IE 009 – Processamento Adaptativo de Sinais.

* Filtro de Kalman

O conceito de filtragem de Kalman está intimamente ligado à noção de estimação recursiva MMSE, tendo como pano de fundo a definição do espaço de estados do sistema dinâmico tratado.

– Definição do Problema

• parâmetros de interesse associados à operação de um sistema dinâmico variam com o tempo.

• estes parâmetros apresentam erros de medição ou não podem ser observados diretamente. Somente podemos recorrer a medidas ruidosas de grandezas disponíveis para obter informações sobre os parâmetros internos.

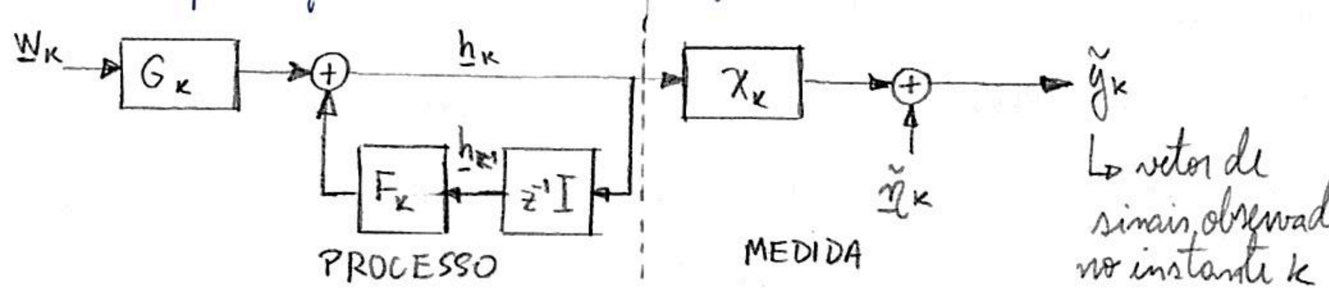
Estado: conjunto mínimo de variáveis que fornece uma representação completa das condições internas do sistema em um dado instante de tempo. Em um instante k_0, o estado juntamente com a entrada para $k \geq k_0$ determina de maneira única a saída do sistema para $k \geq k_0$.

Representação de um sistema dinâmico linear a tempo discreto

(1) $\underline{h}_k = F_k \underline{h}_{k-1} + G_k \underline{w}_k$ Equação de processo → pode incluir uma entrada externa do tipo $B_k \underline{u}_k$

(2) $\breve{y}_k = X_k \underline{h}_k + \underline{\eta}(k)$ Equação de medida

A primeira equação define a dinâmica das variáveis de estado, enquanto a segunda relaciona estas variáveis aos sinais que efetivamente conseguimos observar.



$\underline{h}_k$: vetor de estados que desejamos estimar a cada instante k.

F_k : matriz de transição de estados – descreve a evolução temporal das variáveis de estado entre os instantes k-1 e k

$\underline{w}_k$: vetor de ruído de processo (ou de excitação)

X_k : matriz de medida (observação)

$\tilde{\underline{\eta}}_k$: ruído aditivo de medida.

• os ruídos $\underline{w}_k$ e $\tilde{\underline{\eta}}_k$ são brancos, descorrelacionados, de média nula e matrizes de covariância dadas por $E\{\underline{w}_k \underline{w}_l^H\} = \begin{cases} Q_k, & k=l \\ 0, & k \neq l \end{cases}$

e $E\{\tilde{\underline{\eta}}_k \tilde{\underline{\eta}}_l^H\} = \begin{cases} R_k, & k=l \\ 0, & k \neq l \end{cases}$

<u>Objetivo</u>: a partir do conhecimento da dinâmica do sistema (modelo) e do sinal observado $\tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_n$, queremos estimar o vetor $\underline{h}_i$

* Filtragem : $i = n$
* Predição : $i > n$
* Suavização : $0 \leq i \leq n$

<u>Critério</u> <u>para</u> <u>estimação</u> : mínimo erro quadrático médio (MMSE)

Vimos no tópico inicial deste curso que o estimador MMSE é obtido pela média condicional dos parâmetros desconhecidos (variáveis de estado) dados os sinais observados:

$$\hat{\underline{h}}_{MMSE} = E\{\underline{h} \mid \{\tilde{\underline{y}}\}\}$$

, onde $\{\tilde{y}\}$ denota a sequência observada disponível.

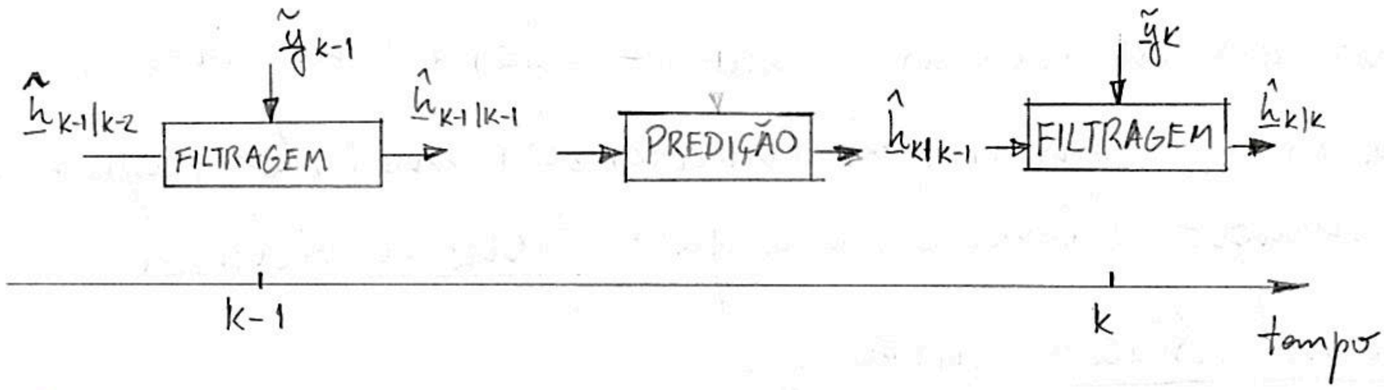
⇒ Como o estado possui natureza dinâmica, é interessante desenvolver algoritmos que calculem a estimativa ótima a cada instante de tempo. Se nos restringirmos aos estimadores lineares - cujas estimativas são dadas por combinações lineares das medidas disponíveis - a solução ótima é dada pelo <u>filtro</u> <u>de</u> <u>Kalman</u>.

<u>Características</u> <u>iniciais</u>:

- fornece uma solução recursiva de mínimo erro quadrático médio para o problema de estimação de estados.
- a recursão evita o armazenamento de todo o histórico de sinais observados, tornando o algoritmo computacionalmente eficiente.
- versatilidade para trabalhar em ambientes estacionários e não-estacionários.

*Derivação do Filtro de Kalman

O processo recursivo da filtragem de Kalman pode ser dividido em duas etapas: (i) na fase de predição, a equação de processo é usada para predizer o estado em um instante k a partir da estimativa do estado calculada no instante k-1; (ii) na fase de filtragem, os valores preditos para as variáveis de estado são corrigidos (atualizados) com base na equação de medida e na observação do instante k.



$\hat{\underline{h}}_{k-1|k-2}$ representa a estimativa do estado no instante k-1 baseada nos sinais observados até o instante k-2.

- a melhor estimativa do estado no instante k, dada a observação dos sinais $\tilde{\underline{y}}$ do instante inicial até k, é calculada como:

$$\hat{\underline{h}}_{k|k} = E\{\underline{h}_k \mid \tilde{\underline{y}}_1, \tilde{\underline{y}}_2, \ldots, \tilde{\underline{y}}_k\}$$

Se, porém, tivermos acesso às medidas até o instante k-1, podemos computar uma estimativa <u>a priori</u> do estado:

$$\hat{\underline{h}}_{k|k-1} = E\{\underline{h}_k \mid \tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_{k-1}\}$$

Usando a equação de estado, podemos escrever:

$$\hat{\underline{h}}_{k|k-1} = E\{F_k \underline{h}_{k-1} + G_k \underline{w}_k \mid \tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_{k-1}\}$$

$$= F_k E\{\underline{h}_{k-1} \mid \tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_{k-1}\} + G_k E\{\underline{w}_k \mid \tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_{k-1}\}$$

Como $\underline{w}_k$ é independente das observações passadas,

$$= F_k E\{\underline{h}_{k-1} \mid \tilde{\underline{y}}_1, \dots, \tilde{\underline{y}}_{k-1}\} + \cancelto{0}{G_k E\{\underline{w}_k\}}$$

$$= \underline{F_k \cdot \hat{\underline{h}}_{k-1|k-1}} \quad (1)$$

O desempenho (ou qualidade) de um estimador pode ser calculado através da matriz de covariância do erro de estimação:

$$P_{k|k-1} = E\{(\underline{h}_k - \hat{\underline{h}}_{k|k-1})(\underline{h}_k - \hat{\underline{h}}_{k|k-1})^H\}$$

$$= E\left\{\left[F_k \underline{h}_{k-1} + G_k \underline{w}_k - F_k \hat{h}_{k-1|k-1}\right]\left[F_k h_{k-1} + G_k \underline{w}_k - F_k \hat{\underline{h}}_{k-1|k-1}\right]^H\right\}$$

$$= E\left\{\left[F_k(\underline{h}_{k-1} - \hat{\underline{h}}_{k-1|k-1}) + G_k \underline{w}_k\right]\left[F_k(\underline{h}_{k-1} - \hat{\underline{h}}_{k-1|k-1}) + G_k \underline{w}_k\right]^H\right\}$$

$$= F_k E\{(\underline{h}_{k-1} - \hat{\underline{h}}_{k-1|k-1})(\underline{h}_{k-1} - \hat{h}_{k-1|k-1})^H\} F_k^H + G_k E\{\underline{w}_k \underline{w}_k^H\} G_k^H$$

$$= \underline{F_k P_{k-1|k-1} F_k^H + G_k Q_k G_k^H} \quad (2)$$

A equações (1) e (2) dependem do conhecimento da dinâmica do sistema e da estimativa do estado no instante anterior ($\hat{h}_{k-1|k-1}$) ⇒ por isso, elas formam a etapa de <u>predição</u>.

O próximo passo é entender como os sinais observados são utilizados para refinar a estimativa produzida na etapa de predição.

• <u>Estimador linear recursivo</u>: $\underline{\hat{h}}_{k|k} = \underline{\hat{h}}_{k|k-1} + K_k\overbrace{(\tilde{\underline{y}}_k - \underbrace{X_k\underline{\hat{h}}_{k|k-1}}_{\text{estimativa da observação } k \text{ usando as informações até } k-1})}^{\text{erro}}$

Qual o valor ótimo de K_k?

$J_k = Tr\{P_{k|k}\}$, onde $P_{k|k}$ é a matriz de covariância do erro de estimação do estado.

Ora, $\underline{e}_{k|k} = \underline{h}_k - \underline{\hat{h}}_{k|k}$

$$= \underbrace{\underline{h}_k - \underline{\hat{h}}_{k|k-1}}_{} - K_k(\tilde{\underline{y}}_k - X_k\underline{\hat{h}}_{k|k-1})$$

$$= \underline{e}_{k|k-1} - K_k(X_k\underline{h}_k + \tilde{\underline{\eta}}_k - X_k\underline{\hat{h}}_{k|k-1})$$

$$= \underline{e}_{k|k-1} - K_k\left(X_k(\underline{h}_k - \underline{\hat{h}}_{k|k-1}) + \tilde{\underline{\eta}}_k\right)$$

$$= \underline{e}_{k|k-1} - K_k X_k \underline{e}_{k|k-1} - K_k \tilde{\underline{\eta}}_k$$

$$= (I - K_k X_k)\,\underline{e}_{k|k-1} - K_k \tilde{\underline{\eta}}_k$$

Então, $P_{k|k} = E\{\underline{e}_{k|k} \cdot \underline{e}_{k|k}^H\}$

$$= E\{[(I - K_k X_k)\underline{e}_{k|k-1} - K_k\tilde{\underline{\eta}}_k)][\cdots]^H\}$$

Os termos cruzados do erro com o ruído serão cancelados pois o ruído é independente do erro e possui média nula. Logo,

$$P_{k|k} = (I - K_k X_k)P_{k|k-1}(I - K_k X_k)^H + K_k R_k K_k^H \quad (*)$$

MÍNIMA COVARIÂNCIA DO ERRO: $\dfrac{\partial J_k}{\partial K_k} = 0$

$$\frac{\partial J_k}{\partial K_k} = \frac{\partial \, Tr\{P_{k|k}\}}{\partial K_k} = 0$$

Propriedade: $\dfrac{\partial \, Tr\{ABA^H\}}{\partial A} = 2AB$ se B é simétrica.

$$\frac{\partial J_k}{\partial K_k} = 2(I - K_k X_k) P_{k|k-1}(-X_k^H) + 2K_k R_k = 0$$

Solução: $\underline{\underline{K_k = P_{k|k-1} X_k^H \left(X_k P_{k|k-1} X_k^H + R_k \right)^{-1}}}$ (3)

Voltando à equação do estimador linear recursivo:

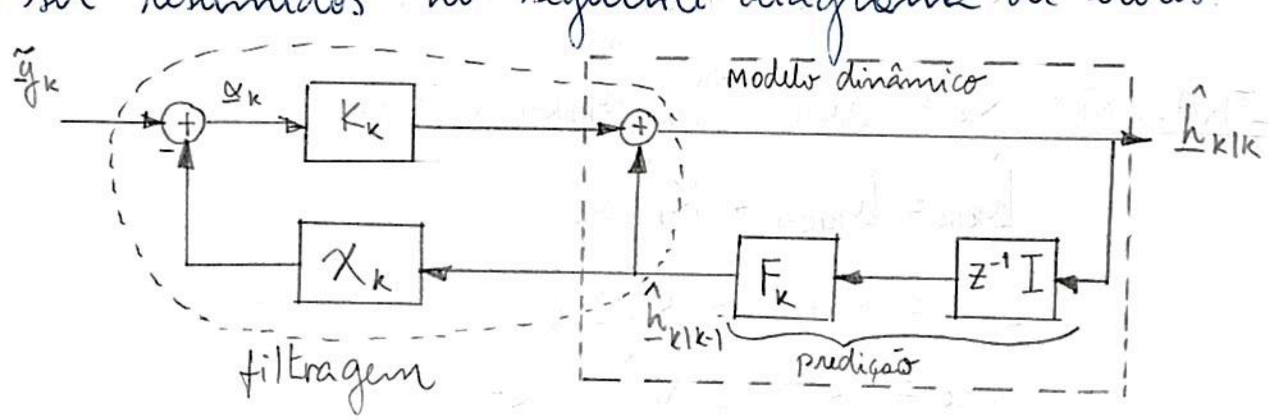
$$\hat{\underline{h}}_{k|k} = \hat{\underline{h}}_{k|k-1} + K_k \underbrace{\left(\tilde{\underline{y}}_k - X_k \hat{\underline{h}}_{k|k-1} \right)}_{\underline{\alpha}_k}$$

erro de predição de $\tilde{\underline{y}}_k$ a partir de $\hat{h}_{k|k-1}$:

$$\underline{\alpha}_k = \tilde{\underline{y}}_k - X_k \hat{\underline{h}}_{k|k-1} \quad (4)$$

Logo, $\underline{\underline{\hat{\underline{h}}_{k|k} = \hat{\underline{h}}_{k|k-1} + K_k \underline{\alpha}_k}}$ (5)

As duas etapas do processo de filtragem de Kalman podem ser resumidos no seguinte diagrama de blocos



Observe que o modelo dinâmico do sistema faz parte do filtro de Kalman. Assim, o termo $K_k \underline{\alpha}_k$ do filtro de Kalman pode ser visto como uma estimativa do ruído de excitação $G_k \underline{w}_k$.

É justamente pelo fato de o filtro de Kalman usar uma "cópia" do modelo dinâmico do sistema que ele é capaz de rastrear a evolução temporal do estado.

A última operação do filtro de Kalman consiste da atualização da covariância do erro de estimação na etapa de filtragem. A partir da equação (*), pode-se mostrar:

$$P_{k|k} = (I - K_k X_k) P_{k|k-1} (I - K_k X_k)^H + K_k R_k K_k^H$$

$$= \left[(P_{k|k-1})^{-1} + X_k^H R_k^{-1} X_k \right]^{-1}$$

$$= \underline{\underline{(I - K_k X_k) P_{k|k-1}}} \quad (6)$$

Reunindo as expressões de (1) a (6), obtemos o filtro de Kalman

PREDIÇÃO:

$$\hat{\underline{h}}_{k|k-1} = F_k \cdot \hat{\underline{h}}_{k-1|k-1}$$

$$P_{k|k-1} = F_k P_{k-1|k-1} F_k^H + G_k Q_k G_k^H$$

FILTRAGEM:

$$K_k = P_{k|k-1} X_k^H \left(X_k P_{k|k-1} X_k^H + R_k \right)^{-1}$$

$$\underline{\alpha}_k = \tilde{\underline{y}}_k - X_k \hat{\underline{h}}_{k|k-1}$$

$$\hat{\underline{h}}_{k|k} = \hat{\underline{h}}_{k|k-1} + K_k \underline{\alpha}_k$$

$$P_{k|k} = (I - K_k X_k) P_{k|k-1}$$

Observações e propriedades:

- o termo de correção $\underline{\alpha}_k$ é chamado de inovação

$$\hat{\underline{\tilde{y}}}_{k|k-1} = E\{\underline{\tilde{y}}_k \mid \underline{\tilde{y}}_1, \ldots, \underline{\tilde{y}}_{k-1}\}$$

$$= E\{X_k \underline{h}_k + \underline{\tilde{\eta}}_k \mid \underline{\tilde{y}}_1, \ldots, \underline{\tilde{y}}_{k-1}\}$$

$$= X_k E\{\underline{h}_k \mid \underline{\tilde{y}}_1, \ldots, \underline{\tilde{y}}_{k-1}\} + E\{\underline{\tilde{\eta}}_k\} = X_k \hat{\underline{h}}_{k|k-1}$$

Logo, $\underline{\alpha}_k = \underline{\tilde{y}}_k - \hat{\underline{\tilde{y}}}_{k|k-1} \Rightarrow \underline{\tilde{y}}_k = \hat{\underline{\tilde{y}}}_{k|k-1} + \underline{\alpha}_k$

O termo $\hat{\underline{\tilde{y}}}_{k|k-1}$ é a parcela da observação atual $\underline{\tilde{y}}_k$ que pode ser predita a partir das observações passadas (possui redundância em relação ao passado). $\underline{\alpha}_k$, por sua vez, é a parcela nova, que não pode ser predita (daí o nome inovação).

Em termos geométricos, $\hat{\underline{\tilde{y}}}_{k|k-1}$ é a projeção ortogonal de $\underline{\tilde{y}}_k$ no espaço gerado por $\{\underline{\tilde{y}}_1, \ldots, \underline{\tilde{y}}_{k-1}\}$, enquanto $\underline{\alpha}_k$ é a componente de $\underline{\tilde{y}}_k$ ortogonal a esse espaço.

É possível mostrar que a seqüência de inovações $\underline{\alpha}_1, \ldots, \underline{\alpha}_k$ forma um processo branco de média nula.

- quando $\underline{w}_k$ e $\underline{\tilde{\eta}}_k$ são processos aleatórios brancos e gaussianos, o filtro de Kalman corresponde à solução ótima MMSE para o problema de estimação de estados.

- o Filtro de Kalman sempre representa o estimador ótimo MMSE linear (que produz uma estimativa a partir de uma combinação linear das medidas) quando os ruídos considerados não são gaussianos, por exemplo.

• se $\underline{w}_k$ e $\tilde{\eta}_k$ forem correlacionados ou coloridos, então o filtro de Kalman pode ser modificado para resolver o problema.

• para sistemas não-lineares, várias formulações de filtros de Kalman não-lineares oferecem soluções aproximadas para o problema.

• O algoritmo RLS pode ser visto como um caso particular do filtro de Kalman.

• $P_{k|k-1}$, K_k e $P_{k|k}$ não dependem de $\tilde{y}_k$, apenas das características do sistema dinâmico (X_k; G_k, Q_k e R_k) - poderiam, portanto, ser pré-calculados e armazenados em uma memória. É possível saber se o filtro terá ou não sucesso de modo antecipado, verificando se $P_{k|k}$ converge para valores pequenos.

• Inicialização da matriz de covariância: $\begin{cases} P_{0|0} = 0, \text{ se conhecemos exatamente } \underline{h}_0 \\ P_{0|0} = \infty I, \text{ se não sabemos nada sobre o estado inicial.} \end{cases}$

• Há alguns motivos que podem levar o filtro de Kalman a não funcionar adequadamente: (i) aritmética de precisão finita, (ii) erros na definição do modelo matemático do sistema dinâmico.

"Truques" interessantes:

- usar um filtro com desvanecimento de memória
- inserir um ruído de processo fictício

} fazem com que o filtro dê mais ênfase às medidas do que ao modelo adotado (eq. de processo)

Referências Complementares:

- D. Simon, "Optimal State Estimation - Kalman, H∞ and Nonlinear Approaches", John Wiley & Sons, 2006.
- B.D.O. Anderson e J.B. Moore, "Optimal Filtering", Prentice Hall, 1979.
- S.M. Kay, "Fundamentals of Statistical Signal Processing, Volume I: Estimation Theory", Prentice Hall, 1993.
- T. Kailath, A.H. Sayed e B. Hassibi, "Linear Estimation", Prentice Hall, 2000.