

FACULDADE DE ENGENHARIA DE CAMPINAS

UNICAMP

JUNHO 1977

ESTUDO COMPARATIVO DE MÉTODOS DE FILTRAGEM
PARA SISTEMAS ESTOCÁSTICOS DISCRETOS
NÃO LINEARES

JOSE CECHIN

Orientador - LUIS GIMENO LATRE

Tese apresentada à Faculdade de
Engenharia de Campinas -UNICAMP,
como parte dos requisitos neces
sários para obtenção do título
de MESTRE EM CIÊNCIAS.

Este trabalho não se teria
concentrado sem o apoio financeiro da
Fundação de Amparo à Pesquisa do Esta
do de São Paulo, e sem o estímulo inte
lectual de Luis Gimeno Latre, Manuel de
Jesus Mendes e Márcio Luis de Andrade
Netto.

ÍNDICE

CAPÍTULO I - INTRODUÇÃO

CAPÍTULO II - RECONSTRUÇÃO NUMÉRICA DAS DENSIDADES

- II.1 - Sistema e Propriedades Estatísticas
- II.2 - Densidade a posteriori via regra de Bayes
- II.3 - Equações do filtro de Kalman
- II.4 - Solução numérica ótima
- II.5 - Observações
- II.6 - Exemplo numérico

CAPÍTULO III- FILTROS SUBÓPTIMOS

- III.1 - Filtros generalizados
 - III.1.1 - Filtro de Kalman extendido
 - III.1.2 - Filtro de segunda ordem
- III.2 - Filtros fitados
 - III.2.1 - Filtro polinomial fitado
 - III.2.2 - Determinação dos coeficientes a_i e b_i
 - III.2.3 - Determinação do tamanho do intervalo D
 - III.2.4 - Filtro polinomial simplificado
- III.3 - Filtros iterados
 - III.3.1 - Filtro iterado com alissagem
 - III.3.2 - Filtros com iterações locais
 - III.3.3 - Filtro extendido modificado
- III.4 - Filtros gaussianos por aproximação direta da densidade a posteriori
 - III.4.1 - Filtro gaussiano por aproximação da densidade a posteriori
 - III.4.2 - Filtro gaussiano por aproximação da função de verossimilhança
 - III.4.3 - Filtro gaussiano por aproximação a priori da função de verossimilhança

CAPÍTULO IV - RESULTADOS NUMÉRICOS

CONCLUSÕES

CAPÍTULO I

INTRODUÇÃO

Processos físicos sujeitos a perturbações aleatórias cujo estado possa ser representado por um vetor de dimensão finita, podem ser modelados por equações vetoriais a diferenças.

$$x_{k+1} = f(x_k, w_k, k) \quad 1.1$$

x_k = vetor n-dimensional representando o estado.

w_k = vetor r-dimensional representando o ruído aleatório do sistema ($r \leq n$).

f = vetor de funções reais n-dimensional, continuamente diferenciável de seus argumentos.

$k = 0, 1, 2, \dots$ representa instantes de tempo.

A equação 1.1 é uma equação a diferenças, estocástica vetorial e representa um sistema dinâmico discreto estocástico. A sequência $\{x_1, x_2, \dots\}$ é uma sequência aleatória pelo fato de $\{w_0, w_1, \dots\}$ ser uma sequência aleatória. A condição inicial é x_0 e pode ser uma constante ou uma variável aleatória com distribuição de probabilidade conhecida ou assumida.

Se o ruído w_k estivesse ausente da equação 1.1, ela seria uma equação a diferenças, ordinária, cuja solução é x_n . Porém, na presença de ruído, é a densidade de probabilidade do estado x_k que é de interesse e para que possa ser determinada é necessário supor uma determinada densidade do ruído.

Se, por exemplo, $\{w_0, w_1, \dots\}$ for uma sequência branca e independente do estado inicial x_0 , então, dado o valor do estado no instante k , o valor do estado no instante de tempo seguinte dependerá somente de w_k que é independente de $\{x_{k-1}, x_{k-2}, \dots, x_0\}$, vale dizer, a solução de 1.1 será uma sequência de Markov. A densidade de probabilidade de x_{k+1} dado x_k , supondo $r = n$ e que o sistema 1.1 possa ser resolvido para w_k , será dada, de acordo com o Teorema Fundamental do cálculo de probabilidade, por:

$$p(x_{k+1}/x_k) = p\left(w_k = f^{-1}(x_k, x_{k+1}, k)\right) \left\| \frac{\partial f^{-1}(x_k, x_{k+1}, k)}{\partial x_{k+1}} \right\|$$

Esta densidade é conhecida como densidade de transição de estado.

O problema que ora se coloca é o da determinação, instante por instante, da densidade do estado do sistema 1.1 quando ele é imperfeitamente observado por medidas ruidosas. Este problema é interessante não apenas por seus méritos próprios, mas é também parte inerente da otimização de sistemas dinâmicos estocásticos.

As observações do estado obedecem à seguinte equação:

$$y_k = h(x_k, v_k, k) \quad 1.2$$

y_k = vetor de medidas m-dimensional

v_k = vetor de ruído aleatório m-dimensional

h = vetor de funções m-dimensional

Dada uma sequência de observações $\{y_0, y_1, \dots, y_\ell\}$ que será denotada brevemente por J_ℓ

$$J_\ell = \{y_0, y_1, \dots, y_\ell\}$$

o problema da estimação consiste em calcular a densidade do estado x_k condicionado à sequência de medidas J_ℓ . O presente trabalho está limitado ao problema conhecido como filtragem que é a estimação do estado no instante k condicionada à sequência de observações J_k .

O problema da filtragem recebeu basicamente duas abordagens metodologicamente distintas. A primeira delas, da qual não se ocupa o presente trabalho, consiste dos métodos estatísticos que têm como principal atrativo o fato de não necessitarem hipóteses explícitas sobre as diversas propriedades estatísticas das variáveis aleatórias envolvidas. As sequências $\{w_0, w_1, \dots\}$ e $\{v_0, v_1, \dots\}$ são consideradas simplesmente como erros de caráter desconhecido e as equações de estado e medida já não são consideradas estocásticas. Por exemplo, o clássico LEAST-SQUARES partindo de um conjun

to de observações imperfeitas $\{y_0, y_1, \dots, y_n\}$, determina uma estimativa de $\{x_0, x_1, \dots, x_n\}$ tal que simultaneamente os erros no sistema e nas observações sejam minimizados. A inconveniência destes métodos reside na dificuldade em atribuir uma interpretação probabilística aos resultados obtidos.

A segunda abordagem é a probabilística. Neste caso, a solução completa do problema da filtragem é a densidade do estado no instante k condicionada ao conjunto de medidas J_k , simplesmente porque esta densidade contém toda a informação sobre o estado trazida pelas medidas e pelo estado inicial. Nenhuma descrição do estado mais completa é possível. A introdução de um critério de estimação reduz a informação a respeito do estado a uma coleção finita de números. Em geral, isto fornecerá uma descrição incompleta e possivelmente inadequada do estado, sendo a única exceção um sistema linear e gaussiano.

A densidade do estado condicionada às medidas pode ser calculada de forma recorrente a cada nova medida realizada, através da regra de Bayes. Esta densidade é chamada de densidade a posteriori de estado.

Para sistemas lineares com estado inicial e ruídos gaussianos, a densidade a posteriori do estado será sempre gaussiana. A recorrência pode, neste caso, ser feita diretamente sobre o valor médio e matriz de covariância da densidade a posteriori, pois estes dois parâmetros especificam completamente uma gaussiana. As equações de evolução da média e covariância constituem o filtro linear de Kalman (10,11). Mas para sistemas não lineares a densidade a posteriori não é em geral redutível a um conjunto finito de parâmetros, isto é, não possui uma estatística suficiente finita. Tão pouco é possível conseguir uma forma analítica fechada para descrevê-la. Faz-se necessário, então, recorrer às aproximações ou calcular numericamente as densidades.

O cálculo numérico das densidades permite alcançar a precisão numérica desejada, mas às custas de uma velocidade de processamento extremamente baixa, inviabilizando o emprego prático desta solução. Diversos são os trabalhos desenvolvidos nesta linha.

A reconstrução numérica "ótima" das densidades pela regra de Bayes empregando "splines" como interpolares foi trabalho desenvolvido por Márcio Andrade Netto et alli (2,3,14). No Capítulo II este algoritmo está sumarizado, pois dele se fará uso para explicar a performance dos filtros linearizados, bem como para obter os resultados "exatos" para fins de comparação.

Nesta categoria podem ser incluídos os trabalhos sugeridos e elaborados por Bucy and Senne (6), E.TSE (17) e Alspach and Sorenson (1,15). Trata-se basicamente de aproximar as densidades intermediárias da fórmula de Bayes por um certo número de funções impulso (TSE) ou por uma soma de gaussianas (Sorenson and Alspach) de variâncias muito pequenas. O teorema de Wiener (18) afirma ser possível aproximar qualquer função densidade por um número finito de termos das somas, a precisão da aproximação depende do número de termos. A densidade a posteriori será obtida operando simultaneamente com todos os termos que compõem as densidades intermediárias. Desta forma, a densidade a posteriori será também descrita por uma soma de funções impulso ou por uma soma de gaussianas. No entanto, pode ocorrer que o número de termos da soma de impulsos deva aumentar para manter a precisão à medida que o algoritmo funciona ou que a variância dos termos da soma de gaussianas aumente invalidando as hipóteses de linearização assumidas, e neste caso também deve ser aumentado o número de termos da soma de gaussianas.

Segundo apontado por TSE, este tipo de aproximação poderá se tornar extremamente valioso com o advento da computação paralela, pois se se puder dispor de um processador para cada termo da soma, a velocidade de execução deste algoritmo será da ordem de grandeza da velocidade dos algoritmos comumente utilizados na prática.

Aparte estas aproximações muito dispendiosas, estão as aproximações que receberam ampla aceitação prática por sua simplicidade e rapidez de cálculo. Todas elas mantêm a estrutura da solução proposta por Kalman para o caso linear (10), o que é possível mediante a hipótese simplificadora de que todas as densidades envolvidas na regra de Bayes são gaussianas. Para isto, as não linearidades são aproximadas por

desenvolvimento em série de Taylor em torno de uma trajetória de referência. Se o desenvolvimento incluir apenas os termos até primeira ordem, a solução resultante é conhecida como filtro de Kalman linearizado (4,12,13). Se incluir os termos de segunda ordem podem resultar diversas soluções, dependendo de como se aproximam os momentos de ordem superior a dois (5,9). Estas soluções são conhecidas como filtros de segunda ordem. Diferem dos anteriores por introduzirem uma correção de polarização nos valores médios e alargarem as variâncias das densidades.

As aproximações até primeira ordem podem incorrer em graves erros de linearização e calcular as derivadas de segunda ordem nem sempre é tarefa simples. Denhan and Pines (7) desenvolveram um procedimento iterativo para reduzir os erros de linearização sem necessitar do cálculo das derivadas de segunda ordem e obter deste modo uma densidade a posteriori aprimorada. A solução resultante é conhecida como filtro iterado de primeira ordem (16).

Pode ser constatado facilmente que as soluções recém descritas situam-se em dois extremos: de um lado estão as soluções numéricas que fornecem precisão às custas de um tempo de cálculo muito elevado inviabilizando sua aplicação prática, de outro as soluções aproximadas cuja simplicidade e rapidez de cálculo tornam-se atrativas para as aplicações práticas mas podem apresentar resultados pouco satisfatórios.

Tentar-se-á neste trabalho esclarecer as condições em que possivelmente as soluções aproximadas conduzem a resultados insatisfatórios. Genericamente pode ser dito que as linearizações por desenvolvimento em série de Taylor dificilmente podem representar o comportamento que não seja estritamente local das não linearidades. Para superar esta limitação propõe-se aproximar as não linearidades não apenas localmente mas considerando um certo intervalo de interesse. A aproximação será feita por funções polinomiais cujos coeficientes são determinados obedecendo à minimização de algum critério de custo. Mas a hipótese de densidades gaussianas não é abandonada e esta solução se enquadra, portanto, no grupo das soluções aproximadas.

Todavia existem situações em que qualquer solução aproximada conduz a densidades gaussianas que não representam adequadamente as densidades verdadeiras ou ignoram parcialmente a informação trazida por uma nova medida. A maneira de evitá-los é sofisticar a solução, isto é, introduzir um procedimento intermediário entre os dois extremos antes mencionados. A solução apresentada neste trabalho é operar diretamente com as densidades que compõem a fórmula de Bayes. A densidade a posteriori resultante é aproximada por uma gaussiana.

CAPÍTULO II

RECONSTRUÇÃO NUMÉRICA DAS DENSIDADES

O trabalho que se desenvolve ao longo deste Capítulo e nos próximos faz referência ao conjunto de hipóteses que aqui se expõem.

II.1 - SISTEMA E PROPRIEDADES ESTATÍSTICAS

As equações de evolução do sistema são:

$$x_{k+1} = f(x_k, k) + \Gamma_k w_k \quad 2.1$$

$$y_k = h(x_k, k) + v_k \quad 2.2$$

As propriedades estatísticas do estado inicial x_0 e dos ruídos de estado e de medida são:

$$E \{ x_0 \} = \hat{x}_0.$$

$$E \{ (x - \hat{x}_0) (x - \hat{x}_0)^T \} = X_0$$

$$E \{ w_k \} = E \{ v_k \} = 0$$

2.3

$$E \{ w_k \cdot w_1^T \} = W_{k,1} \delta_{k,1}$$

$$E \{ v_k \cdot v_1^T \} = V_{k,1} \delta_{k,1}$$

$$E \{ w_k \cdot v_1^T \} = E \{ x_0 \cdot w_k \} = E \{ x_0 \cdot v_k \} = 0$$

Em descrição verbal vale dizer que o estado inicial é um vetor de média \hat{x}_0 e variância X_0 . Os ruídos de estado w_k e de medida v_k são brancos, gaussianos de média zero, mutuamente independentes e independentes do estado inicial x_0 . (E é o operador esperança matemática, T indica transposição e $\delta_{k,1}$ é o delta de Kronecker).

A hipótese de perturbações estocásticas aditivas, acima assumida, é conveniente para simplificar a exposição pois, por exemplo, segue-se que

$$p(y_k/x_k) = p\left(y_k - h(x_k, k)\right) = p(v_k)$$

consequentemente, a densidade condicional da medida é igual à densidade conhecida ou assumida do ruído v_k .

II.2 - DENSIDADE A POSTERIORI VIA REGRA DE BAYES

O problema da estimação do estado no instante $k+1$ baseada na informação disponível a partir das observações até este instante - daqui para a frente referido simplesmente como problema de filtragem de estado - pode considerar-se resolvido, uma vez que se tenha obtido a densidade a posteriori do estado condicionada às observações já realizadas. Esta densidade é obtida recursivamente fazendo-se uso da regra de Bayes. Mediante um procedimento relativamente simples, obtém-se o resultado apresentado a seguir.

$$p(x_{k+1}/J_{k+1}) = \frac{p(x_{k+1}/J_k) \cdot p(y_{k+1}/x_{k+1})}{p(y_{k+1}/J_k)} \quad 2.4$$

Cada termo envolvido na expressão da densidade a posteriori pode ser desdobrado do seguinte modo:

- constante de normalização

$$p(y_{k+1}/J_k) = \int p(x_{k+1}/J_k) \cdot p(y_{k+1}/x_{k+1}) dx_{k+1} \quad 2.5$$

- densidade de previsão da observação

$$p(y_{k+1}/x_{k+1}) = p\left(v_{k+1} = y_{k+1} - h(x_{k+1}, k+1)\right) \quad 2.6$$

- densidade de previsão de estado

$$p(x_{k+1}/J_k) = \int p(x_{k+1}/x_k) \cdot p(x_k/J_k) dx_k \quad 2.7$$

- densidade de transição de estado

$$p(x_{k+1}/x_k) = p\left(w_k = \Gamma_k^{-1} \left(x_{k+1} - f(x_k, k) \right) \right) \|\Gamma_k^{-1}\| \quad 2.8$$

$p(x_k/J_k)$ é a densidade do estado no instante anterior. A inicialização da recorrência tem como condição inicial

$$p(x_0/J_{-1}) = p(x_0)$$

e portanto o processamento da primeira observação se resume a:

$$p(x_0/J_0) = \frac{p(x_0) \cdot p(y_0/x_0)}{p(y_0)} \quad 2.9$$

Se é verdade que a equação 2.4 é a solução completa do problema da filtragem no sentido de que ela contém toda a informação a respeito do estado fornecida pelas observações, também é verdade que nesta forma não pode ser diretamente empregada. É preciso converter esta densidade numa coleção finita de números, seus momentos. Em geral, não existe um conjunto finito de números capaz de descrever completamente a densidade, perdendo-se portanto, informação. Contudo, sob algumas hipóteses restritivas, o cálculo dos momentos é extremamente simplificado. Tal é o caso do filtro de Kalman que, sob a hipótese de linearidade das funções f e h e mantendo válidas as relações 2.3, apresenta uma estatística suficiente e a recorrência é feita diretamente sobre a média e covariância.

Com a finalidade de introduzir e clarificar a notação que será seguida neste estudo, apresenta-se a seguir as equações do filtro de Kalman.

II.3 - EQUAÇÕES DO FILTRO DE KALMAN

Esta seção apresenta o filtro de Kalman para as funções f e h lineares dadas por:

$$f(x_k, k) = F_k \cdot x_k \quad 2.10$$

$$h(x_k, k) = H_k \cdot x_k$$

É uma propriedade notável de uma variável aleatória gaussiana que as transformações de uma densidade a priori pela regra de Bayes para se obter a densidade a posteriori preserva sua forma gaussiana quando as funções f e h são lineares. Assim se o estado inicial for assumido com densidade de probabilidade gaussiana, todos os estados subsequentes terão uma distribuição gaussiana sob as relações 2.3. Neste caso, a média e a covariância constituem uma estatística suficiente e suas equações de recorrência são as equações do filtro de Kalman.

- previsão de estado

$$\hat{x}_{k+1/k} \stackrel{\Delta}{=} E\{x_{k+1}/J_k\}$$

$$X_{k+1/k} \stackrel{\Delta}{=} E\{(x_{k+1} - \hat{x}_{k+1/k})(x_{k+1} - \hat{x}_{k+1/k})^T / J_k\}$$

$\hat{x}_{k+1/k}$ = valor médio da densidade de previsão de estado

$x_{k+1} - \hat{x}_{k+1/k}$ = erro de previsão de estado

$X_{k+1/k}$ = covariância da densidade de previsão de estado

$$\hat{x}_{k+1/k} = F_k \cdot \hat{x}_{k/k}$$

2.11

$$X_{k+1/k} = F_k X_{k/k} F_k^T + \Gamma_k W_k \Gamma_k^T$$

- previsão da observação

$$\hat{y}_{k+1/k} \stackrel{\Delta}{=} E\{y_{k+1}/J_k\}$$

$$Y_{k+1/k} \stackrel{\Delta}{=} E\{(y_{k+1} - \hat{y}_{k+1/k})(y_{k+1} - \hat{y}_{k+1/k})^T / J_k\}$$

$\hat{y}_{k+1/k}$ = é o valor médio da densidade de previsão da medida

$y_{k+1} - \hat{y}_{k+1/k}$ = erro de previsão de saída

$Y_{k+1/k}$ = covariância da densidade de previsão da medida

$$\hat{y}_{k+1/k} = H_{k+1} \hat{x}_{k+1/k}$$

2.12

$$Y_{k+1/k} = H_{k+1} X_{k+1/k} H_{k+1}^T + V_{k+1}$$

- equações de filtragem

$$\hat{x}_{k+1/k} \stackrel{\Delta}{=} E \left\{ x_{k+1} / J_{k+1} \right\}$$

$$X_{k+1/k+1} \stackrel{\Delta}{=} E \left\{ (x_{k+1} - \hat{x}_{k+1/k+1}) (x_{k+1} - \hat{x}_{k+1/k+1})^T / J_{k+1} \right\}$$

$\hat{x}_{k+1/k+1}$ = valor médio da densidade a posteriori

$x_{k+1} - \hat{x}_{k+1/k+1}$ = erro de estimação de estado

$X_{k+1/k+1}$ = covariância da densidade a posteriori

$$\hat{x}_{k+1/k+1} = \hat{x}_{k+1/k} + K_{k+1} (y_{k+1} - \hat{y}_{k+1/k})$$

$$X_{k+1/k+1} = (I - K_{k+1} H_{k+1}) X_{k+1/k} \quad 2.13$$

$$K_{k+1} = X_{k+1/k} H_{k+1}^T Y_{k+1/k}^{-1}$$

Um procedimento frequentemente empregado para sistemas não lineares é a linearização por desenvolvimento em série de Taylor em torno de uma trajetória de referência. Aplicam-se as equações precedentes ao sistema linearizado, as equações assim obtidas constituem uma generalização do filtro de Kalman, conhecidas na literatura como EXTENDED, ITERATED EXTENDED, SECOND ORDER KALMAN FILTERS. Obviamente, esta é uma solução subótima, cujo grande mérito é preservar a estrutura simples do filtro de Kalman, mas que frequentemente conduz a resultados insatisfatórios, justamente porque a linearização por desenvolvimento em série de Taylor

é apenas uma aproximação local da não linearidade, não representativa, portanto, do seu caráter geral. Isto porque a equação de Bayes 2.7 envolve uma integração de $-\infty$ a $+\infty$ idealmente. O Capítulo III descreverá estes algoritmos e no Capítulo IV constarão alguns resultados destas aproximações.

II.4 - SOLUÇÃO NUMÉRICA ÓPTIMA

Dado que para o caso não linear não existe em geral uma solução analítica fechada, desenvolveu-se um algoritmo para aproximação numérica da relação 2.4. Em cada instante se reconstroem numericamente as densidades de previsão de estado e de previsão de saída e a constante de normalização. Consegue-se, com este método, a precisão numérica que se desejar, a custa de maior tempo de processamento.

A partir da densidade a posteriori, podem ser calculados os momentos fornecendo uma referência básica para comparação com os valores estimados fornecidos pelos filtros linearizados.

Na sequência desta exposição, procura-se mostrar, passo a passo, a trajetória de cálculo que leva da densidade a priori, isto é, da densidade a posteriori do instante anterior, para a densidade a posteriori atual, abrangendo a nova informação trazida por uma nova medida. O primeiro passo se refere ao cálculo da densidade de previsão de estado - equação 2.7. Esta expressão requer o conhecimento da densidade de transição de estado - equação 2.8 - que é a responsável pela introdução da não linearidade do sistema. A expressão 2.8 fornece a densidade do estado no instante $k+1$, dado o valor de x_k e será sempre dada pela própria densidade do ruído. Contudo, na expressão 2.7, x_k não é um valor fixado, mas é a própria variável, fazendo com que $p(x_{k+1}/x_k)$ não seja mais densidade, mas uma função $L(x_k:x_{k+1})$ que, por analogia, chamaremos de função de verossimilhança.

Esta função é analiticamente conhecida, valendo para cada $x_{j,k+1}$:

$$\begin{aligned}
 L(x_k : x_{j,k+1}) &= p(x_{j,k+1} / x_k) = p_w \left(w_k = \Gamma_k^{-1} \left(x_{j,k+1} - f(x_k, k) \right) \right) \left| \Gamma_k \right|^{-1} \\
 &= \text{CTE} \cdot \exp \left(-0.5 // \Gamma_k^{-1} \left(x_{j,k+1} - f(x_k, k) \right) // w_k^{-1} \right)^2
 \end{aligned}
 \quad 2.14$$

Portanto, a densidade de previsão de estado é obtida ponto a ponto por uma integração dada por 2.7 e 2.14. Este constitui verdadeiramente o ponto crucial deste tipo de algoritmo, pois cada valor numérico da densidade de previsão, $p(x_{j,k+1} / J_k)$, exige que se efetue uma integração, fazendo com que o tempo de processamento se torne demasiadamente longo.

Apresenta-se a seguir uma ilustração gráfica para o caso escalar, com a finalidade de tornar mais clara a compreensão da "convolução não linear". Pode-se começar pelo gráfico da função f que relaciona o estado x_{k+1} com o estado x_k : dado um certo valor de estado no instante k , $x_{m,k}$, o estado no instante seguinte, x_{k+1} , terá média $f(x_{m,k}, k)$ e uma distribuição dada pela distribuição do ruído do sistema, (Figura 2.1).

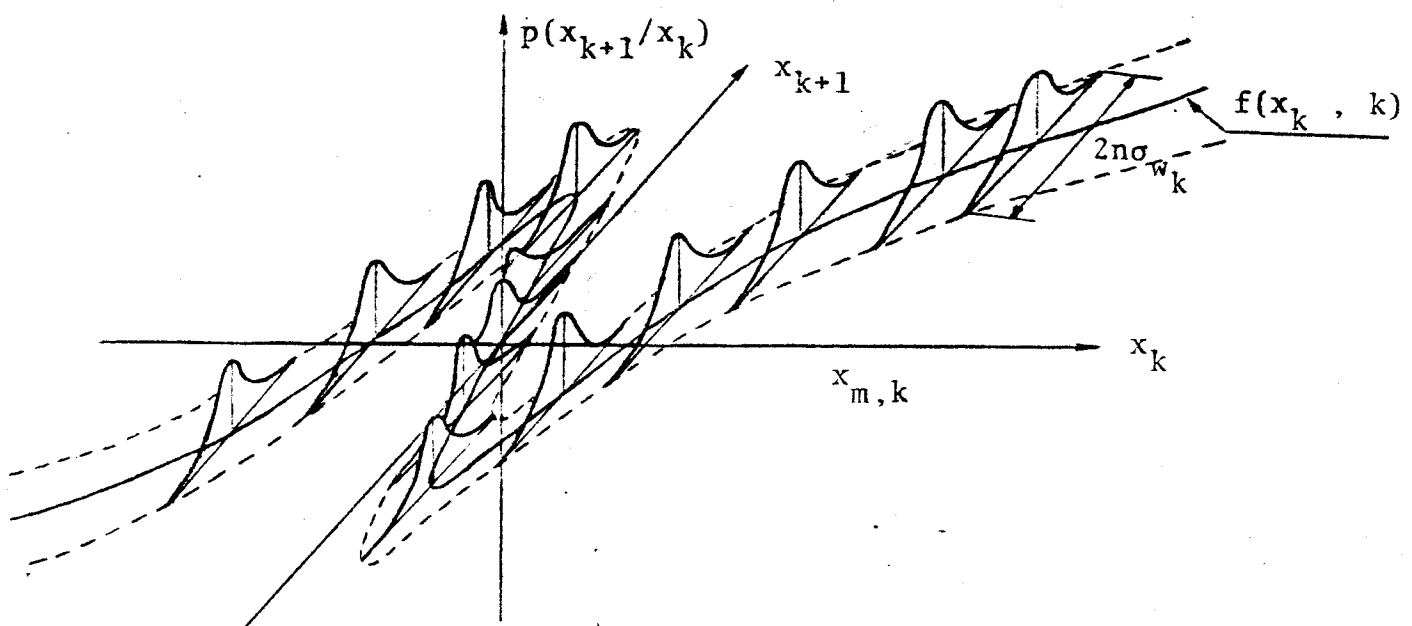


Fig. 2.1 - Densidades de transição de estado considerando apenas valores da densidade de ruído no intervalo $f(x_k, k) \pm \sigma_w$.

Contudo, especificado um certo valor de x_{k+1} , por exemplo $x_{j,k+1}$, $p(x_{j,k+1}/x_k)$ já não é mais nem densidade nem sequer terá a forma gaussiana, mas é a função de verossimilhança $L(x_k : x_{j,k+1})$, podendo ser, inclusive, multimodal.

Algumas situações típicas são apresentadas na Fig. 2.2.

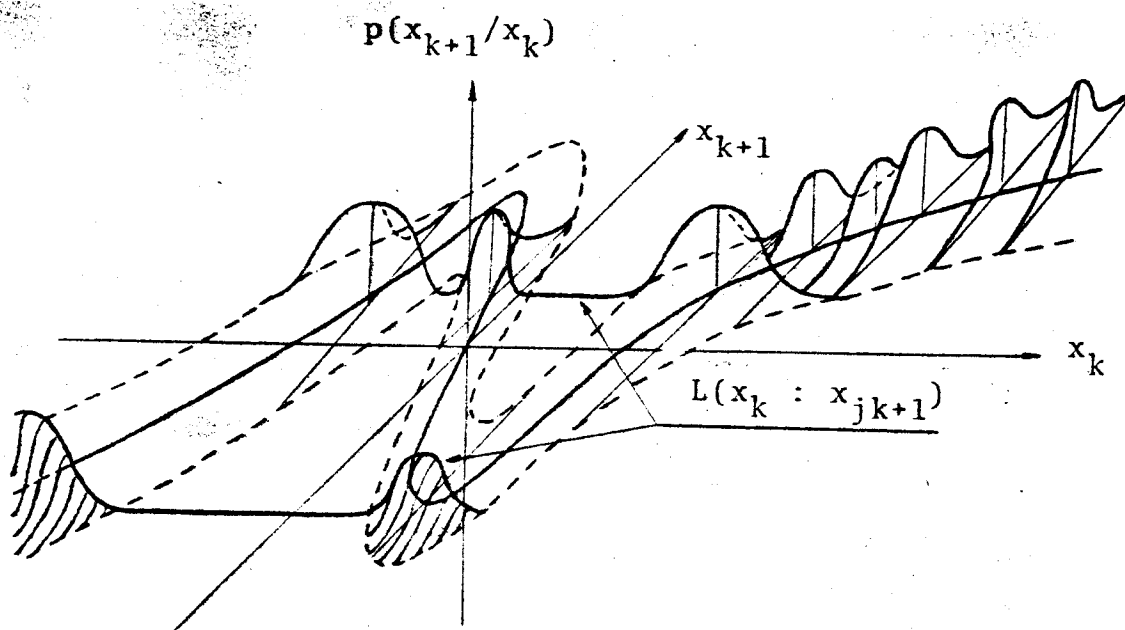
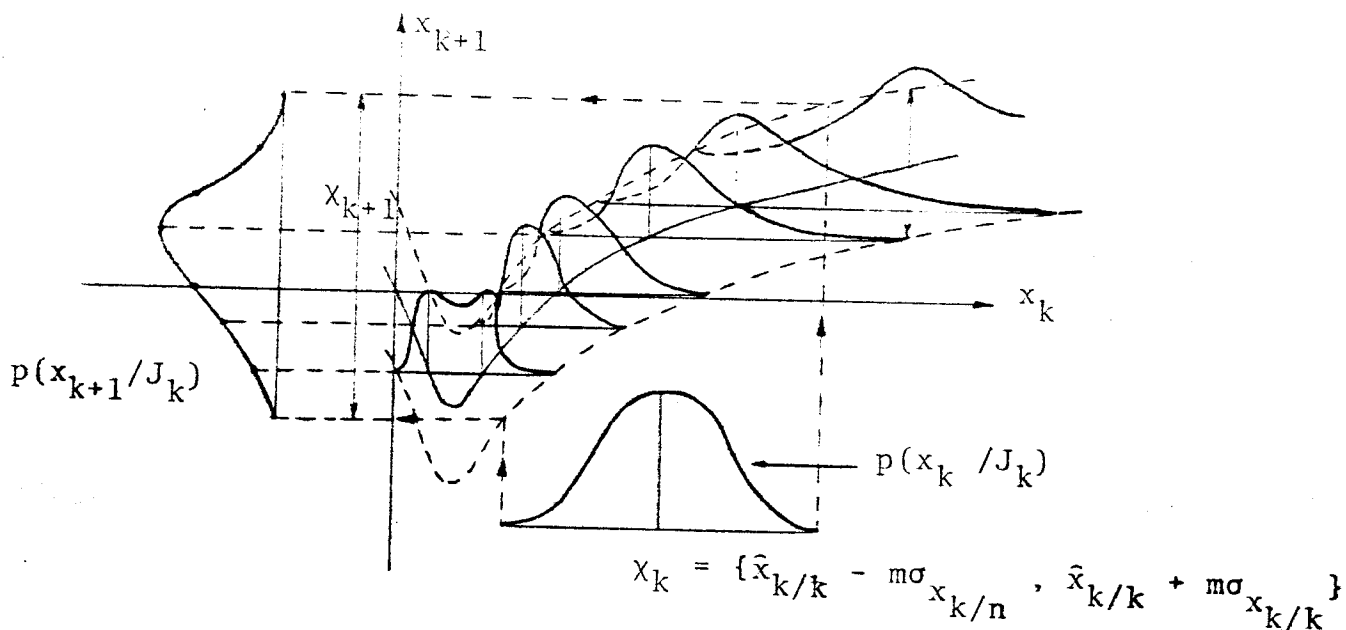


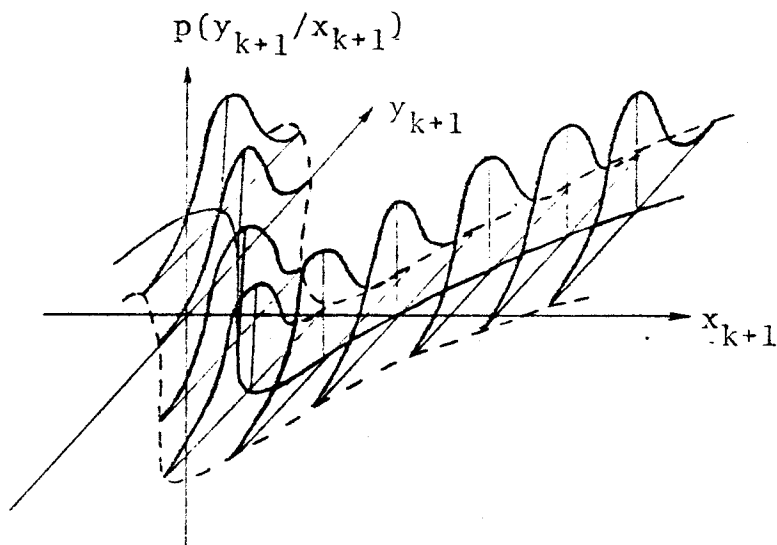
Fig. 2.2 - Funções de verossimilhança

Para se obter a densidade de previsão é necessário definir um domínio x'_{k+1} , no qual tomará valores não desprezíveis, subdividí-lo em M pontos $x_{j,k+1}$ e, para cada ponto, realizar a integração do produto da densidade a posteriori do instante anterior com a função de verossimilhança. Tal domínio se define de acordo com o domínio de valores significativos da densidade a priori x_k e considerando uma faixa $(f(x_k, k) - n\sigma_{w_k}, f(x_k, k) + n\sigma_{w_k})$ em torno da função f .

A Fig. 2.3, que se apresenta a seguir, ilustra a definição do domínio x'_{k+1} .



Obtida a densidade de previsão de estado, necessi-
ta-se agora determinar a densidade de previsão da medida
 $p(y_{k+1}/x_{k+1})$. De modo análogo ao descrito nas linhas prece-
dentes, $p(y_{k+1}/x_{k+1})$, para x_{k+1} especificado, será uma
densidade cuja média é $h(x_{1,k+1}, k+1)$ e cuja distribuição
é a do próprio ruído de medida. Graficamente, está represen-
tado na Fig. 2.4.



Entretanto, na expressão 2.4, $p(y_{k+1}/x_{k+1})$ apresenta x_{k+1} como variável e y_{k+1} como um valor especificado, ou seja, y_{k+1} assume o valor realmente observado: $y_{k+1} = y_{obs}$.

Sendo assim, $p(y_{k+1}/x_{k+1})$ já não é uma densidade, mas outra função de verossimilhança $L_1(x_{k+1} : y_{obs})$ - Fig. 2.5.

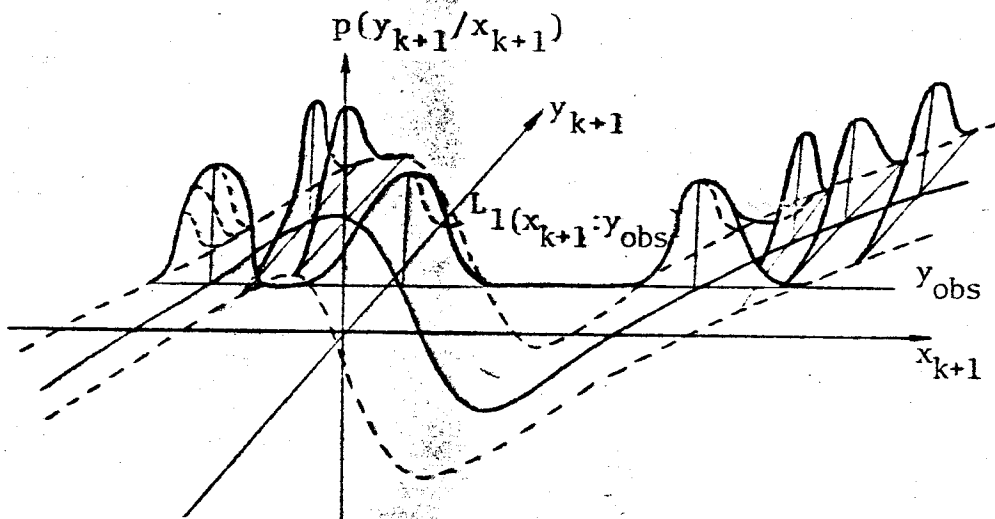


Fig. 2.5 - Função de verossimilhança para $y_{k+1} = y_{obs}$

Finalmente, a densidade a posteriori é obtida realizando o produto da densidade de previsão de estado pela função $L_1(x_{k+1} : y_{obs})$, como mostra a Fig. 2.6, e em seguida normalizando o resultado. O domínio x_{k+1} da densidade a posteriori é finalmente definido como o intervalo em que $p(x_{k+1}/y_{k+1})$ assume valores não desprezíveis.

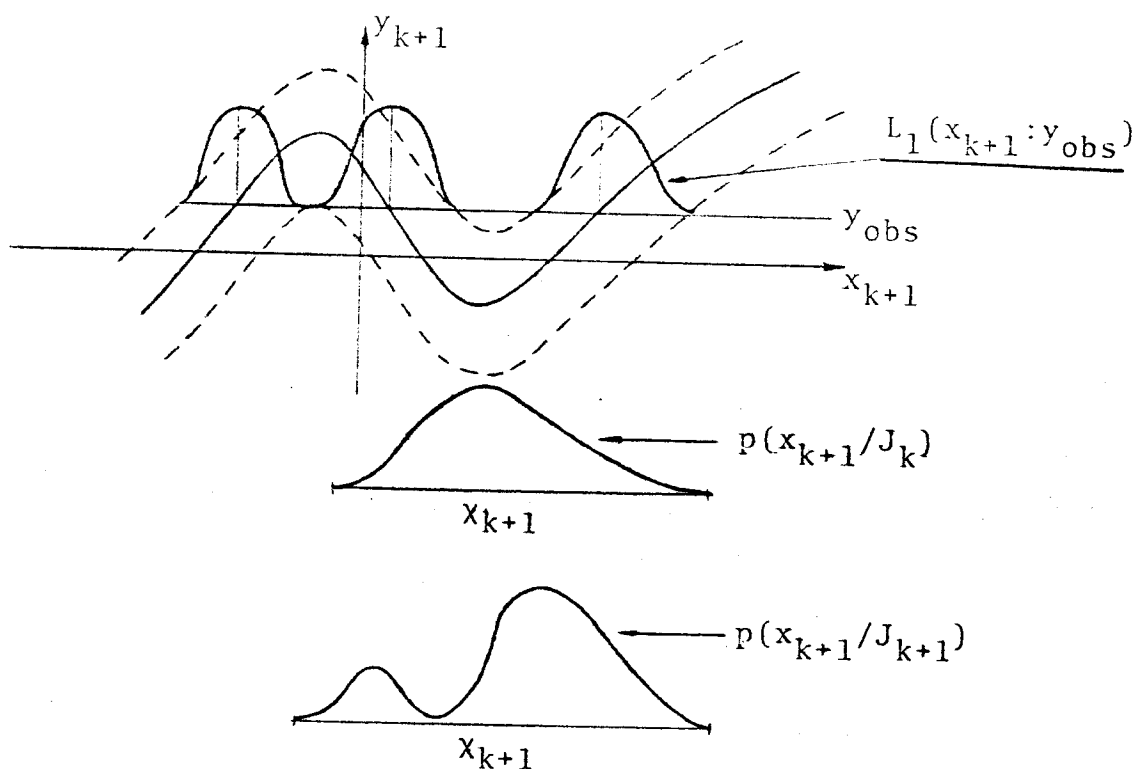


Fig. 2.6 - Obtenção da densidade a posteriori e definição do seu domínio x_{k+1}

II.5 - OBSERVAÇÕES

O cálculo da densidade de previsão de estado requer integração numérica. Como a função de verossimilhança e a densidade a priori são dadas numericamente, para que as densidades não se degradem na convolução faz-se necessário recorrer a interpolações. Em princípio, as possibilidades de escolha de um interpolador particular são muitas, mas o presente algoritmo recorreu aos "splines" por serem um ferramental analítico que combina conveniência para integração e um bom nível de redução de dados. As funções "splines" foram também utilizadas para o cálculo dos momentos (2).

O desenvolvimento da seção precedente não pode ser entendido como mero exercício acadêmico. Além de poder for

necer resultados numéricos com a precisão que se desejar, servindo como um referencial básico para a comparação de desempenho dos diversos algoritmos de filtragem subótimos, provê um poderoso método para a compreensão do comportamento dos filtros linearizados. O próximo Capítulo será dedicado a este estudo. Mas desde logo seja dito que todos os filtros linearizados que se utilizam da estrutura das equações de Kalman (2.11 - 2.13) mantêm a hipótese de que todas as densidades intermediárias para o cálculo da densidade a posteriori são gaussianas. Até que ponto esta hipótese representa uma boa aproximação para as densidades reais? Em que condições deixa de ser válida? Que ações deve tomar o algoritmo frente a estes casos? A informação trazida por uma nova medida é efetivamente introduzida para corrigir o valor predito do estado? De que forma se poderia utilizar melhor esta informação que certamente está disponível? Estas questões nortearão a pesquisa do próximo Capítulo.

Para finalizar esta seção apresentam-se os passos do filtro ótimo para o caso linear acompanhados de algumas ilustrações com o intuito de tornar mais fácil a compreensão dos linearizados.

Sejam as funções f e h lineares dadas por 2.10 com as mesmas propriedades estatísticas assumidas na primeira seção. Torna-se fácil o cálculo da densidade de previsão de estado pois as funções de verossimilhança $L(x_k : x_{k+1})$ tem a forma gaussiana.

$$L(x_k : x_{j,k+1}) = p(x_{j,k+1}/x_k) = p\left(w_k = \Gamma_k^{-1} \cdot (x_{j,k+1} - F_k x_k)\right) ||\Gamma_k^{-1}||$$

$$L(x_k : x_{j,k+1}) = \text{Cte} \cdot \exp\left(-0.5 // x_{j,k+1} - F_k x_k // (\Gamma_k W_k \Gamma_k^T)^{-1}\right)$$

A densidade de previsão de estado é obtida por uma convolução linear como se pode ver pela expressão a seguir e demonstra-se que se a densidade do instante anterior é gaussiana, aquela também o será (Figura 2.7).

$$p(x_{k+1}/J_k) = \int \text{CTE} \cdot \exp \left(-0,5 // F_k \left(F_k^{-1} x_{k+1} - \hat{x}_{k/k} - (x_k - \hat{x}_{k/k}) \right) // W_k^{*-1} \right) \cdot \exp \left(-0,5 // x_k - \hat{x}_{k/k} // X_{k/k}^{-1} \right) dx_k$$

$$\text{com } W^* = \Gamma_k W_k \Gamma_k^T$$

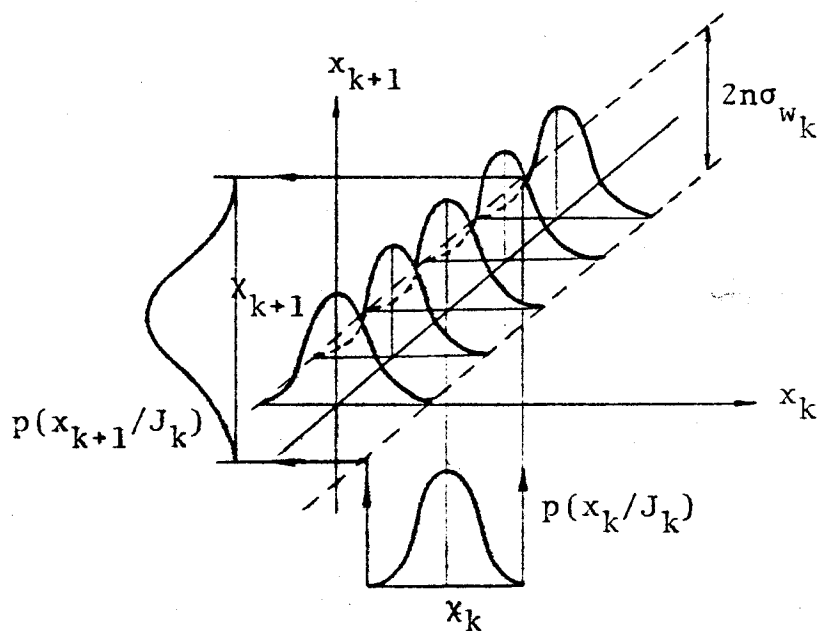


Figura 2.7 - Funções de verossilhança e convolução linear

Para uma função de medidas linear também as funções de verossilhança $L_1(x_{k+1}:y_{\text{obs}})$ têm a forma gaussiana:

$$L_1(x_{k+1}:y_{\text{obs}}) = p(y_{k+1}=y_{\text{obs}}/x_{k+1}) = p(v_{k+1}=y_{\text{obs}}-H_{k+1}x_{k+1})$$

$$L_1(x_{k+1}:y_{\text{obs}}) = \text{Cte} \exp \left(-0,5 // y_{\text{obs}} - H_{k+1} x_{k+1} // V_{k+1}^{-1} \right)$$

A densidade a posteriori é obtida pelo produto desta forma gaussiana com a densidade de previsão de estado, que é gaussiana, resultando outra forma gaussiana que será normalizada. A operação pode ser seguida junto ao gráfico 2.8.

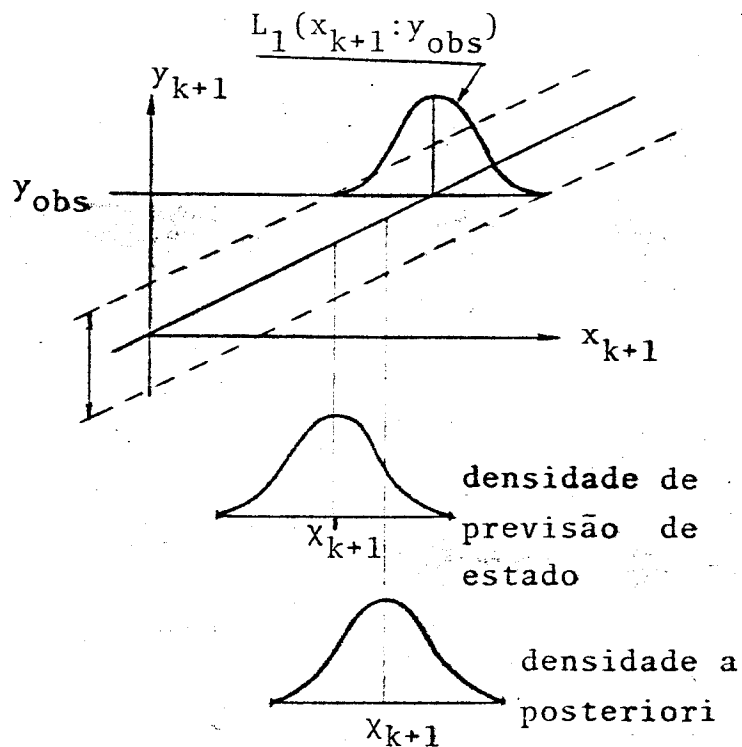


Figura 2.8 - Determinação da densidade a posteriori

O que se pretendeu deixar claro, embora sem demonstração, pois ela é facilmente encontrável na literatura, é que se o filtro é inicializado com uma densidade a priori do estado gaussiana, a densidade de previsão do estado também o será se a função de verossimilhança L tiver uma forma gaussiana. Isto acontece se o estado evoluir linearmente. Do mesmo modo, a densidade a posteriori será gaussiana se, partindo de uma densidade de previsão de estado gaussiana, a função de verossimilhança L_1 tiver uma forma gaussiana, o que será verdade para uma função de medidas h linear. Como a densidade a posteriori atual é a densidade a priori para o próximo instante, fecha-se o círculo garantindo que todas as densidades

a posteriori serão gaussianas. Daqui resulta óbvio que a finalidade da linearização é para aproximar por gaussianas as funções de verossimilhança L e L_1 .

II.6 - EXEMPLO

O algoritmo de reconstrução numérica das densidades foi aplicado ao sistema descrito abaixo:

$$x_{k+1} = 0,9 x_k + \frac{-10 x_k}{1 + x_k^2} - 0,1 K + w_k$$

$$y_k = 0,5 (x_k + 40) + \frac{-10(x_k + 40)}{1 + x_k^2} + v_k$$

O estado inicial é gaussiano e os ruídos são brancos gaussianos, mutuamente independentes e independentes do estado inicial cujas propriedades são:

estado inicial : $N(15,10)$

ruído de estado : $N(0,1)$

ruído de medida : $N(0,1)$

$N(A,B)$ representa uma gaussiana de média A e variância B .

As evoluções do estado e medida são descritas pelas funções não lineares f e h abaixo:

$$f(x) = 0,9x - \frac{10 x}{1 + x^2}$$

$$h(x) = 0,5x - \frac{10(x + 40)}{1 + x^2}$$

As equações de transição de estado apresenta um comportamento fortemente não linear no intervalo compreendido entre os valores -10 e 10 sendo aproximadamente linear fora deste intervalo. A não linearidade da função de medidas é semelhante à da transição de estado, mas o comporta

mento fortemente não linear localiza-se no intervalo (-50 , -30). As regiões de não linearidades acentuadas estão suficientemente separadas de maneira que as medidas sejam aproximadamente lineares no intervalo da não linearidade da transição de estado e esta seja linear onde aquela é fortemente não linear. Tal separação é conveniente pois permite distinguir as diferentes influências de cada não linearidade (Fig. 2.9).

O estado no instante de inicialização do algoritmo é igual a 20 observando-se sua evolução durante 90 instantes subsequentes. Pode-se perceber pela Fig. 2.10 que o estado evolui pela região fortemente não linear entre os instantes 8 e 20 e as medidas entre os instantes 49 e 62.

A densidade a posteriori é aproximadamente gaussiana entre os instantes $k=1$ a 7, $k=21$ a 49 e $k=63$ a 90 quando tanto o estado quanto a respectiva medida se encontram nos intervalos em que as funções f e h são praticamente lineares. Entre os instantes $k=8$ e $k=20$, o estado percorre a região acentuadamente não linear da função f e a densidade a posteriori tem um caráter fortemente assimétrico. Para os instantes $k=49$ a 62, o estado evolui de forma quase linear, mas as observações são acentuadamente não lineares, conduzindo a densidades a posteriori assimétricas ou multimodais. Algumas densidades típicas estão ilustradas na Fig. 2.11.

A Fig. 2.12 apresenta o erro de estimação ϵ_k e o desvio padrão σ_k .

- Erro de estimação:

$$\epsilon_k = x_k^r - \hat{x}_{k,k} \quad (x_k^r \text{ é o estado real})$$

- Desvio padrão:

$$\sigma_k = \sqrt{x_{k,k}}$$

Para evitar a influência de realizações particulares de ruído, as Figs. 2.13 e 2.14 descrevem os resultados provenientes de simulações de Monte Carlo com médias sobre M ($=20$) sequências amostradas:

- Erro médio:

$$\bar{\varepsilon}_k = \frac{1}{M} \sum_{i=1}^M \varepsilon_k^i$$

- Desvio padrão médio:

$$\bar{\sigma}_k = \left\{ \frac{1}{M} \sum_{i=1}^M x_{k/k}^i \right\}^{1/2}$$

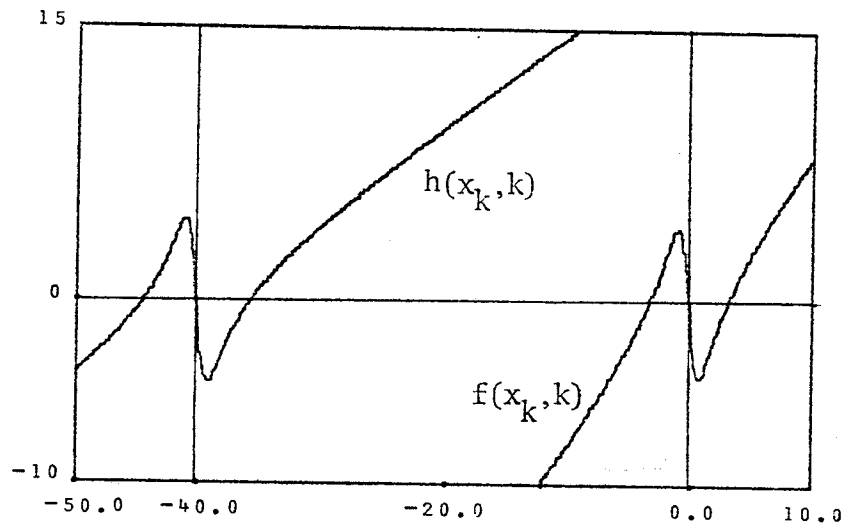


Fig. 2.9 - Não linearidades do estado e da medida

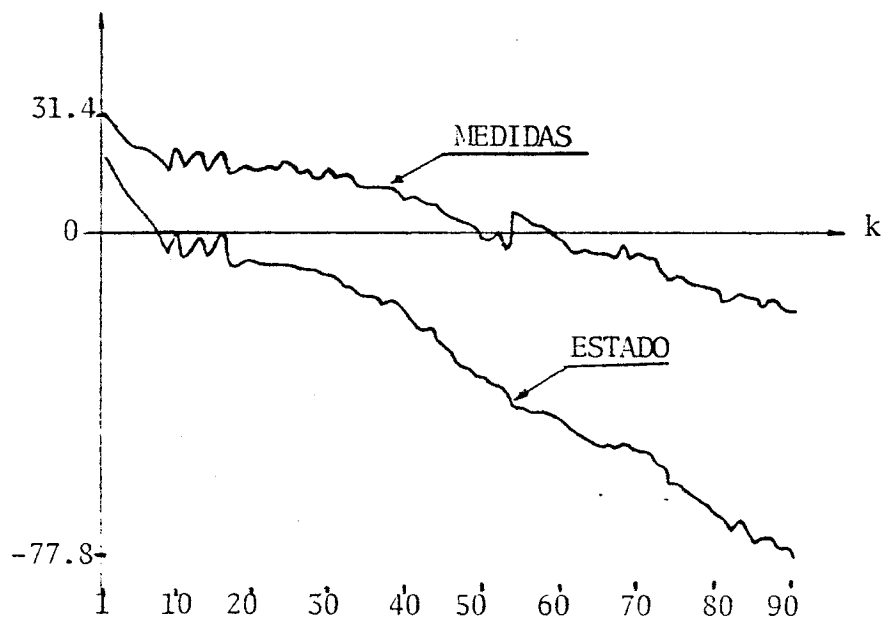
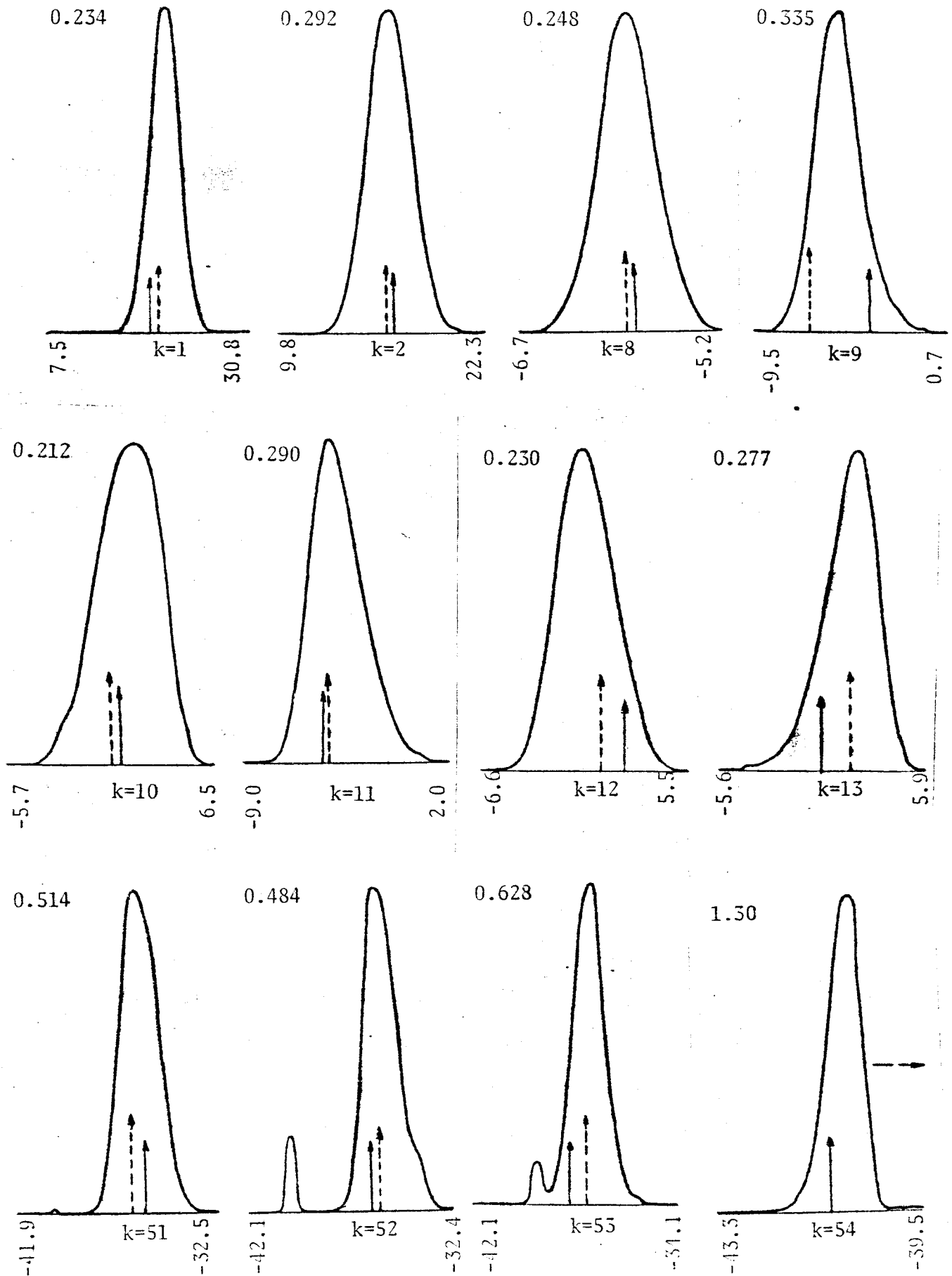
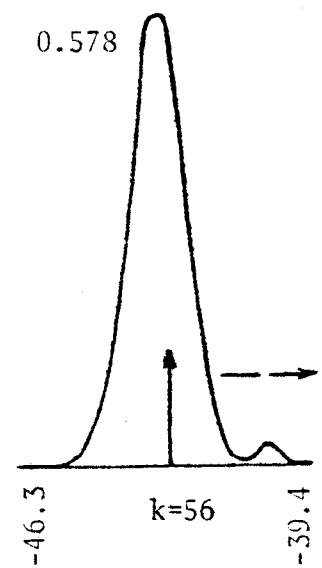
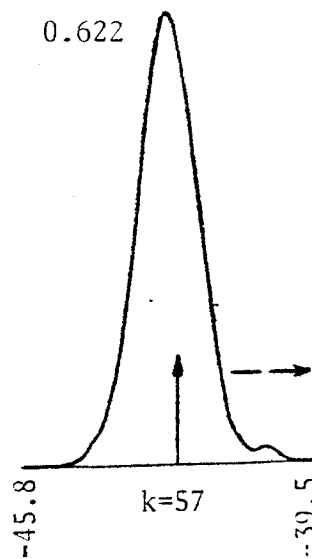
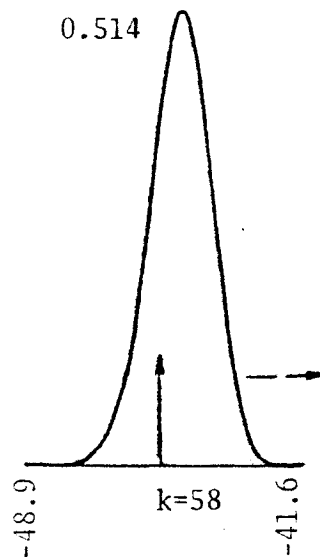
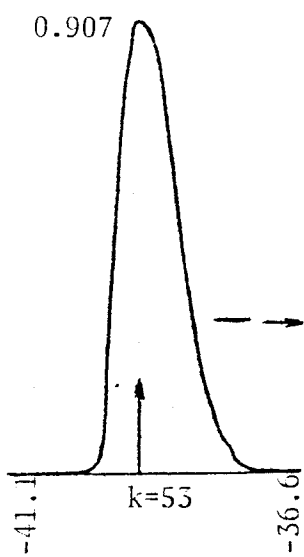
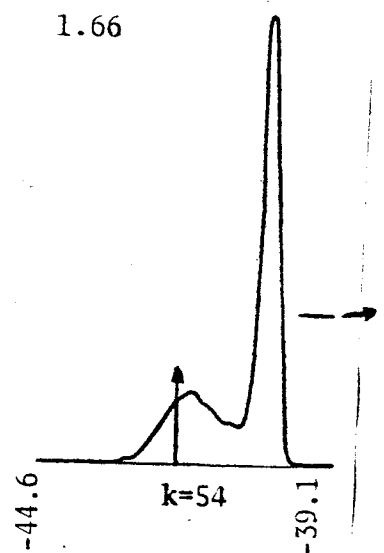
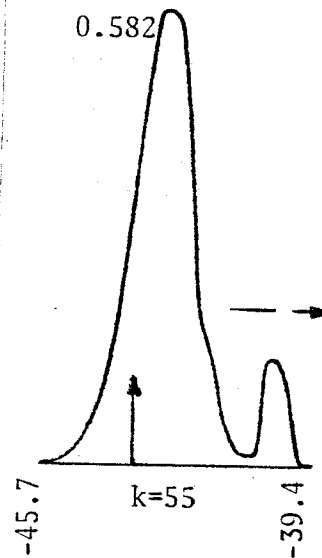
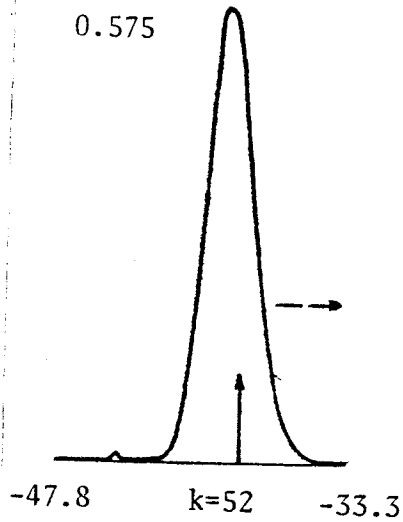
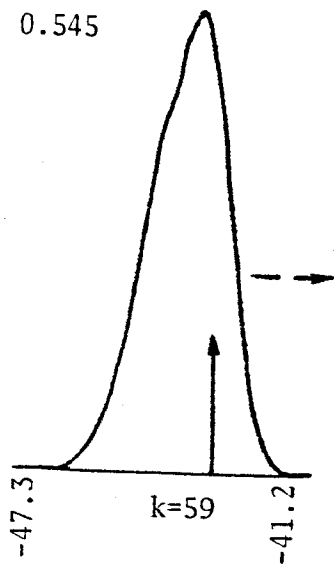
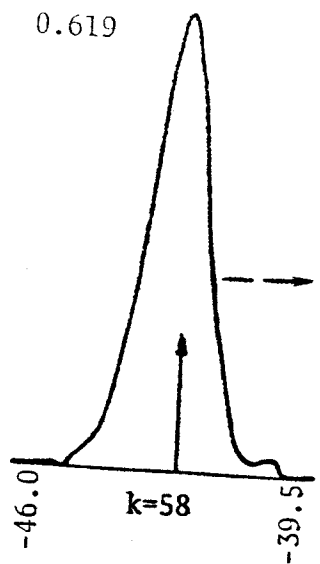
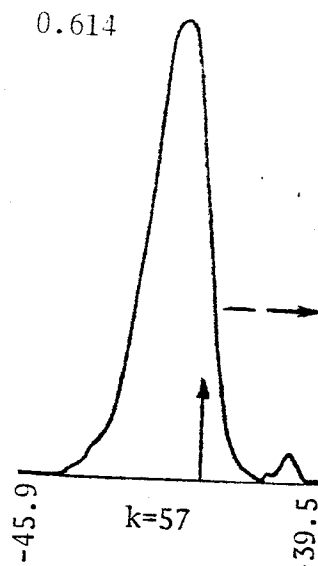
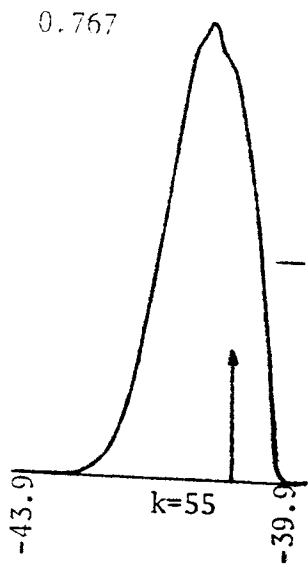


Fig. 2.10 - Evolução do estado e respectiva medida

Fig. 2.11 - Algumas densidades típicas para duas sequências aleatórias

..... estado real
..... estado estimado pelo EXTENDIDO
→ estado estimado pelo EXTENDIDO cai fora da escala





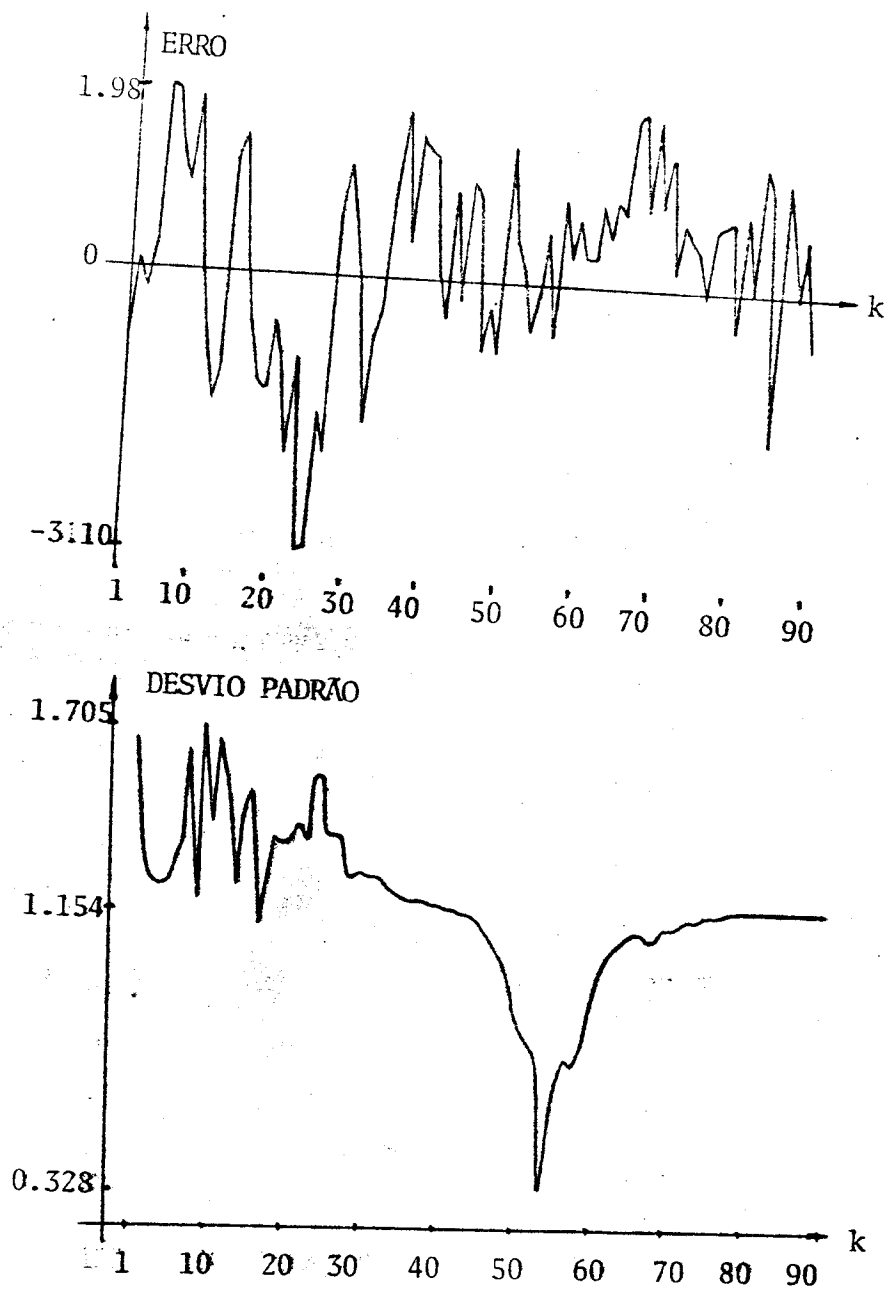


Fig. 2.12 - Erro de estimação e desvio padrão

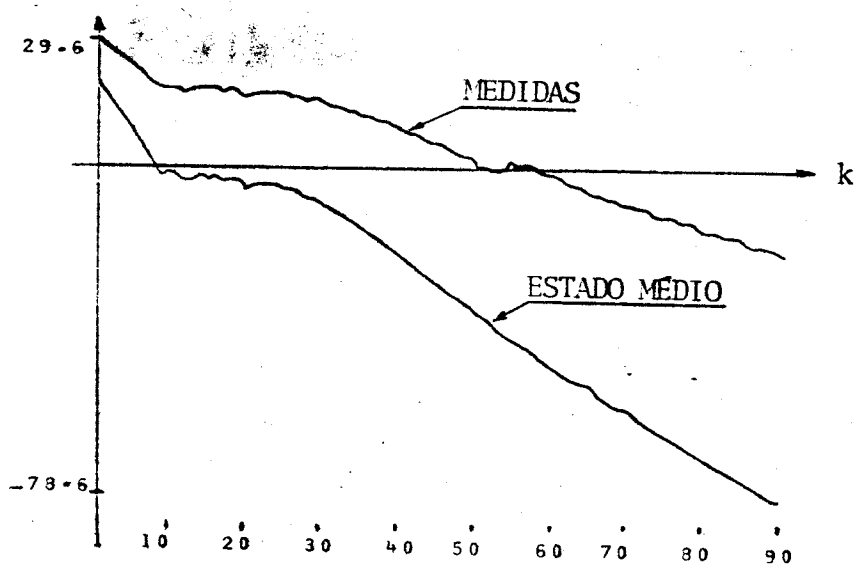


Fig. 2.13 - Evolução média do estado e medidas

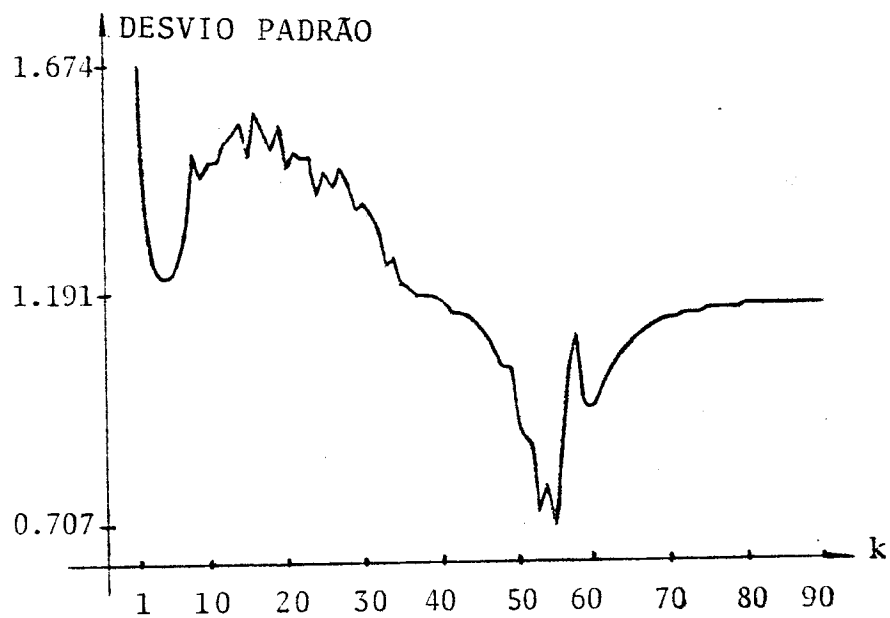
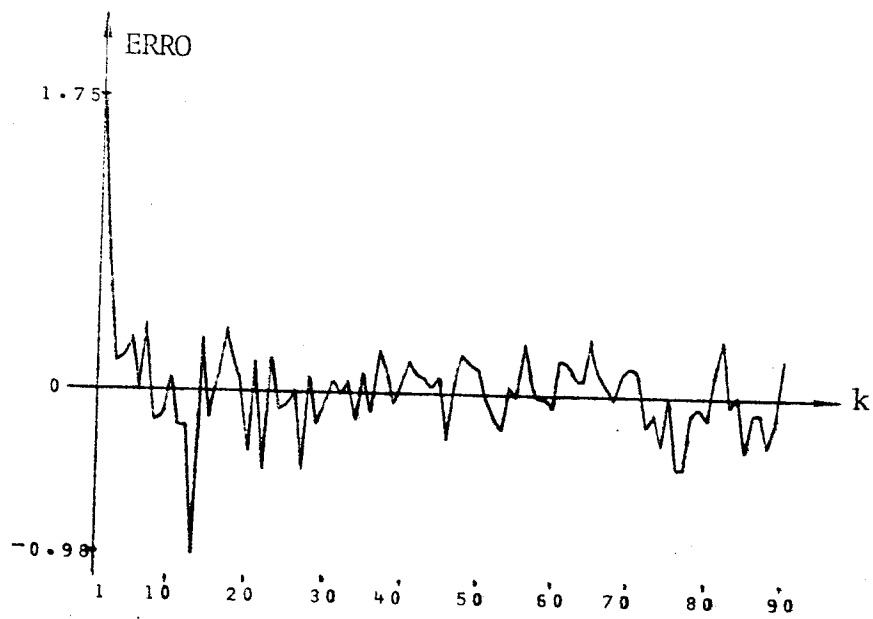


Fig. 2.14 - Erro de estimação médio
e desvio padrão médio

CAPÍTULO III

FILTROS SUBÓPTIMOS

Considere-se um modelo de sistema descrito pelas equações de evolução do estado e medida a seguir:

$$x_{k+1} = f(x_k, k) + \Gamma_k w_k \quad (n \times 1) \quad 3.1$$

$$y_k = h(x_k, k) + v_k \quad (m \times 1) \quad 3.2$$

$$E \{ x_0 \} = \hat{x}_0$$

$$E \{ (x - \hat{x}_0) (x - \hat{x}_0)^T \} = X_0$$

$$E \{ w_k \} = E \{ v_k \} = 0$$

$$E \{ w_k w_k^T \} = W_{k.1} \delta_{k.1} \quad 3.3$$

$$E \{ v_k v_k^T \} = V_{k.1} \delta_{k.1}$$

$$E \{ w_k v_k^T \} = E \{ w_k \hat{x}_0^T \} = E \{ v_k \hat{x}_0^T \} = 0$$

Considerando que a solução numérica ótima não é viável para implementação prática por causa do tempo de processamento demasiadamente longo, é necessário desenvolver métodos aproximados para calcular a densidade a posteriori. O objeto central das aproximações realizadas tem sido o de reter para a solução aproximada, a estrutura das equações de Kalman (2.11 - 2.13) pela sua simplicidade e rapidez de cálculo. A hipótese básica assumida implicitamente neste procedimento é a de que todas as densidades são satisfatoriamente aproximadas por gaussianas. Para isto, as equações que compõem o sistema são linearizadas a cada novo instante por desenvolvimento em série de Taylor em torno de certos valores de referência, como se mostra abaixo:

$$f(x_k, k) = f(\bar{x}_k, k) + F_{\bar{x}} \cdot (x_k - \bar{x}_k) + 1/2 \sum_{i,j=1}^n (x_k - \bar{x}_k)^i (x_k - \bar{x}_k)^j F_{\bar{x}\bar{x}}^{ij} + \dots$$

$$h(x_k, k) = h(\bar{x}_k, k) + H_{\bar{x}} \cdot (x_k - \bar{x}_k) + 1/2 \sum_{i,j=1}^n (x_k - \bar{x}_k)^i (x_k - \bar{x}_k)^j H_{\bar{x}\bar{x}}^{ij} + \dots$$

onde

$$F_x = \left. \frac{\partial f(x_k, k)}{\partial x_k} \right|_{x_k = \bar{x}_k} \quad (n \times n)$$

$$F_{xx}^{ij} = \left. \frac{\partial^2 f(x_k, k)}{\partial x_k^i \partial x_k^j} \right|_{x_k = \bar{x}_k} \quad (n \times 1)$$

3.4

$$H_x = \left. \frac{\partial h(x_k, k)}{\partial x_k} \right|_{x_k = \bar{x}_k} \quad (m \times n)$$

$$H_{xx}^{ij} = \left. \frac{\partial^2 h(x_k, k)}{\partial x_k^i \partial x_k^j} \right|_{x_k = \bar{x}_k} \quad (m \times 1)$$

\underline{i} e \underline{j} indicam componentes do vetor x_k ou do vetor $(x_k - \bar{x}_k)$ cuja dimensão é \underline{n} . (Para seguir o desenvolvimento acima, ver Bruno, Paulo T.M. (5)).

III.1 - FILTROS GENERALIZADOS

Os valores em torno dos quais a linearização é feita, podem, em princípio, ser os valores assumidos pela evolução do estado e respectiva medida na ausência de ruídos. No entanto, é de se esperar que os valores filtrados em cada instante, já que levam em conta as observações já realizadas, estejam mais próximos do verdadeiro valor do estado. Muito pode ser ganho (9) se as linearizações forem feitas em torno destes pontos, vale dizer, se a trajetória de referência for a dos valores estimados até este momento. Então $f(x_k, k)$ será linearizada em torno do valor filtrado do instante anterior $\bar{x} = \hat{x}_{k/k}$ e $h(x_{k+1}, k+1)$ em torno do valor predito do estado $\bar{x}_{k+1} = \hat{x}_{k+1/k}$.

III.1.1 - Filtro de Kalman extendido : EX

Se o desenvolvimento por série de Taylor

das funções \underline{f} e \underline{h} for limitado aos termos de primeira ordem, as equações de recorrência da média e da covariância da densidade a posteriori são conhecidas como filtro de Kalman estendido e será denotado por EX.

As densidades de previsão de estado e medida são descritas por:

$$\hat{x}_{k+1/k} = f(\hat{x}_{k/k}, k) \quad 3.5$$

$$X_{k+1/k} = F_x X_{k/k} F_x^T + \Gamma_k W_k \Gamma_k^T$$

$$\hat{y}_{k+1/k} = h(\hat{x}_{k+1/k}, k) \quad 3.6$$

$$Y_{k+1/k} = H_x X_{k+1/k} H_x^T + V_{k+1}$$

A evolução da densidade a posteriori será dada pelas equações 2.13, pois dado que o sistema foi linearizado, as densidades de previsão são gaussianas e a densidade a posteriori também o será.

Nas linhas seguintes, procura-se discutir a validade da linearização por série de Taylor, abrangendo exclusivamente termos de primeira ordem, e a hipótese subsequente de normalidade, à luz do algoritmo de reconstrução numérica das densidades apresentado no Capítulo precedente.

Quando as não linearidades são aproximadas por linhas retas, as funções de verossimilhança resultantes terão formas gaussianas, com certeza diferentes das verdadeiras. Esta diferença será tanto maior quanto mais pronunciada for a não linearidade, podendo conduzir à "divergências".

Pode-se esperar que o comportamento do filtro estendido não apresente a tendência a diver

gir quando:

- 1) o domínio x_k é tal que as funções de verossimilhança $L(x_k:x_{j,k+1})$ apresentem não mais do que uma moda neste domínio, de tal forma que a hipótese de normalidade para a densidade de previsão é aproximadamente válida, e
- 2) a função $L_1(x_{k+1}:y_{obs})$ tem apenas um máximo no intervalo x'_{k+1} permanecendo igualmente válida a aproximação normal para a densidade a posteriori.

Entretanto, se x_k ou x'_{k+1} forem suficientemente grandes a ponto de apanharem mais do que um máximo das funções $L(x_k:x_{j,k+1})$ ou $L_1(x_{k+1}:y_{obs})$ a linearização por desenvolvimento em série de Taylor não será mais representativa das não linearidades nos respectivos domínios. Assumindo, por exemplo, uma densidade a posteriori anterior gaussiana com domínio x_k no qual as funções L têm mais do que um máximo, e isto com certeza acontece se neste domínio a derivada de primeira ordem da função f mudar de sinal, a linearização, considerando apenas os termos de primeira ordem, tende a reduzir a variância da densidade de previsão de estado em relação à do filtro ótimo e introduz um viés como mostra a Fig. 2.7. Como a densidade a posteriori é obtida pelo produto da densidade de previsão de estado com a função L_1 , o filtro atribui um peso excessivo à previsão do estado, praticamente ignorando a informação trazida pela nova medida e a variância da densidade a posteriori será consequentemente menor do que deveria ser.

Isto pode ser facilmente mostrado através das equações 2.13.

Este efeito e o viés podem conduzir à erros de estimação que se situem além da devida proporção com o erro quadrático previsto pelas equações de Filtragem.

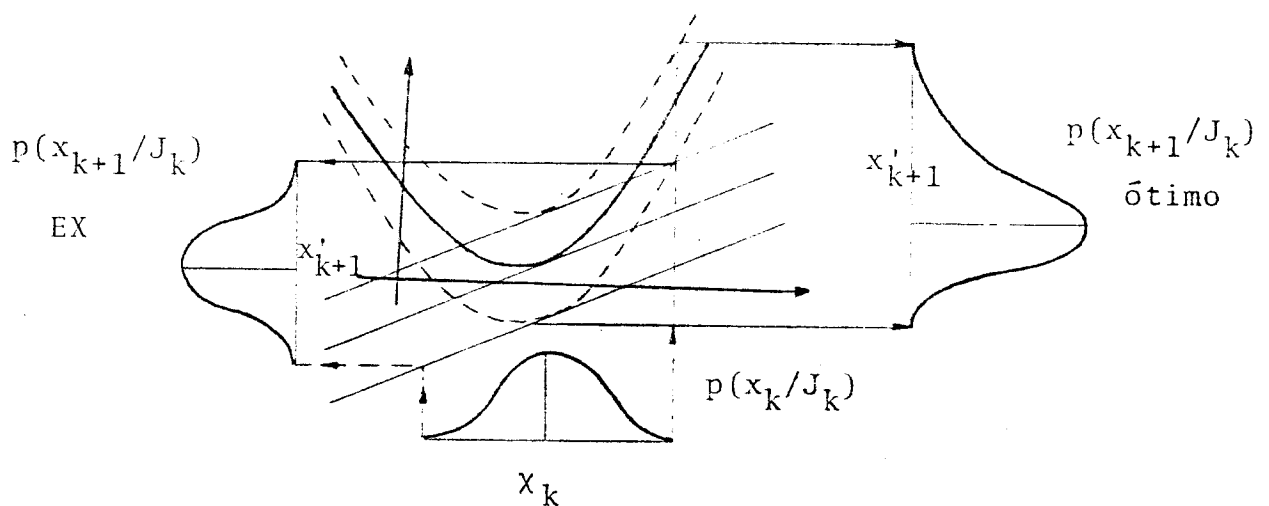


Fig. 3.1 - Influência da não linearidade f no cálculo da densidade de previsão de estado

Por outro lado, partindo de uma densidade de previsão de estado gaussiana com o domínio x'_{k+1} no qual a função de verossimilhança $L_1(x_{k+1}; y_{obs})$ apresenta mais do que um máximo, isto é, neste domínio a derivada primeira da função h muda de sinal, a aproximação por desenvolvimento em série de Taylor desprezando os termos de segunda ordem em diante, fará com que apenas um dos máximos seja considerado, reduzindo sensivelmente a variância da função de verossimilhança linearizada em relação à verdadeira. Novamente o resultado será a diminuição da variância da densidade a posteriori. No entanto, o comportamento do filtro nesta situação é em boa medida oposto ao do caso anterior, pois a ponderação maior será agora exercida pela função de verossimilhança, atribuindo pouco peso à informação acumulada na densidade de previsão de estado. Além do mais, não há razões suficientes para esperar que a linearização conduza à escolha do máximo mais adequado da função L_1 e em geral a

função de verossimilhança resultante da aproximação linear $L_{1,\ell}$ não será uma aproximação adequada para a real (Fig. 2.8). Esta má aproximação pode conduzir ao fenômeno descrito na literatura como "divergência" : "os erros no valor estimado eventualmente assumem valores inteiramente fora da proporção com os erros médios quadráticos previstos pelas equações de filtragem" (8). A explicação mais frequentemente oferecida pela literatura para este fenômeno é que a variância calculada se torna menor do que a real de forma tal que o filtro "confia" demasiadamente nos seus valores estimados, ignorando efetivamente as medidas subsequentes e o estado real se afasta progressivamente do valor estimado. Neste trabalho emprega-se o conceito "divergência" para uma situação um pouco diferente. Diz-se-á que o filtro tende a divergir em um instante particular quando em simulações de Monte Carlo o valor médio quadrático do erro de estimação se torna muito maior do que a variância da densidade a posteriori, mesmo que nos instantes subsequentes o filtro possa se corrigir.

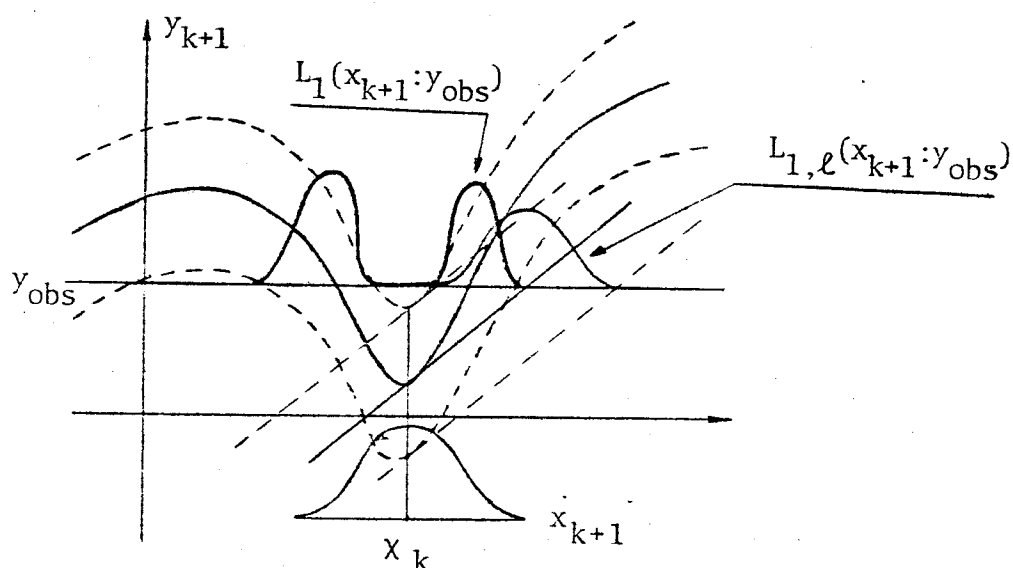


Fig. 3.2 - Influência da não linearidade h na função de verossimilhança

É preciso, neste ponto, esclarecer uma diferença essencial entre o cálculo da densidade de previsão de estado através de uma convolução e a obtenção final da densidade a posteriori através do produto da densidade de previsão com a função $L_1(x_{k+1}:y_{obs})$. Em primeiro lugar, repita-se, existe uma grande diferença de tempo de cálculo para a aproximação numérica ótima: enquanto que para obter a densidade de previsão é necessário efetuar um certo número de integrações, a densidade final é obtida por uma simples multiplicação. Todavia, não é esta a diferença a ressaltar. Porém sustenta-se que o filtro linearizado por série de Taylor com termos até primeira ordem tem um desempenho melhor na previsão do estado do que na obtenção da densidade a posteriori. Compreenda-se o porque.

O filtro ótimo leva em conta a não linearidade do estado no cálculo da densidade de previsão através da função de verossimilhança $L(x_k:x_{j,k+1})$. Suponha-se que a densidade do instante anterior seja aproximadamente normal. Esta densidade pondera as funções de verossimilhança $L(x_k:x_{j,k+1})$. A ponderação exercida é maior na vizinhança de sua média. Mas é justamente nesta vizinhança que a linearização por série de Taylor é mais representativa da não linearidade. À medida que se vai afastando desta vizinhança, a aproximação linear deixa progressivamente de ser válida, a função de verossimilhança linearizada L_ℓ difere cada vez mais da verdadeira L , no entanto, a ponderação exercida pela densidade anterior também foi se tornando menor, reduzindo a influência dos erros de linearização (Fig. 3.3).

Bem distinto é o caso da não linearidade da medida. Esta é introduzida pelo filtro ótimo na função de verossimilhança L_1 que multiplica a densidade de previsão. O valor observado pode diferir bastante do valor predito da saída. Co

Fig. 3.4 .- Linearização e produto

III.1.2 - Filtro de segunda ordem : SO

A aproximação das funções \underline{f} e \underline{h} considerando apenas os termos lineares pode estar in tro duzindo erros de estimação grosseiros por cau sa dos próprios erros de linearização, principal mente quando o estado estimado anterior diferir bastante do estado real. O mesmo é válido se os domínios x_k e x'_{k+1} forem grandes.

Evidentemente, estes erros serão reduzi dos se forem considerados também os termos de se gunda ordem na aproximação das funções \underline{f} e \underline{h} . Para manter as densidades de previsão de estado e medida gaussianas, os momentos de ordem superior a dois ímpares são desprezados. Se os de quarta ordem forem desprezados, obtém-se o filtro trunca do modificado, se forem escritos em função dos de ordem dois obtém-se o filtro gaussiano modificado (5, 9, 12).

- equações de previsão de estado

$$\hat{x}_{k+1/k} = f(\hat{x}_{k/k}, k) + \sum_{ij=1}^n x_{k/k}^{ij} F_{xx}^{ij} \quad 3.7$$

Filtro truncado

$$x_{k+1/k} = F_x x_{k/k} F_x^T - 1/4 \sum_{ijp\ell=1}^n x_{k/k}^{ij} x_{k/k}^{p\ell} F_{xx}^{ij} (F_{xx}^{p\ell})^T + \Gamma_k W_k \Gamma_k^T$$

Filtro gaussiano

$$x_{k+1/k} = F_x x_{k/k} F_x^T + 1/4 \sum_{ijp\ell=1}^n (x_{k/k}^{ip} x_{k/k}^{j\ell} + x_{k/k}^{i\ell} x_{k/k}^{jp}) F_{xx}^{ij} (F_{xx}^{p\ell})^T + \Gamma_k W_k \Gamma_k^T \quad 3.8$$

- equações de previsão de medida

$$\hat{y}_{k+1/k} = h(\hat{x}_{k+1/k}, k) + \sum_{ij=1}^n X_{k+1/k}^{ij} H_{xx}^{ij} \quad 3.9$$

Filtro truncado

$$Y_{k+1/k} = H_x^T X_{k+1/k} H_x^{-1/4} \sum_{ijp\ell=1}^n X_{k+1/k}^{ij} X_{k+1/k}^{p\ell} H_{xx}^{ij} (H_{xx}^{p\ell})^T + V_{k+1}$$

Filtro gaussiano

$$Y_{k+1/k} = H_x^T X_{k+1/k} H_x + 1/4 \sum_{ijp\ell=1}^n (X_{k+1/k}^{ip} X_{k+1/k}^{j\ell} + X_{k+1/k}^{i\ell} X_{k+1/k}^{jp}) H_{xx}^{ij} (H_{xx}^{p\ell})^T + V_{k+1} \quad 3.10$$

com $X_{k/k}^{ij}$: elemento da i-ésima linha e j-ésima coluna da matriz de covariância $X_{k/k}$

$X_{k+1/k}^{ij}$: elemento da i-ésima linha e j-ésima coluna da matriz de covariância $X_{k+1/k}$

F_x, H_x, F_{xx} e H_{xx}^{ij} definidos por 3.4

Em (9) podem ser encontradas outras aproximações de segunda ordem que não serão simuladas neste trabalho. Das duas acima expostas, apenas são simuladas as equações do filtro gaussiano denotado simplesmente por S0, pois a existência de um termo negativo nas equações do filtro truncado pode conduzir a variâncias negativas.

Espera-se que esta aproximação até segunda ordem apresente melhores resultados do que os conseguidos com as aproximações de primeira ordem pelo fato de reduzir os erros de linearização. No entanto, as densidades de previsão foram aproximadas por gauss

sianas e isto equivale efetivamente a uma aproximação linear. Que modificações são, pois, introduzidas pelos termos de segunda ordem? Para facilidade e clareza de exposição, a análise será feita considerando um caso escalar.

Pelas equações 3.7 e 3.9, pode-se ver que os termos de segunda ordem introduzem uma correção nos valores previstos do estado e medida proporcional às derivadas de segunda ordem. Ao mesmo tempo, as respectivas variâncias são aumentadas (3.8 e 3.10). Isto equivale a considerar um sistema linear no qual foram modificadas as propriedades estatísticas dos ruídos

$$x_{k+1} = f(\hat{x}_{k/k}, k) + F_x \cdot (x_k - \hat{x}_{k/k}) + \eta_k$$

$$y_{k+1} = h(\hat{x}_{k+1/k}, k+1) + H_x \cdot (x_{k+1} - \hat{x}_{k+1/k}) + \xi_{k+1}$$

com

$$\bar{\eta}_k = E\{\eta_k\} = 1/2 F_{xx} X_{k/k} \quad E\{(\eta_k - \bar{\eta}_k)^2\} = \Gamma_k^2 W_k + 1/2 F_{xx}^2 X_{k/k}^2$$

$$\bar{\xi}_{k+1} = E\{\xi_{k+1}\} = 1/2 H_{xx} X_{k+1/k} \quad E\{(\xi_{k+1} - \bar{\xi}_{k+1})^2\} = V_{k+1} + 1/2 H_{xx}^2 X_{k/k}^2$$

Graficamente pode ser ilustrado como na Fig. 3.5.

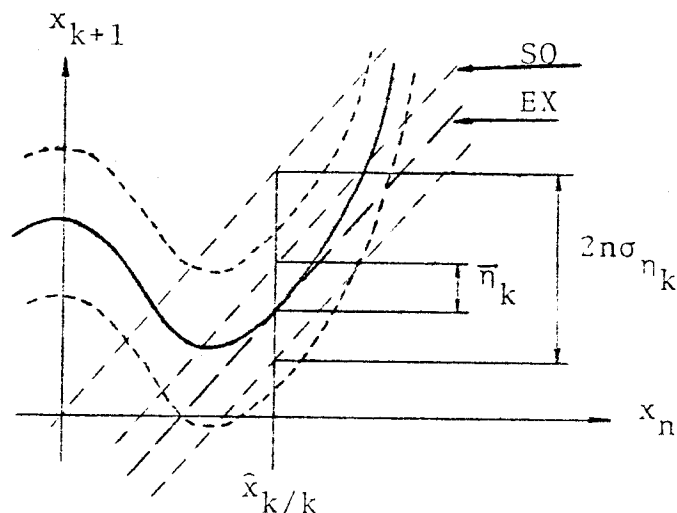


Fig. 3.5 - Sistema linear correspondente à

O filtro SO apresentará uma tendência menor a divergir pelo fato de aumentar a variância da densidade a posteriori; mas, às vezes, este aumento é também o seu defeito. A variância pode se tornar excessivamente grande a ponto de só se poder atribuir muito pouco significado aos valores estimados. O filtro SO assume desta maneira um caráter "conservativo". Outras vezes, e principalmente nas regiões de não linearidades de medida acentuadas, o aumento da variância pode não ser suficiente para evitar a divergência. Isto reforça a hipótese de que a aproximação por desenvolvimento em série de Taylor só representa a não linearidade localmente, sugerindo outro tipo de aproximação que represente melhor as não linearidades nos respectivos domínios de interesse x_k e x'_{k+1} .

III.2 - FILTROS FITADOS

Foi discutido anteriormente que as linearizações por desenvolvimento em série de Taylor não são representativas do caráter geral das não linearidades, principalmente para regiões em que as funções de verossimilhança podem apresentar mais do que um máximo. Com o intuito de diminuir os erros devidos às linearizações, propõe-se uma outra aproximação para as funções f e h nos domínios x_k e x'_{k+1} respectivamente. Ao invés de linearizar as funções f e h por desenvolvimento em série de Taylor em torno do valor médio da densidade a posteriori do instante anterior e do valor médio da densidade de previsão de estado, respectivamente, as funções são aproximadas a cada passo da filtragem por polinômios do tipo $\sum_{i=1}^{p+1} a_i (x-\bar{x})^{i-1}$ onde a_i são coeficientes a determinar, \bar{x} o ponto em torno do qual se fará a aproximação e p é o grau do polinômio. Desenvolve-se a seguir o respectivo algoritmo de filtragem bem como um método para a obtenção dos coeficientes, considerando-se apenas o caso escalar.

III.2.1 - Equações do filtro polinomial fitado : FI

As aproximações para as funções f e h

considerando apenas o caso escalar, são:

$$f(x_k, k) \cong \sum_{i=1}^{p+1} a_i (x_k - \hat{x}_{k/k})^{i-1}$$

$$h(x_{k+1}, k+1) \cong \sum_{i=1}^{q+1} b_i (x_{k+1} - \hat{x}_{k+1/k})^{i-1}$$

Para o desenvolvimento que se fará a seguir, considera-se que a evolução do sistema e da medida sejam descritos pelas equações 3.11 e 3.12, com as propriedades estatísticas do estado inicial e dos ruídos dados por 3.3.

$$x_{k+1} = \sum_{i=1}^{p+1} a_i (x_k - \hat{x}_{k/k})^{i-1} + \Gamma_k \cdot w_k \quad 3.11$$

$$y_{k+1} = \sum_{i=1}^{q+1} b_i (x_{k+1} - \hat{x}_{k+1/k})^{i-1} + v_{k+1} \quad 3.12$$

Apesar das relações 3.11 e 3.12 não serem lineares, deseja-se manter a estrutura de Kalman para o filtro fitado. É fácil demonstrar que se as densidades de previsão de estado e de previsão de medida forem gaussianas, a densidade a posteriori obtida pela normalização do produto das densidades mencionadas será também gaussiana. Obviamente para um sistema linear com medidas lineares e nas hipóteses 3.3, as densidades de previsão de estado e medida são gaussianas. Como no presente caso a aproximação não é linear, as densidades são aproximadas por gaussianas, desprezando os momentos de ordem superior a dois ímpares e escrevendo os de ordem par em função dos de segunda ordem ().

$$E \left\{ (x - E\{x\})^{2n} \right\} = 1.3.5 \dots (2n-1) \left(E \left\{ (x - E\{x\})^2 \right\} \right)^n \quad 3.13$$

Dados os dois primeiros momentos das densidades de previsão de estado e medida, a densidade a posteriori será descrita pelas equações do fil

tro linear de Kalman 2.13. As densidades de previsão são dadas pelas expressões abaixo:

- Equações de previsão de estado:

$$\hat{x}_{k+1/k}^{\Delta} = E \left\{ x_{k+1}/J_k \right\}$$

$$\hat{x}_{k+1/k} = E \left\{ \sum_{i=1}^{p+1} a_i (x_k - \hat{x}_{k/k})^{i-1} + \Gamma_k \cdot w_k / J_k \right\}$$

$$\hat{x}_{k+1/k} = \sum_{i=1}^{p+1} a_i E \left\{ (x_k - \hat{x}_{k/k})^{i-1} / J_k \right\}$$

Desprezando os momentos de ordem ímpar superiores a dois e escrevendo os de ordem par pela relação 3.13 obtém-se para $p=5$:

$$\hat{x}_{k+1/k} = a_1 + a_3 x_{k/k} + 3a_5 x_{k/k}^2 \quad 3.14$$

$$x_{k+1/k}^{\Delta} = E \left\{ (x_{k+1} - \hat{x}_{k+1/k})^2 / J_k \right\}$$

$$x_{k+1/k} = E \left\{ x_{k+1}^2 / J_k \right\} - \hat{x}_{k+1/k}^2$$

$$\begin{aligned} E x_{k+1}^2 / J_k &= E \left\{ \sum_{i,j=1}^6 a_i a_j (x_k - \hat{x}_{k/k})^{i+j-2} / J_k \right\} + \\ &+ E \left\{ \sum_{i=1}^6 a_i (x_k - \hat{x}_{k+1/k})^{i-1} \cdot \Gamma_k w_k / J_k \right\} + \\ &+ E \left\{ \Gamma_k^2 w_k^2 / J_k \right\} \end{aligned}$$

Desprezando os momentos de ordem superior a dois ímpares e escrevendo os de ordem par pela relação 3.12. e lembrando que o segundo termo desta última soma é nulo, pois x_k é independente de w_k , chega-se ao seguinte resultado:

$$E \left\{ x_{k+1}^2 / J_k \right\} = a_1^2 + (a_2^2 + 2a_1 a_3) x_{k/k} + (a_3^2 + 2a_2 a_4 + 2a_1 a_5) 3x_{k/k}^2 +$$

$$+ (a_4^2 + 2a_3a_5 + 2a_2a_6)15x_{k/k}^3 + (a_5^2 + 2a_4a_6)105x_{k/k}^4 +$$

$$+ 945 a_6 x_{k/k}^5 + \Gamma_k^2 \cdot W_k$$

Resolvendo $\hat{x}_{k+1/k}^2$ e subtraindo da expressão anterior:

$$x_{k+1/k} = a_2^2 x_{k/k} + (2a_3^2 + 6a_2a_4)x_{k/k}^2 + (15a_4^2 + 24a_3a_5 + 30a_2a_6)x_{k/k}^3 +$$

$$+ (96a_5^2 + 210a_4a_6)x_{k/k}^4 + 945a_6x_{k/k}^5 + \Gamma_k^2 W_k \quad 3.15$$

- Equações de previsão de medida

Por um procedimento análogo ao anterior, obtêm-se as seguintes expressões para os momentos de previsão de medida para $q=5$:

$$\hat{y}_{k+1/k} = b_1 + b_3 x_{k+1/k} + 3b_5 x_{k+1/k}^2 \quad 3.16$$

$$y_{k+1/k} = b_2^2 x_{k+1/k} + (2b_3^2 + 6b_2b_4)x_{k+1/k}^2 + (15b_4^2 + 24b_3b_5 + 30b_2b_6)x_{k+1/k}^3 +$$

$$+ (96b_5^2 + 210b_4b_6)x_{k+1/k}^4 + 945b_6x_{k+1/k}^5 + V_{k+1} \quad 3.17$$

- Equações de filtragem

Dado que as densidades de previsão de estado e de medida foram aproximadas por gaussianas, os momentos da densidade a posteriori obedecem às equações 2.13.

III.2.2 - Determinação dos coeficientes a_i e b_i

Não é de interesse para o problema de filtragem que as funções f e h sejam descritas por polinômios no intervalo $(-\infty, +\infty)$. Necessita-se descrever polinomialmente a função f apenas no intervalo em que a densidade de previsão de esta

do tem valores significativos e a função h apenas no intervalo em que a densidade de previsão de \underline{me} dida tem valores não desprezíveis. Dado que as densidades mencionadas são supostas gaussianas, é óbvio que o intervalo em que estas densidades têm valores significativos está diretamente relacionado com as respectivas variâncias. Os tamanhos dos intervalos serão portanto definidos como funções das respectivas variâncias. Como as variâncias va riam a cada passo da filtragem, necessita-se uma nova aproximação para cada passo. Portanto, a úni ca diferença entre o cálculo dos coeficientes a_i e b_i reside nos intervalos em que a aproximação é feita e por isto mesmo considera-se a seguir a penas o cálculo dos coeficientes a_i . Para facili tar a notação escreve-se simplificadamente $f(x)$, no lugar de $f(x_k, k)$. O polinômio que se aproxima de $f(x)$ será denotado por $P(x)$. O intervalo em que a aproximação deve ser feita, define-se como:

$$D = [\bar{x} - s\sigma, \bar{x} + s\sigma]$$

em que

\bar{x} : valor médio da densidade a posteriori do instante anterior

σ^2 : variância da referida densidade

s : número real que especifica o tamanho do intervalo

O intervalo D é subdividido num conjun to A de M pontos utilizados para obter os coe ficientes a_i do polinômio $P(x)$ de grau p que me lhor se aproxime de $f(x)$ neste intervalo.

$$P(x) = \sum_{i=1}^{p+1} a_i (x - \bar{x})^{i-1}$$

A literatura apresenta diversos métodos para a obtenção dos coeficientes, podendo todos eles serem expressos como a minimização de uma função de custo do tipo

$$J_M(\hat{a}) = \sum_{j=1}^M \epsilon^2(j, \hat{a})$$

em que $\epsilon(j, \hat{a})$ é o j -ésimo resíduo e \hat{a} é o vetor contendo os coeficientes desejados. A expressão explícita de $\epsilon(j, \hat{a})$ como função de \hat{a} difere segundo os diferentes métodos para calcular \hat{a} .

Um método frequentemente empregado é o LEAST SQUARES. Este método opera simultaneamente com todos os M pontos pertencentes ao conjunto A e a função de custo assume a seguinte forma

$$\epsilon(j, a) = f(x_j) - \sum_{i=1}^{p+1} a_i (x_j - \bar{x})^{i-1} \quad 3.18$$

com $x_j \in A$ e $j=1, 2, \dots, M$

O vetor de parâmetros \hat{a} é determinado de tal modo que a soma dos quadrados dos erros seja a menor possível. Os termos da expressão 3.18 podem ser reagrupados para permitir escrever esta expressão na forma vetorial.

$$\begin{array}{ccc} y = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_M) \end{bmatrix} & \phi = \begin{bmatrix} 1 & x_1 - \bar{x} & \dots & (x_1 - \bar{x})^p \\ 1 & x_2 - \bar{x} & \dots & (x_2 - \bar{x})^p \\ \dots & \dots & \dots & \dots \\ 1 & x_m - \bar{x} & \dots & (x_m - \bar{x})^p \end{bmatrix} & \epsilon = \begin{bmatrix} \epsilon_1 \\ \epsilon_2 \\ \vdots \\ \epsilon_M \end{bmatrix} \\ (M \times 1) & M \times (p+1) & (M \times 1) \end{array}$$

com $\epsilon_j = \epsilon(j, a)$ e $j = 1, 2, \dots, M$

A expressão 3.17 pode ser reescrita simplicadamente como:

$$y = \phi a + \epsilon$$

e a função de custo

$$\sum_{j=1}^M \epsilon^2(j, a) = \epsilon^T \epsilon$$

A minimização de $\epsilon^T \epsilon$ é fácil, resultando para o valor do vetor de parâmetros

$$\hat{a} = (\phi^T \phi)^{-1} \phi^T y \quad 3.19$$

O cálculo dos coeficientes pela expressão 3.18 requer a inversão da matriz $\phi^T \phi$ cuja dimensão é $(p+1) \times (p+1)$. Se o grau do polinômio for elevado, esta inversão necessita elevado tempo de cálculo e pode conduzir â erros consideraveis, principalmente se a matriz for mal condicionada. Para evitar estes problemas inerentes â inversão de matrizes, transforma-se a relação 3.19 numa forma recorrente. Ao invés de proceder a uma minimização considerando simultaneamente todos os M pontos do conjunto A , processa-se ponto por ponto sequencialmente. Com isto a inversão da matriz em 3.19 se reduz a inverter um valor escalar. A literatura descreve este método nomeando-o de Least Squares recursivo.

Neste trabalho, adota-se um método um pouco modificado. Aplicam-se as equações de filtragem de Kalman para o cálculo dos coeficientes. Para isto considera-se o vetor de coeficientes a descrito como um processo estacástico, cujas observações são os valores assumidos pela função f nos pontos pertencentes ao conjunto A . As observações são processadas sequencialmente aprimorando a estimativa do vetor de coeficientes.

Considera-se que os coeficientes evoluem estacionariamente, de acordo com a equação abaixo

$$a_{j+1} = a_j + e_j \quad 3.20$$

com as observações dadas por

$$y_j = \theta_j a_j + \eta_j \quad 3.21$$

com $j = 1, 2, \dots, M$

$$y_j = f(x_j) \quad e \quad x_j \in A$$

$$\theta_j = [1 \ x_j - \bar{x} \ \dots \ (x_j - \bar{x})^p] \quad j\text{-ésima linha da matriz } \theta$$

e_j e η_j são ruídos gaussianos, mutuamente independentes e independentes do valor inicial do vetor de parâmetros, com médias nulas e variâncias R_j e S_j respectivamente.

No caso de $R_j=0$, este método é idêntico ao Least Squares recursivo.

As equações de filtragem para o sistema descrito por 3.20 e 3.21 onde o vetor de coeficientes é assumido como o estado, são:

- Equações de previsão de estado (vetoriais):

$$\hat{a}_{j+1/j} = \hat{a}_{j/j}$$

$$A_{j+1/j} = A_{j/j} + R_j$$

- Equações de previsão de observação (escalares):

$$\hat{y}_{j+1/j} = \theta_{j+1} \cdot \hat{a}_{j+1/j}$$

$$Y_{j+1/j} = \theta_{j+1} \cdot A_{j+1/j} \cdot \theta_{j+1}^T + S_{j+1}$$

- Equações de filtragem (vetoriais):

$$\hat{a}_{j+1/j+1} = \hat{a}_{j+1/j} + K_{j+1} (y_{j+1} - \hat{y}_{j+1/j})$$

$$A_{j+1/j+1} = (I - K_{j+1} \theta_{j+1}^T) A_{j+1/j}$$

$$K_{j+1} = A_{j/j} \theta_j^T Y_{j+1/j}^{-1} \quad (\text{inversão de um escalar})$$

Resta ainda discutir as condições para a inicialização do filtro. Muito pouco pode ser

conhecido a priori sobre o valor do vetor de coeficientes a . Por isto, o filtro é inicializado assumindo que o vetor inicial do estado é gaussiano com média nula e variância infinita. Para tornar praticável o cálculo da matriz A , escolhe-se $A_0 = 1/\epsilon I$ com $\epsilon > 0$, e o vetor de coeficientes procurado será:

$$\hat{a} = \hat{a}_{M/M}$$

III.2.3 - Determinação do tamanho do intervalo D

Resta ainda para especificar o fator s que determina a magnitude da intervalo D . Poder-se-ia, à primeira vista, definir o intervalo D como sendo o domínio em que a densidade de previsão assume valores não desprezíveis, isto é, $D_e = \chi_k$ para a previsão de estado e $D_s = \chi'_{k+1}$ para a previsão de medida, o que corresponde a escolher o mesmo valor de s para os dois casos citados, ou que se consideram como desprezíveis os valores da densidade cuja relação com o máximo é menor do que $e^{-s^2/2}$. Assim, por exemplo, para $s=3$, seriam desprezíveis os valores das densidades menores do que um centésimo do seu máximo.

No entanto, o critério apresenta bons resultados quando adotado para a aproximação da não linearidade da medida, o mesmo não acontecendo para a aproximação da não linearidade do estado. Neste último caso, o fator s deve ser menor. A explicação deste diferente comportamento está diretamente relacionada com o que foi discutido na seção 1.1 deste Capítulo. Basta acrescentar aqui que um fator s pequeno implica que o domínio D_e é relativamente pequeno fazendo com que o polinômio se aproxime com maior precisão da não linearidade do estado justamente onde a densidade a posteriori do instante anterior exerce a maior ponderação sobre as "funções de verossimilhança" para o cálculo da densidade de previsão (Fig. 3.6). A influência da não linearidade de medida se faz sen

tir através da função de verossimilhança $L_1(x_{k+1}:y_{obs})$, que multiplica a densidade de previsão do estado para chegar à densidade a posteriori. Nestas operações, não há nenhuma espécie de ponderação semelhante ao caso da não linearidade do estado e o que verdadeiramente importa aqui é que a função de verossimilhança resultante da aproximação descreva adequadamente a real, e é justamente por isto que é válida a consideração inicial.

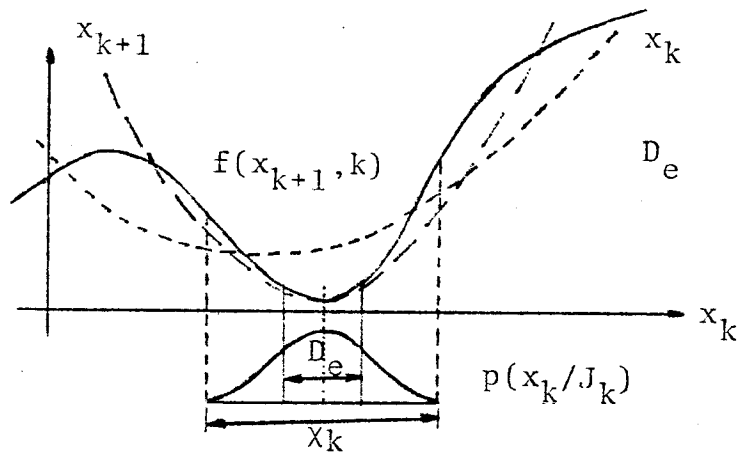


Fig. 3.6 - Influência do intervalo D na aproximação polinomial

III.2.4 - Filtro polinomial simplificado : FS

A determinação dos coeficientes a_i e b_i através da minimização de uma função de custo especificada, demanda um tempo de computação suficientemente grande para fazer com que o filtro polinomial fitado tenha uma velocidade de processamento inferior a dos filtros generalizados. Isto limita a utilização prática deste algoritmo para os casos em que se pode pagar uma maior precisão às custas de tempo de cálculo. No entanto, pode-se reduzir o tempo de processamento se os coeficientes a_i e b_i forem calculados de um modo mais simples. Ao invés de proceder a uma minimização de função de custo, os coeficientes são obtidos especificando-se $p+1$ abscissas do intervalo D, e fazendo passar pelas respectivas ordenadas o polinômio $P(x)$ de grau p . O cálculo dos coeficientes se reduz à inversão de uma matriz quadrada de ordem $p+1$.

Considerando a aproximação da função f ,

o problema pode ser formulado da seguinte forma:

- Determinar os coeficientes a_i do polinômio

$$P(x) = \sum_{i=1}^{p+1} a_i (x - \bar{x})^{i-1}$$

sujeito às restrições

$$P(x_j) = \sum_{i=1}^{p+1} a_i (x_j - \bar{x})^{i-1} = f(x_j)$$

para $j = 1, \dots, p+1$ e $x_j \in D$

Esta última expressão pode ser descrita na forma matricial mediante as seguintes definições:

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_{p+1} \end{bmatrix} = \begin{bmatrix} f(x_1) \\ f(x_2) \\ \vdots \\ f(x_{p+1}) \end{bmatrix} \quad a = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_{p+1} \end{bmatrix} \quad \Omega = \begin{bmatrix} 1 & (x_1 - \bar{x}) & \dots & (x_1 - \bar{x})^p \\ 1 & (x_2 - \bar{x}) & \dots & (x_2 - \bar{x})^p \\ \cdot & \dots & \dots & \dots \\ 1 & (x_{p+1} - \bar{x}) & \dots & (x_{p+1} - \bar{x})^{p+1} \end{bmatrix}$$

Então $y = \Omega a$ cuja solução é:

$$a = \Omega^{-1} y$$

A inversão da matriz Ω não oferece dificuldades pois é uma matriz de Vandermond.

As considerações anteriormente feitas sobre a magnitude do intervalo são válidas para este caso também.

III.3 - FILTROS ITERADOS

Pelo fato das aproximações até primeira ordem incorrerem em pesados erros de linearização, introduziram-se os termos de segunda ordem. Mas o cálculo das derivadas de segunda ordem pode ser, às vezes, tarefa que consome muito tempo de computação (Vide Bruno, Cap.IV). Para corrigir er

ros devidos à linearizações sem recorrer às derivadas de segunda ordem foi desenvolvido um procedimento iterativo, que corresponde à linearizações em torno de valores mais próximos dos verdadeiros (13,16). O valor filtrado no instante $K+1$ está em média mais próximo do valor real do estado do que o valor predito, pois contém a informação trazida pela nova medida. Este valor poderá então ser utilizado para:

- 1) Recalcular o valor filtrado do instante anterior - alissagem; relinearizar as equações de estado para obter um valor estimado atual mais preciso.
- 2) Relinearizar as equações de medida em torno deste valor, para aprimorar o valor estimado atual - iteração local.

O processo pode ser repetido até que a alteração do valor estimado atual seja desprezível.

III.3.1 - Filtro iterado com alissagem : IS

O desenvolvimento deste algoritmo pode ser encontrado em Tozzi (16). As respectivas equações são as que seguem:

$$\hat{x}_{k/k+1}^1 = \hat{x}_{k/k}$$

$$X_{k/k+1}^1 = X_{k/k}$$

3.22

$$\hat{x}_{k+1/k}^j = f(\hat{x}_{k/k}^j, k)$$

$$X_{k+1/k}^j = F_X^j \cdot X_{k/k+1}^j \cdot F_X^{jT} + \Gamma_k W_k \Gamma_k^T$$

$$\hat{y}_{k+1/k}^j = h(\hat{x}_{k+1/k}^j)$$

$$Y_{k+1/k}^j = H_X^j \cdot X_{k+1/k}^j \cdot H_X^{jT} + V_{k+1}$$

$$\hat{x}_{k+1/k+1}^j = \hat{x}_{k+1/k}^j + K_{k+1}^j (y_{obs} - \hat{y}_{k+1/k}^j)$$

$$X_{k+1/k+1}^j = (I - K_{k+1}^j H_X^j) X_{k+1/k}^j$$

$$K_{k+1}^j = X_{k+1/k}^j \cdot H_X^{jT} \cdot (Y_{k+1/k}^j)^{-1}$$

$$\hat{x}_{k/k+1}^{j+1} = \hat{x}_{k/k} + K_{sk}^j (\hat{x}_{k+1/k+1}^j - \hat{x}_{k+1/k}^j)$$

$$X_{k/k+1}^{j+1} = X_{k/k} + K_{sk}^j (X_{k+1/k+1}^j - X_{k+1/k}^j) K_{sk}^{jT}$$

para $j=1,2,\dots,M$ representa a j -ésima iteração do instante $K+1$ e

$$\hat{x}_{k+1/k+1} = \hat{x}_{k+1/k+1}^M \quad \text{e} \quad X_{k+1/k+1} = X_{k+1/k+1}^M$$

$$\text{com } F_X^j = \left. \frac{\partial f(x_k, k)}{\partial x_k} \right|_{x_k = \hat{x}_{k/k+1}^j} \quad H_X^j = \left. \frac{\partial h(x_{k+1}, k+1)}{\partial x_{k+1}} \right|_{x_{k+1} = \hat{x}_{k+1/k}^j}$$

A finalidade deste iterador é corrigir os erros de linearização da equação de estado. A seguir, apresentam-se as equações do algoritmo com iterações locais.

III.3.2 - Filtro com iterações locais : IL

A finalidade destas iterações é corrigir erros de linearização da equação de medidas. As equações de predição de estado são as mesmas do filtro estendido (3.5). As equações de predição de medida e de filtragem obedecem ao procedimento iterativo abaixo.

$$\mu_1 = \hat{x}_{k+1/k}$$

$$\hat{y}_{k+1/k}^i = h(\mu_i) + H_X^i \cdot (\mu_1 - \mu_i)$$

$$Y_{k+1/k}^i = H_X^i \cdot X_{k+1/k} \cdot H_X^{iT} + V_{k+1}$$

$$\mu_{i+1} = \mu_1 + K_{k+1}^i \cdot (y_{\text{obs}} - \hat{y}_{k+1/k}^i)$$

$$K_{k+1}^i = X_{k+1/k} H_X^{iT} (Y_{k+1/k}^i)^{-1}$$

$$\hat{x}_{k+1/k+1} = \mu_N$$

$$X_{k+1/k+1} = (I - K_{k+1}^N H_X^N) X_{k+1/k}$$

$$\text{com } H_X^i = \left. \frac{\partial h(x_{k+1}, k+1)}{\partial x_{k+1}} \right|_{x_{k+1} = \mu_i}$$

e $i=1,2,\dots,N$ denota a i -ésima iteração.

A literatura se refere ao filtro com iterações locais como "excelente para levar em conta a não linearidade da função de medidas" (9). Por isto, é interessante compreender a mecânica de seu funcionamento, o que pode ser conseguido pelas equações 3.23. Partindo do valor médio da densidade de privisão de estado μ_1 , através de um passo idêntico ao do filtro estendido, obtém-se o valor médio da densidade a posteriori μ_2 . Com este valor torna-se a linearizar a equação de medidas para obter em seguida o novo valor médio filtrado (Fig.3.7). O processo é repetido até que $|\mu_{i+1} - \mu_1| < \epsilon$ onde ϵ é uma quantidade arbitrariamente especificada.

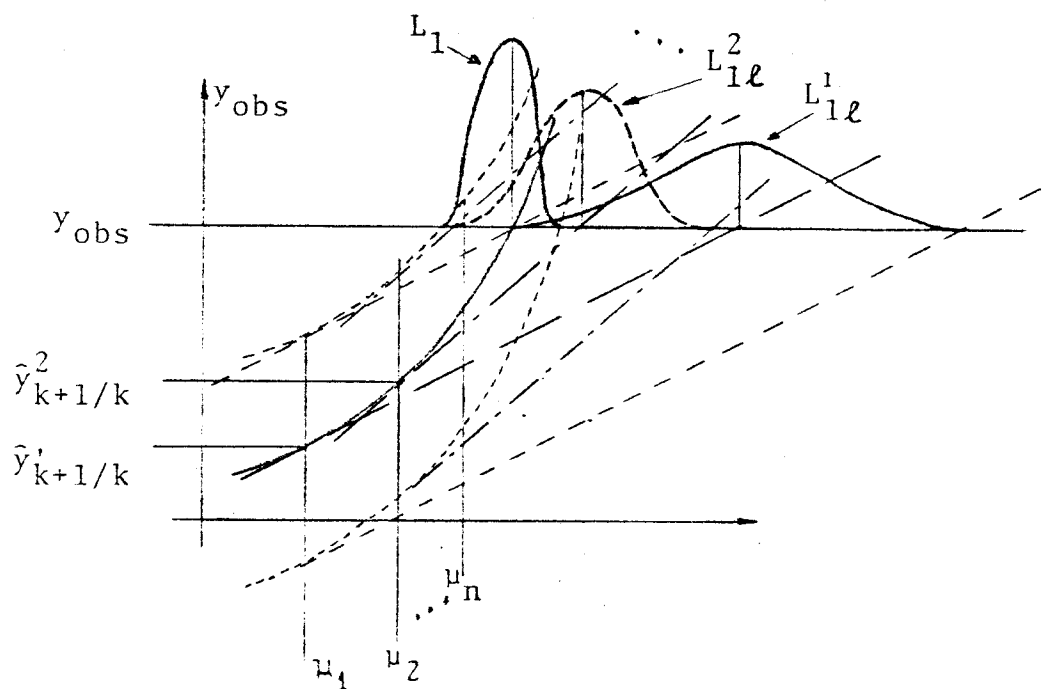


Fig. 3.7 - Iterações Locais

L_i^1 : função de verossimilhança linearizada

Admitindo que o iterador seja convergente, afigura-se como óbvio que a medida que as iterações se sucedem, a densidade a posteriori é recalculada a partir da densidade de previsão de estado, mantida constante, multiplicada pela função de verossimilhança linearizada $L_{1\ell}^i$ que a cada iteração está mais próxima da verdadeira L_1 . Isto é a diferença absoluta entre o valor médio de $L_{1\ell}^i$ e a moda de L_1 diminui a medida que o número de iterações aumenta. No entanto, mesmo que se aumente arbitrariamente o número de iterações N , em geral, $L_{1\ell}^N$ ainda estará afastada de L_1 , no seguinte sentido:

3.24

$$\lim_{N \rightarrow \infty} \frac{\int x_{k+1} L_{1\ell}^N(x_{k+1}:y_{obs}) dx_{k+1}}{\int L_{1\ell}^N(x_{k+1}:y_{obs}) dx_{k+1}} \neq \text{moda } L_1(x_{k+1}:y_{obs})$$

Para que a desigualdade 3.24 possa se converter em igualdade é necessário que $\mu_{i+1} = \mu_i = C_{k+1}$ onde $C_{k+1} = h^{-1}(y_{obs}, k+1)$. Mas para isto é necessário que ou $\mu_1 = C_{k+1}$ e neste caso nenhuma iteração poderá introduzir alguma melhora, ou o filtro ignore completamente a informação contida na densidade de previsão de estado, o que só poderia acontecer se a função de verossimilhança L_1 fosse uma função impulso.

No entanto, a expressão 3.24 pode ser transformada em igualdade, substituindo o algoritmo de iterações locais por um procedimento modificado. Se é intuito do iterador local aproximar $L_{1\ell}^N$ a L_1 , isto pode ser conseguido diretamente determinando analítica ou numericamente o ponto C_{k+1} solução de $y_{obs} = h(x_{k+1}, k+1)$. Com isto se evita procedimento iterativo que requer em cada ciclo a atualização do valor estimado do estado e do valor do ganho de Kalman - cálculos intermediários laboriosos. É deste algoritmo que trata a próxima seção.

III.3.3 - Filtro estendido modificado : EM

O filtro proposto nesta seção tem por finalidade evitar os cálculos intermediários do iterador local para a redução de erros de aproximação de primeira ordem. A linearização da equação de medidas já não é feita em torno do valor previsto do estado, mas do ponto solução para $y_{k+1}=y_{obs}$. Assim, a função de verossimilhança $L_1(x_{k+1}:y_{obs})$ é aproximada por uma forma gaussiana cuja média é a moda de L_1 . Com certeza, esta aproximação é mais adequada do que a do filtro estendido, e mesmo do filtro com iterações locais, pelo fato da função L_{1l} constituir uma melhor aproximação para L_1 .

A densidade de previsão de estado é calculada de acordo com qualquer um dos métodos anteriores. A densidade de previsão de medida é calculada considerando que a equação de medidas é linearizada por desenvolvimento em série de Taylor em torno do ponto C_{k+1} .

$$C_{k+1} = h^{-1}(y_{k+1} = y_{obs}, k+1)$$

$$h(x_{k+1}, k+1) = h(C_{k+1}, k+1) + H_x \cdot (x_{k+1} - C_{k+1}) + \dots$$

$$\text{com } H_x = \left. \frac{\partial h(x_{k+1}, k)}{\partial x_{k+1}} \right|_{x_{k+1} = C_{k+1}}$$

As equações de previsão de medida se tornam:

$$\hat{y}_{k+1/k}^{\Delta} = E \{ y_{k+1} / J_k \}$$

$$\hat{y}_{k+1/k} = E \{ h(C_{k+1}, k+1) + H_x \cdot (x_{k+1} - C_{k+1}) + v_{k+1} / J_k \}$$

$$\hat{y}_{k+1/k} = h(C_{k+1}, k+1) + H_x \cdot (\hat{x}_{k+1/k} - C_{k+1}) \quad 3.25$$

$$Y_{k+1/k}^{\Delta} = E \{ (y_{k+1} - \hat{y}_{k+1/k}) (y_{k+1} - \hat{y}_{k+1/k})^T / J_k \}$$

$$\begin{aligned}
&= E \{ H_X \cdot (x_{k+1} - \hat{x}_{k+1/k}) + v_{k+1} \quad H_X \cdot (x_{k+1} - \hat{x}_{k+1/k}) + v_{k+1} \quad / J_{k+1}^T \} \\
&= H_X^T E \{ (x_{k+1} - \hat{x}_{k+1/k}) (x_{k+1} - \hat{x}_{k+1/k})^T / J_k \} H_X + V_{k+1} \\
Y_{k+1/k} &= H_X \cdot X_{k+1/k} \cdot H_X^T + V_{k+1} \quad 3.26
\end{aligned}$$

A densidade a posteriori é obtida através das equações de Kalman 2.13 com a densidade de previsão de medida dada por 3.25 e 3.26.

Algumas observações são necessárias neste momento. Em primeiro lugar, este estimador, bem como os seus precedentes iterados com alissagem ou locais, são estimadores da moda e por isto mesmo, em geral, são polarizados. Em segundo lugar, o extendido modificado necessita calcular a função inversa $h^{-1}(y_{obs,k+1})$ para determinar o ponto em torno do qual se fará a linearização. Caso h seja uma função biunívoca, a determinação de C_{k+1} , analítica ou por algum processo iterativo, é mais rápida, pois não necessita calcular a cada iteração a matriz de ganho K_{k+1}^i evitando com isto a inversão da matriz de covariância da densidade de previsão da medida. As iterações são feitas direta e exclusivamente sobre a função h . Quando h não for biunívoca, é necessário recorrer a algum critério de escolha entre as possíveis soluções alternativas. A questão recai no que foi anteriormente discutido (seção III.1.1): escolhe-se a solução que esteja mais próxima do valor previsto do estado, isto é, escolhe-se C_{k+1} tal que $C_{k+1} \in \chi'_{k+1}$, a menos que no domínio χ'_{k+1} a função de verossimilhança L_1 tenha mais do que uma moda e neste caso não se pode esperar comportamento satisfatório de nenhum filtro linearizado.

Outro particular surge quando $y_{k+1} = h(x_{k+1,k+1})$ não tem solução para $y_{k+1} = y_{obs}$. Tal caso pode acontecer se a função h apresentar pon

tos de derivada nula no domínio x'_{k+1} . Nestas situações, os filtros extendido e iterados se comportam, na melhor das hipóteses, quase exclusivamente como preditores, como pode ser visto pela expressão e gráfico abaixo.

$$K_{k+1} = X_{k+1/k} H_X^T Y_{k+1/k}^{-1} \approx 0 \text{ pois } H_X = \left. \frac{\partial h}{\partial x_{k+1}} \right|_{x_{k+1} = \hat{x}_{k+1/k}} \approx 0$$

Deste modo:

$$\hat{x}_{k+1/k+1} \approx \hat{x}_{k+1/k}$$

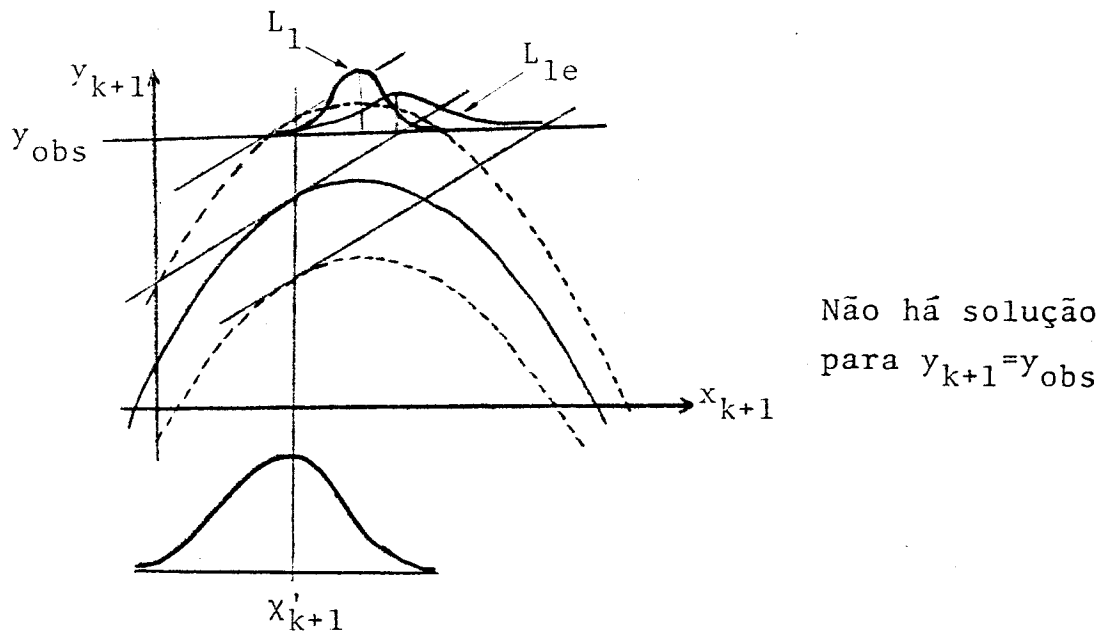


Fig. 3.8

Em outras palavras, estes filtros não absorvem a informação trazida pela nova medida nestas condições. Obviamente isto é pobreza de comportamento, pois mesmo que $H_X = 0$, dada uma observação y_{obs} , a função de verossimilhança $L_1(x_{k+1} : y_{obs})$ continua perfeitamente definida. O que se propõe para o filtro extendido modificado nestas condições é a utilização direta da função de verossimilhança L_1 . Mas este assunto é o tema da próxima seção.

III.4 - FILTROS GAUSSIANOS POR APROXIMAÇÃO DIRETA DA DENSIDADE A POSTERIORI

Os diversos desdobramentos da solução do problema da filtragem não linear apresentados pela literatura situam-se em dois extremos. De um lado, a recorrência numérica sobre a própria densidade a posteriori via regra de Bayes (2.4). O Capítulo II descreve uma possível aproximação deste tipo na qual não são necessárias hipóteses explícitas a respeito das densidades. No entanto, um dos seus passos para a obtenção da densidade a posteriori, o cálculo da convolução não linear (expressão 2.7) limita severamente a aplicação prática deste algoritmo por necessitar de uma integração numérica para cada valor da densidade de previsão de estado em certo instante. De outro lado, situam-se as soluções aproximadas, cuja hipótese central é a de que todas as densidades envolvidas pela regra de Bayes são gaussianas. Tal hipótese é perfeitamente válida para sistemas lineares com densidades a priori e de ruídos gaussianos e neste caso, a recorrência pode ser feita diretamente sobre a média e covariância das densidades do estado conduzindo ao conhecido filtro de Kalman (2.11 - 2.13). Para o caso não linear, a hipótese é verificada mediante linearização das não linearidades, permitindo com isto que a solução retenha a estrutura simples do filtro de Kalman linear. A linearização é feita por desenvolvimento das funções em série de Taylor.

Porém, o desenvolvimento em série de Taylor, desde que limite o número de termos, só representará localmente as não linearidades e a aproximação gaussiana resultante poderá ser suficientemente inadequada para resultar em divergência do filtro. Em vista disto, a seção III.2 propõe que as não linearidades sejam aproximadas num certo intervalo e não apenas a partir de características locais. Para corrigir erros de linearização, a literatura descreve os procedimentos iterados. A iteração tem por finalidade aproximar a média da função de verossimilhança linearizada gaussiana da moda da função de verossimilhança L_1 e o faz calculando valores intermediários dispensáveis. A seção precedente propôs um método alternativo, que dispensa estes cálculos. Além

do mais, pode-se dar uma interpretação do método: a linearização já não se apoia no valor previsto de estado, mas no novo valor da observação. No entanto, os filtros iterados e o modificado são polarizados por serem estimadores da moda. Uma correção de viés pode ser introduzida, considerando os termos de segunda ordem.

A inviabilidade prática do filtro ótimo, repita-se, é devida primordialmente ao alto tempo demandado para realizar a convolução não linear. Para os filtros linearizados, ao contrário, o ponto crucial está na aproximação da equação de medidas, não porque exija muito cálculo, mas porque uma aproximação inadequada pode facilmente conduzir o filtro à divergência. A transformação da convolução não linear em linear, mediante aproximação da função f reduz enormemente o tempo de cálculo e a densidade a posteriori é menos sensível a erros de aproximação da não linearidade de estado do que a erros de aproximação da não linearidade de medida. Portanto, para melhorar o desempenho dos algoritmos de filtragem será necessário investir tempo de cálculo para aproximar mais adequadamente a função de verossimilhança L_1 , evitando aproximações da função de medidas.

A linearização da função h , em geral implica em estimar indevidamente a informação trazida por uma nova medida. Esta informação está totalmente contida na função L_1 , que é analiticamente conhecida. Propõe-se, então, abandonar a linearização de h e calcular a densidade a posteriori, realizando e normalizando o produto da função L_1 com a densidade de previsão do estado que será suposta gaussiana em todos os instantes e descrita por:

$$p(x_{k+1}/J_k) = N(\hat{x}_{k+1/k}, X_{k+1/k}) = \text{CTE} \cdot \exp\left(-\frac{1}{2} \frac{(x_{k+1} - \hat{x}_{k+1/k})^2}{X_{k+1/k}}\right)$$

$\hat{x}_{k+1/k}$ e $X_{k+1/k}$ são calculados por algumas das expressões apresentadas em seções precedentes.

Toda a informação a respeito do estado obtida por uma nova medida, está na função de verossimilhança L_1 , que embora não tendo forma gaussiana, sua expressão analítica

é perfeitamente conhecida:

$$L_1(x_{k+1}:y_{obs}) = p(y_{k+1} = y_{obs}/x_{k+1})$$

$$L_1(x_{k+1}:y_{obs}) = CTE \cdot \exp\left(-0,5 // y_{obs}^{-h(x_{k+1}, k+1)} // V_{k+1}^{-1}\right)$$

A densidade a posteriori será:

$$p(x_{k+1}/J_{k+1}) = \frac{p(y_{k+1}/x_{k+1}) \cdot p(x_{k+1}/J_k)}{p(y_{k+1}/J_k)}$$

$$p(x_{k+1}/J_k) = \frac{L_1(x_{k+1}:y_{obs}) \cdot N(\hat{x}_{k+1/k}, X_{k+1/k})}{p(y_{k+1}/J_k)} \quad 3.27$$

Esta densidade não é gaussiana. No entanto, é objetivo do procedimento que está sendo proposto, manter a densidade a posteriori gaussiana. Isto pode ser realizado em dois níveis:

- 1) aproximando a densidade a posteriori que resulta da relação 3.27 através do cálculo de seus momentos.
- 2) aproximando a função L_1 por uma gaussiana realizando depois o produto.

Entre os dois níveis, o que melhor se aproxima do ótimo é, evidentemente, o primeiro, porque a aproximação só é realizada após o último estágio de cálculo, isto é, depois que a densidade a posteriori em certo instante foi calculada. No entanto, a aproximação gaussiana neste nível só pode ser realizada com o filtro em operação ("ON LINE"). No segundo nível, a aproximação da função L_1 pode ser feita, mediante certas hipóteses, a priori ("OFF LINE"), acarretando um aumento da velocidade de processamento do filtro.

Mas, densidades a posteriori gaussianas pressupõem linearidade na evolução do estado e das medidas. Então, em última instância, também estes filtros correspondem a uma linearização da equação de medidas; mas com certeza esta linearização

rização não dependerá mais exclusivamente da densidade de previsão de estado senão que dependerá também do próprio valor das observações em cada instante.

III.4.1 - Filtro gaussiano por aproximação da densidade a posteriori : PT

A densidade a posteriori é obtida calculando a expressão 3.27 sem recorrer a nenhuma aproximação para a função de verossimilhança L_1 . A densidade a posteriori será dada por:

$$p(x_{k+1}/J_k) = \frac{\exp\left(-0,5/y_{obs} - h(x_{k+1}, k+1)/V_{k+1} - 0,5/x_{k+1} - \hat{x}_{k+1}/k/X_{k+1/l}\right)^2}{\int [\text{numerador}] dx_{k+1}}$$

Os momentos serão calculados por:

$$\hat{x}_{k+1/k+1} = \int x_{k+1} \cdot p(x_{k+1}/J_{k+1}) dx_{k+1}$$

$$X_{k+1/k+1} = \int (x_{k+1} - \hat{x}_{k+1/k+1}) \cdot (x_{k+1} - \hat{x}_{k+1/k+1})^T \cdot p(x_{k+1}/J_{k+1}) dx_{k+1}$$

A densidade a posteriori é aproximada por uma gaussiana definida por aqueles dois momentos:

$$p(x_{k+1}/J_{k+1}) = N(\hat{x}_{k+1/k+1}, X_{k+1/k+1})$$

O procedimento recém descrito introduz na densidade a posteriori toda a informação disponível numa nova observação. Para os instantes subsequentes, uma parcela da informação é perdida evidentemente pelo fato de ignorar os momentos de ordem superior a dois.

III.4.2 - Filtro gaussiano por aproximação da função de verossimilhança : PS

A função de verossimilhança L_1 é aproxi

mada por uma gaussiana, mas sem recorrer à linearização da função de medidas. Ao invés disto, calculam-se os dois primeiros momentos de L_1 ; a gaussiana aproximada é aquela definida por estes dois momentos:

$$L_{1g}(x_{k+1}:y_{obs}) = N(x_{k+1}^*, x_{k+1}^*)$$

com

$$x_{k+1}^* = \frac{1}{c} \int_{x_{k+1}^*} x_{k+1} L_1(x_{k+1}:y_{obs}) dx_{k+1}$$

$$x_{k+1}^* = \frac{1}{c} \int_{x_{k+1}^*} (x_{k+1} - x_{k+1}^*)(x_{k+1} - x_{k+1}^*)^T L_1(x_{k+1}:y_{obs}) dx_{k+1}$$

$$c = \int_{x_{k+1}^*} L_1(x_{k+1}:y_{obs}) dx_{k+1}$$

3.28

Para esta aproximação, é necessário restringir o domínio de integração que já não pode ser todo o espaço R^n , pois isto levaria o filtro a atribuir igual ponderação a todas as modas, caso L_1 seja multimodal, quando a ponderação deve ser distinta. O domínio pode ser definido de acordo com o domínio de valores significativos da densidade de previsão de estado. Um caso particular é ilustrado pela Fig. 3.9.

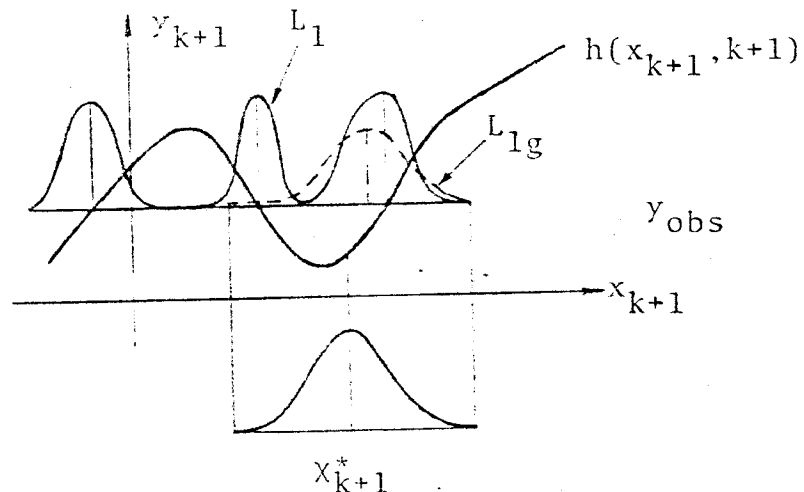


Fig. 3.9 - Aproximação da função de verossimilhança L_1 por uma gaussiana L_{1g} .

A densidade a posteriori se obtém realizando o produto 3.27 que é agora um produto de duas gaussianas, cujo resultado é:

$$\hat{x}_{k+1/k+1} = (X_{k+1/k} \cdot x_{k+1}^* + X_{k+1}^* \cdot \hat{x}_{k+1/k}) (X_{k+1/k+1} + X_{k+1}^*)^{-1}$$

$$X_{k+1/k+1} = X_{k+1}^* \cdot X_{k+1/k} (X_{k+1}^* + X_{k+1/k})^{-1}$$

III.4.3 - Filtro gaussiano por aproximação a priori da função de verossimilhança : MP

A aproximação precedente não apresenta vantagem em termos de velocidade de processamento sobre o da seção III.4.1. Ambas necessitam de três integrações numéricas: uma para calcular a constante de normalização e as outras duas para calcular os dois primeiros momentos.

Não há dúvida de que, com relação ao filtro ótimo, as duas aproximações acima requerem uma quantidade de cálculo extraordinariamente menor, porém, quando comparados com os filtros das seções 4.1 e 4.3, são de execução lenta.

No entanto, a solução proposta em III.4.2 admite uma variante capaz de solucionar o problema da execução lenta. Basta para tanto, que as funções de verossimilhança L_1 sejam aproximadas a priori por gaussianas. De que forma pode ser isto realizado?

Evidentemente, é necessário conhecer aproximadamente o domínio de valores das observações. Para um certo conjunto de pontos deste domínio, calculam-se os valores médios e covariâncias de L_1 de acordo com 3.28, em que é necessário redefinir o domínio de integração, pois x_{k+1}^* não pode ser conhecido a priori por uma função da densidade de previsão de estado. Diversos casos podem ocorrer:

- a) A função de verossimilhança L_1 tem apenas uma moda (ou é possível desprezar as outras): o domínio de integração será o domínio em que a função de verossimilhança assumir valores significativos.
- b) A função de verossimilhança L_1 é multimodal: dado que o domínio de valores significativos da densidade de previsão de estado não é conhecido a priori e considerando que nele a função L_1 pode possuir predominantemente qualquer uma das modas ou uma combinação de modas adjacentes, é conveniente uma aproximação gaussiana para cada uma das alternativas possíveis. O domínio de integração será definido de maneira a incluir a região pertinente a uma moda ou a uma combinação de modas adjacentes, como mostra a Fig. 3.10. (Por região pertinente, entende-se uma vizinhança da moda na qual a função L_1 assume valores significativos.)

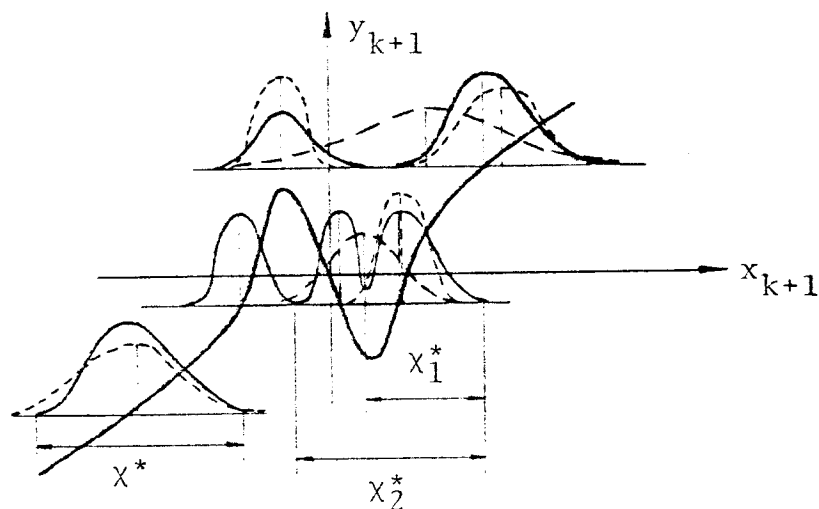


Fig. 3.10

Aproximações alternativas da função de verossimilhança L_1

.... considerando apenas uma moda

---- considerando duas modas

Para cada valor particular y^i da observação, a função de verossimilhança $L_1(x:y^i)$ é aproximada por uma gaussiana em cada um dos possíveis domínios x_j^* ($j = 1, \dots, p$ onde p é determinado pelo número de modas). Organiza-se uma tabela constando de uma linha para cada valor particular y^i e de $p+1$ colunas, contendo o valor y^i , os valores médios e respectivas matrizes de covariância da função $L_1(x,y^i)$ nos diversos domínios x_j^* .

Esta tabela construída a priori é fornecida ao algoritmo que a cada nova medida realizada, procede da seguinte maneira: com o valor observado atual determina a linha da tabela que mais se aproxima de y_{obs} e com o domínio x'_{k+1} da densidade de previsão de estado determina a coluna cujo conteúdo represente o valor médio e a matriz de covariância de L_1 , calculados no domínio x_j^* que inclua as modas de L_1 pertencentes ao domínio x'_{k+1} .

A gaussiana definida por aquela posição da tabela será:

$$L_{1g}(x_{k+1} : y_{obs}) = N(x_{k+1}^* : X_{k+1}^*)$$

Se a discretização do domínio de valores do vetor de medidas para construir a tabela tiver sido grosseira, a gaussiana acima pode não representar suficientemente bem a função L_1 . A aproximação pode ser refinada recorrendo ao conteúdo das posições adjacentes àquela para realizar uma interpolação. A gaussiana interpolada será:

$$L_{1g}(x_{k+1}:y_{obs}) = N(\bar{x}_{k+1}, \bar{X}_{k+1})$$

A densidade a posteriori será definida por:

$$\hat{x}_{k+1/k+1} = (X_{k+1/k} - \bar{x}_{k+1} + \bar{X}_{k+1} \cdot \hat{x}_{k+1/k}) \cdot (X_{k+1/k} + \bar{X}_{k+1})^{-1}$$

$$X_{k+1/k+1} = \bar{X}_{k+1} \cdot X_{k+1/k} \cdot (X_{k+1/k} - \bar{X}_{k+1})^{-1}$$

O procedimento acima descrito, pode ser empregado juntamente com o filtro extendido modificado nos casos em que $y_{k+1}=h(x_{k+1},k+1)$ não tenha solução para $y_{k+1}=y_{obs}$ ou em que a função de verossimilhança L_1 possuir mais do que uma moda no domínio x'_{k+1} .

CAPÍTULO IV

RESULTADOS NUMÉRICOS

Neste Capítulo serão apresentados alguns dos resultados obtidos utilizando os algoritmos de filtragem desenvolvidos no Capítulo anterior. Foram escolhidos três exemplos que simulados permitem evidenciar as vantagens de cada um dos algoritmos propostos. A comparação entre os diversos algoritmos tem como base resultados obtidos a partir de simulações de Monte Carlo com médias sobre 700 realizações. Para fins de comparação observam-se as seguintes variáveis (M=700):

- Erro médio:

$$\bar{\epsilon}_k = \frac{1}{M} \sum_{i=1}^M \epsilon_k^i \quad \text{e} \quad \epsilon_k^i = x_k^i - \hat{x}_{k/k}^i$$

Esta variável é um indicador de polarização, pois a medida que o número de simulações cresce, o erro médio deve tender a zero. O filtro estará polarizado se o erro médio assumir tendencialmente só valores positivos ou só valores negativos.

- Desvio padrão médio:

$$\bar{\sigma}_k = \left\{ \frac{1}{M} \sum_{i=1}^M x_{k/k}^i \right\}^{1/2}$$

- Relação AK:

$$AK_k = \left\{ \frac{1}{M} \sum_{i=1}^M \frac{(\epsilon_k^i)^2}{x_{k/k}^i} \right\}^{1/2}$$

A relação AK, utilizada por Alsparch and Sorenson (1), é um indicador de divergência. Para o filtro de Kalman operando sobre um sistema linear com perturbações gaussianas seu valor fica em torno de 1. Se o valor de AK se tornar muito maior do que 1, diz-se que o filtro divergiu, vale dizer, a variância calculada diminui excessivamente enquanto o erro médio quadrático aumentou. Se AK se tornar muito menor do que 1, significa que o filtro aumentou demasiadamente a variância, apresentando um caráter conservativo.

Considera-se o mesmo exemplo da seção 6 do Capítulo II, cujas equações e propriedades estatísticas estão repetidas abaixo:

$$x_{k+1} = 0,9x_k - \frac{10 x_k}{1+x_k} - 0,1 k + w_k \quad (x_0=20, k=0,1,\dots,90)$$

$$y_k = 0,5 x_k - \frac{10(x_k + 40)}{1+(x_k+40)^2} + v_k$$

Os ruídos são brancos, gaussianos, mtualmente independentes e independentes do estado inicial.

- condição inicial do estado: $N(15,10)$
- ruído de estado: $N(0,1)$
- ruído de medidas: $N(0,1)$

Os resultados apresentados em II.6 mostram que densidade a posteriori é frequentemente multimodal ou fortemente assimétrica quando o estado ou as medidas se encontram em regiões de não linearidades acentuadas (entre os instantes $k=8$ e $k=25$ para o estado e entre $k=49$ e $k=62$ para as medidas. Pode-se esperar, então, que, nestes instantes apenas os filtros que tomem em conta o caráter global das não linearidades tenham um desempenho satisfatório. Nos outros instantes as funções f e h são aproximadamente lineares, as densidades a posteriori são bastante próximas de gaussianas, esperando-se por isto que os filtros aproximados tenham bom desempenho.

FILTROS EX, IL, EM

A Fig. 4.1 representa os resultados das simulações com os filtros extendido (EX), iterado com 4 iterações locais (IL) e extendido modificado (EM). O filtro extendido modificado necessita de algumas observações. Dado que a função h não é biunívoca, podem ocorrer casos em que a função inversa $h^{-1}(y_{obs,k+1})$ tenha apenas uma solução ou tenha duas ou mesmo três soluções. Quando a função inversa h^{-1} tiver três soluções escolhe-se a que se encontrar mais próxima do valor previsto

do estado. A função h^{-1} tem duas soluções, se o valor y_{obs} for um dos dois valores extremos de h . Se o estado previsto estiver mais próximo da solução que é ponto extremo, é conveniente adotar um algoritmo diferente do EM (pois a derivada de h neste ponto seria zero, e o filtro ignoraria a medida realizada), por exemplo, aproximar a função de verossimilhança L_1 por uma gaussiana e obter a densidade a posteriori multiplicando-se pela densidade de previsão de estado (este é o procedimento adotado no filtro EQ abaixo). A mesma precaução deve ser tomada caso h^{-1} tenha apenas uma solução e y_{obs} for um valor próximo a um valor extremo de h e o estado previsto se encontrar nas vizinhanças do correspondente ponto extremo. Nesta situação, o filtro extendido modificado se comporta como o filtro extendido EX.

Os resultados obtidos a partir destes três filtros são praticamente coincidentes, o que era de se esperar, pois as iterações locais têm por finalidade corrigir apenas os erros de linearização da função de medidad e o extendido modificado busca os mesmos resultados simplesmente evitando os cálculos desnecessários do iterador local. Portanto, os três filtros são praticamente idênticos frente à não linearidade acentuada do estado. Frente à não linearidade da função de medidas os filtros IL e EM não constituem melhoras sobre o filtro EX, pois os três filtros ao adotarem uma solução linearizada de primeira ordem não podem levar em conta o caráter multimodal da função de verossimilhança.

Observando-se a Fig. 4.1 constata-se que o desempenho destes filtros na região altamente não linear da equação de transição de estado ($k=8$ a $k=25$) é satisfatório, apresentando apenas um leve aumento da relação AK. No entanto, quando as medidas são pronunciadamente não lineares ($k=48$ a $k=63$) o desempenho destes filtros é completamente insatisfatório e com resultados divergentes. (A explicação deste diferente comportamento foi fornecida no Capítulo III, seção 1.1). A divergência provém do aumento dos erros ao mesmo tempo em que a variância se reduz. A explicação fornecida é a de que a linearização conduz a uma função de verossimilhança com forma gaussiana exclusivamente concentrada próximo de uma das modas da função de verossimilhança verdadeira, reduzindo assim

a variância da densidade a posteriori. Para ainda mais o resultado, pode suceder que a aproximação se localize em torno da moda que não contenha o estado real (ver os exemplos de densidades a posteriori para $k=54$ a 59 - Fig. 2.11) aumentando. É claro, o erro de estimação. O gráfico do erro médio (Fig. 4.1) revela que estes filtros efetivamente são polarizados entre $k=48$ e $k=63$.

FILTROS IS, EQ, PR

IS é o filtro iterado com alissagem de estado combinado com duas iterações locais. O filtro EQ é o extendido modificado com aproximação da função de verossimilhança por gaussiana caso h^- tenha uma ou duas soluções, y_{obs} esteja próximo de um valor extremo de h e o valor previsto do estado se encontra nas vizinhanças do ponto extremo correspondente. O filtro PR é composto pelas equações do extendido para a obtenção da densidade de previsão de estado e a densidade a posteriori é calculada realizando o produto com a função de verossimilhança aproximada a priori (off-line), num domínio suficientemente grande para conter todas as modas de L_1 , por uma gaussiana. Evidentemente, esta aproximação é grossa e o valor estimado conterá um forte viés na região de alta não linearidade de medida. Mas o filtro estará prevenido da divergência pelo aumento substancial da variância (Fig. 4.A, abaixo).

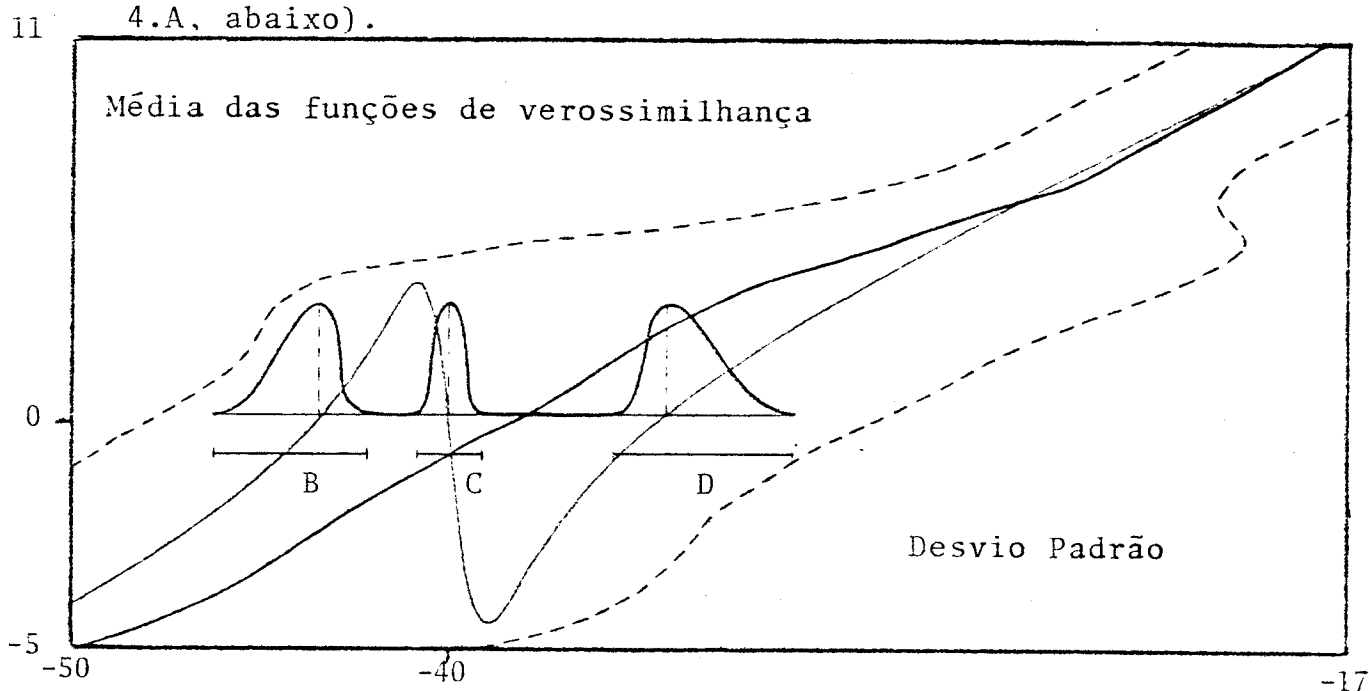


Fig. 4.A - Média e variância das funções de verossimilhança ou parâmetros que definem sua aproximação gaussiana

A partir dos gráficos da Fig. 4.2 pode-se avaliar o desempenho destes filtros. Para a região de forte não linearidade do estado os resultados dos três filtros são semelhantes com pequena superioridade do IS, pois apresenta menor erro médio e menor variância com uma relação AK não maior.

Entre os instantes de medidas fortemente não lineares o erro médio do filtro PR é maior do que o dos filtros IS e EQ.

A forma do gráfico do erro médio do filtro PR pode ser entendido a partir da Fig. 4.A: a linha do valor médio das funções de verossimilhança corresponde a uma aproximação da não linearidade de h e está patente a correlação entre o erro médio de estimação e o erro de aproximação. O problema pode ser visto de outra forma. Uma observação y_A pode ser provocada por um valor de estado real que pertença indistintamente a um dos domínios B, ou C ou D. A densidade de previsão de estado, se correta, estará concentrada no domínio que contém o estado real, por exemplo, no domínio B. A densidade a posteriori será basicamente determinada pelo produto da densidade de previsão (concentrada em B) pela parte da função de verossimilhança concentrada em B, as outras duas partes não exercendo influência significativa. Caso a densidade de previsão estivesse concentrada em C (ou D), seria a parte de L_1 concentrada em C (ou D), que determinaria basicamente a densidade a posteriori. Ora, quando L_1 é aproximada por uma única gaussiana, o filtro discerne com menor nitidez a influência das três modas e a densidade a posteriori estará consequentemente polarizada.

Contudo, a variância calculada por este filtro aumenta adequadamente para impedir a divergência como pode ser visto pelo gráfico da relação AK. O mesmo não sucede com os filtros IS e EQ, que apresentaram divergência. As iterações com alissagem são inefetivas, frente à não linearidade h e os resultados são os mesmos do filtro IL. O filtro EQ também não superam o filtro EM e isto porque mesmo quando a função inversa h^{-1} tem três soluções, a variância da densidade de previsão de estado pode ser suficientemente grande para conter no intervalo 3σ mais do que uma moda de L_1 . A escolha de uma das

das soluções implica na redução da variância da densidade a posteriori quando os erros estão aumentando.

Note-se que à excessão do filtro PR, todos os algoritmos até aqui utilizados trabalham com aproximações locais de primeira ordem. Constata-se que no caso de não linearidade de medida acentuadas em que a função de verossimilhança possui mais do que uma moda no domínio de valores significativos da densidade de previsão nenhum destes filtros poderá ter um comportamento satisfatório. Apenas o filtro PR com uma aproximação global grosseira da não linearidade consegue reduzir o valor de AK.

SO, FI, FS

À continuação apresentam-se os resultados obtidos com algoritmos de filtragem de segunda ordem, utilizando características puramente locais (filtro gaussiano de segunda ordem: SO) ou realizando uma aproximação de segunda ordem da não linearidade num certo domínio (filtros fitados: FI e FS).

Os domínios D_e e D_S em que as funções f e h são respectivamente aproximadas são:

$$D_e = (\hat{x}_{k/k} - s_1 k/k, \hat{x}_{k/k} + s_1 k/k) \quad e \quad s_1 = \sqrt{X_{k/k}}$$

$$D_S = (\hat{x}_{k+1/k} - s_2 k+1/k, \hat{x}_{k+1/k} + s_2 k+1/k) \quad e \quad s_2 = \sqrt{X_{k+1/k}}$$

Foram realizadas diversas simulações para diferentes valores de s_1 e s_2 , observando-se que $s_1 = 0,5$ e $s_2 = 3,0$ forneceram os melhores resultados. Um valor menor para o cálculo da densidade de previsão explica-se pelas razões já apontadas em III. Estes valores foram utilizados nas simulações apresentadas.

A Fig. 4.3 sumariza os resultados fornecidos por estes filtros. Verifica-se claramente que os filtros fitados (FI e FS) têm um desempenho, de modo geral, superior ao do SO.

são fornecidas. As gaussianas que aproximam L_1 são definidas pelos dois primeiros momentos, média e variância de L_1 calculados no domínio $x_{k+1}^* = (\hat{x}_{k+1/k} \pm 4\sigma_{k+1/k})$ e $\sigma_{k+1/k} = \sqrt{\hat{x}_{k+1/k}}$.

O filtro PT aproxima a densidade a posteriori.

A densidade de previsão de estado é calculada pelas equações do extendido para o filtro MP e pelas equações do fitado simplificado para os filtros PS e PT.

Os resultados são descritos pelos gráficos da Fig. 4.4. O que se observa no intervalo da não linearidade do estado é uma comparação entre os filtros EX e FS, pois como neste intervalo as medidas são praticamente lineares, o tratamento dado à nova informação pelo algoritmo MP, PS e PT é muito semelhante. O fato a ressaltar neste intervalo é que o filtro FS consegue uma relação AK mais próxima de 1 do que o filtro EX pelo aumento da variância.

Entre os instantes $k=50$ e $k=63$ (medidas não lineares) o filtro PT apresenta resultados excelentes: erro médio próximo a zero, relação AK muito próxima de 1 com a menor variância. Os resultados do filtro PS são semelhantes aos do PT apresentando apenas uma pequena polarização, que pode ser facilmente compreendida, pois aparece justamente nos instantes em que os valores medidos se encontram próximos dos valores extremos da função h . Provavelmente, poderia ser eliminada por um ajuste conveniente do intervalo x_{k+1}^* em que a aproximação é realizada. Os resultados do filtro MP também são satisfatórios. Comparado com o filtro PR, constitui uma melhora sensível: o erro médio é menor e menos polarizado, a variância aumenta menos e num menor intervalo e a relação AK é levemente inferior. A polarização deste filtro pode ser extendida de maneira análoga à do filtro PR e reduzi-la é tarefa que requer maior refinamento na aproximação da função de verossimilhança, aumentando o tamanho da tabela da qual o filtro fará uso.

EXEMPLO 2

Na região de alta não linearidade do estado a relação AK é bastante próxima de 1 para os três filtros. No entanto, os filtros fitados FI e FS são preferíveis ao SO, pois este tem uma variância extraordinariamente grande, um erro maior e viesado e maior do que o próprio erro médio do filtro EX (Fig. 4.1). A variância dos fitos é levemente maior do que a dos aproximados de primeira conduzindo a uma relação AK mais próxima de 1.

Entre os instantes $k=50$ e $k=63$ (medidas fortemente não lineares), o filtro FI tem o melhor desempenho, pois possui o menor erro médio e a relação AK mais próxima de 1. O filtro FS apresenta uma pequena polarização, porém consegue manter a relação AK suficientemente próxima da unidade aumentando a variância. O filtro SO é polarizado, embora menos do que os algoritmos de primeira ordem, devido aos termos de segunda ordem que corrigem polarizações, e sua variância aumenta em relação à daqueles, mas não o suficiente para impedir a divergência. É interessante observar que os filtros fitados FI e FS, e o filtro PR conseguem ajustar a variância com precisão suficiente para manter a relação AK próxima da unidade e o SO baseado em propriedades locais da não linearidade não o consegue. Dos algoritmos já discutidos neste Capítulo, sobressaem-se os fitados, FI e FS, e o PR, os únicos que conseguem manter AK satisfatoriamente próximo da unidade. Destes é preferível o FI, por apresentar menor variância (excluindo considerações sobre tempo de execuções).

MP, PS, PT

O algoritmo MP é semelhante ao PR, isto é, as funções de verossimilhança L_1 são aproximadas a priori (off-line) por gaussianas, mas a aproximação é mais refinada. Dado o caráter multimodal de L_1 no intervalo de não linearidade de medida acentuada, uma função de verossimilhança particular foi aproximada por diversas gaussianas, cada qual correspondendo a uma das modas ou a uma combinação de modas adjacentes.

O algoritmo PS aproxima a função de verossimilhança por uma gaussiana, levando em consideração a densidade de previsão em cada instante, à medida que as novas observações

Dado que foi possível pelo exemplo antes simulado distinguir os desempenhos dos filtros EX, IL e EM, devido ao fato das funções usadas apresentarem regiões de não linearidades muito acentuadas, estes algoritmos foram aplicados a dois sistemas, cujos estados evoluem linearmente, mas com não linearidades de medida de segundo e terceiro grau.

Para a não linearidade de segundo grau, o estado, as médias e as propriedades estatísticas são:

$$x_{k+1} = 1,02 x_k + 0,01k + w_k \quad (x_0 = 11 \text{ e } k=0,1,2,\dots,54)$$

$$y_k = 0,3 x_k^2 + 1,2 x_k + v_k$$

- estado inicial : N(6,10)
- ruído de estado : N(0,1)
- ruído de medida : N(0,10)

O estado evolui longe do ponto de mínimo da não linearidade da medida.

A Fig. 4.5 descreve os resultados. O gráfico do erro médio fornece boas razões para se afirmar que o filtro EM tem um desempenho superior aos outros dois e o gráfico da relação AK corrobora a afirmação. Os três filtros fornecem a mesma variância, o que se explica pela pequena variação da largura da função de verossimilhança. Nota-se que a melhora proporcionada pelo filtro EM num caso de não linearidade pouco acentuada consiste na eliminação da polarização inerente a qualquer aproximação de primeira ordem.

EXEMPLO 3

Trata-se de uma não linearidade de medida de terceiro grau. As equações de transição de estado e de medidas são:

$$x_{k+1} = 0,96 x_k - 0,02k + w_k \quad (x_0 = 20 \text{ e } k=0,1,\dots,80)$$

$$y_k = 0,2 x_k^3 + 0,1 x_k + v_k$$

As densidades do estado inicial e dos ruídos são:

- estado inicial : $N(15,10)$
- ruído de estado : $N(0,4)$
- ruído de medidas : $N(0,5)$

O estado evolui de forma tal que as primeiras observações se encontram em regiões de derivadas de primeira e segunda ordem da função h elevadas, em seguida atravessam o domínio em que estas derivadas são pequenas e se dirigem finalmente para a outra região de derivadas altas.

Neste exemplo, o filtro estendido foi incapaz de realizar uma estimativa adequada para o estado, isto é, o filtro estendido divergiu. O filtro estendido modificado foi claramente superior ao iterado com 4 iterações locais como mostram os gráficos da Fig. 4.6: erro médio menor, variância menor e relação AK mais próxima de 1. Para que a relação AK do filtro IL assuma valores próximos à relação AK do filtro EM são necessárias em torno de 10 iterações locais.

CONSIDERAÇÕES SOBRE TEMPOS DE CPU

A tabela 4.1 apresenta os tempos de CPU dos diversos filtros aplicados ao exemplo 1 relativos ao filtro estendido que é tomado como padrão. A comparação foi feita sobre os tempos de CPU de cada um dos filtros em simulações de Monte Carlo com 700 realizações. Todos os filtros foram simulados em idênticas condições quanto à inicialização e escritura dos resultados.

O filtro EQ é relativamente mais lento do que o EM pois realiza a aproximação da função de verossimilhança por uma gaussiana durante o processamento (on Line) para os casos em que o valor y_{obs} se situe nas proximidades de um valor extremo da função e o valor do estado se encontre próximo do ponto extremo correspondente. O fato dos filtros PS e PT terem uma velocidade de processamento muito inferior aos outros filtros se deve à rotina de integração utilizada. O tempo de processamento para os filtros PS e PT é determinado principalmente pela rotina de integração numérica que deve ser utilizada.

No presente trabalho não se fez uma seleção do método de integração mais adequado para o tipo particular de funções de que se trata: densidades de probabilidades.

Constitui fato notável, mas não surpreendente, que os filtros PR e MP tenham um tempo de execução muito próximo ao filtro EX. Evidencia-se também que o filtro EM é mais rápido do que o IL com 4 iterações locais.

EX: 1	PR: 1,00	SO: 1,11	FS: 1,54	PS: 73,9
	EM: 1,08	IL: 1,22	EQ: 4,25	PT: 80,4
	MP: 1,09	IS: 1,35	FI: 6,36	NL: 1373

Tabela 4.1 - Tempos de CPU relativos ao tempo
do filtro EXTENDIDO (EX)

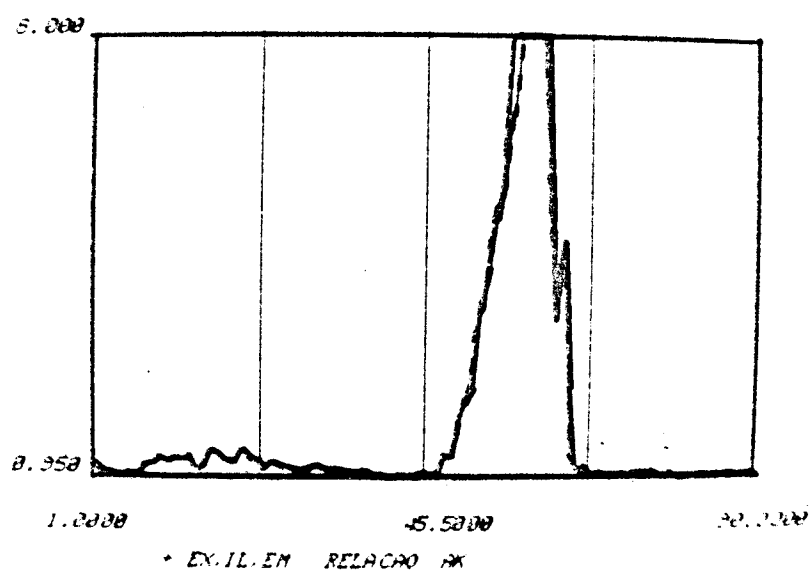
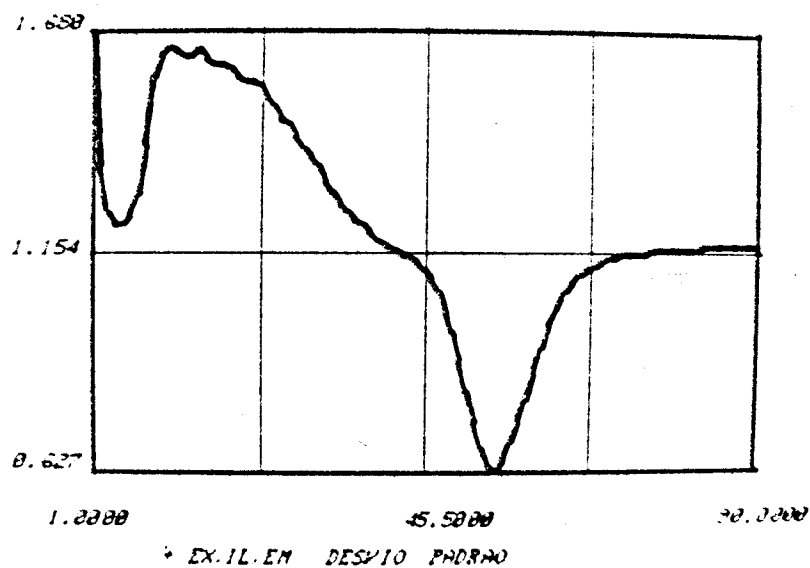
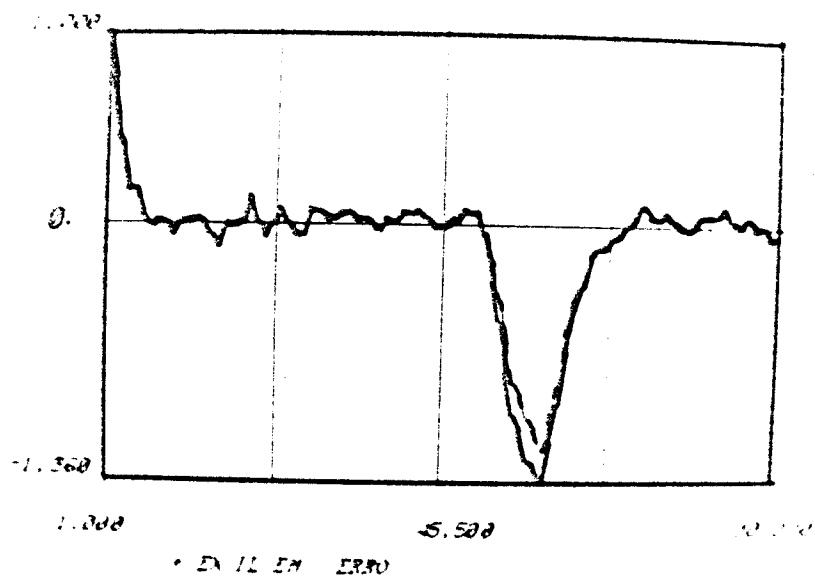


Fig. 4.1 - Filtros* EX, IL e EM aplicados ao exemplo 1

* Convenção adotada para as Figs. 4.1 - 4.6:
 corresponde ao primeiro filtro
 corresponde ao segundo filtro
 corresponde ao terceiro filtro

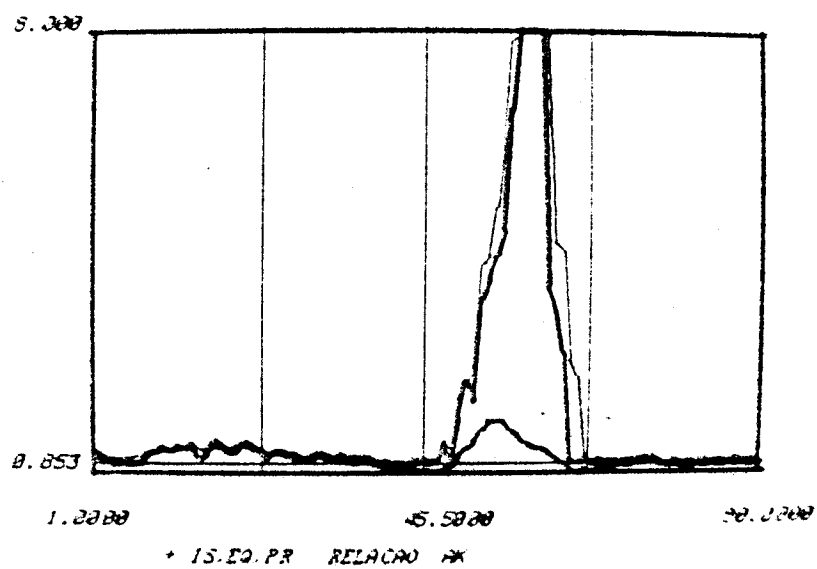
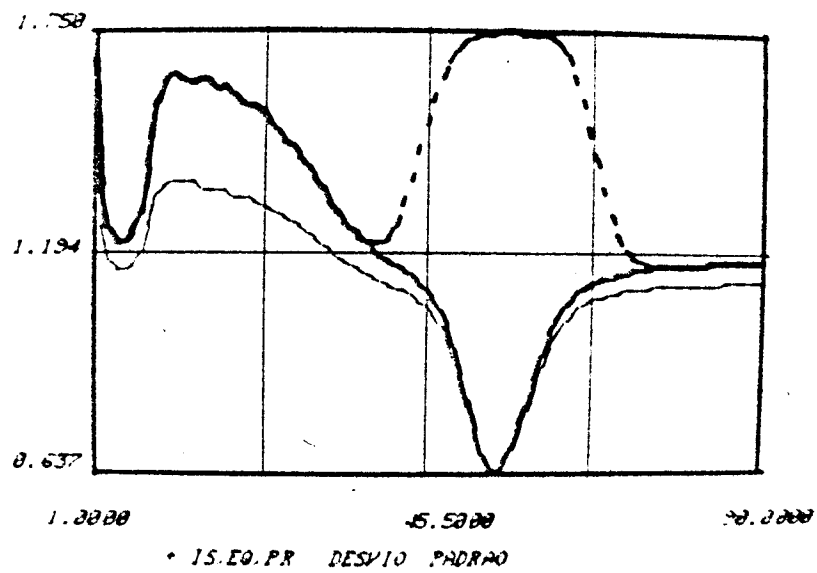
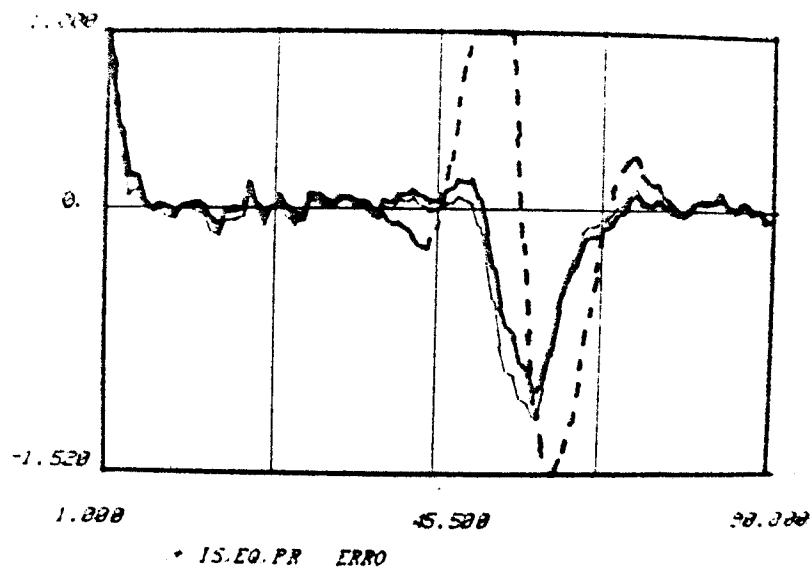


Fig. 4.2 - Filtros IS, EQ e PR aplicados
ao exemplo 1

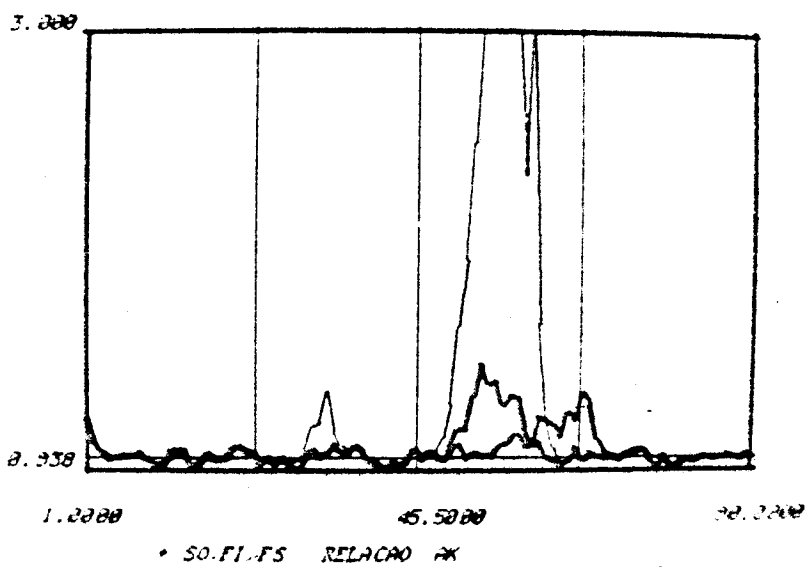
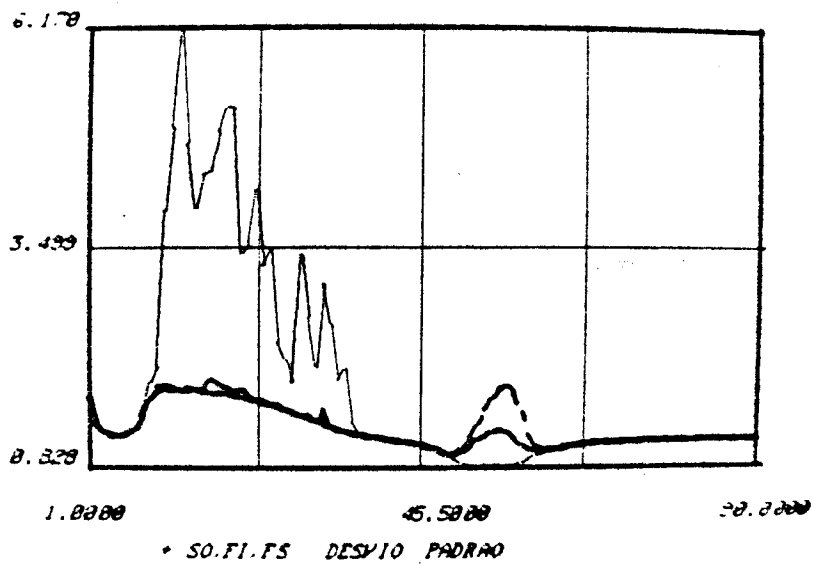
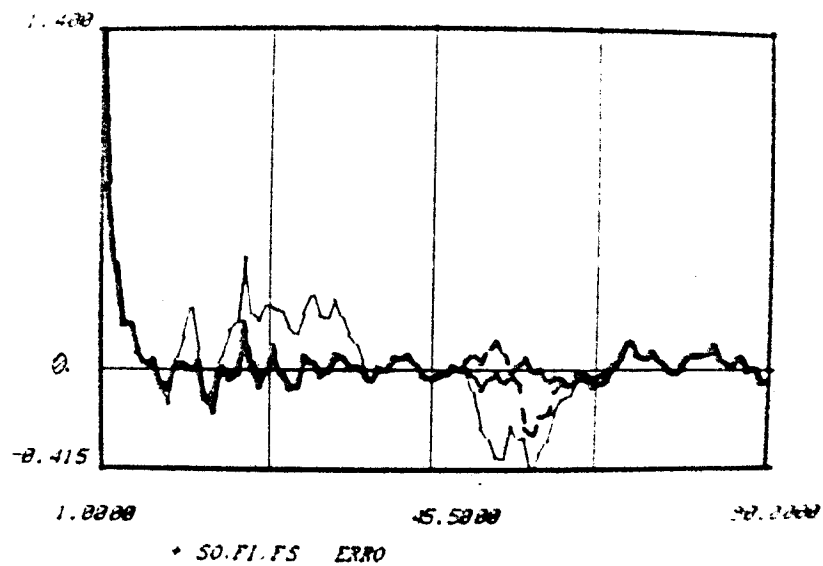


Fig. 4.3 - Filtros SO, FI, e FS aplicados
ao exemplo 1

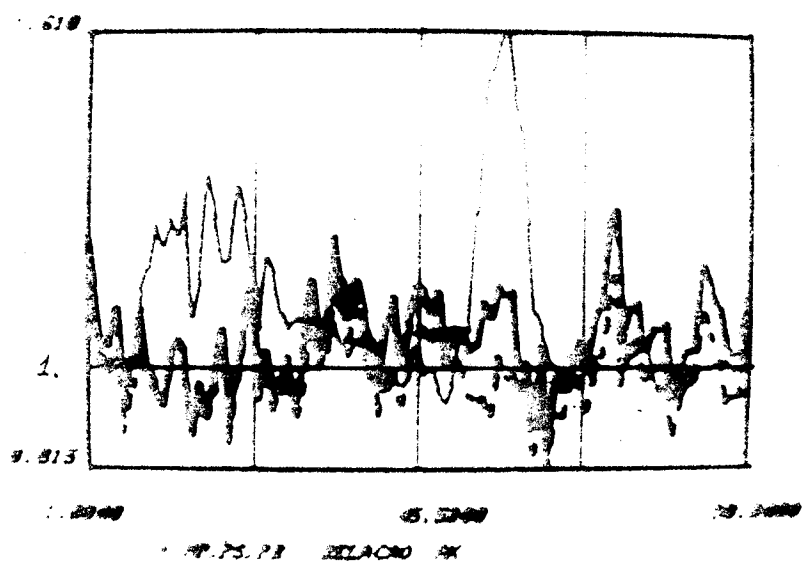
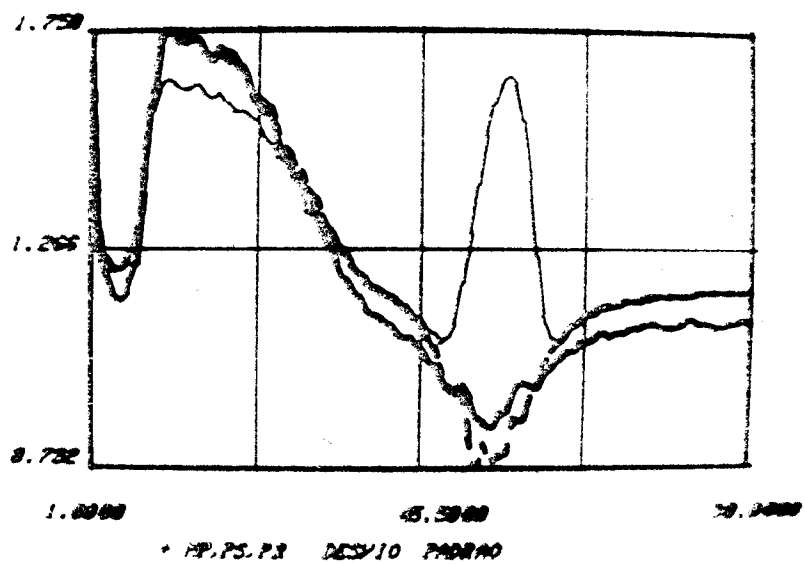
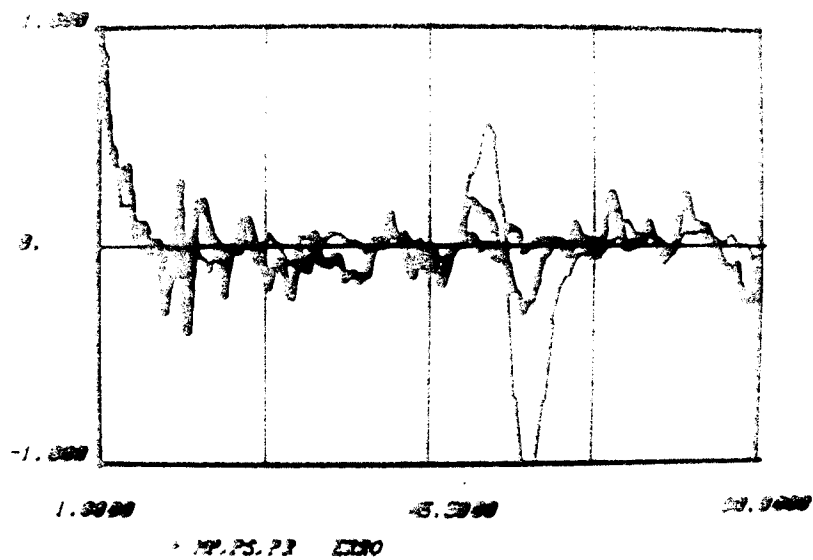


Fig. 4.4 - Filtros MP, PS e PT aplicados
ao exemplo 1

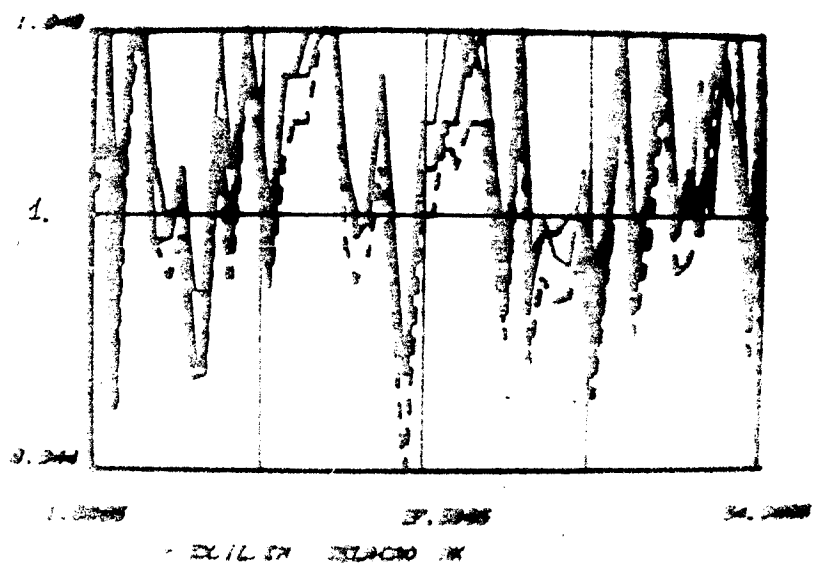
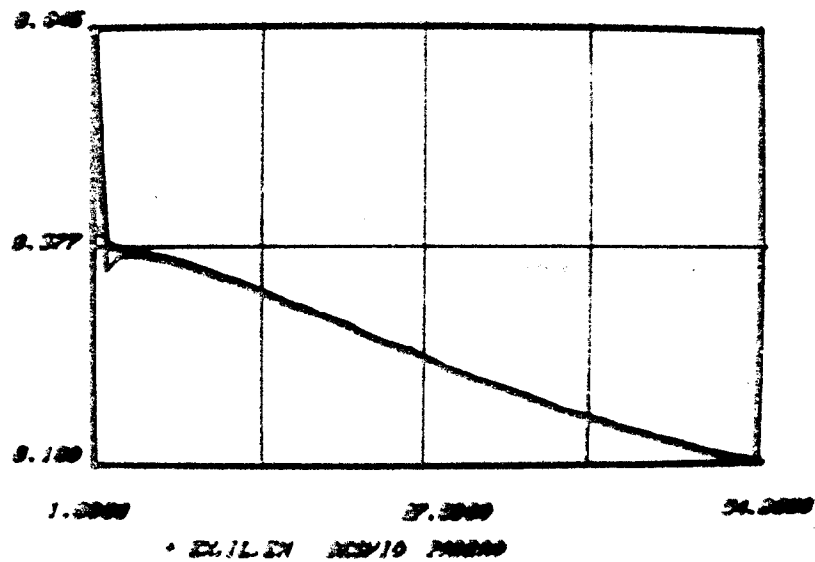
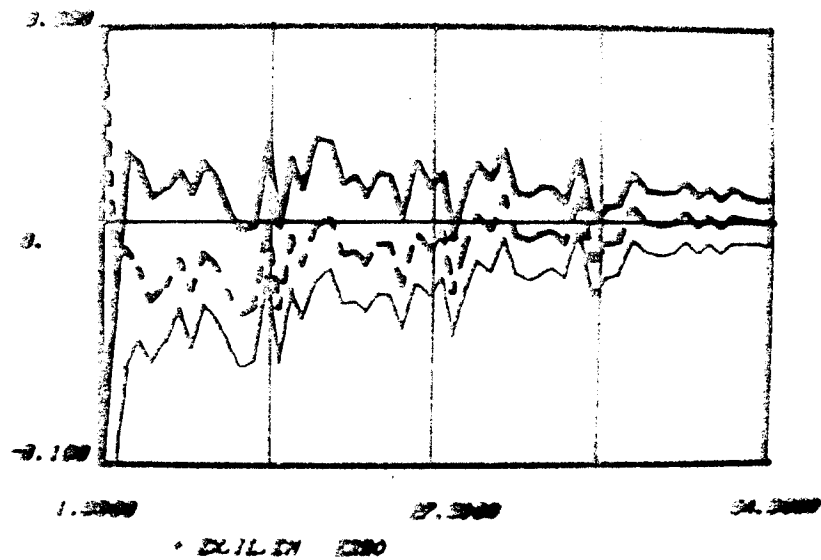


Fig. 4.5 - Filtros EX, IL e EM aplicados ao exemplo 2: não linearidades de medida do segundo grau

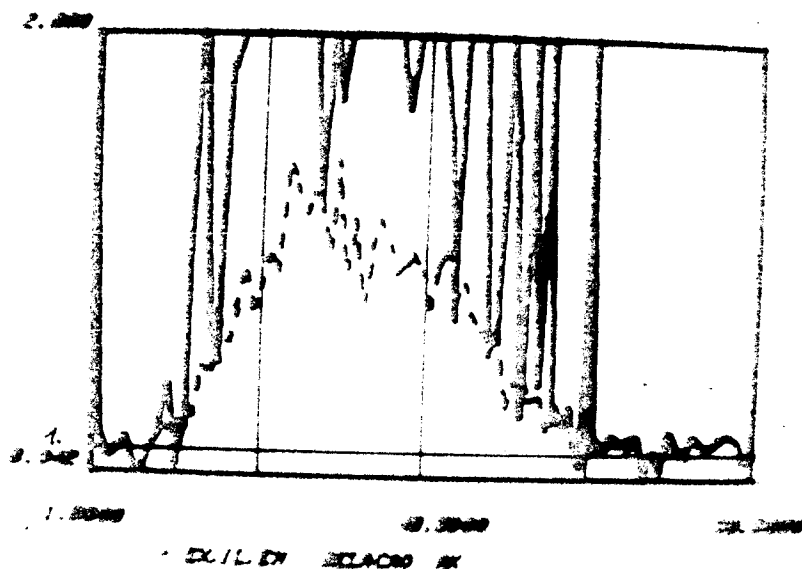
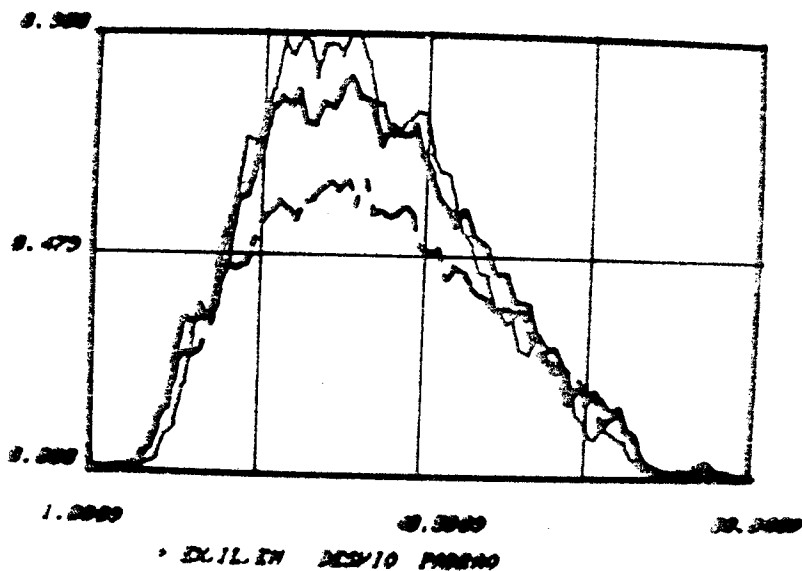
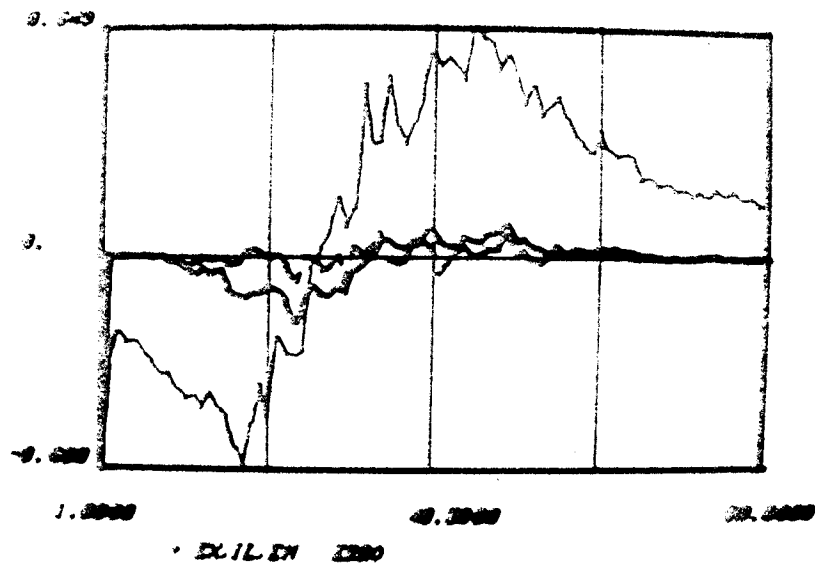


Fig. 4.6 - Filtros EX, IL e EX aplicados ao exemplo 3: não linearidade de medida do terceiro grau

CONCLUSÕES

A partir da análise dos algoritmos realizada no Capítulo III e dos resultados numéricos do Capítulo IV, podem ser sumarizadas as seguintes conclusões:

- 1) Os filtros generalizados e iterados, incluindo o estendido modificado, têm um desempenho satisfatório para não linearidades pouco acentuadas, isto é, quando as funções de verossimilhança L e L_1 tenham uma forma bastante próxima da gaussiana. Estes filtros têm um tempo de execução suficientemente reduzido para torná-los muito atrativos.
- 2) O ponto crucial, em termos de velocidade de processamento, do filtro ótimo é o cálculo da densidade de previsão de estado por necessitar de uma série de integrações numéricas. Por outro lado, dada a densidade de previsão de estado, obter a densidade a posteriori pelo filtro ótimo é relativamente simples, pois basta realizar o produto daquela pela função de verossimilhança L_1 . O cálculo dos momentos da densidade a posteriori necessita exclusivamente três integrações numéricas.
- 3) Os valores estimados são menos sensíveis a erros de linearização da equação de transição de estado do que a erros de linearização da função de medidas.
- 4) Os filtros generalizados e iterados têm desempenho satisfatório frente a não linearidades do estado acentuadas, mas completamente insatisfatório frente a não linearidades de medida acentuadas, ou seja, quando a função de verossimilhança possui mais do que uma moda no intervalo de valores significativos da densidade de previsão de estado.
- 5) Os filtros fitados conseguem um desempenho mais satisfatório nos intervalos de não linearidades

acentuadas, através de uma correção de polarização e aumento de variância mais adequados do que o filtro SO.

A velocidade de processamento do filtro FS não é muito inferior a do EX, mas a do filtro FI é relativamente baixa.

- 6) Os filtros por aproximação da função de verossimilhança a priori são comparáveis em tempos de execução do filtro EX. Apesar de serem polarizados nas regiões de não linearidades de medida acentuadas, conseguem um desempenho satisfatório, mediante o aumento adequado da variância. Os filtros que aproximam a função de verossimilhança durante o processamento, apresentam resultados plenamente satisfatórios, porém apresentam baixa velocidade de processamento quando comparados com o filtro estendido.
- 7) Determinar o ponto solução de $y_{k+1}=h(x_{k+1},k+1)$ para $y_{k+1}=y_{obs}$ para o filtro estendido modificado em geral envolve um processo iterativo, mas será mais rápido do que o processo iterativo utilizado pelo filtro estendido com iterações locais.

As simulações levadas a efeito neste trabalho, limitaram-se ao caso escalar. Evidentemente, esta situação é muito restrita, porquanto os casos reais sempre apresentam ordem superior a 1. Mas não há que se desprezar a exemplificação escalar, pois sua simplicidade fornece elementos importantes para a compreensão do funcionamento destes algoritmos. Esforços já estão sendo envidados para transladar os resultados obtidos para o caso multidimensional. Grandes esperanças são depositadas nos filtros que realizam a aproximação da função de verossimilhança ou que aproximam diretamente a densidade a posteriori, quando as integrações numéricas forem realizadas com o auxílio de "splines" interpoladores devidamente generalizados para situações multidimensionais. Outro ponto que admite um ulterior desenvolvimento é a determinação dos domínios para a filtragem. Neste trabalho, forneceu-se uma explicação para a diferença entre o domínio para aproximação da e

quação de transição de estado e o domínio para aproximação da função de saída, os valores, determinados empiricamente, obedeceram ao critério de apresentar uma relação AK o mais próximo possível de 1. Sugere-se que os domínios sejam determinados si multaneamente com a filtragem, minimizando por exemplo a variância da relação AK que deve se situar em torno de 1.

REFERÊNCIAS BIBLIOGRÁFICAS

- (1) - ALSPACH, D.L. & SORENSON, H.W.
"Nonlinear Bayesian estimation using gaussian sum approximations"
IEEE Trans. on Automatic Control, 17(4):439-447, Aug. 1972
- (2) - ANDRADE NETTO, M.
"Interpolação por splines: aplicações em filtragem de sistemas dinâmicos não lineares"
Faculdade de Engenharia de Campinas, 1976 - Tese (Doutoramento)
FEC - UNICAMP - Campinas
- (3) - ANDRADE NETTO, M.; GIMENO, L.; MENDES, M.J.
"Nonlinear filtering of discrete time systems"
in: Identification and system parameter estimation, Moscow,
Institute of Control Sciences, 1976-p.3, p.319-328
(IFAC Symposium, 4 - Tbilisi, USSR, 1976)
- (4) - AOKI, M.
Optimization of stochastic systems - topics in discrete time systems
New York, Academic Press, 1967
(Mathematics in Science and Engineering, V.32)
- (5) - BRUNO, P.T.M.
"Estudo do desempenho de métodos de filtragem sequencial aplicados a sistemas não lineares com aproximação até segunda ordem"
Campinas - Faculdade de Engenharia de Campinas - 1976 - Tese (Mestre em Ciências) - FEC - UNICAMP - Campinas
- (6) - BUCY, R.S. & SENNE, K.D.
"Realization of optimum Discrete-Time Nonlinear Estimation Theory and its Application"
San Diego, California - Sep. 1970
- (7) - DENHAM, W.F. & PINES, S.
"Sequential estimation when measurement function nonlinearity is comparable to measurement error"
AIAA Journal, 4(6):1071-1076, Jun. 1966
- (8) - FITZGERALD, R.J.
"Divergence of the Kalman Filter"
IEEE Trans. on Automatic Control, 16(6):736-747, Dec. 1971

- (9) - JASWINSKY, A.H.
Stochastic processes and filtering theory
New York, Academic Press, 1970
(Mathematics in Science and Engineering; V.64)
- (10) - KALMAN, R.E.
"A new approach to linear filtering and prediction problems"
Trans. ASME Journal Basic Engineering, New York, V.82:34-45,
Mar. 1960
- (11) - KALMAN, R.E. & BUCY, R.S.
"New results in linear filtering and prediction theory"
Trans. ASME Journal Basic Engineering, New York, V.83:95-107,
Mar. 1961
- (12) - SAGE, A.P. & MESLA, J.L.
Estimation theory with applications to communications and control
Mc Graw-Hill, New York, 1971 (Mc Graw-Hill series in systems
science)
- (13) - SCHMIDT, G.T.
"Linear and Nonlinear Filtering techniques"
in: LEONDES, C.T. ed - Control and dynamic systems; advances in
theory and applications - New York, Academic Press, V.12:62 - 1976
- (14) - SORENSON, H.W.
"An overview of filtering and stochastic control in dynamic
systems"
in: LEONDES, C.T. ed. - Control and dynamic systems; advances
in theory and applications - New York,
Academic Press, V.12:1-61, 1976
- (15) - SORENSON, H.W. & ALSPACH, D.L.
"Recursive Bayesian estimation using Gaussian sums"
Automatica, Great Britain, 7(4):465-479, Jul. 1971
- (16) - TOZZI, C.L.
"Estudo do desempenho de métodos sequenciais de filtragem não
linear usando aproximações iteradas de primeira ordem"
Faculdade de Engenharia de Campinas, 1976 - Tese (Mestre em Ciências)
FEC - UNICAMP - Campinas

(17) - TSE, E.

"Parallel computation of the conditional mean state estimate
for nonlinear systems".

Systems Control, Inc., Palo Alto, California, 1971

(18) - WIENER, N.

The Fourier integral and certain of its applications

New York, Dover Publications, Inc., 1933