

UNIVERSIDADE ESTADUAL DE CAMPINAS
FACULDADE DE ENGENHARIA ELÉTRICA
DEPARTAMENTO DE MÁQUINAS, COMPONENTES E SISTEMAS INTELIGENTES

FILTRO DE KALMAN:
HIERARQUIZAÇÃO E COMPUTAÇÃO PARALELA

Autor :

ROGÉRIO BASTOS QUIRINO *sk*

Orientador :

PROF. DR. CELSO PASCOLI BOTTURA *X*

Tese apresentada à Faculdade de Engenharia
Elétrica da Universidade Estadual de Campinas -
- UNICAMP , como parte dos requisitos exigidos
para a obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA

AGOSTO - 1990

BC/9104471

UNICAMP
BIBLIOTECA CENTRAL

Este exemplar corresponde à edição final da tese
defendida por *Rogério Bastos Quirino*

aprovada pela Comissão

Julgadora em

Celso Pascoli Bottura
Orientador

" O INFINITO SEPARA O NADA DE UM SIMPLES TRABALHO "

Aos

Meus Pais, Gerson e Maria Aparecida

Meus Irmãos, Garibaldi, Gerson, Inês, Glêbson,

Jefferson, Neide, Romero e familiares.

" AGRADECIMENTOS "

Agradeço a todos que direta ou indiretamente contribuíram para a realização deste trabalho.

Meu eterno agradecimento ao Prof. Dr. Celso Pascoli Bottura pela maneira humilde, sincera e perspicaz de orientar e valorizar este trabalho.

A Ingeborg e Philipp que me proporcionaram bons momentos durante a realização deste trabalho.

A Eílson e Otacílio, meus amigos de convívio, que sempre souberam me entender e aceitar.

Aos colegas Tarcísio J. T. Filho, Gilmar Barreto, Celso Munaro, Carlos Richter, Gorgônio Araújo, Miguel Rozsas e Samuel Benchimol, que contribuíram diretamente na realização deste trabalho.

As colegas Ademilde Félix e Elza Aoki e, ao colega Celso R. D. Bueno, pela paciência e contribuição na confecção deste trabalho.

Aos Colegas e Amigos da Pós-Graduação, que via um excelente relacionamento, sempre me ajudaram de alguma maneira.

RESUMO

Neste trabalho realizamos a computação paralela de um algoritmo para filtragem ótima de sistemas dinâmicos lineares interconectados, explorando eficientemente o paralelismo natural da estrutura de cálculo hierárquico, empregando multiprogramação com o sistema operacional Unix.

ABSTRACT

In this work the computational parallelization of an algorithm for optimum filtering of large scale systems is made, efficiently exploiting the natural parallelism there is in a hierarquical calculation structure, via multiprogramming on the operational system Unix.

ÍNDICE

PÁGINA

CAPÍTULO-1 - TEORIA DO FILTRO DE KALMAN DISCRETO

11 -INTRODUÇÃO	1
12 -BASE DA TEORIA CLÁSSICA DE ESTIMAÇÃO	1
1.2.1 - ESTIMAÇÃO MÉDIA QUADRÁTICA (VARIÂNCIA MÍNIMA)	2
1.2.2 - ESTIMAÇÃO MÉDIA QUADRÁTICA DE UMA VARIÁVEL ALEATÓRIA X DADA UMA VARIÁVEL ALEATÓRIA Z: CASO GERAL	3
1.2.3 - PRINCÍPIO DA ORTOGONALIDADE	5
1.2.4 - ESTIMAÇÃO MÉDIA QUADRÁTICA LINEAR DE UMA VARIÁVEL ALEATÓRIA X, DADA UMA VARIÁVEL ALEATÓRIA Z	7
13 -OBSERVADOR DE ESTADO DETERMINÍSTICO	11
1.3.1 - ASPECTOS ESTOCÁSTICOS	18
1.3.1.1 - PROPAGAÇÃO DAS MÉDIAS E COVARIÂNCIAS	19
14 -FILTRO DE KALMAN DISCRETO VIA TEOREMA DA PROJEÇÃO ORTOGONAL	23
15 -FILTRO DE KALMAN DISCRETO VIA TEOREMA DE BAYES	32
16 -FILTRO DE KALMAN DISCRETO VIA CÁLCULO VARIACIONAL E DUALIDADE AO CONTROLE ÓTIMO ...	39

CAPÍTULO-2 - ESTIMAÇÃO MULTINÍVEL DE ESTADO

2.1 -INTRODUÇÃO	43
2.2 -O FILTRO DE KALMAN GLOBAL E O FILTRO MULTINÍVEL DE PEARSON	46
2.2.1 - O FILTRO DE KALMAN GLOBAL	
2.2.2 - O FILTRO MULTINÍVEL DE PEARSON (MÉTODO COORDENAÇÃO POR OBJETIVO)	48
2.3 -O FILTRO MULTINÍVEL DE ARAFEH E SAGE (MÉTODO MAP)	55
2.4 -O FILTRO PARTICIONADO SUBÓTIMO DE SHAH (MÉTODO SPA)	67
2.5 -O FILTRO DE KALMAN MULTINÍVEL (MÉTODO DE PROJEÇÕES MÚLTIPLAS)	73

CAPÍTULO-3 - PROGRAMAÇÃO PARALELA DO FILTRO DE KALMAN MULTINÍVEL

3.1 -INTRODUÇÃO	95
3.2 -DESCRIÇÃO DAS ESTRUTURAS HIERÁRQUICAS DE CÁLCULO DO FILTRO DE KALMAN MULTINÍVEL	95
3.3 -AMBIENTES DE MULTIPROGRAMAÇÃO E PROCESSAMENTO PARALELO DO FILTRO DE KALMAN MULTINÍVEL	107
3.3.1 - INTRODUÇÃO	
3.3.2 - SISTEMAS DE MULTIPROGRAMAÇÃO	

3.3.3 - AMBIENTE DE PROGRAMAÇÃO UNIX	108
3.3.4 - MULTIPROGRAMAÇÃO COM TEMPO COMPARTILHADO E FILTRAGEM DE KALMAN PARALELA	111
3.3.5 - IMPLEMENTAÇÃO EM MULTIPROGRAMAÇÃO COM TEMPO COMPARTILHADO	113
3.4 - ASPECTOS DE CONVERGÊNCIA	122
3.5 - IMPLICAÇÕES TÉCNICAS DA UTILIZAÇÃO DOS ALGORITMOS PARALELOS DE FILTRAGEM	123
3.5.1 - INTRODUÇÃO	
3.5.2 - APLICAÇÕES EM REDES INTEGRADAS MULTISENSORIAIS	125

CAPÍTULO-4 - ANÁLISE DE DESEMPENHO COMPUTACIONAL

4.1 - INTRODUÇÃO	129
4.2 - REQUISICÃO DE TEMPO DE PROCESSAMENTO	129
4.3 - DADOS COMUNICADOS	137
4.4 - APLICAÇÃO E RESULTADOS NUMÉRICOS	140

CAPÍTULO-5 - CONCLUSÃO	143
------------------------------	-----

REFERÊNCIAS BIBLIOGRÁFICAS

CAPÍTULO 1

FILTRO DE KALMAN DISCRETO

1.1. INTRODUÇÃO

Baseado num estudo [1] da história da teoria de filtragem linear, exponho alguns aspectos importantes relacionados ao filtro de Kalman.

Num reexame da teoria clássica de predição e filtragem existentes, Kalman [2] combinou descrições de espaço de estado e a noção de inovações discretas no tempo, como enfocadas por exemplo em [3, Seções XII.1 e XII.3], para apresentar uma completa e elegante solução.

A solução de Kalman também introduziu uma equação de recorrência não linear que é uma solução discreta no tempo da equação diferencial de Riccati, já encontrada por ele em estudos relacionados a problemas de minimização quadrática em controle ótimo; Kalman verificou uma dualidade existente entre problemas de filtragem e controle. Uma imediata consequência da análise do comportamento em regime estacionário da equação de Riccati em controle ótimo foi que, em certas condições de observabilidade e controlabilidade, a solução convergia para uma única solução, num intervalo de tempo finito, independente da condição inicial e dos erros conduzidos durante a computação.

1.2. BASE DA TEORIA CLASSICA DE ESTIMAÇÃO [5,6]

Existem muitos meios para se estimar uma quantidade desconhecida a partir de dados disponíveis.

Neste ítem, nos preocupamos em mostrar parcialmente as bases da teoria clássica de estimação anteriores a 1960, quando Kalman e Bucy desenvolveram uma solução [2,4].

1.2.1. ESTIMAÇÃO MÉDIA QUADRÁTICA (VARIÂNCIA MÍNIMA)

Suponha um vetor x desconhecido e um vetor z disponível relacionado com x de alguma forma, tal que o reconhecimento de z traz alguma informação sobre x . Os vetores x e z são processos estocásticos e as estatísticas conjuntas do x e z são conhecidas à priori.

Dado z , gostaríamos de fazer uma estimação $\hat{x}(z)$ do valor de x tal que o erro de estimação

$$\tilde{x} = x - \hat{x} \quad (1.2.1-1)$$

seja pequeno. Pequeno pode ser definido de muitas formas, conduzindo a diferentes métodos de estimação.

Definimos o erro médio quadrático como o valor esperado do quadrado da norma Euclidiana de $\tilde{x}$.

$$J = E \left[\tilde{x}^T \tilde{x} \right] = E \left[(x - \hat{x})^T (x - \hat{x}) \right] \quad (1.2.1-2)$$

onde $E(\cdot)$ representa a
esperança matemática

O critério de estimação média quadrática pode ser estabelecido da seguinte forma: utilizando todas as informações disponíveis (estatísticas à priori de x e z e o conhecimento de z), seleccionamos uma estimação $\hat{x}$ para minimizar 1.2.1-2.

A estimação média quadrática é um caso especial da estimação de Bayes. O problema de estimação de Bayes é encontrar a estimação que minimize a função risco de Bayes:

$$J = E \left[C(\tilde{x}/z) \right] = \int_{-\infty}^{\infty} C(x - \hat{x}) f_{xz}(x, z) dx dz, \quad (1.2.1-3)$$

onde $c(\cdot)$ é uma função custo e $f_{xz}(x, z)$ é a função densidade de probabilidade conjunta. Selecionando $C(\tilde{x}) = \tilde{x}^T \tilde{x}$ resulta em 1.2.1-2. Um outro caso especial é a estimação máximo a posteriori: $C(\tilde{x}) = 0$ se todas as componentes de $\tilde{x}$ têm magnitude menor do que um dado ϵ e $C(\tilde{x}) = 1$ caso contrário.

1.2.2. ESTIMAÇÃO MÉDIA QUADRÁTICA DE UMA VARIÁVEL ALEATÓRIA X, DADA UMA VARIÁVEL ALEATÓRIA Z: CASO GERAL

Supomos que a estatística à priori do valor x é conhecida e o valor Z é disponível por um processo de observação. Nenhuma relação entre x e Z é suposta. Supomos que a função densidade de probabilidade conjunta $f_{xz}(x, z)$ de x e z é dada. A estimação média quadrática depende da medida Z (conhecida), tal que podemos escrever $\hat{x}(z)$.

Para um valor fixo $Z = z$, encontramos $\hat{x}(z)$, reescrevendo 1.2.1-2 como:

$$J = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [x - \hat{x}(z)]^T [x - \hat{x}(z)] f_{xz} dx dz$$

ou, utilizando a regra de Bayes:

$$J = \int_{-\infty}^{\infty} f_z(z) \int_{-\infty}^{\infty} [x - \hat{x}(z)]^T [x - \hat{x}(z)] f_{x/z}(x/z) dx dz,$$

onde $f_{x/z}$ representa a função densidade de probabilidade condicional. Desta forma, J pode ser minimizada para cada z , se o valor do erro médio quadrático condicional

$$J' = E \int_{-\infty}^{\infty} [x - \hat{x}(z)]^T [x - \hat{x}(z)] f_{x/z}(x/z) dx \quad (1.2.2-1)$$

for minimizado.

Utilizando a média condicional, podemos escrever:

$$J' = \left\{ \left[x - \hat{x}(z) \right]^T \left[x - \hat{x}(z) \right] / z \right\}$$

$$= E \left[x^T x / z \right] - 2 \hat{x}^T(z) E [x/z] + \hat{x}^T(z) \hat{x}(z) \quad (1.2.2-2)$$

Minimizando a função em relação a $\hat{x}(z)$, encontramos que:

$$\frac{\partial J'}{\partial \hat{x}(z)} = -2 E [x/z] + 2 \hat{x}(z) = 0$$

e finalmente:

$$\hat{x}_{\text{m.m.}}(z) = E [x/z] , \quad (1.2.2-3)$$

onde a notação m.m.
significa mínimos quadrados

Se x e z são independentes, então $f_{xz} = f_x f_z$ tal que

$$E [x/z] = \int_{-\infty}^{\infty} x f_{x/z}(x/z) dx = \int_{-\infty}^{\infty} x \frac{f_{xz}(x, z)}{f_z(z)} dx = E [x] \quad (1.2.2-4)$$

Isto quer dizer que o conhecimento de z não provê qualquer informação sobre x .

Observe que para encontrar $E [x/z]$ requer o conhecimento de f_{xz} , a qual pode não ser disponível. Em alguns casos, conhecemos somente o primeiro e segundo momentos.

A equação 1.2.2-3 representa a estimação média quadrática generalizada.

A estimação de x representada pela equação 1.2.2-3 é não polarizada se:

$$E \left[\tilde{x}/z \right] = 0, \quad \text{ou ainda}$$

$$E \left[(x - \hat{x})/z \right] = 0 \Rightarrow E [x/z] = E \left[\hat{x}/z \right]$$

A precisão da estimação é dada pela covariância do erro dessa estimação.

Para encontrar a covariância do erro associada com a estimação representada pela equação 1.2.2-3 escrevemos:

$$\begin{aligned} P_{\tilde{x}} &= E \left[\tilde{x} \tilde{x}^T \right] = E \left[(x - \hat{x}) (x - \hat{x})^T \right] \\ &= E \left[x - E(\hat{x}/z) \right] \left[x - E(\hat{x}/z) \right]^T / z \\ &= E \left[P_{x/z} \right] \end{aligned} \quad (1.2.2-5)$$

onde $P_{x/z}$ é a função densidade de probabilidade condicional.

1.2.3. O PRINCÍPIO DA ORTOGONALIDADE

Este princípio é básico e de fundamental importância para a teoria de estimação probabilística.

Sejam x e z variáveis aleatórias distribuídas conjuntamente. Então para qualquer função $g(\cdot)$,

$$E \left\{ g(z) \cdot \left[x - E(x/z) \right]^T \right\} = 0$$

Isto é, qualquer função de z é ortogonal a x subtraída da média condicional de x . Para mostrar isto, escrevemos

$$\begin{aligned} E \left\{ g(z) \cdot [x - E(x/z)]^T \right\} &= E \left\{ g(z) \cdot x^T - E \left[g(z) \cdot E(x^T/z) \right] \right\} \\ &= E \left\{ g(z) \cdot x^T \right\} - E \left\{ E \left[g(z) (x^T/z) \right] \right\} \\ &= E \left\{ g(z) \cdot x^T \right\} - E \left\{ g(z) \cdot x^T \right\} = 0 \end{aligned}$$

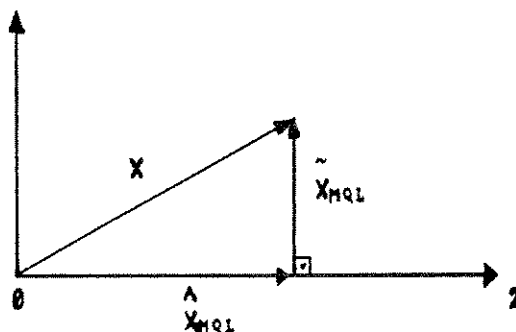
O princípio da ortogonalidade pode ser ilustrado como na Figura (1.1). Todas as funções de z são representadas na direção de z .

O princípio da ortogonalidade diz que a variável aleatória

$$\tilde{x} = x - E(x/z) \quad (1.2.3-1)$$

que é o erro de estimação, é ortogonal a todas as variáveis aleatórias $g(z)$. A Figura (1.1) é simplesmente heurística, pois se a função $g(\cdot)$ for não linear, então as variáveis aleatórias da forma $g(z)$ não formam um subespaço.

A Figura (1.1) nada mais é do que um dispositivo mnemônico.



MQL - Mínimos Quadrados Linear

Figura 1.1 - Princípio da Ortogonalidade

Tal princípio se aplica da seguinte forma à teoria de estimação: se $g(\cdot)$ é qualquer função, então

$$E \|x - E(x/z)\| \leq E \|x - g(z)\| , \quad (1.2.3-2)$$

onde $\|\cdot\|$ é a norma Euclidiana. Portanto, nenhuma outra função de z provê uma melhor aproximação para x no sentido probabilístico do que $E(x/z)$.

1.2.4. ESTIMAÇÃO MÉDIA QUADRÁTICA LINEAR DE UMA VARIÁVEL ALEATÓRIA x , DADA UMA VARIÁVEL ALEATÓRIA z

Restringimos a estimação $\hat{x}(z)$ a uma dependência linear de z , isto é

$$\hat{x}(z) = Az + b , \quad (1.2.4-1)$$

para uma matriz constante $A \in \mathbb{R}^{n \times p}$ e um vetor $b \in \mathbb{R}^n$. A e b devem ser escolhidos de forma a minimizar o erro médio quadrático 1.2.1-2.

Podemos escrever 1.2.1-2 como sendo:

$$\begin{aligned} J &= \text{tr } E \left\{ (x - \hat{x}) (x - \hat{x})^T \right\} \\ &= \text{tr } E \left\{ (x - Az - b) (x - Az - b)^T \right\} \\ &= \text{tr } E \left\{ \begin{bmatrix} x - E(x) \end{bmatrix} - \begin{bmatrix} Az + b - E(x) \end{bmatrix} \right\} \\ &\quad \left\{ \begin{bmatrix} x - E(x) \end{bmatrix} - \begin{bmatrix} A(z + b - E(x)) \end{bmatrix}^T \right\} \end{aligned}$$

Fazendo algumas manipulações, resulta que:

$$J = \text{tr} \left\{ P_x + A \left[P_z + E(z) \cdot E(z^T) \right] \cdot A^T + \left[b - E(x) \right] \left[b - E(x) \right]^T + 2AE(z) \cdot \left[b - E(x) \right]^T - 2A P_{xz} \right\}$$

Como $\frac{d}{dA} \text{tr}(ABA^T) = 2AB$ e $\frac{d}{dA} \text{tr}(BAC) = B^T C^T$,

$$\frac{\partial J}{\partial b} = 2 \left[b - E(x) \right] + 2A E(z) = 0 \quad (1.2.4-2a)$$

onde, P_x é a covariância de x ,

P_z é a covariância de z , e

P_{xz} é a covariância cruzada entre x e z ;

e

$$\frac{\partial J}{\partial A} = 2A \left\{ P_z + E(z) \cdot \left[E(z) \right]^T \right\} - 2 P_{xz} + 2 \left[b - E(x) \right] \cdot \left[E(z) \right]^T = 0 \quad (1.2.4-2b)$$

Então, de 1.2.4-2a resulta:

$$b = E(x) - AE(z)$$

e substituindo em 1.2.4.2b resulta:

$$A = P_{xz} P_z^{-1}$$

Utilizando A e b em 1.2.4-1 a estimação média quadrática linear é $\hat{x}_{MQL}$ é determinada da forma:

$$\hat{x}_{MQL} = E(x) + P_{xz} \cdot P_z^{-1} [z - E(z)] \quad (1.2.4-5)$$

Podemos considerar $E(z)$ como uma estimação para a medida z e definimos o resíduo como:

$$\tilde{z} = z - E(z), \quad (1.2.4-6)$$

$E(x)$ como uma estimação à priori para x e, o segundo termo é uma correção baseada no resíduo. Se a medida for exatamente o valor estimado, ou seja, $z = E(z)$, então o valor estimado de x é o valor estimado à priori e $\hat{x} = E(x)$. A ponderação $P_{xx} P_{xx}^{-1}$ através da qual o erro de medida $[z - E(z)]$ é incorporado na estimação de x é dependente das estatísticas conjuntas de segunda ordem. Portanto, se P_{xx} é "muito pequeno", então x e z possuem um grau de dependência reduzido e a medida de z traz pouca informação sobre x .

Se x e z são independentes, então $P_{xx} = 0$, de forma que a melhor estimação de x para este caso é o seu valor médio.

Se P_{xx} é "muito grande", então confiamos pouco na medida z e portanto o resíduo $[z - E(z)]$ deve ser menos ponderado na estimação de x .

Compare as equações 1.2.4-5 e 1.2.2-3. Para encontrar a estimação média quadrática generalizada $\hat{x}_{MQL}$, todas as estatísticas conjuntas de x e z são necessárias desde que a função densidade de probabilidade (ou igualmente f_{xx} e f_z) é necessária, enquanto que a estimação linear x_{MQL} só requer o conhecimento das estatísticas de primeira (média) e segunda (covariância) ordem.

Verificando a polarização da estimação, escrevemos:

$$\begin{aligned} E \left\{ \hat{x}_{MQL} \right\} &= E \left\{ E(x) \right\} + P_{xx} P_{xx}^{-1} E [z - E(z)] \\ &= E(x) + P_{xx} P_{xx}^{-1} [E(z) - E(z)] \\ &= E(x) \end{aligned}$$

A matriz de covariância do erro de estimação pode ser escrita como:

$$\begin{aligned} P_{\hat{x}} &= E \left\{ (\hat{x} - x) (\hat{x} - x)^T \right\} \\ &= E \left\{ \left[x - E(x) \right] - P_{xz} P_z^{-1} \left[z - E(z) \right] \right\} \cdot \left\{ \left[x - E(x) \right] - P_{xz} P_z^{-1} \left[z - E(z) \right] \right\}^T \\ &= P_x - P_{xz} P_z^{-1} P_{zx} - P_{xz} P_z^{-1} P_{zx} + P_{xz} P_z^{-1} P_{zx} P_z^{-1} P_{zx} \end{aligned}$$

$$P_{\hat{x}} = P_x - P_{xz} P_z^{-1} P_{zx} \quad (1.2.4-7)$$

Na equação 1.2.4-7, P_x representa a covariância à priori do erro e o segundo termo representa uma correção devido à incerteza da medida.

Suponhamos o caso em que x e z fossem processos Gaussianos relacionadas por um modelo de medida Gaussiana linear

$$z = Hx + v$$

onde $x \sim N(E(x), P_x)$, $v \sim N(E(v), R)$, H a matriz de observação conhecida e constante, $R > 0$, e x e v ortogonais, ou seja, $E(x \cdot v^T) = 0$.

Seja $x \in \mathbb{R}^n$, $v \in \mathbb{R}^m$, tal que H é uma matriz $p \times n$ e $z \in \mathbb{R}^p$, então a estimação com erro médio quadrático é linear, visto que a média condicional $E(x/z)$ e a covariância condicional $P(x/z)$ para este caso são idênticas às equações 1.2.4-5 e 1.2.4-7. De forma que o valor estimado de x , $\hat{x} = E(x/z)$ é a projeção ortogonal de x no espaço linear das observações passadas [2].

Em outras palavras, a estimação ótima pode ser considerada como a saída de um filtro linear, onde as entradas são os valores das variáveis aleatórias observáveis que ocorrem durante o processo de filtragem.

1.3. OBSERVADOR DE ESTADO DETERMINÍSTICO [7]

Para introduzir as idéias básicas utilizadas no filtro de Kalman discreto, falaremos um pouco do observador de estado determinístico.

Admitamos a seguinte planta dinâmica discreta:

$$x_{k+1} = A x_k + B u_k, \quad \hat{x}(0) = \hat{x}_0 \quad (1.3-1)$$

$$y_k = H x_k \quad (1.3-2)$$

onde o estado $x \in \mathbb{R}^n$, a entrada de controle $U \in \mathbb{R}^m$, a saída $Z \in \mathbb{R}^r$ e A , B e H são matrizes constantes conhecidas de dimensão apropriada. Todas as variáveis são determinísticas, tal que se o estado inicial x_0 for conhecido então 1.3-1 e 1.3-2 podem ser resolvidas para $k \geq 0$.

O problema de estimação resume-se em projetar um estimador, cuja saída $\hat{x}_k$ converge com k para um estado atual x_k de 1.3-1 quando o estado inicial x_0 é desconhecido, mas u_k e y_k são exatamente conhecidos.

Um método de estimação de estado que podemos estabelecer é construir um modelo da equação dinâmica da planta

$$\hat{x}_{k+1} = A \hat{x}_k + B u_k \quad (1.3-3)$$

onde $\hat{x}$ denota a estimação do estado atual x .

Conhecidas A , B e $U(k)$, a estimação se desenvolveria corretamente se fossem também conhecidas precisamente as condições iniciais $x(0)$ e ajustássemos $\hat{x}(0)$ de forma idêntica.

Se definirmos o erro na estimação como

$$\tilde{x} = x - \hat{x} \quad (1.3-4)$$

então a dinâmica do erro é descrita por:

$$\tilde{\mathbf{x}}_k = \mathbf{A} \tilde{\mathbf{x}}_{k-1} \quad (1.3-5)$$

Portanto, se $\tilde{\mathbf{x}}(0)$ é desconhecido, então a dinâmica do erro é a dinâmica da planta não compensada (malha aberta) como mostrada na Figura (1.2).

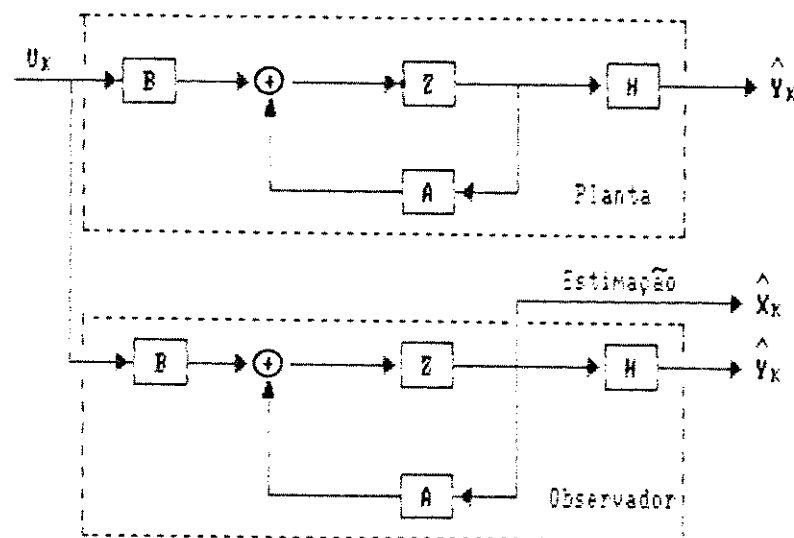


Figura 1.2 - Observador de Estado em Malha Aberta

Entretanto, se o sistema for instável (dependendo dos autovalores de $\mathbf{A}$), então o erro $\tilde{\mathbf{x}}(k+1)$ tornar-se-á arbitrariamente grande quando k tender para infinito, não importando o quão pequeno seja o erro de estimação do estado inicial.

Mesmo se o sistema fosse estável, mas alguns autovalores tivessem partes reais que são muito pequenas, o efeito do erro na estimação inicial levaria muito tempo para ser eliminado.

Visto que a dinâmica do erro representada pela equação 1.3-5 é indesejável, um procedimento clássico proposto para se conseguir a dinâmica desejada, ou seja, anular o erro da estimação, resume-se em considerar na dinâmica do modelo do sistema uma realimentação do erro na estimação através de um termo proporcional. Mas, como determinar esse erro $x(.) - \hat{x}(.) = \tilde{x}(.)$, se $x(.)$ não nos é disponível?

A saída $y(.)$ do sistema não utilizada na solução em malha aberta, está relacionada com a quantidade $x(.)$, que nos interessa da seguinte forma: $y(.) = H x(.)$ e $y(.)$ é uma medida a que se tem acesso.

Portanto, o sinal de erro pode ser gerado como sendo:

$$y_k - \hat{y}_k = y_k - H \hat{x}_k = H \left[x_k - \hat{x}_k \right] = H \tilde{x}_k$$

e pode ser utilizada para produzir a equação do estimador da forma (Ver Figura 1.3)

$$\hat{x}_{k+1} = A \hat{x}_k + B U_k + G[y_k - H \hat{x}_k], \quad \hat{x}(0) = \hat{x}_0 \quad (1.3-6)$$

onde

$\hat{x}_0$ = um valor do estado inicial estimado

e

G = um vetor do ganho de realimentação, adequadamente escolhido.

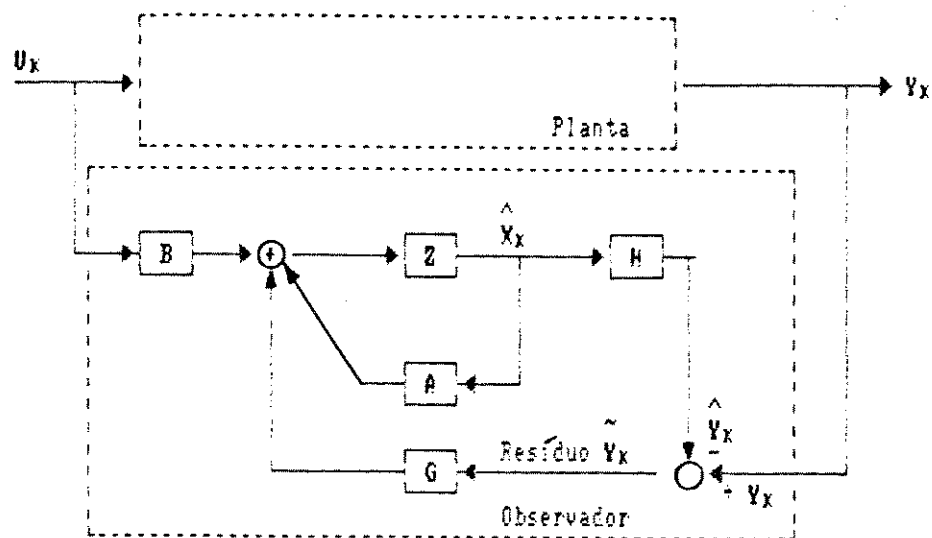


Figura 1.3 - Observador de Estado em Malha Fechada

O papel da realimentação do erro $G [y(.) - \hat{y}(.)]$ na equação 1.3-6 é de nos fornecer um controle sobre o comportamento do erro $\tilde{x}(.)$.

As matrizes A , B e H são dadas de forma que devemos encontrar a matriz G tal que o erro de estimação $\tilde{x}_k = x_k - \hat{x}_k$ vá para zero com k para todo x_0 .

Note que o observador consiste de duas partes, das matrizes A , B e H que representam o modelo da planta e uma parcela de correção de erro descrita por $G(y_k - \hat{y}_k) = G \tilde{y}_k$, onde a estimação de saída é $\hat{y}_k = H \hat{x}_k$ e $\tilde{y}_k = y_k - \hat{y}_k$ o resíduo.

Para escolher um ganho G que minimize o erro de estimação $\tilde{x}_k = x_k - \hat{x}_k$, é necessário examinar a dinâmica do erro de estimação $\tilde{x}_k$.

Sabemos que,

$$\begin{aligned}\tilde{x}_{k+1} &= x_{k+1} - \hat{x}_{k+1} \\ &= Ax_k + Bu_k - \left\{ A\hat{x}_k + G \left[y_k - H\hat{x}_k \right] + Bu_k \right\} \\ &= A \left[x_k - \hat{x}_k \right] - G \left[Hx_k - H\hat{x}_k \right]\end{aligned}$$

Resultando que

$$\tilde{x}_{k+1} = (A-GH)\tilde{x}_k \quad (1.3-7)$$

descreve a dinâmica do erro de estimação.

A equação 1.3-7 nos permite estudar as propriedades de convergência do observador. A condição inicial da equação 1.3-7 é $\tilde{x}_0 = x_0 - \hat{x}_0$, que é o erro inicial desconhecido. É possível observarmos que para que $\tilde{x}_k$ se anule com k para qualquer $\tilde{x}_0$, o ganho G do observador deve ser selecionado tal que $(A-GH)$ seja estável. Especificamos as localizações das raízes do estimador no plano z de forma que G é unicamente determinado, provido que y seja um escalar e o sistema seja observável.

Dois métodos podem ser empregados para a computação de G . O primeiro é expandir o determinante e igualar os coeficientes em potências de z dos dois membros de

$$[ZI - A + GH] = (Z-p_1) \dots (Z-p_n) \quad (1.3-8)$$

onde os p_i s são localizações desejadas das raízes do estimador e estabelecem a rapidez da convergência da estimação de estado para o valor do estado correto da planta.

O segundo é a utilização da fórmula de Ackermann [7]

$$G = \mathcal{L}e(A) \cdot \rho_n^{-1} \cdot e_n \quad (1.3-9)$$

onde $\mathcal{L}e(A)$ é o polinômio característico do sistema,

$$\rho_n = \begin{bmatrix} H \\ HA \\ \vdots \\ HA^{n-1} \end{bmatrix}$$

é a matriz observabilidade e e_n é a n-ésima coluna da matriz identidade $n \times n$.

A equação de estimação 1.3-6 fornece o valor estimado $\hat{x}(k+1)$ baseado nas medidas até o instante k . Isto significa que o valor estimado $\hat{x}(k+1)$ não depende do valor do erro nas observações no instante $(k+1)$, e então, do ponto de vista da aplicação de um controle por realimentação do estado, poderíamos não ter um controle tão preciso quanto desejaríamos. Pensando em sistemas de grande porte controlados por computador com reduzida velocidade de processamento e que exijam altas taxas de amostragem, o atraso entre fazer uma observação e processá-la para o controle pode resultar numa atuação insatisfatória. Portanto, formulou-se um estimador alternativo $\hat{x}(k)$ que se baseia na medida $y(k)$. Se temos uma estimação em k , $\hat{x}(k)$, então predizemos o próximo estado como sendo:

$$\bar{x}_{k+1} = A \hat{x}_k + B U_k, \quad (1.3-10)$$

Agora, no instante $k+1$, observamos a saída $y(k+1)$ e através dela corrigimos $\bar{x}$, de forma a resultar na seguinte equação de estimação:

$$\hat{x}_{k+1} = \bar{x}_{k+1} + G \left(Y_{k+1} - H \bar{x}_{k+1} \right) \quad (1.3-11)$$

Na prática, tal estimador não pode ser implementado com a devida precisão, já que é impossível amostrar $y(k+1)$ e utilizá-lo na estimação sem que haja algum atraso. No entanto, é uma das equações que constituem o filtro de Kalman (quando G minimiza a variância do erro de estimação), pois combinar a estimação prévia $\bar{x}(k+1)$ com a medida presente $y(k+1)$ foi uma das idéias básicas para a geração das equações deste filtro.

A equação do erro para o estimador representado pela equação 1.3-11 é similar à equação do erro 1.3-7 para o estimador preditor.

A equação do erro do estimador 1.3-11 é obtida, tomando-se $\tilde{x} = x - \hat{x}$, no que resulta em:

$$\tilde{x}_{k+1} = [A - GHA] \tilde{x}_k \quad (1.3-12)$$

Portanto, a matriz G de ganho é obtida exatamente como antes, exceto que H é substituído por HA . Os dois métodos representados pelas equações 1.3-8 e 1.3-9 são escritos na forma:

A) Igualando os coeficientes de:

$$|ZI - A + GHA| = \mathcal{Z}e(z)$$

onde $\mathcal{Z}e$ é o polinômio característico desejado do erro do estimador, ou

B) Utilizando a fórmula de Ackermann:

$$G = \mathcal{Z}e(A) \cdot \begin{bmatrix} HA^4 \\ HA^3 \\ \vdots \\ HA^n \end{bmatrix}^{-1} \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}$$

1.3.1. ASPECTOS ESTOCASTICOS

Consideremos a situação mostrada na Figura 1.4. A planta foi adicionado o ruído do processo ω e à saída um ruído v . Sob estas circunstâncias mais reais, o observador projetado deterministicamente como anteriormente mostrado, não fará com que $\tilde{x}(k) \rightarrow 0$ para $k \rightarrow \infty$. Entretanto, utilizando uma estrutura virtualmente idêntica à Figura 1.3, o erro pode ser levado a zero (mínimo) se forem convenientemente considerados os aspectos estocásticos.

Devemos da mesma forma encontrar o ganho do observador $G(k)$ variante no tempo, baseado nas considerações estocásticas.

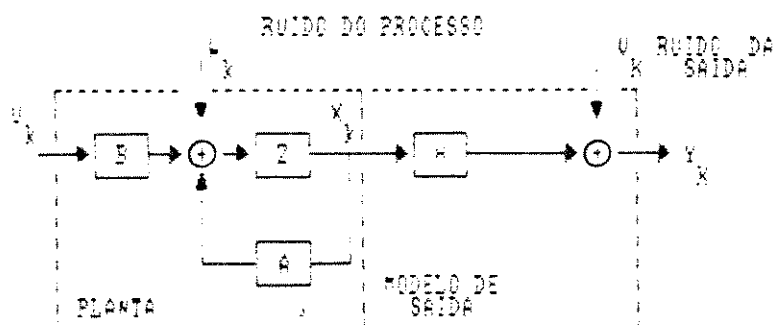


Figura 1.4 - Representação da Planta Estocástica

Para descrever o sistema da Figura 1.4, escrevemos a equação dinâmica estocástica discreta:

$$x_{k+1} = Ax_k + Bu_k + \Gamma\omega_k \quad (1.3.1-1a)$$

$$y_k = Hx_k + v_k \quad (1.3.1-1b)$$

onde $x_k \in \mathbb{R}^n$, $U(k) \in \mathbb{R}^m$, $\omega(k) \in \mathbb{R}^L$, $Y(k) \in \mathbb{R}^r$. Suponha o ruído do processo $\omega(k)$ com média $E\{\omega(k)\} = 0$ e covariância $E\{\omega(k) \omega^T(k)\} = Q$, representadas por $\omega(k) \sim (0, Q)$ e o ruído da saída da forma $V(k) \sim (0, R)$. Ambos os ruídos são supostamente processos ruído branco com suas funções de covariância satisfazendo

$$P_{\omega_{k1} \omega_{k2}} = E \left\{ \omega_{k1} \omega_{k2}^T \right\} = Q \delta(k1 - k2) \quad (1.3.1-2)$$

onde $\delta(k)$ é o delta de Kronecker; e a versão correspondente para v_k .

Suponha que o estado inicial $x(0)$ é desconhecido, mas é sabido à priori que $x(0)$ tem a forma $x(0) \sim [E(x_0), P_{x_0}]$. Suponhamos que $x(0)$, $\omega(k)$ e $v(k)$ são mutuamente não correlatos, tal que $E(\omega_k v_j^T) = 0$ para todo j e k .

P_{x_0} , Q e R são matrizes simétricas e semidefinida positivas. O sinal $u(k)$ é uma entrada de controle determinística.

Sob estas circunstâncias, o estado $x(k)$ é uma variável aleatória com média $E(x_k)$ e covariância P_{x_k} . Similarmente, a saída $y(k)$ é aleatória da forma $y(k) \sim [E(x_k), P_{x_k}]^{y_k}$. Note que (x_k) e (z_k) são processos estocásticos, em geral não estacionários.

1.3.1.1. Propagação das Médias e Covariâncias

Verifiquemos como as médias e covariâncias de x_k e z_k propagam-se no tempo sob a influência da dinâmica da equação do estado 1.3.2-1a.

Para a média do estado escrevemos:

$$\begin{aligned} E(X_{k+1}) &= E(Ax_k + BU_k + \Gamma W_k) \\ &= AE(x_k) + BE(U_k) + \Gamma E(W_k), \end{aligned}$$

ou,

$$E(X_{k+1}) = AE(x_k) + BU \quad (1.3.1.1-1)$$

A condição inicial $E(x_0)$ de 1.3.1.1-1 é dada.

Portanto, é possível notar que a média propaga-se de acordo com a dinâmica determinística 1.3.1.1-1.

Verificando como as covariâncias do estado propagam-se, escreveremos:

$$\begin{aligned} P_{x_{k+1}} &= E \left\{ \left[x_{k+1} - E(x_{k+1}) \right] \left[x_{k+1} - E(x_{k+1}) \right]^T \right\} \\ &= E \left\{ \left[A(x_k - E(x_{k+1})) + \Gamma \omega_k \right] \left[A(x_k - E(x_{k+1})) + \Gamma \omega_k \right]^T \right\} \\ &= A \cdot E \left\{ \left[x_k - E(x_{k+1}) \right] \cdot \left[x_k - E(x_{k+1}) \right]^T \right\} \cdot A^T + \\ &+ \Gamma E \left\{ \omega_k \left[x_k - E(x_k) \right]^T \right\} A^T + A \cdot E \left\{ \left[x_k - E(x_k) \right] \right\} \cdot \\ &\cdot \omega_k^T \Gamma^T + \Gamma \cdot E \left\{ \omega_k \omega_k^T \right\} \cdot \Gamma^T, \end{aligned}$$

ou

$$P_{x_{k+1}} = A P_{x_k} A^T + \Gamma P_{\omega_k x_k} A^T + A P_{x_k \omega_k} \cdot \Gamma^T + \Gamma Q \Gamma^T \quad (1.3.1.1-2)$$

Portanto,

$$P_{x_{k+1}} = A P_{x_k} A^T + \Gamma Q \Gamma^T \quad (1.3.1.1-3)$$

Os segundo e terceiro termos da equação (1.3.1.1-2) se anulam devido aos seguintes argumentos: o estado x depende de x e do ruído do processo ω_j para $j \leq k-1$, porém ω_k não é correlato com x e com todos os ω_j ($j \leq k-1$), portanto $P_{\omega_k x_k} = 0$.

A equação 1.3.1.1-3 é uma equação de Lyapunov para P_{x_k} . A condição inicial é P_{x_0} , a qual é dada.

Com relação a média e covariância de saída, teremos:

$$E \{y_k\} = H \cdot E \{x_k\}$$

A covariância cruzada entre estado e saída é

$$\begin{aligned} P_{x_k y_k} &= E \left\{ \begin{bmatrix} x_k - E(x_k) \end{bmatrix} \begin{bmatrix} y_k - E(y_k) \end{bmatrix}^T \right\} \\ &= E \left\{ \begin{bmatrix} x_k - E(x_k) \end{bmatrix} \begin{bmatrix} H(x_k - E(x_k)) + v_k \end{bmatrix}^T \right\}, \end{aligned}$$

ou

$$P_{x_k y_k} = P_{x_k} H^T \quad (1.3.1.1-4)$$

devido a não correlação entre x_k e v_k .

Para a covariância da saída,

$$\begin{aligned} P_{y_k} &= E \left\{ \begin{bmatrix} y_k - E(y_k) \end{bmatrix} \begin{bmatrix} y_k - E(y_k) \end{bmatrix}^T \right\} \\ &= E \left\{ \begin{bmatrix} H(x_k - E(x_k)) + v_k \end{bmatrix} \begin{bmatrix} H(x_k - E(x_k)) + v_k \end{bmatrix}^T \right\} \end{aligned}$$

Portanto,

$$P_{y_k} = H P_{x_k} H^T + R, \quad (1.3.1.1-5)$$

onde mais uma vez foram utilizadas as suposições de não correlação.

Neste momento, é de fundamental importância notar que x_k e z_k são variáveis aleatórias conjuntamente distribuídas com momentos dados por

$$\begin{aligned} \begin{bmatrix} x_k \\ y_k \end{bmatrix} &\sim \begin{bmatrix} E(x_k) \\ E(y_k) \end{bmatrix}, \begin{bmatrix} P_{x_k} & P_{x_k w_k} \\ P_{w_k x_k} & P_{w_k} \end{bmatrix} \\ &= \begin{bmatrix} E(x_k) \\ H E(x_k) \end{bmatrix}, \begin{bmatrix} P_{x_k} & P_{x_k} H^T \\ H P_{x_k} & H P_{x_k} H^T + R \end{bmatrix} \end{aligned} \quad (1.3.1.1.-6)$$

onde $E(x_k)$ e P_{x_k} são determinadas recursivamente por 1.3.1.1-1 e 1.3.1.1-3. Se a sequência y_k for medida e a variável aleatória relacionada é desconhecida, então podemos aplicar os resultados obtidos do item 1.2.2 ou 1.2.4 para encontrar uma estimação $\hat{x}_k$. Neste caso, poderíamos supor estatísticas arbitrárias e conhecer os momentos de primeira e segunda ordem para x_k , w_k e v_k , de forma que o estimador mínimos quadrados linear pudesse ser utilizado.

Todas as derivações vistas até agora têm sido para o caso de estatísticas gerais. Se fizermos as suposições que x_k , w_k e v_k são processos Gaussianos, então desde que o sistema 1.3.1-1 é linear, x_k e y_k são também processos Gaussianos para todo k .

Dessa forma, os primeiro e segundo momentos especificam completamente a função densidade de probabilidade conjunta, tal que a informação completa sobre a interdependência do estado e a saída é dada por 1.3.1.1-6.

Se todas as estatísticas são Gaussianas, então x_k é descrito pela função densidade da probabilidade condicional.

$$f_{x_k}(x, k) = \frac{1}{\sqrt{(2\pi)^n |P_{x_k}|}} \exp \left\{ -\frac{1}{2} [x - E(x)]^T \cdot P_{x_k}^{-1} [x - E(x_k)] \right\}$$

dados $E(x_k)$, P_{x_k} e as entradas $U_0, U_1, \dots, U_{k-1}$, onde n é a ordem de x_k .

1.4. FILTRO DE KALMAN DISCRETO VIA TEOREMA DA PROJEÇÃO ORTOGONAL

Neste item, mostraremos o desenvolvimento do Filtro de Kalman discreto com base no teorema da projeção ortogonal. A motivação para a versão discreta do Filtro de Kalman foi, em parte, devida ao desenvolvimento da teoria de sistemas a dados amostrados durante a década de 50 como um resultado do uso de computadores digitais nas áreas da comunicação e controle.

Consideraremos o desenvolvimento de um algoritmo de mínima variância do erro não polarizado para a estimação discreta no tempo.

Consideremos a equação a diferença vetorial:

$$X_{k+1} = A(k+1, k) X_k + \Gamma_k W_k \quad (1.4-1)$$

e a equação de observação:

$$Y_k = H_k X_k + V_k \quad (1.4-2)$$

onde:

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

$$E \left\{ x_{k_0} \right\} = \bar{x}_{k_0}$$

$$\text{Var} \left\{ x_{k_0} \right\} = P_{k_0}$$

$$E \left\{ \begin{matrix} w_k & w_j^T \end{matrix} \right\} = Q_k \delta_{k-j}, \quad \delta \text{ é o delta do Kronecker}$$

e,

$$\text{COV} \left\{ v_k, x_{k_0} \right\} = \text{COV} \left\{ w_k, x_{k_0} \right\} = \text{COV} \left\{ v_k, w_j \right\} = 0$$

w_k e v_k são ruídos brancos Gaussianos

x_{k_0} é uma distribuição Gaussiana

O modelo para nosso problema é mostrado na Figura 1.5. Veremos que os algoritmos gerados descrevem a propagação da média e covariância de x_k .

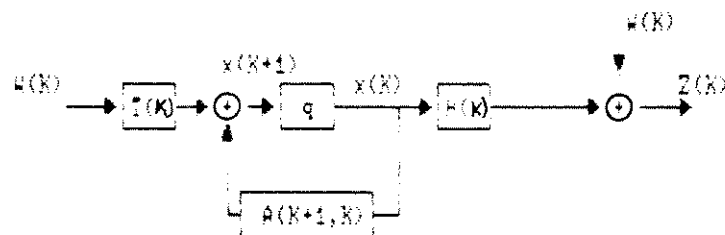


Figura 1.5 - Modelos da Planta e Observação para a Estimação Ótima Discreta.

A variância mínima do erro linear de x , dado um subespaço de observação linear Y pertencente ao espaço de Hilbert H , é dada pela projeção ortogonal de x sobre Y , isto é, $\hat{x} = E \{ x/Y \}$ [6].

Seja a sequência ortogonal $\{ y_i \}_{i=1}^m$ formando uma base de Y . Podemos escrever a projeção de x no subespaço Y como:

$$E \{ x/Y \} = \sum_{i=1}^m E \left\{ x y_i^T \right\} \left\{ E \left[y_i y_i^T \right] \right\}^{-1} y_i \quad (1.4-3)$$

Desse mesmo resultado segue que, para um dado $y^* \in H$ ortogonal a Y , isto é, $E \{ Y^T y^* \} = 0$, então,

$$E \left\{ x/Y, y^* \right\} = E \left\{ x/Y \right\} + E \left\{ x/y^* \right\} \quad (1.4-4)$$

Procedemos por indução no desenvolvimento de uma equação recursiva para $\hat{x}_{k+1/k}$ supondo que $\hat{x}_{k/k-1}$ é conhecido e calculamos $\hat{x}_{k+1/k}$ em termos de $\hat{x}_{k/k-1}$ e y_k . Visando a utilização da equação 1.4-4, devemos determinar um y^* que seja ortogonal a Y_{k-1} .

Como a inovação do processo

$$\tilde{y}_k = y_k - H_k \hat{x}_{k/k-1}$$

é ortogonal a Y_{k-1} e os espaços vetoriais lineares gerados por $\{ y_k, Y_{k-1} \}$ e $\{ \tilde{y}_k, Y_{k-1} \}$ são idênticos, então com base na equação 1.4-4, podemos escrever que

$$\begin{aligned}
\hat{x}_{k+1/k} &= E \{x(k+1) / Y(k)\} \\
&= E \left\{ x(k+1) / \left[\tilde{y}(k) \quad , \quad Y(k-1) \right] \right\} \\
&= E \{x(k+1) / Y(k-1)\} + E \left\{ x(k+1) / \tilde{y}(k) \right\}
\end{aligned}
\tag{1.4-5}$$

Segue das equações 1.4-1 e 1.4-2 e das suposições iniciais que:

$$E \left\{ x_{k+1} / Y_{k-1} \right\} = A_{k+1, k} \cdot \hat{x}_{k/k-1} \tag{1.4-6}$$

O primeiro termo do 2º membro da equação 1.4-5 pode ser representado em termos de $\hat{x}(k/k-1)$.

Desenvolvendo o segundo termo do 2º membro da equação 1.4-5, com base na equação 1.4-3, temos:

$$E \left\{ x(k+1) / \tilde{y}(k) \right\} = E \left\{ x(k+1) / \tilde{y}^T(k) \right\} \cdot \left\{ E \left[\tilde{y}(k) \tilde{y}^T(k) \right] \right\}^{-1} \cdot \tilde{y}(k) \tag{1.4-7}$$

Sabendo que $\tilde{y}(k) = H(k) \cdot \tilde{x}(k/k-1) + V(k)$, obtemos

$$\begin{aligned}
E \left\{ x(k+1) / \tilde{y}^T(k) \right\} &= E \left\{ [A(k+1, k) \cdot x(k) + \Gamma(k)\omega(k)] \cdot \right. \\
&\quad \left. [H(k)\tilde{x}(k/k-1) + V(k)]^T \right\} =
\end{aligned}$$

$$\begin{aligned}
A(k+1, k) &= E \left\{ x(k) \tilde{x}^T(k/k-1) \right\} \cdot H^T(k) + \\
A(k+1, k) &= E \left\{ x(k) \cdot V^T(k) \right\} + \Gamma(k) \cdot \\
E \left\{ \omega(k) \tilde{x}^T(k/k-1) \right\} &= H^T(k) + \Gamma(k) E \left\{ \omega(k) V^T(k) \right\} \\
&= A(k+1, k) \cdot E \left\{ x(k) \tilde{x}^T(k/k-1) \right\} H^T(k) ,
\end{aligned}
\tag{1.4-8}$$

onde todos os outros termos são zero devido às suposições sobre os ruídos.

Como

$$\tilde{x}(k+1/k) = x(k+1) - \hat{x}(k+1/k) ,$$

então podemos escrever que:

$$\begin{aligned}
E \left\{ x(k) \tilde{x}^T(k/k-1) \right\} &= E \left\{ \hat{x}(k/k-1) \tilde{x}^T(k/k-1) \right\} + E \left\{ \tilde{x}(k/k-1) \cdot \right. \\
&\quad \left. \tilde{x}^T(k/k-1) \right\} = E \left\{ \tilde{x}(k/k-1) \tilde{x}^T(k/k-1) \right\}
\end{aligned}
\tag{1.4-9a}$$

Já que $E \left\{ \hat{x}(k/k-1) \tilde{x}^T(k/k-1) \right\} = 0$ pelo teorema da projeção ortogonal.

Definindo,

$$P(k/k-1) = E \left\{ \tilde{x}(k/k-1) \cdot \tilde{x}^T(k/k-1) \right\}
\tag{1.4-9b}$$

podemos escrever a equação 1.4-8 como sendo:

$$E \left\{ x(k+1) / \tilde{y}^T(k) \right\} = A(k+1, k) P(k/k-1) \cdot H^T(k) \quad (1.4-10)$$

Similarmente ao desenvolvimento anterior, resulta que:

$$E \left\{ \tilde{y}(k) \tilde{y}^T(k) \right\} = H(k) P(k/k-1) H^T(k) + R(k) \quad (1.4-11)$$

Substituindo as equações 1.4-10 e 1.4-11 em 1.4-7, resulta que:

$$E \left\{ x(k+1) / \tilde{y}(k) \right\} = A(k+1, k) P(k/k-1) H^T(k) \cdot \left[H(k) P(k/k-1) H^T(k) + R(k) \right]^{-1} \tilde{y}(k) \quad (1.4-12)$$

E finalmente, substituindo 1.4-6 e 1.4-12 em 1.4-5, resulta:

$$\hat{x}(k+1/k) = A(k+1, k) \hat{x}(k/k-1) + G(k) \left[y(k) - H(k) \hat{x}(k/k-1) \right] \quad (1.4-13)$$

onde o ganho de Kalman $G(k)$ é dado por:

$$G(k) = A(k+1, k) P(k/k-1) H^T(k) \left[H(k) P(k/k-1) H^T(k) + R(k) \right]^{-1} \quad (1.4-14)$$

Vemos da equação 1.4-14 que o cálculo do ganho do filtro requer a determinação da matriz $P(k/k-1)$, que pode ser escrita recursivamente, como veremos a seguir.

é fácil verificarmos que $\tilde{x}(k+1/k)$ satisfaz a equação recursiva

$$\tilde{x}(k+1/k) = A(k+1,k) \tilde{x}(k/k-1) + \Gamma(k) \omega(k) - G(k) \tilde{y}(k) \quad (1.4-15)$$

Portanto, utilizando a definição de $P(k+1/k)$ e as equações 1.4-11, 1.4-13 e as equações precedentes, verificamos que $P(k+1/k)$ satisfaz a equação recursiva

$$\begin{aligned} P(k+1/k) = & A(k+1,k) P(k/k-1) A^T(k+1,k) + \Gamma(k) Q(k) \Gamma^T + \\ & G(k) \left[H(k) P(k/k-1) H^T(k) + R(k) \right] G^T(k) - \\ & A(k+1,k) P(k/k-1) H^T(k) G^T(k) - \\ & G(k) H(k) P(k/k-1) A^T(k+1,k) \end{aligned} \quad (1.4-16)$$

Há muitas formas de descrever a recursividade 1.4-16. Por exemplo, da substituição de $G(k)$, definido na equação 1.4-14, na equação 1.4-16 resulta, depois de algumas manipulações algébricas, que:

$$\begin{aligned} P(k+1/k) = & A(k+1,k) P(k/k-1) A^T(k+1,k) - A(k+1,k) P(k/k-1) \\ & H^T(k) \left[H(k) P(k/k-1) H^T(k) + R(k) \right]^{-1} \cdot \\ & H(k) P(k/k-1) A(k+1,k) + \Gamma(k) Q(k) \Gamma^T(k) \end{aligned} \quad (1.4-17)$$

conhecida como a equação de Riccati discreta.

A equação 1.4-13 representa uma descrição recursiva de um preditor de variância mínima linear, o qual requer o conhecimento da covariância do erro à priori $P(k+1/k)$, representado pelas equações 1.4-16 ou 1.4-17. No entanto, o cálculo recursivo de $P(k+1/k)$ pode ser feito "on line", dado que não exige o conhecimento da sequência de observações $Y(k)$. Já o cálculo do preditor ótimo requer o conhecimento direto das observações.

As condições iniciais das equações 1.4-13 e 1.4-17 são dadas por:

$$\hat{x}(k_0+1/k_0) = A(k_0+1, k_0) \bar{x}(k_0)$$

e

$$P(k_0+1/k_0) = A(k_0+1, k_0) P(k_0) A^T(k_0+1, k_0) + \Gamma(k_0) Q(k_0) \Gamma^T(k_0)$$

Entretanto, o que nos interessa é a determinação de $\hat{x}(k/k) = \hat{x}(k) = E \{x(k) / Y(k)\}$ para o preditor $\hat{x}(k+1/k)$.

Pela propagação da média de x é possível escrevermos que:

$$E\{x(k+1)\} = \bar{x}(k+1) = A(k+1, k) \bar{x}(k)$$

e, portanto, originar a seguinte relação:

$$\hat{x}(k+1/k) = A(k+1, k) \hat{x}(k/k) \quad (1.4-18)$$

Substituindo a equação 1.4-18 na equação 1.4-13 e pré-multiplicando ambos os membros por $A^{-1}(k+1, k)$, teremos:

$$\hat{\mathbf{x}}(k/k) = \hat{\mathbf{x}}(k/k-1) + \mathbf{G}^*(k) \left[\mathbf{Y}(k) - \mathbf{H}(k) \hat{\mathbf{x}}(k/k-1) \right]$$

onde:

$$\mathbf{G}^*(k) = \mathbf{P}(k/k-1) \mathbf{H}^T(k) \cdot \left[\mathbf{H}(k) \mathbf{P}(k/k-1) \mathbf{H}^T(k) + \mathbf{R}(k) \right]^{-1}$$

com a condição inicial $\hat{\mathbf{x}}[k_0] = \bar{\mathbf{x}}[k_0]$ e a respectiva covariância do erro:

$$\mathbf{P}(k/k) = E \left\{ \tilde{\mathbf{x}}(k/k) \tilde{\mathbf{x}}^T(k/k) \right\}$$

$$\mathbf{P}(k/k) = \mathbf{P}(k/k-1) - \mathbf{P}(k/k-1) \mathbf{H}^T(k) \left[\mathbf{H}_k \mathbf{P}(k/k-1) \mathbf{H}_k^T + \mathbf{R}_k \right]^{-1} \mathbf{H}_k \mathbf{P}(k/k-1)$$

ou ainda

$$\mathbf{P}(k/k) = \left[\mathbf{I} - \mathbf{G}^*(k) \mathbf{H}(k) \right] \mathbf{P}(k/k-1)$$

As equações do filtro podem ser representadas como mostrado na Figura 1.6.

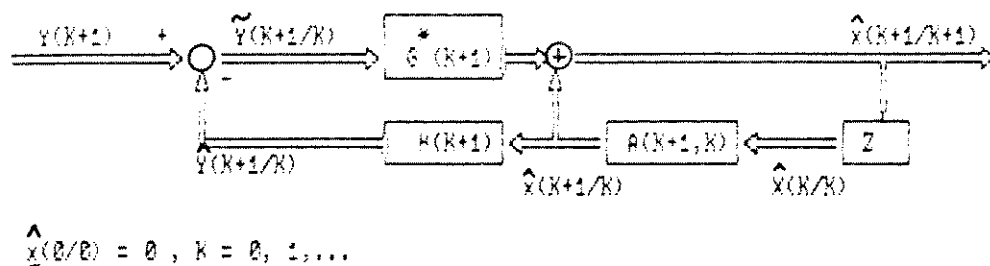


Figura 1.6 - Diagrama de Blocos do Filtro de Kalman.

1.5. FILTRO DE KALMAN DISCRETO VIA TEOREMA DE BAYES

Consideremos ainda as equações 1.4-1, 1.4-2 e suas respectivas suposições.

Nesta formulação do problema de filtragem, a estimação ótima é realizada pela determinação de $x(k)$, que maximiza a função densidade de probabilidade

$$P [x(k)/Y(k)] , \text{ isto é, calculando } \left. \frac{\partial p [x(k)/Y(k)]}{\partial x(k)} \right|_{x(k) = \hat{x}(k) = 0}$$

e demonstrando que:

$$\left. \frac{\partial}{\partial x(k)} \left\{ \frac{\partial p [x(k)/Y(k)]}{\partial x(k)} \right\} \right|_{x(k) = \hat{x}(k)} < 0 \quad (1.5-1)$$

A satisfação da equação 1.5-1, a qual requer que a matriz resultante da derivada segunda seja definida negativa, garante que a função $p [x(k)/Y(k)]$ foi maximizada.

Esse método de estimação é conhecido como estimação Máximo à posteriori (MAP) ou estimação máxima verossimilhança à posteriori.

Para encontrarmos a expressão para $p [x(k)/Y(k)]$, utilizamos o teorema de Bayes, o qual estabelece que:

$$p [x(k)/Y(k)] = \frac{p [x(k), Y(k)]}{p [Y(k)]} \quad (1.5-2)$$

Como $Y(k)$ é a combinação da nova observação $y(k)$ e das observações passadas $Y(k-1)$, então a equação 1.5-2 torna-se:

$$p [x(k)/Y(k)] = \frac{p [x(k) , y(k) , Y(k-1)]}{p [y(k) , Y(k-1)]} \quad (1.5-3)$$

Ainda utilizando o teorema de Bayes, podemos escrever o numerador desta expressão como:

$$\begin{aligned} p [x(k), y(k), Y(k-1)] &= p [y(k)/x(k), Y(k-1)] p [x(k), Y(k-1)] \\ &= p [y(k)/x(k), Y(k-1)] p [x(k)/Y(k-1)] p[Y(k-1)] \end{aligned}$$

Resultando que:

$$p [x(k), y(k), Y(k-1)] = p [y(k)/x(k)] p [x(k)/Y(k-1)] p [Y(k-1)] \quad (1.5-4)$$

pelo fato de o conhecimento de $x(k)$ dispensar a sequência $Y(k-1)$.

Com relação ao denominador da expressão 1.5-2, podemos escrever com base no mesmo teorema que:

$$p [y(k) , Y(k-1)] = p [y(k)/Y(k-1)] \cdot p [Y(k-1)]$$

E, portanto, a função densidade da probabilidade condicional torna-se:

$$p [x(k)/Y(k)] = \frac{p [y(k)/x(k)] p [x(k)/Y(k-1)]}{p [y(k) / Y(k-1)]} \quad (1.5-5)$$

Para a determinação da densidade de probabilidade de $x(k)$, dado $Y(k)$, devemos calcular cada uma das expressões de probabilidade da equação 1.5-5.

Consideraremos cada termo separadamente, mostrando que uma das densidades de probabilidade é Gaussiana e então, encontrados os dois primeiros momentos que caracterizam a distribuição.

Examinemos primeiramente $p[y(k)/x(k)]$. Visto que $y(k)$ é dado por $y(k) = H(k) x(k) + V(k)$ e V é Gaussiano, a densidade de probabilidade $p[y(k)/x(k)]$ é Gaussiana ($y(k)$ é composto de um processo Gaussiano mais uma constante $H(k) x(k)$).

A média do processo é

$$E[y(k)/x(k)] = E[H(k) x(k) + v(k)/x(k)] = H(k) x(k), \quad (1.5-6)$$

pois $E[v(k)/x(k)] = 0$. A variância do processo é

$$\text{Var}[y(k)/x(k)] = E\left\{ [y(k) - H(k) x(k)] [y(k) - H(k) x(k)]^T \right\} = R(k) \quad (1.5-7)$$

Portanto, a densidade de probabilidade $p[y(k)/x(k)]$ é dada por:

$$p[y(k)/x(k)] = \left\{ (2\pi)^m \det[R(k)] \right\}^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2} [y(k) - H(k) x(k)]^T R^{-1}(k) [y(k) - H(k) x(k)] \right\} \quad (1.5-8)$$

Utilizando a equação de observação do modelo, podemos escrever o denominador de 1.5-5 como:

$$p [y(k)/Y(k-1)] = p [H(k) x(k) + v(k)/Y(k-1)]$$

Da especificação original do problema, sabemos que $v(k)$ é uma distribuição Gaussiana e é independente de $Y(k-1)$. Como $p [x(k)/Y(k-1)]$ é Gaussiana para todo $k \geq 0$, então $p [y(k)/Y(k-1)]$ é também Gaussiana para todo $k \geq 0$, pelo fato de $y(k)$ ser uma função linear de distribuições Gaussianas.

A média de $p [y(k)/Y(k-1)]$ é dada por:

$$\begin{aligned} E [y(k)/Y(k-1)] &= H(k) E [x(k)/Y(k-1)] + E [v(k)/Y(k-1)] \\ &= H(k) \hat{x}(k/k-1) \end{aligned}$$

A variância do processo é

$$\begin{aligned} \text{Var} [y(k)/Y(k-1)] &= E \left\{ \left[y(k) - H(k) \hat{x}(k/k-1) \right] \right. \\ &\quad \left. \left[y(k) - H(k) \hat{x}(k/k-1) \right]^T \right\} \end{aligned}$$

Se substituirmos $y(k)$ e utilizarmos a notação $\tilde{x}(k/k-1) = x(k) - \hat{x}(k/k-1)$, teremos:

$$\text{Var} [y(k)/Y(k-1)] = H(k) \text{var} \left[\tilde{x}(k/k-1) \right] H^T(k) + \text{var} [v(k)]$$

tal que:

$$\text{Var} [y(k)/Y(k-1)] = H(k) P(k/k-1) H^T + R(k)$$

A densidade de probabilidade de $y(k)$ dado $Y(k-1)$ é portanto

$$p [y(k)/Y(k-1)] = \left\{ \left(2\pi \right)^m \det \left[H(k) P(k/k-1) H^T(k) + R(k) \right] \right\}^{-\frac{1}{2}} \\ \exp \left\{ -\frac{1}{2} \left[y(k) - H(k) \hat{x}(k/k-1) \right]^T \right. \\ \left. \left[H(k) P(k/k-1) H^T(k) + R(k) \right]^{-1} \left[y(k) - H(k) \hat{x}(k/k-1) \right] \right\} \quad (1.5-9)$$

onde m é a dimensão do vetor de saída.

Como $p [x(k)/Y(k-1)]$ é Gaussiano, então só precisamos encontrar a média e variância.

A média é exatamente a solução da predição e é dada pela equação 1.4-18:

$$E [x(k)/Y(k-1)] = A(k, k-1) \hat{x}(k-1/k-1) = \hat{x}(k/k-1)$$

Também do nosso desenvolvimento anterior, sabemos que a variância de $x(k)$, dado $Y(k-1)$ é a variância do erro à priori $P(k/k-1)$.

Portanto, temos da equação 1.4-9b,

$$\text{Var} [x(k)/Y(k-1)] = A(k, k-1) P(k-1/k-1) A^T(k, k-1) \\ + \Gamma(k-1) R(k-1) \Gamma^T(k-1)$$

Dessa forma, a densidade de probabilidade $p [x(k)/Y(k-1)]$ pode ser escrita

$$p [x(k)/Y(k-1)] = \left\{ (2\pi)^N \det [P(k/k-1)] \right\}^{-\frac{1}{2}} \exp \left\{ -\frac{1}{2} [x(k) - \hat{x}(k/k-1)]^T P^{-1}(k/k-1) [x(k) - \hat{x}(k/k-1)] \right\} \quad (1.5-10)$$

onde N é a dimensão do vetor $x(k)$.

Finalmente, calculados os termos, podemos calcular $p [x(k)/Y(k)]$, substituindo 1.5-8, 1.5-9 e 1.5-10 em 1.5-5, teremos:

$$p[x(k)/Y(k)] = \left\{ (2\pi)^N \det [R(k)] \left\{ \det [H(k) P(k/k-1) + H^T(k) + R(k)] \right\}^{-1} \cdot \det [P(k/k-1)] \right\}^{-\frac{1}{2}} \cdot \exp \left\{ -\frac{1}{2} [y(k) - H(k) x(k)]^T R_k^{-1}(k) [y(k) - H(k) x(k)] + [x(k) - \hat{x}(k/k-1)]^T P^{-1}(k/k-1) [x(k) - \hat{x}(k/k-1)] - [y(k) - H(k) \hat{x}(k/k-1)]^T [H(k) P(k/k-1) H^T(k) + R(k)]^{-1} [y(k) - H(k) \hat{x}(k/k-1)] \right\} \quad (1.5-11)$$

Esta expressão pode ser escrita

$$p [x(k)/Y(k)] = A \exp \left\{ \frac{1}{2} [x(k) - \hat{x}(k/k)]^T P^{-1}(k/k) [x(k) - \hat{x}(k/k)] \right\}$$

onde A não é uma função de $x(k)$,

$$\hat{x}(k/k) = E [x(k)/Y(k)] = \hat{x}(k/k-1) + G^* [y(k) - H(k) \hat{x}(k/k-1)] \quad (1.5-12)$$

$$\begin{aligned} G^*(k) &= p(k/k-1) H^T(k) [H(k) P(k/k-1) H^T(k) + R(k)]^{-1} \\ &= p(k/k) H^T(k) R^{-1}(k), \end{aligned}$$

$$\begin{aligned} P(k/k) &= \text{VAR} \left\{ \tilde{x}(k) \right\} = [I - G^*(k) H(k)] P(k/k-1) \\ &= P(k/k-1) - P(k/k-1) H^T(k) [H(k) P(k/k-1) H^T(k) + R(k)]^{-1} \\ &\quad H(k) P(k/k-1) = [H^T(k) R^{-1}(k) H(k) + P^{-1}(k/k-1)]^{-1} \end{aligned}$$

e,

$$A = \left\{ (2\pi)^N \det [P(k/k)] \right\}^{-\frac{1}{2}} = \left\{ \frac{(2\pi)^N \det [R(k)] \det [P(k/k-1)]}{\det [H(k) P(k/k-1) H^T(k) + R(k)]} \right\}^{-\frac{1}{2}}$$

Estamos interessados em encontrar o valor de $\hat{x}(k)$ que maximiza $p[x(k)/Y(k)]$, ou seja:

$$\left. \frac{\partial}{\partial x(k)} \left\{ \ln p[x(k)/Y(k)] \right\} \right|_{x(k) = \hat{x}_{\text{MAP}(k)}} = 0 \quad (1.5-13)$$

Já que $\ln p[x(k)/Y(k)]$ alcançará seu valor máximo para o mesmo valor de $\hat{x}_{\text{MAP}}$.

Substituindo a equação 1.5-11 em 1.5-13, resulta:

$$H^T(k) R^{-1}(k) \left[y(k) - H(k) \hat{x}_{MAP}(k) \right] - \\ P^{-1}(k/k-1) \cdot \left[\hat{x}_{MAP}(k) - \hat{x}(k/k-1) \right] = 0$$

Agrupando os termos envolvendo $\hat{x}_{MAP}(k)$, obtemos:

$$\left[H^T(k) R^{-1}(k) H(k) + P^{-1}(k/k-1) \right] \cdot \hat{x}_{MAP}(k) = \\ P^{-1}(k/k-1) \hat{x}(k/k-1) + H^T(k) R^{-1}(k) y(k),$$

resultando finalmente que:

$$\hat{x}_{MAP}(k) = \left[H^T(k) \cdot R^{-1}(k) H(k) + P^{-1}(k/k-1) \right]^{-1} \cdot \\ \left[P^{-1}(k/k-1) \cdot \hat{x}(k/k-1) \right] + H^T(k) R^{-1}(k) y(k) \quad (1.5-14)$$

Aplicando o lema da inversão [12], à equação 1.5-14, obteremos a equação de estimação 1.5-12. Portanto, foi possível saber deste ítem que a aproximação Bayesana origina as equações do filtro de Kalman para um processo discreto utilizando estatísticas Gaussianas.

1.6. FILTRO DE KALMAN DISCRETO VIA CALCULO VARIACIONAL E DUALIDADE AO CONTROLE ÓTIMO

Neste ítem, mostraremos que o desenvolvimento do filtro de Kalman pode ser feito utilizando as mesmas técnicas de controle ótimo, ou seja, encontrar um valor de $x(k+1)$ em cada iteração, que minimize a função custo,

$$J = \frac{1}{2} \left\| x(k_0) - \bar{x}(k_0) \right\|_{P^{-1}(k_0)}^2 + \frac{1}{2} \sum_{k=k_0}^{K-1} \left\| y(k+1) - H(k+1) x(k+1) - \omega(k) \right\|_{Q^{-1}(k)}^2 \quad (1.6-1)$$

sujeito às restrições de igualdade,

$$x(k+1) = A(k+1, k) x(k) + \Gamma(k) \omega(k)$$

se resume num problema de estimação, e resulta exatamente na mesma solução do problema de maximizar $p[x(k+1)/Y(k+1)]$ como mostrado no item anterior.

Escrevendo o Hamiltoniano correspondente a 1.6-1 como

$$H = \frac{1}{2} \left\| y(k+1) - H(k+1) A(k+1, k) x(k) - H(k+1) \Gamma(k) \omega(k) \right\|_{R^{-1}(k+1)}^2 + \frac{1}{2} \left\| \omega(k) \right\|_{Q^{-1}(k)}^2 + \lambda^T(k+1) [A(k+1, k) x(k) + \Gamma(k) \omega(k)]$$

Aplicando o princípio do máximo [9], resulta na determinação de

$$\frac{\partial H}{\partial \omega(k)} = 0, \quad \frac{\partial H}{\partial x(k)} = \lambda(k)$$

como condições de otimalidade.

Com as condições de transversalidade,

$$\lambda(k_0) = -P^{-1}(k_0) \left[x(k_0) - \bar{x}(k_0) \right], \quad \lambda(k_f) = 0$$

A solução deste problema resulta num problema de dupla condição de contorno representado pelas equações discretas lineares:

$$\hat{x}(k+1) = A(k+1, k) \hat{x}(k) + \Gamma(k) Q(k) \Gamma^T(k) A^{-T}(k+1, k) \lambda(k) \quad (1.6-2)$$

$$\lambda(k+1) = \left[I - H^T(k+1) R^{-1}(k+1) H(k+1) \Gamma(k) Q(k) \Gamma^T(k) \right] \cdot$$

$$A^{-T}(k+1, k) \lambda(k) + H^T(k+1) \cdot R^{-1}(k+1) \cdot$$

$$\left[y(k+1) - H(k+1) A(k+1, k) x(k) \right]$$

$$\lambda(k_0) = P^{-1}(k_0) \hat{x}(k_0), \quad \lambda(k_f) = 0$$

A solução desse problema de dupla condição de contorno utilizando programação dinâmica, nos fornece o melhor valor de $\hat{x}(k)$ para $k_0 \leq k \leq k_f$.

No entanto, para a maioria das aplicações, a solução de 1.6-2 fica limitada a um procedimento "off-line", já que há a necessidade do conhecimento à priori das observações.

A técnica do embutimento invariante [9] é aplicada à equação 1.6-2, gerando um algoritmo sequencial de estimação "on-line", idêntico aos algoritmos obtidos nos itens 1.4 e 1.5.

Kalman demonstrou [2] a dualidade entre controle ótimo e o problema de estimação de estado pelo fato de a solução para o problema de filtragem ótima ser em forma, a mesma solução do problema de regulação linear quadrática considerando as devidas correspondências entre as matrizes do sistema para os dois problemas, ou seja, ambas soluções se reduzem a um problema de dupla condição de contorno linear. Este fato é de grande importância no teorema da separação, quando controle e estimação são combinados.

CAPÍTULO 2

ESTIMAÇÃO MULTINÍVEL DE ESTADO

2.1. INTRODUÇÃO

Todos os sistemas reais operam em um ambiente estocástico onde estão sujeitos a perturbações e em adição a isso, o controlador, na prática, tem de atuar baseado em medidas imperfeitas. No caso de sistemas dinâmicos lineares, estas dificuldades são superadas, pelo menos em princípio, com a utilização do filtro de Kalman [2], visto que este provê uma estimação de estado ótima. Sabemos do teorema da separação que o controle ótimo estocástico pode ser projetado em combinação com o filtro ótimo como mostrado na Figura 2.1.

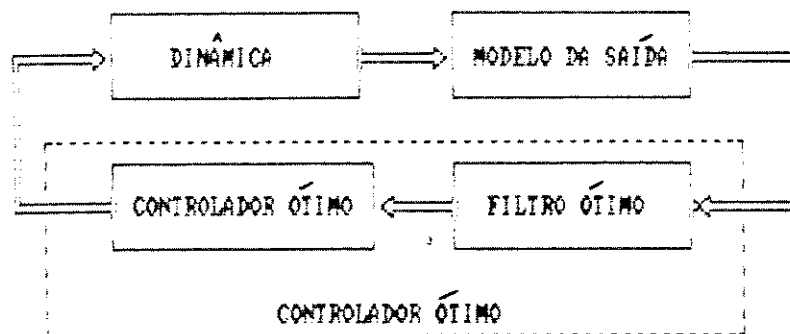


FIGURA 2.1 - Sistema de Controle Ótimo Estocástico.

A desvantagem da aplicação do filtro de Kalman global é que o esforço computacional associado com a sua implementação é proporcional ao cubo da ordem do sistema [41].

Consequentemente, se tal filtro fosse utilizado com um controlador hierárquico [16], então a vantagem computacional adquirida na implementação do controlador seria sub-utilizada. Com o objetivo de contornar esse problema, houve a necessidade de se estudar procedimentos para a decomposição do filtro de Kalman global.

As principais contribuições no estudo dos métodos para superar as dificuldades encontradas na implementação da teoria do filtro de Kalman, foram:

1. Métodos de Redução do Modelo do Sistema:

Visto que a dimensão do filtro de Kalman-Bucy é da mesma ordem do sistema, métodos propondo uma redução do modelo foram desenvolvidos.

Porém, para sistemas de ordem > 100 , a aplicação dos métodos de redução do modelo originava modelos de ordem reduzida (> 20), que ainda, exigiam um grande esforço computacional para sua implementação, aplicando a teoria do filtro de Kalman. Portanto, em geral, os métodos de redução do modelo são limitados.

2. Construção e Uso de Observadores de Ordem Reduzida:

Dado um vetor de estado de dimensão n e um vetor de medidas de dimensão m , ambos sem perturbação, o observador de ordem $(n-m)$ é construído para estimar $(n-m)$ estados e os outros m estados são estimados diretamente das m medidas. Este método foi estendido para o caso de sistemas lineares estocásticos.

Em geral, em situações práticas, m é consideravelmente muito menor do que n e, portanto, $(n-m)$ ainda é muito elevado, de forma que o problema da dimensão ainda permanece. Além disso, para sistemas variantes no tempo, este método é inconveniente, já que as matrizes de transformação devem ser calculadas em cada iteração.

3. Método de Perturbações Singulares:

A teoria de perturbação singular foi aplicada para o problema de filtragem ótima. O sistema é particionado em dois subsistemas, tal que os autovalores do primeiro subsistema são muito maiores em magnitude do que os autovalores do segundo subsistema.

Utilizando a teoria de perturbação singular, um filtro de Kalman-Bucy desacoplado é construído para o segundo subsistema.

Os valores estimados para o segundo subsistema são tratados como entradas na filtragem de Kalman para o primeiro subsistema. Dessa forma surge, naturalmente, uma estrutura descentralizada na formulação do problema. Este método é limitado a uma classe de subsistemas fracamente acoplados, que podem ser particionados em dois conjuntos que apresentem uma grande diferença em magnitude dos autovalores correspondentes.

4. Métodos de Partição:

Diante das dificuldades computacionais (relacionadas com a ordem do modelo) para a solução das equações recursivas do filtro de Kalman, um procedimento interessante para superar tais dificuldades foi o de particionar o sistema em um conjunto de subsistemas de menor ordem. Tais métodos para a solução das equações do filtro ignoraram as interações entre os sistemas particionados, o que, em geral, conduz a um vetor estimado de estado distante do vetor real de estado do sistema.

SHAH [17], desenvolveu um filtro sub-ótimo (filtro S.P.A. de SHAH) que apresenta um menor esforço computacional em relação ao filtro de PEARSON. O filtro de SHAH é baseado no método de filtragem S.P.A. (Supplemented Partitioning Approach), onde o sistema é particionado em um conjunto de subsistemas, e estimadores de variância mínima lineares são construídos para cada um dos subsistemas.

A principal característica do método S.P.A. é que todos os cálculos dos filtros dos subsistemas podem ser executados em paralelo e as interações entre os vários subsistemas são considerados através da adição de termos derivados das interações entre os vários subsistemas às matrizes de covariância dos ruídos do sistema e da observação.

PEARSON [20, 21], em particular, sugeriu um filtro multinível ótimo, mas no entanto, tal filtro requer um maior esforço computacional do que o filtro de Kalman Global (centralizado).

Apesar disso, a formulação de PEARSON tem sua importância pelo fato de delinear as dificuldades de se implementar um filtro multinível ótimo.

O problema de decomposição e coordenação do filtro ótimo de Kalman recebeu atenção da parte de muitos estudiosos [17-20, 38-40]. Porém, dentre as aproximações existentes, a maioria não apresenta uma solução ótima.

No item a seguir, falaremos de algumas dessas aproximações.

2.2. O FILTRO GLOBAL DE KALMAN E O FILTRO MULTINÍVEL DE PEARSON

2.2.1. O FILTRO DE KALMAN [2]

Devido a sua importância, iniciaremos com uma breve descrição do filtro ótimo de KALMAN, visto no Capítulo 1. Considere o sistema:

$$\underline{X}_{k+1} = A \underline{X}_k + B \underline{U}_k + \Gamma \underline{W}_k$$

$$\underline{Y}_{k+1} = H \underline{X}_{k+1} + \underline{V}_k$$

onde $\underline{X}$ e $\underline{Y}$ são, respectivamente, os vetores de estado e saída. $\underline{W}_k$ e $\underline{V}_k$ são supostos ruídos brancos Gaussianos com média nula e covariâncias.

$$E \left\{ \begin{bmatrix} \underline{W}_k & \underline{W}_j^T \end{bmatrix} \right\} = Q \delta_{kj}$$

$$E \left\{ \underline{v}_k \underline{v}_j^T \right\} = R \delta_{kj}$$

onde δ é o delta de Kronecker. $\underline{v}_k$, $\underline{u}_k$ e $\underline{x}_0$ são supostos não correlatos entre si para todo k . Para tal sistema, dadas as estimacões iniciais $E \left\{ \underline{x}_0 \right\} = \hat{\underline{x}}_0$ e $P(0) = P_0$, a estimacão ótima $\hat{\underline{x}}$ do estado $\underline{x}$ é expressa por:

$$\hat{\underline{x}}_{k+1} = A \underline{x}_k + B \underline{u}_k + G_k \left(\underline{y}_k - H \hat{\underline{x}}_k \right)$$

onde G é uma matriz de ganho definida por:

$$G_k = P_k H_k^T R_k^{-1}$$

e onde P_k é a matriz de covariância do erro de estimacão e é uma solução da equação de Riccati matricial:

$$P_{k+1} = A \left[P_k - P_k \cdot H^T \left(H P_k H^T + R \right)^{-1} \cdot H P_k \right]$$

$$A^T + \Gamma Q \Gamma ; P(0) = P_0$$

Portanto, a principal parte do esforço computacional está na computação da equação de Riccati matricial, de ordem determinada pelo sistema. Claramente, se $\underline{x}$ puder ser particionado em vetores de menor ordem, o esforço computacional para a solução da matriz de covariância será reduzido, e este é em essência o princípio básico no qual se baseia o filtro de PEARSON, comentado a seguir.

2.2.2. O FILTRO MULTINÍVEL DE PEARSON [20-22]:

O filtro de PEARSON foi talvez a primeira formulação do problema de estimação de estado multinível para sistemas dinâmicos lineares conectados. Mas, veremos a seguir que tal filtro apresenta dificuldades para sua implementação.

A idéia básica é uma extensão direta de algumas idéias dos algoritmos de controle hierárquico [16].

A versão discretizada do modelo (conforme Figura 2.2) no qual PEARSON se baseou, tem uma estrutura particular. Esse modelo consiste de subsistemas que são conectados através de uma matriz de conexão L , pelas saídas $\underline{Y}$ de outros subsistemas, ou seja:

$$\underline{Z}_k = L \underline{Y}_k \quad (2.2.2-1)$$

onde $\underline{Z}_k$ é uma entrada de conexão e L é uma matriz formada de elementos um e zero.

As entradas de conexão são supostas completamente mensuráveis com a presença de perturbações, isto é,

$$\underline{W}_k = \underline{Z}_k + \underline{\xi}_k$$

onde $\underline{\xi}$ é um vetor ruído Gaussiano com média nula e covariância conhecida R .

$$E \left[\underline{\xi}(j) \underline{\xi}^T(k) \right] = R \delta_{jk}$$

onde δ é a função delta de Kronecker.

Além disso, a equação de estado é expressa por:

$$\underline{X}_{k+1} = A \underline{X}_k + B \underline{M}_k + C \underline{Z}_k$$

onde $\underline{Z}$ provê a conexão dos outros subsistemas e $\underline{M}$ é uma perturbação suposta um processo Gaussiano com média zero e covariância conhecida S , isto é,

$$E \left\{ \underline{M}_j \cdot \underline{M}_k^T \right\} = S \delta_{jk}$$

O sistema tem também um número de saídas $\underline{Y}$ que são utilizadas em (2.2.2-1) para definir as entradas $\underline{Z}$, e estas estão relacionadas com os estados $\underline{X}$ pela equação:

$$\underline{Y}_k = m \underline{X}_k + N \underline{M}_k$$

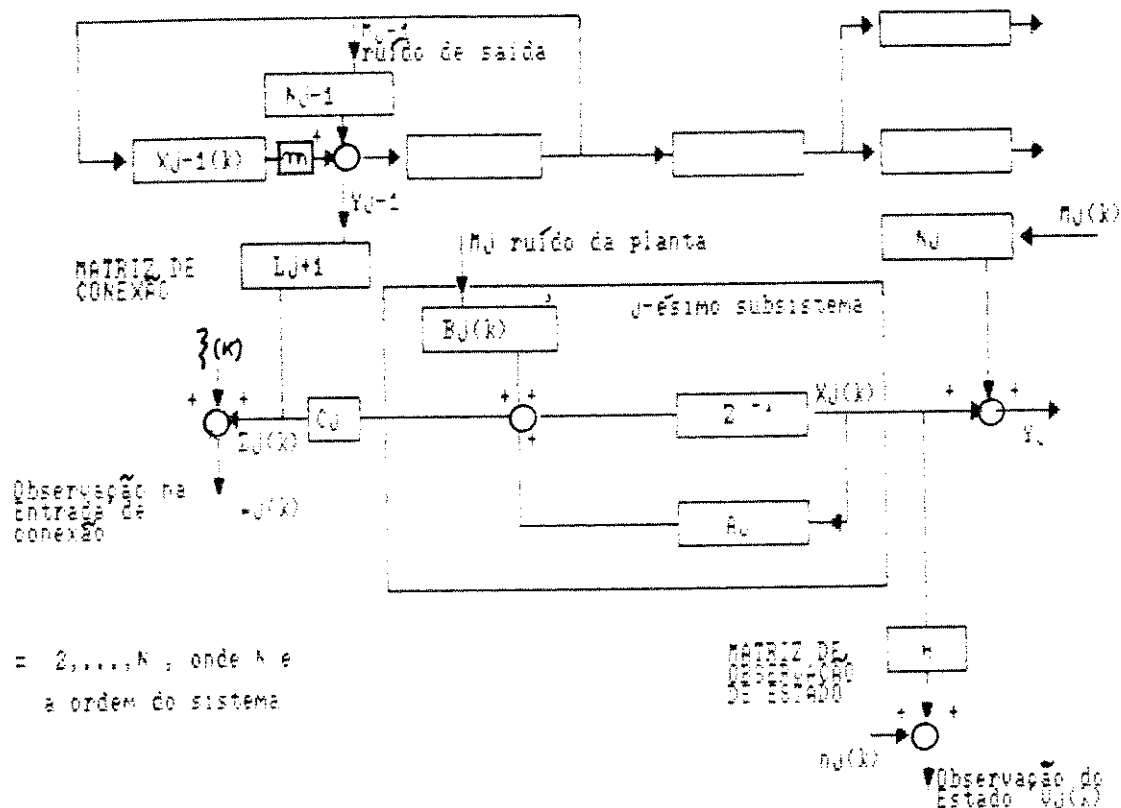


FIGURA 2.2 - Modelo Discretizado do Sistema Dinâmico.

Para a observação de estado, foi postulado um processo $\underline{V}$, onde:

$$\underline{V}_k = H \underline{X}_k + \underline{\eta}_k$$

e $\underline{\eta}_k$ é também um processo Gaussiano com média nula e covariância conhecida Q , isto é,

$$E \{ \underline{\eta}_j \underline{\eta}_k^T \} = Q \delta_{jk}$$

Pela exposição do modelo da Figura 2.2, é possível notarmos que as interações entre os sub-sistemas existem somente devido a matriz de conexão L e que as matrizes A , B , C , m , N , Q , R , S são todas bloco diagonais.

A base do método se resume em: através de um problema de minimização funcional, encontrar a melhor estimação de $\underline{X}$, utilizando métodos de otimização hierárquica.

Dessa forma, as equações que representam o modelo integrado são:

$$\underline{X}_{k+1} = (A + CLm) \underline{X}_k + (B + CLN) \underline{M}_k \quad (2.2.2-2a)$$

$$\begin{bmatrix} \underline{V}_k \\ \underline{W}_k \end{bmatrix} = \begin{bmatrix} H \\ Lm \end{bmatrix} \underline{X}_k + \begin{bmatrix} \underline{\eta}_k \\ LNM_k + \underline{\zeta}_k \end{bmatrix} \quad (2.2.2-2b)$$

isto é,

$$\underline{X}_{k+1} = A^* \underline{X}_k + B^* \underline{M}_k$$

$$\underline{V}_k^* = H^* \underline{X}_k + \underline{\eta}_k^*$$

onde $\underline{M}$ e $\underline{\eta}^*$ estão correlacionados da forma:

$$E \left[\begin{bmatrix} \underline{M}_j \\ \underline{\eta}_j^* \end{bmatrix} \begin{bmatrix} \underline{M}_k \\ \underline{\eta}_k^* \end{bmatrix}^T \right] = \begin{bmatrix} S & O & SN^T L^T \\ O & Q & O \\ LNS & O & R + LNSN^T L^T \end{bmatrix} \delta_{jk}$$

$$= \begin{bmatrix} S & (S^*)^T \\ S^* & Q^* \end{bmatrix} \delta_{jk}$$

onde:

$$A^* = A + CLm ; \quad H^* = \begin{bmatrix} H \\ Lm \end{bmatrix} ; \quad \eta^* = \begin{bmatrix} \eta \\ LNm + \xi \end{bmatrix} ; \quad B^* = B + CLN ;$$

$$V^* = \begin{bmatrix} V \\ W \end{bmatrix} ; \quad S^* = \begin{bmatrix} LNS & O \end{bmatrix} ; \quad Q^* = \begin{bmatrix} Q & O \\ O & R + LNSN^T L^T \end{bmatrix}$$

A função densidade de probabilidade Gaussiana a posteriori, tem um argumento exponencial negativo descrito por:

$$\frac{1}{2} \left\| \underline{X}_o - E \left(\underline{X}_o \right) \right\|_{P_o^{-1}}^2 +$$

(2.2.2-3)

$$+ \frac{1}{2} \sum_{k=0}^T \left\{ \left\| V^* - H^* X \right\|^2 \begin{bmatrix} S & (S^*)^T \\ S^* & Q^* \end{bmatrix}^{-1} \right\}$$

onde X e M satisfazem a equação 2.2.2-2.

Portanto, a estimação por máxima verossimilhança de X , pode ser obtida pela minimização da expressão na equação 2.2.2-3, com relação a $\underline{X}_k$ e $\underline{M}_k$; $k=0, \dots, T$.

Com o intuito de mostrar explicitamente as interações entre os subsistemas na função de custo (2.2.2-3), objetivando a decomposição de 2.2.2-3, relativa aos subsistemas (tendo em vista que a matriz de conexão L possui uma estrutura arbitrária), manipulamos a forma quadrática na Equação 2.2.2-3, relacionada com a equação 2.2.2-2, como mostrado a seguir.

Pelo método de eliminação de GAUSS, podemos escrever:

$$\begin{bmatrix} S & O & SN^T L^T \\ O & Q & O \\ LNS & O & R + LNSN^T L^T \end{bmatrix}^{-1} = \begin{bmatrix} S^{-1} + N^T L^T R^{-1} LN & O & N^T L^T R^{-1} \\ O & Q^{-1} & O \\ -R^{-1} LN & O & R^{-1} \end{bmatrix}$$

de forma que o somatório na equação 2.2.2-3 pode ser escrito como:

$$\begin{bmatrix} M & V - HX & W - LmX \end{bmatrix} \begin{bmatrix} S^{-1} + N^T L^T R^{-1} LN & O & N^T L^T R^{-1} \\ O & Q^{-1} & O \\ -R^{-1} LN & O & R^{-1} \end{bmatrix}$$

$$\begin{bmatrix} M \\ V - HX \\ W - LmX \end{bmatrix} = \left\{ \begin{bmatrix} M (S^{-1} + N^T L^T R^{-1} LN) - (W - LmX) R^{-1} LN \end{bmatrix} \right.$$

$$\left. \begin{bmatrix} (V - HX) Q^{-1} \end{bmatrix} \begin{bmatrix} (W - LmX) R^{-1} - MN^T L^T R^{-1} \end{bmatrix} \begin{bmatrix} M \\ V - HX \\ W - LmX \end{bmatrix} \right\}$$

$$= \left\| M \right\|_{S^{-1}}^2 + \left\| V - HX \right\|_{Q^{-1}}^2 + \left[MN^T L^T R^{-1} LNM - (W - LmX) \right. \\ \left. \cdot R^{-1} LNM + (W - LmX) R^{-1} (W - LmX) \cdot M (N^T L^T R^{-1}) (W - LmX) \right]$$

Sabemos que $Z = LNM + LmX$ e, portanto, $LmX = Z - LNM$. Desse modo, o somatório do funcional torna-se:

$$\left\| M \right\|_{S^{-1}}^2 + \left\| V - HX \right\|_{Q^{-1}}^2 + \left[MN^T L^T - W + Z - LNM \right] \\ \cdot R^{-1} LNM + \left[W - Z + LNM - LNM \right] R^{-1} (W - LmX) \\ = \left\| M \right\|_{S^{-1}}^2 + \left\| V - HX \right\|_{Q^{-1}}^2 + \left\{ \left[Z - W \right] \cdot R^{-1} LNM - \left[Z - W \right] R^{-1} \right. \\ \left. \cdot \left[W - Z + LNM \right] \right\} = \left\| M \right\|_{S^{-1}}^2 + \left\| V - HX \right\|_{Q^{-1}}^2 + \left\| Z - W \right\|_{R^{-1}}^2$$

tal que encontrar um valor estimado $\hat{X}$ corresponde a minimizar a função de custo:

$$J = \frac{1}{2} \left\| \underline{X}_0 - E(\underline{X}_0) \right\|_{P_0^{-1}}^2 + \frac{1}{2} \sum_{k=0}^T \\ \cdot \left\{ \left\| \underline{M}_k \right\|_{Q^{-1}}^2 + \left\| \underline{V}_k - H_k \underline{X}_k \right\|_{R^{-1}}^2 + \left\| \underline{Z}_k - \underline{W}_k \right\|_{S^{-1}}^2 \right\}$$

sujeito às restrições dinâmicas:

$$\underline{X}_{k+1} = A \underline{X}_k + B \underline{M}_k + C \underline{Z}_k$$

e

$$\underline{Z}_k = Lm\underline{X}_k + LNM_k$$

Nesta formulação, as conexões entre os subsistemas são atribuídas somente à matriz L de conexão, já que as matrizes A , B , C , Q , R , S e H são bloco diagonais.

Portanto, é possível decompor este problema pelo método de coordenação por objetivo que se resume em maximizar a função dual $\theta(\lambda)$ com relação a λ , onde:

$$\theta(\lambda) = \min_{\underline{X}, \underline{M}} \left[\begin{array}{ll} L(\underline{X}, \underline{M}, \underline{\lambda}) & \text{s.a} \\ \underline{X}_{k+1} = A \underline{X}_k + B \underline{M}_k + C \underline{Z}_k \end{array} \right]$$

onde o Lagrangeano L é dado por:

$$L(\underline{X}, \underline{M}, \underline{\lambda}) = \frac{1}{2} \left\| \underline{X}_0 - E(\underline{X}_0) \right\|_{P_0^{-1}}^2 + \sum_{k=0}^T$$

$$\left[\frac{1}{2} \left\| \underline{M}_k \right\|_{Q^{-1}}^2 + \frac{1}{2} \left\| \underline{V}_k - H_k \underline{X}_k \right\|_{R^{-1}}^2 + \right.$$

(2.2.2-4)

$$\left. \frac{1}{2} \left\| \underline{Z}_k - \underline{W}_k \right\|_{S^{-1}}^2 + \underline{\lambda}_k^T \left(LHX_k + LNM_k - \underline{Z}_k \right) \right]$$

Para um dado valor de multiplicadores de Lagrange λ_k , podemos decompor o Lagrangeano (2.2.2-4) pelo fato de L ser aditivamente separável e, dessa forma, em cada instante K, minimizar os subLagrangeanos no nível dos subsistemas da estrutura hierárquica e, no nível de coordenação, calcular a trajetória λ_k utilizando uma técnica de gradiente.

Esta solução via decomposição por objetivo, tanto em dois quanto em três níveis, resulta numa estimativa ótima $\hat{x}$ de x.

Do ponto de vista prático, tal solução não é interessante para uma estimativa em tempo real, visto que ela exige o conhecimento à priori, não disponível das perturbações v_k e w_k , $k=0, \dots, T$, havendo a necessidade de observá-las durante a evolução do processo, e o cálculo da trajetória λ_k envolvendo um procedimento de otimização (cálculos de comprimento do passo e da direção de busca que garantam convergência) que sobrecarrega computacionalmente a implementação dessa solução e torna a sequência de valores estimados não factível.

Para contornar essas dificuldades, em [20], é proposto um estimador sub-ótimo para uma classe de problemas mais complexos. No próximo item, descrevemos tal estimador.

2.3. FILTRO MULTINÍVEL DE ARAFEH E SAGE [20]

ARAFEH e SAGE consideraram um sistema discreto não linear de n-ésima ordem, dado por:

$$\underline{x}_{k+1} = A \left[\underline{x}_k, k \right] + \underline{w}_k \quad (2.3-1a)$$

$$\underline{z}_k = H \left[\underline{x}_k, k \right] + \underline{v}_k \quad (2.3-1b)$$

onde $\underline{X}_k$ é um vetor de dimensão n de estados e parâmetros a serem estimados, $A \left[\underline{X}_k, k \right]$ é um vetor de dimensão n de funções não lineares que descrevem a estrutura do sistema incluindo as entradas determinísticas, $\underline{W}$ é um vetor de dimensão n representando ruído de planta que é suposto um ruído branco Gaussiano com:

$$E \left(\underline{W}_k \right) = \underline{q}_k \quad ; \quad E \left(\underline{W}_k \underline{W}_j^T \right) = Q_k \delta_{kj}$$

onde δ é o delta de Kronecker, Q é uma matriz bloco diagonal definida positiva, $\underline{Z}_k$ é um vetor de observação de dimensão m , $H \left[\underline{X}_k, k \right]$ é um vetor de dimensão m de funções não lineares da equação de observação.

O ruído $\underline{V}_k$ na observação é também suposto um ruído branco Gaussiano com:

$$E \left(\underline{V}_k \right) = \underline{r}_k \quad ; \quad E \left(\underline{V}_k \underline{V}_j^T \right) = R_k \delta_{kj}$$

onde R_k é uma matriz bloco diagonal definida positiva.

O vetor de estado inicial é também suposto Gaussiano com:

$$E \left(\underline{X}_0 \right) = \underline{X}_0 \quad e \quad E \left(\underline{X}_0 \underline{X}_0^T \right) = P_0$$

onde P_0 é bloco diagonal.

Para esta modelagem, a estimação por máximo a posteriori pode ser obtida com a minimização de:

$$J_{\underline{W}_k} = \frac{1}{2} \left\| \underline{X}_0 - \underline{X}_0 \right\|_{P_0^{-1}}^2 + \frac{1}{2} \sum_{k=k_0}^{kf-1}$$

$$\left\{ \left\| \underline{z}_{k+1} - \underline{r}_{k+1} - H \left[\underline{x}_{k+1}, k+1 \right] \right\|_{R_{k+1}^{-1}}^2 + \left\| \underline{w}_k - \underline{q}_k \right\|_{Q_k^{-1}}^2 \right\} \quad (2.3-2)$$

sujeito às restrições dinâmicas dadas pela equação (2.3-1) com k_0 e k_f fixados.

Podemos escrever o Hamiltoniano deste problema de minimização como:

$$H = \frac{1}{2} \left\{ \left\| \underline{z}_{k+1} - \underline{r}_{k+1} - H \left[\underline{x}_{k+1}, k+1 \right] \right\|_{R_{k+1}^{-1}}^2 + \frac{1}{2} \left\| \underline{w}_k - \underline{q}_k \right\|_{Q_k^{-1}}^2 + \underline{\lambda}_{k+1}^T \left\{ A \left[\underline{x}_k, k \right] + \underline{w}_k \right\} \right\}$$

onde $\underline{\lambda}$ é o vetor de coestado.

Utilizando o princípio do máximo [9], as condições necessárias para otimalidade produzem um problema de dupla condição de contorno não linear, descrito por:

$$\underline{x}_{k+1} = A \left[\underline{x}_k, k \right] + \underline{q}_k - Q_k \left[\underline{\lambda}_{k+1} - \underline{A}_k^T h_{k+1}^T R_{k+1}^{-1} \underline{y}_{k+1} \right] \quad (2.3-3a)$$

$$\underline{\lambda}_k = - \underline{A}_k^* \underline{H}_{k+1}^T \underline{R}_{k+1}^{-1} \underline{Y}_{k+1} + \underline{A}_k^T \underline{\lambda}_{k+1} \quad (2.3-3b)$$

onde:

$$\underline{Y}_{k+1} = \underline{Z}_{k+1} - \underline{r}_{k+1} - \underline{H} \left[\underline{X}_{k+1}, k+1 \right]$$

$$\underline{A}_k^* = \frac{\partial \underline{A}^T \left[\underline{X}_k, k \right]}{\partial \underline{X}_k}$$

$$\underline{h}_{k+1} = \frac{\partial \underline{H}^T \left[\underline{X}_{k+1}, k+1 \right]}{\partial \underline{X}_{k+1}}$$

e das condições de transversalidade:

$$\underline{\lambda}_{k_0} = \underline{P}_{k_0}^{-1} \left[\underline{X}_{k_0} - \underline{\bar{X}}_{k_0} \right]$$

$$\underline{\lambda}_{k_f} = \underline{0}$$

A solução da equação 2.3-3 fornece a estimação de x baseada na sequência de observações $(Z_0, Z_1, \dots, Z_{k_f})$. Entretanto, a solução deste problema de dupla condição de contorno só pode ser encontrada "off-line", visto que a sequência de observações $(Y_k; k_0 \leq k \leq k_f)$ precisa ser conhecida à priori.

Buscando a decomposição do sistema, introduzimos as variáveis de conexão da forma:

$$\underline{\Pi}_i(k) = \underline{g}_i \left[\underline{X}_j, k \right] = \sum_{j \neq i}^N g_{ij} \left[\underline{X}_j \right] \quad (2.3-4)$$

e, portanto, decompos a função de custo de forma que:

$$J = \sum_{i=1}^N J_i$$

onde:

$$J_i = \frac{1}{2} \left\| \underline{X}_i(0) - \underline{Y}_i(0) \right\|_{P_i^{-1}(0)}^2 + \sum_{k=k_0}^{k_f-1} \left\{ \frac{1}{2} \left\| \underline{Z}_i(k+1) - \underline{r}_i(k+1) - H_i \left[\underline{X}_i, \underline{\Pi}_i, k+1 \right] \right\|_{R_{k+1}^{-1}}^2 + \frac{1}{2} \left\| \underline{W}_i(k) - q_i(k) \right\|_{Q_i^{-1}(k)}^2 \right\}$$

e o modelo do sistema decomposto é descrito por:

$$\underline{X}_i(k+1) = A_i \left[\underline{X}_i, \underline{\Pi}_i, k \right] + \underline{W}_i(k)$$

$$\underline{Z}_i(k) = H_i \left[\underline{X}_i, \underline{\Pi}_i, k \right] + \underline{V}_i(k) \quad \text{para } i = 1, 2, \dots, N$$

Introduzindo o vetor multiplicador de Lagrange β_i para acrescentar a restrição de igualdade 2.3-4, o Hamiltoniano do i -ésimo subsistema pode ser escrito como:

$$H_i = \frac{1}{2} \left\| \underline{Z}_i(k+1) - \underline{r}_i(k+1) - H_i \left[\underline{X}_i, \underline{\Pi}_i, k+1 \right] \right\|_{R_{k+1}^{-1}(k+1)}^2 + \frac{1}{2} \left\| \underline{W}_i(k) - q_i(k) \right\|_{Q_i^{-1}(k)}^2$$

$$q_i(k) = \left[\begin{array}{c} z \\ Q_i^{-1}(k) \end{array} + \beta_i^T(k) \underline{\Pi}_i(k) - \sum_{j \neq i}^N \beta_j^T(k) q_{ji}(X_i) \right. \\ \left. + \underline{\lambda}_i^T(k+1) \left[\underline{A}_i \left(\underline{X}_i, \underline{\Pi}_i, k \right) + \underline{W}_i(k) \right] \right]$$

As condições necessárias para otimalidade originam as seguintes equações:

$$\hat{\underline{X}}_i(k+1) = \underline{A}_i \left[\hat{\underline{X}}_i, \underline{\Pi}_i, k \right] + q_i(k) - Q_i(k) A_i^{*-T}(k)$$

$$\left[\underline{\lambda}_i(k) - \underline{B}_i(k) \right]$$

$$\underline{\lambda}_i(k+1) = A_i^{*-T} \left[\underline{\lambda}_i(k) - \underline{B}_i(k) \right] +$$

$$\underline{h}_i^T(k+1) R_i^{-1}(k+1) \underline{Y}_i(k+1)$$

onde:

$$A_i^{*}(k) = \frac{\partial \underline{A}_i \left[\hat{\underline{X}}_i, \underline{\Pi}_i, k \right]}{\partial \underline{X}_i(k)}$$

$$\underline{h}_i(k+1) = \frac{\partial \underline{H}_i \left[\hat{\underline{X}}_i, \underline{\Pi}_i, k \right]}{\partial \underline{X}_i(k+1)}$$

$$\underline{Y}_i(k+1) = \underline{Z}_i(k+1) - \underline{r}_i(k+1) - \underline{H}_i \left[\hat{\underline{X}}_i, \underline{\Pi}_i, k+1 \right]$$

$$\underline{B}_i(k) = \frac{\partial}{\partial \hat{\underline{X}}_i(k)} \left[\underline{\beta}_i^T(k) \underline{\Pi}_i(k) - \sum_{j=1}^N \underline{\beta}_j^T(k) \mathbf{a}_{ji}(\hat{\underline{X}}_i) \right]$$

com:

$$\underline{\lambda}_i(k_0) = \underline{P}_i^{-1}(k_0) \left[\hat{\underline{X}}_i(k_0) - \underline{X}_i(k_0) \right]$$

$$\underline{\lambda}_i(k_f) = 0$$

O método de predição de interação [16] pode ser utilizado para a solução do problema decomposto. A solução baseada nesse método está na predição de novos vetores $\underline{\Pi}_i(k)$ e $\underline{\beta}_i(k)$ à nível de coordenação. Esse método se mostra mais eficiente com relação ao de coordenação por objetivo, visto que [22]:

- O algoritmo de coordenação é mais simples, pois não há a necessidade de se fazer buscas para o cálculo de trajetória ótima, como no método de coordenação por objetivo;
- A formulação do problema não inclui o termo quadrático em $\underline{\Pi}$ (sem significado físico real, mas para evitar singularidades) como no método de coordenação por objetivo;
- Apresenta convergência extremamente rápida.

Baseados na técnica do embutimento invariante (*invariant imbedding*) [9], de programação dinâmica, ARAFEH e SAGE desenvolveram o algoritmo sequencial mostrado na Tabela 2.3-1.

Modelo do Sistema	$X_{i(k+1)} = A_i \left[X_i, \Pi_i, k \right] + w_i(k)$
Modelo da Observação	$Y_{i(k)} = H_i \left[X_i, \Pi_i, k \right] + v_i(k)$
Estatísticas à priori	$E \left\{ X_{i(k_0)} \right\} = \bar{X}_{i(k_0)}$ $E \left\{ w_i(k) \right\} = q_i(k) ; E \left\{ v_i(k) \right\} = r_i(k)$ $COV \left\{ w_i(k), w_i(l) \right\} = Q_i(k) S_k(k-l)$ $E \left\{ v_i(k), v_i(l) \right\} = R_i(k) S_k(k-l)$ $Var \left\{ X_{i(k_0)} \right\} = P_i(k_0)$
Inovações do Processo	$\hat{v}_i(k+1) = Y_i(k+1) - H_i \left[\hat{X}_i(k+1/k), \Pi_i(k+1), k+1 \right] - r_i(k+1)$
Definições	$\beta(k) \text{ e } \underline{\Pi}(k) \text{ são Variáveis de Coordenação e}$ $\Pi_i(k) = g_i \left[\hat{X}_j(k), k \right]$ $A_{i(k)}^* = \frac{\partial A_i \left[\hat{X}_i, \Pi_i, k \right]}{\partial \hat{X}_i(k)}$

continua...

Definições	$h_i(k+1/k) = \frac{\partial H_i \left[\hat{X}_i(k+1/k), \Pi_i(k+1), k+1 \right]}{\partial \hat{X}_i(k+1/k)}$
Algoritmo	$M_i(k+1) = h_i^T(k+1) \cdot R_i^{-1}(k+1) \cdot \hat{V}_i(k+1)$ $N_i(k+1) = - \left\{ P_i(k+1) \cdot \left[I - \frac{\partial M_i(k+1)}{\partial \hat{X}_i(k+1/k)} \cdot Q_i(k) \right] \right.$ $\left. - Q_i(k) \right\} A_i^{*-T}(k)$
Intermediário	$C_i(k+1) = \frac{\partial}{\partial \hat{X}_i(k)} \left\{ A_i^{*-T}(k) \cdot B_i(k) - \frac{\partial M_i(k+1)}{\partial \hat{X}_i(k+1/k)} \right.$ $\left. \cdot Q_i(k) \cdot A_i^{*-T}(k) \cdot B_i(k) \right\} \cdot P_i(k) \cdot A_i^{*T}(k)$ $D_i(k) = \frac{\partial}{\partial \hat{X}_i(k)} \left\{ Q_i(k) \cdot A_i^{*-T}(k) \cdot B_i(k) \right\} \cdot$ $P_i(k) \cdot A_i^{*T}(k)$

continua...

Continuação Tabela 2.3-1.

Preditor	$\hat{x}_i(k+1/k) = A_i \left[\hat{x}_i(k) \cdot \Pi_i(k), k \right] + q_i(k)$
Filtro	$\hat{x}_i(k+1) = \hat{x}_i(k+1/k) + G_i(k+1) \cdot \hat{v}_i(k+1) + N_i(k+1) \cdot B_i(k)$
Ganho	$G_i(k+1) = P_i(k+1) \cdot h_i^T(k+1) \cdot R_i^{-1}(k+1)$
Variância do erro à priori	$P_i(k+1/k) = A_i^*(k) \cdot P_i(k) \cdot A_i^{*T}(k) + Q_i(k)$
Variância do erro à posteriori	$P_i(k+1) = \left[P_i(k+1/k) + D_i(k) \right] \cdot \left[I - \frac{\partial M_i(k+1)}{\partial \hat{x}_i(k+1/k)} \cdot P_i(k+1/k) + C_i(k+1) \right]^{-1}$
Condições Iniciais	$\hat{x}_i(k_0) = E \left\{ X_i(k_0) \right\} ; P_i(k_0) = \text{Var} \left\{ X_i(k_0) \right\}$

TABELA 2.3-1 - Algoritmo para Estimação MAP Discreta para β e Π conhecidas e fornecidas pela Unidade de Coordenação.

Nesse método, uma técnica de coordenação hierárquica sequencial foi desenvolvida, com o objetivo de coordenar as unidades à nível de subsistemas, tal que o processo de estimação pudesse ser feito em tempo real, através da predição e correção das variáveis de coordenação $\Pi_{(k)}$ e $\beta_{(k)}$, como mostrado na Tabela 2.3-2.

Preditor	$\beta_i(k+1/k) = \beta_i + \beta_i^*(k) ; \beta_i(k_0) = 0$ $\beta_i^*(k) = \beta_i(k) - \beta_i(k-1)$ $\hat{x}_i(k+1/k) = A_i \left[\hat{x}_i(k), \Pi_i(k), k \right] + q_i(k) ;$ $\hat{x}_i(k_0) = E \{ X_i(k_0) \} ; p/ i = 1, 2, \dots, N$
Corretor	$\beta_i(k+1) = \beta_i(k+1/k) + K_{\beta_i} \frac{\partial H_i}{\partial \beta_i}$ $\Pi_i(k+1) = g_i \left[\hat{x}_j(k+1/k), k \right]$ onde $K_{\beta_i} = \alpha$ (comprimento fixo do passo) $\frac{\partial H_i}{\partial \beta_i} = \Pi_i(k) - g_i \left[\hat{x}_j(k), k \right]$ $p/ i = 1, 2, \dots, N$

TABELA 2.3-2 - Algoritmo Preditor-Corretor Sequencial Discreto.

A Figura 2.3, mostra o fluxograma do processo de coordenação hierárquica sequencial pelo método de predição de interação, proposto por SAGE e ARAFEH.

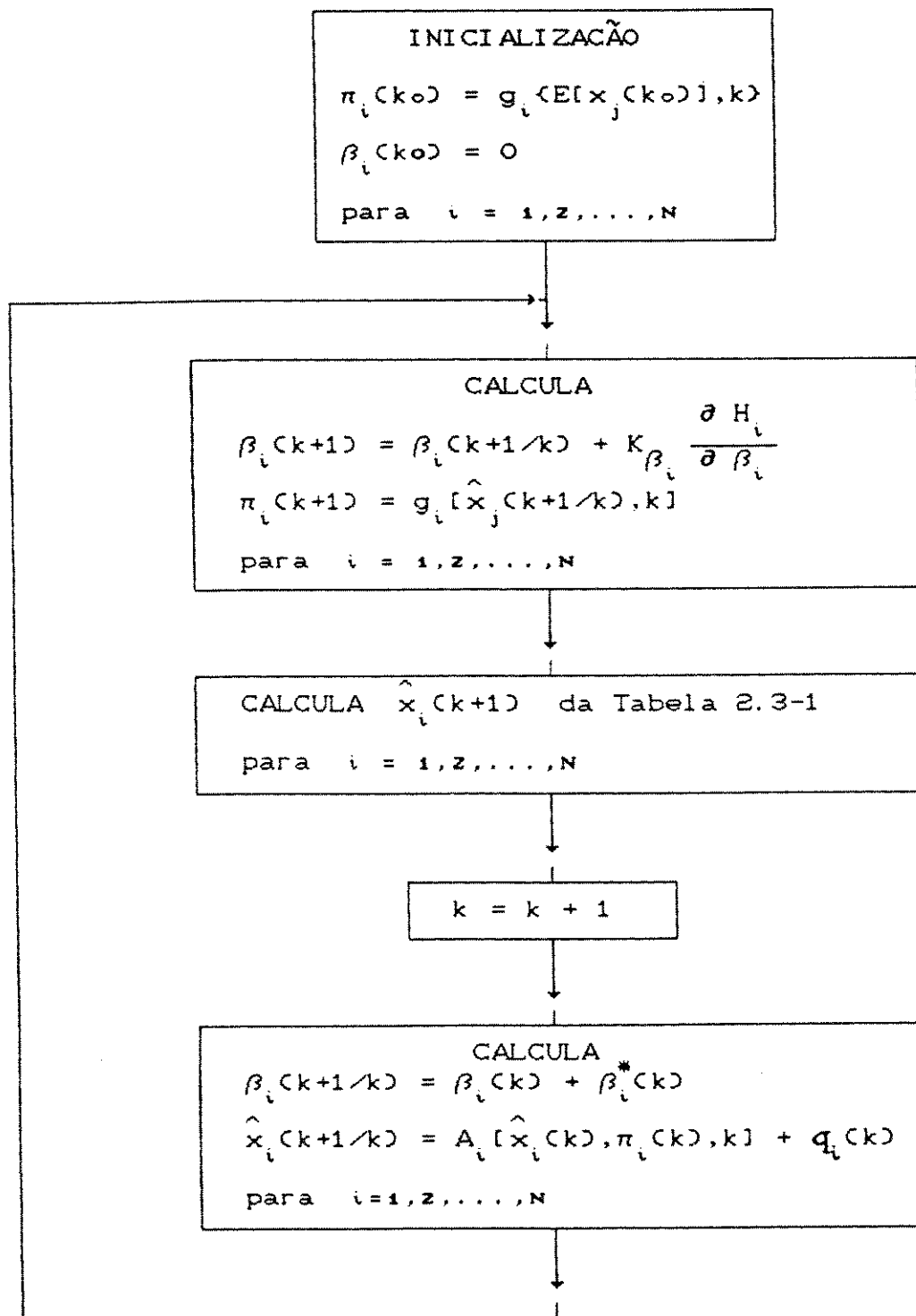


FIGURA 2.3 - Fluxograma do Processo de Coordenação Hierárquica Sequencial.

O método MAP juntamente com a técnica de coordenação sequencial desenvolvidos por ARAFEH e SAGE foram utilizados com sucesso em problemas de estimação de sistemas de potência conectados [24].

2.4. FILTRO PARTICIONADO SUB-ÓTIMO DE SHAH [17]

Considere o modelo descrito pelas equações:

$$\begin{matrix} \underline{X}_{k+1} \\ n \times 1 \end{matrix} = \begin{matrix} A_k \\ n \times n \end{matrix} \cdot \begin{matrix} \underline{X}_k \\ n \times 1 \end{matrix} + \begin{matrix} \underline{W}_k \\ n \times 1 \end{matrix} \quad (2.4-1a)$$

$$\begin{matrix} Y_k \\ m \times 1 \end{matrix} = \begin{matrix} H_k \\ m \times n \end{matrix} \cdot \begin{matrix} \underline{X}_k \\ n \times 1 \end{matrix} + \begin{matrix} \underline{V}_k \\ m \times 1 \end{matrix} \quad (2.4-1b)$$

onde,

$\underline{X}_k$ = Vetor de estado

A_k = Matriz de transição do estado

$\underline{W}_k$ = Vetor de ruído do sistema com média nula e covariância Q_k

Y_k = Vetor de observação

H_k = Matriz de observação

$\underline{V}_k$ = Vetor de ruído na observação com média nula e covariância R_k

e

$E \left\{ \underline{W}_k \cdot \underline{V}_k^T \right\} = 0$, isto é, $\underline{W}_k$ e $\underline{V}_k$ não correlatos,

P_0 e X_0 conhecidos.

Para melhor entendimento, particularizaremos o filtro no caso geral para o caso de dois subsistemas.

Particionando o sistema de equações (2.4-1) em dois subsistemas, teremos:

Subsistema 1:

$$\underline{X}_1(k+1) = A_{11}(k) \underline{X}_1(k) + A_{12}(k) \underline{X}_2(k) + \underline{W}_1(k)$$

$$\underline{Y}_1(k) = H_{11}(k) \underline{X}_1(k) + H_{12}(k) \underline{X}_2(k) + \underline{V}_1(k)$$

Subsistema 2:

$$\underline{X}_2(k+1) = A_{22}(k) \underline{X}_2(k) + A_{21}(k) \underline{X}_1(k) + \underline{W}_2(k)$$

$$\underline{Y}_2(k) = H_{22}(k) \underline{X}_2(k) + H_{21}(k) \underline{X}_1(k) + \underline{V}_2(k)$$

onde,

$\underline{X}_1(k)$ é um vetor de estado $n_1 \times 1$

$\underline{X}_2(k)$ é um vetor de estado $n_2 \times 1$

$\underline{Y}_1(k)$ é um vetor de observação $m_1 \times 1$

$\underline{Y}_2(k)$ é um vetor de observação $m_2 \times 1$

$$n_1 + n_2 = n$$

$$m_1 + m_2 = m$$

e as matrizes de transição e observação são de dimensões apropriadas.

O método de partição não pode ser diretamente aplicado às equações do sistema e de observação e, ao mesmo tempo, considerar as interações.

Portanto, o problema é reformulado de forma que aos ruídos do sistema e da observação, são adicionados termos que consideram as interações entre os subsistemas através das equações do sistema e de observação.

Ou seja,

$$\underline{X}_{1(k+1)} = A_{11}(k) \underline{X}_{1(k)} + A_{12}(k) \hat{\underline{X}}_{2(k/k)} + \underline{W}_{1(k)}^*$$

$$Z_{11}(k) = H_{11}(k) \underline{X}_{1(k)} + \underline{V}_{1(k)}^*$$

onde,

$$\underline{W}_{1(k)}^* = \underline{W}_{1(k)} + A_{12}(k) \tilde{\underline{X}}_{2(k/k)} \quad (2.4-2a)$$

$$\underline{V}_{1(k)}^* = \underline{V}_{1(k)} + H_{12}(k) \tilde{\underline{X}}_{2(k/k-1)} \quad (2.4-2b)$$

são por aproximação considerados ruídos brancos com média zero e covariância $Q_{11}^*(k)$ e $R_{11}^*(k)$.

Se $\hat{\underline{X}}_{2(k/k)}$ e $\hat{\underline{X}}_{2(k/k-1)}$ são conhecidos, então a teoria do filtro de KALMAN pode ser aplicada, resultando que:

$$\hat{\underline{X}}_{1(k+1/k)} = A_{11}(k) \hat{\underline{X}}_{1(k/k)} + A_{12}(k) \hat{\underline{X}}_{2(k/k)}$$

$$\hat{\underline{X}}_{1(k/k)} = \hat{\underline{X}}_{1(k/k-1)} + G_{11}(k)$$

$$\cdot \left[Z_{11}(k) - H_{11}(k) \hat{\underline{X}}_{1(k/k-1)} \right]$$

Processando o vetor de medidas $\underline{Y}_2(k)$ para o subsistema 1, teremos:

$$Z_{12}(k) = \underline{Y}_2(k) - H_{22}(k) \hat{\underline{X}}_{2(k/k)}$$

$$= H_{21}(k) \underline{X}_{1(k)} + \underline{V}_2(k) + H_{22}(k) \tilde{\underline{X}}_{2(k/k)}$$

resultando nas seguintes equações de filtragem para o subsistema 1:

$$\hat{\underline{X}}_{11}(k/k) = \hat{\underline{X}}_{11}(k/k-1) + G_{11}^*(k)$$

$$\cdot \left[Z_{12}(k) - H_{21}(k) \hat{\underline{X}}_{11}(k/k-1) \right]$$

onde,

$$G_{11}(k) = P_{11}(k/k-1) H_{11}^T(k) \left[H_{11}(k) P_{11}(k/k-1) \right.$$

$$\left. H_{11}^T(k) + R_{11}^*(k) \right]^{-1}$$

$$P_{11}(k+1/k+1) = A_{11}(k) P_{11}(k/k) A_{11}^T(k) + Q_{11}(k) +$$

$$A_{12}(k) \cdot P_{22}(k/k) + A_{12}^T(k)$$

$$P_{11}(k/k-1) = \left[I - G_{11}(k) H_{11}(k) \right] P_{11}(k/k-1)$$

$$\cdot \left[I - G_{11}(k) H_{11}(k) \right]^T + G_{11}(k) R_{11}^*(k) G_{11}^T(k)$$

$$G_{11}^*(k) = P_{11}(k/k-1) H_{21}^T(k) \left[H_{21}(k) P_{11}(k/k-1) H_{21}^T(k) \right.$$

$$\left. + R_{22}(k) + H_{22}(k) P_{22}(k/k-1) H_{22}^T(k) \right]^{-1}$$

$$P_{11}(k/k) = \left[I - G_{11}^*(k) H_{21}(k) \right] P_{11}(k/k-1)$$

$$\left[I - G_{11}^* (k) H_{21} (k) \right]^T + G_{11}^* (k) \left[R_{22} (k) + \right.$$

$$\left. H_{22} (k) \cdot P_{22} (k / k - 1) H_{22}^T (k) \right] G_{11}^{*T} (k)$$

$$Q_{11}^* (k) = Q_{11} (k) + A_{12} (k) P_{22} (k / k) A_{12}^T (k)$$

$$R_{11}^* (k) = R_{11} (k) + H_{12} (k) P_{22} (k / k - 1) H_{12}^T (k)$$

As equações de filtragem para o subsistema 2 são derivadas de forma análoga às do subsistema 1.

O modelo do sistema particionado para a estimação dos estados pelo método SPA e a estrutura do filtro SPA são mostrados na Figura 2.4 e Figura 2.5.

Os métodos de partição são dos mais variados [17], tais como: método de análise da correlação, método de análise da correlação parcial, método de transformação linear. O método S.P.A., é um método factível (tanto para sistemas fracos, como fortemente acoplados), para o projeto de filtros e controladores. Foi derivado de um procedimento heurístico (aproximação intuitiva), tornando as equações 2.4-2a e 2.4-2b ruídos brancos e posteriormente aplicando as equações do filtro de Kalman Global.

A sub-otimalidade desse filtro é devida ao fato dele não considerar os termos não diagonais (cruzados) da matriz de covariância durante o processo de estimação.

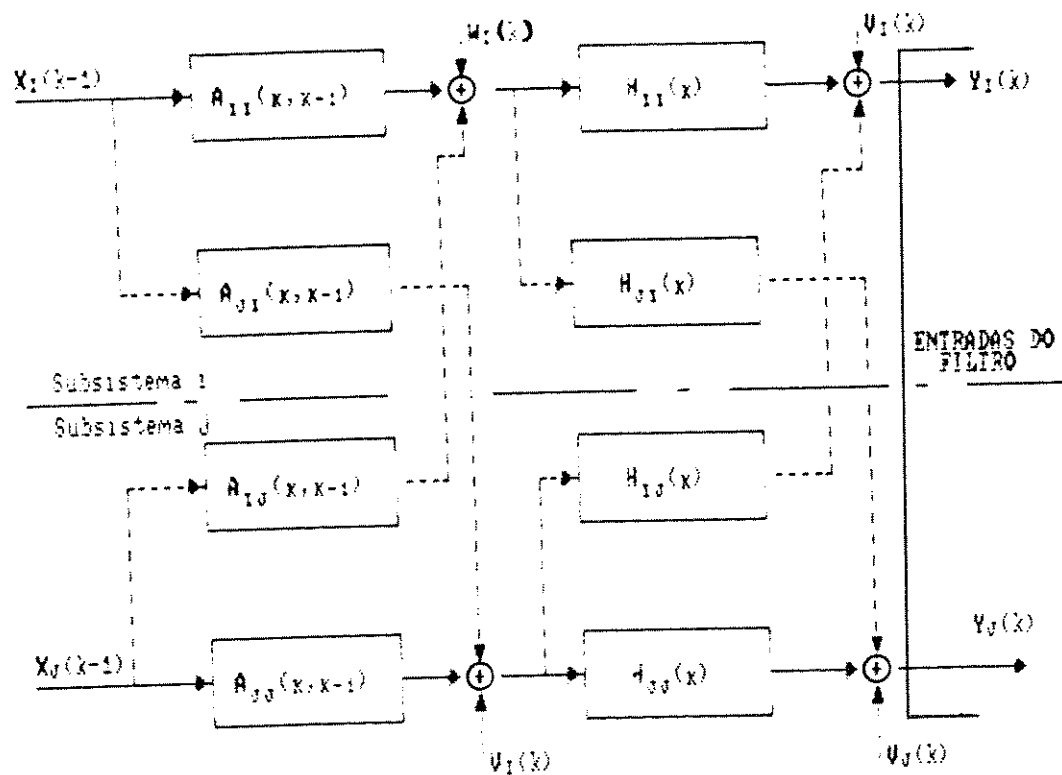


FIGURA 2.4 - Modelo do Sistema Particionado.

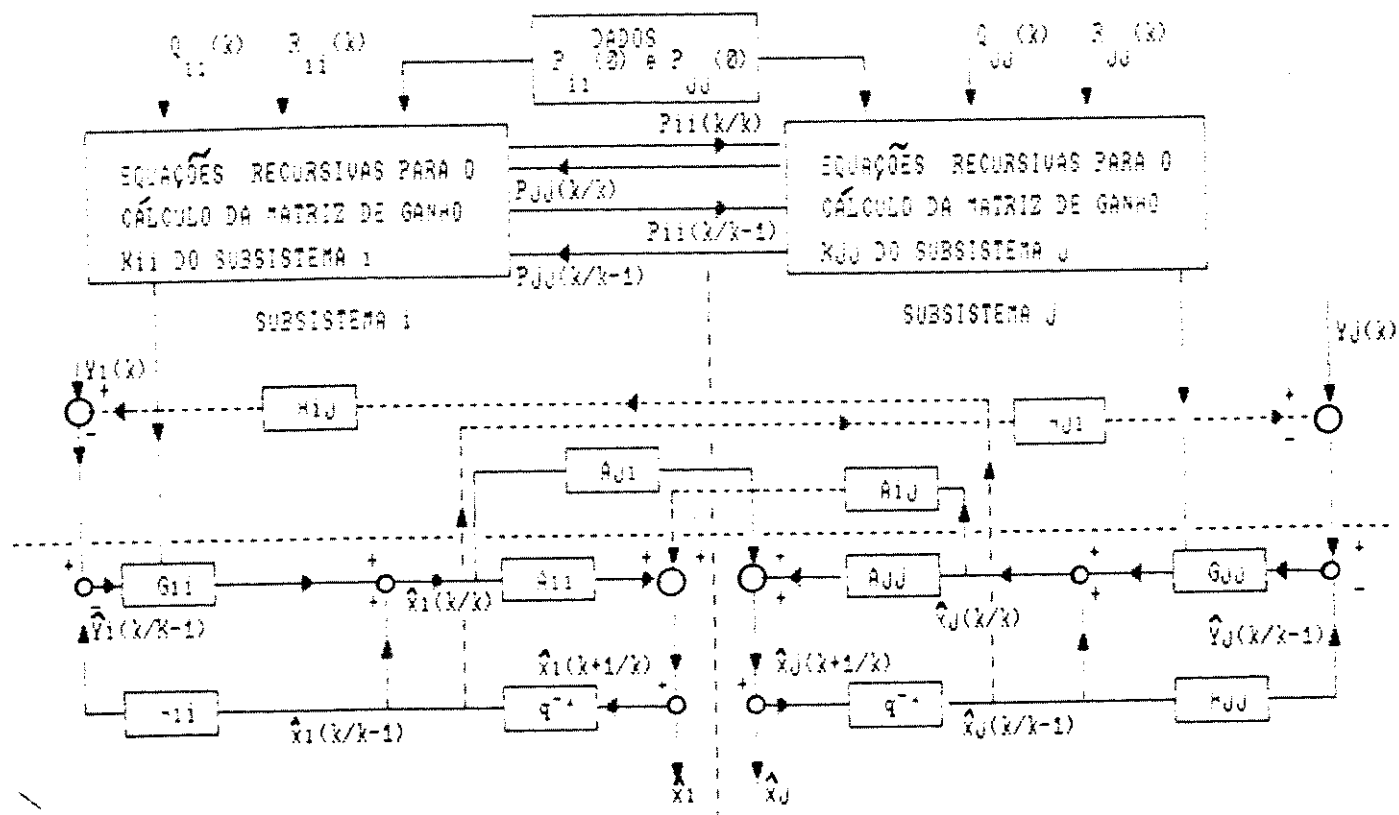


FIGURA 2.5 - Método SPA .

2.5. O FILTRO DE KALMAN MULTINÍVEL [26-29]: MÉTODO DAS PROJEÇÕES MÚLTIPLAS

2.5.1. INTRODUÇÃO

Em princípio, a teoria de filtragem linear desenvolvida por KALMAN [2], pode ser aplicada diretamente ao problema de estimação de estado de sistemas lineares de grande porte, porém, a inconveniência de sua utilização nesses sistemas, surge pelo fato de para um sistema de ordem n , a requisição de memória ser da ordem de n^2 e o tempo de processamento ser da ordem de n^3 [41].

Dentro dessa ótica, a decomposição do algoritmo de filtragem de KALMAN procura contornar esses dois problemas.

Uma gama de estudos foi realizada no que se refere a estes problemas [17-28, 38-40], mas os resultados obtidos são sub-ótimos, ou implicam em dificuldades de implementação em tempo real. Por exemplo, o filtro de NOTON [19], bem como o filtro de PEARSON, anteriormente mostrado, exigem mais cálculos que o filtro de Kalman global.

O filtro de SHAH [17], também mostrado anteriormente, apesar de sub-ótimo, é atraente pelo fato de exigir uma menor quantidade de cálculo em relação ao filtro de Kalman global e, finalmente, o filtro de ARAFEH e SAGE [20], que também é um filtro sub-ótimo.

Será mostrado a seguir, o desenvolvimento do algoritmo multinível, com base no princípio teórico do filtro de Kalman Global.

O princípio está em construir um estimador linear com variância mínima para cada subsistema, utilizando-se todas as medidas dos subsistemas. Será também mostrado que o estimador ótimo de Kalman é obtido, após um número finito de interações entre o primeiro nível e o coordenador. Uma qualidade deste método reside no fato de todos os acoplamentos entre os subsistemas serem considerados.

A estrutura de cálculo utilizada é uma estrutura hierárquica em dois níveis e o filtro é aplicável a sistemas compostos de N subsistemas. Além disso, visto que os cálculos são feitos nos subsistemas de menor ordem que o sistema global, estes cálculos serão mais precisos, o que conduzirá a um filtro numericamente mais estável [27] (ver ítem 3.4).

2.5.2. A BASE DA ESTRUTURA ALGÉBRICA DO FILTRO MULTINÍVEL

Do ponto de vista prático, uma das mais interessantes propriedades do filtro de Kalman global é a sua natureza recursiva. Dentro do aspecto recursivo, a estimação baseia-se na nova observação que é ortogonal à observação passada disponível, conduzindo a seguinte interpretação geométrica do filtro de Kalman:

Considere um conjunto de vetores reais de observações aleatórias dado por $\underline{Y}_{k_0}, \underline{Y}_{k_0+1}, \dots, \underline{Y}_k, \underline{Y}_{k+1}$. O conjunto de todas as combinações lineares dessas variáveis aleatórias com os coeficientes reais a_i :

$$\sum_{i=k_0}^{k+1} a_i \underline{Y}_{ki}$$

forma um espaço vetorial denotado por $\mathcal{V}_{k+1}$.

Visto que $k+1$ não é fixado, $\mathcal{V}_{k+1}$ pode ser considerado como um sub-espaço de dimensão finita do espaço, de todas as observações possíveis. Então, segundo o teorema de Kalman, o estimador ótimo, $\hat{\underline{X}}_{(k+1/k+1)}$, de $\underline{X}_{(k+1)}$ conhecendo as observações $\underline{Y}_{k_0}, \dots, \underline{Y}_{k+1}$, é dado pela projeção ortogonal de $\underline{X}_{k+1}$ sobre $\mathcal{V}_{k+1}$.

Se os processos estocásticos em questão são Gaussianos, então a projeção ortogonal coincide com a esperança matemática condicional, como visto anteriormente no ítem 1.4, ou seja:

$$\hat{\underline{X}}_{k+1} / k+1 = E \left\{ \underline{X}_{k+1} / \mathcal{Y}_{k+1} \right\} \quad (2.5.2-1)$$

e a forma recursiva do filtro de Kalman é baseada no seguinte teorema [6]:

- TEOREMA 1:

Seja $\underline{X}$ um elemento de um espaço fechado H de variáveis aleatórias, e seja $\hat{\underline{X}}_1$ sua projeção ortogonal sobre um sub-espaço $\mathcal{Y}_1$ de H ($\hat{\underline{X}}_1$ é a melhor estimacão de $\underline{X}$ dentro de $\mathcal{Y}_1$). Seja $\underline{Y}_2$ um vetor de dimensão m de variáveis aleatórias gerando um sub-espaço $\mathcal{Y}_2$ de H e seja $\hat{\underline{Y}}_2$ o vetor de dimensão m , a projeção de $\underline{Y}_2$ sobre $\mathcal{Y}_1$ ($\hat{\underline{Y}}_2$ é a melhor estimacão de $\underline{Y}_2$ dentro do $\mathcal{Y}_1$).

$$\text{Seja } \tilde{\underline{Y}}_2 = \underline{Y}_2 - \hat{\underline{Y}}_2$$

Então, a projeção de $\underline{X}$ sobre o sub-espaço $\mathcal{Y}_1 \oplus \mathcal{Y}_2$, denotado por $\hat{\underline{X}}$, é dada por [6]:

$$\hat{\underline{X}} = \hat{\underline{X}}_1 + E \left[\underline{X} \tilde{\underline{Y}}_2^T \right] \left[E \left[\tilde{\underline{Y}}_2 \tilde{\underline{Y}}_2^T \right] \right]^{-1} \cdot \tilde{\underline{Y}}_2 \quad (2.5.2-2)$$

Considerando que $\mathcal{Y}_1 = \mathcal{Y}_{(k)}$ e $\mathcal{Y}_2$ o sub-espaço gerado por $\underline{Y}_{(k+1)}$, a equação 2.5.2-1 pode ser reescrita como

$$\hat{\underline{X}}_{(k+1/k+1)} = E \left\{ \underline{X}_{(k+1)} / \mathcal{V}_{(k+1)} \right\} =$$

(2.5.2-3)

$$E \left\{ \underline{X}_{(k+1)} / \mathcal{V}_{(k)} \right\} + E \left\{ \underline{X}_{(k+1)} / \tilde{\underline{Y}}_{(k+1/k)} \right\}$$

Portanto, o estimador ótimo de $\underline{X}_{(k+1)}$ é a soma da projeção ortogonal de $\underline{X}_{(k+1)}$ sobre o sub-espaço $\mathcal{V}_{(k)}$ [denotado por $\hat{\underline{X}}_{(k+1/k)}$] e a projeção ortogonal de $\underline{X}_{(k+1)}$ sobre o sub-espaço $\tilde{\underline{Y}}_{(k+1/k)}$, conhecido como espaço vetorial da inovação.

Este último espaço é formado pelo sub-espaço vetorial gerado pelo vetor de observação $\underline{Y}_{(k+1)}$ e pelo sub-espaço gerado pela transformação linear da projeção ortogonal de $\underline{X}_{(k+1)}$ sobre $\mathcal{V}_{(k)}$.

É importante salientar que o espaço vetorial da inovação, não é ortogonal ao vetor de erro $\tilde{\underline{X}}_{(k+1/k)}$, visto que a projeção de $\underline{X}_{(k+1)}$ sobre o espaço da inovação é diferente daquela sobre o espaço $\mathcal{V}_{(k)}$. Por outro lado, o vetor de erro $\tilde{\underline{X}}_{(k+1/k+1)}$ gerado pela projeção de $\underline{X}_{(k+1)}$ no espaço $\mathcal{V}_{(k+1)}$ é ortogonal ao espaço $\mathcal{V}_{(k+1)}$ ou $\mathcal{V}_{(k)} \oplus \tilde{\underline{Y}}_{(k+1/k)}$.

Utilizaremos esta idéia na seqüência do desenvolvimento do algoritmo do filtro de Kalman multinível para sistemas de grande porte.

Considere o sistema discreto linear composto de N sistemas dinâmicos lineares conectados, definido por:

$$\begin{aligned} \underline{X}_i(k+1) = & A_{ii}(k+1, k) \underline{X}_i(k) + \sum_{j=1, i \neq j}^S \\ & A_{ij}(k+1, k) \underline{X}_j(k) + \underline{W}_i(k) \end{aligned} \quad (2.5.2-4)$$

com as equações de observações dadas por:

$$\begin{aligned} \underline{Y}_i(k+1) &= H_i \underline{X}_i(k+1) + \underline{V}_i(k+1) \\ i &= 1, 2, \dots, S \quad \text{e} \quad S \leq N \end{aligned}$$

onde: N é a ordem do sistema;

$\underline{X}_i$ é o vetor de estado $\in R^{n_i}$;

$\underline{Y}_i$ é o vetor de observação $\in R^{m_i}$;

A_{ij} , $i = 1 \dots S$, $j = 1 \dots S$ são as matrizes de transição do sistema;

H_i é a matriz de observação (sub-matriz da matriz H para o sistema global, que é suposta bloco diagonal);

$\underline{W}_i$, $\underline{V}_i$, são vetores ruídos brancos Gaussianos não correlatos, médias nulas e covariâncias Q_i e R_i respectivamente.

Considere o espaço de HILBERT $\mathcal{Y}_{k+1}$ formado à partir de todas as observações do sistema até o instante $(k+1)$.

A estimação ótima (Variância Mínima de Erro) é dada pela equação 2.5.2-3, que traduz algebricamente o resultado geométrico do teorema 1.

Pode-se notar que o primeiro termo da equação 2.5.2-3 pode ser determinado diretamente com base nas observações até o instante (K).

A idéia básica do filtro multinível é decompor o segundo termo da equação 2.5.2-3 de forma que a estimação seja a projeção ortogonal de $\underline{X}_i(k+1)$ sobre o espaço de Hilbert gerado por:

$$\mathcal{V}_{(k)} \oplus \tilde{\underline{Y}}_1(k+1/k) \oplus \tilde{\underline{Y}}_2^1(k+1/k+1) \oplus$$

$$\tilde{\underline{Y}}_3^2(k+1/k+1) \oplus \dots \oplus \tilde{\underline{Y}}_s^{s-1}(k+1/k+1)$$

onde $\tilde{\underline{Y}}_s^{s-1}$ é o sub-espaço gerado pelo sub-espaço das observações $\underline{Y}_s(k+1)$ e por sua projeção sobre o sub-espaço gerado por $\mathcal{V}_{(k)} \oplus \underline{Y}_1(k+1) \oplus \underline{Y}_2(k+1) \oplus \dots \oplus \underline{Y}_{s-1}(k+1)$.

Para aplicar a idéia acima, reescrevemos 2.5.2-3 como:

$$\hat{\underline{X}}_i(k+1/k+1) = E \left\{ \underline{X}_i(k+1) / \mathcal{V}_{(k)}, \right. \\ \left. \underline{Y}_1(k+1), \dots, \underline{Y}_s(k+1) \right\}$$

Manipulando esta expressão, teremos:

$$E \left\{ \underline{X}_i(k+1) / \mathcal{V}_{(k+1)} \right\} = E \left\{ \underline{X}_i(k+1) / \mathcal{V}_{(k)}, \right. \\ \left. \underline{Y}_1(k+1), \dots, \underline{Y}_i(k+1), \underline{Y}_{i+1}(k+1), \right.$$

$$\dots, \underline{Y}_{s-1}(k+1) \} + E \left\{ \underline{X}_i(k+1) / \right.$$
(2.5.2-5)

$$\underline{\tilde{Y}}_{s-1}^{s-1}(k+1 / k+1) \} = E \left\{ \underline{X}_i(k+1) / \underline{Y}(k) \right\} +$$

$$E \left\{ \underline{X}_i(k+1) / \underline{\tilde{Y}}_1(k+1 / k) \right\} +$$

$$\sum_{r=2}^s E \left\{ \underline{X}_i(k+1) / \underline{\tilde{Y}}_r^{r-1}(k+1 / k+1) \right\}$$

onde,

$$\underline{\tilde{Y}}_r^{r-1}(k+1 / k+1) = \underline{Y}_r(k+1) - E \left\{ \underline{Y}_r(k+1) / \right.$$
(2.5.2-6)

$$\underline{Y}(k), \underline{Y}_1(k+1), \dots, \underline{Y}_{r-1}(k+1) \}$$

de forma que, nesta manipulação, foi mostrado essencialmente o particionamento do vetor de observação até o instante $(k+1)$, $\underline{Y}(k+1)$, segundo suas componentes para cada subsistema.

O resultado dessa manipulação, prova o seguinte teorema [27, 28]:

- TEOREMA 2:

A estimação ótima $\underline{X}_i(k+1 / k+1)$ para o i -ésimo subsistema é dada pela soma da projeção ortogonal de $\underline{X}_i(k+1)$ sobre o espaço gerado por todas as medidas até o instante K $\left[\underline{Y}(k) \right]$ com a projeção de $\underline{X}_i(k+1)$ sobre o sub-espaço definido por:

$$\tilde{\underline{Y}}_1 (k+1 / k) \oplus \tilde{\underline{Y}}_2^1 (k+1 / k+1) \oplus \dots \oplus \tilde{\underline{Y}}_n^{n-1} (k+1 / k+1)$$

O teorema fornece a estrutura de base do filtro multinível que, como se pode notar, utiliza a idéia de ortogonalizações sucessivas sobre os sub-espacos definidos por 2.5.2-6, para a obtenção da estrutura algébrica do filtro multinível.

Portanto, as projeções sucessivas de $\underline{X}_i (k+1)$ sobre $\tilde{\underline{Y}}_r^{r-1} (k+1 / k+1)$ adicionadas as projeções representadas pelos dois primeiros termos de 2.5.2-5, fornece a estimacão ótima $\hat{\underline{X}}_i (k+1 / k+1)$ de $\underline{X}_i (k+1)$.

Este processo de ortogonalizações sucessivas pode ser feito de forma independente para cada subsistema, ensejando sua implementacão sobre um sistema de multiprocessamento, como propomos a seguir.

2.5.3. EQUAÇÕES RECURSIVAS DO FILTRO MULTINÍVEL

Neste ítem, será mostrada a obtenção da estrutura algébrica hierarquizada do filtro de Kalman.

Com base no teorema da variância mínima [6],

$$E \left\{ \underline{X} / \underline{Y} \right\} = E \left[\underline{X} \cdot \underline{Y}^T \right] \cdot \left[E \left(\underline{Y} \underline{Y}^T \right) \right]^{-1} \cdot \underline{Y}$$

podemos reescrever 2.5.2-5 como:

$$\hat{\underline{X}}_{i \ (k+1 / k+1)} = \hat{\underline{X}}_{i \ (k+1 / k)} +$$

$$P \underline{X}_{i \ \tilde{Y}_1 \ (k+1 / k)} \cdot P^{-1} \tilde{Y}_1 \tilde{Y}_1 \ (k+1 / k) \cdot$$

(2.5.3-1)

$$\tilde{Y}_1 \ (k+1 / k) + \sum_{r=2}^s P \underline{X}_{i \ \tilde{Y}_{r-1} \ (k+1 / k+1)} \cdot$$

$$P^{-1} \tilde{Y}_{r-1} \tilde{Y}_{r-1} \ (k+1 / k+1) \cdot \underline{\tilde{Y}}_{r-1}^{r-1} \ (k+1 / k+1)$$

O primeiro termo da equação 2.5.3-1 é a predição ótima do estado para o i-ésimo subsistema, de forma que, aplicando-se a média condicional à equação 2.5.2-4 e considerando as condições iniciais, teremos:

$$\hat{\underline{X}}_{i \ (k+1 / k)} = E \left\{ \underline{X}_{i \ (k+1)} / \mathcal{Y}_{(k)} \right\} =$$

$$A_{ii} \ (k+1, k) \cdot E \left\{ \underline{X}_{i \ (k)} / \mathcal{Y}_{(k)} \right\} +$$

(2.5.3-2)

$$\sum_{j=1, i \neq j}^s A_{ij} \ (k+1, k) \cdot E \left\{ \underline{X}_{j \ (k)} / \mathcal{Y}_{(k)} \right\}$$

$$\hat{\underline{X}}_{i \ (k+1 / k)} = A_{ii} \ (k+1, k) \hat{\underline{X}}_{i \ (k / k)} +$$

$$\sum_{j=1, i \neq j}^s A_{ij} \ (k+1, k) \cdot \hat{\underline{X}}_{j \ (k / k)}$$

já que a média do ruído do sistema é nula.

A matriz de covariância do erro de predição é expressa por:

$$\begin{aligned}
 E \left\{ \tilde{\underline{X}}_{i(k+1/k)} \tilde{\underline{X}}_{i(k+1/k)}^T \right\} &= E \left\{ \left[\underline{X}_{i(k+1)} - \hat{\underline{X}}_{i(k+1/k)} \right] \cdot \left[\underline{X}_{i(k+1)} - \hat{\underline{X}}_{i(k+1/k)} \right]^T \right\} \\
 &= E \left\{ \left[A_{ii}(k+1, k) \tilde{\underline{X}}_{i(k/k)} + \sum_{j=1, i \neq j}^n A_{ij}(k+1, k) \tilde{\underline{X}}_{j(k/k)} + \underline{w}_{i(k)} \right] \cdot \left[A_{ii}(k+1, k) \tilde{\underline{X}}_{i(k/k)} + \sum_{j=1, i \neq j}^n A_{ij}(k+1, k) \tilde{\underline{X}}_{j(k/k)} + \underline{w}_{i(k)} \right]^T \right\}
 \end{aligned}$$

$$\begin{aligned}
 P_{X_i X_i(k+1/k)} &= \left[A_{ii}(k+1, k) P_{X_i X_i(k/k)} + \sum_{j=1, i \neq j}^n A_{ij}(k+1, k) \cdot P_{X_i X_j(k/k)}^T \right] \\
 &\quad + A_{ii}^T(k+1, k) + A_{ii}(k+1, k)
 \end{aligned}$$

(2.5.3-3)

$$\sum_{j=1, i \neq j}^n P_{X_i X_j(k/k)} \cdot A_{ij}^T(k+1, k) +$$

$$\sum_{j=1, i \neq j}^n A_{ij}(k+1, k) \cdot \sum_{j=1, i \neq j}^n P X_j X_j(k/k) \\ \cdot A_{ij}^T(k+1, k) + Q_i(k)$$

Os termos que correlacionavam a perturbação $\underline{w}_i(k)$ e os estados, são nulos devido a suposição inicial de não correlação.

A matriz de covariância cruzada do erro de predição é deduzida da seguinte forma:

$$P X_i X_j(k+1/k) = E \left\{ \tilde{X}_i(k+1/k) \tilde{X}_j^T(k+1/k) \right\} \\ = E \left\{ \left[A_{ii}(k+1, k) \tilde{X}_i(k/k) + \sum_{j=1, i \neq j}^n A_{ij}(k+1, k) \cdot \tilde{X}_j(k/k) + \underline{w}_i(k) \right] \right. \\ \cdot \left[A_{jj}(k+1, k) \cdot \tilde{X}_j(k/k) + \sum_{j=1, i \neq j}^n A_{ji}(k+1, k) \cdot \tilde{X}_i(k/k) + \underline{w}_j(k) \right]^T \Big\} \\ = \left[A_{ii}(k+1, k) P X_i X_i(k/k) + \sum_{j=1, i \neq j}^n A_{ij}(k+1, k) P^T X_i X_j(k/k) \right] \cdot \quad (2.5.3-4)$$

$$A^T_{j,i}(k+1, k) + \left[A_{i,i}(k+1, k) \right.$$

$$\sum_{j=1, i \neq j}^n P_{X_i} X_j(k/k) + \sum_{j=1, i \neq j}^n A_{i,j}(k+1, k) \cdot$$

$$\left. \sum_{j=1, i \neq j}^n P_{X_j} X_j(k/k) \cdot A^T_{i,j}(k+1, k) \right] \cdot A^T_{j,j}(k+1, k)$$

Buscando uma distribuição do cálculo da equação 2.5.3-1 dentro do processo recursivo, escrevemos a equação 2.5.3-1 como sendo:

$$\hat{X}_i(k+1/k+1) = \hat{X}_i(k+1/k+1)_l +$$

$$\sum_{r=l+1}^n P_{X_i} \tilde{Y}_r^{r-1}(k+1/k+1) \cdot P^{-1} \tilde{Y}_r^{r-1} \tilde{Y}_r^{r-1}(k+1/k+1)$$

(2.5.3-1a)

$$\cdot \tilde{Y}_r^{r-1}(k+1/k+1)$$

onde:

$$\hat{X}_i(k+1/k+1)_l = \hat{X}_i(k+1/k) + P_{X_i} \tilde{Y}_1(k+1/k)$$

$$P^{-1} \tilde{Y}_1 \tilde{Y}_1(k+1/k) \tilde{Y}_1(k+1/k) +$$

$$\sum_{B=2}^l P X_i \tilde{Y}_B^{B-1} (k+1 / k+1)$$

$$P^{-1} \tilde{Y}_B^{B-1} \tilde{Y}_B^{B-1} (k+1 / k+1) \tilde{Y}_B^{B-1} (k+1 / k+1)$$

ou, ainda:

$$\hat{X}_i (k+1 / k+1)_l = \hat{X}_i (k+1 / k+1)_{l-1} + G_{i_l}^{l-1} (k+1)$$

(2.5.3-5)

$$\tilde{Y}_l^{l-1} (k+1 / k+1)$$

onde:

$$G_{i_l}^{l-1} (k+1) = P X_i \tilde{Y}_l^{l-1} (k+1 / k+1)$$

$$P^{-1} \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1)$$

Por definição:

$$\tilde{X}_i (k+1 / k+1)_l = \tilde{X}_i (k+1) - \hat{X}_i (k+1 / k+1)_l$$

que, a partir de 2.5.3-5, pode ser reescrita como:

$$\tilde{X}_i (k+1 / k+1)_l = \tilde{X}_i (k+1 / k+1)_{l-1} -$$

$$G_{i_l}^{l-1} (k+1) \tilde{Y}_l^{l-1} (k+1 / k+1)$$

Dai, obtemos:

$$P \tilde{X}_i \tilde{X}_i (k+1 / k+1)_l = E \left\{ \left[\begin{array}{c} \tilde{X}_i (k+1 / k+1)_{l-1} \\ G_{i_l}^{l-1}(k+1) \tilde{Y}_l^{l-1}(k+1 / k+1) \end{array} \right] \cdot \left[\begin{array}{c} \tilde{X}_i^T (k+1 / k+1)_{l-1} \\ \tilde{Y}_l^{T l-1}(k+1 / k+1) G_{i_l}^{T l-1}(k+1) \end{array} \right] \right\} =$$

(2.5.3-5a)

$$P \tilde{X}_i \tilde{X}_i (k+1 / k+1)_{l-1} - P \tilde{X}_i^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) \cdot$$

$$G_{i_l}^{T l-1}(k+1) - G_{i_l}^{l-1}(k+1) P \tilde{Y}_l^{l-1} \tilde{X}_i^{l-1} (k+1 / k+1) +$$

$$G_{i_l}^{l-1}(k+1) P \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) \cdot G_{i_l}^{T l-1}(k+1)$$

Considerando os segundo e quarto termos da equação 2.5.3-5a, teremos:

$$\left[\begin{array}{c} G_{i_l}^{l-1}(k+1) P \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) - \\ P \tilde{X}_i^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) \end{array} \right] \cdot G_{i_l}^{T l-1}(k+1) =$$

(2.5.3-5b)

$$\left[P \begin{matrix} \tilde{X}_i^{l-1} & \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right] = G_{i_l}^{T_{l-1}}(k+1)$$

Mas,

$$P \begin{matrix} \tilde{X}_i^{l-1} & \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} = E \left\{ \begin{matrix} \tilde{X}_i^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right. \\ \left. \begin{matrix} \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right\} = E \left\{ \left[\begin{matrix} \hat{\tilde{X}}_i^{l-1} \\ (k+1 \text{ } / \text{ } k+1)_{l-1} \end{matrix} \right. \right. \\ \left. \left. \begin{matrix} \tilde{X}_i^{l-1} \\ (k+1 \text{ } / \text{ } k+1)_{l-1} \end{matrix} \right] \cdot \begin{matrix} \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right\} =$$

(2.5.3-5c)

$$E \left\{ \begin{matrix} \hat{\tilde{X}}_i^{l-1} \\ (k+1 \text{ } / \text{ } k+1)_{l-1} \end{matrix} \cdot \begin{matrix} \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right\} =$$

$$P \begin{matrix} \tilde{X}_i^{l-1} & \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} ,$$

$$\text{já que } E \left\{ \begin{matrix} \hat{\tilde{X}}_i^{l-1} \\ (k+1 \text{ } / \text{ } k+1)_{l-1} \end{matrix} \cdot \begin{matrix} \tilde{Y}_l^{l-1} \\ (k+1 \text{ } / \text{ } k+1) \end{matrix} \right\} = 0$$

Portanto, utilizando as relações 2.5.3-5b e 2.5.3-5c na equação 2.5.3-5a, obtemos:

$$P \tilde{X}_i \tilde{X}_i (k+1 / k+1)_l = P \tilde{X}_i \tilde{X}_i (k+1 / k+1)_{l-1} -$$

(2.5.3-5d)

$$G_{i_l}^{l-1}(k+1) \cdot P \tilde{Y}_l^{l-1} \tilde{X}_i^{l-1}(k+1 / k+1)$$

A matriz de covariância cruzada do erro de estimação de estado no nível l é dada por:

$$P \tilde{X}_i \tilde{X}_j (k+1 / k+1)_l = E \left\{ \left[\tilde{X}_i (k+1 / k+1)_{l-1} - \right. \right.$$

$$\left. G_{i_l}^{l-1}(k+1) \cdot \tilde{Y}_l^{l-1}(k+1 / k+1) \right] \cdot$$

$$\left[\tilde{X}_j (k+1 / k+1)_{l-1} - G_{j_l}^{l-1}(k+1) \cdot \tilde{Y}_l^{l-1}(k+1 / k+1) \right]^T \Bigg\}$$

$$= P \tilde{X}_i \tilde{X}_j (k+1 / k+1)_{l-1} - E \left\{ \tilde{X}_i (k+1 / k+1)_{l-1} \right.$$

$$\cdot \tilde{Y}_l^{T l-1}(k+1 / k+1) \cdot G_{j_l}^{T l-1}(k+1) \Bigg\} - E \left\{ G_{i_l}^{l-1}(k+1) \right.$$

$$\cdot \tilde{Y}_l^{l-1}(k+1 / k+1) \cdot \tilde{X}_j^{T}(k+1 / k+1)_{l-1} \Bigg\} +$$

$$E \left\{ G_{i_l}^{l-1}(k+1) \cdot P_{\tilde{Y}_l}^{l-1} \tilde{Y}_l^{l-1}(k+1 / k+1) \cdot G_{j_l}^{T l-1}(k+1) \right\}$$

$$P_{\tilde{X}_i \tilde{X}_j}^{l-1}(k+1 / k+1)_l = P_{\tilde{X}_i \tilde{X}_j}^{l-1}(k+1 / k+1)_{l-1} -$$

$$P_{\tilde{X}_i}^{l-1} \tilde{Y}_l^{l-1}(k+1 / k+1) \cdot G_{j_l}^{T l-1}(k+1) - G_{i_l}^{l-1}(k+1)$$

(2.5.3-5e)

$$\cdot P_{\tilde{Y}_l}^{l-1} \tilde{X}_j^{l-1}(k+1 / k+1) + G_{i_l}^{l-1}(k+1)$$

$$\cdot P_{\tilde{Y}_l}^{l-1} \tilde{Y}_l^{l-1}(k+1 / k+1) \cdot G_{j_l}^{T l-1}(k+1)$$

onde $i = 1, 2, \dots, s$; $j = 1, 2, \dots, s$, $i \neq j$

A matriz de covariância cruzada do erro de previsão de saída é expressa por:

$$P_{\tilde{Y}_i \tilde{Y}_j}^{l-1}(k+1 / k) = E \left\{ \left[\underline{Y}_i(k+1) - H_i(k+1) \cdot \right. \right.$$

$$\left. \hat{\underline{X}}_i(k+1 / k) \right] \cdot \left[\underline{Y}_j(k+1) - H_j(k+1) \cdot \right.$$

(2.5.3-5f)

$$\left. \left. \hat{\underline{X}}_j(k+1 / k) \right]^T \right\} = H_i(k+1) \cdot P_{\tilde{X}_i \tilde{X}_j}^{l-1}(k+1 / k) \cdot H_j^T(k+1)$$

Podemos escrever que:

$$\tilde{Y}_{l-1}^{l-1}(k+1 / k+1) = Y_{l-1}(k+1) - E \left\{ Y_{l-1}(k+1) / \mathcal{V}_{(k)} \right\},$$

$$\dots, Y_{l-1}(k+1) \left\} = Y_{l-1}(k+1) - E \left\{ Y_{l-1}(k+1) / \right.$$

$$\mathcal{V}_{(k)}, \dots, Y_{l-2}(k+1) \left\} = E \left\{ Y_{l-1}(k+1) / \right.$$

$$\tilde{Y}_{l-1}^{l-2}(k+1 / k+1) \left\} = Y_{l-1}(k+1) - \hat{Y}_{l-1}^{l-2}(k+1 / k+1) -$$

$$P_{Y_{l-1}} \tilde{Y}_{l-1}^{l-2}(k+1 / k+1) \cdot P^{-1} \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2}(k+1 / k+1)$$

$$\cdot \tilde{Y}_{l-1}^{l-2}(k+1 / k+1)$$

$$\tilde{Y}_{l-1}^{l-1}(k+1 / k+1) = \tilde{Y}_{l-1}^{l-2}(k+1 / k+1) -$$

$$H_{l-1}(k+1) P_{X_{l-1}} \tilde{Y}_{l-1}^{l-2}(k+1 / k+1) \cdot P^{-1} \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2}(k+1 / k+1)$$

$$\cdot \tilde{Y}_{l-1}^{l-2}(k+1 / k+1)$$

Seja:

$$\underline{X}_{l-1}(k+1) = \hat{\underline{X}}_{l-1}(k+1 / k+1)_{l-2} + \tilde{\underline{X}}_{l-1}(k+1 / k+1)_{l-2}$$

Dessa relação, podemos escrever que:

$$\tilde{Y}_l^{l-1} (k+1 / k+1) = \tilde{Y}_l^{l-2} (k+1 / k+1) -$$

$$P \tilde{Y}_l^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) \cdot P^{-1} \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1)$$

$$\therefore \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) = \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) -$$

$$G_{l-1}^{l-2} (k+1) \cdot \tilde{Y}_{l-1}^{l-2} (k+1 / k+1)$$

onde:

$$G_{l-1}^{l-2} (k+1) = P \tilde{Y}_l^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) \cdot$$

$$P^{-1} \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1)$$

De onde resulta que:

$$P \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) = P \tilde{Y}_l^{l-2} \tilde{Y}_l^{l-2} (k+1 / k+1) -$$

$$P \tilde{Y}_l^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) \cdot G_{l-1}^{T l-2} (k+1) - G_{l-1}^{l-2} (k+1)$$

(2.5.3-6)

$$\therefore P \tilde{Y}_{l-1}^{l-2} \tilde{Y}_l^{l-2} (k+1 / k+1) + G_{l-1}^{l-2} (k+1) \cdot$$

$$P \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2} (k+1 / k+1) \cdot G_{l-1}^{T l-2} (k+1)$$

Substituindo $G_l^{l-2}(k+1)$ no último termo da equação 2.5.3-6, teremos:

$$P \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1}(k+1/k) = P \tilde{Y}_l^{l-2} \tilde{Y}_l^{l-2}(k+1/k) - G_l^{l-2}(k+1) \cdot P \tilde{Y}_{l-1}^{l-2} \tilde{Y}_{l-1}^{l-2}(k+1/k) \quad (2.5.3-6a)$$

No nível de recursividade l , podemos escrever:

$$\tilde{Y}_l(k+1/k) = \underline{Y}_l(k+1) - \hat{\underline{Y}}_l(k+1/k)$$

$$\tilde{Y}_l(k+1/k) = H_l(k+1) \cdot \underline{X}_l(k+1) + \underline{V}_l(k+1) -$$

$$H_l(k+1) \hat{\underline{X}}_l(k+1/k) = H_l(k+1/k) \cdot \tilde{\underline{X}}_l(k+1/k) + \underline{V}_l(k+1)$$

Então,

$$P \tilde{Y}_l \tilde{Y}_l(k+1/k) = E \left\{ \left[H_l(k+1) \cdot \tilde{\underline{X}}_l(k+1/k) + \underline{V}_l(k+1) \right] \cdot \left[H_l(k+1) \cdot \tilde{\underline{X}}_l(k+1/k) + \underline{V}_l(k+1) \right]^T \right\}$$

$$P \tilde{Y}_l \tilde{Y}_l(k+1/k) = H_l(k+1) \cdot P \tilde{\underline{X}}_l \tilde{\underline{X}}_l(k+1/k) \cdot$$

$$H_l^T(k+1) + R_l(k+1)$$

Além disso,

$$P \tilde{X}_i^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) = E \left\{ \tilde{X}_i (k+1 / k+1)_{l-1} \right. \\ \left. \tilde{Y}_l^{T l-1} (k+1 / k+1) \right\} = P \tilde{X}_i^{l-1} \tilde{X}_l^{l-1} (k+1 / k+1) \cdot H_l^T (k+1) \quad (2.5.3-7)$$

e:

$$\tilde{X}_j (k+1 / k+1)_l = \tilde{X}_j (k+1 / k+1)_{l-1} - G_j^{l-1} (k+1) \\ \tilde{Y}_l^{l-1} (k+1 / k+1) \quad (2.5.3-8)$$

onde:

$$G_j^{l-1} (k+1) = P \tilde{X}_j^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1) \cdot P^{-1} \tilde{Y}_l^{l-1} \tilde{Y}_l^{l-1} (k+1 / k+1)$$

Então, à partir da equação 2.5.3.-8, obtemos:

$$P \tilde{X}_i \tilde{X}_j (k+1 / k+1)_l = E \left\{ \left[\tilde{X}_i (k+1 / k+1)_{l-1} - \right. \right. \\ \left. \left. G_i^{l-1} (k+1) \cdot \tilde{Y}_l^{l-1} (k+1 / k+1) \right] \cdot \left[\tilde{X}_j^T (k+1 / k+1)_{l-1} - \right. \right. \\ \left. \left. \tilde{Y}_l^{T l-1} (k+1 / k+1) \cdot G_j^{T l-1} (k+1) \right] \right\}$$

$$P \tilde{X}_i^{l-1} \tilde{X}_j^{l-1} (k+1 / k+1) = P \tilde{X}_i^{l-1} \tilde{X}_j^{l-1} (k+1 / k+1) - G_{i,l}^{l-1} (k+1)$$

(2.5.3-9)

$$P \tilde{Y}_l^{l-1} \tilde{X}_j^{l-1} (k+1 / k+1)$$

Completando, dessa forma, as equações recursivas do filtro que seguem as seguintes etapas de execução:

1. A partir das equações 2.5.3-2 à 2.5.3-4, calculamos a predição ótima e as matrizes de covariância correspondentes;
2. Fazemos $l = 1$ (frisando que $\hat{\underline{X}}_i (k+1 / k+1)_0 = \underline{X}_i (k+1 / k)$ e $P \tilde{\underline{X}}_i^0 \tilde{\underline{X}}_j^0 (k+1 / k+1) = P \tilde{\underline{X}}_i \tilde{\underline{X}}_j (k+1 / k)$ para $i = 1$ à S e $j = 1$ à S) e utilizando as equações 2.5.3-5 à 2.5.3-9, calculamos a estimação ótima de estado e as matrizes de covariância correspondentes;
3. Se $l = s$, então foram estimados todos os vetores de estado relativos aos subsistemas. Senão, faça $l = l + 1$ e volte à etapa 2.

É importante observar que o algoritmo do filtro multinível aqui mostrado e o filtro global de Kalman, são algebricamente equivalentes, porém, veremos no item 3.4 que as propriedades numéricas do algoritmo multinível são mais interessantes, já que os cálculos são realizados sobre sistemas de menor ordem.

CAPÍTULO 3

PROGRAMAÇÃO PARALELA DO FILTRO DE KALMAN MULTINÍVEL

3.1. INTRODUÇÃO

Este capítulo aborda principalmente o ambiente para estudo do paralelismo dos algoritmos multiníveis de estimação de estado e a viabilização da paralelização do Filtro de Kalman multinível num ambiente de multiprogramação com tempo compartilhado bem como um procedimento para a sua implementação numa arquitetura de multiprocessamento.

3.2. DESCRIÇÃO DAS ESTRUTURAS HIERÁRQUICAS DE CÁLCULO DO FILTRO DE KALMAN MULTINÍVEL

Para efeito de simplificação da descrição, particularizaremos a aplicação do algoritmo, ao caso do particionamento do sistema global em dois subsistemas, como mostrado a seguir.

Obedecendo as seqüências de execução do algoritmo, podemos escrever as equações 2.5.3-2 à 2.5.3-4 para dois subsistemas como sendo:

A predição ótima:

$$\hat{\underline{x}}_i(k+1/k) = A_{i,i}(k+1,k) \hat{\underline{x}}_i(k/k) + A_{i,j}(k+1,k) \hat{\underline{x}}_j(k/k) \quad (3.2-1)$$

e as matrizes de covariância do erro de predição:

$$P_{x_1, x_1}^{\sim \sim} (k+1/k) = \left[\sigma_{11} (k+1/k) + \sigma_{1j}^* (k+1/k) \right] A_{11}^T (k+1, k) + \sigma_{1j} (k+1/k) A_{1j}^T (k+1, k) + Q_{11}^{**} (k) \quad (3.2-2)$$

$$P_{x_1, x_2}^{\sim \sim} (k+1/k) = \left[\sigma_{11} (k+1/k) + \sigma_{1j}^* (k+1/k) \right] A_{21}^T (k+1, k) + \left[\sigma_{12} (k+1/k) + \sigma_{12}^{**} (k+1, k) \right] A_{22}^T (k+1, k) \quad (3.2-3)$$

onde:

$$\sigma_{11} (k+1/k) = A_{11} (k+1, k) \cdot P_{x_1, x_1}^{\sim \sim} (k/k)$$

$$\sigma_{1j} (k+1/k) = A_{1j} (k+1, k) \cdot P_{x_1, x_j}^{\sim \sim} (k/k)$$

$$\sigma_{1j}^* (k+1/k) = A_{1j} (k+1, k) \cdot P_{x_1, x_j}^{T \sim \sim} (k/k)$$

e,

$$Q_{11}^{**} (k) = Q_{11} (k) + Q_{11}^* (k) \quad (3.2-3a)$$

onde:

$$Q_{11}^* (k) = \sigma_{1j}^{**} (k+1/k) \cdot A_{1j}^T (k+1, k) \quad (3.2-3b)$$

e,

$$\sigma_{1j}^{**} (k+1/k) = A_{1j} (k+1, k) \cdot P_{x_j, x_j}^{\sim \sim} (k/k)$$

Explorando o paralelismo natural dos cálculos da equação 3.2-1 à 3.2-3, podemos representá-lo segundo as etapas 1 e 2, numa estrutura hierárquica em dois níveis como segue:

ETAPA 1 -

Cada subsistema envia ao coordenador as matrizes $A_{ij}(k+1, k)$ e $P_{x_i, x_j}^{\sim \sim}(k/k)$, onde $i=1, 2$; $j=1, 2$, e $i \neq j$.

O coordenador calcula as matrizes $Q_i^*(k)$ ($i=1, 2$) expressas por 3.2-3b e envia ao subsistema i , que calcula a matriz $Q_i^{**}(k)$, expressa por 3.2-3a. Esta etapa é representada na Figura 3.1.

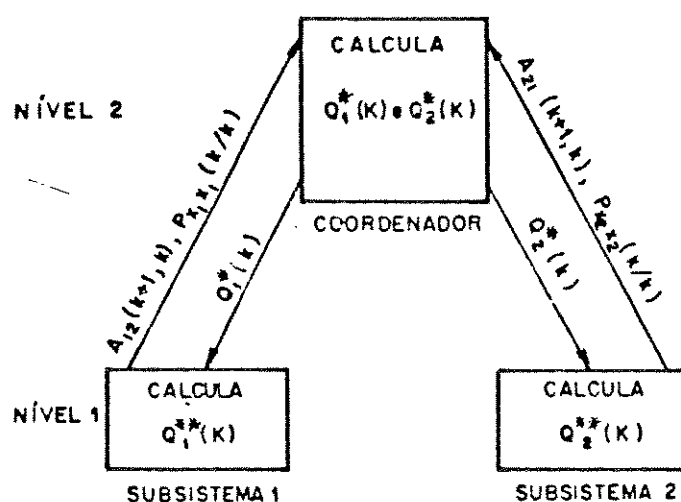


Figura 3.1

ETAPA 2 -

Esta etapa é concernente a predição ótima, e como mostra a Figura 3.2, o coordenador comunica do subsistema j para o subsistema i , a predição filtrada ótima $\hat{x}_j(k/k)$ e fornece-lhes a matriz de covariância cruzada $P_{x_i, x_j}^{\sim \sim}(k/k)$, ($i \neq j$).

Cada subsistema calcula a predição ótima $\hat{x}_i(k+1/k)$ dada por 3.2-1 e a matriz de covariância $P_{x_i, x_i}^{\sim \sim}(k+1/k)$ dada por 3.2-3 correspondente ao erro de predição.

As matrizes $\begin{bmatrix} \sigma_{11}(k+1/k) + \sigma_{12}^*(k+1/k) \\ \sigma_{12}(k+1/k) \end{bmatrix}$, $\sigma_{12}(k+1/k)$, $H_1(k+1)$ e $P_{x_1 x_1}(k+1/k)$, são enviadas ao coordenador pelo subsistema 1, enquanto ao mesmo tempo, o subsistema 2, envia-lhe a matriz $A_2(k+1, k)$. Finalmente, o coordenador calcula a matriz de covariância cruzada do erro de predição $P_{x_1 x_2}(k+1/k)$.

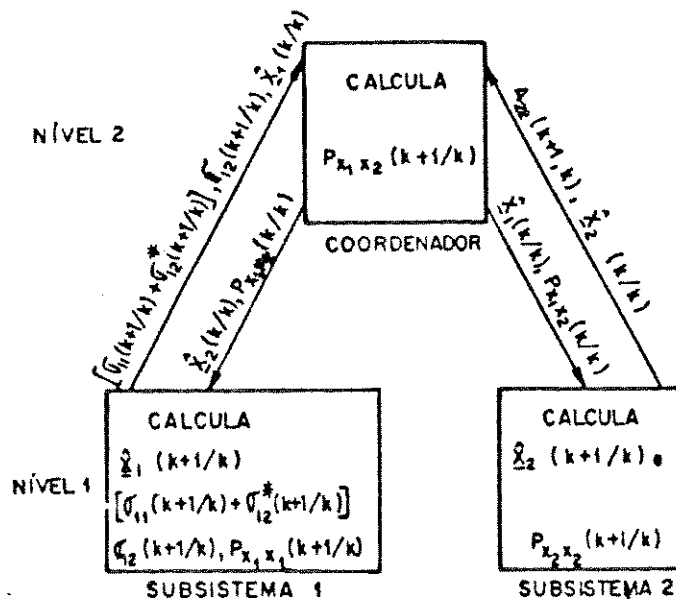


Figura 3.2

Fazendo $\ell=1$:

Para o subsistema 1, a equação 2.5.3-5, pode ser reescrita

como:

$$\hat{x}_1(k+1/k+1) = \hat{x}_1(k+1/k) + G_{11}^u(k+1) \tilde{y}_1(k+1/k) \quad (3.2-4a)$$

onde:

$$\begin{aligned} G_{11}^u(k+1) &= P_{x_1 x_1}(k+1/k) \cdot P_{x_1 x_1}^{-1}(k+1/k) \\ &= P_{x_1 x_1}(k+1/k) \cdot H_1^T(k+1) \cdot P_{x_1 x_1}^{-1}(k+1/k) \end{aligned}$$

$$\tilde{y}_a(k+1/k) = y_a(k+1) - H_a(k+1) \hat{x}_a(k+1/k), \text{ então}$$

$$\begin{aligned} \tilde{y}_a(k+1/k) &= H_a(k+1) x_a(k+1) + v_a(k+1) - H_a \hat{x}_a(k+1/k) \\ &= H_a(k+1) \cdot \tilde{x}_a(k+1/k) + v_a(k+1) \end{aligned}$$

Resultando que:

$$P_{\tilde{x}_a \tilde{x}_a}(k+1/k) = E \left\{ \left[H_a(k+1) \cdot \tilde{x}_a(k+1/k) + v_a(k+1) \right] \cdot \right.$$

$$\left. \left[H_a(k+1) \tilde{x}_a(k+1/k) + v_a(k+1) \right]^T \right\}$$

$$= H_a(k+1) P_{\tilde{x}_a \tilde{x}_a}(k+1/k) H_a^T(k+1) + R_a(k+1)$$

De 2.5.3-5d,

$$P_{\tilde{x}_a \tilde{x}_a}(k+1/k+1) = G_{aa}(k+1) \cdot P_{\tilde{x}_a \tilde{x}_a}(k+1/k) \quad (3.2-5a)$$

onde:

$$G_{aa}(k+1) = I - G_{aa}^u(k+1) \cdot H_a(k+1)$$

A matriz de covariância cruzada do erro da estimação de estado no nível k dada pela equação 2.5.3-5e resulta em:

$$P_{\tilde{x}_a \tilde{x}_j}(k+1/k+1) = G_{aa}(k+1) \cdot P_{\tilde{x}_a \tilde{x}_j}(k+1/k) \cdot G_{jj}^T(k+1) \quad (3.2-6a)$$

A matriz de covariância cruzada do erro de previsão da saída dada pela equação 2.5.3-5f é:

$$P_{x_1 x_1}^{j,j}(k+1/k) = H_1^T(k+1) \cdot P_{x_1 x_1}^{j,j}(k+1/k) \cdot H_1(k+1) \quad (3.2-7a)$$

Para o j-ésimo subsistema:

$$\hat{x}_j(k+1/k+1) = \hat{x}_j(k+1/k) + G_{j1}^o(k+1) \cdot \tilde{y}_1(k+1/k) \quad (3.2-4b)$$

onde:

$$\begin{aligned} G_{j1}^o(k+1) &= P_{x_j x_1}^{j,j}(k+1/k) \cdot P_{x_1 x_1}^{-1}(k+1/k) \\ &= P_{x_j x_1}^{j,j}(k+1/k) \cdot H_1^T(k+1) \cdot P_{x_1 x_1}^{-1}(k+1/k) \end{aligned}$$

De 2.5.3-5d,

$$P_{x_j x_j}^{j,j}(k+1/k+1) = G_{jj}(k+1) P_{x_j x_j}^{j,j}(k+1/k) \quad (3.2-5b)$$

onde:

$$G_{jj}(k+1) = I - G_{j1}^o(k+1) \cdot H_1(k+1)$$

As matrizes de covariância cruzada dos erros de estimação no nível l e de previsão de saída para o subsistema j são:

$$P_{x_j x_1}^{j,j}(k+1/k+1) = P_{x_1 x_j}^T(k+1/k+1) \quad (3.2-6b)$$

$$P_{x_j x_1}^{j,j}(k+1/k) = P_{x_1 x_j}^T(k+1/k) \quad (3.2-7a)$$

ETAPA 3:

Nesta etapa, completamos o cálculo hierarquizado em dois níveis do primeiro termo da equação 2.5.3-1a para ambos subsistemas, conforme Figura 3.3.

O subsistema 1 calcula $\hat{x}_1(k+1/k+1)$ e a matriz $P_{x_1 x_1}(k+1/k+1)$ dados por 3.2-4a e 3.2-5a, enquanto o subsistema 2 calcula $\hat{x}_2(k+1/k+1)$ e $P_{x_2 x_2}(k+1/k+1)$ dados por 3.2-4b e 3.2-5b e finalmente é enviado ao nível 2 $G_{11}(k+1)$ e $H_1(k+1)$, $(i=1,2)$, para que este calcule as matrizes $P_{x_1 x_2}(k+1/k+1)$ e $P_{y_1 y_2}(k+1/k)$ dadas pelas equações (3.2-6) e (3.2-7).

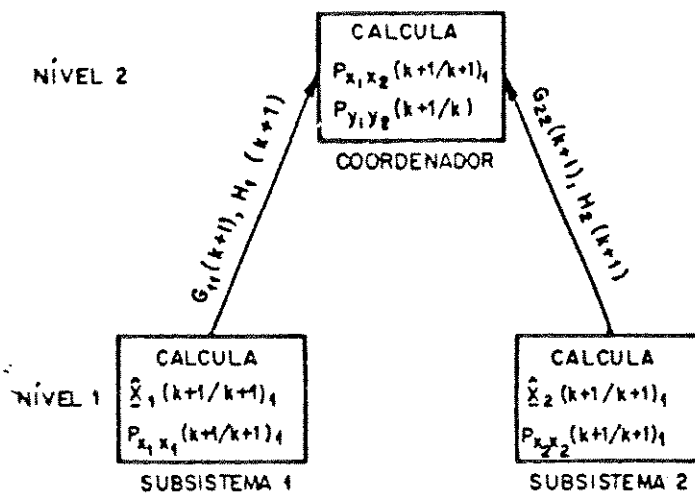


Figura 3.3

Referente ao segundo termo da equação 2.5.3-1a, podemos escrever que:

Para o subsistema 1:

$$\begin{aligned}
 P_{\tilde{x}_1, \tilde{z}_2}(k+1/k+1) &= P_{\tilde{x}_1, \tilde{z}_2}(k+1/k) \\
 &= P_{\tilde{x}_1, \tilde{z}_2}(k+1/k) \cdot H_z^T(k+1) \\
 &= E \left\{ \left[\tilde{x}_1(k+1/k) - G_{1,1}^o \tilde{y}_1(k+1/k) \right] \cdot \left[\tilde{z}_2(k+1/k) - \right. \right. \\
 &\quad \left. \left. G_{2,1}^o(k+1) \cdot \tilde{y}_1(k+1/k) \right]^T \right\} \cdot H_z^T(k+1) \\
 &= \left\{ P_{\tilde{x}_1, \tilde{z}_2}(k+1/k) - E \left[\tilde{x}_1(k+1/k) \tilde{y}_1^T(k+1/k) \cdot \right. \right. \\
 &\quad \left. \left. G_{2,1}^{T^o}(k+1) \right] - E \left[G_{1,1}^o(k+1) \cdot \tilde{y}_1^T(k+1/k) \cdot \right. \right. \\
 &\quad \left. \left. \tilde{x}_2^T(k+1/k) \right] + E \left[G_{1,1}^o(k+1) \cdot \tilde{y}_1^T(k+1/k) \cdot \right. \right. \\
 &\quad \left. \left. \tilde{y}_1^T(k+1/k) \cdot G_{2,1}^{T^o}(k+1) \right] \cdot H_z^T(k+1) \right\} \\
 &= \left\{ P_{\tilde{x}_1, \tilde{z}_2}(k+1/k) - P_{\tilde{x}_1, \tilde{z}_1}(k+1/k) \cdot G_{2,1}^{T^o}(k+1) - \right. \\
 &\quad \left. G_{1,1}^o(k+1) P_{\tilde{z}_1, \tilde{z}_2}(k+1/k) + \right. \\
 &\quad \left. G_{1,1}^o(k+1) P_{\tilde{z}_1, \tilde{z}_1}(k+1/k) G_{2,1}^{T^o}(k+1) \right] \cdot \\
 &\quad \left. H_z^T(k+1) - P_{\tilde{x}_1, \tilde{z}_1}(k+1/k) \right\}
 \end{aligned}$$

$$P_{\tilde{x}_1, \tilde{z}_2}(k+1/k+1) = G_{1,1}^o(k+1) \cdot P_{\tilde{x}_1, \tilde{z}_2}(k+1/k) \cdot H_z^T(k+1) \quad (3.2-8a)$$

$$\begin{aligned}
\tilde{y}_2^1(k+1/k+1) &= y_2(k+1) - E\left\{y_2(k+1)/Y(k), y_1(k+1)\right\} \\
&= \tilde{y}_2(k+1/k) + \hat{y}_2(k+1/k) - \hat{y}_2(k+1/k) - \\
&\quad E\left\{y_2(k+1)/\tilde{y}_1(k+1/k)\right\} \\
&= \tilde{y}_2(k+1/k) - P_{\substack{22 \\ 12}}^{\sim \sim}(k+1/k) \cdot P_{\substack{11 \\ 11}}^{-1} \tilde{y}_1(k+1/k) - \\
&\quad \tilde{y}_1(k+1/k)
\end{aligned}$$

$$\tilde{y}_2^1(k+1/k+1) = \tilde{y}_2(k+1/k) - G_{21}^*(k+1) \tilde{y}_1(k+1/k) \quad (3.2-9a)$$

onde:

$$G_{21}^*(k+1) = P_{\substack{12 \\ 11}}^{\sim \sim}(k+1/k) \cdot P_{\substack{11 \\ 11}}^{-1} \tilde{y}_1(k+1/k)$$

E portanto:

$$\begin{aligned}
P_{\substack{12 \\ 22}}^{\sim \sim}(k+1/k+1) &= E\left\{\left[\tilde{y}_2(k+1/k) - G_{21}^*(k+1) \tilde{y}_1(k+1/k)\right] \cdot \right. \\
&\quad \left. \left[\tilde{y}_2(k+1/k) - G_{21}^*(k+1) \tilde{y}_1(k+1/k)\right]^T\right\} \\
&= P_{\substack{22 \\ 22}}^{\sim \sim}(k+1/k) - P_{\substack{22 \\ 21}}^{\sim \sim}(k+1/k) \cdot G_{21}^{*T}(k+1) - \\
&\quad G_{21}^*(k+1) P_{\substack{12 \\ 11}}^{\sim \sim}(k+1/k) + G_{21}^*(k+1) P_{\substack{11 \\ 11}}^{\sim \sim}(k+1/k) \cdot \\
&\quad G_{21}^{*T}(k+1)
\end{aligned}$$

$$P_{z_1 z_2}^{-1} (k+1/k+1) = P_{z_1 z_2}^{-1} (k+1/k) - G_{z_1}^{*} (k+1) P_{z_1 z_2}^{-1} (k+1/k) \quad (3.2-10a)$$

Utilizando o mesmo procedimento, podemos escrever:

$$P_{z_1 z_2}^{-1} (k+1/k+1) = G_{z_1}^{*} (k+1) P_{z_1 z_2}^{-1} (k+1/k) \cdot G_{z_2}^{*T} (k+1) - P_{z_1 z_2}^{-1} (k+1/k) \quad (3.2-11a)$$

Para o subsistema j utiliza-se o mesmo procedimento, de forma que podemos descrever o cálculo hierarquizado das equações 3.2-8 à 3.2-11, conforme a Etapa 4.

ETAPA 4:

Esta etapa relata o cálculo das equações 3.2-8 à 3.2-11 e a troca de informação necessária para a sua execução entre o primeiro e o segundo nível como ilustra a Figura 3.4.

O coordenador comunica do subsistema j para o subsistema i o vetor do erro de predição da saída $\tilde{y}_j(k+1/k)$, a matriz de observação $H_j(k+1)$ e a matriz de covariância $P_{z_j z_j}^{-1}(k+1/k)$ e fornece ao nível 1 as matrizes $P_{z_i z_j}^{-1}(k+1/k)$ e $P_{z_j z_i}^{-1}(k+1/k)$, ($i \neq j$).

Assim, o subsistema i calcula o vetor de erro $\tilde{y}_i^1(k+1/k+1)$ e sua matriz de covariância $P_{z_i z_i}^{-1}(k+1/k+1)$ e ainda a matriz $P_{z_i z_j}^{-1}$.

Finalmente, cada subsistema envia ao nível 2 a matriz G_{ij}^{*} para que este calcule a matriz de covariância $P_{z_i z_j}^{-1}(k+1/k+1)$.

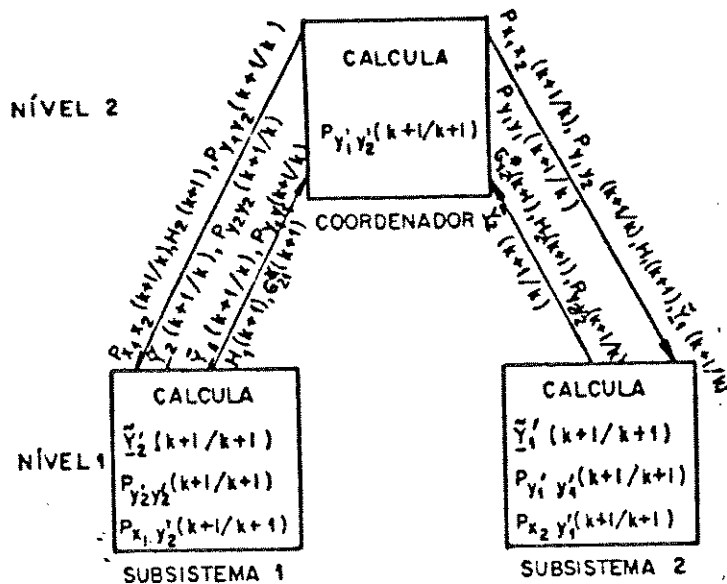


Figura 3.4

Finalizando, calcula-se a equação 2.5.3-1a, ou seja:

$$\hat{\underline{x}}_i(k+1/k+1) = \hat{\underline{x}}_i(k+1/k+1)_a + G_{i,j}(k+1) \cdot \tilde{\underline{y}}_j'(k+1/k+1) \quad (3.2-12)$$

onde:

$$G_{i,j}(k+1) = P_{\tilde{\underline{x}}_i, \tilde{\underline{y}}_j}(k+1/k+1) \cdot P_{\tilde{\underline{y}}_j, \tilde{\underline{y}}_j}^{-1}(k+1/k+1)$$

e as matrizes de covariância correspondentes:

$$P_{\tilde{\underline{x}}_i, \tilde{\underline{x}}_i}(k+1/k+1) = P_{\tilde{\underline{x}}_i, \tilde{\underline{x}}_i}(k+1/k+1)_a + G_{i,j}(k+1) \cdot \left[P_{\tilde{\underline{y}}_j, \tilde{\underline{y}}_j}(k+1/k+1) \right]$$

$$G_{i,j}^T(k+1) = P_{\tilde{\underline{x}}_i, \tilde{\underline{y}}_j}^T(k+1/k+1) \cdot P_{\tilde{\underline{y}}_j, \tilde{\underline{y}}_j}^{-1}(k+1/k+1) \cdot G_{i,j}^T(k+1)$$

(3.2-13)

$$\begin{aligned}
 P_{x_1 x_2}^{\sim \sim}(k+1/k+1) &= P_{x_1 x_2}^{\sim \sim}(k+1/k+1) + \left[P_{x_1 x_2}^{\sim \sim}(k+1/k+1) \cdot G_{11}^{*T}(k+1) + \right. \\
 &\quad \left. G_{12}(k+1) \cdot P_{x_2 x_2}^{\sim \sim}(k+1/k+1) \right] \cdot G_{11}^T(k+1) + \\
 &\quad G_{12}(k+1) \cdot G_{11}^{*T} \cdot P_{x_2 x_2}^T(k+1/k+1)
 \end{aligned} \quad (3.2-14)$$

A distribuição dos cálculos das equações 3.2-13 à 3.2-14, concluindo o processo de estimação, é ilustrada na Figura 3.5 e relatada na seguinte etapa:

ETAPA 5:

Cada subsistema estima o seu vetor de estado e envia ao coordenador as matrizes necessárias para o cálculo da matriz de covariância cruzada $P_{x_1 x_2}^{\sim \sim}(k+1/k+1)$, finalizando assim o processo de estimação, ilustrado pela Figura 3.5.

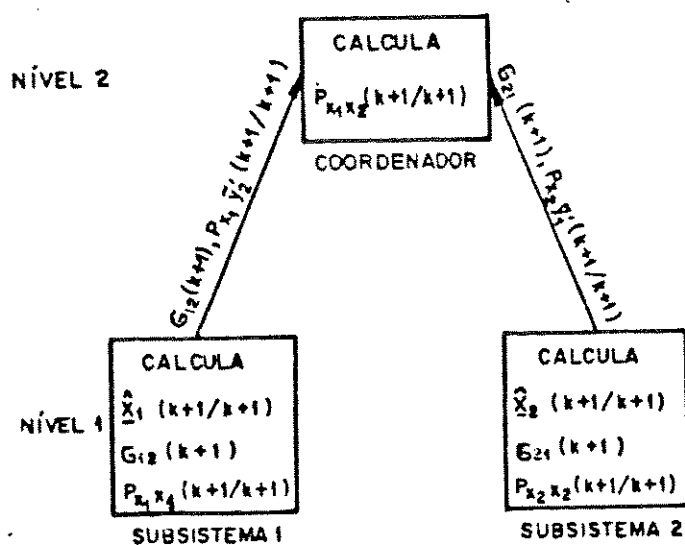


Figura 3.5

é importante notar que:

- 1) A troca de informação entre os subsistemas ocorre nas etapas 2 e 4;
- 2) A comunicação entre os subsistemas poderia ser feita diretamente entre eles e não via coordenador.

3.3. AMBIENTES DE MULTIPROGRAMAÇÃO E PROCESSAMENTO DO FILTRO DE KALMAN MULTINÍVEL

3.3.1. INTRODUÇÃO

As estruturas dos sistemas de programação e de processamento paralelo para a estimação multinível do filtro de Kalman, variam conforme hardware e software de suporte correspondente [16, 31-34]. A escolha dessas estruturas deve ser feita, tomando-se como critério, a que oferece melhores mecanismos para a aplicação das estratégias de coordenação e decomposição do problema e dos requisitos de comunicação delas conseqüente. Neste item, nos restringimos às explanações do sistema Unix que fornece suporte para programação concorrente e com o qual o filtro de Kalman paralelo foi implementado.

3.3.2. SISTEMAS DE MULTIPROGRAMAÇÃO

Os sistemas de multiprogramação permitem que vários processos (programas em execução) utilizem concorrentemente um único processador.

Na concorrência desses processos o tempo alocado pela CPU a cada um deles, varia e é determinado pelas interrupções, pela política de gerenciamento do processador e pelo uso do tempo do processador pelo processo.

3.3.3. AMBIENTE UNIX PARA PROGRAMAÇÃO CONCORRENTE

O sistema Unix [30] é simples e dispõe de um interessante ambiente de programação.

A Figura 3.6 mostra a arquitetura de um sistema Unix. O hardware no centro da figura, dá suporte ao sistema operacional com serviços básicos. O sistema Kernel, na camada superior, está isolado dos programas para o usuário.

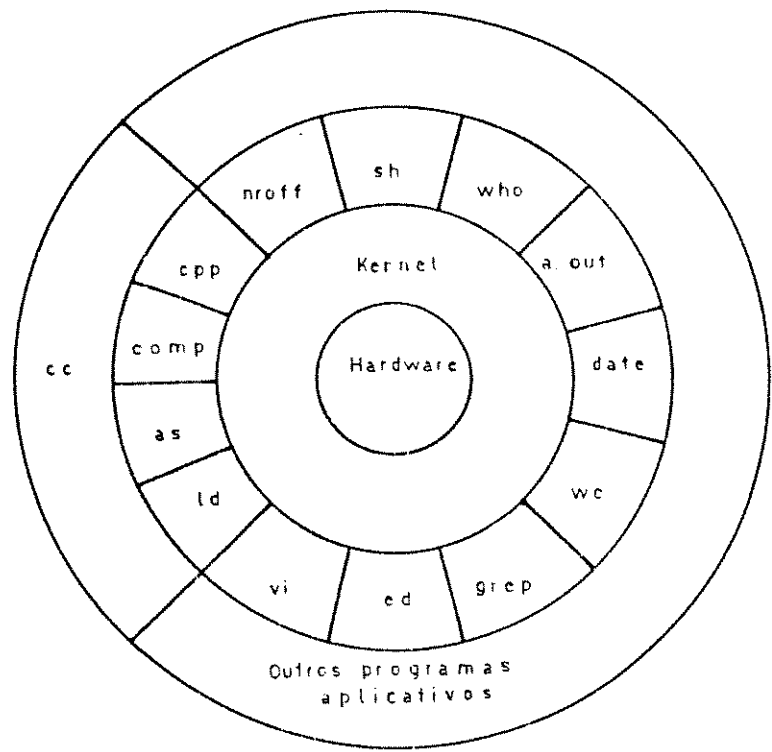


Figura 3.6 - Arquitetura do Sistema Operacional Unix

3.3.3.1. O Shell

O Shell é um interpretador tanto de comandos interativos quanto de linguagem de programação a nível de comando. Ele provê a interface com o usuário e dispõe de estruturas de controle sofisticadas para repetição de sequência de instruções e para o teste de condições, possibilitando a construção de programas eficientes.

Um programa é qualquer comando ou sequência de comandos armazenados num arquivo executável, e um processo é uma instância do programa em execução. Muitos processos podem ser executados concorrentemente sobre sistemas Unix (essa característica é chamada de multiprogramação ou multitarefa).

3.3.3.2. Sistemas de Arquivo

Os arquivos podem ser utilizados como uma área comum de dados por vários programas, fornecendo meios de estabelecer uma comunicação entre programas concorrentes. Os arquivos estão organizados em diretórios numa estrutura hierárquica. Os arquivos presentes em um diretório podem ser arquivos ordinários, arquivos diretórios e arquivos especiais (Dispositivos de E/S).

Os arquivos e dispositivos são tratados da mesma forma. Consequentemente, os programas podem escrever ou ler em um dispositivo E/S (terminal, por exemplo), arquivos ou um outro programa. Este recurso é utilizado em [31] para computação distribuída.

Um arquivo ordinário é uma sequência de bytes. Os arquivos não têm uma estrutura interna. Os bytes existentes num arquivo são apenas aqueles postos pelo usuário ou programa. Não há prealocação de espaço para arquivos e um arquivo é tão grande quanto seja necessário. O acesso a um arquivo pode ser tanto sequencial quanto aleatório.

Embora os arquivos não tenham uma estrutura, o(s) programa(s) que interagem com outro(s) programa(s) pode(m) impor uma estrutura para tratar um fluxo de dados.

3.3.3.3. O Pipeline

O pipe conecta a saída padrão de um programa à entrada padrão de outro programa. Por exemplo, se incluíssemos o vetor de controle nas equações de estado 2.5.2-4, poderíamos calculá-lo "off-line" utilizando um procedimento de controle ótimo e utilizá-lo no cálculo do filtro multinível da seguinte forma:

```
cat CONTROLE | FKM
```

O programa cat para o usuário ler os vetores de controle armazenados no arquivo CONTROLE e o símbolo |, instrui o interpretador shell para criar um pipe que os transmite para o programa FKM (Filtro de Kalman Multinível).

Os pipes são uni-direcionais e os programas por eles conectados são executados concorrentemente. O sistema operacional se encarrega da sincronização e "bufferização" durante a conexão dos programas.

3.3.3.4. Redirecionamento de Entrada e de Saída

A entrada e a saída padrões são normalmente conectadas ao terminal de computador.

O caracter ">>" permite redirecionar a saída à primeira posição livre de um arquivo especificado.

O redirecionamento de entrada permite que um programa receba dados de um arquivo especificado.

PROGRAMA 1 >> ARQUIVO R, o caractere ">>" faz com que o PROGRAMA 1 posicione seqüencialmente seus dados de saída no arquivo R.

3.3.3.5. Programas Utilitários

Os programas utilitários do sistema Unix (Ver Figura 3.6), oferecem grandes vantagens, dentre elas, a de contabilizar com certa precisão (décimo ou centésimo de segundos, dependendo da versão Unix) o tempo de execução de um programa.

Por exemplo: `time PROGRAMA 1`. Ao usuário será apresentado após o término da execução do PROGRAMA 1, o seguinte resultado:

t_1 real t_2 user t_3 sys

onde: t_2 tempo gasto pela CPU somente com a execução do PROGRAMA 1.

t_1 tempo gasto pela CPU, servindo todos usuários (o ambiente Unix é também um sistema multi-usuário)

t_3tempo gasto pelo sistema para execução do PROGRAMA 1.

3.3.4. MULTIPROGRAMAÇÃO EM TEMPO COMPARTILHADO E FILTRAGEM DE KALMAN PARALELA

Pela análise das estruturas hierárquicas de cálculo do Filtro de Kalman Multinível descritas no item 3.2, é possível notar a existência de um paralelismo natural conseqüente da teoria de projeções múltiplas utilizada para o desenvolvimento do filtro. Dessa forma, podemos implementar um filtro de Kalman paralelo utilizando multiprogramação e/ou multiprocessamento. Num sistema operacional de tempo compartilhado, o escalonamento dos processos é um dos elementos principais.

O compartilhamento do tempo é a suspensão de um processo e o seu reinício posteriormente. No Unix, como na maioria dos sistemas de tempo compartilhado, a atividade de suspensão e reinício ocorre muitas vezes por segundo, deixando transparecer externamente que a CPU está executando vários programas simultaneamente. Quanto maior o número de programas a serem executados, menor será o tempo alocado a cada um deles. O sistema operacional calcula dinamicamente as prioridades dos processos para determinar qual processo inativo, mas pronto para execução será executado quando o processo atual ativo for suspenso.

Esta forma de multiplexagem do processador permite a execução de cada programa sobre um processador virtual: um processador abstrato que implementa uma unidade de processamento paralelo, Figura 3.7.

Portanto, a multiprogramação é importante por fornecer um ambiente abstrato para o estudo de paralelismo, sem complicações desnecessárias com detalhes de implementação que apesar de importante em outras ocasiões, podem ser ignorados nesta fase.

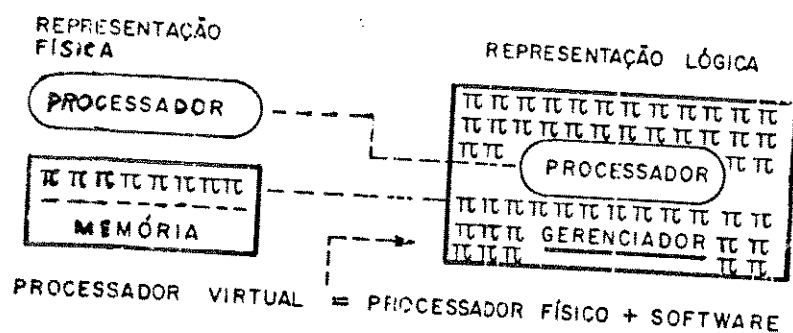


Figura 3.7 - Ambiente de Multiprogramação

O conceito chave para multiprogramação com tempo compartilhado é o seguinte:

3.3.4.1 - Execuções em retaguarda e vanguarda de programas

O interpretador Shell pode iniciar a execução de um programa e deixá-lo executando, enquanto outros programas podem também iniciar suas execuções. Os nomes retaguarda e vanguarda especificam a ordem em que os programas estão sendo executados.

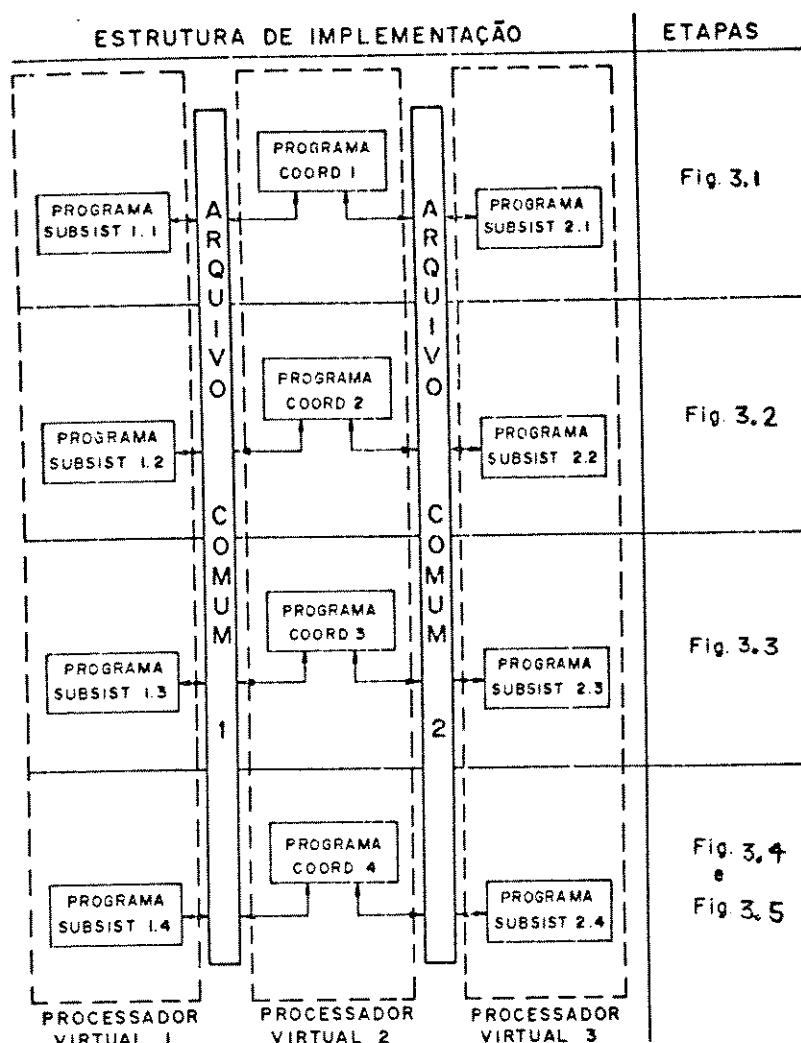
Os caracteres a seguir, são utilizados para determinar uma ordem de execução:

- a) ";" indica uma execução sequencial. Por exemplo: PROGRAMA 1; PROGRAMA 2. O PROGRAMA 2 só será executado, após o término da execução do PROGRAMA 1.
- b) "&" indica uma execução assíncrona (concorrente). Por exemplo: PROGRAMA 1 & PROGRAMA 2. O PROGRAMA 1 será executado em retaguarda e o PROGRAMA 2 em vanguarda, concorrentemente.

3.3.5. IMPLEMENTAÇÃO EM MULTIPROGRAMAÇÃO COM TEMPO COMPARTILHADO

Para a paralelização do filtro de Kalman, aproveitamos as estruturas hierárquicas descritas no item 3.2.

Os subproblemas do nível inferior e o coordenador foram resolvidos sequencialmente numa primeira fase. Na segunda fase, associamos a sequência das cinco etapas descritas no item 3.2 para, formar a seguinte estrutura de implementação:



Nesta estrutura, os arquivos ordinários são utilizados como áreas comuns de dados.

Para a execução da implementação do filtro de Kalman multinível, via multiprogramação com tempo compartilhado, utilizamos o computador Digirede 8000, com CPU 68020 e sistema operacional DIGIX (versão 1.4a-5) e a linguagem de programação Pascal.

Na linguagem de programação Shell do referido Unix, o programa que viabiliza a estrutura de implementação é o seguinte:

ESTIMAÇÃO

gerar arquivos de inicialização;

WHILE lista de comando

DO

(subsist 1.1 & subsist 2.1); coord1

(subsist 1.2 & subsist 2.2); coord2

(subsist 1.3 & subsist 2.4); coord3

(subsist 1.4 & subsist 2.4); coord4

DONE

1ª Fase:

Implementação sequencial

(linguagem Pascal)

PROGRAMA ESTIMAÇÃO; ...

PROCEDURE ETAPA 1;

BEGIN ...

calcular $Q_i^*(k)$, $Q_i^{**}(k)$ ($i=1,2$)

END;

PROCEDURE ETAPA 2; ...

BEGIN ...

calcular $\hat{x}_i(k+1/k)$, $\left[\sigma_{i,i} + \sigma_{i,i}^*\right]$, $\sigma_{i,i}$, $P_{x_i, x_i}(k+1/k)$, $P_{x_i, x_j}(k+1/k)$

($i=1,2$; $i \neq j$)

END;

PROCEDURE ETAPA 3; ...

BEGIN ...

calcular $\hat{x}_i(k+1/k+1)$, $P_{x_i, x_i}(k+1/k+1)$, $P_{x_i, x_j}(k+1/k+1)$,
 $P_{x_i, x_j}(k+1/k)$ (i=1,2; i≠j)

END;

PROCEDURE ETAPA 4; ...

BEGIN ...

calcular $\tilde{y}_i^j(k+1/k+1)$, $P_{y_i, y_i}(k+1/k+1)$, $P_{y_i, y_j}(k+1/k+1)$,
 $P_{y_i, y_j}(k+1/k+1)$ (i=1,2; i≠j)

END;

PROCEDURE ETAPA 5; ...

BEGIN ...

calcular $\hat{x}_i(k+1/k+1)$, $G_{ij}(k+1)$, $P_{x_i, x_i}(k+1/k+1)$, $P_{x_i, x_j}(k+1/k+1)$
 (i=1,2; i≠j)

END;

(** PROGRAMA PRINCIPAL **)

BEGIN

inicializar variáveis

REPEAT

ETAPA 1;

ETAPA 2;

ETAPA 3;

ETAPA 4;

ETAPA 5;

UNTIL n

END.

onde: n é o número de iterações desejado.

Nesta 1ª fase (implementação seqüencial) as comunicações representadas nas etapas dispensam o uso de arquivos ordinários.

Devido a simetria das tarefas paralelas e dos dados comunicados entre os níveis, descreveremos para a 2ª fase, apenas os programas relativos ao subsistema 1 e ao coordenador.

2ª Fase:

Implementação via multiprogramação com tempo compartilhado.

PROGRAMA S1.1;

VAR

arq1, arqR: FILE OF REAL;

BEGIN

ler P_{11} (k/k) do arq L;

escreve A_{12} e P_{11} no arq 1;

END.

PROGRAMA S1.2;

VAR

arq1, arqL: FILE OF REAL;

BEGIN

ler $Q_1^*(k), P_{x_1 x_2}(k/k)$ do arq1;

calcula $\hat{x}_1(k+1/k), [\sigma_{11} + \sigma_{12}^w], \sigma_{12}, P_{x_1 x_1}(k+1/k)$

escreve $\left\{ \begin{array}{l} [\sigma_{11} + \sigma_{12}^w], \sigma_{12}, \hat{x}_1(k/k), A_{11} \text{ no arq 1;} \\ P_{x_1 x_1}(k+1/k); \hat{x}_1(k+1/k) \text{ no arq L;} \end{array} \right.$

END.

PROGRAMA S1.3;

VAR

arq1, arqL: FILE OF REAL;

BEGIN

ler $P_{x_1 x_1}(k+1/k), \hat{x}_1(k+1/k)$ do arq L;

calcula $G_{11}, \tilde{y}_1(k+1/k), P_{w_1 w_1}(k+1/k);$

escreve $G_{11}, H_1, \tilde{y}_1(k+1/k), P_{w_1 w_1}(k+1/k)$ no arq 1;

calcula $\hat{x}_1(k+1/k+1), P_{x_1 x_1}(k+1/k+1);$

END.

PROGRAMA S1.4;

VAR

arq1, arqL: FILE OF REAL;

BEGIN

ler $P_{x_1 x_2}(k+1/k)$, $\hat{H}_z$, $P_{x_1 x_2}(k+1/k)$, $\tilde{y}_z(k+1/k)$, $P_{x_2 x_2}(k+1/k)$ no

arq 1;

calcula G_{z1}^* e escreve no arq 1;

calcula $\tilde{y}_z^1(k+1/k+1)$, $P_{x_2 x_2}^1(k+1/k+1)$, $P_{x_1 x_2}^1(k+1/k+1)$, $G_{12}(k+1)$;

escreve $G_{12}(k+1)$ e $P_{x_1 x_2}^1(k+1/k+1)$ no arq 1;

calcule $\hat{x}_1(k+1/k+1)$, $P_{x_1 x_1}(k+1/k+1)$ e escreve no arq L;

END.

PROGRAMA C1;

VAR

arq1, arq2, arqB: file of real;

BEGIN

ler $\begin{cases} A_{12} \text{ e } P_{x_1 x_1}(k/k) \text{ do arq 1;} \\ A_{21} \text{ e } P_{x_2 x_2}(k/k) \text{ do arq 2;} \end{cases}$

calcula Q_1^* e Q_2^* ;

ler $P_{x_1 x_2}^1(k/k)$ do arq B;

escreve $\begin{cases} Q_1^*(k) \text{ e } P_{x_1 x_2}^T(k/k) \text{ no arq 1;} \\ Q_2^*(k) \text{ e } P_{x_1 x_2}(k/k) \text{ no arq 2;} \end{cases}$

END.

PROGRAMA C2;

VAR

arq1, arq2, arqB: FILE OF REAL;

BEGIN

ler $\left\{ \begin{array}{l} \left[\sigma_{11} + \sigma_{12}^* \right], \sigma_{12}, \hat{x}_1(k/k), A_{11} \text{ do arq 1;} \\ A_{22}, \hat{x}_2(k/k) \text{ do arq 2;} \end{array} \right.$

calcula $P_{x_1 x_2}(k+1/k)$ e escreve no arq B;

escreve $\left\{ \begin{array}{l} \hat{x}_2 \text{ no arq 1;} \\ \hat{x}_1 \text{ no arq 2;} \end{array} \right.$

END.

PROGRAMA C3;

VAR

arq1, arq2, arqB: FILE OF REAL;

BEGIN

ler $\left\{ \begin{array}{l} P_{x_1 x_2}(k+1/k) \text{ do arq B;} \\ G_{11}, H_1, \tilde{y}_1(k+1/k), P_{u_1 u_1}(k+1/k) \text{ do arq 1;} \\ G_{22}, H_2, \tilde{y}_2(k+1/k), P_{u_2 u_2}(k+1/k) \text{ do arq 2;} \end{array} \right.$

calcula $P_{x_1 x_2}(k+1/k), P_{x_1}(k+1/k);$

$$\left\{ \begin{array}{l} P_{x_1 x_2} (k+1/k), H_{x_1 x_2} (k+1/k), \hat{y}_2 (k+1/k), \\ P_{x_2 x_2} (k+1/k) \text{ no arq 1;} \\ \\ P_{x_1 x_2} (k+1/k), H_{x_1 x_2} (k+1/k), \hat{y}_1 (k+1/k), \\ P_{x_1 x_1} (k+1/k) \text{ no arq 2;} \end{array} \right.$$
 escreve

END.

PROGRAMA C4;

VAR

arq1, arq2, arqB: FILE OF REAL;

BEGIN

$$\text{ler} \left\{ \begin{array}{l} G_{x_1}^*, G_{x_2}^*, P_{x_1 x_2}^* (k+1/k+1) \text{ do arq 1;} \\ G_{x_2}^*, G_{x_1}^*, P_{x_2 x_1}^* (k+1/k+1) \text{ do arq 2;} \end{array} \right.$$

calcula $P_{x_1 x_2}^* (k+1/k+1), P_{x_2 x_1}^* (k+1/k+1)$

ler $P_{x_1 x_2}^* (k+1/k+1), P_{x_1 x_2}^* (k+1/k)$ do arq B;

calcula $P_{x_1 x_2}^* (k+1/k+1)$ e escreve no arq B;

END.

onde: os arquivos B e L simulam memórias locais.

3.4. ASPECTOS DE CONVERGÊNCIA DO FILTRO MULTINÍVEL

Neste item comentamos um aspecto de grande importância no processo de estimação conhecido como "convergência". A "convergência" num processo de estimação pode ser caracterizada de muitas formas. Por exemplo, em [35] a convergência da estimação com o emprego do Filtro de Kalman Global é caracterizada com o fato de durante o período de operação do filtro, os erros nas estimações assumirem valores abaixo dos seus valores eficazes (rms) estimados pelas equações de filtragem. A explicação mais comum, para isto é que a matriz de covariância calculada não atinge valores fora das proporções reais, impossibilitando ponderações indevidas nas estimações e assim evitando por exemplo que as observações subseqüentes sejam ignoradas. O efeito da divergência pode ocorrer devido a inúmeras causas. As mais comuns são: imprecisão do modelo utilizado e erros acumulados nas operações computacionais. A fonte dessas duas causas pode englobar imprecisões na modelagem das variáveis utilizadas nas equações do filtro, não linearidades, polarizações que não são compensadas por falta de conhecimento dos modelos estatísticos para a geração do filtro, e o acúmulo de erros de arredondamento e truncamento durante as computações. Em [35] são investigados os mecanismos pelos quais tais fontes de erro degradam o desempenho do filtro de Kalman Global, e sugeridas algumas técnicas que provêem uma redução dessas degradações.

Em [36, 37], são feitas análises teóricas da propagação do erro devido ao arredondamento numérico, para quatro diferentes implementações do filtro de Kalman: filtro de Kalman Global, filtro raiz quadrada da covariância, filtro raiz quadrada da informação e o filtro raiz quadrada de Chandrasekhar que foram desenvolvidos com o objetivo de evitar os problemas numéricos presentes na implementação do Filtro de Kalman Global.

Apesar de não dispor de um estudo de convergência para o Filtro de Kalman Multinível, é possível afirmar [27] que este apresenta melhores propriedades numéricas do que o Filtro de Kalman Global, pelo fato dos cálculos serem realizados sobre sistemas de menor ordem, reduzindo o acúmulo de erros durante as operações computacionais. Portanto, além da redução do tempo de processamento, este fator torna a aplicação do Filtro de Kalman Multinível preponderante em relação ao Filtro de Kalman Global, particularmente em se tratando de aplicações a sistemas de grande porte.

3.5. IMPLICAÇÕES TÉCNICAS DA UTILIZAÇÃO DAS ESTRUTURAS MULTINÍVEIS PARA FILTRAGEM PARALELA

3.5.1 INTRODUÇÃO

Uma grande classe de problemas em controle e processamento de sinais requer uma computação em tempo real rápida e precisa de uma grande quantidade de dados como uma consequência das ordens elevadas dos modelos e altas taxas de amostragem. Com o rápido desenvolvimento em sistemas multiprocessadores, tornou-se de grande interesse, obter elevadas velocidades do processamento através do paralelismo em algoritmos até então considerados sequenciais por natureza. Recentes desenvolvimentos em sistemas com rede de sensores integrados têm incentivado ainda mais a busca de algoritmos que permitam o processamento descentralizado dos sinais. O paralelismo surge naturalmente em sistemas onde os dados são obtidos por diferentes subsistemas localizados em posições geograficamente distribuídas. Em tais estruturas multiníveis ocorre uma quantidade de processamento local nos processadores locais ou nós, e os resultados desses são comunicados para um processador central (também referenciado como coordenador central ou centro de fusão) de maior grau na hierarquia. Este processador central ou de fusão combina as informações dos vários processadores locais para produzir uma estimação global.

Em [40] é citada uma estrutura paralela do filtro de Kalman para sistemas multisensores onde as informações obtidas dos sensores locais são combinadas para gerar a estimação global, porém apresentando a desvantagem pelo fato de cada processador utilizar a estimação global do estado do sistema no passo anterior para gerar sua estimação presente. Ou seja, existe a necessidade de comunicação do processador central para os processadores locais no nível mais baixo da hierarquia.

Isto dificulta um processamento paralelo nos nós locais, e portanto, provoca uma redução na velocidade do processamento.

É possível também notar que o algoritmo baseado em ortogonalizações sucessivas sobre os subespaços das medidas de cada subsistema, descrito no item 2.5, apresenta uma desvantagem no sentido de que cada processador local precisa conhecer implicitamente observações de outros processadores para gerar sua própria estimação.

Em [40], é também citado um algoritmo de estimação hierárquica, onde cada processador local, em cada iteração, gera sua própria estimação baseado somente nos seus próprios dados, comunicando essa estimação bem como sua matriz de covariância do erro associada, para o processador central, o qual combina todas as estimações locais e suas covariâncias do erro correspondentes para gerar a estimação global e a covariância do erro. Porém, neste algoritmo, a hierarquização da equação de predição e covariância do erro correspondente no filtro de Kalman não é considerada.

Uma outra aproximação para paralelização das equações do filtro de Kalman é dispersar a sequência de medidas em diferentes segmentos e processar cada segmento em paralelo e posteriormente, num nível mais alto da hierarquia, combinar a informação obtida de cada segmento para produzir a estimação global. Os resultados de [38, 39] podem ser citados como referência.

3.5.2. APLICAÇÕES EM REDES INTEGRADAS MULTISENSORIAIS

Neste item abordamos os cenários físicos das redes e, suas implicações quanto ao algoritmo paralelo de filtragem utilizado.

Consideremos um ambiente multisensor onde muitos sensores, em geral, observam o mesmo sistema dinâmico, onde cada sensor é ligado a um processador local. Uma questão a se considerar, no sistema de rede integrada, é o instante da disponibilidade da informação em cada um desses sensores (time-alignment). Na maioria dos casos, supõe-se que as medidas são disponíveis de forma síncrona, isto é, os sensores são amostrados no mesmo instante. Enquanto esta suposição de sincronismo é razoável se os sensores estiveram geograficamente concentrados, torna-se menos justificável no caso de sensores geograficamente dispersos.

Em tais casos, é mais apropriado supor que as observações estejam disponíveis quando as saídas dos diferentes sensores forem amostradas ciclicamente em sucessão, numa taxa uniforme.

As diferentes situações são identificadas como segue:

- 1) Todos os sensores concentrados: quando todos os sensores são colocados na mesma vizinhança como mostrado na Figura 3.8. Supondo a independência entre os ruídos das medidas em cada sensor, a estrutura do filtro de Kalman paralelo do item 2.5 pode ser aplicada.
- 2) Todos os sensores dispersos: esta situação corresponde ao caso onde há sensores colocados em partes distantes no processo, como mostrado na Figura 3.9.

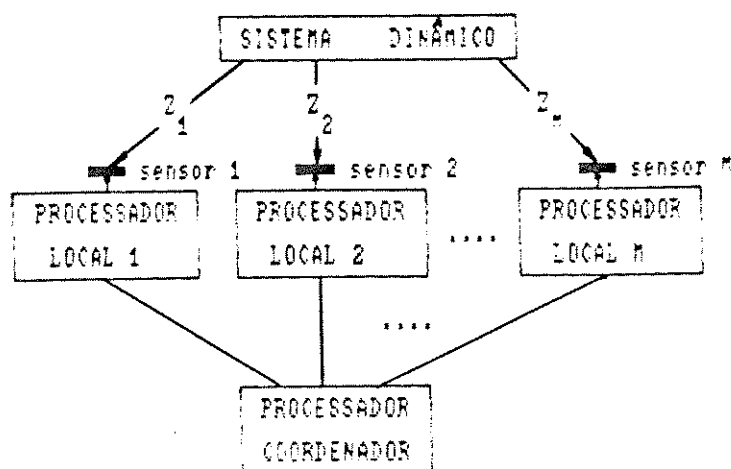


Figura 3.8

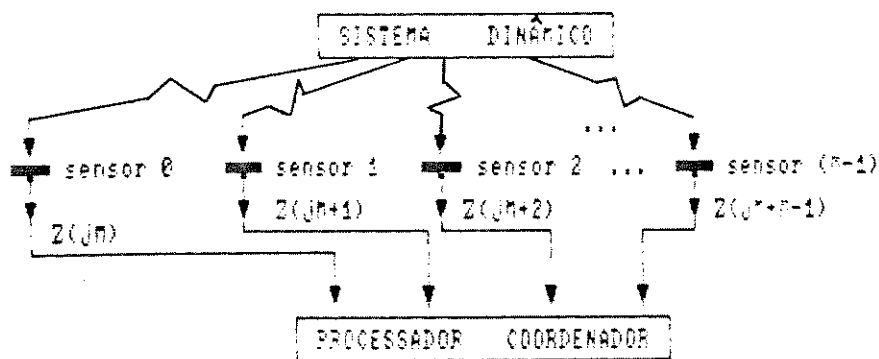


Figura 3.9

Neste cenário, as observações são obtidas sequencialmente no i -ésimo sensor nos instantes $((jn+i), j=0,1,2,\dots)$. O algoritmo do item 2.5 ainda pode ser aplicado a este tipo de cenário, porém, com mais dificuldade do que no cenário anterior, pelo fato de não dispormos sincronamente das observações. Já que a estrutura do cálculo do filtro é linear, teoricamente é possível processar a

estimação com base nas observações já recebidas de alguns sensores, e a medida que as observações dos outros sensores fossem recebidas, fariam parte da estimação nos termos restantes das projeções múltiplas. Entretanto, deve ser elaborado um estudo das implicações práticas dessa alternativa para verificar sua viabilidade.

Os cenários 1) e 2) combinados dão origem a um 3) mostrado na Figura 3.10.

A estrutura hierárquica da Figura 3.10 baseia-se nas seguintes suposições:

- 1) As K_i observações na i -ésima localidade ($i=1, \dots, M$) são obtidas sincronamente, onde K_i é o número de sensores na i -ésima localidade.
- 2) Porém, as medidas, e conseqüentemente as estimações do estado, nas diferentes localizações são supostas assíncronas (sequenciais no tempo).
- 3) Os K_i ruídos nas observações na i -ésima localidade são supostos não correlatos.

Com base nesses três cenários, e uma série de suposições quanto as equações e matrizes do sistema, em [40] é desenvolvido um algoritmo de estimação para cada um desses cenários de forma que os processadores locais não se comunicam e a comunicação entre os níveis é somente no sentido dos processadores locais para o processador central.

Um fator importante que ainda deve ser ressaltado quanto à aplicação dos algoritmos de estimação de estados a estes cenários é o da existência de um algoritmo que não conduza à uma degradação do processo de estimação como um todo, quando a informação de um ou alguns desses sensores não for mais disponível, e/ou quando o módulo coordenador não estiver operando por alguma falha do sistema.

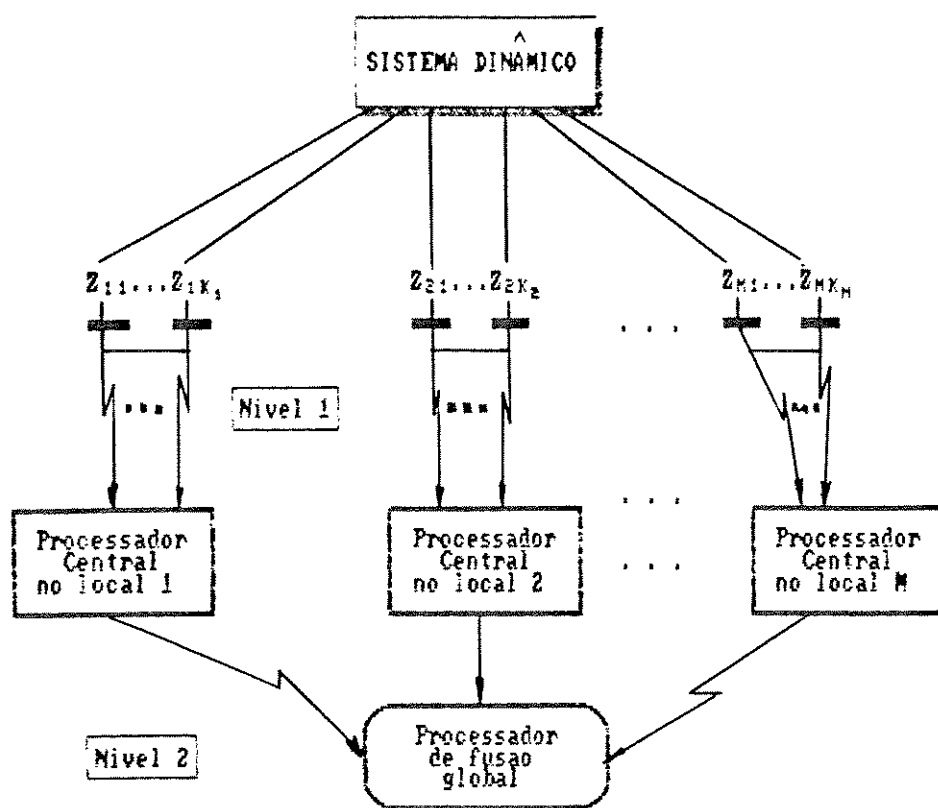


Figura 3.10

CAPÍTULO 4

ANÁLISE DE DESEMPENHO COMPUTACIONAL

4.1. INTRODUÇÃO

O estudo da eficiência computacional do filtro multinível em relação ao filtro de Kalman global é dividido em duas categorias, uma é quanto a requisição de memória, e a outra é quanto ao tempo de processamento. Sob algumas hipóteses, desenvolvemos expressões para avaliação do desempenho computacional do algoritmo paralelo sobre sistemas multiprocessadores, tomando como parâmetros o tempo computacional e a quantidade de dados comunicados.

A requisição de memória do filtro de Kalman multinível é aproximadamente a mesma para o filtro de Kalman global, porém a requisição de memória do filtro multinível pode ser convenientemente distribuída num sistema multiprocessador.

Um estudo dos requisitos computacionais (tempo de computação por iteração e memória requerida) na implementação do filtro de Kalman Global, como funções das dimensões de importantes matrizes do sistema, é desenvolvido em [41]. Nele, são discutidos dois tipos de processamento das observações: simultâneo e sequencial, com o intuito de principalmente reduzir o tempo de processamento das equações do filtro.

4.2. REQUISIÇÃO DE TEMPO DE PROCESSAMENTO

Para efeito de comparação do tempo de processamento sequencial entre o filtro multinível e o filtro de Kalman Global, tomaremos o número de multiplicações necessárias para cada um deles.

Supondo H bloco diagonal e cada subsistema com o mesmo número de estados e saídas, obtemos com a utilização da Tabela 4.2-1 e das equações do item 2.5 que:

A. Número de multiplicações para o filtro de Kalman Global:

$$M_{FKG} = 1,5 n^2 + 1,5 n^3 + nm \left(\frac{1}{N} + \frac{m+1}{2N} + m+1 + \frac{n+1}{2} \right) + \frac{m^2 (3m+1)}{2}$$

B. Número de multiplicações para o filtro de Kalman Multinível:

$$M_{FKM} = 1,5 n^2 + 1,5 n^3 + N \left\{ \frac{mn}{N^2} + \frac{mn(m+N)}{2N^3} + m^2 \frac{\left(\frac{3m+1}{N} \right)}{2N^2} \right. \\ \left. + N \left[\frac{n^2 m}{N^3} + \frac{nm^2}{N^3} + \frac{nm}{N^2} + \frac{nm(n+N)}{2N^3} \right] \right. \\ \left. + \frac{N(N-1)}{2} \frac{n^2 m}{N^3} \right\}$$

onde: N é o número de subsistemas.

Variável	Equação **	Computações	Número de	
			Multipl.	Adições
$\hat{x}(k+1/k)$	$A(k+1,k)\hat{x}(k/k)$	$A\hat{x}$	n^2	n^2-n
$P(k+1/k)$	$A(k+1,k)P(k/k)$ $A^T(k+1,k)+Q(k)$	PA^T $A(PA^T)$ $A(PA^T)+Q$	n^3 n^3 0	n^3-n n^3-n^2 n^2
$G(k+1)$	$P(k+1/k)H^T(k+1) [$ $H(k+1)P(k+1/k)H^T$ $(k+1)+R(k+1)]^{-1}$	PH^T $H(PH^T)$ $H(PH^T)+R$ $[H(PH^T)+R]^{-1}$	n^2m nm^2 0 m^3	n^2m-nm nm^2-m^2 m^2 m^3
$\hat{x}(k+1/K+1)$	$\hat{x}(k+1/k)+G(k+1)$ $[Y(k+1)-H(k+1)$ $x(k+1/k)]$	$H\hat{x}$ $Y-(H\hat{x})$ $G[Y-(H\hat{x})]$ $\hat{x}+G[Y-(H\hat{x})]$	nm 0 nm 0	$nm-m$ m $mn-n$ n
$P(k+1/k+1)$	$[I-G(k+1)H(k+1)]$ $P(k+1/k)$	GH $I-(GH)$ $[I-(GH)]P$	n^2m 0 n^3	n^2m-n^3 n^2 n^3-n^2
TOTAL				

TABELA 4.2-1 - Número de Operações Elementares para o FILTRO de KALMAN GLOBAL.

* Para manter as propriedades de definida positiva e simétrica, a forma alternativa é utilizada

$$P(K+1/K+1) = [I-G(K+1)H(K+1)] P(K+1/K) [I-G(K+1)H(K+1)]^T + G(K+1) R(K+1) G^T(K+1)$$

e então os números de multiplicações e adições aumentarão de $n^3 + n^2m + nm$ e $n^3 + n^2m - n$ respectivamente.

** Todas as variáveis que aparecem nestas equações são definidas na Tabela 4.2-2.

VARIÁVEL	DEFINIÇÃO	DIMENSÃO
$\hat{x}(k/k)$	Estado estimado em k dado Y_k	$n \times 1$
$P(k/k)$	Matriz de covariância do erro em $\hat{x}(k/k)$	$n \times n$
$A(k+1, k)$	Matriz de transição de estado k para k+1	$n \times n$
$Q(k)$	Matriz de covariância do ruído do sistema	$n \times n$
$\hat{x}(k+1/k)$	Estado estimado em k+1 dado Y_k	$n \times 1$
$P(k+1/k)$	Matriz de covariância do erro em $\hat{x}(k+1/k)$	$n \times n$
$H(k+1)$	Matriz de observação	$m \times n$
$R(k+1)$	Matriz de covariância do ruído na observação	$m \times m$
$G(k+1)$	Matriz de ganho do filtro de KALMAN em k+1	$n \times m$
$Y(k+1)$	Vetor de observação em k+1	$n \times 1$

TABELA 4.2-2 - Variáveis do Filtro.

Portanto, é possível verificar que para os sistemas de grande porte, o filtro multinível apresenta um menor tempo de cálculo se m e n são aproximadamente iguais. Mas, se $n \gg m$, ambos os filtros exigem aproximadamente o mesmo tempo de cálculo.

Consideremos o seguinte exemplo:

Seja $m = n/2$, então o número de multiplicações para o filtro de Kalman global é:

$$M_{FKG} = 1,5 n^2 + 1,5 n^3 + \frac{n^2}{2} \left[\frac{1}{N} + \frac{n}{N} + \frac{n}{2} + 1 + \frac{n+1}{2} \right] + \frac{n^2 (3n + 2)}{16}$$

e o número de multiplicações para o filtro multinível é:

$$M_{FKM} = 1,5 n^2 + 1,5 n^3 + N \left\{ \frac{n^2}{2N^2} + \frac{n^2 \left(\frac{3n}{2N} + 1 \right)}{8N^2} + \frac{n^3}{2N^3} + \right. \\ \left. + N \left[\frac{n^2}{2N^3} + \frac{n^3}{4N^3} + \frac{n^2}{2N^2} + \frac{n^2(N+n)}{4N^3} \right] + \frac{N(N-1)}{2} \cdot \frac{n^3}{2N^3} \right\} \quad (4.2-1)$$

a) Para $n=10$ e $N=2$, teremos:

$M = 2.700$ multiplicações

$M_{FKG} = 2.553$ multiplicações

Um ganho de: $\left[(M - M_{FKG}) / M_{FKG} \right] 100 = 5,5 \%$

Particionando num número maior de subsistemas, ou seja:

b) Para $n=10$ e $N=5$, teremos:

$M_{FKG} = 2.535$ multiplicações

$M_{FKM} = 2.165$ multiplicações

Um ganho de: $\left[(M_{FKG} - M_{FKM}) / M_{FKG} \right] 100 \approx 14,5 \%$

A Figura 4.2-1 nos fornece o ganho no tempo de processamento seqüencial em função de N .

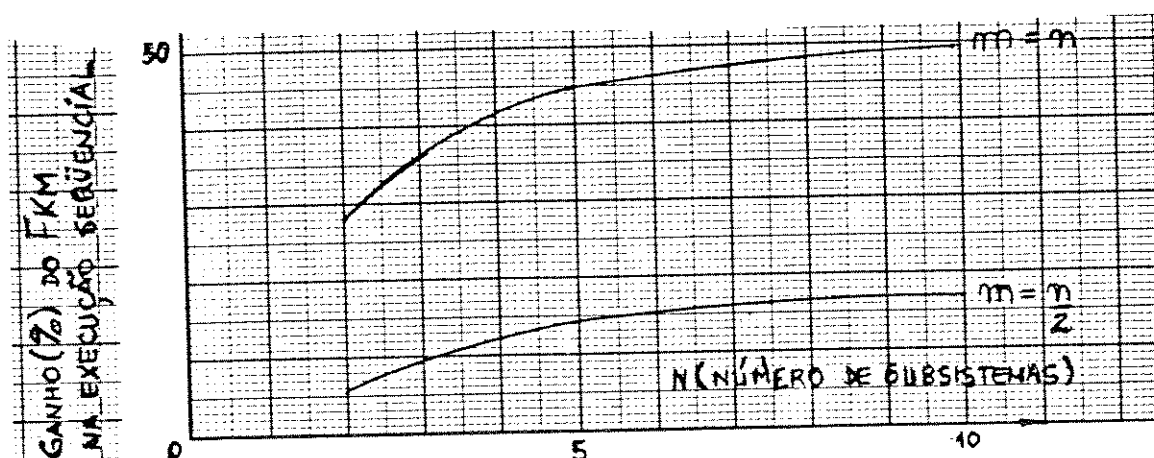


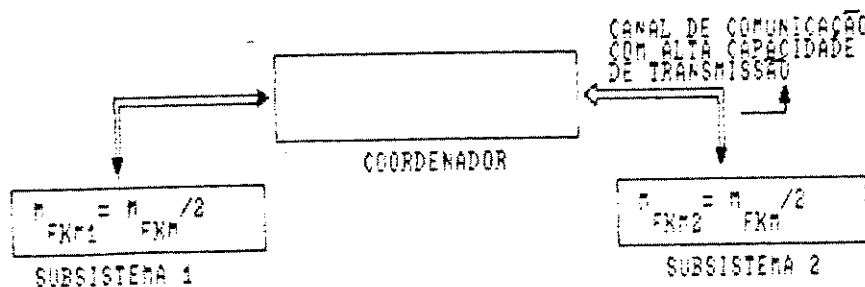
Figura 4.2-1 - Ganho no Processamento Seqüencial do FKM versus o Número de Subsistemas.

É possível notar que para $m=n$, o ganho com o aumento do número de subsistemas particionados, aumenta numa taxa maior do que o ganho para $m=n/2$ e que para valores de $m < n/2$, praticamente não teremos ganho algum, mesmo com aumento do número de subsistemas.

Entretanto, se utilizarmos um sistema multiprocessador, o tempo de cálculo é dividido por aproximadamente N (n^0 de subsistemas).

Para se ter uma noção do ganho que pode-se obter num sistema multiprocessador, façamos a seguinte aproximação:

O cálculo da equação 4.2-1 concernente ao subsistema é alocado somente ao processador relativo a esse subsistema e supondo os tempos de comunicação e espera pouco significativos em relação ao tempo de execução, ou seja:



Dessa forma, teríamos uma redução pela metade do número de multiplicações, que implicaria num ganho do sistema multiprocessador da ordem de $G(\%) = (2700 - 1276,5) / 2700 \approx 52,7\%$ em relação ao FKG e uma eficiência de $E(3) = (52,7 / 3) \cdot 100 = 17,6\%$ o que podemos considerar baixa.

Como uma extensão dessa aproximação, é possível constatar que para sistemas onde $m \geq n$, este ganho aumentaria significativamente.

Vemos, portanto, que o tempo de execução de um subproblema é uma função não linear da dimensão do respectivo subsistema (já que $n_i = n_j$, $\forall i, j$) podemos escrever $t_i = f(n_i)$, $i = 1, 2, \dots, N$ e então considerando as fases de multiprogramação e processamento paralelo, podemos originar os seguintes itens:

a) Implementação sequencial: O tempo t_s gasto no cálculo de todos os N subproblemas numa iteração é:

$$t_s = \sum_{i=1}^N f(n_i)$$

b) Multiprogramação: O tempo gasto para resolver os N subproblemas numa iteração é:

$$t_M(N) = \max [f(n_1), \dots, f(n_N)]$$

c) Processamento Paralelo: O tempo gasto para calcular os N subproblemas numa iteração é:

$$t_{PP}(N) = \max [f(n_1), f(n_2), \dots, f(n_N)]$$

O tempo total de execução para cada item mencionado é:

$$T_M = \sum_{i=1}^N f(n_i) + t_c$$

onde: $t = g(n_i)$, g uma função não linear e t_c é o tempo gasto para execução do programa coordenador.

$$T_{PP}(N) = T_M(N) = \max [f(n_1), \dots, f(n_N)] + t_c$$

Já que todos os subsistemas têm o mesmo número de estados, ou seja, $n_1 = n_2 = \dots = n_N = n$, podemos reescrever T_{PP} e T_M como:

$$\begin{aligned} T(N) &= N f(n) + t_c \\ T_{PP}(N) &= T_M(N) = f(n) + t_c \end{aligned}$$

Verificando o ganho do algoritmo paralelo, teremos:

$$g(N) = \frac{T(N)}{T_{PP}(N)} \quad \therefore \quad g(N) \simeq \frac{t_{SUB} + t_c}{\frac{t_{SUB}}{N} + t_c}, \quad \text{onde } t_{SUB} = N f(n)$$

ou ainda,

$$g(N) \approx \frac{N}{1 + (N-1)\beta} \quad , \text{ onde } \beta = \frac{t_c}{t_{sub} + t_c} \quad e$$

g é o ganho no tempo de processamento.

Portanto, é possível verificarmos que uma redução no tempo de execução do programa coordenador implicaria numa elevação do desempenho da execução paralela.

4.3. DADOS COMUNICADOS

Neste ítem nos preocupamos em enfocar a quantidade de dados comunicados em cada uma das etapas do item 3.2. para o caso geral (N subsistemas), obtendo-se expressões que permitam o cálculo aproximado do tempo gasto na comunicação entre os subproblemas no mesmo nível e entre os níveis na estrutura hierárquica. Considerando os subsistemas com os mesmos números de estados e saídas, podemos escrever:

A. QUANTIDADE DE DADOS COMUNICADOS NA ETAPA 1:

$$D(N) = B \cdot N \cdot (N+1) \cdot n^2$$

onde:

B = número de bytes necessários para armazenar uma variável;

N = número de subsistemas;

$n = n_i$ para $i = 1, 2, \dots, N$.

$m = m_i$ para $i = 1, 2, \dots, N$.

B. QUANTIDADE DE DADOS COMUNICADOS NA ETAPA 2:

$$D(N) = B \left[D_1(N) + D_2(N) \right]$$

onde:

$D_1(N)$ = qtde. de dados comunicados entre os níveis;

$D_2(N)$ = qtde. de dados comunicados entre os N subsistemas.

e tem os seguintes valores:

$$D_1(N) = B_N \left[(mn + m^2 + m) + (N-1) n \right]$$

$$D_2(N) = B \left[N^2 - 1 + (P(N,2)/2) \right] n^2$$

onde:

$P(N,2)$ = Permutação de N subsistemas dois a dois.

C. QUANTIDADE DE DADOS COMUNICADOS NA ETAPA 3:

$$D(N) = D_1(N) = B_N (nm + n^2)$$

D. QUANTIDADE DE DADOS COMUNICADOS NA ETAPA 4:

$$D_1(N) = B_N \left[(N-1) (2m^2 + n^2) \right]$$

$$D_2(N) = B \left[2Nnm + N(N-1) (m + m^2) \right]$$

Na execução do cálculo das matrizes e vetores que constituem o terceiro (ortogonalizações sucessivas) termo da equação 2.5.3-1, o número de informações enviadas do coordenador para os N -subsistemas é de:

$$N (N-1) (m^2 + n^2) \text{ bytes.}$$

E. QUANTIDADE DE DADOS COMUNICADOS NA ETAPA 5:

A estimação finalmente é concluída na etapa 5, com a requisição do seguinte número de informações entre os níveis:

$$D(N) = D_1(N) = B 2N \ln n$$

Nessa etapa o envio de mensagens é somente dos N subsistemas para o coordenador.

Portanto, como podemos ver, o número de informações em bytes que são transmitidas tanto entre os níveis como entre os subsistemas no mesmo nível, aumenta não linearmente com o número de subsistemas particionados, concluindo dessa forma que o número de particionamento em subsistemas pode se tornar proibitivo por exigir um elevado tempo de comunicação entre os módulos da estrutura hierárquica.

Numa implementação do algoritmo via PP (processador preferencial) [16], onde a comunicação entre os subproblemas é feita através de uma memória global por meio de um único barramento, o tempo de comunicação para um N elevado pode se tornar significativamente maior quando comparado ao tempo de execução dos subproblemas.

Além disso, o esforço de programação na elaboração do mecanismo de sincronização (o que aumenta o tempo de espera para acesso à memória global) pode se tornar exagerado.

A implementação desse algoritmo numa arquitetura paralela que permita uma configuração com ligações (canais de comunicação) entre os N processadores, pode contornar esse problema.

4.4. APLICAÇÃO E RESULTADOS NUMÉRICOS

Realizamos a solução paralelizada do filtro de Kalman para o seguinte modelo de quarta ordem [45] discretizado:

$$\begin{bmatrix} X(k+1) \\ X^1(k+1) \\ X^2(k+1) \\ X^3(k+1) \\ X^4(k+1) \end{bmatrix} = \begin{bmatrix} 0,18 & 0,00 & 0,00 & 0,00 \\ -0,25 & 0,27 & 0,00 & 0,00 \\ 0,55 & 0,00 & 0,18 & 0,00 \\ 0,00 & 0,55 & -0,25 & 0,27 \end{bmatrix} \begin{bmatrix} X(k) \\ X^1(k) \\ X^2(k) \\ X^3(k) \\ X^4(k) \end{bmatrix} + \begin{bmatrix} 4,50 \\ 6,15 \\ 2,00 \\ 2,65 \end{bmatrix} + \begin{bmatrix} W(k) \\ W^1(k) \\ W^2(k) \\ W^3(k) \\ W^4(k) \end{bmatrix}$$

$$\begin{bmatrix} Y(k+1) \\ Y^1(k+1) \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X(k+1) \\ X^1(k+1) \\ X^2(k+1) \\ X^3(k+1) \\ X^4(k+1) \end{bmatrix} + \begin{bmatrix} V(k+1) \\ V^1(k+1) \end{bmatrix}$$

$$\underline{X}(0) = [5 \ 5 \ 5 \ 5]^T$$

$$\hat{\underline{X}}(0) = [0 \ 0 \ 0 \ 0]^T$$

Os vetores de ruído $\underline{W}(k)$, $k = 0,1,\dots,N$ e $\underline{V}(k)$, $k = 0,1,\dots,N$, bem como as condições iniciais do sistema $\underline{X}(0)$, são supostos ruídos brancos Gaussianos com média nula e não correlatos entre si. A matriz de covariância Q do ruído do sistema é $5I_4$ onde I_4 é a matriz identidade de quarta ordem. A matriz de covariância R do ruído na observação é I_2 onde I_2 é a matriz identidade de segunda ordem e a matriz de covariância do estado inicial $\underline{X}(0)$ é $P(0) = 25 I_4$. A Tabela 4.4-1 mostra os resultados para análise da computação paralela do filtro de Kalman nos cinco primeiros instantes de amostragem.

As Figuras 4.4-1 e 4.4-2 mostram o comportamento das saídas do sistema e as Figuras 4.4-3 à 4.4-6 apresentam as estimações das variáveis de estado do sistema. São plotados os valores reais bem como os valores estimados.

PROGRAMAÇÃO PARALELA	Nº DE LINHAS	PROG. EXECUTÁVEL
PROGRAMA SUBSIST 1	285	33580 BYTES
PROGRAMA SUBSIST 2	285	33560 BYTES
PROGRAMA COORD	268	30420 BYTES

TAB. I		PROG. EXECUTÁVEL
PROGRAMAÇÃO SEQUENCIAL		21544 BYTES
Nº DE LINHAS : 424	T ₁	T ₂
TEMPO RESOLUÇÃO	06	10

T₁ - TEMPO GASTO EXECUTANDO O PROGRAMA
T₂ - TEMPO GASTO PELO SISTEMA EM BENEFÍCIO DO PROGRAMA
OBS: TEMPO EM SEGUNDOS COM PRECISÃO DE 1/10 DE SEGUNDOS

TEMPO	ITERAÇÃO					
	1	2	3	4	5	
T ₁	00	00	00	00	00	SUBSIST 1.1
	00	00	00	00	00	SUBSIST 1.2
	00	00	01	00	00	SUBSIST 1.3
	00	00	00	01	00	SUBSIST 1.4
	00	00	00	00	00	SUBSIST 2.1
	00	00	00	00	00	SUBSIST 2.2
	01	00	01	01	00	SUBSIST 2.3
	01	00	00	00	00	SUBSIST 2.4
	01	00	00	00	00	COORD 1
	00	00	00	00	01	COORD 2
	00	00	00	00	01	COORD 3
	01	00	00	00	00	COORD 4
T ₂	01	01	00	00	01	SUBSIST 1.1
	01	01	02	01	01	SUBSIST 1.2
	04	03	01	03	02	SUBSIST 1.3
	01	01	01	02	01	SUBSIST 1.4
	02	02	04	03	01	SUBSIST 2.1
	06	04	02	04	05	SUBSIST 2.2
	07	07	05	06	02	SUBSIST 2.3
	04	05	04	05	04	SUBSIST 2.4
	01	01	01	01	01	COORD 1
	02	01	02	01	00	COORD 2
	02	02	04	01	00	COORD 3
	00	02	00	01	01	COORD 4

Tabela 4.4-1

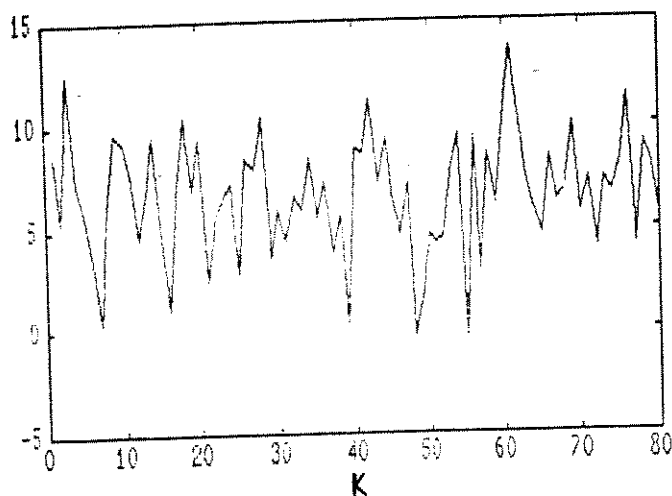


Figura 4.4-1
Saída Y₁

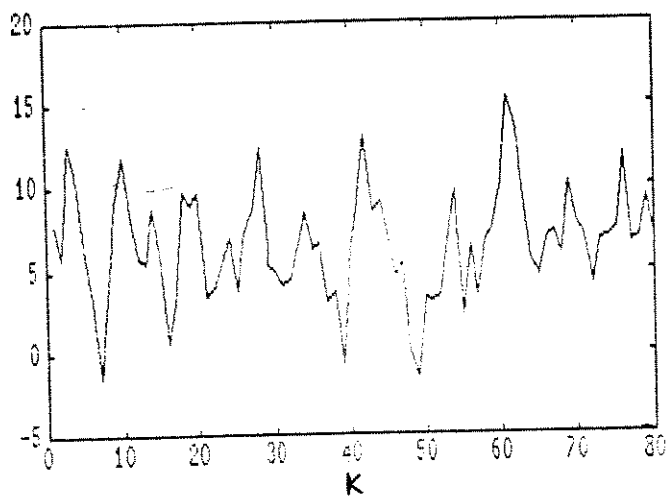


Figura 4.4-2
Saída Y₂

Na Figura 4.4-3 podemos observar a pobreza na estimação do estado X_1 , o mesmo ocorrendo com a estimação do estado X_2 , devido à não observabilidade dos mesmos. Porém, a condição de estimação não polarizada é satisfeita para todos os estados do sistema.

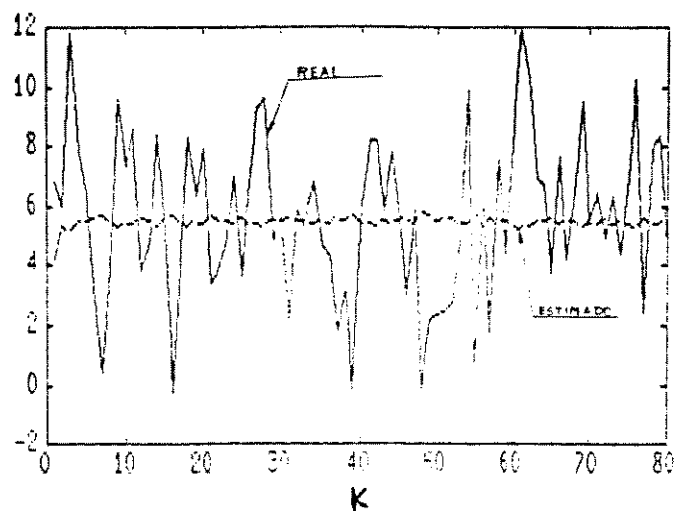


Figura 4.4-3
Variável de estado X_1

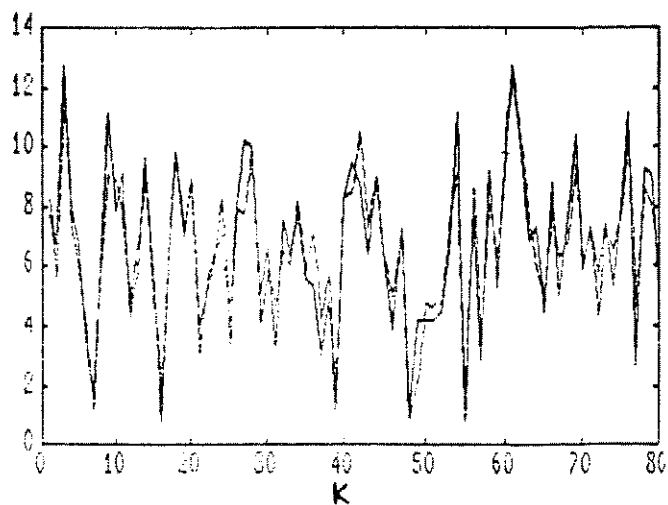


Figura 4.4-4
Variável de estado X_2

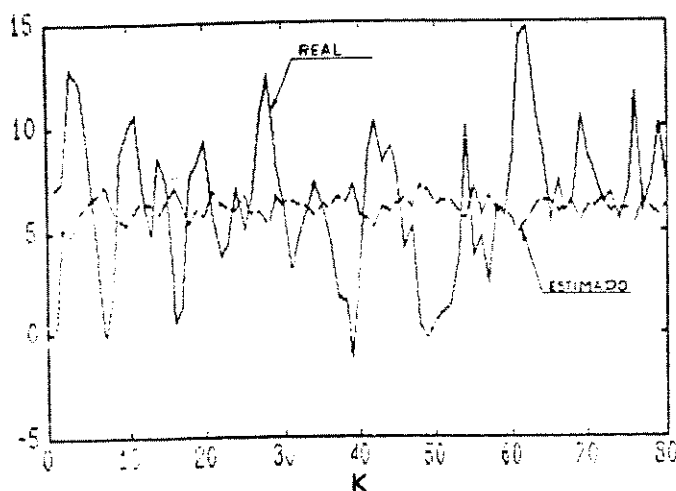


Figura 4.4-5
Variável de estado X_3

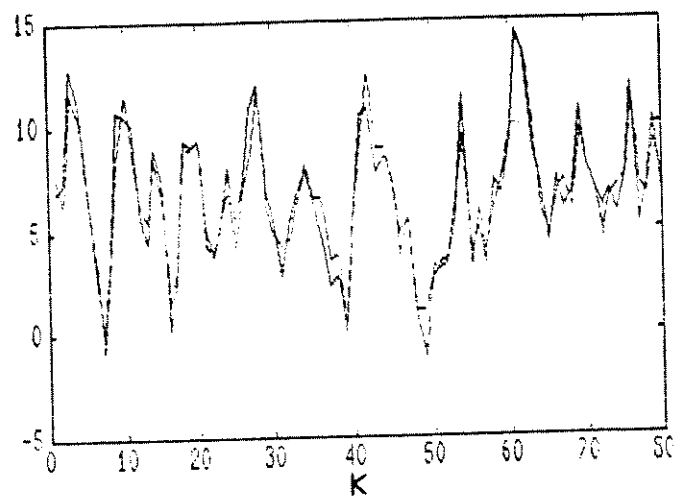


Figura 4.4-6
Variável de estado X_4

CAPÍTULO 5

CONCLUSÕES

Dos resultados obtidos neste trabalho ressaltamos, além de aspectos já comentados nos diversos capítulos, e além da análise e da síntese sobre metodologia de hierarquização do Filtro de Kalman e da realização da sua computação paralela, os seguintes aspectos relativos a esta:

- 1) A implementação sequencial exige mais memória para os cálculos dos níveis em um único computador, enquanto o armazenamento dos programas compilados no ambiente de multiprogramação exige maior quantidade ainda de memória.
- 2) O tempo gasto na resolução sequencial dos subproblemas foi maior do que o tempo gasto no subproblema mais lento.
- 3) No ambiente de multiprocessamento, pode-se adicionar novos subsistemas, aumentando a capacidade do sistema, sendo possível que o tempo de processamento do nível inferior se mantenha o mesmo, o que não ocorrerá no processamento sequencial, já que a relação tempo de processamento / requisição de memória pode tornar-se proibitiva.
- 4) A minimização do tempo de comunicação assume uma importância significativa em relação ao tempo do processamento total de forma que o estudo de estratégias para redução da quantidade de dados comunicados entre os processadores diante da arquitetura paralela utilizada (configuração dos processadores e capacidade dos canais de comunicação entre eles), pode conduzir a bons resultados.

5)O procedimento de paralelização aqui proposto e implementado, utilizando o ambiente Unix, mostrou-se viável, útil e viabilizador de novas indagações e de novos procedimentos certamente importantes para a implementação em tempo real em sistemas de múltiplos processadores do filtro de Kalman paralelizado, um instrumento de importância fundamental para o estudo, estimação e controle de sistemas de grande porte.

REFERÊNCIAS BIBLIOGRÁFICAS

- [1] Kailath, T., "A View of Three Decades of Linear Filtering Theory", IEEE Trans. Infor. Theory, Vol. IT-20, No.2, pp.146-181, Março/1974.
- [2] Kalman, R.E., "A New Approach to Linear Filtering and Prediction Problems", Trans. ASME, J. Basic Engng., vol. 82, pp. 34-45, Março/1960.
- [3] Doob, J.L., "Stochastic Processes", John Wiley & Sons, Inc., New York, N.Y., 1955.
- [4] Kalman, R.E. & Bucy, R.S., "New Results in Linear Filtering and Prediction Theory", Trans. A.S.M.E., J. Basic Engng. Sc. D., vol. 83, pp. 95-107, Dezembro/1961.
- [5] Lewis, F.L., "Optimal Estimation: With an Introduction to Stochastic Control Theory", John Wiley & Sons, N.Y., 1986.
- [6] Luenberger, D.G., "Optimization by Vector Space Methods", Wiley, New York, 1969.
- [7] Franklin, F.G. & Powell D.J., "Digital Control of Dynamic Systems", Addison-Wesley Publishing, Inc., 1980.
- [8] Sage, A.P. & Melsa, J.L., "Estimation Theory With Application to Communication and Control", McGraw Hill, 1971
- [9] Sage, A.P., "Optimum System Control", Prentice-Hall, Englewood Cliffs, N.J.: Prentice-Hall, 1979
- [10] Amaral, Wagner C., "Identificação Recursiva de Sistemas Multivariáveis", Tese de Doutorado, Campinas, UNICAMP/FEE, 1980.

- [11]Cechim J., "Estudo Comparativo de Métodos de Filtragem para Sistemas Estocásticos Discretos Não Lineares", Tese de Mestrado, Campinas, UNICAMP/FEE, 1977.
- [12]B.D.O. Anderson & J.B. Moore, "Optimal Filtering", Englewood Cliffs, N.J.: Prentice-Hall, 1979.
- [13]Rhodes, Ian B., "A Tutorial Introduction to Estimation and Filtering", IEEE Transactions on Automatic Control, Vol. 16, No.6, December, 1971.
- [14]Jazwinski, A.H., "Stochastic Processes and Filtering Theory", Academic Press, 1970.
- [15]Sage, A.P. & Melsa, J.L., "System Identification", Academic Press, N.Y., 1971.
- [16]Costa Filho, J.T., "Processamento Paralelo de Algoritmos de Controle Hierárquico". Tese de Mestrado, Campinas, UNICAMP/FEE, 1988.
- [17]Shah, M.M., "Supplemented Partitioning Approach to Linear Filtering Theory", Proc. IFAC Symposium, Udine, Italy, June, 16-2-, 1976, pp. 379-400.
- [18]Singh, M.G., "Multi-level State Estimation", Int. J. Systems Sci., vol. 6, No. 6, pp. 533-555, 1975.
- [19]Noton, A.R.M., "A Two-level Form of the Kalman Filter", IEEE Trans. Automatic Control, Vol. 16, No 2, pp. 128-132, 1971.
- [20]Arafteh, S. & Sage A.P., "Multilevel Discrete Time Identification on Large Scale Systems", Int. J. Syst. Sci., vol. 5, 1974.

- [21] Sage, A.P. & Smith, N.J., "A Sequential Method for System Identification in Hierarchical Structure", *Automatica*, vol. 9, pp. 677-688, 1973.
- [22] Sage, A.P. & Guinzy, N.J., "System Identification in Large Scale Systems with Hierarchical Structures", *Comput. & Elect. Engng.*, vol. 1, pp. 23-42, 1973.
- [23] Roberts, P.D., "Multilevel Approaches to the Combined Problem of System Optimization and Parameter Identification", *Int. J. Systems Sci.*, vol. 8, No. 3, pp. 273-299, 1977.
- [24] Sage, A.P. & Arafeh S.A., "Hierarchical System Identification of States and Parameters in Interconnected Power Systems", *Int. J. Systems Sci.*, vol. 5, No. 9, 817-846, 1974.
- [25] Hassan, M.F., "New Multi-level Algorithm for Linear Dynamical Systems", *Int. J. Systems Sci.*, vol. 19, No. 12, pp. 2631-2642, 1988.
- [26] Singh, M.G. & Titli, A., "Systems: Decomposition, Optimization and Control", Pergamon Press, 1978.
- [27] Hassan, M.F., G. Sautl., M.G. Singh, A. Titli, "A Decentralized Computational Algorithm for the Global Kalman Filter", *IEEE Trans. Aut. Control*, vol. 23, 1978.
- [28] Hassan, M.F. "Optimisation et Commande Hierarchisées des Systèmes Dynamiques Interconnectés", Tese de Doutorado, Université Paul Sabatier, Toulouse/França, 1978.
- [29] Bottura, C.P., Costa Filho, J.T., Quirino, R.B., "Programação Paralela do Filtro de Kalman Multinível", 9^o Congresso Brasileiro de Automática, Setembro/1990.

- [38] Desai, V.B. & Das B., "Parallel Algorithms for Kalman Filtering", in Proc. Amer. Contr. Conf., Boston, MA, June 1985, pp. 920-921.
- [39] Meyer G.G. and Weinert W.H., "Parallel Algorithms and Computational Structures for Linear Estimation Problems", in Statistical Signal Processing, E.J. Wegman and J.G. Smith, Eds. N.Y.: Marcel-Dekker, 1984, pp. 507-516
- 40] Hashemipour H.R., Roy S., Laub A.J., "Decentralized Structures for Parallel Kalman Filtering", IEEE Trans. on Automatic Control, vol. 33, No. 1, January 1988.
- 41] Mendel, J.M., "Computational Requirements for a Discrete Kalman Filter", IEEE Trans. on Automatic Control, vol. 16, No. 6, December, 1971.
- 42] Lainiotis, D.G. & Anagnostou, K.E., "Comparative Computational Analysis of a New Per-Sample Partitioning Linear Filter and Kalman Filter", Int. J. Systems Sci., vol. 18, No. 2, pp. 351-370, 1987.
- 43] Chin L., "Advances in Computational Efficiencies of Linear Filtering", Leondes, C.T. (Ed.), Control and Dynamic Systems, Academic Press, vol. 19, pp. 125-188, 1983.
- 44] Gomide, F.A.C., "Análise e Implementação de Algoritmos de Controle Hierárquico de Sistemas Dinâmicos", Tese de Mestrado, UNICAMP/FEE, Campinas, 1979.
- 45] Tamura, H., "A Discrete Dynamic Model with Distributed Transport Delays and its Hierarchical Optimization for Preserving Stream Quality", IEEE Trans. on Systems, Man and Cybernetics, vol. SMC 4, nº 5, 1974.
- 46] Manual Digirede 8000; S.O Digix, v.1.4a-5.