

FILTRAGEM DE SISTEMAS

&

E FILTRO DE KALMAN

João B. R. do Val e Wagner C. do Amaral

Depto. de Telemática – FEEC – UNICAMP

Depto. de Computação e Automação – FEEC – UNICAMP

IA-856 IDENTIFICAÇÃO E FILTRAGEM

18 de novembro de 2010

arquivo filtro_kalman_trans05.tex

[FULLSCREEN](#) [NEXT](#)

FILTRO DE KALMAN

O PROBLEMA DE FILTRAGEM

■ Filtragem ou estimação de estado.

Considere o sistema não-linear com ruído nas variáveis de estado e na equação de observação:

$$\begin{cases} x_{k+1} = f_k(x_k, e_k) & (\text{processo de estado}) \\ y_k = h_k(x_k, v_k) & (\text{processo de observação}) \end{cases} \quad (1)$$

onde:

- $f_k(\cdot, \cdot)$ e $h_k(\cdot, \cdot)$ são funções que descrevem respectivamente, a evolução dinâmica e a observação a cada instante $k = 0, 1, \dots$;
- $\{e_k\}$ e $\{v_k\}$ são seqüências de vetores de v.a.'s (ruídos) independentes e identicamente distribuídos (i.i.d.);
- O estado inicial x_0 tem uma distribuição conhecida;

O PROBLEMA DE FILTRAGEM

- y_k forma a sequência de observações do sistema, i.e., não há observação do vetor de estados $\{x_k\}$, a não ser indiretamente através de $\{y_k\}$.
- PROBLEMA DE FILTRAGEM: conhecendo-se $\mathbf{y}^k = \{y_0, y_1, \dots, y_k\}$ obter estimativas filtrada para x_k , o que significa obter

$$p(x_k|\mathbf{y}^k)$$

1. Suponha que $p(x_{k-1}|\mathbf{y}^{k-1})$ fosse conhecida no instante $k-1$. Então a PREDIÇÃO UM ESTÁGIO A FRENTE seria calculada como:

$$p(x_k|\mathbf{y}^{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|\mathbf{y}^{k-1})dx_{k-1}$$

pois $p(x_k|x_{k-1}, \mathbf{y}^{k-1}) = p(x_k|x_{k-1})$, definida pelo processo de estado.

[BACK](#) [NEXT](#)

3/51

O PROBLEMA DE FILTRAGEM

2. No instante k , quando a medida y_k é obtida, um ESTÁGIO DE ATUALIZAÇÃO deve ser feito. Usando a regra de Bayes,

$$p(x_k|\mathbf{y}^k) = p(x_k|y_k, \mathbf{y}^{k-1}) = \frac{p(y_k|x_k, \mathbf{y}^{k-1})p(x_k|\mathbf{y}^{k-1})}{p(y_k|\mathbf{y}^{k-1})} = \frac{p(y_k|x_k)p(x_k|\mathbf{y}^{k-1})}{p(y_k|\mathbf{y}^{k-1})}$$

3. A constante de normalização é obtida como

$$p(y_k|\mathbf{y}^{k-1}) = \int p(y_k|x_k)p(x_k|\mathbf{y}^{k-1})dx_k$$

depende da verossimilhança $p(y_k|x_k)$ dada pelo processo de observação em (1).

- As equações no passo de predição em 1. e de atualização em 2. formam a base da solução ótima do filtro utilizando a fórmula de Bayes. O conhecimento de $p(x_k|\mathbf{y}^k)$ permite o cálculo do estimador de mínimo erro quadrático médio, a partir da média condicional de x_k

$$E[x_k|\mathbf{y}^k] = \int x_k p(x_k|\mathbf{y}^k)dx_k$$

[BACK](#) [NEXT](#)

4/51

SISTEMAS LINEARES ESTOCÁSTICOS

Sistema linear com ruído aditivo nas variáveis de estado e na equação de observação

$$\begin{cases} x_{k+1} = Ax_k + Fe_k \\ y_k = Cx_k + Dv_k \end{cases}$$

onde:

- A, F, C, D são matrizes de dimensão $n \times n, n \times \ell, r \times n, r \times \ell$ respectivamente,
- $\{e_k\}$ é uma seqüência de vetores de v.a.'s gaussianas de média nula, não correlacionados e $\text{cov}(e_k) = E[e_k e_k'] = Q = Q' > 0, \quad e_k \sim N(0, Q);$
- $\{v_k\}$ é uma seqüência de vetores de v.a.'s gaussianas de média nula, não correlacionados e $\text{cov}(v_k) = E[v_k v_k'] = R = R' > 0, \quad v_k \sim N(0, R);$

[BACK](#)
[NEXT](#)

5/51

SISTEMAS LINEARES ESTOCÁSTICOS

- $x_0 \sim N(\bar{x}_0, P_0);$
- * O estado inicial x_0 , $\{e_k\}$ e $\{v_k\}$ são não correlacionados ($\rightarrow$ independentes) entre si;
- * y_k forma a sequência de observações do sistema, i.e., não há observação do vetor de estados $\{x_k\}$, a não ser indiretamente através de $\{y_k\}$.

[BACK](#)
[NEXT](#)

6/51

CÁLCULOS DE PREDIÇÃO DE MÉDIA E COVARIÂNCIA

Suponha que $E[x_0] = \bar{x}_0$, $\text{cov}(x_0) = P_0$. Então $P(k) := \text{cov}(x_k)$ satisfaz

$$E[x_k] := \bar{x}_k \text{ satisfaz } \bar{x}_{k+1} = A\bar{x}_k, \bar{x}_0 = E[x_0] \quad (2)$$

e

$$P(k+1) = AP(k)A' + FQF', P(0) = P_0 \quad (3)$$

e

$$\text{cov}(y_k, y_{k-j}) = \begin{cases} CP(k)C' + DRD', & \text{se } j = 0 \\ CA^jP(k-j)C', & \text{se } j = 1, \dots, k \end{cases} \quad (4)$$

PROVA: (2) é trivial.

[BACK](#)
[NEXT](#)

7/51

CÁLCULOS DE PREDIÇÃO DE MÉDIA E COVARIÂNCIA

PROVA: (3): Definindo

$$\tilde{x}_k := x_k - \bar{x}_k, \quad \tilde{y}_k = y_k - \bar{y}_k, \quad \bar{y}_k = E[y_k]$$

podemos concluir que

$$\tilde{x}_k = A\tilde{x}_{k-1} + Fe_{k-1} = A^j\tilde{x}_{k-j} + A^{j-1}Fe_{k-j} + \dots + Fe_{k-1}$$

Como $\tilde{x}_k$ e e_k são não correlacionados

$$\text{cov}(x_{k+1}) = P(k+1) = E[(A\tilde{x}_k + Fe_k)(A\tilde{x}_k + Fe_k)'] = AP(k)A' + FQF'$$

e (3) ✓. Agora (4)

$$\text{cov}(y_k) = E[(C\tilde{x}_k + Dv_k)(C\tilde{x}_k + Dv_k)'] = CP(k)C' + DRD'$$

[BACK](#)
[NEXT](#)

8/51

CÁLCULOS DE PREDIÇÃO DE MÉDIA E COVARIÂNCIA

Como $\tilde{x}_{k-j}$ é não correlacionado com $e_{k-1}, \dots, e_{k-j}$,

$$E[\tilde{x}_k \tilde{x}'_{k-j}] = A^j P(k-j), \quad E[\tilde{x}_k e'_{k-j}] = A^{j-1} F Q$$

Assim podemos avaliar:

$$\begin{aligned} \text{cov}(y_k y_{k-j}) &= E[\tilde{y}_k \tilde{y}'_{k-j}] \\ &= E[(C\tilde{x}_k + Dv_k)(C\tilde{x}_{k-j} + Dv_{k-j})'] = CA^j P(k-j)C' \end{aligned}$$

e (4) ✓

[BACK](#)
[NEXT](#)

9/51

FILTRO DE KALMAN

- Assuma adicionalmente que $DD' > 0$.

Isso implica que $\ell \geq r$. Senão, teríamos um vetor λ tal que $\lambda'D = 0$, e assim,

$$\lambda'y_k = \lambda'Cx_k$$

e certas combinações lineares de componentes do vetor de estados x_k poderiam ser medidas precisamente, sem ruído.

NOTAÇÃO:

- $\hat{x}_{k|j}$ o melhor estimador linear de x_k dada a seqüência de observações $y^j = (y_0, \dots, y_j)$;
- $\tilde{x}_{k|j} = x_k - \hat{x}_{k|j}$ o erro de estimação;
- $\hat{x}_{k|k-1}$ estimador “um passo a frente”.

Distribuições “um passo a frente”: $p(x_{k+1}|x_k, u_k)$ e $p(y_k|x_k)$?

[BACK](#)
[NEXT](#)

10/51

DISTRIBUIÇÃO UM PASSO A FRENTE

Calculando $E[Fe_k] = FE[e_k] = 0$, $\text{cov}(Fe_k) = E(Fe_k - 0)(Fe_k - 0)' = FE[e_k e_k']F' = FQF'$ portanto

$$Fe_k \sim N(0, FQF')$$

e portanto,

$$p(x_{k+1}|x_k) \sim N(Ax_k, FQF')$$

$$p(y_k|x_k) \sim N(Cx_k, DQD')$$

Qual é a probabilidade de transição em múltiplos estágios:

$$p(x_{k+m}|x_k,)?$$

Podemos escrever:

$$x_{k+m} = A^m x_k + \sum_{t=0}^{m-1} A^{m-1-t} Fe_{k+t}$$

[BACK](#)
[NEXT](#)

11/51

DISTRIBUIÇÃO UM PASSO A FRENTE

então,

$$p(x_{k+m}|x_k,) \equiv N(A^m x_k, \Sigma_{k+m|k})$$

onde

$$\begin{aligned} \Sigma_{k+m|k} &= \text{cov}\left(\sum_{t=0}^{m-1} A^{m-1-t} Fe_{k+t}\right) = \text{cov}\left(\sum_{t=0}^{m-2} A^{m-1-t} Fe_{k+t} + Fe_{k+m-1}\right) \\ &= \text{cov}\left(A\left[\sum_{t=0}^{m-2} A^{m-2-t} Fe_{k+t}\right] + Fe_{k+m-1}\right) = A\Sigma_{k+m-1|k}A' + FQF' \end{aligned}$$

com $\Sigma_{k|k} = 0$. Quando $m = 1$

$$p(x_{k+1}|x_k) \equiv N(Ax_k, FQF')$$

[BACK](#)
[NEXT](#)

12/51

DISTRIBUIÇÃO M-PASSOS A FRENTE

Observações:

- 1) A distribuição condicional de x_{k+1} é sempre gaussiana.
- 2) A distribuição não condicional não é gaussiana em geral.

Exemplo: $x_{k+1} = x_k + u_k + w_k$ (sistema controlado)

$$x_0 \sim N(0, 1), w_0 \sim N(0, 1), x_1 = x_0 + u_0 + w_0$$

Suponha que $u_0 = x_0^2$. Então, $x_1 = x_0 + x_0^2 + w_0$ e $p(x_1)$ não é gaussiana. Porém $p(x_1|x_0, u_0)$ é gaussiana.

[BACK](#)
[NEXT](#)

13/51

DISTRIBUIÇÃO M-PASSOS A FRENTE

■ Em resumo, se

$$\begin{cases} x_{k+1} = Ax_k + Fe_k \\ y_k = Cx_k + Dv_k \end{cases}$$

Média: $\bar{x}_{k+m} = A^m \bar{x}_k$

Covariância: $E[(x_{k+m} - \bar{x}_{k+m})(x_{k+m} - \bar{x}_{k+m})'] := \Sigma_{k+m}$

como

$$(x_{k+m} - \bar{x}_{k+m}) = A^m(x_k - \bar{x}_k) + \sum_{t=0}^{m-1} A^{m-1-t} Fe_{k+t}$$

então

$$\Sigma_{k+m} = A^m \Sigma_k A^{m'} + \sum_{t=0}^{m-1} A^{m-1-t} F Q F' (A^{m-1-t})'$$

[BACK](#)
[NEXT](#)

14/51

DISTRIBUIÇÃO M-PASSOS A FRENTE

- Para a variável de saída:

$$\bar{y}_k = C\bar{x}_k$$

definindo $\Sigma_k^y := \text{cov}((y_k - \bar{y}_k)(y_k - \bar{y}_k)')$, então

$$\Sigma_k^y = C\Sigma_k C' + DRD'$$

e também:

$$E[(y_{k+m} - \bar{y}_{k+m})(y_k - \bar{y}_k)'] := CA^m \Sigma_k C'$$

↪ Caso x_0, e_k, v_k não sejam gaussianas, as fórmulas de média e covariância ainda valem.

[BACK](#) [NEXT](#)

15/51

ESTIMAÇÃO DE ESTADO PARA SISTEMAS LINEARES

$$\begin{cases} x_{k+1} = A_k x_k + F_k e_k \\ y_k = C_k x_k + D_k v_k \end{cases}$$

- $x_0, w_0, w_1, \dots, v_0, v_1, \dots$ independentes
- $x_0 \sim N(\bar{x}_0, \Sigma_0) \quad e_k \sim N(0, Q) \quad v_k \sim N(0, R)$
- Sistema Linear Gaussiano variante no tempo $k \rightarrow (A_k, F_k, C_k, D_k)$ seqüências de matrizes conhecidas.

Já vimos que:

1. $p(x_k | y^k)$ é gaussiano. Aqui:

$$p(x_k | y^k) \sim N(\hat{x}_{k|k}, \Sigma_{k|k})$$

com $\hat{x}_{k|k}$ = média condicional, $\Sigma_{k|k}$ = covariância condicional.

[BACK](#) [NEXT](#)

16/51

ESTIMAÇÃO DE ESTADO PARA SISTEMAS LINEARES

2. Do curso de probabilidade: “Se x e y são dois vetores de v.a.’s gaussianas conjuntas com $x \sim N(\bar{x}, \Sigma_x)$, $y \sim N(\bar{y}, \Sigma_y)$ e $\text{cov}(x, y) = \Sigma_{xy}$ então

$$\hat{x} = E[x|y] = \bar{x} + \Sigma_{xy}\Sigma_y^{-1}(y - \bar{y})$$

$$\text{cov}(x|y) = \Sigma_x - \Sigma_{xy}\Sigma_y^{-1}\Sigma_{yx}$$

(vide exercício).

3. Notação $\tilde{x} = x - \hat{x}$. Temos que $E[\tilde{x}] = 0$, e então $\text{cov}(\tilde{x}|y) = \text{cov}(x|y)$. Além disso,

$$E[\tilde{x}(y - \bar{y})'] = E\{[(x - \bar{x}) - \Sigma_{xy}\Sigma_y^{-1}(y - \bar{y})](y - \bar{y})'\} =$$

$$\Sigma_{xy} - \Sigma_{xy}\Sigma_y^{-1}\Sigma_y = 0$$

Portanto $\tilde{x}$ e y são não correlatos, portanto independentes (pois gaussianos).

[BACK](#)
[NEXT](#)

17/51

FILTRO DE KALMAN

Notação:

$$\hat{x}_{k|k} = E[x_k | y^k] \quad (\text{melhor estimador de } x_k \text{ dada } y^k)$$

$$\hat{x}_{k+1|k} = E[x_{k+1} | y^k] \quad (\text{estimador de } x \text{ “um passo” a frente})$$

$$\tilde{x}_{k|k} = x_k - \hat{x}_{k|k} \quad (\text{erro de estimação de } x_k)$$

$$\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k} \quad (\text{erro de estimação “um passo” a frente})$$

$$\Sigma_{k|k} = \text{cov}(\tilde{x}_{k|k}) \quad (\text{matriz de covariância do erro de estimação})$$

$$\Sigma_{k+1|k} = \text{cov}(\tilde{x}_{k+1|k}) \quad (\text{covariância do erro “um passo” a frente})$$

[BACK](#)
[NEXT](#)

18/51

FILTRO DE KALMAN

PASSO 1 (atualização temporal do estado):

$$\begin{aligned}\hat{x}_{k+1|k} &= E[x_{k+1}|y^k] && \text{(média)} \\ &= E[A_k x_k + F_k e_k | y^k] = A_k \hat{x}_{k|k}\end{aligned}$$

Então

$$\begin{aligned}\tilde{x}_{k+1|k} &= x_{k+1} - \hat{x}_{k+1|k} \\ &= A_k x_k + F_k e_k - A_k \hat{x}_{k|k} = A_k \tilde{x}_{k|k} + F_k e_k\end{aligned}$$

e portanto

$$\Sigma_{k+1|k} = A_k \Sigma_{k|k} A_k' + F_k Q F_k' \quad \text{(covariância do erro)}$$

[BACK](#) [NEXT](#)

19/51

FILTRO DE KALMAN

PASSO 2 (atualização temporal da saída):

$$y_{k|k-1} = E[y_k | y^{k-1}] = E[C_k x_k + D_k v_k | y^{k-1}] = C_k \hat{x}_{k|k-1}$$

$$\begin{aligned}\tilde{y}_{k|k-1} &= y_k - y_{k|k-1} \\ &= C_k x_k + D_k v_k - C_k \hat{x}_{k|k-1} = C_k \tilde{x}_{k|k-1} + D_k v_k\end{aligned}$$

e portanto,

$$\text{cov}(\tilde{y}_{k|k-1}) = \Sigma_{k|k-1}^y = C_k \Sigma_{k|k-1} C_k' + D_k R D_k'$$

PASSO 3 (correlação cruzada):

$$\begin{aligned}E[x_k \tilde{y}_{k|k-1}'] &= E[x_k (C_k \tilde{x}_{k|k-1} + D_k v_k)'] \\ &= E[(\tilde{x}_{k|k-1} + \hat{x}_{k|k-1})(C_k \tilde{x}_{k|k-1} + D_k v_k)'] = \Sigma_{k|k-1} C_k' + 0\end{aligned}$$

[BACK](#) [NEXT](#)

20/51

FILTRO DE KALMAN

PASSO 4 (atualização da medida) A informação contida em (y^{k-1}, y_k) é a mesma que aquela em $(y^{k-1}, \tilde{y}_{k|k-1})$.

$$\begin{aligned}\hat{x}_{k|k} &= E[x_k | y^k] = E[x_k | y^{k-1}, y_k] = E[x_k | y^{k-1}, \tilde{y}_{k|k-1}] \\ &= \hat{x}_{k|k-1} + E[x_k \tilde{y}'_{k|k-1}] \left(\Sigma_{k|k-1}^y \right)^{-1} \tilde{y}_{k|k-1} \\ &= \hat{x}_{k|k-1} + \Sigma_{k|k-1} C_k' \left(C_k \Sigma_{k|k-1} C_k' + D_k R D_k' \right)^{-1} (y_k - C_k \hat{x}_{k|k-1})\end{aligned}\tag{5}$$

(Atualização de Medida $\leadsto$ média)

[BACK](#) [NEXT](#)

21/51

FILTRO DE KALMAN

Segue também de (5) que

$$\tilde{x}_{k|k} = \tilde{x}_{k|k-1} - \Sigma_{k|k-1} C_k' \left(\Sigma_{k|k-1}^y \right)^{-1} \tilde{y}_{k|k-1}$$

então

$$\text{cov}(\tilde{x}_{k|k}) = \text{cov}(\tilde{x}_{k|k-1}) - E[x_k \tilde{y}'_{k|k-1}] \left(\Sigma_{k|k-1}^y \right)^{-1} E[x_k \tilde{y}'_{k|k-1}]'$$

o que leva a:

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} C_k' \left(C_k \Sigma_{k|k-1} C_k' + D_k R D_k' \right)^{-1} C_k \Sigma_{k|k-1}$$

(Atualização de Medida $\leadsto$ covariância)

[BACK](#) [NEXT](#)

22/51

FILTRO DE KALMAN

Note que as eqs. de atualização de medida permitem expressar:

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} = A_k \hat{x}_{k|k-1} + A_k L_k (y_k - C_k \hat{x}_{k|k-1}), \quad (6)$$

com $\hat{x}_{0|-1} = \bar{x}_0$ e L_k definido por:

$$L_k = \Sigma_{k|k-1} C_k' \left(C_k \Sigma_{k|k-1} C_k' + D_k R D_k' \right)^{-1} \quad (7)$$

e

$$\begin{aligned} \Sigma_{k+1|k} &= A_k \Sigma_{k|k} A_k' + F_k Q F_k' \\ &= A_k \Sigma_{k|k-1} A_k' - A_k \Sigma_{k|k-1} C_k' \left(C_k \Sigma_{k|k-1} C_k' + D_k R D_k' \right)^{-1} C_k \Sigma_{k|k-1} A_k' + \\ &\quad + F_k Q F_k', \end{aligned} \quad (8)$$

com $\Sigma_{0|-1} = \Sigma_0$. Equações (6), (7) e (8) definem o Filtro de Kalman.

[BACK](#)
[NEXT](#)

23/51

FILTRO DE KALMAN: OBSERVAÇÕES

Obs1: $\tilde{y}_{k|k-1} = y_k - C_k \hat{x}_{k|k-1}$ é chamado de PROCESSO DE INOVAÇÃO, $\tilde{y}_{k|k-1}$ é independente de y^{k-1} .

Obs2: Se o que deseja é:

$\hat{x}_{k|k} \rightsquigarrow$ Problema de Filtragem

$\hat{x}_{k|k-1} \rightsquigarrow$ Problema de Predição

$\hat{x}_{k|k+n} \rightsquigarrow$ Problema de Interpolação ou Smoothing

[BACK](#)
[NEXT](#)

24/51

FILTRO DE KALMAN: OBSERVAÇÕES

Fazendo referência à solução conceitual de filtragem, para o FK temos:

- PREDIÇÃO UM ESTÁGIO A FRENTE, calculada como:

$$p(x_k|\mathbf{y}^{k-1}) = \int p(x_k|x_{k-1})p(x_{k-1}|\mathbf{y}^{k-1})dx_{k-1}$$

Está associada às distribuições normais:

$$\begin{aligned} p(x_{k-1}|\mathbf{y}^{k-1}), \quad x_{k-1} &\sim N(\hat{x}_{k-1|k-1}, \Sigma_{k-1|k-1}) \\ p(x_k|\mathbf{y}^{k-1}), \quad x_k &\sim N(\hat{x}_{k|k-1}, \Sigma_{k|k-1}) \end{aligned}$$

- ESTÁGIO DE ATUALIZAÇÃO, feito no instante que a medida y_k é obtida:

$$p(x_k|\mathbf{y}^k) = \frac{p(y_k|x_k)p(x_k|\mathbf{y}^{k-1})}{p(y_k|\mathbf{y}^{k-1})}$$

está associado à distribuição

$$p(x_k|\mathbf{y}^k), \quad x_k \sim N(\hat{x}_{k|k}, \Sigma_{k|k})$$

[BACK](#)
[NEXT](#)

25/51

FILTRO DE KALMAN: EXEMPLO

Baseado na rotina de demonstração do Matlab `kalmdemo`.

```
% Given the following discrete plant
A = [ 1.1269    -0.494    0.1129
      1         0         0
      0         1         0 ]
B = [ -0.3832
      0.5919
      0.5191 ]

% Design a Kalman filter to estimate the output y based on the
% noisy measurements yv[n] = C x[n] + v[n]
% A time-varying Kalman filter is able to perform well even
% when the noise covariance is not stationary. However for this
% demonstration, we will use stationary covariance, matrices Q and R.
```

[BACK](#)
[NEXT](#)

26/51

FILTRO DE KALMAN: EXEMPLO

```
% The time varying Kalman filter has the following update equations
%
% Time update:      x[n+1|n] = Ax[n|n] + Bu[n]
%
%                  P[n+1|n] = AP[n|n]A' + B*Q*B'
%
%
% Measurement update:
%                  x[n|n] = x[n|n-1] + M[n](yv[n] - Cx[n|n-1])
%                                  -1
%                  M[n] = P[n|n-1] C' (CP[n|n-1]C'+R)
%
%                  P[n|n] = (I-M[n]C) P[n|n-1] % function KALMAN.
% This function determines the optimal
% steady-state filter gain M based on the process noise
% covariance Q and the sensor noise covariance R.
% First specify the plant + noise model.
% CAUTION: set the sample time to -1 to mark plant
%          as discrete

sys = ss(A,B,C,D,-1);
```

[BACK](#)[NEXT](#)

27/51

FILTRO DE KALMAN: EXEMPLO

```
% Enter the process noise covariance:
Q = 1;
% Enter the sensor noise covariance:
R = 2;

% Generate a sinusoidal input vector (known)
t = [0:100]';
u = sin(t/5);
% Generate process noise and sensor noise vectors
randn('seed',0)
w = sqrt(Q)*randn(length(t),1);
v = sqrt(R)*randn(length(t),1);

% Now simulate the response using LSIM
% Generate the noisy plant response
y = lsim(sys,u+w); % w = process noise
yv = y + v;       % v = meas. noise
```

[BACK](#)[NEXT](#)

28/51

FILTRO DE KALMAN: EXEMPLO

```

% Now implement the filter recursions in a FOR loop:
P=B*Q*B';          % Initial error covariance
x=zeros(3,1);      % Initial condition on the state
ye = zeros(length(t),1);
ycov = zeros(length(t),1);

for i=1:length(t)
% Measurement update
  Mn = P*C'/(C*P*C'+R);
  x = x + Mn*(yv(i)-C*x); % x[n|n]
  P = (eye(3)-Mn*C)*P;

  ye(i) = C*x;
  errcov(i) = C*P*C';

% Time update
  x = A*x + B*u(i);      % x[n+1|n]
  P = A*P*A' + B*Q*B';
end

```

[BACK](#)[NEXT](#)

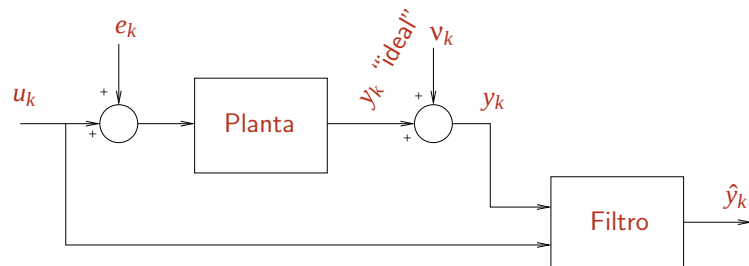
29/51

FILTRO DE KALMAN: EXEMPLO

```

% Let's see how this filter works by generating some data and
% comparing the filtered response with the true plant response.

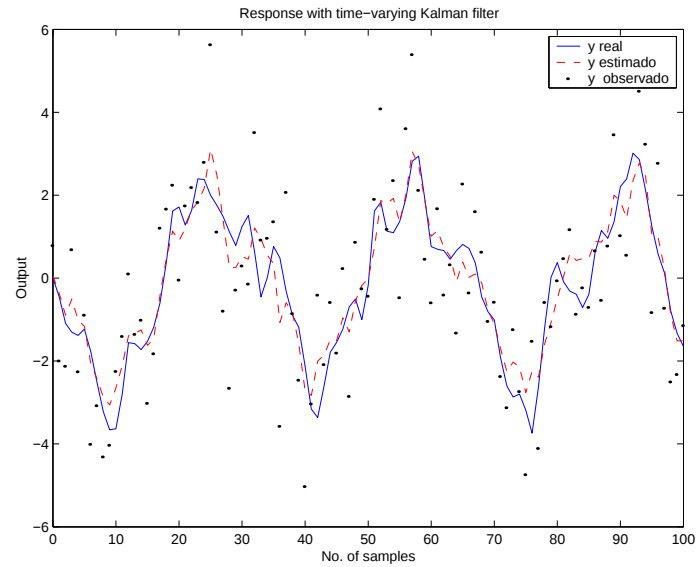
```

[BACK](#)[NEXT](#)

30/51

FILTRO DE KALMAN: EXEMPLO

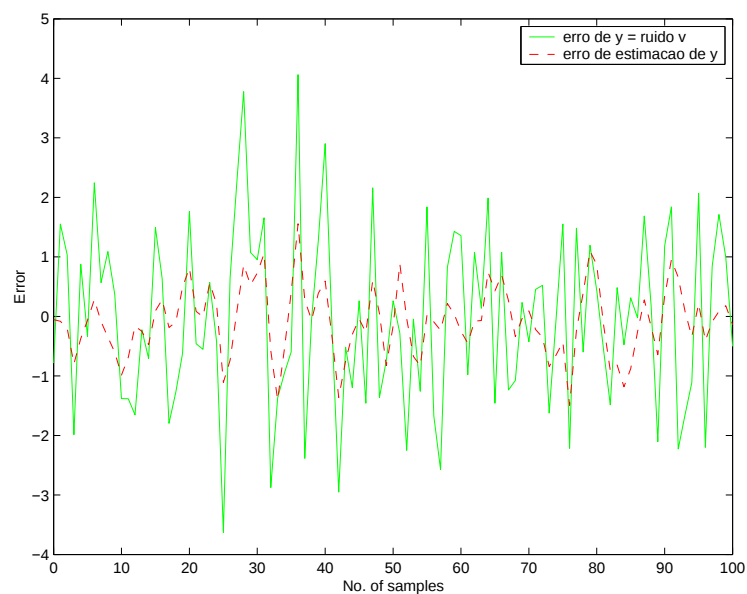
```
% Compare true response with filtered response
figure, plot(t,y,'b',t,ye,'r--',t,yv,'k.'),
xlabel('No. of samples'), ylabel('Output')
legend('y real','y estimado','y observado')
title('Response with time-varying Kalman filter')
```


[BACK](#)
[NEXT](#)

31/51

FILTRO DE KALMAN: EXEMPLO

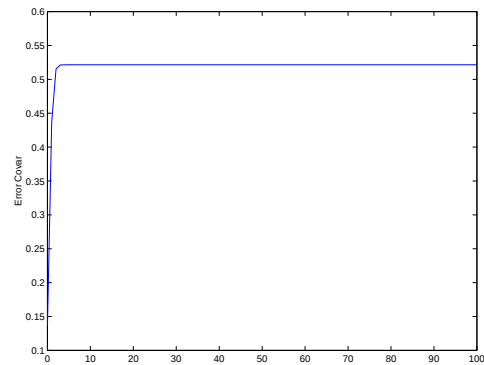
```
figure, plot(t,y-yv,'g',t,y-ye,'r--'),
legend('erro de y = ruido v','erro de estimacao de y')
xlabel('No. of samples'), ylabel('Error')
```


[BACK](#)
[NEXT](#)

32/51

FILTRO DE KALMAN: EXEMPLO

```
% The time varying filter also estimates the output covariance
% during the estimation. Let's plot it to see if our filter reached
% steady state (as we would expect with stationary input noise).
figure
plot(t,errcov), ylabel('Error Covar'),
```



```
% From the covariance plot we see that the output covariance did
% indeed reach a steady state in about 5 samples. From then on,
% our time varying filter has the same performance as the steady
% state version.
```

[BACK](#)
[NEXT](#)

33/51

FILTRO DE KALMAN: EXEMPLO

```
% Compare covariance errors
MeasErr = y-yv;
MeasErrCov = sum(MeasErr.*MeasErr)/length(MeasErr);
EstErr = y-ye;
EstErrCov = sum(EstErr.*EstErr)/length(EstErr);

% Covariance of error before filtering (measurement error)
MeasErrCov
>> MeasErrCov = 2.2277
% Covariance of error after filtering (estimation error)
EstErrCov
>> EstErrCov = 0.3394
```

[BACK](#)
[NEXT](#)

34/51

FILTRAGEM LINEAR

IDÉIA BÁSICA: aproximação em espaços de Hilbert apropriado

Seja $(\Omega, \mathcal{F}, P)$ um espaço de probabilidade. Para os nossos objetivos, considere o subconjunto $\mathcal{H}$ de $(\Omega, \mathcal{F}, P)$ contendo todas as v.a.'s quadrado-integráveis (o espaço ℓ_2 , contendo todas as seqüências numéricas infinitas $x(0), x(1), \dots$), isto é

$$E[X^2] < \infty \quad \left[\sum_{j=0}^{\infty} x(j)^2 < \infty \right]$$

O espaço $\mathcal{H}$ é um ESPAÇO LINEAR e podemos definir a NORMA de $X \in \mathcal{H}$ ($x \in \ell_2$) como

$$\|X\|_2 = E[X^2]^{1/2} \quad \left[\|x\|_2 = \left(\sum_{j=0}^{\infty} x(j)^2 \right)^{1/2} \right]$$

[BACK](#) [NEXT](#)

35/51

FILTRAGEM LINEAR: CONSTRUÇÃO DO ESPAÇO DE HILBERT

- A distância entre X e Y é

$$\|X - Y\|_2 = E[(X - Y)^2]^{1/2}$$

- Uma seqüência $X_n, n = 0, 1, \dots$ converge para X significa que

$$E[(X_n - X)^2] \rightarrow 0, \text{ quando } n \rightarrow \infty \quad (\text{noção: convergência na média quadrática}).$$

Além disso o espaço $\mathcal{H}$ é COMPLETO, isto é, todas as seqüências de Cauchy são convergentes.

- Definindo o produto interno

$$\langle X, Y \rangle = E[XY]$$

o espaço de v.a.'s quadrado integráveis é um ESPAÇO DE HILBERT.

[BACK](#) [NEXT](#)

36/51

APROXIMAÇÃO EM ESPAÇOS DE HILBERT

Já vimos anteriormente que neste espaço tem-se a noção de ORTOGONALIDADE ' $X \perp Y$ ':

$$X, Y \in \mathcal{H} \text{ e } X \perp Y \Leftrightarrow \langle X, Y \rangle = 0$$

Um elemento $X \in \mathcal{H}$ e um subespaço fechado $S \subset \mathcal{H}$ são ortogonais se

$$X \perp Y \text{ para todo } Y \in S \quad (X \perp S)$$

TEOREMA Para $\mathcal{H}$ e $S \subset \mathcal{H}$ subespaço fechado. Então $X \in \mathcal{H}$ pode ser decomposto unicamente como

$$X = W + V$$

com $W \in S$ e $V \in S^\perp$ determinados unicamente por S . Em outras palavras

$$\mathcal{H} = S \oplus S^\perp$$

[BACK](#) [NEXT](#)

37/51

APROXIMAÇÃO EM ESPAÇOS DE HILBERT: PROJEÇÃO

- Noção de PROJEÇÃO ORTOGONAL: a projeção de X sobre Y é $\alpha = \langle X, Y \rangle$

Generalizando, a PROJEÇÃO ORTOGONAL de um $X \in \mathcal{H}$ sobre $S \subset \mathcal{H}$ é um único $Y \in S$ tal que $(X - Y) \perp S$.

- O elemento Y é a PROJEÇÃO DE X EM S e representa-se por $\pi(X|S)$.

DEFINIÇÃO Um subespaço fechado S representado por

$$\alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n$$

para $X_1, X_2, \dots, X_n \in \mathcal{H}$ e escalares $\alpha_1, \alpha_2, \dots, \alpha_n$ quaisquer é o RANGE LINEAR DE $X_1, X_2, \dots, X_n$, representado por

$$\mathcal{L}\{X_1, X_2, \dots, X_n\}$$

Assim $\pi(X|S)$ corresponde a uma escolha de escalares tal que

$$(X - \alpha_1 X_1 + \alpha_2 X_2 + \cdots + \alpha_n X_n) \perp S$$

[BACK](#) [NEXT](#)

38/51

APROXIMAÇÃO EM ESPAÇOS DE HILBERT: BASE

Uma coleção infinita $X_1, X_2, \dots$ de elementos em $\mathcal{H}$ (ou qualquer outro espaço de Hilbert) tal que $\langle X_i, X_j \rangle = 0, \forall i \neq j$ forma uma BASE em $\mathcal{H}$, no sentido que qualquer elemento $X \in \mathcal{H}$ pode ser escrito em termos de projeção sobre a tal base:

$$X = \sum_{j=1}^{\infty} \frac{\langle X, X_j \rangle}{\langle X_j, X_j \rangle} X_j = \sum_{j=1}^{\infty} \langle X, X_j \rangle \frac{X_j}{\|X_j\|_2^2}$$

Para uma BASE ORTONORMAL, temos adicionalmente que $\|X_j\|_2^2 = 1$ e a representação fica escrita como

$$X = \sum_{j=1}^{\infty} \alpha_j X_j, \quad \alpha_j = \langle X, X_j \rangle$$

APROXIMAÇÃO DE ORDEM n é a projeção, obtida truncando-se a representação

$$X \approx \sum_{j=1}^n \alpha_j X_j = \pi(X|S)$$

onde $S = \{Y \in \mathcal{H} : Y = \sum_{j=1}^n \alpha_j X_j\}$.

[BACK](#)
[NEXT](#)

39/51

PROPRIEDADES DAS PROJEÇÕES

1. Linearidade $\pi(X + Y|S) = \pi(X|S) + \pi(Y|S)$

2. Minimiza a norma induzida do erro:

$$\|X - \pi(X|S)\|_2 \leq \|X - Y\|_2, \text{ para todo } Y \in S$$

3. Erro ortogonal

$$(X - \pi(X|S)) \perp S$$

4. Igualdade do produto interno

$$\langle X, Y \rangle = \langle \pi(X|S), Y \rangle, \text{ para todo } Y \in S$$

[BACK](#)
[NEXT](#)

40/51

CASO DE VETORES DE V.A.'S

■ Espaço $\mathcal{H}^d$ dos vetores de v.a.'s quadrado integráveis definidos num espaço de probabilidade $(\Omega, \mathcal{F}, P)$, as vezes também denotado por $\mathcal{H}^d = L_2((\Omega, \mathcal{F}, P); \mathbb{R}^n)$. Aqui x e y são vetores de variáveis aleatórias:

- Produto Interno em $\mathcal{H}^d$:

$$\begin{aligned} \langle x, y \rangle &= E[(x - \bar{x})(y - \bar{y})'] = \sum_{i=1}^n \text{cov}(X_i, Y_i) \\ &= \sum_{i=1}^n \langle X_i - \bar{X}_i, Y_i - \bar{Y}_i \rangle \end{aligned}$$

esta última noção em $\mathcal{H}$.

- Norma em $\mathcal{H}^d$: $\|x\| = \langle x, x \rangle^{1/2} = \left(\sum_{i=1}^n \|X_i - \bar{X}_i\|_2^2 \right)^{1/2}$ é uma medida do desvio padrão do vetor

[BACK](#) [NEXT](#)

41/51

CASO DE VETORES: PRODUTO INTERNO E EXTERNO

- Produto Externo $[x, y] = E[(x - \bar{x})(y - \bar{y})']$ é a matriz de covariância

Note que

$$[x, y] = \begin{bmatrix} \text{cov}(X_1, Y_1) & \text{cov}(X_1, Y_2) & \cdots & \text{cov}(X_1, Y_n) \\ \text{cov}(X_2, Y_1) & \text{cov}(X_2, Y_2) & \cdots & \text{cov}(X_2, Y_n) \\ \vdots & \vdots & \ddots & \vdots \\ \text{cov}(X_n, Y_1) & \text{cov}(X_n, Y_2) & \cdots & \text{cov}(X_n, Y_n) \end{bmatrix} = \begin{bmatrix} \langle X_1, Y_1 \rangle & \langle X_1, Y_2 \rangle & \cdots & \langle X_1, Y_n \rangle \\ \langle X_2, Y_1 \rangle & \langle X_2, Y_2 \rangle & \cdots & \langle X_2, Y_n \rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle X_n, Y_1 \rangle & \langle X_n, Y_2 \rangle & \cdots & \langle X_n, Y_n \rangle \end{bmatrix}$$

DEFINIÇÃO $x, y \in \mathcal{H}^d$ são ortogonais se $[x, y] = 0$.

Note que em $\mathcal{H}^d$,

$$\langle x, y \rangle = E[(x - \bar{x})'(y - \bar{y})] = E[\text{traço}\{(x - \bar{x})(y - \bar{y})'\}] = \text{traço}\{[x, y]\}$$

e a noção de ortogonalidade aqui é mais forte que exigir que $\langle x, y \rangle = 0$.

[BACK](#) [NEXT](#)

42/51

CASO DE VETORES: PRODUTO EXTERNO

Dados $x, y, z \in \mathcal{H}^d$ e matrizes A, B de dimensões apropriadas

- $[x + y, z] = [x, z] + [y, z]$
- $[Ax, By] = A[x, y]B'$
- x, y independentes $\rightarrow [x, y] = 0$

[BACK](#)
[NEXT](#)

43/51

O PROBLEMA DE FILTRAGEM LINEAR

É visto como um problema de projeção em $\mathcal{H}^d$

- Estimativa $\hat{x}_{k|k}$ como função do range linear das observações $\mathcal{L}\{y_0, y_1, \dots, y_k\}$
- Ruídos e_k e v_k gaussianos não correlatados. Iremos expressar

$$E[x_k | y^k] = \pi(x_k | \mathcal{L}\{y_0, y_1, \dots, y_k\})$$

- Estratégia: encontrar uma base ortonormal para $\mathcal{L}\{y_0, y_1, \dots, y_k\}$ utilizando ortogonalização de Gram-Schmidt em $\mathcal{H}^d$.

[BACK](#)
[NEXT](#)

44/51

ORTOGONALIZAÇÃO DA SEQUÊNCIA DE OBSERVAÇÃO

A k -ésima observação pode ser representada por

$$\begin{aligned} y_k &= \pi(y_k | \mathcal{L}\{y_0, \dots, y_{k-1}\}) + \pi(y_k | \mathcal{L}\{y_0, \dots, y_{k-1}\}^\perp) \\ &= \hat{y}_{k|k-1} + \tilde{y}_{k|k-1} \end{aligned}$$

■ Base ortogonal para $\mathcal{L}\{y_0, \dots, y_{k-1}\}$

1. $\hat{y}_0 = y_0$. Define-se $\tilde{y}_1 = y_1 - \hat{y}_1$ e

$$E[\tilde{y}_1 \tilde{y}_0'] = [(y_1 - \hat{y}_1), (y_0 - \hat{y}_0)] = [(y_1 - \hat{y}_1), 0] = 0 \quad \text{e } \tilde{y}_1 \perp \tilde{y}_0$$

2. Verifica-se que

$$E[\tilde{y}_2 \tilde{y}_1' | \tilde{y}_0, \tilde{y}_1] = [(y_2 - \hat{y}_2), (y_1 - \hat{y}_1)] = E[(y_2 - \bar{y}_2 - \Sigma_{21} \Sigma_1^{-1} (y_1 - \bar{y}_1)) \tilde{y}_1'] = 0$$

portanto, $\tilde{y}_2 \perp \tilde{y}_1$

BACK NEXT

45/51

ORTOGONALIZAÇÃO DA SEQUÊNCIA DE OBSERVAÇÃO

3. Podemos concluir por indução que

$$E[\tilde{y}_3 \tilde{y}_2' | \tilde{y}_0, \dots, \tilde{y}_2] = [\tilde{y}_3, \tilde{y}_2] = 0 \implies \tilde{y}_3 \perp \mathcal{L}\{\tilde{y}_0, \tilde{y}_1, \tilde{y}_2\}$$

e geral

$$E[\tilde{y}_k \tilde{y}_{k-1}' | \mathcal{L}\{y_0, \dots, y_{k-1}\}] = [\tilde{y}_k, \tilde{y}_{k-1}] = 0 \implies \tilde{y}_k \perp \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\}$$

Adota-se $\tilde{y}_0, \tilde{y}_1, \dots$ (PROCESSO DE INOVAÇÃO) como a base de referência.

4. Pela ortogonalidade temos que

$$\mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_k\} = \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\} \oplus \mathcal{L}\{\tilde{y}_k\}$$

e

$$\begin{aligned} \pi(x_k | \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_k\}) &= \pi(x_k | \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\}) + \pi(x_k | \mathcal{L}\{\tilde{y}_k\}) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + \pi(x_k | \mathcal{L}\{\tilde{y}_k\}) \end{aligned} \tag{9}$$

é o estimador a posteriori

BACK NEXT

46/51

ESTIMADOR A PRIORI

$$\begin{aligned}
\hat{x}_{k|k-1} &= \pi(x_k | \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\}) \\
&= \pi(Ax_{k-1} + Fe_{k-1} | \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\}) \\
&= Ax_{k-1|k-1} + \pi(Fe_k | \mathcal{L}\{\tilde{y}_0, \dots, \tilde{y}_{k-1}\}) = Ax_{k-1|k-1}
\end{aligned} \tag{10}$$

PROJEÇÃO SOBRE A INOVAÇÃO

$$\begin{aligned}
\pi(x_k | \mathcal{L}\{\tilde{y}_k\}) &= \pi(\hat{x}_{k|k-1} + \tilde{x}_k | \mathcal{L}\{\tilde{y}_k\}) = \pi(\tilde{x}_k | \mathcal{L}\{\tilde{y}_k\}) \quad (\text{pois } \hat{x}_{k|k-1} \perp \tilde{y}_k) \\
&= [\tilde{x}_k, \tilde{y}_k] [\tilde{y}_k, \tilde{y}_k]^{-1} \tilde{y}_k \\
&= \mathcal{K}_k \tilde{y}_k
\end{aligned} \tag{11}$$

na qual $\mathcal{K}_k$ é o operador projeção de $\mathcal{L}\{\tilde{y}_k\}$ em $\mathcal{L}\{\tilde{x}_k\}$

[BACK](#) [NEXT](#)

47/51

DETERMINAÇÃO DA PROJEÇÃO

O filtro fica reduzido ao cálculo de duas matrizes de covariância $[\tilde{x}_k, \tilde{y}_k]$ e $[\tilde{y}_k, \tilde{y}_k]$

- Definindo $\Sigma_{k|k} = [\tilde{x}_k, \tilde{x}_k] = E[\tilde{x}_k \tilde{x}_k']$, calcula-se

$$\begin{aligned}
[x_k, \tilde{y}_k] &= [\hat{x}_k + \tilde{x}_k, \tilde{y}_k] = [\hat{x}_k, \tilde{y}_k] + [\tilde{x}_k, \tilde{y}_k] \\
&= 0 + [\tilde{x}_k, C\tilde{x}_k + Dv_k] \\
&= [\tilde{x}_k, C\tilde{x}_k] + [\tilde{x}_k, Dv_k] \\
&= \Sigma_{k|k} C' + [x_k - \hat{x}_k, v_k] D' \\
&= \Sigma_{k|k} C' + ([x_k, v_k] - [\hat{x}_k, v_k]) D' \\
&= \Sigma_{k|k} C' + [Ax_{k-1} + Fe_{k-1}, v_k] - 0' = \Sigma_{k|k} C'
\end{aligned}$$

[BACK](#) [NEXT](#)

48/51

DETERMINAÇÃO DA PROJEÇÃO

e

$$\begin{aligned} [\tilde{y}_k, \tilde{y}_k] &= [C\tilde{x}_k + Dv_k, C\tilde{x}_k + Dv_k] \\ &= C\Sigma_{k|k}C' + DRD' \end{aligned}$$

Portanto,

$$\mathcal{K}_k = \Sigma_{k|k}C'(C\Sigma_{k|k}C' + DRD')^{-1} \quad (12)$$

e

$$\hat{x}_{k|k} = \hat{x}_{k|k-1} + \Sigma_{k|k}C'(C\Sigma_{k|k}C' + DRD')^{-1}\tilde{y}_k$$

[BACK](#) [NEXT](#)

49/51

PROPAGAÇÃO DAS ESTIMATIVAS

• covariância

$$\begin{aligned} \tilde{x}_{k|k} &= x_k - \hat{x}_{k|k} \\ &= x_k - (\hat{x}_{k|k-1} + \mathcal{K}_k\tilde{y}_k) \\ &= \tilde{x}_{k|k-1} - \mathcal{K}_k(C\tilde{x}_k + Dv_k) \\ &= (I - \mathcal{K}_kC)\tilde{x}_k - \mathcal{K}_kDv_k \end{aligned}$$

e

$$\begin{aligned} \Sigma_{k|k} &= (I - \mathcal{K}_kC)\Sigma_{k|k-1}(I - \mathcal{K}_kC)' + \mathcal{K}_kDRD'\mathcal{K}_k' \\ &= (I - \mathcal{K}_kC)\Sigma_{k|k-1} \end{aligned} \quad (13)$$

tendo (12) em vista, mostre isso!

[BACK](#) [NEXT](#)

50/51

PROPAGAÇÃO DAS ESTIMATIVAS

- propagação da covariância a posteriori da apriori

$$\begin{aligned}
 \Sigma_{k|k-1} &= [\tilde{x}_{k|k-1}, \tilde{x}_{k|k-1}] \\
 &= [Ax_{k-1} + Fe_{k-1} - A\hat{x}_{k-1|k-1}, Ax_{k-1} + Fe_{k-1} - A\hat{x}_{k-1|k-1}] \\
 &= [A\tilde{x}_{k-1|k-1} + Fe_{k-1}, A\tilde{x}_{k-1|k-1} + Fe_{k-1}] \\
 &= A\Sigma_{k-1|k-1}A' + FQF'
 \end{aligned} \tag{14}$$

Combinando (13) e (14) temos a expressão

$$\Sigma_{k|k} = (I - \mathcal{K}_k)(A\Sigma_{k-1|k-1}A' + FQF')$$

e finalmente,

$$\begin{cases} x_{k|k} = Ax_{k-1|k-1} + \mathcal{K}_k\tilde{y}_k \\ \mathcal{K}_k = (A\Sigma_{k-1|k-1}A' + FQF')C'[C(A\Sigma_{k-1|k-1}A' + FQF')C' + DRD']^{-1} \end{cases}$$