54
4.4.2	Problema de identificação: resultados para o caso <i>2-aditivo</i>	56
4.5	Sumário	61
5	ANÁLISE DE UNICIDADE	64
5.1	Análise de unicidade das soluções	64
5.1.1	Análise de unicidade para sistemas sobredeterminados	65
5.1.2	Análise de unicidade para sistemas subdeterminados	65
5.2	Análise da unicidade: resultados numéricos	67
5.2.1	Resultados numéricos para sistemas sobredeterminados	68
5.2.2	Resultados numéricos para sistemas subdeterminados	69
5.3	Sumário	72
6	CONCLUSÃO	74

REFERÊNCIAS 76

1 Introdução

Uma das problemáticas no apoio à tomada de decisão é estabelecer, a partir de um conjunto de informações, uma ordem entre um conjunto de alternativas. Quando esse processo de ordenamento deve ser feito levando em conta diferentes critérios, o problema resultante pode ser avaliado através dos métodos de apoio à decisão multicritério (MCDA, do inglês *multiple criteria decision aiding*) (BOUYSSOU et al., 2006; GRECO; FIGUEIRA; EHRGOTT, 2005). Essas técnicas encontram aplicações em muitos domínios, como finanças e telecomunicações (GRECO; FIGUEIRA; EHRGOTT, 2005). Por exemplo, em (DEHE; BAMFORD, 2015), dois métodos de MCDA foram usados para abordar um problema de decisão relacionado à localização de hospitais, clínicas e centros de saúde. Em (XU; OUENNICHE, 2012), três métodos de MCDA foram aplicados para ordenar modelos de previsão dos preços do petróleo bruto.

Uma das abordagens possíveis para a problemática de ordenamento em MCDA é definir um operador de agregação. Dada uma alternativa e um conjunto de avaliações dessas alternativas com relação a diferentes critérios, o objetivo do operador de agregação é combinar essas avaliações em uma única avaliação. Então, esses valores agregados podem ser usados para estabelecer uma ordem entre as alternativas consideradas (GRABISCH, 1996). O exemplo mais fundamental dos operadores de agregação é a média ponderada, que possui uma natureza linear. Embora sejam simples em termos de implementação, os operadores de agregação linear trabalham sob a hipótese de que não há qualquer tipo de relação entre os diferentes critérios escolhidos. Tal hipótese pode ser vista como uma simplificação, pois em muitas aplicações práticas, pode haver interações entre os critérios. Nesse cenário, a descrição correta de comportamentos racionais de decisão pode requerer a presença de elementos não-lineares no modelo considerado (GRECO; FIGUEIRA; EHRGOTT, 2005).

Um dos operadores não-lineares de agregação que vem sendo intensamente utilizado em MCDA é a integral de Choquet (GRABISCH, 1996; GRABISCH, 1997). Esse operador é capaz de representar diferentes tipos de interações entre critérios, tais como sinergia e redundância. Além disso, a integral de Choquet pode ser parametrizada por variáveis que admitem interpretações interessantes no contexto de tomada de decisão (GRABISCH, 1996; GRABISCH, 1997; GRABISCH; LABREUCHE, 2019). Uma parametrização muito útil da integral de Choquet está baseada na noção de medidas *fuzzy* (ou capacidades de Choquet), as quais associam uma medida não aditiva a cada critério e a cada coalizão de critérios.

Um aspecto central no uso da integral de Choquet em MCDA consiste no problema da estimação das medidas *fuzzy* associadas com a integral, problema esse que

também é conhecido como identificação das capacidades de Choquet. Há muitos métodos para estimar os parâmetros da integral de Choquet e eles podem ser divididos em três abordagens: não-supervisionada, semi-supervisionada e supervisionada. Na abordagem supervisionada, que é o foco desse trabalho, se conhecem os dados de entrada e saída e se deseja obter as capacidades de Choquet. Tem-se então um problema inverso, que pode ser resolvido determinando-se uma função custo para aferir a qualidade da representação fornecida pelo modelo para explicar as saídas observadas. Portanto, a identificação das capacidades é baseada em problemas de otimização (GRABISCH; KOJADINOVIC; MEYER, 2008; BELIAKOV; JAMES; WU, 2019).

Uma das dificuldades na identificação das capacidades é a de que o problema inverso resultante é mal-posto. Portanto, a solução obtida após a otimização pode não ser única, ou seja, diferentes conjuntos de capacidade de Choquet podem resultar num mesmo valor após a etapa de agregação. Tal dificuldade pode limitar o uso da integral de Choquet, especialmente em aplicações nas quais a interpretação dos parâmetros é útil, tais como aprendizado de preferência e regressão Choquística (TEHRANI et al., 2012a). Além disso, problemas inversos mal-postos podem levar a desafios relacionados às soluções numéricas dos problemas de otimização resultantes (ASTER; BORCHERS; THURBER, 2012).

Uma abordagem comum para resolver problemas inversos mal-postos é considerar regularização. Nessa abordagem, em adição a uma função de perda, tal como mínimos quadrados, a função custo a ser otimizada abrange um termo de regularização que é uma função dos parâmetros do modelo. Na regularização de Tikhonov (ASTER; BORCHERS; THURBER, 2012), por exemplo, o termo de regularização é dado pelo quadrado da norma Euclidiana dos parâmetros do modelo. Outra maneira muito popular de se considerar regularização é assumindo que os parâmetros do modelo são dados por um vetor esparso, no qual a maioria dos elementos são zero (ou muito próximos de zero) (PLUMBLEY et al., 2010; STARCK; MURTAGH; FADILI, 2010). Regularização baseada em esparsidade tem se mostrada uma abordagem eficiente em diversas subáreas de ciência de dados.

Em trabalhos recentes, a regularização esparsa foi explorada no problema de identificação das capacidades de Choquet (ANDERSON; PRICE; HAVENS, 2014; ADEYEBA; ANDERSON; HAVENS, 2015). Na presente tese, nós também investigamos o uso de regularização esparsa (BARANIUK, 2007) para a identificação dos parâmetros da integral de Choquet. Entretanto, diferente do trabalho apresentado em (ADEYEBA; ANDERSON; HAVENS, 2015), que considera a esparsidade no domínio da medida *fuzzy*, nós exploramos a esparsidade em outro domínio, conhecido como representação de índice de interação (GRABISCH, 2000). De fato, como será discutido ao longo do texto, assumir a esparsidade no domínio do índice de interação é mais natural no contexto de decisão, do que considerá-la no domínio das medidas *fuzzy*.

Uma segunda contribuição desse trabalho está relacionada com aspectos de

unicidade do problema de identificação das capacidades. Um modelo possui solução única quando os dados observados podem ser explicados por um conjunto único de parâmetros do modelo. Tal propriedade é importante em problemas de tomada de decisão pela integral de Choquet, pois os parâmetros do modelo nesse caso podem prover interpretações relevantes de cada critério, assim como das coalizões de critérios (GRABISCH; LABREUCHE, 2019).

Neste trabalho, as análises de unicidade do problema de identificação das capacidades são abordadas para duas situações. A primeira situação aborda o caso sobre-determinado, no qual o número de observações disponíveis, que no nosso problema corresponde ao número de alternativas, é maior que o número de parâmetros do modelo. Para esta situação, nossa análise se baseia no cálculo do posto da matriz de decisão aumentada, em que cada coluna está relacionada a um parâmetro da integral de Choquet. A segunda situação aborda o caso subdeterminado, no qual o número de observações é menor que o número de parâmetros do modelo. Nesse caso, se torna necessário recorrer a abordagens alternativas, como o cálculo do *spark* da matriz de decisão aumentada (ELAD; FIGUEIREDO; MA, 2010). Tal abordagem teve sua origem na área de *compressed sensing* e tem sido aplicado em problemas mais gerais de representação esparsa (ELDAR; KUTYNIOK, 2012b).

A tese está organizada da seguinte maneira:

- Capítulo 2: Métodos de apoio à decisão multicritério baseados na integral de Choquet
Neste capítulo é feita uma breve revisão sobre Métodos de Apoio à Decisão Multicritério, dando ênfase ao processo de agregação. Também é apresentada a integral de Choquet, assim como suas propriedades.
- Capítulo 3: Identificação das capacidades de Choquet
Neste capítulo é apresentado o problema de estimação dos parâmetros da integral de Choquet. Mais especificamente, é apresentada a abordagem utilizada nesse trabalho, que é a identificação supervisionada das capacidades da integral de Choquet, assim como uma maneira de se regularizar o problema por meio da norma l_1 , que leva a uma solução esparsa.
- Capítulo 4: Identificação no domínio do índice de interação
Neste capítulo é apresentada a motivação para trabalhar no domínio do índice de interação, assim como o problema de identificação dos parâmetros da integral de Choquet, considerando agora esse novo domínio. Então, são apresentados os resultados numéricos que mostram que a utilização da hipótese de esparsidade se mostra mais adequada quando o domínio do índice de interação é considerado, sendo essa a primeira contribuição original da tese.
- Capítulo 5: Análise de unicidade

Neste capítulo é apresentada a segunda contribuição original da tese, que é a análise que mostra sob quais condições se terá uma solução única para o problema de identificação das capacidades de Choquet. Também são apresentados resultados numéricos para diferentes situações.

- Capítulo 6: Conclusão

Neste último capítulo, encerra-se a tese com uma conclusão sobre os resultados obtidos e algumas perspectivas de trabalhos futuros.

2 Métodos de apoio à decisão multicritério baseados na integral de Choquet

Esse capítulo está dividido em três seções. Na primeira seção é feita uma breve revisão sobre Métodos de Apoio à Decisão Multicritério, sendo apresentado sua definição, os cenários e problemáticas onde se aplica. Na segunda seção é dada ênfase ao processo de agregação, que é uma das ferramentas que auxiliam na tomada de decisão. Por último, na terceira seção, é introduzida a integral de Choquet, ferramenta que será estudada por todo esse trabalho.

2.1 Métodos de apoio à decisão multicritério

Decidir tem sido objeto de reflexão de muitos pensadores desde os tempos antigos. Grandes filósofos, como Aristóteles, Platão e Tomás de Aquino discutiram a capacidade do ser humano de decidir e, de alguma maneira, disseram que esta habilidade é o que nos distingue dos animais (GRECO; FIGUEIRA; EHRGOTT, 2005).

A tomada de decisão está relacionada com a comparação de uma pluralidade de pontos de vista, os quais podem ser definidos como critérios. Quando se observa os prós e os contras de cada um desses pontos de vista, ou seja, quando se observam diferentes critérios para se tomar uma decisão, se está no domínio dos Métodos de Apoio à Decisão Multicritério (MCDA, *Multiple Criteria Decision Analysis*) (GRECO; FIGUEIRA; EHRGOTT, 2005).

Tomar uma decisão pode ser uma tarefa difícil de se realizar, pois ela é diferente de um sistema mecânico, o qual é possível modelar e se conhecem as entradas e saídas. Por exemplo, na Figura 2.1 é mostrado um sistema mecânico composto por massa, mola e amortecedor, no qual se conhece a entrada $F(t)$, que é uma força dependente do tempo, e se conhece a saída $x(t)$, que é um deslocamento dependente do tempo (SINGIRESU et al., 1995). Vê-se que, para um sistema mecânico, existe um modelo que associa a entrada e a saída e, portanto, é possível determinar qual será a consequência de uma dada ação antes mesmo dela ser realizada.

Um cenário de análise de decisão muitas vezes é complexo, o que o torna difícil de modelar. A princípio, não se tem muita certeza de quais critérios devemos levar em consideração, ou seja, não se conhecem muito bem as entradas do sistema. Consequentemente, o resultado dessa análise é, na maioria dos casos, desconhecido *a priori*. Pode-se ter três cenários em processos de tomada de decisão:

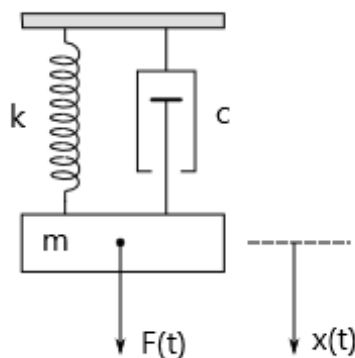


Figura 2.1 – Sistema mecânico.

1. Cenários de certeza: no qual as consequências do processo decisório são precisamente conhecidas;
2. Cenários de risco: no qual as consequências do processo decisório e suas probabilidades são conhecidas;
3. Cenários de incerteza: no qual as consequências do processo decisório não são conhecidas ou, quando são, as suas probabilidades não são conhecidas.

Frequentemente, importantes decisões devem ser feitas em cenários de risco e, até mesmo, em cenários de incerteza. Alguns exemplos simples de tomada de decisão são (BOUYSSOU et al., 2006):

1. Quando um diretor de uma escola deve escolher três alunos para receber um prêmio, dentre uma sala com trinta alunos, ele pede para os seus professores avaliarem seus alunos por meio de notas, calcula uma média dessas notas e escolhe aqueles que possuem maior média. Nesse caso, as notas são os critérios utilizados para se tomar a decisão correta;
2. Quando um banco deve decidir se um determinado cliente obterá crédito ou não, geralmente se faz uma análise de crédito, onde se analisam alguns critérios, como restrições cadastrais do cliente, comprometimento da renda do cliente, entre outros;
3. Quando se precisa fazer um ordenamento de algum curso universitário, se analisam alguns critérios como, por exemplo, quantidade de alunos ingressantes e concluintes, capacidade de produção científica, entre outros.

Devido às dificuldades inerentes ao processo de tomada de decisão, o uso de ferramentas de MCDA pode ser interessante. Para se tomar uma decisão baseada em MCDA, a qual apresenta diversas abordagens, métodos e técnicas, deve-se definir alguns

elementos básicos: um conjunto finito ou infinito de ações (alternativas, soluções...), ao menos dois critérios e ao menos um tomador de decisão. Através desses elementos, o objetivo da MCDA é auxiliar na tomada de decisão, principalmente em três problemáticas: de escolha, de classificação ou de ordenamento das ações (GRECO; FIGUEIRA; EHRGOTT, 2005). Como exemplo, considere um processo de tomada de decisão no qual podemos ter 5 alternativas, como mostrado na Figura 2.2. Quando se trabalha com a problemática de escolha, o objetivo do tomador de decisão é selecionar um subconjunto de boas alternativas, como mostrado na Figura 2.3. A Figura 2.4 mostra a problemática de classificação, na qual o objetivo do tomador de decisão é atribuir cada alternativa em uma classe mais apropriada, dentre uma família de classes pré-definidas. Por último, na problemática de ordenamento, mostrada na Figura 2.5, o objetivo do tomador de decisão é estabelecer uma ordem entre as alternativas, da seguinte maneira: $A_3 \succ A_4 \succ A_1 \succ A_2 \succ A_5$, onde o símbolo \succ denota uma relação binária relacionada à preferência estrita (GRECO; FIGUEIRA; EHRGOTT, 2005). Nos exemplos citados acima, temos um caso prático para as três problemáticas: o diretor faz uma escolha entre seus alunos, o banco classifica o cliente como apto ou não para receber o crédito e se faz um ordenamento entre os cursos universitários.

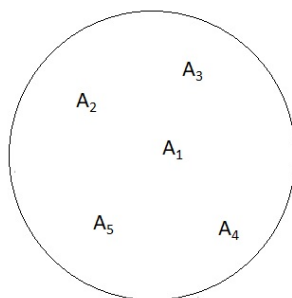


Figura 2.2 – Conjunto com 5 alternativas.

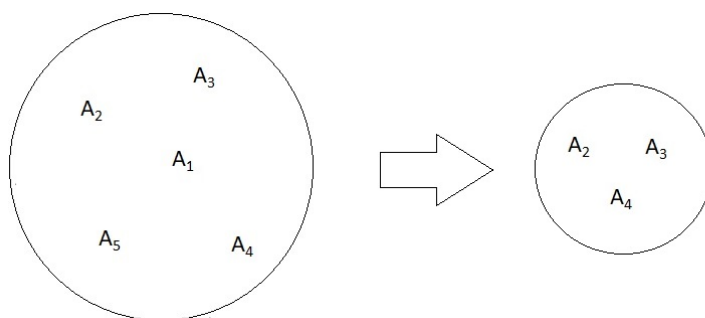


Figura 2.3 – Problemática de escolha.

O ordenamento das alternativas que é feito estabelecendo preferências entre as alternativas possíveis é conhecido como elicitacão de preferências. Essa tarefa pode não ser trivial, pois as preferências expressas pelo tomador de decisão podem ser imprecisas, conflitantes, instáveis e dependentes do tempo (MOUSSEAU; PIRLOT, 2015).

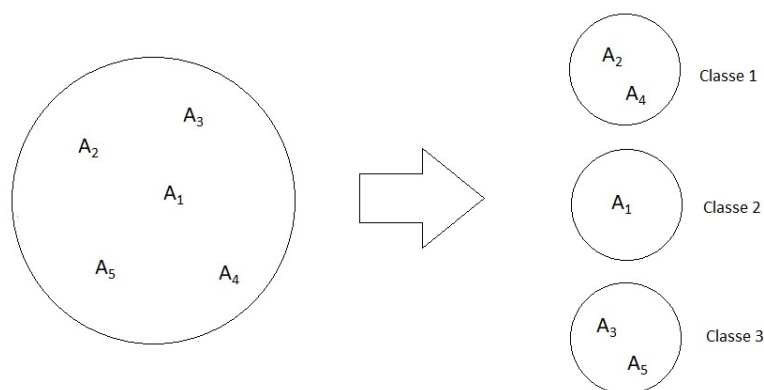


Figura 2.4 – Problemática de classificação.

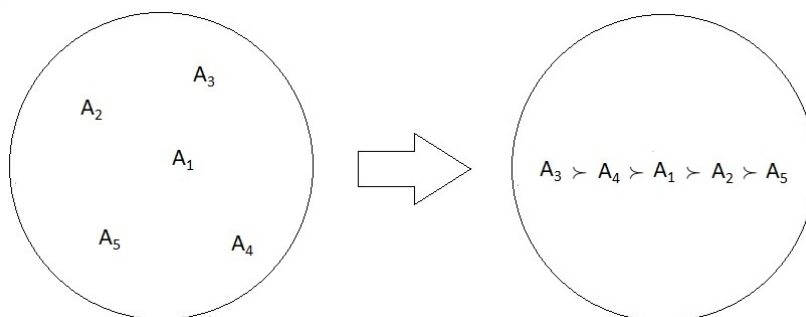


Figura 2.5 – Problemática de ordenamento.

Na eliciação de preferências, em problemas com um único critério, admite-se uma ordem naturalmente. Por exemplo, consideremos duas alternativas **a** e **b** e as comparemos levando em consideração apenas um critério. Obviamente, para escolher entre **a** ou **b** é necessário apenas verificar se **a** tem preferência sobre **b** ou vice-versa, com relação a esse único critério. Em problemas com mais de um critério não é possível admitir ordem naturalmente e o tomador de decisão recorre, portanto, ao conceito de dominância de Pareto (GRECO; FIGUEIRA; EHRGOTT, 2005), no qual **a** domina **b** se **a** tem pelo menos um critério maior que **b**, de modo que todos os outros são, no pior dos casos, iguais. Em situações práticas, porém, é comum que as alternativas sejam todas não-dominadas. Por exemplo, consideremos as mesmas duas alternativas **a** e **b** e as comparemos levando em consideração n critérios. Provavelmente, veremos que **a** terá preferência sobre **b** com relação a alguns critérios e que **b** terá preferência sobre **a** com relação a outros, portanto elas são não-dominadas. Nesse caso, se torna necessário realizar um processo de agregação dos diferentes critérios, levando em conta preferências estabelecidas sobre a relevância dos diferentes critérios no processo decisório (ROY, 2005).

Na próxima seção será explicado melhor o processo de agregação, apresentando o operador de agregação, assim como algumas de suas propriedades. Também será mostrado

um exemplo de agregação linear, que abrirá caminho para a apresentação da integral de Choquet, que é um conhecido operador de agregação não-linear.

2.2 O processo de agregação

Considere $\mathbf{P} \in \mathbb{R}^{n \times m}$ como uma matriz de decisão, onde n é o número de alternativas, pertencentes ao conjunto $A = \{a_1, a_2, \dots, a_n\}$, e m é o número de critérios, pertencentes ao conjunto $C = \{c_1, c_2, \dots, c_m\}$. Cada elemento P_{ij} da matriz representa a utilidade da alternativa i com relação ao critério j . O objetivo do processo de agregação é encontrar, para uma dada alternativa a_i , a sua utilidade, que é dada por u_i . Isso é feito utilizando-se um operador de agregação $\mathcal{F}(\cdot)$, que é aplicado nos critérios referentes a essa alternativa. Esses critérios estão associados a um conjunto de parâmetros \mathbf{w} , que podem representar, por exemplo, as ponderações de cada critério no processo de agregação. A Figura 2.6 ilustra esse processo.

$$\begin{bmatrix} P_{11} & P_{12} & \cdots & P_{1m} \\ P_{21} & P_{22} & \cdots & P_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & \cdots & P_{nm} \end{bmatrix} \rightarrow u_i = \mathcal{F}(P_{i1}, \dots, P_{im}, \mathbf{w})$$

Figura 2.6 – Processo de agregação.

O objetivo do tomador de decisão é modelar as preferências entre as alternativas, da seguinte maneira (GRABISCH; KOJADINOVIC; MEYER, 2008):

$$a_i \succ a_{i'} \leftrightarrow u_i > u_{i'}, \forall a_i, a_{i'} \in A, \quad (2.1)$$

onde A é o conjunto de alternativas, a_i é a alternativa i , $a_{i'}$ é a alternativa i' , u_i é a utilidade da alternativa i e $u_{i'}$ é a utilidade da alternativa i' . As utilidades podem ser vistas então como uma representação numérica da relação de preferência e podem ser usadas em aplicações como um modelo da expertise do tomador de decisão (GRABISCH; KOJADINOVIC; MEYER, 2008).

Os operadores de agregação possuem algumas propriedades, que são definidas a seguir (GRABISCH, 1996). Consideramos sem perda de generalidade que P_{ij} está no intervalo $[0, 1]$:

- Propriedades de valores extremos: $\mathcal{F}(0, \dots, 0) = 0$ e $\mathcal{F}(1, \dots, 1) = 1$;
- Idempotência: $\mathcal{F}(P_{ij}, P_{ij}, \dots, P_{ij}) = P_{ij}, \forall P_{ij} \in [0, 1]$;

- Continuidade: $\lim_{x \rightarrow x'} \mathcal{F}(x, x, \dots, x) = \mathcal{F}(x', x', \dots, x') = x'$;
- Monotonicidade: $\mathcal{F}(P_{i1}, P_{i2}, \dots, P_{im}) \leq \mathcal{F}(P_{i1} + \delta_1, P_{i2} + \delta_2, \dots, P_{im} + \delta_m)$;
- Comutatividade: $\mathcal{F}(P_{i1}, P_{i2}) = \mathcal{F}(P_{i2}, P_{i1})$;
- Decomposabilidade: $\mathcal{F}^{(m)}(P_{i1}, P_{ik}, P_{i(k+1)}, \dots, P_{im}) = \mathcal{F}^{(m)}(P_{ij}, \dots, P_{ij}, P_{i(k+1)}, \dots, P_{im})$, onde $P_{ij} = \mathcal{F}^{(k)}(P_{i1}, \dots, P_{ik})$, para todo (P_{i1}, \dots, P_{im}) ;
- Estabilidade na transformação linear: $\mathcal{F}(rP_{i1} + t, \dots, rP_{im} + t) = r\mathcal{F}(P_{i1}, \dots, P_{im}) + t$, $\forall r > 0, \forall t \in \mathbb{R}$

Os operadores de agregação mais comuns são a média, a mediana e funções de máximo ou mínimo (GRABISCH, 1996). Um operador de agregação de interesse para esse trabalho, pois pode ser visto como um caso particular da integral de Choquet, é a média ponderada ordenada (OWA, *Ordered weighted averaging*), que é definida como (YAGER, 1988; FILEV; YAGER, 1998; YAGER, 2009):

$$OWA_i(P_{i1}, \dots, P_{im}) = \sum_{j=1}^m w_j P_{i(j)}, \quad (2.2)$$

onde $\sum_j w_j = 1$ e $P_{i(1)} \leq P_{i(2)} \leq \dots \leq P_{i(m)}$. Entre os membros dessa classe de operadores de agregação estão os operadores de máximo (*max*), mínimo (*min*) e a média aritmética (*MA*). Eles podem ser emulados, escolhendo-se diferentes pesos, como mostrado abaixo (YAGER, 1988):

$$OWA_i(P_{i1}, \dots, P_{im}) = \max(P_{i1}, \dots, P_{im}), \quad (2.3)$$

se $w_m = 1$ e $w_j = 0, \forall j \neq m$.

$$OWA_i(P_{i1}, \dots, P_{im}) = \min(P_{i1}, \dots, P_{im}), \quad (2.4)$$

se $w_1 = 1$ e $w_j = 0, \forall j \neq 1$.

$$OWA_i(P_{i1}, \dots, P_{im}) = MA(P_{i1}, \dots, P_{im}), \quad (2.5)$$

se $w_j = \frac{1}{m}, \forall j \in [1, m]$.

Todos esses operadores assumem que não há nenhum tipo de relação entre os critérios na etapa de agregação, o que nem sempre é verdade, pois algumas interações

entre critérios podem ser observadas em aplicações reais (GRABISCH, 1996; BELIAKOV et al., 2004), como será discutido no exemplo a seguir.

Exemplo 2.2.1. Ordenação de estudantes (GRABISCH, 1996)

Podemos, por exemplo, verificar como se comporta o caso de uma agregação num problema real que é a ordenação de estudantes segundo um certo índice de desempenho. Essa medida pode ser calculada como uma média ponderada (MP), do seguinte modo:

$$MP_i = \sum_{j=1}^m w_j P_{ij}, \quad (2.6)$$

onde P_{ij} corresponde à nota do aluno i na disciplina j , normalizada entre 0 e 1. O peso w_j dá a importância relativa do critério j no processo de decisão, sendo que $\sum_j w_j = 1$. Nesse exemplo, w_j corresponde a carga horária relativa, para cada disciplina j .

Na Tabela 2.1 pode-se ver as notas de três alunos genéricos, A, B e C, em três disciplinas: Física, Cálculo e Ética. A média ponderada também aparece calculada conforme a Equação (2.6). Para o aluno A, por exemplo, MP é dada por:

$$MP_A = \left(\frac{60}{150}\right) \left(\frac{9}{10}\right) + \left(\frac{60}{150}\right) \left(\frac{8,5}{10}\right) + \left(\frac{30}{150}\right) \left(\frac{4,5}{10}\right) = 0,79.$$

Aplicando a agregação, considerando a média ponderada, obtém-se a seguinte ordenação Aluno A \succ Aluno C \succ Aluno B, ou seja, o Aluno A tem preferência sobre o Aluno C, que tem preferência sobre o Aluno B. A posição do Aluno B se mostra coerente, pois ele obteve as notas mais baixas, entre os três, em todas as disciplinas. Porém, entre o Aluno A e o Aluno C, o Aluno A obteve preferência, mesmo tendo sua nota em Ética tão baixa quanto o Aluno B. Já o Aluno C manteve uma regularidade de notas em todas as disciplinas, o que pode ser interessante quando se deseja um aluno com habilidades diferentes. A dificuldade em se escolher a melhor ordenação nesse caso é o fato de que, apesar de termos três critérios, dois deles estão fortemente relacionados. Um aluno que tem um bom desempenho em Física, provavelmente também tem um bom desempenho em Cálculo, pois as duas são disciplinas da área de exatas, ou seja, há uma forte relação entre as duas. Portanto, há basicamente duas habilidades envolvidas nesse processo de ordenamento. Essa forte relação entre critérios gera um viés para a ordenação dos alunos.

Fim do exemplo

No Exemplo 2.2.1 foi utilizado um operador de agregação linear, a média ponderada, que mostrou um problema de enviesamento no ordenamento, pois havia dois critérios fortemente relacionados. A fim de se evitar que critérios relacionados enviesem a

Tabela 2.1 – Ordenação dos estudantes pela média ponderada.

Aluno	Física (60 horas)	Cálculo (60 horas)	Ética (30 horas)	MP
A	9	8,5	4,5	0,79
B	5	6	4,5	0,53
C	7	7	8	0,72

ordenação das alternativas, podem ser escolhidos outros operadores de agregação, como os não-lineares. Um operador de agregação não-linear, bem conhecido em MCDA, é a integral de Choquet (GRABISCH, 1996), que será definida na seção seguinte.

2.3 A integral de Choquet

A integral de Choquet foi apresentada na década de 1950 por Gustave Choquet (CHOQUET, 1954; GRABISCH; LABREUCHE, 2010) e começou a ser aplicada em MCDA na década de 1990, onde os primeiros trabalhos permaneceram num nível intuitivo e experimental (GRABISCH; LABREUCHE, 2010). As ideias apresentadas em (MUROFUSHI, 1992) e (MUROFUSHI; SONEDA, 1993) iniciaram o desenvolvimento de uma base teórica para o uso da integral de Choquet em MCDA e o trabalho de (GRABISCH, 1996) mostrou como estava esse desenvolvimento em 1996, popularizando o uso da integral de Choquet em MCDA (GRABISCH; LABREUCHE, 2010).

A integral de Choquet é calculada, para uma alternativa i , do seguinte modo:

$$\gamma_i^\mu(P_{i1}, \dots, P_{im}) = \sum_{j=1}^m (P_{i(j)} - P_{i(j-1)}) \mu(D_{(j)}), \quad (2.7)$$

onde o operador (\cdot) representa uma permutação da seguinte maneira $P_{i(0)} \leq P_{i(1)} \leq P_{i(2)} \leq \dots \leq P_{i(m)}$ com $P_{i(0)} = 0$, de modo que o subconjunto de critérios $D_{(j)} = \{c_{(j)}, \dots, c_{(m)}\}$ contém do j -ésimo menor critério até o maior critério. O parâmetro μ é uma medida *fuzzy* (MARICHAL, 2000), também conhecida como capacidade de Choquet, que é definida não somente para os critérios individuais, mas também para todos os subconjuntos de critérios. Devido ao caráter combinatório dessa medida, para um problema com m critérios, teremos 2^m elementos que a compõem.

Podemos definir a forma vetorial da integral de Choquet, que para uma alternativa i é dada por:

$$\gamma_i^\mu(P_{i1}, \dots, P_{im}) = \mathbf{m}_i^t \boldsymbol{\mu}, \quad (2.8)$$

onde

$$\mathbf{m}_i = \begin{bmatrix} 0 \\ \vdots \\ P_{i(m)} - P_{i(m-1)} \\ \vdots \\ P_{i(2)} - P_{i(1)} \\ \vdots \\ P_{i(1)} - P_{i(0)} \end{bmatrix}. \quad (2.9)$$

cuja dimensão é 2^m . As diferenças $P_{i(j)} - P_{i(j-1)}$, que aparecem no vetor \mathbf{m}_i , devem estar localizadas nas mesmas posições correspondentes às de suas capacidades no vetor de capacidades, que pode ser definido, codificado numa sequência lexicográfica, da seguinte forma:

$$\boldsymbol{\mu} = [\mu(\emptyset), \mu(\{c_1\}), \mu(\{c_2\}), \dots, \mu(\{c_1, c_2\}), \dots, \mu(\{c_1, c_2, \dots, c_m\})]^t. \quad (2.10)$$

Para simplificar a notação, podemos assumir $\mu(\{c_j\}) = \mu_j$. Então, tem-se:

$$\boldsymbol{\mu} = [\mu_\emptyset, \mu_1, \mu_2, \dots, \mu_{12}, \dots, \mu_{12\dots m}]^t. \quad (2.11)$$

Podemos ainda definir a forma matricial da integral de Choquet, considerando o conjunto com todas as alternativas, da seguinte forma:

$$\gamma^\mu(P_{i1}, \dots, P_{im}) = \mathbf{M}^t \boldsymbol{\mu}, \quad (2.12)$$

onde \mathbf{M} é dada por:

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(m)} - P_{1(m-1)} & P_{2(m)} - P_{2(m-1)} & \dots & P_{n(m)} - P_{n(m-1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(2)} - P_{1(1)} & P_{2(2)} - P_{2(1)} & \dots & P_{n(2)} - P_{n(1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(1)} - P_{1(0)} & P_{2(1)} - P_{2(0)} & \dots & P_{n(1)} - P_{n(0)} \end{bmatrix}, \quad (2.13)$$

cuja dimensão é $2^m \times n$. Como dito acima, as diferenças $P_{i(j)} - P_{i(j-1)}$, devem estar localizadas nas mesmas posições correspondentes às de suas capacidades no vetor de

capacidades, portanto não necessariamente as posições dos elementos para cada coluna devem ser as mesmas, pois cada critério terá um ordenamento diferente.

As propriedades da medida *fuzzy* aplicadas no conjunto de critérios C são as seguintes:

- $\mu(\emptyset) = 0, \mu(C) = 1$;
- $G \subset H \subset C$ implica $\mu(G) \leq \mu(H)$ (monotonicidade).

A propriedade de monotonicidade, já definida anteriormente, significa que o fato de se adicionar um novo critério em uma coalizão não pode diminuir o valor da capacidade para aquela coalizão (GRABISCH, 2000). Por exemplo, usando a notação apresentada em (2.11), temos que $\mu_1 \leq \mu_{12}$.

As capacidades de Choquet podem ser vistas como os análogos dos pesos presentes na agregação linear, pois permitem ao usuário obter uma quantificação das importâncias de cada um dos critérios, bem como das coalizões entre os critérios (GRABISCH, 1996). Uma possibilidade interessante ao se trabalhar com medidas *fuzzy* é a de se conseguir modelar as interações entre os critérios, da seguinte maneira:

- Se $\mu(G, H) = \mu(G) + \mu(H)$, a medida *fuzzy* é chamada de aditiva e representa uma situação na qual não há interação entre os critérios;
- Se $\mu(G, H) \leq \mu(G) + \mu(H)$, a medida *fuzzy* é chamada de subaditiva e representa uma situação na qual há uma interação negativa entre os critérios, ou seja, os critérios apresentam redundância entre eles;
- Se $\mu(G, H) \geq \mu(G) + \mu(H)$, a medida *fuzzy* é chamada de superaditiva e representa uma situação na qual há uma interação positiva entre os critérios, ou seja, os critérios apresentam sinergia entre eles.

Uma propriedade interessante da integral de Choquet é que ela generaliza o agregador linear e o OWA. De fato, a média aritmética com pesos $w_j = \frac{1}{m}$ coincide com uma integral de Choquet com medida *fuzzy* aditiva, isto é, com as capacidades definidas por (GRABISCH, 2000):

$$\mu(G) = \sum_{j=1}^{|G|} w_j, \quad (2.14)$$

onde $|G|$ representa a cardinalidade do conjunto G . Da mesma maneira, podemos dizer

que o operador OWA com pesos w_1, \dots, w_m é uma integral de Choquet, cujas capacidades são definidas por (GRABISCH, 1996):

$$\mu(G) = \sum_{j=0}^{|G|-1} w_{m-j}, \quad (2.15)$$

onde $|G|$ representa a cardinalidade do conjunto G . Nos dois exemplos a seguir, será verificado como se dá essa generalização da média aritmética e do OWA pela integral de Choquet.

Exemplo 2.3.1. A relação entre a integral de Choquet e a média aritmética

Consideremos um exemplo em que temos $w_1 = w_2 = w_3 = \frac{1}{3}$, $P_1 = 1$, $P_2 = 3$ e $P_3 = 5$. Utilizando a equação (2.5), a média aritmética (MA) é:

$$MA = \left(\frac{1}{3} \times 1\right) + \left(\frac{1}{3} \times 3\right) + \left(\frac{1}{3} \times 5\right) = 3.$$

Para calcular a integral de Choquet, encontramos as capacidades através da equação (2.14):

$$\begin{aligned} \mu(\{c_1\}) &= \mu(\{c_2\}) = \mu(\{c_3\}) = \sum_{j=1}^1 w_j = w_1 = \frac{1}{3}; \\ \mu(\{c_1, c_2\}) &= \mu(\{c_1, c_3\}) = \mu(\{c_2, c_3\}) = \sum_{j=1}^2 w_j = w_1 + w_2 = \frac{2}{3}; \\ \mu(\{c_1, c_2, c_3\}) &= \sum_{j=1}^3 w_j = w_1 + w_2 + w_3 = 1. \end{aligned}$$

A partir de (2.7), podemos então calcular a integral de Choquet da seguinte maneira:

$$\gamma^\mu(P_1, P_2, P_3) = (P_{(1)} - P_{(0)})\mu(\{c_1, c_2, c_3\}) + (P_{(2)} - P_{(1)})\mu(\{c_2, c_3\}) + (P_{(3)} - P_{(2)})\mu(\{c_3\})$$

$$\gamma^\mu(P_1, P_2, P_3) = (1 - 0) \times 1 + (3 - 1) \times \frac{2}{3} + (5 - 3) \times \frac{1}{3} = 3.$$

Portanto, verificamos que $MA = \gamma^\mu(P_1, P_2, P_3) = 3$.

Fim do exemplo

Exemplo 2.3.2. A relação entre a integral de Choquet e o OWA

Para esse exemplo, consideremos que $w_1 = 0, 2$, $w_2 = 0, 3$, $w_3 = 0, 5$, $P_{(1)} = 1$, $P_{(2)} = 3$ e $P_{(3)} = 5$. Utilizando a equação (2.2), o OWA é:

$$OWA = (0, 2 \times 1) + (0, 3 \times 3) + (0, 5 \times 5) = 3, 6.$$

Para calcular a integral de Choquet, encontramos as capacidades através da equação (2.15):

$$\begin{aligned}\mu(\{c_1\}) &= \mu(\{c_2\}) = \mu(\{c_3\}) = \sum_{j=0}^0 w_{3-j} = w_3 = 0, 5; \\ \mu(\{c_1, c_2\}) &= \mu(\{c_1, c_3\}) = \mu(\{c_2, c_3\}) = \sum_{j=0}^1 w_{3-j} = w_3 + w_2 = 0, 8; \\ \mu(\{c_1, c_2, c_3\}) &= \sum_{j=0}^2 w_{3-j} = w_3 + w_2 + w_1 = 1.\end{aligned}$$

A partir de (2.7), podemos então calcular a integral de Choquet da seguinte maneira:

$$\gamma^\mu(P_1, P_2, P_3) = (P_{(1)} - P_{(0)})\mu(\{c_1, c_2, c_3\}) + (P_{(2)} - P_{(1)})\mu(\{c_2, c_3\}) + (P_{(3)} - P_{(2)})\mu(\{c_3\})$$

$$\gamma^\mu(P_1, P_2, P_3) = (1 - 0) \times 1 + (3 - 1) \times 0, 8 + (5 - 3) \times 0, 5 = 3, 6.$$

Portanto, verificamos que $OWA = \gamma^\mu(P_1, P_2, P_3) = 3, 6$.

Fim do exemplo

Exemplo 2.3.3. Ordenação de estudantes aplicando a integral de Choquet

Voltemos ao exemplo da ordenação de estudantes, aplicando agora a formulação da integral de Choquet. O objetivo aqui é verificar como o ordenamento vai se comportar, utilizando uma modelagem que leva em consideração as interações entre os critérios. Inicialmente, vamos definir as capacidades de Choquet, que terão o mesmo papel dos pesos no agregador linear. Como temos 3 critérios, teremos que definir 8 ($= 2^3$) capacidades de Choquet.

1. $\mu(\emptyset) = 0$, por definição;

Para as capacidades de Choquet dos critérios individuais, vamos utilizar os mesmos pesos que no caso linear:

2. $\mu(\{Física\}) = \frac{60}{150} = 0, 4$;

3. $\mu(\{Cálculo\}) = \frac{60}{150} = 0, 4$;

$$4. \mu(\{\acute{E}tica\}) = \frac{30}{150} = 0,2;$$

Para a coalizão de dois critérios, devemos levar em consideração se há sinergia ou redundância entre eles:

$$5. \mu(\{Física, Cálculo\}) = 0,5 < 0,4 + 0,4, \text{ pois há redundância entre esse dois critérios;}$$

$$6. \mu(\{Física, \acute{E}tica\}) = 0,9 > 0,4 + 0,2, \text{ pois há sinergia entre esse dois critérios;}$$

$$7. \mu(\{Cálculo, \acute{E}tica\}) = 0,9 > 0,4 + 0,2, \text{ pois também há sinergia entre esse dois critérios;}$$

$$8. \mu(\{Física, Cálculo, \acute{E}tica\}) = 1, \text{ por definição.}$$

Vamos agora calcular, para cada aluno, a integral de Choquet, utilizando a Equação 2.7. Para o aluno A ($i = 1$), temos:

$$\gamma_1^\mu(P_{11}, P_{12}, P_{13}) = (P_{1(1)} - P_{1(0)})\mu(D_{(1)}) + (P_{1(2)} - P_{1(1)})\mu(D_{(2)}) + (P_{1(3)} - P_{1(2)})\mu(D_{(3)})$$

Ordenando as notas, temos: $P_{1(0)} = 0$, $P_{1(1)} = 4,5$ ($\acute{E}tica$), $P_{1(2)} = 8,5$ ($Cálculo$) e $P_{1(3)} = 9$ ($Física$). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{\acute{E}tica, Cálculo, Física\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{Cálculo, Física\}) = 0,5;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{Física\}) = 0,4.$$

Então, temos para o aluno A:

$$\gamma_1^\mu(P_{11}, P_{12}, P_{13}) = \frac{(4,5 - 0)}{10} \times 1 + \frac{(8,5 - 4,5)}{10} \times 0,5 + \frac{(9 - 8,5)}{10} \times 0,4 = 0,71.$$

A Figura 2.7 mostra a interpretação geométrica da agregação linear vista no Exemplo 2.2.1 e da agregação feita utilizando-se a integral de Choquet, para o aluno A. Nela, aparecem os valores dos critérios para cada caso, assim como os pesos para o agregador linear. Na interpretação geométrica da integral de Choquet, os critérios já foram previamente ordenados.

Para o aluno B ($i = 2$), temos:

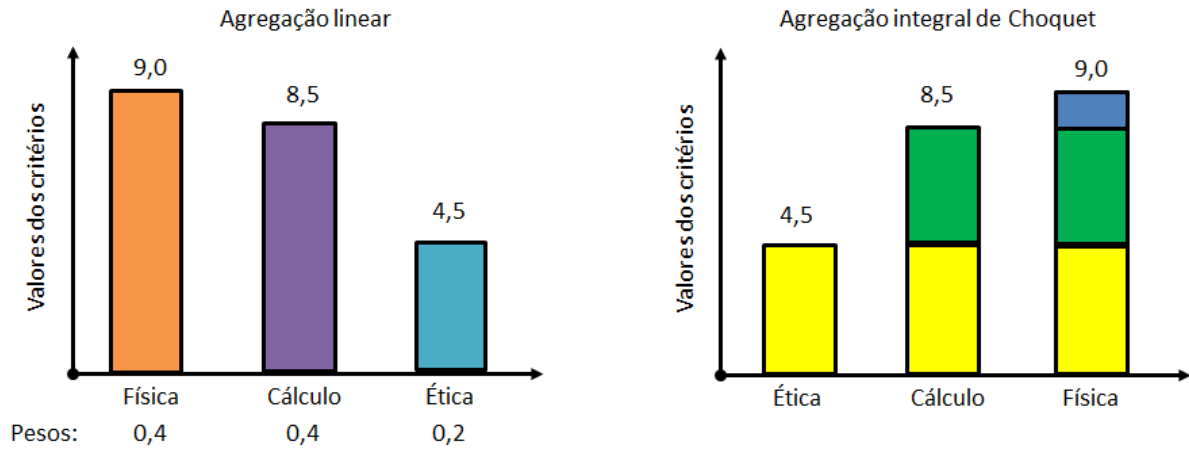


Figura 2.7 – Interpretação geométrica para a agregação linear e para a agregação feita utilizando a integral de Choquet.

$$\gamma_2^\mu(P_{21}, P_{22}, P_{23}) = (P_{2(1)} - P_{2(0)})\mu(D_{(1)}) + (P_{2(2)} - P_{2(1)})\mu(D_{(2)}) + (P_{2(3)} - P_{2(2)})\mu(D_{(3)})$$

Ordenando as notas, temos: $P_{2(0)} = 0$, $P_{2(1)} = 4,5$ (Ética), $P_{2(2)} = 5$ (Física) e $P_{2(3)} = 6$ (Cálculo). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{\text{Ética}, \text{Física}, \text{Cálculo}\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{\text{Física}, \text{Cálculo}\}) = 0,5;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{\text{Cálculo}\}) = 0,4.$$

Então, temos para o aluno B:

$$\gamma_2^\mu(P_{21}, P_{22}, P_{23}) = \frac{(4,5 - 0)}{10} \times 1 + \frac{(5 - 4,5)}{10} \times 0,5 + \frac{(6 - 5)}{10} \times 0,4 = 0,515.$$

Para o aluno C ($i = 3$), temos:

$$\gamma_3^\mu(P_{31}, P_{32}, P_{33}) = (P_{3(1)} - P_{3(0)})\mu(D_{(1)}) + (P_{3(2)} - P_{3(1)})\mu(D_{(2)}) + (P_{3(3)} - P_{3(2)})\mu(D_{(3)})$$

Ordenando as notas, temos: $P_{3(0)} = 0$, $P_{3(1)} = 7$ (Física), $P_{3(2)} = 7$ (Cálculo) e $P_{3(3)} = 8$ (Ética). Vamos agora verificar quais capacidades de Choquet utilizaremos nesse caso:

$$\mu(D_{(1)}) = \mu(\{c_{(1)}, c_{(2)}, c_{(3)}\}) = \mu(\{Física, Cálculo, Ética\}) = 1;$$

$$\mu(D_{(2)}) = \mu(\{c_{(2)}, c_{(3)}\}) = \mu(\{Cálculo, Ética\}) = 0,9;$$

$$\mu(D_{(3)}) = \mu(\{c_{(3)}\}) = \mu(\{Ética\}) = 0,2.$$

Então, temos para o aluno C:

$$\gamma_3^\mu(P_{31}, P_{32}, P_{33}) = \frac{(7-0)}{10} \times 1 + \frac{(7-7)}{10} \times 0,9 + \frac{(8-7)}{10} \times 0,2 = 0,72.$$

Podemos verificar como fica a forma matricial da integral de Choquet para esse exemplo. A matriz \mathbf{M} é dada por:

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & 0 \\ 9-8,5 & 0 & 0 \\ 0 & 6-5 & 0 \\ 0 & 0 & 8-7 \\ 8,5-4,5 & 5-4,5 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 7-7 \\ 4,5-0 & 4,5-0 & 7-0 \end{bmatrix},$$

e o vetor de capacidades é dado por:

$$\boldsymbol{\mu} = \begin{bmatrix} 0 \\ 0,4 \\ 0,4 \\ 0,2 \\ 0,5 \\ 0,9 \\ 0,9 \\ 1 \end{bmatrix},$$

de tal modo que:

$$\boldsymbol{\gamma} = \mathbf{M}^t \boldsymbol{\mu} = \begin{bmatrix} 0,71 \\ 0,515 \\ 0,72 \end{bmatrix}.$$

A Tabela 2.2 mostra novamente os três alunos A, B e C, com suas notas nas três disciplinas: Física, Cálculo e Ética. Nela aparece os valores calculados utilizando a integral de Choquet e, também, a média ponderada, a fim de se comparar os resultados.

Tabela 2.2 – Ordenação dos estudantes por meio da integral de Choquet.

Aluno	Física (60 horas)	Cálculo (60 horas)	Ética (30 horas)	MP	γ
A	9	8,5	4,5	0,79	0,71
B	5	6	4,5	0,53	0,515
C	7	7	8	0,72	0,72

Vemos que, utilizando a integral de Choquet, a ordem ficou Aluno C \succ Aluno A \succ Aluno B, ou seja, o Aluno C tem preferência sobre o Aluno A, que tem preferência sobre o Aluno B. Portanto, utilizando a formulação da integral de Choquet, que permite modelar as interações que existem entre critérios, vemos que a ordem foi alterada, privilegiando aquele aluno que manteve uma regularidade de notas nas três disciplinas.

Fim do exemplo

Um problema que encontramos nesse cenário é o de identificação das capacidades de Choquet. Pelos Exemplos 2.2.1 e 2.3.3 pode-se verificar que para um operador de agregação como a média ponderada temos m pesos, enquanto que para a integral de Choquet temos 2^m coeficientes *fuzzy*. Sabe-se, pelas propriedades da medida *fuzzy*, que $\mu(\emptyset) = 0$ e $\mu(C) = 1$ (ROWLEY; GESCHKE; LENZEN, 2015; KOJADINOVIC, 2004), portanto num problema de identificação temos $(2^m - 2)$ coeficientes para determinar, levando-se em consideração a monotonicidade existente entre esse coeficientes. No próximo capítulo, será mostrada uma maneira de se abordar o problema de identificação de capacidades.

3 Identificação das capacidades de Choquet

Nesse capítulo será apresentado o problema de estimação dos parâmetros da integral de Choquet. Inicialmente, serão apresentadas brevemente as abordagens existentes para estimar os parâmetros da integral de Choquet. Em seguida, será apresentada a abordagem utilizada nesse trabalho, que é a identificação supervisionada das capacidades da integral de Choquet. Por último, serão discutidas as características de um problema mal-posto, assim como uma maneira de se regularizar o problema por meio da norma l_1 . Essa regularização, como será vista, leva a uma solução esparsa, conceito esse que também será discutido.

3.1 Estimação dos parâmetros da integral de Choquet

Há muitos métodos para estimar os parâmetros da integral de Choquet e eles podem ser divididos em três abordagens: não-supervisionada, semi-supervisionada e supervisionada.

Na abordagem não-supervisionada, os ajustes das capacidades são conduzidos por uma solução *data-driven*. Isso significa que a única informação requerida é a matriz de decisão \mathbf{P} e, portanto, não há necessidade de o tomador de decisão fazer qualquer hipótese sobre as preferências, a princípio. Os métodos não-supervisionados são fortemente baseados em ferramentas estatísticas, dado que as interações entre os critérios devem ser estimadas dos dados observados (\mathbf{P}). Portanto, métodos não-supervisionados são usados para lidar com problemas em que as interações entre critérios existem, mas elas não são claras para o tomador de decisão (DUARTE, 2018).

Na abordagem semi-supervisionada, o tomador de decisão precisa apenas definir valores para modelar a importância dos critérios e das interações entre eles, mas não há a necessidade de definir um pré-ordenamento entre as alternativas (DUARTE, 2018). Exemplos de métodos semi-supervisionados podem ser encontrados em (WU et al., 2014) e (WU et al., 2015).

Por fim, na abordagem supervisionada, que é o foco desse trabalho, o tomador de decisão deve estabelecer a importância de todos os critérios, assim como as interações entre eles, e deve estabelecer um pré-ordenamento entre as alternativas. Uma vez que as entradas são determinadas, o problema de identificação das capacidades consiste num problema de otimização com restrições, no qual a função custo é dada por uma distância entre as saídas da integral de Choquet e os valores que explicam as preferências do tomador de decisão (DUARTE, 2018). Os primeiros métodos supervisionados foram propostos

em (MARICHAL; ROUBENS, 2000) e (GRABISCH, 1996). Em (ANDERSON; PRICE; HAVENS, 2014) e (TEHRANI et al., 2012b) foram apresentadas abordagens nas quais se introduziu termos de regularização na função custo. Uma revisão dos métodos supervisionados pode ser encontrada em (GRABISCH; KOJADINOVIC; MEYER, 2008). Os métodos supervisionados podem ser muito restritivos em problemas práticos de decisão, como a elicitación de preferências, devido à necessidade de se ter informações iniciais (GRABISCH; LABREUCHE, 2010; DUARTE, 2018). Entretanto, eles têm uma relevância grande em problemas de desagregação (DOUMPOS; ZOPOUNIDIS, 2011), podendo ser aplicado, por exemplo, em sistemas de recomendação.

3.1.1 Aprendizado supervisionado da integral de Choquet

O presente trabalho considera uma abordagem supervisionada para identificação das capacidades de Choquet, na qual é assumido o conhecimento da matriz de decisão \mathbf{P} e do vetor de utilidades \mathbf{u} . Tem-se então um problema inverso, que pode ser resolvido através da minimização da soma dos erros quadráticos, calculados considerando a saída da integral de Choquet e a utilidade para cada uma das alternativas. Em termos matemáticos, tal problema pode ser formulado da seguinte maneira (GRABISCH; NGUYEN; WALKER, 2013; ADEYEBA; ANDERSON; HAVENS, 2015):

$$\min_{\boldsymbol{\mu}} \sum_{i=1}^n (\gamma_i^{\boldsymbol{\mu}}(P_{i1}, \dots, P_{im}) - u_i)^2, \quad (3.1)$$

onde u_i é a utilidade da alternativa i .

A partir da Equação (2.8), pode-se escrever (3.1) de forma matricial, do seguinte modo:

$$\min_{\boldsymbol{\mu}} \sum_{i=1}^n (\mathbf{m}_i^t \boldsymbol{\mu} - u_i)^2, \quad (3.2)$$

onde \mathbf{m}_i^t é dado por (2.9) e $\boldsymbol{\mu}$ é dado por (2.11). Após um desenvolvimento elementar da Equação (3.2), temos que:

$$\min_{\boldsymbol{\mu}} \sum_{i=1}^n (\boldsymbol{\mu}^t \mathbf{m}_i \mathbf{m}_i^t \boldsymbol{\mu} - 2u_i \mathbf{m}_i^t \boldsymbol{\mu} + u_i^2). \quad (3.3)$$

Podemos, então, escrever o seguinte problema de otimização quadrática:

$$\begin{aligned} \min_{\boldsymbol{\mu}} \quad & \frac{1}{2} \boldsymbol{\mu}^t \mathbf{Q} \boldsymbol{\mu} + \mathbf{f}^t \boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{R} \boldsymbol{\mu} \leq \mathbf{s} \\ & 0 \leq \boldsymbol{\mu} \leq 1, \end{aligned} \tag{3.4}$$

onde $\mathbf{Q} = 2 \sum_{i=1}^n \mathbf{m}_i \mathbf{m}_i^t$ e $\mathbf{f} = \sum_{i=1}^n (-2u_i \mathbf{m}_i)$. Note que a utilidade u_i é uma constante pois é um dado do problema e, portanto, o problema de otimização expresso em (3.3) é equivalente ao problema expresso em (3.4). As restrições em (3.4) estão relacionadas com as propriedades da medida *fuzzy*: $\mu(\emptyset) = 0$, $\mu(C) = 1$ e monotonicidade, apresentadas na Seção 2.3. Elas podem ser representadas por (ADEYEBA; ANDERSON; HAVENS, 2015):

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ -1 & 0 & 0 & \dots & 0 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & -1 & 0 & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & 0 & -1 & \dots & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 1 & -1 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & -1 \end{bmatrix}. \tag{3.5}$$

cuja dimensão é $(m(2^{m-1} - 1) + 4) \times 2^m$. O vetor \mathbf{s} é dado por:

$$\mathbf{s} = \begin{bmatrix} 0 & 0 & \dots & 0 & 0 & 1 & -1 \end{bmatrix}^t, \tag{3.6}$$

cujos tamanho é $(m(2^{m-1} - 1) + 4)$.

Exemplo 3.1.1. Vamos verificar como são construídos \mathbf{R} e \mathbf{s} para um caso com 3 critérios. O vetor de capacidades é dado por:

$$\boldsymbol{\mu} = [\mu_{\emptyset}, \mu_1, \mu_2, \mu_3, \mu_{12}, \mu_{13}, \mu_{23}, \mu_{123}]^t,$$

e as restrições são dadas por:

$$\begin{aligned}
\mu_1 - \mu_{12} &\leq 0, \\
\mu_1 - \mu_{13} &\leq 0, \\
\mu_2 - \mu_{12} &\leq 0, \\
\mu_2 - \mu_{23} &\leq 0, \\
\mu_3 - \mu_{13} &\leq 0, \\
\mu_3 - \mu_{23} &\leq 0, \\
\mu_{12} - \mu_{123} &\leq 0, \\
\mu_{13} - \mu_{123} &\leq 0, \\
\mu_{23} - \mu_{123} &\leq 0.
\end{aligned}$$

Além disso, temos $\mu_\emptyset = 0$, que pode ser escrito como: $\mu_\emptyset \leq 0$ e $-\mu_\emptyset \leq 0$. Temos também $\mu_{123} = 1$, que pode ser escrito como $\mu_{123} \leq 1$ e $-\mu_{123} \leq -1$. Portanto, a matriz \mathbf{R} e o vetor \mathbf{s} é dado por:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

$$\mathbf{s} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix}^t.$$

Fim do exemplo

Como dito anteriormente, o problema de otimização expresso em (3.4) está associado a um problema inverso que, na maioria dos casos, é mal-posto (SARKAR; WEINER; JAIN, 1981). Na seção seguinte, serão apresentadas brevemente as condições em que problemas inversos são mal-postos e as suas consequências, assim como algumas abordagens de solução.

3.1.1.1 Problemas inversos

Um caso particular de problema inverso é um sistema linear do tipo $\mathbf{Ax} = \mathbf{y}$, em que se conhecem a entrada \mathbf{A} e a saída \mathbf{y} , e quer se estimar \mathbf{x} . No início dos anos 1900, o matemático francês Hadamard definiu que, para um problema inverso ser bem-posto, ele deve respeitar as três condições abaixo (SARKAR; WEINER; JAIN, 1981):

1. Existe solução \mathbf{x} ;
2. A solução \mathbf{x} é única;
3. Pequenas perturbações em \mathbf{y} resultam em pequenas perturbações na solução \mathbf{x} , sem a necessidade de se impor restrições adicionais ao problema. Ou seja, a solução tem uma dependência contínua com os dados de entrada.

Se alguma das condições acima for violada, o problema passa a ser mal-posto (SARKAR; WEINER; JAIN, 1981). Os exemplos abaixo mostram problemas mal-postos, em que as condições 1, 2 e 3 são violadas, respectivamente.

Exemplo 3.1.2. Problema mal-posto pela violação da condição 1

Suponha o seguinte sistema linear $\mathbf{Ax} = \mathbf{y}$, onde:

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \\ 1 & 3 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ e } \mathbf{y} = \begin{bmatrix} 3 \\ 3 \\ 5 \end{bmatrix}.$$

Esse sistema linear possui 3 equações independentes e 2 incógnitas, portanto não possui solução. Como a condição 1 é violada, esse problema é mal-posto.

Fim do exemplo

Exemplo 3.1.3. Problema mal-posto pela violação da condição 2

Suponha, novamente, o seguinte sistema linear $\mathbf{Ax} = \mathbf{y}$, mas agora com:

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \end{bmatrix}, \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ e } \mathbf{y} = \begin{bmatrix} 8 \end{bmatrix}.$$

Esse sistema linear possui 1 equação e 2 incógnitas, portanto possui infinitas soluções. Como a condição 2 é violada, esse problema é mal-posto.

Fim do exemplo

Exemplo 3.1.4. Problema mal-posto pela violação da condição 3

Por último, suponha o seguinte sistema linear $\mathbf{Ax} = \mathbf{y}$, com:

$$\mathbf{A} = \begin{bmatrix} 36,8624 & 51,2393 \\ 51,2393 & 71,2235 \end{bmatrix} \text{ e } \mathbf{y} = \begin{bmatrix} 192,09 \\ 267,02 \end{bmatrix}.$$

A solução para esse sistema é: $\mathbf{x} = \begin{bmatrix} -459,6179 \\ 334,4056 \end{bmatrix}$.

Vamos verificar agora o que acontece com a solução, quando alteramos \mathbf{y} para:

$$\mathbf{y} = \begin{bmatrix} 192,08 \\ 267,02 \end{bmatrix}.$$

A solução agora é: $\mathbf{x} = \begin{bmatrix} -999,4069 \\ 722,7385 \end{bmatrix}$.

Claramente, vemos que pequenas perturbações em \mathbf{y} resultam em grandes variações em \mathbf{x} . Essa instabilidade é uma consequência do mal-condicionamento da matriz \mathbf{A} . Como a condição 3 é violada, esse problema é mal-posto.

Fim do exemplo

A maioria dos problemas inversos é mal-posta sob as 3 condições de Hadamard (SARKAR; WEINER; JAIN, 1981). De fato, o problema expresso por (3.4) precisará de algumas restrições para satisfazer as 3 condições de Hadamard. No Capítulo 5 será feita uma análise da unicidade da solução. Em particular, apresentaremos quais as condições para que se tenha uma solução única para o problema apresentado.

Tendo um problema mal-posto, existem algumas maneiras de se definir um problema associado que é bem-posto. Esse procedimento é chamado de regularização do problema mal-posto (SARKAR; WEINER; JAIN, 1981). Na seção seguinte, será mostrado como isso pode ser feito.

3.2 Regularização do problema mal-posto

Considere, novamente, o problema inverso $\mathbf{Ax} = \mathbf{y}$ em que se conhecem a entrada \mathbf{A} e a saída \mathbf{y} , e se quer estimar o vetor \mathbf{x} . Quando um problema desse tipo possui infinitas soluções, ou seja, quando é violada a condição 2 de Hadamard, é praticamente impossível identificar qual dessas soluções é a correta sem qualquer informação adicional (CANDES; WAKIN; BOYD, 2008). Nesse caso, é necessário estabelecer um critério adicional para a escolha de uma solução bem definida. Uma maneira familiar de se fazer isso é adicionando-se um termo de regularização, com o objetivo de se dar preferência a uma solução com alguma propriedade desejada. Uma das possibilidades para se fazer a regularização é considerar um termo de regularização quadrático, como (TIKHONOV et al., 2013; BRUCKSTEIN; DONOHO; ELAD, 2009):

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_2 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{y}, \end{aligned} \tag{3.7}$$

onde $\|\mathbf{x}\|_2 = (\sum_{i=1}^n |x_i|^2)^{1/2}$ é a norma l_2 . O problema (3.7) tem, de fato, solução única $\hat{\mathbf{x}}$ que é dada por (TIKHONOV et al., 2013; BRUCKSTEIN; DONOHO; ELAD, 2009):

$$\hat{\mathbf{x}} = \mathbf{A}^+ \mathbf{y} = \mathbf{A}^t (\mathbf{A} \mathbf{A}^t)^{-1} \mathbf{y}. \quad (3.8)$$

Outra maneira de se fazer a regularização de um problema mal-posto é adicionando um termo de modo a promover esparsidade na solução. A seguir, será discutido o que significa uma solução esparsa de um sistema, assim como suas implicações.

3.2.1 Soluções Esparsas

Existem problemas inversos para os quais as variáveis desconhecidas apresentam alguma propriedade conhecida *a priori*. Uma propriedade frequentemente explorada é a esparsidade, que significa que o vetor a ser estimado depende de um pequeno número de parâmetros desconhecidos. O uso da esparsidade como regularização implica, assim, na preferência pela solução mais simples (CANDES; WAKIN; BOYD, 2008).

Um vetor \mathbf{x} é esparsa quando há uma proporção significativa de elementos nulos ou próximos de zero, ou seja, há poucos elementos não nulos entre as possíveis entradas de \mathbf{x} . Uma maneira de se quantificar esparsidade é através da norma l_0 , que corresponde ao número de elementos não-nulos de um vetor. Por exemplo, dado o vetor $\mathbf{x}^T = [4, 0, 0, 6, 3, 0, 0, 0]$, mostrado na Figura 3.1, sua norma l_0 é $\|\mathbf{x}\|_0 = 3$ (OLIVEIRA, 2015; DUARTE et al., 2011; ELAD, 2010).

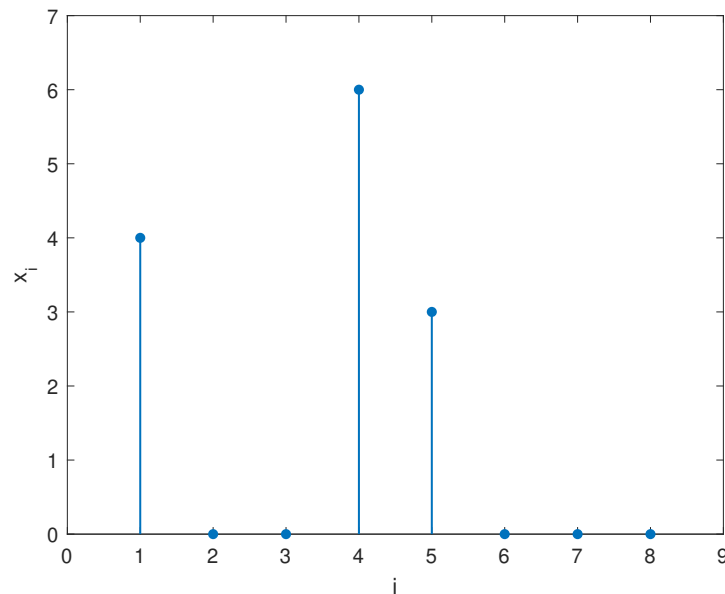


Figura 3.1 – Vetor esparsa.

Embora o termo norma l_0 seja constantemente utilizado, ela não é uma norma verdadeira, pois não obedece a uma das duas condições exigidas para que uma medida seja uma norma verdadeira. Ela obedece à desigualdade triangular, portanto, também obedece à desigualdade triangular reversa, ou seja, dado dois vetores \mathbf{x} e \mathbf{y} , então $\|\mathbf{x} - \mathbf{y}\|_0 \geq |\|\mathbf{x}\|_0 - \|\mathbf{y}\|_0|$. Entretanto, ela não é positiva homogênea, ou seja, dada uma constante $\alpha \neq 0$, então $\|\alpha\mathbf{x}\|_0 \neq |\alpha| \|\mathbf{x}\|_0$. Nesse trabalho, por simplicidade, será usado o termo norma (OLIVEIRA, 2015; DUARTE et al., 2011; ELAD, 2010).

Uma maneira particular de se calcular a norma l_0 de um vetor é a seguinte (OLIVEIRA, 2015; DUARTE et al., 2011; ELAD, 2010):

$$\|\mathbf{x}\|_0 = n - \sum_{i=1}^n I_0(x_i), \quad (3.9)$$

onde n é o número de elementos do vetor e $I_0(x_i)$ é um indicador dado pela Equação (3.10) e visualizado na Figura 3.2.

$$I_0(x_i) = \begin{cases} 0, & \text{se } (x_i \neq 0) \\ 1, & \text{se } (x_i = 0) \end{cases}. \quad (3.10)$$

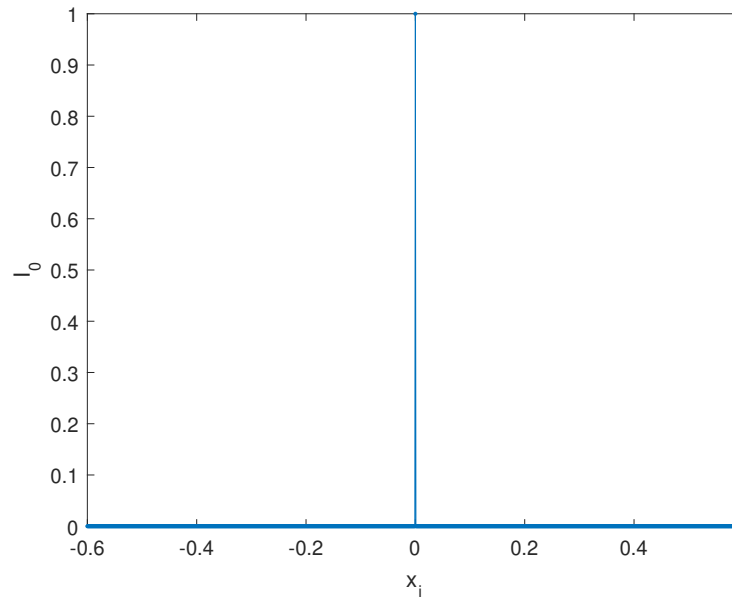


Figura 3.2 – Indicador para o cálculo da norma l_0 .

O uso de esparsidade é justificado, pois leva à solução mais simples possível. Essa busca pela solução mais simples está baseada no princípio da parcimônia, ou princípio da navalha de Occam. Em sua forma original, ele dizia 'Nunquam ponenda est pluralitas sin necessitate', que, aproximadamente traduzido, significa 'As entidades não devem ser

multiplicadas além da necessidade’. A ideia por trás desse princípio, que é vastamente usado atualmente em metodologia científica, é que se duas hipóteses podem explicar um fenômeno, provavelmente a mais simples será a mais adequada (DOMINGOS, 1999). Uma das maneiras de se encontrar a solução mais simples possível, que explica os dados, é resolvendo o seguinte problema de otimização (CANDES; WAKIN; BOYD, 2008):

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_0 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{y}. \end{aligned} \quad (3.11)$$

A solução encontrada nesse caso será esparsa, pois terá o menor número de elementos não-nulos possível. Entretanto, o problema expresso em (3.11) não é convexo, se tratando de um problema *NP-hard*, portanto, geralmente impossível de se resolver (CANDES; WAKIN; BOYD, 2008). Uma alternativa comum é considerar um problema convexo, como:

$$\begin{aligned} \min_{\mathbf{x}} \quad & \|\mathbf{x}\|_1 \\ \text{s.t.} \quad & \mathbf{Ax} = \mathbf{y}, \end{aligned} \quad (3.12)$$

onde $\|\mathbf{x}\|_1 = \sum_{i=1}^n |x_i|$ é a norma l_1 . Diferente de (3.11), este problema é convexo e pode ser resolvido eficientemente através de métodos de otimização linear. Os problemas (3.11) e (3.12) diferem na escolha da função objetivo, com o último usando a norma l_1 como uma relaxação para a norma l_0 , mas existem condições que garantem uma equivalência entre os dois (CANDES; WAKIN; BOYD, 2008). Portanto, consegue-se uma solução esparsa por meio da minimização da norma l_1 . Esta é a ideia que será usada para a regularização, mostrada na próxima seção. Com o objetivo de ilustrar como as normas l_p possuem diferentes propriedades, a Figura 3.3 mostra os lugares geométricos de raio unitário, $\{x : \|\mathbf{x}\|_p = 1\}$, para $p = 1, 2, \infty$ e $p = \frac{1}{2}$, no \mathbb{R}^2 (ELDAR; KUTYNIOK, 2012a).

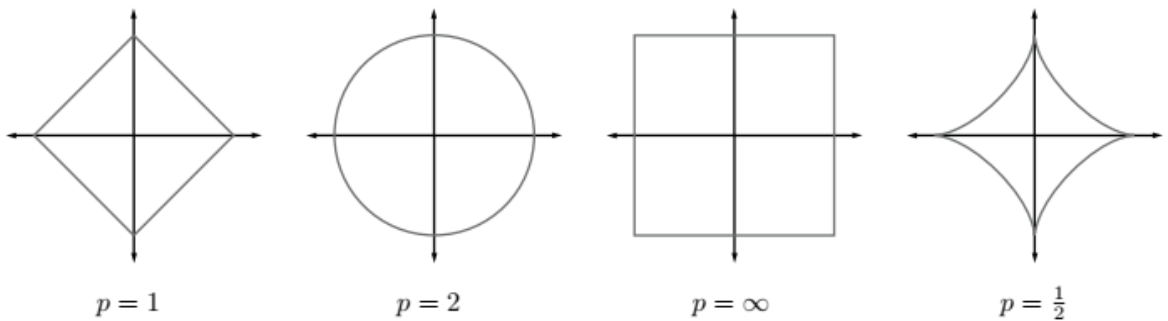


Figura 3.3 – Lugares geométricos de raio unitário no \mathbb{R}^2 para as normas l_p com $p = 1, 2, \infty$, e para a quasi-norma l_p com $p = \frac{1}{2}$. Extraído de (ELDAR; KUTYNIOK, 2012a).

Para resolver o problema inverso $\mathbf{Ax} = \mathbf{y}$ em que se conhecem a entrada \mathbf{A} e a saída \mathbf{y} , o objetivo é estimar um vetor \mathbf{x} , que seja único e que depende da norma l_p que se está considerando. Ou seja, o objetivo é encontrar \mathbf{x} que minimize $\|\mathbf{x}\|_p$, sujeito à restrição $\mathbf{Ax} = \mathbf{y}$. Para um caso de duas dimensões, como mostrado na Figura 3.4, em que x_1 e x_2 são os elementos do vetor \mathbf{x} , a solução deve estar contida na reta, que está associada à restrição. O que se verifica, então, é que, para $p < 2$, o ponto da reta que corresponde ao menor valor da norma l_p corresponde a uma solução esparsa, pois algum dos elementos de \mathbf{x} é nulo. Para $p > 2$ a tendência é se ter uma solução cada vez mais “anti-esparsa”, ou seja, os valores de x_1 e x_2 aparecem uniformemente distribuídos, até chegar em $p = \infty$, em que se tem uma solução com mesma amplitude para x_1 e x_2 . Por fim, para $p = 2$ não há nenhum tipo de “posicionamento preferencial”; a solução vai depender da inclinação da reta, portanto a norma l_2 não promove nem esparsidade nem não-esparsidade.

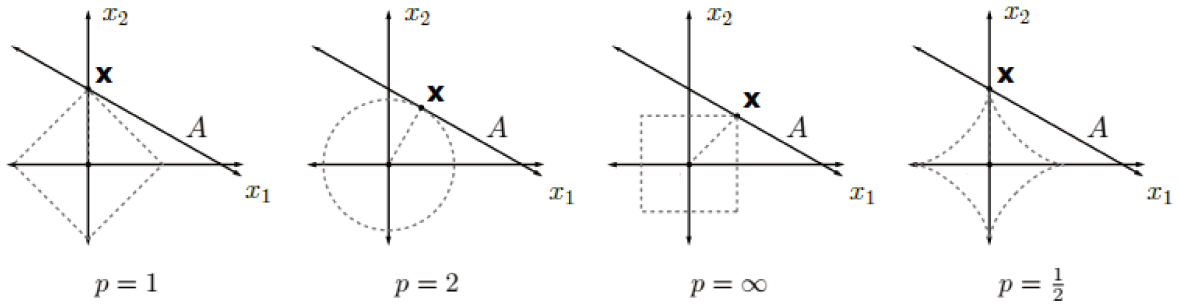


Figura 3.4 – Solução de um problema inverso através da minimização das normas l_p com $p = 1, 2, \infty$, e da quasi-norma l_p com $p = \frac{1}{2}$. Adaptado de (ELDAR; KUTYNIOK, 2012a).

3.2.2 Otimização do problema com um termo de regularização

A fim de se regularizar o problema, será acrescentado na formulação expressa na Equação (3.4) um termo da norma l_1 do vetor de variáveis em questão, $\boldsymbol{\mu}$, como apresentado em (ADEYEBA; ANDERSON; HAVENS, 2015). Como visto na seção anterior, a minimização da norma l_1 do vetor de variáveis tende a resultar num vetor esparsa, no qual há uma proporção significativa de elementos nulos ou próximos de zero (TIBSHIRANI et al., 2005; ADEYEBA; ANDERSON; HAVENS, 2015).

Portanto, o problema dado por (3.4) se torna:

$$\begin{aligned}
 \min_{\boldsymbol{\mu}} \quad & \frac{1}{2} \boldsymbol{\mu}^t \mathbf{Q} \boldsymbol{\mu} + \mathbf{f}^t \boldsymbol{\mu} + \lambda \|\boldsymbol{\mu}\|_1 \\
 \text{s.t.} \quad & \mathbf{R} \boldsymbol{\mu} \leq \mathbf{s} \\
 & \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1},
 \end{aligned} \tag{3.13}$$

onde λ é um parâmetro não negativo que controla a regularização. Esse problema também é conhecido como regressão com penalização (OSBORNE; PRESNELL; TURLACH, 2000), pois o objetivo do termo de regularização é manter as estimativas de $\boldsymbol{\mu}$ as mais esparsas possível. Quanto maior o valor de λ , maior a penalização, ou seja, aparecerão mais termos nulos em $\boldsymbol{\mu}$.

O problema expresso por (3.13) é equivalente ao problema de otimização conhecido como LASSO (do inglês, *Least Absolute Shrinkage and Selection Operator*) (TIBSHIRANI, 1996). Nele, o termo de regularização da norma l_1 que aparece na função objetivo será definido como uma restrição do problema, da seguinte maneira (OSBORNE; PRESNELL; TURLACH, 2000):

$$\begin{aligned} \min_{\boldsymbol{\mu}} \quad & \frac{1}{2} \boldsymbol{\mu}^t \mathbf{Q} \boldsymbol{\mu} + \mathbf{f}^t \boldsymbol{\mu} \\ \text{s.t.} \quad & \mathbf{R} \boldsymbol{\mu} \leq \mathbf{s} \\ & \mathbf{0} \leq \boldsymbol{\mu} \leq \mathbf{1} \\ & \|\boldsymbol{\mu}\|_1 \leq t, \end{aligned} \tag{3.14}$$

onde t é um parâmetro de ajuste não negativo que controla o nível de esparsidade. A equivalência entre (3.13) e (3.14) se dá, pois, para um dado $\lambda \geq 0$, existe um $t \geq 0$, tal que os dois problemas possuem a mesma solução. Do mesmo modo, para um dado $t \geq 0$, existe $\lambda \geq 0$, tal que os dois problemas possuem a mesma solução (OSBORNE; PRESNELL; TURLACH, 2000).

Uma solução muito eficiente para o problema apresentado em (3.14) é conhecida como NNVM (do inglês, *Non-Negative Variable Method*) (TIBSHIRANI, 1996). A ideia nesse método é linearizar a norma l_1 por meio de uma representação na qual se dobra o número de variáveis de $\boldsymbol{\mu}$, ou seja, $[\mu_0^+, \mu_1^+, \mu_2^+, \dots, \mu_0^-, \mu_1^-, \mu_2^-, \dots]$, onde $\mu_j = \mu_j^+ - \mu_j^-$, e se acrescentam as seguintes restrições:

$$\begin{aligned} \mu_j^+ &\geq 0, \mu_j^- \geq 0, \\ \sum_{j=1}^{2m} (\mu_j^+ + \mu_j^-) &\leq t. \end{aligned}$$

É importante notar que, para o subíndice j , também temos que considerar a coalizão de critérios, tais como: $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, etc. O problema de otimização pode ser escrito,

então, da seguinte maneira:

$$\begin{aligned}
 & \min_{\hat{\boldsymbol{\mu}}} \quad \frac{1}{2} \hat{\boldsymbol{\mu}}^t \widehat{\mathbf{Q}} \hat{\boldsymbol{\mu}} + \hat{\mathbf{f}}^t \hat{\boldsymbol{\mu}} \\
 & \text{s.t.} \quad \widehat{\mathbf{R}} \hat{\boldsymbol{\mu}} \leq \mathbf{s} \\
 & \quad \quad \mathbf{0} \leq \hat{\boldsymbol{\mu}} \leq \mathbf{1} \\
 & \quad \quad \|\hat{\boldsymbol{\mu}}\|_1 \leq t,
 \end{aligned} \tag{3.15}$$

onde: $\hat{\boldsymbol{\mu}} = [\mu_{\emptyset}^+, \mu_1^+, \mu_2^+, \dots, \mu_{\emptyset}^-, \mu_1^-, \mu_2^-, \dots]$, $\widehat{\mathbf{Q}} = \begin{bmatrix} \mathbf{Q} & -\mathbf{Q} \\ -\mathbf{Q} & \mathbf{Q} \end{bmatrix}$, $\hat{\mathbf{f}} = \begin{bmatrix} \mathbf{f} \\ -\mathbf{f} \end{bmatrix}$ e $\widehat{\mathbf{R}} = \begin{bmatrix} \mathbf{R} & -\mathbf{R} \end{bmatrix}$.

Como mencionado anteriormente, a formulação expressa em (3.15) tende a promover soluções esparsas no vetor de capacidades de Choquet. Entretanto, como será discutido no próximo capítulo, esparsidade não é uma propriedade esperada para o vetor de capacidades de Choquet, pois ele não apresenta uma alta proporção de elementos nulos. O que será feito, então, é explorar a esparsidade em uma outra representação das capacidades de Choquet, conhecida como índice de interação.

4 Identificação no domínio do índice de interação

No capítulo 2, foi definida a medida *fuzzy* e vimos que ela pode ser usada para modelar a importância de um critério, assim como da coalizão de dois ou mais critérios. Entretanto, mesmo que ela tenha sido largamente estudada de um ponto de vista matemático, há algumas limitações que podem dificultar o seu uso em aplicações práticas. Isso acontece devido a dois fatos. O primeiro é a complexidade exponencial que existe no número de parâmetros de uma medida *fuzzy*. A segunda limitação é a dificuldade de se conseguir uma interpretação do significado preciso de $\mu(G)$. Por exemplo, pode-se ter $\mu(\{j\}) = 0$, sugerindo que o elemento j não é importante, mas pode se verificar que para algum subconjunto $G \subset C$, $\mu(G \cup \{j\})$ é muito maior que $\mu(G)$, sugerindo que j é, de fato, importante. Esta é uma dificuldade que há em se obter a medida *fuzzy* somente pela elicitación do tomador de decisão (GRABISCH, 1997).

Para auxiliar na interpretação de uma medida *fuzzy* e ajudar a resolver a complexidade computacional do problema, podem ser usadas representações alternativas, que são transformações da medida *fuzzy* feitas por operadores lineares inversíveis. Um exemplo é a transformada Möbius, conhecida há tempos na teoria combinatória (ROTA, 1964). Outro exemplo é a representação pelos índices de interação, que será estudada nesse capítulo (GRABISCH, 1997). Primeiramente, será apresentada a motivação para trabalhar no que chamaremos de domínio do índice de interação. Em seguida, será feita uma abordagem teórica sobre os índices de interação. Então, será apresentado o problema de identificação dos parâmetros da integral de Choquet, trabalhando agora nesse novo domínio. Por último, serão apresentados os resultados numéricos que mostram que a utilização da hipótese de esparsidade se mostra mais adequada quando o domínio do índice de interação é considerado.

4.1 Motivação para trabalhar no domínio do índice de interação

Um caso de particular interesse para o presente trabalho, como visto na Seção 2.3, é quando uma medida *fuzzy* é aditiva, ou seja, quando $\mu(G) = \sum_{j \in G} \mu(\{j\})$ para todo $G \subset C$ (GRABISCH, 2000). Tal situação representa o caso particular da agregação linear e, logo, pode ser vista como a expressão mais simples da integral de Choquet.

Com o intuito de ilustrar a motivação do presente trabalho, que é o fato do vetor não ser esparso quando representado na medida *fuzzy*, consideremos um problema de decisão em que se deve avaliar as alternativas sobre 4 critérios via um processo de

agregação linear no qual os pesos são dados por $[1/4; 1/4; 1/4; 1/4]$. Tal situação pode ser representada por uma integral de Choquet cuja medida *fuzzy* associada é aditiva. Assim, na Figura 4.1, apresentamos essa medida, na ordem lexicográfica, calculada para todas $2^4 = 16$ coalizões.

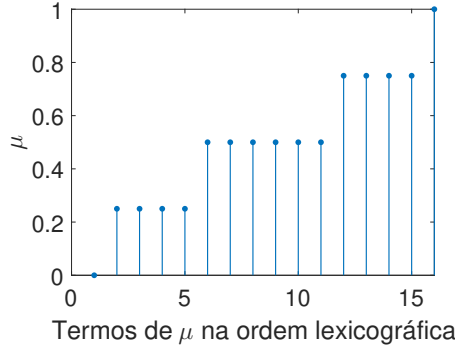


Figura 4.1 – Valores das capacidades de Choquet para o caso aditivo.

É interessante notar que, mesmo numa situação simples de agregação linear, a medida *fuzzy* associada é claramente um sinal não-esparso, ou seja, a medida *fuzzy* não apresenta uma proporção significativa de elementos nulos. Assim sendo, ao se regularizar o problema (3.4), adicionando um termo da norma l_1 do vetor μ , se encontrará uma solução esparsa. Entretanto, essa solução não representará com fidelidade a medida *fuzzy* que se está tentando estimar.

A seguir, será apresentado o domínio do índice de interação, com o objetivo de encontrar uma representação na qual a esparsidade faz mais sentido.

4.2 Índice de interação generalizado

Na *representação de interação*, a medida *fuzzy* é linearmente transformada em índices de interação generalizados. Dados os subconjuntos $G \subset C$ e $H \subset C$, o índice de interação generalizado $\eta(G)$, definido para todas as coalizões (incluindo o conjunto vazio), é dado por (GRABISCH, 2000):

$$\eta(G) = \sum_{K \subset C \setminus G} \frac{(m - |K| - |G|)! |K|!}{(m - |G| + 1)!} \times \sum_{H \subset G} (-1)^{|G| - |H|} \mu(K \cup H), \forall G \subset C, \quad (4.1)$$

onde $|Z|$ representa a cardinalidade do conjunto Z . De particular interesse são os seguintes valores (GRABISCH, 2000):

- $\eta(\{j\}) = \phi_j$, conhecido como o índice de Shapley ou valor Shapley de um jogo, um conceito importante na teoria de jogos (SHAPLEY, 1953). Ele computa a contribuição média do elemento j em todas as coalizões. A ideia de utilizar o índice de Shapley em MCDA se deve a Murofushi (MUROFUSHI, 1992);
- $\eta(\{j, k\}) = \eta_{jk}$, conhecido como índice de interação, foi proposto por Murofushi e Soneda (MUROFUSHI; SONEDA, 1993). Ele modela a ideia de interação entre os elementos j e k . Pode ser positivo, quando a cooperação entre os elementos é efetiva; negativo, quando a cooperação entre os elementos é inefetiva; ou zero, quando não há cooperação entre os elementos, ou seja, não há interação. Mais detalhes do uso do índice de interação em MCDA pode ser encontrado em (GRABISCH, 1997) e (GRABISCH; ROUBENS, 2000).

A relação entre o índice de interação generalizado η e a medida *fuzzy* μ é uma transformação linear inversível. Considerando (4.1), os valores das medidas *fuzzy* podem ser encontrados da seguinte maneira (GRABISCH, 2000):

$$\mu(G) = \sum_{H \subset C} \beta_{|G \cap H|}^{|H|} \eta(H), \forall G \subset C, \quad (4.2)$$

com

$$\beta_q^r = \sum_{p=0}^q \binom{q}{p} B_{r-p}, \quad (4.3)$$

onde B_q são os números de Bernoulli, com os primeiros termos $B_1 = -1/2$, $B_2 = 1/6$, $B_3 = 0$, $B_4 = -1/30$, $B_5 = 0$, etc. Os primeiros valores de β_q^r são (GRABISCH, 2000):

q \ r	0	1	2	3	4
0	1	-1/2	1/6	0	-1/30
1		1/2	-1/3	1/6	-1/30
2			1/6	-1/6	2/15
3				0	-1/30
4					-1/30

Como a transformação é inversível, é equivalente definir $\mu(G)$ para todo $G \subset C$ ou $\eta(G)$ para todo $G \subset C$ (GRABISCH, 2000).

Do mesmo modo como definimos um vetor de capacidades $\boldsymbol{\mu}$ em (2.11), podemos definir um vetor $\boldsymbol{\eta}$, que contém os índices de interação generalizado para todas as coalizões, também codificado numa sequência lexicográfica. Tal vetor é dado por:

$$\boldsymbol{\eta} = [\eta_{\emptyset}, \eta_1, \eta_2, \dots, \eta_{12}, \dots, \eta_{12\dots m}]^t. \quad (4.4)$$

Com isso, podemos representar a transformação linear entre os vetores $\boldsymbol{\mu}$ e $\boldsymbol{\eta}$, expressa em (4.2), definindo uma matriz de transformação \mathbf{T} de dimensão $2^m \times 2^m$, de tal modo que:

$$\boldsymbol{\mu} = \mathbf{T}\boldsymbol{\eta}. \quad (4.5)$$

Na aplicação da integral de Choquet em MCDA, há muito interesse na subclasse de modelos conhecida como *k-aditiva*, na qual $\eta(G) = 0$ para todas as coalizões de mais que k elementos. Nesse caso deve existir ao menos G de exatamente k elementos tal que $\eta(G) \neq 0$ (GRABISCH, 2000). Os modelos *k-aditivo* são muito úteis na prática, principalmente porque eles permitem modelar interações não-lineares entre critérios, enquanto possuem menos parâmetros não-nulos, quando comparado com a integral de Choquet completa.

É interessante notar que, para os modelos *k-aditivo*, os índices de interação generalizados podem ser representados por um vetor, no qual todos os elementos relacionados às coalizões de mais que k critérios são nulos. Ou seja, quando representado no domínio do índice de interação, modelos *k-aditivo* são relacionados com representações esparsas. Isso é verdade, é claro, para valores pequenos de k com relação ao número de critérios, m . Na Seção 4.4.1 nós exploramos o caso *1-aditivo* em que, para todas as coalizões de dois ou mais elementos, $\eta(G) = 0$. Na Seção 4.4.2 nós exploramos o caso *2-aditivo* em que, para todas as coalizões de três ou mais elementos, $\eta(G) = 0$.

Podemos agora verificar como fica a transformação para o índice de interação generalizado da medida *fuzzy*, calculada na Seção 4.1. Como esse caso é um problema de decisão em que se avaliam as alternativas sobre 4 critérios, temos $2^4 = 16$ índices de interação generalizado e a matriz de transformação \mathbf{T} , calculada usando (4.2), é dada por:

$$\mathbf{T} = \begin{bmatrix} 1 & -1/2 & -1/2 & -1/2 & -1/2 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & -1/30 \\ 1 & 1/2 & -1/2 & -1/2 & -1/2 & -1/3 & -1/3 & -1/3 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & -1/30 \\ 1 & -1/2 & 1/2 & -1/2 & -1/2 & -1/3 & 1/6 & 1/6 & -1/3 & -1/3 & 1/6 & 1/6 & 1/6 & 0 & 1/6 & -1/30 \\ 1 & -1/2 & -1/2 & 1/2 & -1/2 & 1/6 & -1/3 & 1/6 & -1/3 & 1/6 & -1/3 & 1/6 & 0 & 1/6 & 1/6 & -1/30 \\ 1 & -1/2 & -1/2 & -1/2 & 1/2 & 1/6 & 1/6 & -1/3 & 1/6 & -1/3 & -1/3 & 0 & 1/6 & 1/6 & 1/6 & -1/30 \\ 1 & 1/2 & 1/2 & -1/2 & -1/2 & 1/6 & -1/3 & -1/3 & -1/3 & -1/3 & 1/6 & -1/6 & -1/6 & 1/6 & 1/6 & 2/15 \\ 1 & 1/2 & -1/2 & 1/2 & -1/2 & -1/3 & 1/6 & -1/3 & -1/3 & 1/6 & -1/3 & -1/6 & 1/6 & -1/6 & 1/6 & 2/15 \\ 1 & 1/2 & -1/2 & -1/2 & 1/2 & -1/3 & -1/3 & 1/6 & 1/6 & -1/3 & -1/3 & 1/6 & -1/6 & -1/6 & 1/6 & 2/15 \\ 1 & -1/2 & 1/2 & 1/2 & -1/2 & -1/3 & -1/3 & 1/6 & 1/6 & -1/3 & -1/3 & -1/6 & 1/6 & 1/6 & -1/6 & 2/15 \\ 1 & -1/2 & 1/2 & -1/2 & 1/2 & -1/3 & 1/6 & -1/3 & -1/3 & 1/6 & -1/3 & 1/6 & -1/6 & 1/6 & -1/6 & 2/15 \\ 1 & -1/2 & -1/2 & 1/2 & 1/2 & 1/6 & -1/3 & -1/3 & -1/3 & -1/3 & 1/6 & 1/6 & 1/6 & -1/6 & -1/6 & 2/15 \\ 1 & 1/2 & 1/2 & 1/2 & -1/2 & 1/6 & 1/6 & -1/3 & 1/6 & -1/3 & -1/3 & 0 & -1/6 & -1/6 & -1/6 & -1/30 \\ 1 & 1/2 & 1/2 & -1/2 & 1/2 & 1/6 & -1/3 & 1/6 & -1/3 & 1/6 & -1/3 & -1/6 & 0 & -1/6 & -1/6 & -1/30 \\ 1 & 1/2 & -1/2 & 1/2 & 1/2 & -1/3 & 1/6 & 1/6 & -1/3 & -1/3 & 1/6 & -1/6 & -1/6 & 0 & -1/6 & -1/30 \\ 1 & -1/2 & 1/2 & 1/2 & 1/2 & -1/3 & -1/3 & -1/3 & 1/6 & 1/6 & 1/6 & -1/6 & -1/6 & -1/6 & 0 & -1/30 \\ 1 & 1/2 & 1/2 & 1/2 & 1/2 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 1/6 & 0 & 0 & 0 & 0 & -1/30 \end{bmatrix}.$$

Na Figura 4.2, apresentamos os índices calculados por meio da Equação (4.5). Os valores calculados são:

- $\eta(\{\emptyset\}) = 0,5$;
- $\eta(\{1\}) = \eta(\{2\}) = \eta(\{3\}) = \eta(\{4\}) = 0,25$

O restante dos termos, que apresentam coalizões de dois ou mais elementos, são todos iguais a zero. Este é o caso *1-aditivo* que, como dito anteriormente, pode ser representado por um vetor esparso.

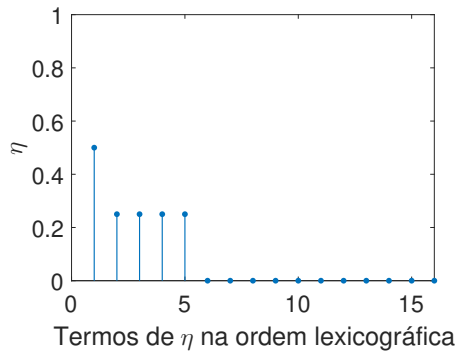


Figura 4.2 – Valores dos índices de interação para o caso aditivo.

Motivado pelo fato de que modelos *k-aditivo*, quando representados no domínio do índice de interação, são descritos por vetores esparsos, o método proposto para esse trabalho tem o objetivo de identificar as capacidades de Choquet por meio do problema de otimização visto no capítulo anterior, porém trabalhando no domínio do índice de interação, no qual a esparsidade faz mais sentido, ou seja, os vetores transformados para esse domínio tendem a ter uma proporção significativa de elementos nulos.

4.3 Regularização esparsa no domínio do índice de interação

Como visto nas Seções 4.1 e 4.2, o vetor de índices de interação generalizado é um vetor esparso e, portanto, faz mais sentido resolver o problema de otimização proposto em (3.4), trabalhando no domínio de η , ao invés de trabalhar no domínio de μ .

Para se realizar a otimização quadrática no domínio de η , deve-se fazer a substituição expressa em (4.5) no problema (3.4), que pode ser reescrito da seguinte maneira:

$$\begin{aligned} \min_{\eta} \quad & \frac{1}{2} \eta^t \tilde{\mathbf{Q}} \eta + \tilde{\mathbf{f}}^t \eta \\ \text{s.t.} \quad & \tilde{\mathbf{R}} \eta \leq \mathbf{s}. \end{aligned} \tag{4.6}$$

onde $\tilde{\mathbf{Q}} = \mathbf{T}^t \mathbf{Q} \mathbf{T}$, $\tilde{\mathbf{f}} = \mathbf{f}^t \mathbf{T}$ e $\tilde{\mathbf{R}} = \mathbf{R} \mathbf{T}$.

O problema apresentado em (4.6) também representa um problema inverso e, devido às condições vistas na seção 3.1.1.1, também é mal-posto. Uma das maneiras de se resolver um problema mal-posto é através da regularização, que pode ser feita adicionando-se um termo de modo a promover esparsidade na solução. Do mesmo modo que foi feito no domínio de $\boldsymbol{\mu}$, o termo de regularização a ser adicionado nesse caso é dado pela norma l_1 do vetor $\boldsymbol{\eta}$, o que leva a soluções esparsas. O problema pode então ser escrito como:

$$\begin{aligned} \min_{\boldsymbol{\eta}} \quad & \frac{1}{2} \boldsymbol{\eta}^t \tilde{\mathbf{Q}} \boldsymbol{\eta} + \tilde{\mathbf{f}}^t \boldsymbol{\eta} + \lambda \|\boldsymbol{\eta}\|_1 \\ \text{s.t.} \quad & \tilde{\mathbf{R}} \boldsymbol{\eta} \leq \mathbf{s}, \end{aligned} \tag{4.7}$$

onde λ é um parâmetro não negativo que controla a regularização. Como dito anteriormente, esse problema também é conhecido como regressão com penalização (OSBORNE; PRESNELL; TURLACH, 2000), pois o objetivo do termo de regularização é manter as estimativas de $\boldsymbol{\eta}$ as mais esparsas possível. Quanto maior o valor de λ , maior a penalização, ou seja, aparecerão mais termos nulos em $\boldsymbol{\eta}$.

Como vimos no capítulo anterior, o problema em (4.7) equivale ao LASSO, que pode ser expresso da seguinte maneira:

$$\begin{aligned} \min_{\boldsymbol{\eta}} \quad & \frac{1}{2} \boldsymbol{\eta}^t \tilde{\mathbf{Q}} \boldsymbol{\eta} + \tilde{\mathbf{f}}^t \boldsymbol{\eta} \\ \text{s.t.} \quad & \tilde{\mathbf{R}} \boldsymbol{\eta} \leq \mathbf{s} \\ & \|\boldsymbol{\eta}\|_1 \leq t, \end{aligned} \tag{4.8}$$

onde t é um parâmetro de ajuste não negativo que controla o nível de esparsidade. Novamente, a solução para o problema apresentado em (4.8) é o NNVM (TIBSHIRANI, 1996), cuja ideia é linearizar a norma l_1 por meio de uma representação na qual se dobra o número de variáveis de $\boldsymbol{\eta}$, ou seja, $[\eta_0^+, \eta_1^+, \eta_2^+, \dots, \eta_0^-, \eta_1^-, \eta_2^-, \dots]$, onde $\eta_j = \eta_j^+ - \eta_j^-$, e se acrescentam as seguintes restrições:

$$\begin{aligned} \eta_j^+ &\geq 0, \eta_j^- \geq 0, \\ \sum_{j=1}^{2^m} (\eta_j^+ + \eta_j^-) &\leq t. \end{aligned}$$

Novamente, é importante notar que, para o subíndice j , também é necessário considerar

as coalizões de critérios, tais como: $\{1, 2\}$, $\{1, 3\}$, $\{2, 3\}$, etc. O problema de otimização pode ser escrito, então, da seguinte maneira:

$$\begin{aligned} \min_{\hat{\boldsymbol{\eta}}} \quad & \frac{1}{2} \hat{\boldsymbol{\eta}}^t \hat{\mathbf{Q}} \hat{\boldsymbol{\eta}} + \hat{\mathbf{f}}^t \hat{\boldsymbol{\eta}} \\ \text{s.t.} \quad & \hat{\mathbf{R}} \hat{\boldsymbol{\eta}} \leq \mathbf{s} \\ & \|\hat{\boldsymbol{\eta}}\|_1 \leq t. \end{aligned} \quad (4.9)$$

onde: $\hat{\boldsymbol{\eta}} = [\eta_{\emptyset}^+, \eta_1^+, \eta_2^+, \dots, \eta_{\emptyset}^-, \eta_1^-, \eta_2^-, \dots]$, $\hat{\mathbf{Q}} = \begin{bmatrix} \tilde{\mathbf{Q}} & -\tilde{\mathbf{Q}} \\ -\tilde{\mathbf{Q}} & \tilde{\mathbf{Q}} \end{bmatrix}$, $\hat{\mathbf{f}} = \begin{bmatrix} \tilde{\mathbf{f}} \\ -\tilde{\mathbf{f}} \end{bmatrix}$ e $\hat{\mathbf{R}} = \begin{bmatrix} \tilde{\mathbf{R}} & -\tilde{\mathbf{R}} \end{bmatrix}$.

4.4 Problema de identificação: resultados numéricos

Com o objetivo de validar o método proposto, nós realizamos um conjunto de experimentos, conforme ilustrado na Figura 4.3. Eles consistem na identificação das capacidades de Choquet $\boldsymbol{\mu}$ ou dos índices de interação generalizados $\boldsymbol{\eta}$, conhecendo-se a matriz de decisão \mathbf{P} e o vetor de utilidades \mathbf{u} . O objetivo é fazer com que a saída da integral de Choquet, $\boldsymbol{\gamma}$, seja a mais próxima das utilidades, \mathbf{u} .

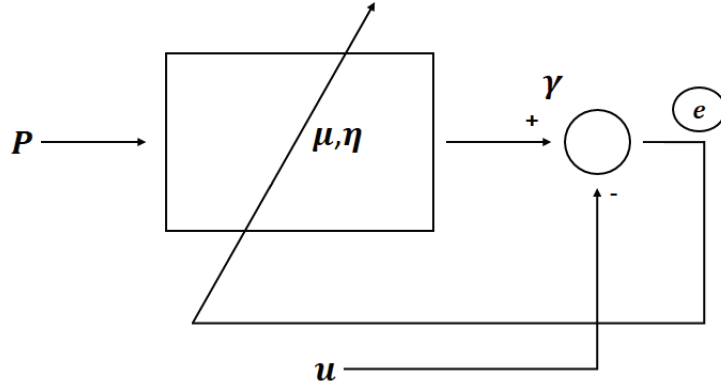


Figura 4.3 – Esquema representativo dos experimentos realizados.

Os experimentos foram realizados utilizando-se as quatro estratégias a seguir:

- Otimização no domínio de $\boldsymbol{\mu}$ sem regularização, conforme o problema de otimização (3.4);
- Otimização no domínio de $\boldsymbol{\mu}$ com um termo de regularização, conforme o problema de otimização (3.15).

Nessas duas primeiras estratégias, se obtém as estimativas das capacidades de Choquet $\boldsymbol{\mu}$. Para se obter as estimativas dos índices de interação generalizados $\boldsymbol{\eta}$, se utiliza a equação (4.5).

- (c) Otimização no domínio de $\boldsymbol{\eta}$ sem regularização, conforme o problema de otimização (4.6);
- (d) Otimização no domínio de $\boldsymbol{\eta}$ com um termo de regularização, conforme o problema de otimização (4.9).

Nesses duas últimas estratégias, se obtém as estimativas dos índices de interação generalizados $\boldsymbol{\eta}$. Para se obter as estimativas das capacidades de Choquet $\boldsymbol{\mu}$, se utiliza a equação (4.5).

Nas seções seguintes, serão apresentados experimentos nos quais os dados foram gerados utilizando-se os modelos *1-aditivo* e *2-aditivo*.

4.4.1 Problema de identificação: resultados para o caso *1-aditivo*

Para o caso *1-aditivo*, nós geramos uma matriz de decisão \mathbf{P} com distribuição uniforme de probabilidade no intervalo $[0, 1]$; nós consideramos 7 critérios e 100 alternativas. O vetor de utilidade \mathbf{u} foi gerado linearmente, da seguinte maneira: $\mathbf{u} = \mathbf{w}^t \mathbf{P}$, onde $\mathbf{w} = [1/7; 1/7; 1/7; 1/7; 1/7; 1/7; 1/7]$ é o vetor de pesos. Com relação ao valor do parâmetro que controla a regularização (t), ele possui uma dependência direta com o número de critérios: quanto maior o número de critérios, maior o valor de t para que se consiga uma melhor regularização do problema. Para o nosso caso, após alguns experimentos, verificou-se que se esse valor é grande ($t > 10$), seria como se o problema não estivesse sujeito à restrição de regularização. Se esse valor é pequeno ($t < 3$), os valores das soluções tendem todos a zerar. Nesse sentido, a melhor configuração observada foi com $t = 5$. Todos os experimentos foram resolvidos com o *solver* de otimização quadrática do *Matlab*.

Para comparação, foram calculados o erro médio quadrático (MSE - do inglês *mean squared error*) e o erro médio absoluto (MAE - do inglês *mean absolute error*, que são dados por (DEGROOT; SCHERVISH, 2012):

$$MSE = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbf{x}} - \mathbf{x})^2, \quad (4.10)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |\hat{\mathbf{x}} - \mathbf{x}|, \quad (4.11)$$

onde $\hat{\mathbf{x}}$ e \mathbf{x} são, respectivamente, os valores estimados e os valores teóricos. Nós calculamos o MSE e o MAE para as estratégias (a), (b), (c) e (d) considerando três casos:

1. entre o vetor de utilidade \mathbf{u} e o valor estimado da integral de Choquet $\boldsymbol{\gamma}$ (erro de representação);

2. entre os valores teóricos e estimados das capacidades de Choquet;
3. entre os valores teóricos e estimados dos índices de interação generalizado.

A Figura 4.4 mostra a dispersão bi-dimensional representando a correlação entre o vetor de utilidade \mathbf{u} e o valor estimado da integral de Choquet γ para as quatro estratégias. A Tabela 4.1 mostra o MSE e o MAE para os mesmos casos. Como pode ser notado, a integral de Choquet foi capaz de explicar as utilidades observadas em todos os casos, exceto no caso em que o problema de otimização foi realizado no domínio de μ com um termo de regularização. Além disso, os valores de MSE e MAE nesse caso são maiores do que nos outros três casos. Este baixo desempenho pode ser atribuído ao fato de que, mesmo para o caso *1-aditivo* considerado aqui, μ não é um vetor esparsos. Portanto, ao se regularizar o problema, adicionando um termo que impõe esparsidade ao vetor μ , se encontra uma solução esparsa, a qual não representa com fidelidade a medida *fuzzy* que se está tentando estimar.

Tabela 4.1 – Caso *1-aditivo*: MSE e MAE entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	2.01×10^{-7}	8.68×10^{-2}	3.00×10^{-7}	1.95×10^{-8}
MAE	3.10×10^{-4}	2.80×10^{-1}	4.10×10^{-4}	1.09×10^{-4}

A Figura 4.5 ilustra os valores das capacidades de Choquet identificados por cada estratégia. Novamente, podemos notar que, devido à restrição de esparsidade, a estratégia (b) (otimização em μ com um termo de regularização) não foi capaz de apresentar uma boa estimativa das capacidades que foram usadas no processo de geração dos dados. Podemos notar que há muitos termos iguais a zero e isso acontece por causa da restrição imposta pela regularização do problema. Esta restrição leva a uma solução esparsa que não condiz com os parâmetros utilizados na geração dos dados. Na Tabela 4.2 percebemos novamente que os valores de MSE e MAE no domínio de μ com um termo de regularização são maiores que nos outros três cenários. Por outro lado, as estratégias (a), (c) e (d) foram capazes de apresentar uma boa estimativa das capacidades, sendo que a estratégia (d) apresentou melhores resultados.

A Figura 4.6 mostra os valores dos índices de interação generalizados identificados pelos métodos de otimização nas quatro estratégias. É interessante notar que a otimização quadrática no domínio de η com um termo de regularização resultou num vetor esparsos, o qual é muito próximo ao do modelo considerado no processo de geração dos dados. De fato, os dados foram gerados linearmente, com pesos iguais, portanto os primeiros sete índices de interação generalizados (sem considerar o termo relativo ao

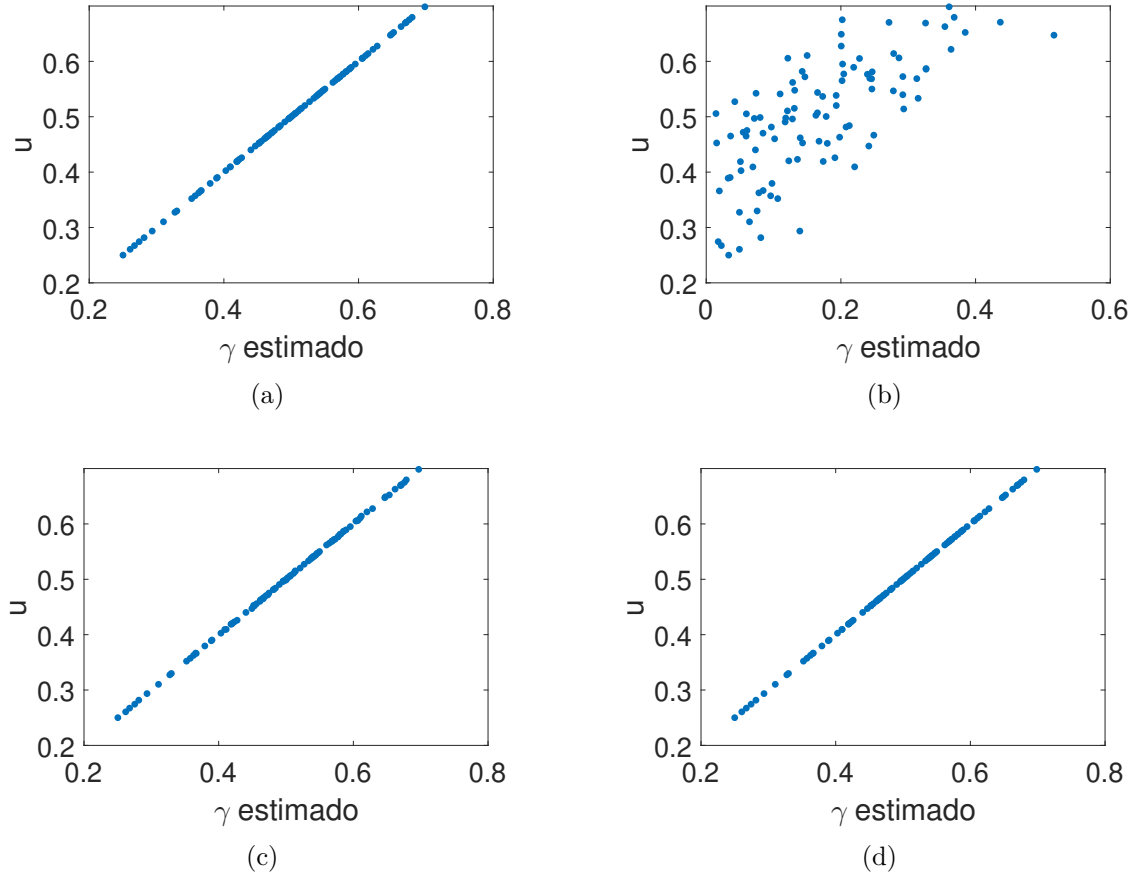


Figura 4.4 – Caso 1-aditivo: Dispersão bi-dimensional entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

Tabela 4.2 – Caso 1-aditivo: MSE e MAE entre os valores teóricos das capacidades de Choquet e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	5.54×10^{-4}	2.51×10^{-1}	2.20×10^{-4}	1.68×10^{-5}
MAE	1.74×10^{-2}	4.65×10^{-1}	1.15×10^{-2}	2.93×10^{-3}

conjunto vazio), devem ser iguais a $1/7$. Os índices de interação que descrevem as coalizões de dois ou mais critérios devem ser todos nulos. Na Tabela 4.3, percebemos que os valores de MSE e MAE no domínio de η com um termo de regularização são menores do que nas outras três estratégias.

4.4.2 Problema de identificação: resultados para o caso 2-aditivo

Com o objetivo de validar o método proposto para o caso 2-aditivo, nós realizamos simulações com as mesmas condições que para o caso 1-aditivo. Nós geramos

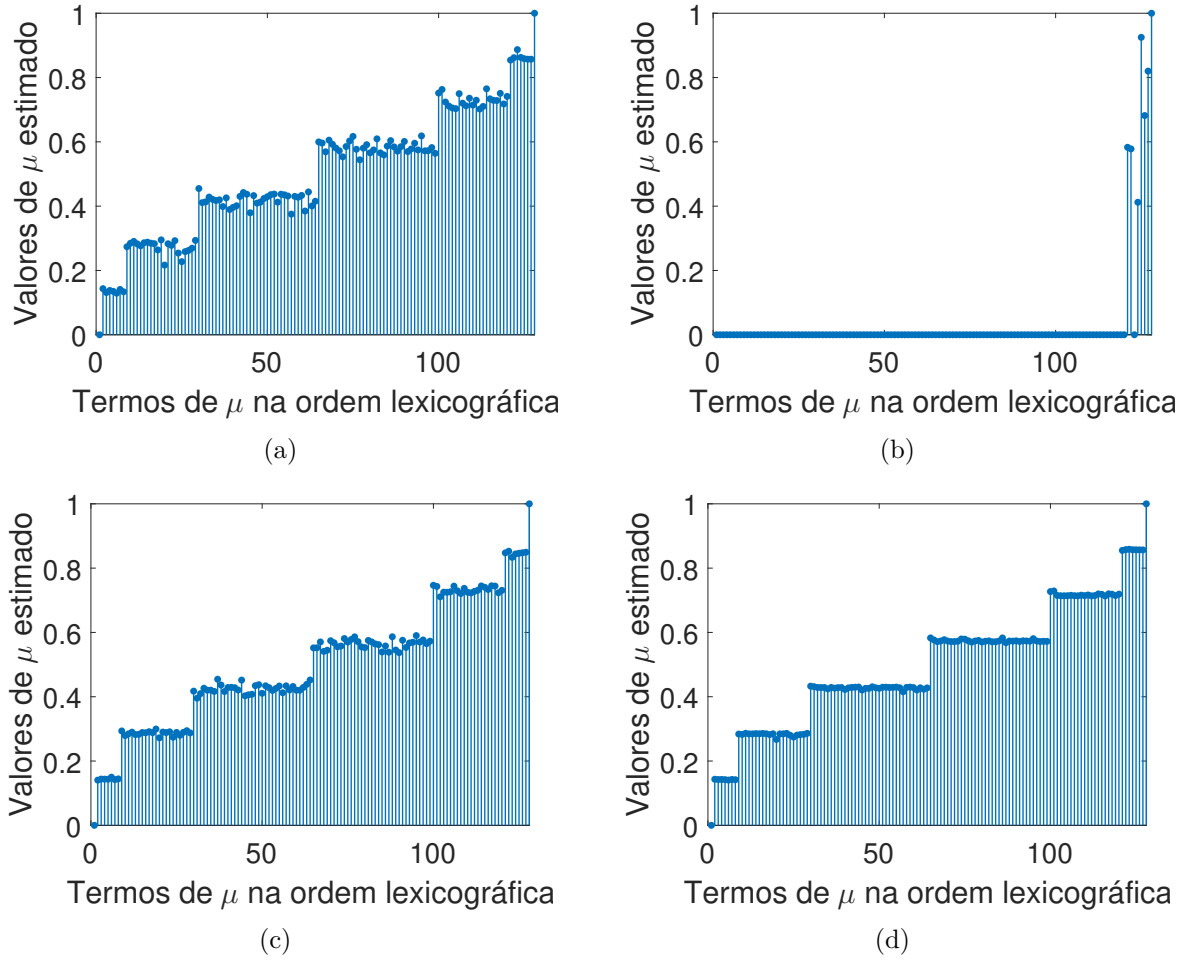


Figura 4.5 – Caso 1-aditivo: Valores das capacidades de Choquet identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

Tabela 4.3 – Caso 1-aditivo: MSE e MAE entre os valores teóricos dos índices de interação generalizados e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	6.14×10^{-2}	2.39×10^{-2}	6.35×10^{-3}	1.55×10^{-5}
MAE	1.72×10^{-1}	1.29×10^{-1}	3.07×10^{-2}	3.00×10^{-3}

aleatoriamente uma matriz de decisão \mathbf{P} com distribuição uniforme de probabilidade no intervalo $[0, 1]$, tendo 7 critérios e 100 alternativas.

O vetor com os índices de interação, η , foram gerados da seguinte maneira: os primeiros sete índices de interação generalizados (sem considerar o termo correspondente ao conjunto vazio), que são os índices de Shapley, foram iguais a $1/7$. Os índices de interação correspondentes a duas coalizões foram construídos aleatoriamente com distribuição uniforme de probabilidade, sujeito à seguinte restrição (GRABISCH; LABREUCHE, 2010;

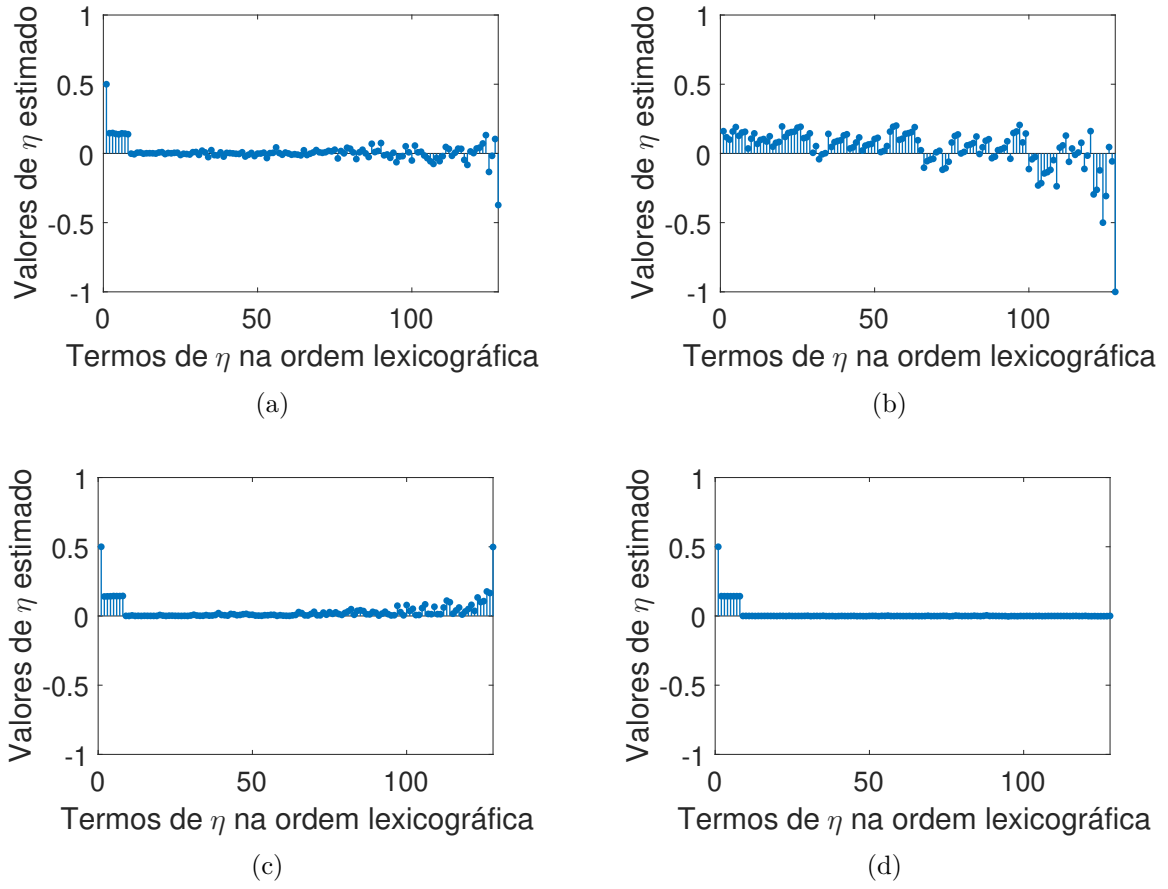


Figura 4.6 – Caso *1-aditivo*: Valores dos índices de interação generalizados identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

DUARTE, 2018):

$$\eta_j - \frac{1}{2} \sum_{k \neq j} |\eta_{jk}| \geq 0, \forall j. \quad (4.12)$$

Essa restrição garante a monotonicidade no caso *2-aditivo*. Os índices correspondentes aos termos com mais de duas coalizões foram todos iguais a zero e o vetor de utilidade \mathbf{u} foi gerado da seguinte maneira (GRABISCH; LABREUCHE, 2010; DUARTE, 2018):

$$u_i = \sum_{\eta_{jk} > 0} (P_{ij} \wedge P_{ik}) \eta_{jk} + \sum_{\eta_{jk} < 0} (P_{ij} \vee P_{ik}) |\eta_{jk}| + \sum_{j=1}^m P_{ij} \left(\eta_j - \frac{1}{2} \sum_{k \neq j} |\eta_{jk}| \right). \quad (4.13)$$

Essa expressão é uma representação da integral de Choquet para o caso *2-aditivo*, na qual pode se verificar que os índices de interação η_{jk} representam o caráter não linear no modelo de agregação, enquanto que os índices de Shapley η_j representam a parte linear do modelo (GRABISCH; LABREUCHE, 2010; DUARTE, 2018).

Todos os experimentos foram resolvidos com o *solver* de otimização quadrática do *Matlab* e, após alguns experimentos, a melhor configuração para o valor do parâmetro que controla a regularização foi $t = 3$. Novamente, em nossas análises, foram calculados o MSE e o MAE para as estratégias (a), (b), (c) e (d) considerando três casos:

1. entre o vetor de utilidade \mathbf{u} e o valor estimado da integral de Choquet γ (erro de representação);
2. entre os valores teóricos e estimados das capacidades de Choquet;
3. entre os valores teóricos e estimados dos índices de interação generalizado.

A Figura 4.7 ilustra a dispersão bi-dimensional entre o vetor de utilidade \mathbf{u} e o valor estimado da integral de Choquet γ para as estratégias (a), (b), (c) e (d). A Tabela 4.4 mostra o MSE e o MAE para o mesmo caso. Como no caso *1-aditivo*, a integral de Choquet foi capaz de explicar os dados em todos as estratégias, exceto para a (c). Novamente, a imposição da esparsidade no domínio de μ limita a capacidade do modelo de explicar os dados observados. Os valores de MSE e MAE no domínio de μ com um termo de regularização são maiores que nos outros três cenários. Isto mostra que, mesmo no caso *2-aditivo*, não é razoável considerar esparsidade no domínio de μ .

Tabela 4.4 – Caso *2-aditivo*: MSE e MAE entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	5.94×10^{-9}	5.70×10^{-2}	5.87×10^{-7}	6.13×10^{-9}
MAE	4.79×10^{-5}	2.23×10^{-1}	5.70×10^{-4}	6.22×10^{-5}

A Figura 4.8 ilustra os valores das capacidades de Choquet identificados nas quatro estratégias. Novamente, somente a estratégia (b) (otimização no domínio de μ com regularização) não foi capaz de apresentar uma boa estimativa das capacidades que foram usadas no processo de geração dos dados. Como antes, a razão para isso é a restrição imposta pela regularização do problema, que leva a uma solução esparsa. Mesmo no caso *2-aditivo*, podemos ver que a solução não é esparsa. Na Tabela 4.5 vemos novamente que os valores de MSE e MAE no domínio de μ com um termo de regularização são maiores que nas outras três estratégias.

Os índices de interação generalizados obtidos são apresentados na Figura 4.9. Nós podemos ver que somente a estratégia (d) (otimização no domínio de η com regularização) forneceu um vetor η que se assemelha aos parâmetros de um modelo *2-aditivo*. Na Tabela 4.6, vemos novamente que os valores de MSE e MAE no domínio de η com um

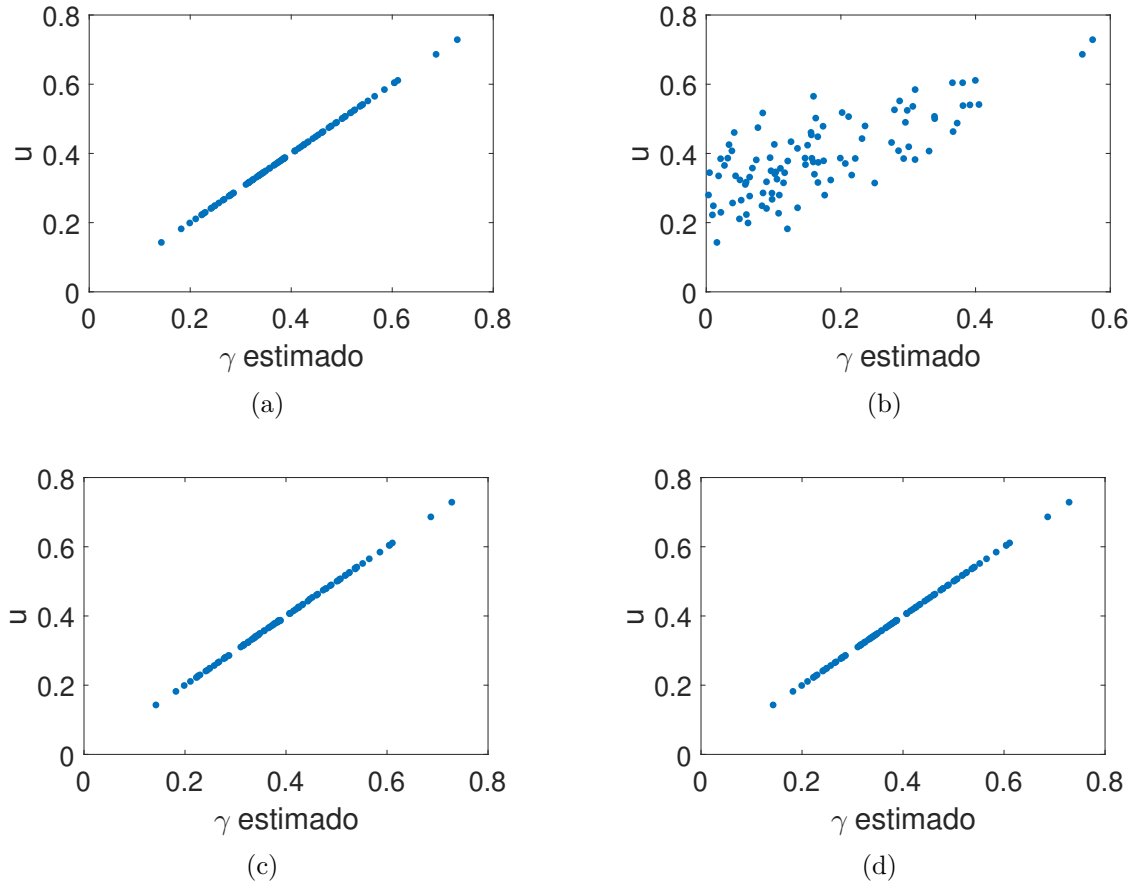


Figura 4.7 – Caso 2-aditivo: Dispersão bi-dimensional entre o vetor de utilidades \mathbf{u} e a estimativa da integral de Choquet γ para a otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

Tabela 4.5 – Caso 2-aditivo: MSE e MAE entre os valores teóricos das capacidades de Choquet e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	7.38×10^{-4}	1.31×10^{-1}	8.44×10^{-4}	3.56×10^{-4}
MAE	2.25×10^{-2}	3.20×10^{-1}	2.33×10^{-2}	1.88×10^{-2}

termo de regularização são menores que nas outras três estratégias. Neste caso, os primeiros sete índices de interação generalizados (sem considerar o termo relativo ao conjunto vazio) devem ser iguais a $1/7$. Os índices de interação que descrevem as coalizões de dois critérios são não nulos. Os outros índices de interação, que descrevem as coalizões de mais de dois elementos, devem ser todos nulos, e somente nesse caso eles são muito próximos de zero.

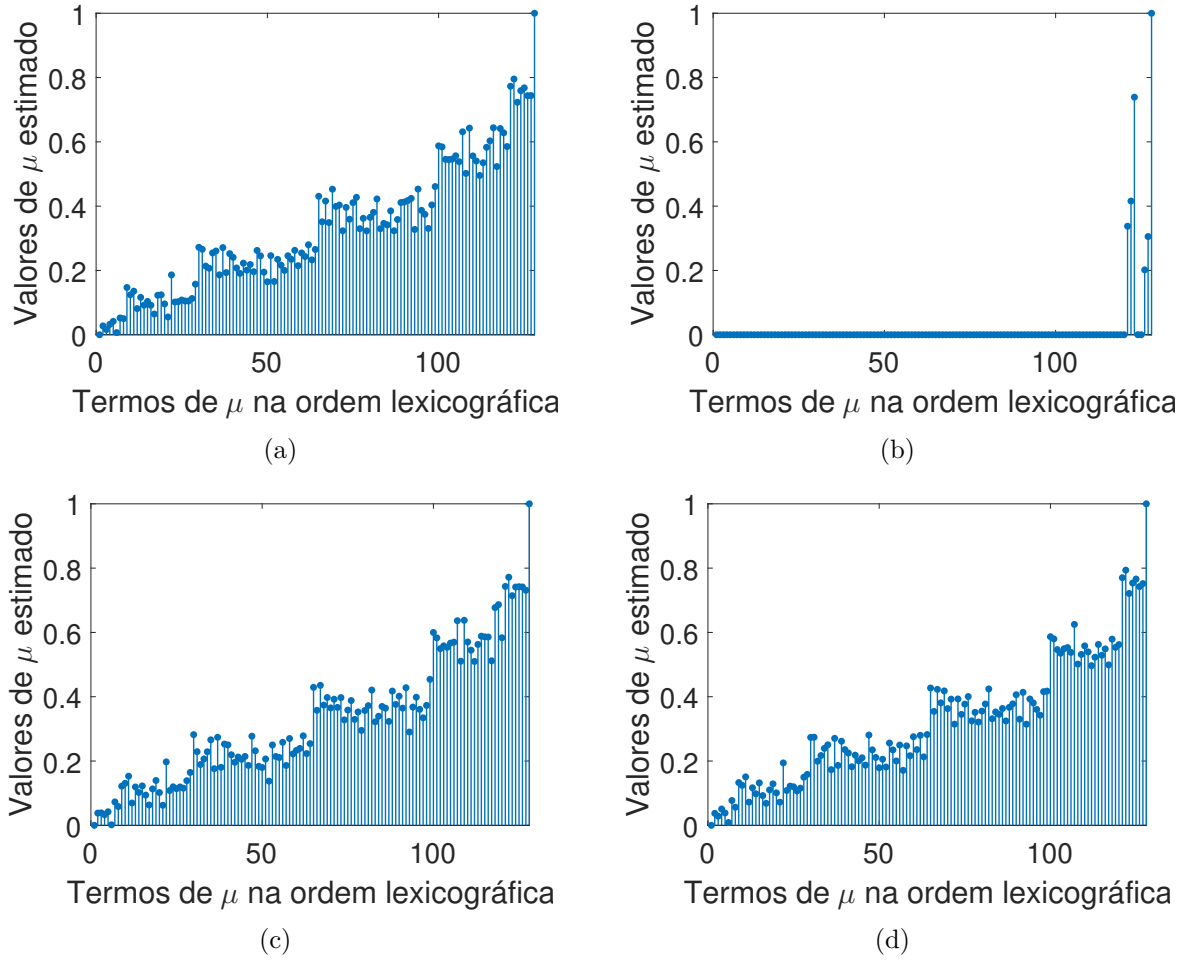


Figura 4.8 – Caso 2-aditivo: Valores das capacidades de Choquet identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

Tabela 4.6 – Caso 2-aditivo: MSE e MAE entre os valores teóricos dos índices de interação generalizados e aqueles identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

	(a)	(b)	(c)	(d)
MSE	3.58×10^{-2}	7.89×10^{-3}	7.87×10^{-3}	3.33×10^{-6}
MAE	1.52×10^{-1}	7.29×10^{-2}	4.33×10^{-2}	7.50×10^{-4}

4.5 Sumário

No presente capítulo, apresentamos a principal contribuição dessa tese, que é a identificação dos parâmetros da integral de Choquet, baseada em representações esparsas. Porém, ao invés de trabalhar no domínio da medida *fuzzy*, a nossa proposta abordou o problema no domínio do índice de interação. Realizamos um conjunto de experimentos numéricos nos quais os dados foram gerados utilizando-se os modelos 1-aditivo e 2-aditivo. Em ambos os casos, verificamos que a nossa proposta alcançou melhores resultados para o

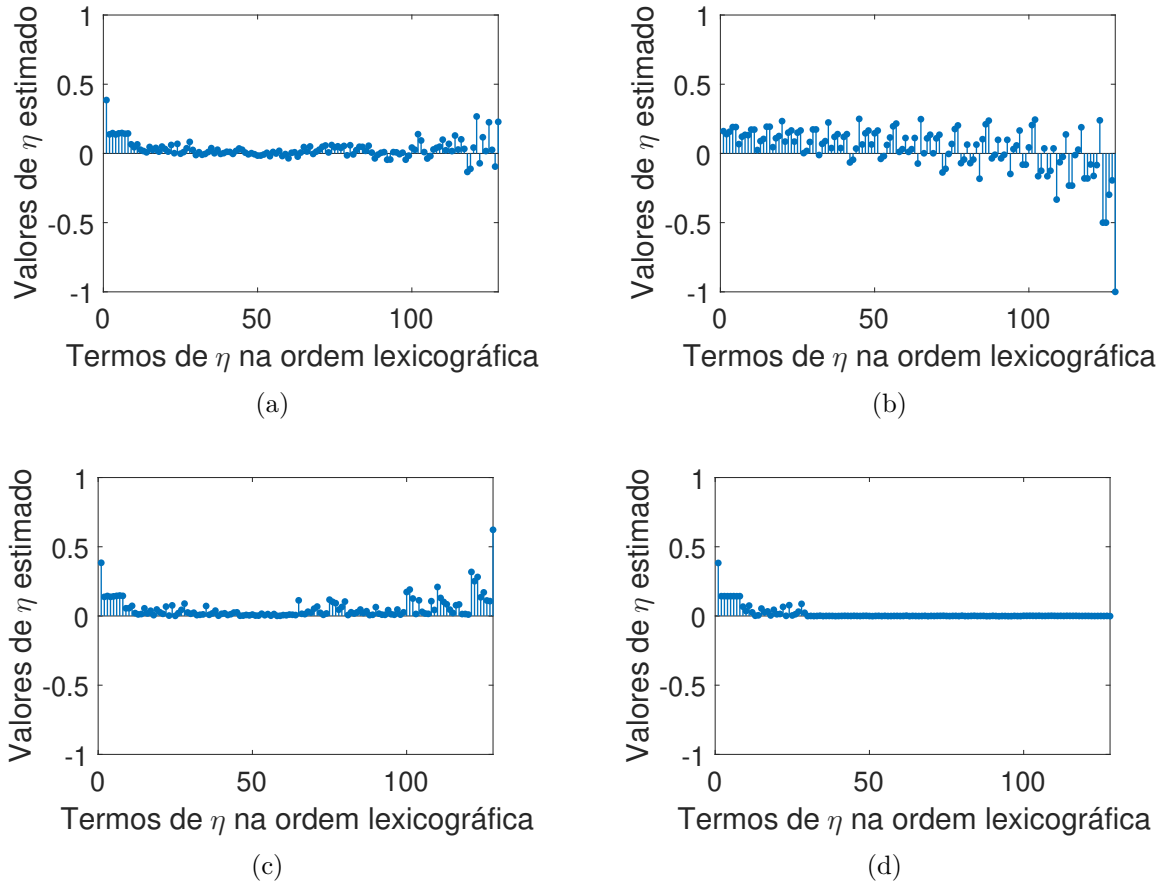


Figura 4.9 – Caso *2-aditivo*: Valores dos índices de interação generalizados identificados na otimização quadrática no domínio de: (a) μ , (b) μ com um termo de regularização, (c) η , (d) η com um termo de regularização.

problema de identificação.

De fato, mesmo em modelos mais simples como o *1-aditivo* e o *2-aditivo*, o vetor representado no domínio da medida *fuzzy* não apresenta, via de regra, uma proporção significativa de elementos nulos. Deste modo, ao se considerar um termo de regularização baseado na norma l_1 na formulação parametrizada pela medida *fuzzy*, a solução do problema de otimização promove soluções esparsas que, portanto, não são capazes de estimar corretamente a medida *fuzzy* que foi utilizada na geração dos dados.

Por outro lado, no domínio do índice de interação, os vetores gerados utilizando-se os modelos *1-aditivo* e *2-aditivo* apresentam uma proporção significativa de elementos nulos. Portanto, ao se resolver o problema de identificação com regularização neste domínio, a solução encontrada representou corretamente o vetor de índices de interação que se estava tentando estimar.

Com relação à complexidade do problema resolvido para 7 critérios e 100 alternativas, verificamos que o tempo de execução dos experimentos foi de aproximadamente 30 minutos para as quatro estratégias consideradas: otimização no domínio de μ sem

regularização, otimização no domínio de μ com um termo de regularização, otimização no domínio de η sem regularização e otimização no domínio de η com um termo de regularização.

5 Análise de unicidade

Na Seção 3.1.1.1, foi visto que uma das condições de Hadamard para um problema inverso ser considerado bem-posto é ele possuir solução única. Nesse capítulo será apresentada uma análise que mostra sob quais condições se terá uma solução única para o problema de identificação das capacidades de Choquet. Essa análise leva em consideração se o sistema linear em questão é sobredeterminado ou subdeterminado, e as ferramentas utilizadas em cada caso são, respectivamente, o posto e o *spark* de uma matriz¹. Primeiramente, discutiremos as situações nas quais o sistema linear é sobredeterminado ou subdeterminado e as análises de unicidade para cada caso. Então, serão apresentados resultados numéricos, considerando diferentes casos.

5.1 Análise de unicidade das soluções

O problema de identificação apresentado nos capítulos anteriores visa encontrar um vetor de capacidades que consiga explicar os dados conhecidos: a matriz de decisão \mathbf{P} e o vetor de utilidades \mathbf{u} . Isto é, o objetivo é encontrar um vetor de capacidades $\boldsymbol{\mu}$, tal que $\mathbf{M}^t \boldsymbol{\mu} = \mathbf{u}$. Como visto em 3.1.1.1, esse é um problema inverso que demanda uma análise da unicidade da solução. Para realizar tal análise, o problema pode ser escrito como um sistema linear, conforme visto na Equação (2.12) e replicado aqui:

$$\gamma^\mu(P_{i1}, \dots, P_{im}) = \mathbf{M}^t \boldsymbol{\mu}, \quad (5.1)$$

onde \mathbf{M} é dada por:

$$\mathbf{M} = \begin{bmatrix} 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(m)} - P_{1(m-1)} & P_{2(m)} - P_{2(m-1)} & \dots & P_{n(m)} - P_{n(m-1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(2)} - P_{1(1)} & P_{2(2)} - P_{2(1)} & \dots & P_{n(2)} - P_{n(1)} \\ 0 & 0 & \dots & 0 \\ \vdots & \vdots & & \vdots \\ P_{1(1)} - P_{1(0)} & P_{2(1)} - P_{2(0)} & \dots & P_{n(1)} - P_{n(0)} \end{bmatrix}, \quad (5.2)$$

¹ O caso determinado será estudado juntamente com o caso sobredeterminado, com o objetivo de separarmos o problema em dois casos apenas: aqueles abordados pelo posto da matriz e aqueles abordados pelo *spark* da matriz.

cuja dimensão é $2^m \times n$. Pela análise da matriz \mathbf{M}^t , que é a transposta da matriz \mathbf{M} e cuja dimensão é $n \times 2^m$, é possível identificar se a solução encontrada é única. Pode-se ter dois casos distintos:

1. Se $2^m \leq n$, o sistema linear apresentado é um sistema sobredeterminado, quando se tem mais equações do que incógnitas, ou determinado, quando o número de equações é igual ao número de incógnitas. Nesse caso, a análise será realizada através do posto da matriz \mathbf{M}^t ;
2. Se $2^m > n$, o sistema linear apresentado é um sistema subdeterminado, pois se tem mais incógnitas do que equações. Nesse caso, a análise será realizada através do *spark* da matriz \mathbf{M}^t .

Esses dois casos, serão explicados nas Seções 5.1.1, que apresenta uma abordagem clássica, e 5.1.2, que apresenta uma abordagem mais recente para a análise de unicidade.

5.1.1 Análise de unicidade para sistemas sobredeterminados

Segundo o teorema de *Rouché-Capelli* (SHAFAREVICH; REMIZOV, 2012), dado um sistema linear $\mathbf{Ax} = \mathbf{b}$, com matriz de coeficientes \mathbf{A} , de tamanho $p \times q$, e matriz aumentada $\mathbf{A}|\mathbf{b}$, pode-se chegar a três conclusões com relação ao número de soluções de um sistema linear:

1. O sistema linear é inconsistente (não tem solução) se e somente se $\text{posto}(\mathbf{A}) < \text{posto}(\mathbf{A}|\mathbf{b})$;
2. o sistema linear tem solução única se e somente se $\text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}|\mathbf{b}) = q$;
3. o sistema linear tem infinitas soluções se e somente se $\text{posto}(\mathbf{A}) = \text{posto}(\mathbf{A}|\mathbf{b}) < q$.

Para o problema em questão, $\mathbf{M}^t \boldsymbol{\mu} = \mathbf{u}$, no qual \mathbf{M}^t tem dimensão $n \times 2^m$, podemos definir o seguinte critério para garantir a unicidade da solução encontrada:

- Se $\text{posto}(\mathbf{M}^t) = \text{posto}(\mathbf{M}^t|\mathbf{u}) = 2^m$, pode-se concluir que o vetor de capacidades é único.

5.1.2 Análise de unicidade para sistemas subdeterminados

Consideremos agora o mesmo sistema linear $\mathbf{Ax} = \mathbf{b}$, com matriz de coeficientes \mathbf{A} , de tamanho $p \times q$. Quando esse sistema possui $p < q$, ele é subdeterminado e, portanto, possui infinitas soluções. Como visto na Seção 3.2.1, quando um sistema linear apresenta

a existência de mais de uma solução, há a necessidade de se estabelecer um critério para a escolha de uma solução, o que pode ser promovido pela inserção de um termo de regularização, que geralmente explora algum tipo de informação *a priori* sobre o vetor de interesse. Uma das maneiras de se fazer isso é considerar que a solução é esparsa, ou seja, apresenta proporção significativa de elementos nulos, o que leva a observação da unicidade (BARANIUK, 2007).

Quando o sistema linear é subdeterminado, a análise de unicidade pode ser feita utilizando o conceito de *spark*, originalmente desenvolvido na área de *compressed sensing* (DONOHO; ELAD, 2003), (DUARTE; ELDAR, 2011). O *spark* de uma matriz é definido como (DUARTE; ELDAR, 2011):

Definição 5.1.1. Dada uma matriz \mathbf{A} , de tamanho $p \times q$, o $\text{spark}(\mathbf{A})$ é definido como o menor número k tal que exista um conjunto de k colunas de \mathbf{A} que são linearmente dependentes.

Matematicamente, pode-se definir o *spark* da seguinte maneira (ELAD, 2010):

$$\text{spark}(A) = \min_{\mathbf{x} \neq \mathbf{0}} \|\mathbf{x}\|_0 \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{0}, \quad (5.3)$$

onde a norma $\|\mathbf{x}\|_0$, dada pela equação (3.9), denota o número de elementos não nulos de \mathbf{x} . Algumas propriedades do *spark* são (ELAD, 2010):

- $\text{spark}(A) = \infty \Leftrightarrow \text{posto}(A) = q$: o *spark* é igual a infinito se e somente se a matriz tem posto completo, ou seja, se todas as colunas de A são linearmente independentes;
- $\text{spark}(A) = 1$ se e somente se a matriz A tem uma coluna nula;
- se $\text{spark}(A) \neq \infty$, então $\text{spark}(A) \leq \text{posto}(A) + 1$.

O exemplo a seguir mostra como calcular o *spark* de uma matriz.

Exemplo 5.1.1. Consideremos a seguinte matriz:

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 2 & 0 & 2 \\ 1 & 2 & 0 & 3 \\ 1 & 0 & -3 & 4 \end{bmatrix}.$$

Podemos verificar que:

- não há nenhuma coluna nula;

- não há nenhum conjunto de 2 colunas que sejam linearmente dependentes;
- mas há um conjunto de 3 colunas que são linearmente dependentes: as três primeiras colunas são linearmente dependentes, pois $\mathbf{c}_1 - \frac{1}{2}\mathbf{c}_2 + \frac{1}{3}\mathbf{c}_3 = \mathbf{0}$.

Portanto, podemos concluir que $\text{spark}(A) = 3$.

Fim do exemplo

Para que se garanta a unicidade da solução em um sistema linear subdeterminado $\mathbf{A}\mathbf{x} = \mathbf{b}$, é necessário que se tenha $\|\mathbf{x}\|_0 < \frac{\text{spark}(A)}{2}$. Portanto, para o problema em questão, $\mathbf{M}^t\boldsymbol{\mu} = \mathbf{u}$, quando $2^m > n$, podemos definir o seguinte critério para garantir a unicidade da solução encontrada:

- Se $\frac{\text{spark}(\mathbf{M}^t)}{2} > \|\boldsymbol{\mu}\|_0$, pode-se concluir que o vetor de capacidades é único.

5.2 Análise da unicidade: resultados numéricos

Nesta seção, a análise de unicidade será detalhada por meio de simulações numéricas. Os experimentos foram realizados em quatro cenários:

1. No primeiro, foram calculados o *posto* e o *spark* para um conjunto de matrizes \mathbf{M}^t , geradas por simulação. O objetivo é verificar a unicidade da solução para o sistema $\mathbf{M}^t\boldsymbol{\mu} = \mathbf{u}$.
2. No segundo cenário, foram inseridas as restrições de igualdade $\mu(\emptyset) = 0$ e $\mu(C) = 1$ ao problema. De fato, as restrições, que vêm dos axiomas da integral de Choquet, agem como um termo de regularização. Portanto, esse caso é mais informativo, no sentido que é mais fácil obter uma solução única quando comparado ao caso sem as restrições. Não foram consideradas aqui as restrições que estão relacionadas à propriedade de monotonicidade da medida *fuzzy*. A análise nesse cenário com as restrições pode ser feita considerando-se a seguinte matriz aumentada:

$$\mathbf{N} = \begin{bmatrix} \mathbf{M}^t \\ \mathbf{R} \end{bmatrix}, \quad (5.4)$$

em que a matriz \mathbf{R} codifica as restrições e, portanto, é dada por:

$$\mathbf{R} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 0 & \dots & 1 \end{bmatrix}, \quad (5.5)$$

cuja dimensão é 2×2^m .

Calculou-se, então, o *posto* e o *spark* para um conjunto de matrizes \mathbf{N} , geradas por simulação, para verificar a unicidade da solução do sistema $\mathbf{N}\boldsymbol{\mu} = \mathbf{u}$.

3. No terceiro cenário, calculou-se o *posto* e o *spark* da matriz $\tilde{\mathbf{M}}$, que é a transformação da matriz \mathbf{M}^t para o domínio do índice de interação, no qual a solução é representada por $\boldsymbol{\eta}$. O objetivo aqui é verificar se há melhora nas condições de unicidade, trabalhando no domínio de $\boldsymbol{\eta}$. A transformação é dada por:

$$\tilde{\mathbf{M}} = \mathbf{M}^t \mathbf{T}, \quad (5.6)$$

onde \mathbf{T} é a matriz de transformação, dada por (4.2).

Então, para um conjunto de matrizes $\tilde{\mathbf{M}}$, geradas por simulação, verificamos a unicidade da solução do sistema $\tilde{\mathbf{M}}\boldsymbol{\eta} = \mathbf{u}$.

4. Por último, no quarto cenário, foram inseridas as restrições de igualdade $\mu(\emptyset) = 0$ e $\mu(C) = 1$, mas agora no domínio de $\boldsymbol{\eta}$. O objetivo foi construir uma matriz $\tilde{\mathbf{N}}$, que possui mais informação do que no caso sem as restrições. Não foram consideradas aqui as restrições que estão relacionadas à propriedade de monotonicidade da medida *fuzzy*. A matriz $\tilde{\mathbf{N}}$ é dada por:

$$\tilde{\mathbf{N}} = \mathbf{N} \mathbf{T}. \quad (5.7)$$

Calculou-se, então, o *posto* e o *spark* de um conjunto de matrizes $\tilde{\mathbf{N}}$, geradas por simulação, para verificar a unicidade da solução do sistema $\tilde{\mathbf{N}}\boldsymbol{\eta} = \mathbf{u}$.

5.2.1 Resultados numéricos para sistemas sobredeterminados

Para a investigação de sistemas sobredeterminados (baseados no *posto*) foram analisados dois casos: no primeiro, nós fixamos o número de critérios em 3 e variamos o número de alternativas de 8 a 12. No segundo, nós fixamos o número de critérios em 4 e consideramos o seguinte número de alternativas: 16, 17, 18, 19, 20, 25 e 40. Para cada caso, foi gerada uma matriz de decisão \mathbf{P} com distribuição uniforme de probabilidade, no intervalo $[0, 1]$. Também foi gerado linearmente um vetor de utilidades \mathbf{u} , com pesos dados por $1/m$, em que m é o número de critérios. Tendo \mathbf{P} , encontrou-se a matriz \mathbf{M}^t , utilizando (5.2). Foram realizadas 1000 simulações para cada caso.

Uma análise detalhada das 1000 simulações, para um sistema com 3 critérios (portanto, $2^m = 8$) e 8 alternativas, mostrou que elas chegaram a apenas 3 situações diferentes para os valores do *posto* como mostrado na Tabela 5.1. Como pode ser observado nessa tabela e, de acordo com nossas discussões prévias, somente na situação 3 e para o cenário 4, o critério de unicidade foi alcançado. Nas situações 1 e 2 mostradas na Tabela 5.1, vemos que o critério do *posto* não foi satisfeito em nenhum dos cenários, pois nenhuma das

matrizes analisadas apresentaram *posto* completo. Já na situação 3 mostrada na Tabela 5.1, vemos que somente quando as restrições de igualdade são adicionadas, o critério do *posto* é satisfeito. Isto acontece porque passa-se a ter um sistema com mais informações, levando as matrizes em questão a terem *posto* completo, o que garante a unicidade da solução.

Tabela 5.1 – Valores do *posto* para os quatro cenários para o caso com 3 critérios e 8 alternativas.

	Cenário 1		Cenário 2		Cenário 3		Cenário 4	
	\mathbf{M}^t	$\mathbf{M}^t \mathbf{u}$	\mathbf{N}	$\mathbf{N} \mathbf{u}$	$\tilde{\mathbf{M}}$	$\tilde{\mathbf{M}} \mathbf{u}$	$\tilde{\mathbf{N}}$	$\tilde{\mathbf{N}} \mathbf{u}$
Situação 1	5	5	6	6	5	5	6	6
Situação 2	6	6	7	7	6	6	7	7
Situação 3	7	7	8	8	7	7	8	8

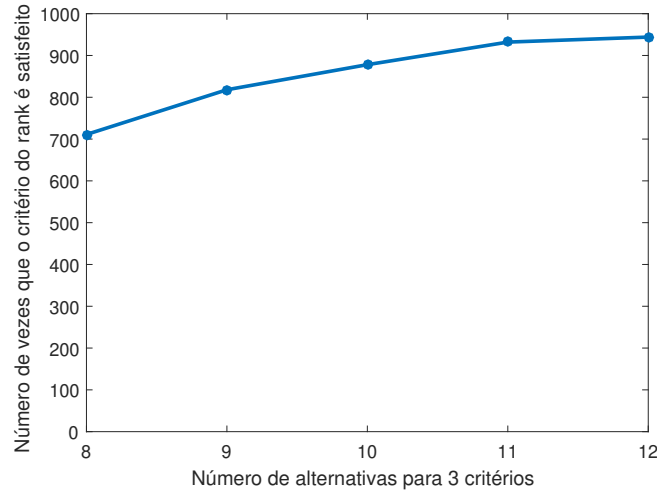
Numa segunda análise, nós investigamos o número de simulações (dentre as 1000) para o qual o critério de unicidade foi satisfeito. Essa investigação foca no cenário 4, que é o que apresenta maior probabilidade de ter o critério de unicidade satisfeito. A Figura 5.1 mostra o número de vezes em que o critério do *posto* foi satisfeito, para cada conjunto de critério e alternativa. Em ambos os casos verificamos que, quando o número de alternativas aumenta, o número de vezes que o critério do *posto* é satisfeito também aumenta. Isto era esperado, pois quando o número de alternativas aumenta, o número de equações do sistema resultante também aumenta, enquanto o número de parâmetros a ser estimado permanece o mesmo.

5.2.2 Resultados numéricos para sistemas subdeterminados

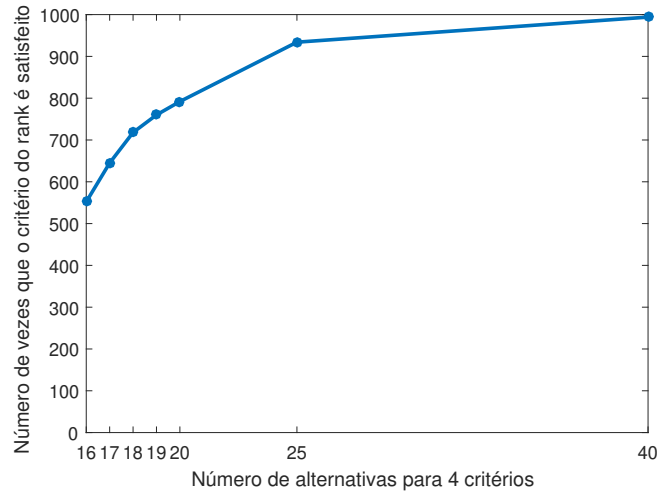
Para as análises do caso subdeterminado, que são baseadas no *spark*, também foram analisados dois casos: no primeiro nós fixamos o número de critérios em 3 e variamos o número de alternativas de 5 a 7. No segundo caso, nós fixamos o número de critérios em 4 e variamos o número de alternativas de 5 a 15. Novamente, para cada caso, foi gerada uma matriz de decisão \mathbf{P} com distribuição uniforme de probabilidade, no intervalo $[0, 1]$. Também foi gerado linearmente um vetor de utilidades \mathbf{u} , com pesos dados por $1/m$, em que m é o número de critérios. Tendo \mathbf{P} , encontrou-se a matriz \mathbf{M}^t , utilizando (5.2). Foram realizadas 1000 simulações para cada caso.

Uma análise detalhada das 1000 simulações, para um sistema 3 critérios (portanto, $2^m = 8$) e 6 alternativas, mostrou que elas chegaram a 14 diferentes situações para os valores de *spark* como mostrado na Tabela 5.2. Como pode ser observado nessa tabela, somente na situação 14 e para o cenário 4, o critério de unicidade foi satisfeito.

Vale lembrar que a condição de unicidade para sistemas subdeterminados, depende do valor do *spark* da matriz do sistema linear, assim como do nível de esparsidade para μ ou η , dependendo do domínio que se está trabalhando. Com relação ao nível de esparsidade, como os dados gerados vêm de um processo de agregação linear, a medida *fuzzy*



(a)



(b)

Figura 5.1 – Número de vezes que o critério do *posto* é satisfeito (considerando um cenário com 1000 simulações), no cenário 4, para diferentes números de alternativas e para: (a) 3 critérios e (b) 4 critérios.

resultante é aditiva, mas não esparsa, pois os elementos de μ não devem necessariamente ser nulos. Mais precisamente, para um modelo de agregação linear, se observa $\|\mu\|_0 = 7$, desde que o único elemento nulo aqui é o μ_\emptyset . Por outro lado, no domínio do índice de interação, a solução η resultante tem necessariamente uma norma $\|\eta\|_0 = 4$ pois, para o caso de agregação linear, $\eta(G) = 0$ para todas as coalizões de mais de um elemento.

Com base na discussão prévia e, de acordo com a condição de unicidade dada pelo *spark*, as seguintes condições devem ser satisfeitas para garantir a unicidade nos casos testados:

- $\text{spark}(\mathbf{M}^t) > 14$;

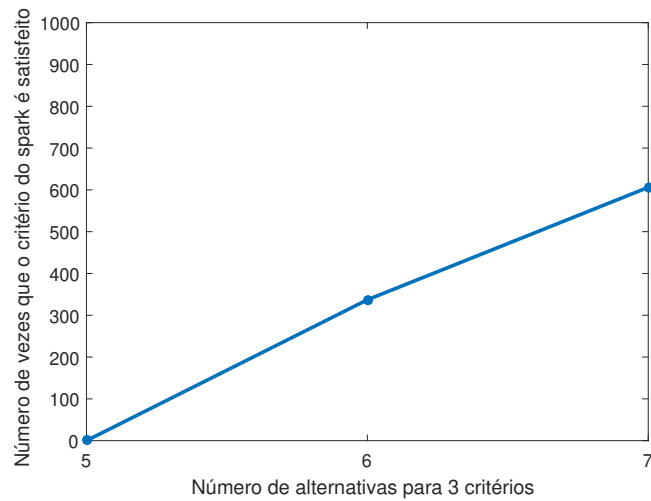
- $\text{spark}(\mathbf{N}) > 14$;
- $\text{spark}(\tilde{\mathbf{M}}) > 8$;
- $\text{spark}(\tilde{\mathbf{N}}) > 8$.

Tabela 5.2 – Valores do spark para os quatro cenários para o caso com 3 critérios e 6 alternativas.

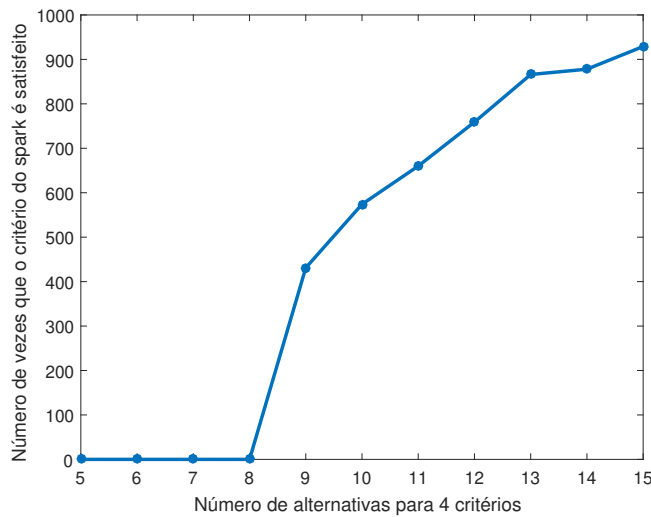
	Cenário 1	Cenário 2	Cenário 3	Cenário 4
	\mathbf{M}^t	\mathbf{N}	$\tilde{\mathbf{M}}$	$\tilde{\mathbf{N}}$
Situação 1	1	1	2	4
Situação 2	1	1	4	4
Situação 3	1	1	4	6
Situação 4	1	1	4	7
Situação 5	1	1	4	8
Situação 6	1	1	5	6
Situação 7	1	1	5	7
Situação 8	1	1	5	8
Situação 9	1	2	6	7
Situação 10	1	2	6	8
Situação 11	1	3	6	8
Situação 12	1	3	7	8
Situação 13	1	4	7	8
Situação 14	1	9	7	9

Nas situações 1 a 13 mostradas na Tabela 5.2, nota-se que o critério do spark não foi satisfeito em nenhum dos cenários e que, portanto, a unicidade não é verificada nos sistemas correspondentes. Na situação 14, temos que somente $\tilde{\mathbf{N}}$ satisfaz o critério de spark . Como observado na Tabela 5.2, os melhores resultados se apresentam quando esparsidade é considerada no domínio de $\boldsymbol{\eta}$, no qual efetivamente se tem um vetor com uma alta proporção de elementos nulos. Na situação 14, mesmo que o valor do spark de \mathbf{N} seja igual ao valor do spark de $\tilde{\mathbf{N}}$, não podemos dizer que o critério do spark foi satisfeito para \mathbf{N} , pois a identificação nesse caso é conduzida no domínio de $\boldsymbol{\mu}$ e, para que o critério do spark seja satisfeito nesse domínio, é necessário que se tenha $\text{spark}(\mathbf{N}) > 14$.

Numa última análise, nós investigamos o número de simulações (dentre as 1000) para o qual o critério de unicidade é satisfeito. Novamente, nossa investigação tem como foco o cenário 4. A Figura 5.2 mostra o número de vezes que o critério de spark foi satisfeito, para cada conjunto de critério e alternativa. Novamente, em ambos os casos, quando o número de alternativas aumenta, o número de vezes que o critério de spark é satisfeito também aumenta.



(a)



(b)

Figura 5.2 – Número de vezes que o critério do *spark* é satisfeito, no cenário 4, para diferentes números de alternativas e para: (a) 3 critérios e (b) 4 critérios.

5.3 Sumário

No presente capítulo, apresentamos a segunda contribuição dessa tese, que é a análise de unicidade do problema de identificação das capacidades de Choquet. Foram estudados dois casos de interesse. O primeiro caso é quando o sistema é sobredeterminado, para o qual se utiliza uma abordagem clássica para a análise, baseada no posto da matriz. O segundo caso é quando o sistema é subdeterminado, para o qual se utiliza uma abordagem mais recente para a análise, baseada no *spark* da matriz.

As análises foram realizadas para o domínio da medida *fuzzy* e para o domínio do índice de interação, além de que foram considerados casos nos quais foram inseridas as restrições de igualdade $\mu(\emptyset) = 0$ e $\mu(C) = 1$ ao problema. Essas restrições agem como um

termo de regularização, levando os problemas a terem maior probabilidade de alcançar solução única, quando comparado ao caso sem as restrições. Entretanto, mesmo chegando a condições para a unicidade, o problema poderia ser ainda melhor condicionado se fossem consideradas as restrições que estão relacionadas à propriedade de monotonicidade da medida *fuzzy*.

6 Conclusão

Neste trabalho, investigamos o problema da identificação dos parâmetros da integral de Choquet. Nossa proposta foi baseada em representações esparsas, com a busca por soluções no domínio do índice de interação.

Num primeiro momento, fizemos uma breve revisão sobre Métodos de Apoio à Decisão Multicritério, definindo o conceito e apresentando os cenários e problemáticas na qual esses métodos se aplicam. Em seguida abordamos o processo de agregação, definindo as propriedades do operador de agregação e apresentando a dificuldade que pode existir quando trabalhamos apenas com operadores de agregação linear. Nesse contexto, apresentou-se a integral de Choquet, como uma solução para essa dificuldade, pois ela é um operador de agregação não-linear.

Então, apresentamos o problema de identificação das capacidades de Choquet, comentando brevemente sobre as abordagens não-supervisionada, semi-supervisionada e supervisionada. Em nosso trabalho, escolhemos a abordagem supervisionada para resolver o problema e vimos que ela conduz a um problema de otimização associado a um problema inverso mal-posto. Nesse caso, faz-se necessário estabelecer um critério adicional que regularize o problema e, para esse trabalho, escolhemos fazer a regularização adicionando um termo que tende a promover soluções esparsas no vetor de capacidades de Choquet.

Nesse contexto, aparece a principal contribuição desse trabalho que é realizar a identificação das capacidades da integral de Choquet por meio de representações esparsas, mas trabalhando no domínio do índice de interação. A fim de comprovar a eficácia de nossa proposta, realizamos experimentos que consistiram em comparar a identificação realizada no domínio do índice de interação com a abordagem que realiza a identificação no domínio da medida *fuzzy*. Nossa proposta alcançou melhores resultados, pois a esparsidade não é uma propriedade esperada quando consideramos uma representação baseada numa medida *fuzzy*. Por outro lado, a esparsidade se torna significativa quando o problema é formulado no domínio do índice de interação e quando modelos *k-aditivos* são considerados.

Uma segunda contribuição desse trabalho foi a análise de unicidade realizada, na qual mostramos em quais situações é possível encontrar solução única para o problema. No caso sobredeterminado, foi mostrado que o critério do posto é satisfeito apenas quando as restrições de igualdade do problema são adicionadas. No caso subdeterminado, verificou-se que o critério do *spark* é satisfeito somente no domínio do índice de interação, no qual a esparsidade faz sentido no contexto dos modelos *k-aditivos*. Para ambos os casos, foi observado que quando o número de alternativas aumenta, o número de vezes em que o critério de unicidade é satisfeito também aumenta.

Para finalizar esse trabalho, gostaríamos de expor algumas perspectivas para trabalhos futuros:

- Realizar experimentos para o problema de identificação considerando dados gerados por modelos *k-aditivos* com $k > 2$ e variando o número de critérios e alternativas;
- Considerar a análise de unicidade para modelos *k-aditivos* com $k > 1$, nos quais os cálculos do *spark* se tornam desafiadores em termos computacionais;
- Estudar a unicidade, considerando outros conceitos comuns na área de *compressed sensing*, tais como a coerência de uma matriz e também a propriedade de isometria restrita (RIP, do inglês *restricted isometry property*) (DUARTE; ELDAR, 2011);
- Verificar o desempenho da solução proposta em cenários práticos.

Referências

- ADEYEBA, T. A.; ANDERSON, D. T.; HAVENS, T. C. Insights and characterization of l_1 -norm based sparsity learning of a lexicographically encoded capacity vector for the choquet integral. In: IEEE. *Fuzzy Systems (FUZZ-IEEE), 2015 IEEE International Conference on*. [S.l.], 2015. p. 1–7. Citado 4 vezes nas páginas 16, 36, 37 e 44.
- ANDERSON, D. T.; PRICE, S. R.; HAVENS, T. C. Regularization-based learning of the choquet integral. In: IEEE. *Fuzzy Systems (FUZZ-IEEE), 2014 IEEE International Conference on*. [S.l.], 2014. p. 2519–2526. Citado 2 vezes nas páginas 16 e 36.
- ASTER, R. C.; BORCHERS, B.; THURBER, C. H. *Parameter estimation and inverse problems*. [S.l.]: Academic Press, 2012. Citado na página 16.
- BARANIUK, R. G. Compressive sensing [lecture notes]. *IEEE signal processing magazine*, IEEE, v. 24, n. 4, p. 118–121, 2007. Citado 2 vezes nas páginas 16 e 66.
- BELIAKOV, G. et al. Choquet integral-based aggregation of interface usability parameters: Identification of fuzzy measure. In: UNIVERSITY OF BALLARAT. *ICOTA 2004: Proceedings of the 6th International Conference on Optimization Techniques and Applications*. [S.l.], 2004. Citado na página 25.
- BELIAKOV, G.; JAMES, S.; WU, J.-Z. *Discrete Fuzzy Measures*. [S.l.]: Springer, 2019. Citado na página 16.
- BOUYSSOU, D. et al. *Evaluation and decision models with multiple criteria: Stepping stones for the analyst*. [S.l.]: Springer Science & Business Media, 2006. v. 86. Citado 2 vezes nas páginas 15 e 20.
- BRUCKSTEIN, A. M.; DONOHO, D. L.; ELAD, M. From sparse solutions of systems of equations to sparse modeling of signals and images. *SIAM review*, SIAM, v. 51, n. 1, p. 34–81, 2009. Citado 2 vezes nas páginas 40 e 41.
- CANDES, E. J.; WAKIN, M. B.; BOYD, S. P. Enhancing sparsity by reweighted l_1 minimization. *Journal of Fourier analysis and applications*, Springer, v. 14, n. 5-6, p. 877–905, 2008. ISSN 1069-5869. Citado 3 vezes nas páginas 40, 41 e 43.
- CHOQUET, G. Theory of capacities. In: *Annales de l'institut Fourier*. [S.l.: s.n.], 1954. v. 5, p. 131–295. Citado na página 26.
- DEGROOT, M. H.; SCHERVISH, M. J. *Probability and statistics*. [S.l.]: Pearson Education, 2012. Citado na página 54.
- DEHE, B.; BAMFORD, D. Development, test and comparison of two multiple criteria decision analysis (mcda) models: A case of healthcare infrastructure location. *Expert Systems with Applications*, Elsevier, v. 42, n. 19, p. 6717–6727, 2015. Citado na página 15.
- DOMINGOS, P. The role of occam's razor in knowledge discovery. *Data mining and knowledge discovery*, Springer, v. 3, n. 4, p. 409–425, 1999. Citado na página 43.

- DONOHU, D. L.; ELAD, M. Optimally sparse representation in general (nonorthogonal) dictionaries via ℓ_1 minimization. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 100, n. 5, p. 2197–2202, 2003. Citado na página 66.
- DOUMPOS, M.; ZOPOUNIDIS, C. Preference disaggregation and statistical learning for multicriteria decision support: A review. *European Journal of Operational Research*, Elsevier, v. 209, n. 3, p. 203–214, 2011. Citado na página 36.
- DUARTE, L. T. A novel multicriteria decision aiding method based on unsupervised aggregation via the choquet integral. *IEEE Transactions on Engineering Management*, IEEE, v. 65, n. 2, p. 293–302, 2018. Citado 4 vezes nas páginas 35, 36, 57 e 58.
- DUARTE, L. T. et al. Blind extraction of sparse components based on ℓ_0 -norm minimization. In: IEEE. *Statistical Signal Processing Workshop (SSP), 2011 IEEE*. [S.l.], 2011. p. 617–620. Citado 2 vezes nas páginas 41 e 42.
- DUARTE, M. F.; ELDAR, Y. C. Structured compressed sensing: From theory to applications. *IEEE Transactions on Signal Processing*, IEEE, v. 59, n. 9, p. 4053–4085, 2011. Citado 2 vezes nas páginas 66 e 75.
- ELAD, M. *Sparse and Redundant Representations - From Theory to Applications in Signal and Image Processing*. [S.l.: s.n.], 2010. ISBN 978-1-4419-7010-7. Citado 3 vezes nas páginas 41, 42 e 66.
- ELAD, M.; FIGUEIREDO, M. A.; MA, Y. On the role of sparse and redundant representations in image processing. *Proceedings of the IEEE*, IEEE, v. 98, n. 6, p. 972–982, 2010. Citado na página 17.
- ELDAR, Y. C.; KUTYNIOK, G. *Compressed sensing: theory and applications*. [S.l.]: Cambridge university press, 2012. Citado 3 vezes nas páginas 9, 43 e 44.
- ELDAR, Y. C.; KUTYNIOK, G. *Compressed sensing: theory and applications*. [S.l.]: Cambridge university press, 2012. Citado na página 17.
- FILEV, D.; YAGER, R. R. On the issue of obtaining owa operator weights. *Fuzzy sets and systems*, Elsevier, v. 94, n. 2, p. 157–169, 1998. Citado na página 24.
- GRABISCH, M. The application of fuzzy integrals in multicriteria decision making. *European journal of operational research*, Elsevier, v. 89, n. 3, p. 445–456, 1996. Citado 8 vezes nas páginas 15, 23, 24, 25, 26, 28, 29 e 36.
- GRABISCH, M. Alternative representations of discrete fuzzy measures for decision making. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, World Scientific, v. 5, n. 05, p. 587–607, 1997. Citado 3 vezes nas páginas 15, 47 e 49.
- GRABISCH, M. A graphical interpretation of the choquet integral. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 8, n. 5, p. 627–631, 2000. Citado 6 vezes nas páginas 16, 28, 47, 48, 49 e 50.
- GRABISCH, M.; KOJADINOVIC, I.; MEYER, P. A review of methods for capacity identification in choquet integral based multi-attribute utility theory: Applications of the kappalab r package. *European journal of operational research*, Elsevier, v. 186, n. 2, p. 766–785, 2008. Citado 3 vezes nas páginas 16, 23 e 36.

GRABISCH, M.; LABREUCHE, C. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *Annals of Operations Research*, Springer, v. 175, n. 1, p. 247–286, 2010. Citado 4 vezes nas páginas 26, 36, 57 e 58.

GRABISCH, M.; LABREUCHE, C. Interpretation of multicriteria decision making models with interacting criteria. In: *New Perspectives in Multiple Criteria Decision Making*. [S.l.]: Springer, 2019. p. 151–176. Citado 2 vezes nas páginas 15 e 17.

GRABISCH, M.; NGUYEN, H. T.; WALKER, E. A. *Fundamentals of uncertainty calculi with applications to fuzzy inference*. [S.l.]: Springer Science & Business Media, 2013. v. 30. Citado na página 36.

GRABISCH, M.; ROUBENS, M. Application of the choquet integral in multicriteria decision making. *Fuzzy Measures and Integrals-Theory and Applications*, Berlin: Springer, p. 348–374, 2000. Citado na página 49.

GRECO, S.; FIGUEIRA, J.; EHRGOTT, M. Multiple criteria decision analysis. *Springer's International series*, Springer, 2005. Citado 4 vezes nas páginas 15, 19, 21 e 22.

KOJADINOVIC, I. Estimation of the weights of interacting criteria from the set of profiles by means of information-theoretic functionals. *European Journal of Operational Research*, Elsevier, v. 155, n. 3, p. 741–751, 2004. Citado na página 34.

MARICHAL, J.-L. An axiomatic approach of the discrete choquet integral as a tool to aggregate interacting criteria. *IEEE transactions on fuzzy systems*, IEEE, v. 8, n. 6, p. 800–807, 2000. Citado na página 26.

MARICHAL, J.-L.; ROUBENS, M. Determination of weights of interacting criteria from a reference set. *European journal of operational Research*, Elsevier, v. 124, n. 3, p. 641–650, 2000. Citado na página 36.

MOUSSEAU, V.; PIRLOT, M. Preference elicitation and learning. *EURO Journal on Decision Processes*, v. 3, n. 1, p. 1–3, jun. 2015. ISSN 2193-9446. Disponível em: <<https://doi.org/10.1007/s40070-015-0044-2>>. Citado na página 21.

MUROFUSHI, T. A technique for reading fuzzy measures (i): the shapley value with respect to a fuzzy measure. In: *2nd fuzzy workshop*. [S.l.: s.n.], 1992. v. 1, n. 1, p. 39–48. Citado 2 vezes nas páginas 26 e 49.

MUROFUSHI, T.; SONEDA, S. Techniques for reading fuzzy measures (iii): interaction index. In: SAPPORO,, JAPAN. *9th fuzzy system symposium*. [S.l.], 1993. p. 693–696. Citado 2 vezes nas páginas 26 e 49.

OLIVEIRA, H. E. de. *Um método para separação cega de fontes esparsas utilizando evolução diferencial*. Dissertação (Mestrado) — FCA/UNICAMP, 2015. Citado 2 vezes nas páginas 41 e 42.

OSBORNE, M. R.; PRESNELL, B.; TURLACH, B. A. On the lasso and its dual. *Journal of Computational and Graphical statistics*, Taylor & Francis, v. 9, n. 2, p. 319–337, 2000. Citado 2 vezes nas páginas 45 e 52.

PLUMBLEY, M. D. et al. Sparse representations in audio and music: from coding to source separation. *Proceedings of the IEEE*, IEEE, v. 98, n. 6, p. 995–1005, 2010. Citado na página 16.

- ROTA, G.-C. On the foundations of combinatorial theory i. theory of möbius functions. *Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete*, Springer-Verlag, v. 2, n. 4, p. 340–368, 1964. Citado na página 47.
- ROWLEY, H. V.; GESCHKE, A.; LENZEN, M. A practical approach for estimating weights of interacting criteria from profile sets. *Fuzzy Sets and Systems*, Elsevier, v. 272, p. 70–88, 2015. Citado na página 34.
- ROY, B. Paradigms and challenges. In: *Multiple criteria decision analysis: state of the art surveys*. [S.l.]: Springer, 2005. p. 3–24. Citado na página 22.
- SARKAR, T.; WEINER, D.; JAIN, V. Some mathematical considerations in dealing with the inverse problem. *IEEE transactions on antennas and propagation*, IEEE, v. 29, n. 2, p. 373–379, 1981. ISSN 0018-926X. Citado 3 vezes nas páginas 38, 39 e 40.
- SHAFAREVICH, I. R.; REMIZOV, A. O. *Linear algebra and geometry*. [S.l.]: Springer Science & Business Media, 2012. Citado na página 65.
- SHAPLEY, L. S. A value for n-person games. *Contributions to the Theory of Games*, v. 2, n. 28, p. 307–317, 1953. Citado na página 49.
- SINGIRESU, S. R. et al. *Mechanical vibrations*. [S.l.]: Addison Wesley, 1995. Citado na página 19.
- STARCK, J.-L.; MURTAGH, F.; FADILI, J. M. *Sparse image and signal processing: wavelets, curvelets, morphological diversity*. [S.l.]: Cambridge university press, 2010. Citado na página 16.
- TEHRANI, A. F. et al. Learning monotone nonlinear models using the choquet integral. *Machine Learning*, Springer, v. 89, n. 1-2, p. 183–211, 2012. Citado na página 16.
- TEHRANI, A. F. et al. Learning monotone nonlinear models using the choquet integral. *Machine Learning*, Springer, v. 89, n. 1-2, p. 183–211, 2012. Citado na página 36.
- TIBSHIRANI, R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society. Series B (Methodological)*, JSTOR, p. 267–288, 1996. Citado 2 vezes nas páginas 45 e 52.
- TIBSHIRANI, R. et al. Sparsity and smoothness via the fused lasso. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Wiley Online Library, v. 67, n. 1, p. 91–108, 2005. Citado na página 44.
- TIKHONOV, A. N. et al. *Numerical methods for the solution of ill-posed problems*. [S.l.]: Springer Science & Business Media, 2013. v. 328. Citado 2 vezes nas páginas 40 e 41.
- WU, J. et al. 2-additive capacity identification methods from multicriteria correlation preference information. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 23, n. 6, p. 2094–2106, 2015. Citado na página 35.
- WU, J. et al. Compromise principle based methods of identifying capacities in the framework of multicriteria decision analysis. *Fuzzy Sets and Systems*, Elsevier, v. 246, p. 91–106, 2014. Citado na página 35.

XU, B.; OUENNICHE, J. Performance evaluation of competing forecasting models: A multidimensional framework based on mcdm. *Expert Systems with Applications*, Elsevier, v. 39, n. 9, p. 8312–8324, 2012. Citado na página 15.

YAGER, R. R. On ordered weighted averaging aggregation operators in multicriteria decisionmaking. *IEEE Transactions on systems, Man, and Cybernetics*, IEEE, v. 18, n. 1, p. 183–190, 1988. Citado na página 24.

YAGER, R. R. Norms induced from owa operators. *IEEE Transactions on Fuzzy Systems*, IEEE, v. 18, n. 1, p. 57–66, 2009. Citado na página 24.