4 A Teoria de Filtragem

Neste capítulo será abordado o conhecimento necessário para a implementação do filtro utilizado neste trabalho, conforme [27]. O filtro de Kalman é formulado matematicamente em termos de variáveis de estado e sua solução é computada recursivamente, ou seja, cada estado estimado atualizado é computado a partir de valores anteriores estimados e dos novos dados, porem só sendo necessário armazenar o valor estimado anterior. Sua implementação é simples, o que o levou a ser utilizado em muitos problemas em diferentes áreas de conhecimento. De fato, o filtro de Kalman é um exemplo da família de filtros adaptativos conhecidos como *"Filtros Recursivos de Mínimos Quadrados"*, para o caso estocástico. Mas o filtro de Kalman precisa do conhecimento dos parâmetros da planta e dos valores das matrizes de covariância do ruído de sistema e de medida. Existem alguns métodos utilizados para a implementação de algoritmos que permitam fornecer os valores ótimos destas matrizes de ruído, que serão tratados adiante.

4.1. Conceitos Probabilísticos

Considere um sistema discreto estocástico dinâmico:

$$x_k = \Phi(x_{k-1}, k-1) + \Gamma(x_{k-1}, k-1) w1_{k-1}, \qquad k = 1,2,3....... \qquad (57)$$

onde:

x_k = Um n-vetor de estado *"verdadeiro"* do sistema no tempo *k,*

Φ = Uma função n-vetorial,

Γ = Uma função matriz *n x r*,

$w1_{k-1}$ = Ruído branco e gaussiano do sistema.

Da equação (57) pode-se ver que x_k é um vetor aleatório, visto que w1 é aleatório também. A saída do sistema pode ser escrita:

$$y_k = H(x_k, k) + w2_k \qquad k = 1,2,3..... \qquad (58)$$

onde:

y_k = Vetor *p* de observações no tempo *k*

H = Vetor p,

w2 = ruído branco e gaussiano de medição.

7

A equação (58) mostra como as observações do estado verdadeiro estão sujeitas aos erros de medida. O registro de valores passados de saída será representado por Y_ℓ :

$$Y_\ell = \{y_1, y_2, y_3, \ldots\ldots y_\ell\} \tag{59}$$

O problema da filtragem discreta é computar uma estimativa de x_ℓ, linear e não viciada, dadas as realizações de Y_ℓ ; neste caso minimizando a variância do erro de estimação. A solução do problema de filtragem assim definido pode ser estabelecida em termos da função de densidade probabilística condicional de x_l dado Y_ℓ ,

$$f((x, \ell)|Y_\ell) \tag{60}$$

se ela contem toda a informação estatística sobre x . No caso linear gaussiano, a solução está caracterizada pela média e a covariância da densidade condicional (60). No caso não linear a solução é muito mais complexa, porque a densidade não pode ser caracterizada por um conjunto finito de parâmetros (não sendo, em geral, gaussiana) [22].

Se a densidade (60) é unimodal (um só pico) e simétrica com respeito a sua média, então, para a maioria dos sistemas (Sherman 1955), o estimador ótimo é simplesmente a média condicional:

$$\hat{x}_\ell = E[x_\ell | Y_\ell] \tag{61}$$

4.2. O Filtro de Kalman

O filtro de Kalman (a tempo discreto) fornece um algoritmo para computar a estimativa ótima e o erro de covariância para um sistema linear dinâmico discreto estocástico excitado por ruídos gaussianos. Seja um sistema linear descrito pela seguinte equação:

$$x_k = \Phi_{k-1} x_{k-1} + \Gamma_{k-1} w1_{k-1} \tag{62}$$

onde:

x_k = vetor *n* de estado no tempo *k,*

Φ = Matriz *n x n* de transição de estado não singular,

w1 = Vetor r de seqüência de ruído branco gaussiano.

O vetor x_k representa as variáveis de estado do sistema dinâmico.

O ruído ser branco gaussiano significa que seus valores ao longo do tempo são variáveis aleatórias gaussianas descorrelatadas e de média nula, sendo completamente representados com sua covariância.

O ruído w1 é suposto branco gaussiano, o que significa:

$$E[w1_k \, w1_\ell] = \begin{cases} Q, & \ell = k \\ 0, & \ell \neq k \end{cases} \tag{63}$$

Para o sistema de observação, de novo utilizamos:

$$y_k = H_k x_k + w2_k \tag{64}$$

com as características do ruído do observação (suposto branco gaussiano) dadas por:

$$E[w2_k \, w2_\ell] = \begin{cases} R, & \ell = k \\ 0, & \ell \neq k \end{cases} \tag{65}$$

Além disso, os dois ruídos são supostos independentes, o que implica que sua covariância é nula.

Suponha conhecidas as observações da saída:

$$Y_k = \{y_1, y_2, y_3, ...y_k\} \tag{66}$$

Define-se o valor estimado $\hat{x}_{k|k-1}$ como o valor estimado de x_k baseado em todas as observações até *k-1*, isto é, $E[x_k|Y_{k-1}]$; e $\hat{x}_{k|k}$ como o valor estimado de x_k baseado em todas as observações incluindo o tempo *k*, isto é, $E[x_k|Y_k]$. Por simplicidade, $\hat{x}_{k|k}$ será denotado apenas como $\hat{x}_k$ e $\hat{x}_{k|k-1}$ como $\hat{x}_{k,k-1}$. Definem-se as matrizes de covariância dos erros de estimação por:

$$P_{k|k-1} = E\left[(x_k - \hat{x}_{k|k-1})(x_k - \hat{x}_{k|k-1})^T\right] \tag{67}$$

$$P_{k|k} = E\left[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T\right] \equiv P_k \tag{68}$$

4.2.1. A Inovação

A inovação é o processo estocástico definido pelo resíduo entre os valores observados ou medidos e os valores estimados de um vetor de estado. O desempenho do filtro é refletido nas propriedades estatísticas da seqüência de inovação. A inovação vem definida da seguinte forma:

$$v_k = y_k - \hat{y}_{k|k-1} \tag{69}$$

onde:

$$\hat{y}_{k|k-1} = H_k \hat{x}_{k|k-1} \tag{70}$$

sendo, no caso do funcionamento ótimo, um processo gaussiano branco de média nula. Pode ser vista como a saída de um filtro de predição de ordem *k-1,* onde sua entrada é uma série de $y_1, y_2, \ldots\ldots\ldots\ldots, y_k$.

4.2.1.1. Propriedades da Inovação

A inovação tem propriedades importantes, que são descritas a seguir:

Propriedade 1: *A inovação α_k associado com a variável observável y_k é ortogonal às observações passadas* $y_1, y_2, \ldots\ldots\ldots\ldots, y_k$ *:*

$$E\left[v_k y_k^T\right] = 0 \tag{71}$$

Propriedade 2: *As inovações v_1, v_2,........ v_k, são ortogonais entre si:*

$$E\left[v_k v_k^T\right] = 0 \tag{72}$$

Propriedade 3: *Existe uma correspondência uma a uma entre os dados observados* $\{y_1, y_2, \ldots\ldots\ldots, y_k\}$ *e as inovações* $\{v_1, v_2, \ldots\ldots\ldots, v_k\}$ *onde, uma seqüência pode ser obtida de outra por meio de um filtro inversível causal sem perda de informação*

$$\{y_1, y_2, \ldots\ldots\ldots, y_k\} \Leftrightarrow \{v_1, v_2, \ldots\ldots\ldots, v_k\} \tag{73}$$

4.2.2. Algoritmo do Filtro de Kalman

O algoritmo desenvolvido por Kalman está baseado na solução recursiva do problema de estimação quadrática mínima, utilizando a definição do processo de inovação, aproveitando as propriedades de correlação deste processo estocástico especial. A seguir será apresentado o algoritmo:

- Vetor de entrada (medida ou observada):

$Y_k = \{y_1, y_2, \ldots\ldots\ldots y_k\}$

- Parâmetros conhecidos:

Φ: matriz de transição de estado

H: matriz de saída

Q: matriz de covariância do ruído de sistema

R: matriz de covariância do ruído de medida

Estimação de estado:

$$\hat{x}_k = \hat{x}_{k,k-1} + K_k^{KB}\left[y_k - H_k \hat{x}_{k,k-1}\right] \tag{74}$$

onde K_k^{KB} é o ganho de Kalman utilizado para corrigir o valor previsto a partir do valor da nova medida y_k.

Atualização da matriz de covariância do erro de estimação:

$$P_k = P_{k,k-1} - K_k^{KB} H_k P_{k,k-1} \tag{75}$$

Computação do ganho do filtro de Kalman:

$$K_k^{KB} = P_{k,k-1} H_k^T \left[H_k P_{k,k-1} + R_k\right]^{-1} \tag{76}$$

Predição de estado:

$$\hat{x}_{k,k-1} = \Phi_k \hat{x}_k \tag{77}$$

Estimação da Matriz de Covariância do erro:

$$P_{k,k-1} = \Phi_{k-1} P_k \Phi_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \tag{78}$$

4.3. Filtro Estendido de Kalman

O filtro de Kalman pode ser estendido ao problema não linear através de um procedimento de linearização, conhecido como filtro estendido de Kalman. O algoritmo tem similitude com o filtro de Kalman, mas com algumas modificações. Corresponde ao filtro de Kalman aplicado ao sistema linearizado em torno do último valor estimado, sendo, depois, o sistema linearizado em torno deste novo valor, para nova aplicação do filtro, e assim recursivamente (Ver Anexo A).

Estimação de estado:

$$\hat{x}_k = \hat{x}_{k,k-1} + K_k^{KB}\left[y_k - H_k \hat{x}_{k.k-1}\right] \tag{79}$$

Atualização da Matriz de covariância do erro de estimação:

$$P_k = P_{k,k-1} - K_k^{KB} H1_k P_{k,k-1} \tag{80}$$

$$H1_k = \left.\frac{\partial H}{\partial x}\right|_{x = x_{k-1}} \tag{81}$$

Computação do ganho do filtro de Kalman:

$$K_k^{KB} = P_{k,k-1} H1_k^T \left[H1_k P_{k,k-1} + R_k\right]^{-1} \tag{82}$$

Predição de estado:

$$\hat{x}_{k,k-1} = \Phi_k\left(k, x_{k-1}, u_k\right) \tag{83}$$

Estimação da Matriz de Covariância do erro de estimação:

$$P_{k,k-1} = \Phi 1_{k-1} P_k \Phi_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \tag{84}$$

$$\Phi 1_k = \left.\frac{\partial \Phi}{\partial x}\right|_{x = x_{k-1}} \tag{85}$$

4.4.Filtro Sub-ótimo

Para a formulação do filtro de Kalman a partir de um sistema dinâmico estocástico (62), (64) é necessário conhecer as matrizes Ψ, Γ, H e as matrizes de covariância dos ruídos de observação e do sistema Q e R. Se um filtro está baseada no conhecimento incompleto de alguma destas matrizes, será dito um filtro sub-ótimo.

Neste trabalho será considerado o caso no qual o filtro sub-ótimo está projetado conforme as equações do filtro de Kalman, mas baseado nas estatísticas de ruído erradas Q^e e R^e. Assim, o algoritmo ficará:

Estimação de estado:

$$\hat{x}_k = \hat{x}_{k,k-1} + K_k^{KB}\left[y_k - H_k \hat{x}_{k,k-1}\right] \quad (86)$$

Atualização da matriz de covariância do erro de estimação, que, para o filtro sub-ótimo, será denotada por S:

$$S_k = S_{k,k-1} - K_k^{KB} H_k S_{k,k-1} \quad (87)$$

Computação do ganho do filtro:

$$K_k^{KB} = S_{k,k-1} H_k^T \left[H_k S_{k,k-1} + R_k^e\right]^{-1} \quad (88)$$

Predição de estado:

$$x_k^f = \Phi_k \hat{x}_k \quad (89)$$

Estimação da matriz de erro de covariância:

$$S_{k,k-1} = \Phi_{k-1} S_k \Phi_{k-1}^T + \Gamma_{k-1} Q_{k-1}^e \Gamma_{k-1}^T \quad (90)$$

É bom lembrar que será utilizado $S_{k,k-1}$, S_k para ressaltar que estas matrizes não podem ser interpretadas como as estimações dos erros de covariâncias verdadeiras $P_{k,k-1}$, P_k. O filtro (74) – (78) produzirá uma estimativa $\hat{x}_k$, mas que não será ótima. Porém este filtro é ótimo para o sistema estocástico com características de segundo ordem Q^e e R^e.

4.5. Filtragem Adaptativa

Existem varias técnicas disponíveis pelas quais um filtro pode ser provido de alguma característica que permita a correção de erros a partir de novas informações. Em nosso estudo será abordada a situação onde não se conhece corretamente os valores das matrizes de covariância de erro Q e R [23].

4.5.1. Estimação Bayesiana

Nesta abordagem, todos os parâmetros desconhecidos são representados por um vetor α. Em principio α pode incluir parâmetros relacionados ao sistema dinâmico assim como as matrizes de covariâncias do ruído Q e R.

Suponha que todos os possíveis valores de α podem ser representados por um número finito de realizações, possivelmente muito grande, de um processo $\{\alpha^i, i = 1,.....N\}$, onde a probabilidade inicial de cada realização é dada por $pr\{\alpha = \alpha^i\}$. Neste caso é possível derivar, via Teorema de Bayes, um conjunto de recursões das matrizes de covariância da densidade de probabilidade condicional de α^i, assim como dos estimadores $\hat{x}^i = E\left[x \middle| \alpha = \alpha^i\right], i = 1,.......N$. Os estimadores x^i serão ponderados pela densidade de probabilidade condicional para produzir uma estimativa final de x [27].

Esta abordagem é atrativa por sua generalidade, mas o procedimento é computacionalmente pesado para grandes sistemas, porque contém N filtros recorrentes.

4.5.2. Estimação de Máxima Verossimilhança

Métodos deste tipo são obtidos pela *maximização da densidade de probabilidade condicional* de um vetor desconhecido α, dado um conjunto de observações Y Na forma mais geral, a abordagem de máxima verossimilhança encontra um conjunto de equações diferenciais não lineares a cada passo, o qual deve ser resolvido iterativamente. Somente quando o sistema é invariante no tempo e o filtro é suposto em estado permanente, pode-se aplicar o método [27].

4.5.3. Adequação da covariância

Este método é aplicado quando se desconhece Q e R e está baseado na idéia de que a covariância da inovação:

$$v_k = y_k - H_k \hat{x}_{k,k-1} \tag{91}$$

tem um valor teórico que está diretamente relacionado a Q e R. Assim, a atual covariância de α_k pode ser estimada a partir da saída do filtro. Esquemas deste tipo podem ser desenvolvidos para ajustar Q e R até que sejam consistentes com os valores observados. A convergência destes métodos apresenta um desempenho baixo porque os ajustes de Q e R são interdependentes e porque estes ajustes a cada passo estão baseados numa simples amostragem da inovação [27].

4.5.4. Métodos de Correlação

Estes métodos são aplicáveis aos filtros onde Q e R são desconhecidos. O método de correlação de saída está baseado na seqüência de auto-correlação observada $\{y_k\}$; o método de correlação de inovação está baseado na autocorrelação da seqüência de inovação $\{v_k\}$. Estes métodos são semelhantes ao caso anterior, mas estão baseados nas propriedades estatísticas de um segmento de tempo da seqüência $\{y_k\}$ ou $\{v_k\}$.

Variantes destes algoritmos têm sido produzidas para gerar um conjunto de equações que relacionam as funções de autocorrelação amostradas com os parâmetros desconhecidos. Estas equações poderiam ser resolvidas em forma algébrica para produzir os estimadores de Q e R ou, alternativamente, produzir diretamente um estimador da matriz de ganho ótima. Mas estes métodos são computacionalmente muito trabalhosos e tem a desvantagem de estarem baseados numa só função de autocorrelação amostrada.

O problema foi resolvido por Bélanger (1974), para o caso especial em que as matrizes Q e R são independentes no tempo e dependem linearmente de um conjunto de parâmetros desconhecidos. Ele mostrou que, para essa condição, a função de autocorrelação verdadeira da seqüência de inovação é também uma função linear deste conjunto de parâmetros.

O método de Bélanger é atrativo por duas condições:

- Ele toma em conta as incertezas associadas com a amostragem da autocorrelação da seqüência de inovação. Os parâmetros desconhecidos são estimados a partir de observações do ruído da função de autocorrelação. Este algoritmo pode ser formulado como um filtro secundário.

- Em muitas aplicações, é razoável supor que o projetista conhece alguma coisa da estrutura de Q e R. O método de Bélanger tira vantagem desta informação inicial para reduzir significativamente o custo computacional e garantir a convergência.

4.5.5. A seqüência de Inovação

O desempenho de um filtro está relacionado com a propriedade estatística da seqüência de inovação $\{v_k\}$:

$$v_k = y_k - H_k \hat{x}_{k,k-1} \tag{92}$$

Como foi dito anteriormente na lista de propriedades da inovação do filtro de Kalman, o fato de que esta seqüência seja um ruído branco significa que o filtro linear já retirou toda a informação linear dele. Ou melhor, que toda a informação linear (a obtida por estimadores lineares, obtidos por projeções no espaço gerado pelas medidas diretamente) já foi obtida. No caso contrário, se a inovação fosse correlacionada no tempo, ele ainda teria informação útil não utilizada pelo filtro. Os algoritmos adaptativos que utilizam esta característica para melhorar o desempenho do filtro são chamados de métodos de correlação da inovação [27].

A seqüência de inovação está relacionada com o erro de estimação $\tilde{e}_k$ da seguinte forma:

$$\tilde{e}_k = x_k - \hat{x}_{k,k-1}$$

$$\begin{aligned} v_k &= y_k - H_k \hat{x}_{k,k-1} \\ v_k &= H_k \tilde{e}_k + w2_k \end{aligned} \tag{93}$$

Podemos obter uma recursão do vetor de previsão da equação do filtro para uma matriz de ganho arbitrária:

$$\hat{x}_{k,k-1} = \psi_{k-1}\hat{x}_k$$

$$\hat{x}_k = \hat{x}_{k,k-1} + K_k\left[y_k - H_k\hat{x}_{k,k-1}\right] \tag{94}$$

$$\hat{x}_{k,0} = E[x_0]$$

$$P_{k,k-1} = \Phi_{k-1}P_k\Phi_{k-1}^T + \Gamma_{k-1}Q_{k-1}\Gamma_{k-1}^T$$

$$P_k = [I - k_kH_k]P_{k,k-1}[I - k_kH_k]^T + k_kR_kK_k^T \tag{95}$$

$$P_{k,0} = E\left[(x_0 - \hat{x}_{k,0})(x_0 - \hat{x}_{k,0})^T\right]$$

A segunda equação de (95) é muito importante porque fornece o erro de covariância do estimador $\hat{x}_k$ para uma matriz de ganho arbitrária. Por exemplo, se uma matriz de ganho K_k é computada como uma aproximação de K_k^{KB} por razões de eficiência computacional, então esta segunda equação pode ser usada para computar o erro de covariância P_k.

Neste caso, sem supor que K_k é a matriz do filtro ótimo, a recursão do vetor de previsão ficará:

$$\begin{aligned} \hat{x}_{k,k-1} &= \Phi_{k-1}\hat{x}_k \\ &= \Phi_{k-1}\left[\hat{x}_{k,k-1} + K_{k-1}H_{k-1}\tilde{e}_{k-1}\right] + \Phi_{k-1}K_{k-1}w2_{k-1} \end{aligned} \tag{96}$$

onde x_k representa não mais a estimativa ótima (dada pelo filtro de Kalman), mas uma estimativa sub-ótima, dada pelo novo filtro calculado a partir de uma matriz de ganho arbitrária.

A recursão para o erro de previsão passa a ser:

$$\begin{aligned} x_k &= \Phi_{k-1}x_{k-1} + \Gamma_{k-1}w1_{k-1} \\ \tilde{e}_k &= \Phi_{k-1}[I - K_{k-1}H_{k-1}]\tilde{e}_{k-1} \\ &\quad + \Gamma_{k-1}w1_{k-1} - \Psi_{k-1}K_{k-1}w2_{k-1} \end{aligned} \tag{97}$$

Tomando ℓ recursões atrás:

$$\begin{aligned}\tilde{e}_k &= \phi^f_{k,k-\ell}\tilde{e}_{k-\ell} \\ &\quad + \sum_{j=k-\ell}^{k-1}\phi^f_{k,j+1}\left[\Gamma_j w1_j - \Phi_j K_j w2_j\right]\end{aligned} \tag{98}$$

onde $\phi_{k1,k2}$ é a matriz de transição de estado do erro do sistema, dada por:

$$\begin{aligned}\phi^f_{k1,k2} &\equiv \Phi_{k1-1}\left[I - K_{k1-1}H_{k1-1}\right]\Phi_{k1-2}\left[I - K_{k1-2}H_{k1-2}\right]\ldots \\ &\qquad \times \Phi_{k2}\left[I - K_{k2}H_{k2}\right] \\ \phi^f_{k1,k1} &\equiv I\end{aligned} \tag{99}$$

Utilizando (93):

$$\begin{aligned}E\left[v_k v^T_{k-\ell}\right] = E&\left[\left(H_k\tilde{e}_k + w2_k\right)\left(H_{k-\ell}\tilde{e}_{k-\ell} + w2_{k-\ell}\right)^T\right] \\ = H_k E&\left[\tilde{e}_k\tilde{e}^T_{k-\ell}\right]H^T_{k-\ell} \\ &+ H_k E\left[\tilde{e}_k\left(w2_{k-\ell}\right)^T\right] \\ &+ E\left[w2_k\, x_{k-\ell}\right]H^T_{k-\ell} \\ &+ E\left[w2_k\left(w2_{k-\ell})\right)^T\right]\end{aligned} \tag{100}$$

O primeiro e o segundo termos da equação (100) podem ser resolvidos com a equação (98), junto com as estatísticas do ruído do sistema e da observação. O terceiro e quarto termo são nulos a partir das hipóteses sobre os ruídos.

Assim:

$$E\left[v_k v^T_k)\right] = H_k P_{k,k-1} H^T_k + R_k \tag{101}$$

$$\begin{aligned}E\left[v_k v^T_{k-\ell}\right] &= H_k \phi^f_{k,k-\ell} P_{k,k-\ell} H^T_{k-\ell} \\ &\quad - H_k \phi^f_{k,k-\ell+1}\Psi_{k-\ell} K_{k-\ell} R_{k-\ell}, \quad \ell > 0\end{aligned} \tag{102}$$

As equações (103) e (104) expressam a autocorrelação da seqüência de inovação do filtro sub-ótimo em termos das matrizes de covariância do erro associadas.

Para mostrar que o lado direito da equação (102) se anula quando o filtro é ótimo, reescreveremos:

$$E\left[v_k v_{k-\ell}^T\right] = H_k \phi_{k,k-\ell+1}^f \Phi_{k-\ell} \times \left\{ \left[I - K_{k-\ell} H_{k-\ell} \right] P_{k,k-\ell} H_{k-\ell}^T - K_{k-\ell} R_{k-\ell} \right\} \tag{103}$$

$$\begin{aligned} E\left[v_k v_{k-\ell}^T\right] &= H_k \phi_{k,k-\ell+1}^f \Phi_{k-\ell} \times \left\{ P_{k,k-\ell} H_{k-\ell} - K_{k-\ell} \left[H_{k-\ell} P_{k,k-\ell} H_{k-\ell}^T + R_{k-\ell} \right] \right\} \\ &= H_k \phi_{k,k-\ell+1}^f \Phi_{k-\ell} \times \left\{ K_{k-\ell}^{KB} - K_{k-\ell} \right\} \left[H_{k-\ell} P_{k,k-\ell} H_{k-\ell}^T + K_{k-\ell} \right] \end{aligned} \tag{104}$$

A equação (104) encontra imediatamente a propriedade de inovação do filtro KB:

$$E\left[v_k v_{k-\ell}^T\right] = 0 \text{ para } \ell > 0 \text{ e para todo } k \tag{105}$$

Para um filtro sub-ótimo, a autocorrelação da seqüência de inovação é diretamente proporcional à diferença entre a matriz de ganho K e a matriz de ganho de Kalman, pelas equações (103) e (104). Se K é o ganho do Filtro de Kalman, então a seqüência de inovação é branca e sua covariância é dada por (101). As inovações formam uma seqüência gaussiana se o ruído do sistema e o ruído de observação são gaussianos e o algoritmo do filtro é linear.

4.5.6. Algoritmo de Bélanger

O algoritmo de Bélanger supõe que Q e R são independentes no tempo e que são combinações lineares de matrizes conhecidas Q_i e Q_i:

$$Q = \sum_{i=1}^{k} \alpha_i Q_i \tag{106}$$

$$R = \sum_{i=1}^{k} \alpha_i R_i \tag{107}$$

Será projetado um filtro sub-ótimo na base de um *estimativa inicial* α^e do vetor $\alpha = [\alpha_1,...,\alpha_k]^T$, fornecendo estimativas iniciais para as matrizes Q e R:

$$Q^e = \sum_{i=1}^{N} \alpha_i^e Q_i \tag{108}$$

$$R^e = \sum_{i=1}^{N} \alpha_i^e R_i \tag{109}$$

Agora o algoritmo do filtro de Kalman será reescrito incluindo a seqüência de inovação:

Estimação de estado:

$$\hat{x}_k = \hat{x}_{k,k-1} + K_k^{KB}\left[y_k - H_k \hat{x}_{k,k-1}\right] \tag{110}$$

Atualização da matriz de covariância de erro de estimação:

$$P_k = P_{k,k-1} - K_k^{KB} H_k P_{k,k-1} \tag{111}$$

Computação do ganho do filtro de Kalman:

$$K_k^{KB} = P_{k,k-1} H_k^T \left[H_k P_{k,k-1} + R_k\right]^{-1} \tag{112}$$

Predição de estado:

$$\hat{x}_{k,k-1} = \Phi_k \hat{x}_k \tag{113}$$

Estimação da matriz de covariância de erro de estimação:

$$P_{k,k-1} = \Phi_{k-1} P_k \Phi_{k-1}^T + \Gamma_{k-1} Q_{k-1} \Gamma_{k-1}^T \tag{114}$$

Inovação:

$$v_k = y_k - H_k \hat{x}_{k,k-1} \tag{115}$$

$$\begin{aligned} \hat{x}_{k,0} &= E[x_0] \\ S_{k,0} &= 0 \end{aligned} \tag{116}$$

As matrizes S^f e S^a serão definidas pelas recursões:

$$S_{i,k-1} = \Phi_{k-1} S_{i,k} \Phi_{k-1}^T + \Gamma_{k-1}, \qquad i = 1,......N \tag{117}$$

$$S_{i,k} = [I - K_k H_k] S_{i,k-1} [I - K_k H_k]^T + K_k R_i K_k^T, \qquad i = 1,...N \tag{118}$$

com:

$$S_{i,k,0} = 0, \qquad i = 1,.....N \tag{119}$$

Portanto podemos escrever:

$$S_{k,k-1} = \sum_{i=1}^{N} \alpha_i^e S_{i,k-1} \tag{120}$$

$$S_k = \sum_{i=1}^{N} \alpha_i^e S_{i,k} \tag{121}$$

Suponhamos que: $\hat{x}_k$ é exata:

$$\hat{x}_{k,0} = x_0 \tag{122}$$

$$P_{k,0} = 0 \tag{123}$$

assim, $S_{k,0} = P_{k,0}$ e:

$$P_{k,k-1} = \sum_{i=1}^{N} \alpha_i \, S_{i,k-1} \tag{124}$$

$$P_k = \sum_{i=1}^{N} \alpha_i \, S_{i,k} \tag{125}$$

Para o caso especial (122) com ajuda de (101) - (105), obtemos:

$$E\left[v_k(k) v_{k-\ell}^T)\right] = \sum_{i=1}^{N} \alpha_i F_{i,k,\ell} \tag{126}$$

onde:

$$F_{i,k,0} = H_k S_{i,k-1} H_k^T + R_i \tag{127}$$

$$\begin{aligned} F_{i,k,\ell} &= H_k \phi_{k,k-\ell+1}^f \Phi_{k-\ell} \\ &\times \left\{ \left[I - K_{k-\ell} H_{k-\ell} \right] S_{i,k-\ell} H_{k-\ell}^T - K_{k-\ell} R_i \right\} \qquad , \ell > 0 \end{aligned} \tag{128}$$

As matrizes $F_{i,k,\ell}$ podem ser computadas recursivamente como um sub-produto do algoritmo do filtro. Tipicamente, o estimador inicial não é exato, e as relações lineares (123) e (124) entre $P_{k,k-1}$ e $S_{k,k-1}$ e entre P_k e S_k não são mantidas inicialmente. Porém elas serão mantidas para um valor muito grande de k devido à estabilidade do filtro. O filtro “esquece” as especificações iniciais de S_k exponencialmente rápido [27]. Portanto é possível escrever:

$$P_{k,k-1} = \sum_{i=1}^{N} \alpha_i S_{i,k-1} + \tilde{S}_k \tag{129}$$

onde:

$$\tilde{S}_k = \phi_{k,1}^f \Phi_0 P_{k,0} \Phi_0^T \left(\phi_{k,1}^f\right)^T \tag{130}$$

Um erro similar pode se apresentar para a equação (125) quando $P_k \neq 0$:

$$E\left[v_k v_{k-\ell}^T\right] = \sum_{i=1}^{N} \alpha_i F_{i,k,\ell} + \tilde{F}_{k,\ell} \tag{131}$$

onde:

$$\tilde{F}_{k,0} = H_k \tilde{S}_k H_k^T \tag{132}$$

$$\tilde{F}_{k,\ell} = H_k \phi_{k,k-\ell}^f \tilde{S}_{k-\ell} H_{k-\ell}^T \quad \text{for} \quad \ell > 0 \tag{133}$$

O termo $\tilde{F}_{k,\ell}$ será nulo quando k for incrementado. $\tilde{F}_{k,\ell}$ irá mais rapidamente para zero que $\tilde{S}_k$ devido à presença de H_k.

4.6. O Filtro Quadrático

Esta seção é uma adaptação dos resultados de Dee (1974) ao problema da presente tese. A equação (125) expressa o fato de que Q e R são funções lineares dos componentes de α, ou seja, a autocorrelação da seqüência de inovação é também linear em α. A abordagem agora será ver (125) como um modelo de observação de α e formular o estimador como um filtro sub-ótimo de Kalman para este parâmetro. A equação (125) pode ser escrita:

$$v_k v_k^T = \sum_{i=1}^{N} \alpha_i F_{i,k,\ell} + \xi_{k,\ell} \tag{134}$$

com:

$$E\left[\xi_{k,\ell}\right] = 0 \tag{135}$$

A equação (134) fornece um *"modelo de observação"* do vetor α desconhecido. As "*observações"* $v_k v_{k-\ell}^T$ e as *"observações padrões"* $F_i(k,l)$ serão obtidas do filtro sub-ótimo. Isto exige então que o algoritmo que estima α possa ser formulado como um *"filtro secundário"* do filtro sub-ótimo.

Antes de descrever o filtro quadrático, devemos escrever o modelo de observação de α em forma de vetor. Será usada a notação a^j para salientar a coluna j^l da matriz pxp A. O vetor vec[A] será o vetor p^2x1 obtido pela concatenação das colunas de A.

$$\mathrm{vec}[A] = \begin{bmatrix} a^1 \\ a^2 \\ . \\ . \\ . \\ . \\ a^p \end{bmatrix} \tag{136}$$

Note que, se A é simétrica, vec[A] terá elementos duplicados. Será definida a matriz T para que Tvec[A] seja a representação de [A] sem elementos duplicados.

$$T = \begin{bmatrix} T^{11} & 0 & ....0 \\ 0 & T^{22} & ....0 \\ . & 0 & ....0 \\ . & . & ....0 \\ 0 & 0 & ....T^{PP} \end{bmatrix} \tag{137}$$

onde cada bloco T^{rs} é uma matriz (p-r+1)xp

$$\begin{aligned} & T^{rs} = 0 \qquad \text{se} \quad r \neq s \\ & T^{rr} = [\underbrace{\;0\;}_{r-1} \qquad \underbrace{\;I\;}_{p-r+1}\,]\, \}p-r+1 \end{aligned} \tag{138}$$

Seja:

$$\sigma_{k,\ell} \begin{cases} \mathrm{Tvec}\left[v_k v_{k-\ell}^T)\right] & \text{se} \quad \ell = 0 \\ \mathrm{vec}\left[v_k v_{k-\ell}^T)\right] & \text{se} \quad \ell > 0 \end{cases} \tag{139}$$

$$Fs_{k,\ell}\begin{cases}\left[Tvec[F_{1,k,0}];Tvec[F_{2,k,0}];.....Tvec[F_{N,k,0}]\right] & \text{se} \quad \ell = 0\\ \\ \left[vec[F_{1,k,\ell})];vec[F_{2,k,\ell})];.....vec[F_{N,k,\ell}\right] & \text{se} \quad \ell > 0\end{cases} \tag{140}$$

$$\eta_{k,\ell} = \begin{cases}Tvec\left[\xi_{k,0})\right] & \text{se} \quad \ell = 0\\ \\ vec\left[\xi_{k,\ell}\right] & \text{se} \quad \ell > 0\end{cases} \tag{141}$$

Assim, o modelo de observação (134) de α em forma vetorial é:

$$\sigma_{k,\ell} = Fs_{k,\ell}\alpha + \eta_{k,\ell} \tag{142}$$

com:

$$E\left[\eta_{k,\ell}\right] = 0 \tag{143}$$

As *observações* $\sigma_{k,\ell}$ e as observações padrões $Fs_{k,\ell}$ são obtidas do filtro primário. A seqüência de inovação $\{v(k)\}$ gera os produtos:

$$\left\{v_k v_{k-\ell}^T; \qquad k = 1,2....; l = 0,1,...k-1\right\}$$

Na prática, deve-se especificar um máximo retardo “L”. O seu tamanho depende da aplicação particular e das considerações computacionais.

As observações $\sigma_{k,\ell}$ são geradas no seguinte ordem:

$$\left\{\sigma_{1,0};\sigma_{2,0};\sigma_{2,1};\sigma_{3,0};\sigma_{3,1};\sigma_{3,2};....\sigma_{k,0};\sigma_{k,1};...\sigma_{k,\ell}\right\}$$

Será definido o contador *t* para simplificar a notação das observações quadráticas, percorrendo os índices acima de forma crescente:

$$\sigma_t = Fs_t\alpha + \eta_t \tag{144}$$

com:

$$E[\eta_t] = 0 \tag{145}$$

$$E\left[\eta_t \eta_s^T)\right] = \nu(t,s,\alpha) \tag{146}$$

De fato, a seqüência de ruído de observação $\{\eta_t\}$ não é branca e seu segundo momento depende de α.

O abordagem de Bélanger é estimar α pela solução de um problema de mínimos quadrados: encontrar o vetor $\hat{\alpha}(t)$ que minimiza:

$$J_t = \sum_{j=0}^{t} \left(\sigma_j - Fs_j\alpha\right)^T W_j^{-1}\left(\sigma_j - Fs_j\alpha\right) + (\sigma - \sigma_0)^T \theta_0^{-1}(\sigma - \sigma_0) \tag{147}$$

onde W_j e θ_0 são matrizes de peso positiva definida e α_0 é um estimador inicial de α. Suponha, por enquanto, que $W^e{}_j$ e θ_0 são matrizes positivas definidas especificadas arbitrariamente. A solução recursiva do problema (147) com $W_j = W_j^e$ é exatamente o filtro KB do modelo.

$$\alpha_t = \alpha_{t-1} \tag{148}$$

$$\sigma_t' = Fs_t\alpha_t + \eta'_t \tag{149}$$

com:

$$E[\eta'_t] = 0 \tag{150}$$

$$E\left[\eta'_t\,(\eta'_s)^T\right] = \delta_{ts} W_t^e \tag{151}$$

e:

$$E[\alpha_0] = \alpha 0 \tag{152}$$

$$E\left[(\alpha_0 - \alpha 0)(\alpha_0 - \alpha 0)^T\right] = \theta_0 \tag{153}$$

Este abordagem aproximará o modelo verdadeiro de α (145) (146) pelo modelo (148) (149), os quais tem ruído de observação com estatísticas especificadas por (150) (151). O filtro ótimo do modelo será:

$$K^{\alpha} = \theta_{t-1} Fs_t^T \left[Fs_t \theta_{t-1} Fs_t^T + W_t^e\right]^{-1} \tag{154}$$

$$\theta_t = \left[I - K_t^{\alpha} Fs_t\right] \theta_{t-1} \tag{155}$$

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} + K_t^{\alpha}\left[\sigma_t - Fs_t \hat{\alpha}_{t-1}\right] \tag{156}$$

$$\hat{\alpha}_0 = E\left[\alpha_0\right] = \alpha 0 \tag{157}$$

$$\theta_0 = E\left[(\alpha_0 - \alpha 0)(\alpha_0 - \alpha 0)^T\right] = \theta 0 \tag{158}$$

Pode-se ver que o desempenho do filtro α dependerá de quão perto o modelo inicial está do modelo verdadeiro. Assim a especificação de W^e e θ^e são essenciais para a taxa de convergência do algoritmo.

4.6.1. Escolha de W^e e θ^e

O filtro α será ótimo se a seqüência de ruído $\{\eta(t)\}$ for branca e se:

$$\alpha 0 = E[\alpha] \tag{159}$$

$$\theta 0 = E\left[(\alpha - \alpha 0)(\alpha - \alpha 0)^T\right] \tag{160}$$

$$W_t^e = E\left[\eta_t \eta_t^T\right] \tag{161}$$

No caso real:

$$\begin{aligned} E\left[\eta_t \eta_t^T\right] &= \nu(t,s,\alpha) \\ &= E\left[\left[\sigma_{k1,\ell 1} - E(\sigma_{k1,\ell 1})\right]\left[\sigma_{k2,\ell 2} - E\sigma_{k2,\ell 2}\right]^T\right] \end{aligned} \tag{162}$$

com:

$$\sigma_{k1,\ell 1} = \begin{cases} T\ \mathrm{vec}\left[v_{k1} v_{k1}^T\right] & \text{se} \quad \ell_1 = 0 \\ \\ \mathrm{vec}[v_{k1}) v_{k1-\ell 1})] & \text{se} \quad \ell_1 > 0 \end{cases} \tag{163}$$

Assim, a auto-correlação de η_t é uma função de momento de quarto ordem da seqüência de inovação e estes momentos não podem ser especificados sem o conhecimento de α. Fazendo uma seleção de W_t^e independentemente de α, será introduzida uma aproximação linear ao problema.

A atual covariância é dada por:

$$W_t = E\left[\eta_t \eta_t^T\right] = \begin{cases} TW_{k,0} & \text{se} \quad \ell = 0 \\ \\ W_{k,\ell} & \text{se} \quad \ell > 0 \end{cases} \tag{164}$$

onde:

$$W_{k,\ell} = E\Bigg[\left(\mathrm{vec}\left[v_k v_k^T\right] - E\left[\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right]\right) \times \left(\mathrm{vec}\left[v_k v_k^T\right] - E\left[\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right]\right)^T\Bigg]$$

$$\begin{aligned} W_{k,\ell} = {} & E\left[\left(\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right)\left(\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right)^T\right] \\ & - E\left[\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right] E\left[\left(\mathrm{vec}\left[v_k v_{k-\ell}^T\right]\right)^T\right] \end{aligned} \tag{165}$$

Portanto, $W_{k,\ell}$ é uma matriz p^2 x p^2 .

$$W_{k,\ell} = \begin{bmatrix} W^{11} & W^{12} \;..............W^{1p} \\ W^{21} & W^{22}.................W^{2p} \\ . & . \quad .... \\ . & . \quad .... \\ . & . \quad .... \\ W^{p1} & W^{p2}................W^{pp} \end{bmatrix} \qquad (166)$$

onde cada bloco W^{rs} é uma matriz *p x p* com:

$$\begin{aligned} W_{ij}^{rs}(k,\ell) &= E\left[v_{i,k} v_{r,k-\ell} v_{j,k} v_{s,k-\ell}\right] \\ &\quad - E\left[v_{i,k} v_{r,k-\ell}\right] E\left[v_{j,k} v_{s,k-\ell}\right] \end{aligned} \qquad (167)$$

Utilizando o fato de que v_k é gaussiana, podemos usar a regra do produto (relacionando momentos de quarta e de segunda ordem de variáveis gaussianas) para expressar os momentos de quarta ordem (166) em termos de momentos de segunda ordem:

$$\begin{aligned} W_{ij}^{rs}(k,\ell) &= E\left[v_{i,k} v_{j,k}\right] E\left[v_{r,k-\ell} v_{s,k-\ell}\right] \\ &\quad + E\left[v_{i,k} v_{s,k-\ell}\right] E\left[v_{r,k-\ell} v_{j,k}\right] \end{aligned} \qquad (168)$$

Agora será introduzida uma aproximação, fazendo $\alpha=\alpha^e$ (Bélanger 1974), com isto a seqüência de inovação será branca (conforme seção 4.5.5), desde que suponhamos (por simplificação) que $\{\eta(t)\}$ é um ruído branco (o que modifica (162) (163)). Com estas novas modificações, e fazendo $W=W^e$:

$$\begin{aligned} \left(W_{ij}^{rs}\right)^e &= E\left[v_{i,k}^e v_{j,k}^e\right] E\left[v_{r,k-\ell}^e v_{s,k-\ell}^e\right] \\ &\quad + \delta_{l0} E\left[v_{i,k}^e v_{s,k}^e\right] E\left[v_{r,k}^e v_{j,k}^e\right] \end{aligned} \qquad (169)$$

onde $\{v^e(k)\}$é a seqüência de inovação quando $\alpha = \alpha^e$. O lado direito da equação (169) pode ser avaliado usando a equação (134):

$$\begin{aligned} E\left[v_k^e \left(v_k^e\right)^T\right] &= \sum_{i=1}^{N} \alpha_i^e F_{i,k,0} \\ &= H_k S_{k,k-1} H_k^t + R^e \end{aligned} \qquad (170)$$

Agora completamos a especificação do filtro α:

- O estimador recursivo é dado por:

$$K^{\alpha} = \theta_{t-1} Fs_t^T \left[Fs_t \theta_{t-1} Fs_t^T + W_t^e \right]^{-1} \tag{154}$$

$$\theta_t = \left[I - K_t^{\alpha} Fs_t \right] \theta_{t-1} \tag{155}$$

$$\hat{\alpha}_t = \hat{\alpha}_{t-1} + K_t^{\alpha} \left[\sigma_t - Fs_t \hat{\alpha}_{t-1} \right] \tag{156}$$

$$\hat{\alpha}_0 = E[\alpha_0] = \alpha 0 \tag{157}$$

$$\theta_0 = E\left[(\alpha_0 - \alpha 0)(\alpha_0 - \alpha 0)^T \right] = \theta 0 \tag{158}$$

- As observações quadráticas $\sigma_{k,\ell}$:

$$\sigma_{k,\ell} \begin{cases} Tvec\left[v_k v_{k-\ell}^T) \right] & se \quad \ell = 0 \\ vec\left[v_k v_{k-\ell}^T) \right] & se \quad \ell > 0 \end{cases} \tag{139}$$

- As observações padrões $Fs(k,l)$:

$$Fs_{k,\ell} \begin{cases} \left[Tvec[F_{1,k,0}]; Tvec[F_{2,k,0}]; \ldots..Tvec[F_{N,k,0}] \right] & se \quad \ell = 0 \\ \left[vec[F_{1,k,\ell})]; vec[F_{2,k,\ell})]; \ldots..vec[F_{N,k,\ell} \right] & se \quad \ell > 0 \end{cases} \tag{140}$$

- Finalmente a matriz de ponderação W^e é definida pelas equações (164) - (170).