



Estimação e Filtragem Estocástica
Curso de Pós-Graduação em Engenharia Elétrica
Departamento de Engenharia Elétrica
Universidade de Brasília

FILTRAGEM ESTOCÁSTICA LINEAR

Geovany A. Borges
gaborges@ene.unb.br

Modelos estocásticos

- Denominação conforme o estado e o tempo:

Estado	Tempo	Classificação
contínuo	contínuo	modelo estocástico em tempo contínuo
contínuo	discreto	modelo estocástico em tempo discreto
discreto	contínuo	cadeia estocástica em tempo contínuo
discreto	discreto	cadeia estocástica em tempo discreto

com $t_k = kT$, sendo T o passo de discretização do tempo.

- Principal interesse com os modelos no contexto da filtragem estocástica: determinar $p(\mathbf{x})$, ou seja, a função densidade de probabilidade do estado.

Modelos estocásticos

- Componentes usadas na modelagem de sistemas:
 - Entradas: vetor $\mathbf{u}$, de dimensão p
 - Saídas: vetor $\mathbf{y}$, de dimensão m
 - Estado (ou variáveis internas não observadas): vetor $\mathbf{x}$, de dimensão n
- Fontes de incerteza:
 - Conhecimento limitado acerca do modelo
 - Conhecimento limitado acerca das entradas de perturbação/interferência
 - Interferência agindo nas entradas ou saídas
- A que se destina um modelo:
 - Representar algum sistema físico (caso mais comum)
 - Representar alguma relação matemática para explicar um fenômeno

Prof. Geovany A. Borges - Estimação e Filtragem Estocástica - 2008.1

1

Modelos estocásticos (tempo contínuo)

- Modelo estocástico em tempo contínuo:

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, \mathbf{w}_t, t), \quad t \geq t_0 \quad (1)$$

com $\mathbf{x}_t \in \mathbb{R}^n$, $\mathbf{f}(\cdot, \cdot, t)$ é uma função vetorial não-linear e $\mathbf{w}_t$ é um processo estocástico de distribuição conhecida.

- Interesse em $p(\mathbf{x}_t)$

- Caso especial:

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{f}(\mathbf{x}_t, t) + \mathbf{G}(\mathbf{x}_t, t)\mathbf{w}_t, \quad t \geq t_0 \quad (2)$$

com $\mathbf{f}(\cdot, t)$ sendo uma função vetorial não-linear, $\mathbf{G}(\cdot, t)$ sendo uma função matricial e $E\{\mathbf{x}_{t_0}\mathbf{w}_t^T\} = \mathbf{0}$. Em geral,

$$\mathbf{w}_t \sim \frac{d\beta_t}{dt} \quad (3)$$

com β_t sendo um vetor de movimentos Brownianos independentes e $E\{\mathbf{w}_t\} = \mathbf{0}$, $E\{\mathbf{w}_t \mathbf{w}_\tau^T\} = \mathbf{Q}(t)\delta(t - \tau)$, $\mathbf{Q} > 0$.

■ Solução de (2):

$$\mathbf{x}_t = \mathbf{x}_{t_0} + \int_{t_0}^t [\mathbf{f}(\mathbf{x}_t, t) + \mathbf{G}(\mathbf{x}_t, t)\mathbf{w}_t] dt \quad (4)$$

$$= \mathbf{x}_{t_0} + \int_{t_0}^t \mathbf{f}(\mathbf{x}_t, t) dt + \mathbf{I}(t) \quad (5)$$

com

$$\mathbf{I}(t) = \int_{t_0}^t \mathbf{G}(\mathbf{x}_t, t) d\beta_t \quad (6)$$

sendo a integral estocástica de Itô.

Modelos estocásticos lineares (tempo contínuo)

■ Modelo estocástico linear em tempo contínuo:

$$\frac{d\mathbf{x}_t}{dt} = \mathbf{F}(t)\mathbf{x}_t + \mathbf{G}(t)\mathbf{w}_t \quad (8)$$

com $\mathbf{x}_{t_0} \sim \mathcal{N}(\bar{\mathbf{x}}_{t_0}, \mathbf{P}_{t_0})$ independente de $\mathbf{w}_t$. Este é um processo Gaussiano (distribuição normal) e Markoviano, pois

$$p(\mathbf{x}_t | \mathbf{x}_\tau, \tau \leq t^* < t) = p(\mathbf{x}_t | \mathbf{x}_{t^*}, t^* < t)$$

uma vez que, dado o modelo (8) e $\mathbf{x}_{t^*}$, por integração é possível chegar a $\mathbf{x}_t$. Assim, $\mathbf{x}_t$ é também chamado de processo de Gauss-Markov.

■ $p(\mathbf{x}_t)$ é normal, sendo caracterizado pela sua média e matriz de covariâncias:

$$\bar{\mathbf{x}}_t = E\{\mathbf{x}_t\}, \quad (9)$$

$$\mathbf{P}_t = E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)(\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\}. \quad (10)$$

■ Evolução de $p(\mathbf{x}_t)$ por meio da equação de Kolmogorov direta:

$$\frac{\partial p(\mathbf{x}_t)}{\partial t} = - \sum_{i=1}^n \frac{\partial(p(\mathbf{x}_t) \cdot [\mathbf{f}(\mathbf{x}_t, t)]_i)}{\partial \mathbf{x}_{i,t}} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2(p(\mathbf{x}_t) \cdot [\mathbf{G}\mathbf{Q}\mathbf{G}^T]_{ij})}{\partial \mathbf{x}_{i,t} \partial \mathbf{x}_{j,t}} \quad (7)$$

com $[\cdot]_{ij}$ sendo o operador correspondendo ao elemento de posição (i, j) .

A equação de Kolmogorov direta se torna

$$\frac{\partial p(\mathbf{x}_t)}{\partial t} = - \sum_{i=1}^n \frac{\partial(p(\mathbf{x}_t) \cdot [\mathbf{F}(t)\mathbf{x}_t]_i)}{\partial \mathbf{x}_{i,t}} + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2(p(\mathbf{x}_t) \cdot [\mathbf{G}\mathbf{Q}\mathbf{G}^T]_{ij})}{\partial \mathbf{x}_{i,t} \partial \mathbf{x}_{j,t}}$$

Pode-se mostrar que

$$\frac{\partial p(\mathbf{x}_t)}{\partial t} = -p(\mathbf{x}_t) \text{tr}(\mathbf{F}(t)) - \frac{\partial p(\mathbf{x}_t)}{\partial \mathbf{x}_t} \mathbf{F}(t) \mathbf{x}_t + \frac{1}{2} \text{tr} \left(\mathbf{G}\mathbf{Q}\mathbf{G}^T \frac{\partial^2(p(\mathbf{x}_t))}{\partial \mathbf{x}_t \partial \mathbf{x}_t} \right) \quad (11)$$

A evolução da média é dada por

$$\frac{d\bar{\mathbf{x}}_t}{dt} = \frac{dE\{\mathbf{x}_t\}}{dt} = E\left\{\frac{d\mathbf{x}_t}{dt}\right\} \quad (12)$$

$$= E\{\mathbf{F}(t)\mathbf{x}_t + \mathbf{G}(t)\mathbf{w}_t\} \quad (13)$$

$$= \mathbf{F}(t)\bar{\mathbf{x}}_t \quad (14)$$

pois $E\{\mathbf{w}_t\} = \mathbf{0}$. A evolução de $\mathbf{P}_t$ segue abaixo:

$$\begin{aligned}\frac{d\mathbf{P}_t}{dt} &= \frac{d}{dt}E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)(\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\} \\ &= E\left\{\left[\frac{d}{dt}(\mathbf{x}_t - \bar{\mathbf{x}}_t)\right](\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\right\} + E\left\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)\left[\frac{d}{dt}(\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\right]\right\}\end{aligned}\quad (15)$$

Sendo

$$\frac{d}{dt}(\mathbf{x}_t - \bar{\mathbf{x}}_t) = \mathbf{F}(t)\mathbf{x}_t + \mathbf{G}(t)\mathbf{w}_t - \mathbf{F}(t)\bar{\mathbf{x}}_t = \mathbf{F}(t)(\mathbf{x}_t - \bar{\mathbf{x}}_t) + \mathbf{G}(t)\mathbf{w}_t \quad (17)$$

então

$$\frac{d\mathbf{P}_t}{dt} = E\{[\mathbf{F}(t)(\mathbf{x}_t - \bar{\mathbf{x}}_t) + \mathbf{G}(t)\mathbf{w}_t](\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\} \quad (18)$$

$$+ E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)[\mathbf{F}(t)(\mathbf{x}_t - \bar{\mathbf{x}}_t) + \mathbf{G}(t)\mathbf{w}_t]^T\} \quad (19)$$

$$= \mathbf{F}(t)\mathbf{P}_t + E\{\mathbf{G}(t)\mathbf{w}_t(\mathbf{x}_t - \bar{\mathbf{x}}_t)^T\} + \mathbf{P}_t^T\mathbf{F}^T(t) \quad (20)$$

$$+ E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)\mathbf{w}_t^T\mathbf{G}^T(t)\} \quad (21)$$

como $E\{\mathbf{w}_t\mathbf{w}_\tau^T\} = \mathbf{Q}(t)\delta(t - \tau)$,

$$E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)\mathbf{w}_t^T\mathbf{G}^T(t)\} = \int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{Q}(t)\delta(t - \tau)\mathbf{G}^T(t)d\tau \quad (27)$$

$$= \Phi(t, t)\mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) \quad (28)$$

$$= \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t) \quad (29)$$

e então

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{F}(t)\mathbf{P}_t + \mathbf{P}_t^T\mathbf{F}^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t), \quad (30)$$

que é a chamada Equação Diferencial de Riccati (EDR). Observe que $\mathbf{P}_{t_0}$ é dado como condição inicial.

A solução $\mathbf{x}_t$ é dada por

$$\mathbf{x}_t = \Phi(t, t_0)\mathbf{x}_{t_0} + \int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{w}_\tau d\tau \quad (22)$$

com $\Phi(t, t_0)$ sendo a matriz de transição de estados do sistema ($\Phi(t, t) = \mathbf{I}$). Assim

$$\bar{\mathbf{x}}_t = E\{\mathbf{x}_t\} = \Phi(t, t_0)\bar{\mathbf{x}}_{t_0} \quad (23)$$

e

$$(\mathbf{x}_t - \bar{\mathbf{x}}_t) = \Phi(t, t_0)(\mathbf{x}_{t_0} - \bar{\mathbf{x}}_{t_0}) + \int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{w}_\tau d\tau \quad (24)$$

Desta forma, como $\mathbf{w}_t$ e $\mathbf{x}_{t_0}$ são independentes,

$$E\{(\mathbf{x}_t - \bar{\mathbf{x}}_t)\mathbf{w}_t^T\mathbf{G}^T(t)\} = E\left\{\left[\int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)\mathbf{w}_\tau d\tau\right]\mathbf{w}_t^T\mathbf{G}^T(t)\right\} \quad (25)$$

$$= \int_{t_0}^t \Phi(t, \tau)\mathbf{G}(\tau)E\{\mathbf{w}_\tau\mathbf{w}_t^T\}\mathbf{G}^T(t)d\tau \quad (26)$$

■ Observa-se que existe um certo grau de dificuldade na determinação de $p(\mathbf{x}_t)$, pois envolve a resolução das seguintes equações diferenciais:

$$\frac{d\bar{\mathbf{x}}_t}{dt} = \mathbf{F}(t)\bar{\mathbf{x}}_t$$

$$\frac{d\mathbf{P}_t}{dt} = \mathbf{F}(t)\mathbf{P}_t + \mathbf{P}_t^T\mathbf{F}^T(t) + \mathbf{G}(t)\mathbf{Q}(t)\mathbf{G}^T(t)$$

de forma a obter

$$p(\mathbf{x}_t) = \frac{1}{(2\pi)^{n/2} \det(\mathbf{P}_t)^{1/2}} \exp\left\{-\frac{1}{2}(\mathbf{x}_t - \bar{\mathbf{x}}_t)^T \mathbf{P}_t^{-1}(\mathbf{x}_t - \bar{\mathbf{x}}_t)\right\}$$

Modelos estocásticos lineares (tempo discreto)

- Modelo estocástico em tempo discreto:

$$\mathbf{x}_{k+1} = \mathbf{f}(\mathbf{x}_k, \mathbf{w}_{k+1}, t_k), \quad k = 0, 1, \dots \quad (31)$$

com $\mathbf{x}_k \in \mathbb{R}^n$, $\mathbf{f}(\cdot, \cdot, t)$ é uma função vetorial não-linear, $\mathbf{w}_k$ é um PE de distribuição conhecida e, em geral, $t_k = kT$ com T sendo o período de amostragem.

- Interesse em $p(\mathbf{x}_k)$: sendo $\mathbf{w}_k$ uma sequência Gaussiana branca e independente de $\mathbf{x}_0$, então

$$p(\mathbf{w}_{k+1} | \mathbf{w}_k, \mathbf{w}_{k-1}, \dots, \mathbf{w}_0) = p(\mathbf{w}_{k+1}) \quad (32)$$

e

$$p(\mathbf{x}_{k+1} | \mathbf{x}_k, \mathbf{x}_{k-1}, \dots, \mathbf{x}_0) = p(\mathbf{x}_{k+1} | \mathbf{x}_k) \quad (33)$$

$$= p_{\mathbf{w}_{k+1}}(\mathbf{w}_{k+1}^*) \cdot |\det \mathbf{J}|^{-1} \quad (34)$$

Modelos estocásticos lineares (tempo discreto)

- Caso linear: sendo

$$\mathbf{f}(\mathbf{x}_k, \mathbf{w}_{k+1}, t_k) = \mathbf{A}_k \mathbf{x}_k + \mathbf{\Gamma}_{k+1} \mathbf{w}_{k+1} \quad (37)$$

com $\mathbf{w}_{k+1} \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_{k+1})$ e $\mathbf{x}_0 \sim \mathcal{N}(\bar{\mathbf{x}}_0, \mathbf{P}_0)$ pode-se mostrar que é um processo de Gauss-Markov tal que

$$\mathbf{x}_k \sim \mathcal{N}(\bar{\mathbf{x}}_k, \mathbf{P}_k) \quad (38)$$

com parâmetros

$$\bar{\mathbf{x}}_k = \mathbf{A}_{k-1} \bar{\mathbf{x}}_{k-1} \quad (39)$$

$$\mathbf{P}_k = \mathbf{A}_{k-1} \mathbf{P}_{k-1} \mathbf{A}_{k-1}^T + \mathbf{\Gamma}_k \mathbf{Q}_k \mathbf{\Gamma}_k^T \quad (40)$$

com

$$\mathbf{w}_{k+1}^* = \mathbf{f}^{-1}(\mathbf{x}_k, \mathbf{x}_{k+1}, t_k) \quad (35)$$

$$\mathbf{J} = \frac{\partial \mathbf{f}(\mathbf{x}_k, \mathbf{w}_{k+1}^*, t_k)}{\partial \mathbf{w}_{k+1}^*} \quad (36)$$

se a função vetorial $\mathbf{f}$ for inversível e existir apenas uma solução única $\mathbf{w}_{k+1}^*$.

Modelos estocásticos lineares (tempo discreto)

Exemplo 1: *SDED Linear com entrada exógena*

O que dizer do processo

$$\mathbf{x}_{k+1} = \mathbf{A}_k \mathbf{x}_k + \mathbf{B}_k \mathbf{u}_k + \mathbf{\Gamma}_{k+1} \mathbf{w}_{k+1} \quad (41)$$

com $\mathbf{u}_k$ sendo uma entrada externa descorrelacionada de $\mathbf{w}_{k+1}$ e $\mathbf{x}_k$?

Modelos estocásticos lineares (tempo discreto)

Exemplo 2: Modelamento de um foguete (simulação *sded_foguete*)

Sendo um foguete com empuxo vertical dado pela força F_k e massa M :

$$\mathbf{z}_k = \mathbf{A}\mathbf{z}_{k-1} + \mathbf{B}\mathbf{u}_k + \mathbf{w}_k \quad (42)$$

com

$$\mathbf{z}_k = \begin{pmatrix} y_k \\ \dot{y}_k \end{pmatrix}, \quad \mathbf{A} = \begin{pmatrix} 1 & T \\ 0 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 0 \\ T \end{pmatrix}, \quad \mathbf{u}_k = \frac{F_k}{M} - g$$

e g sendo a aceleração das gravidade, T sendo o período de amostragem, $\dot{y}_k \triangleq dy_k/dt_k$ e

$$\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k) \quad (43)$$

$$\mathbf{Q}_k = \begin{pmatrix} \sigma_{w_1}^2 & 0 \\ 0 & \sigma_{w_2}^2 \end{pmatrix} \frac{1}{20} t_k \quad (44)$$

■ Esse tipo de modelo se aplica a sistemas que mudam de regime de operação ou de estrutura conforme a ocorrência de algum evento que dita a evolução de θ_k .

■ É comum modelar θ_k como uma cadeia de Markov, em que a distribuição de probabilidade é dada por

$$\boldsymbol{\pi}_k = \begin{pmatrix} \pi_{1,k} & \pi_{2,k} & \cdots & \pi_{S,k} \end{pmatrix}^T$$

com $\pi_{i,k} = \Pr\{\theta_k = s_i\}$.

■ A evolução de $\boldsymbol{\pi}_k$ é ditada por um sistema dinâmico autônomo. Para tanto, define-se a matriz de transição de probabilidades

$$\mathbf{P}_k = \{p_{ij,k}\}, i, j = 1, \dots, S$$

tal que

$$p_{ij,k} = \Pr\{\theta_k = s_i | \theta_{k-1} = s_j\} \quad (47)$$

Modelos estocásticos lineares com salto Markoviano (tempo discreto)

■ Esse tipo de modelo é um híbrido entre estado contínuo e discreto:

$$\mathbf{x}_{k+1} = \mathbf{A}_{\theta_k} \mathbf{x}_k + \mathbf{B}_{\theta_k} \mathbf{u}_k + \boldsymbol{\Gamma}_{\theta_k} \mathbf{w}_{k+1} \quad (45)$$

$$\mathbf{y}_k = \mathbf{C}_{\theta_k} \mathbf{x}_k + \mathbf{D}_{\theta_k} \mathbf{u}_k + \boldsymbol{\Psi}_{\theta_k} \mathbf{v}_k \quad (46)$$

em que θ_k é um processo Markoviano de distribuição $p(\theta_k)$, podendo assumir S valores discretos do conjunto $\mathcal{S} = \{s_1, \dots, s_S\}$.

■ Em inglês, usa-se a denominação *Markov Jump Linear System* (MJLS).

■ A variável de estado do sistema é dada pelo par $\{\mathbf{x}_k, \theta_k\}$.

■ Enquanto θ_k permanecer com o mesmo valor, a dinâmica do sistema não se altera.

■ Observa-se que

$$\sum_{i=1}^S p_{ij,k} = 1, \quad \forall j = 1, \dots, S \quad (48)$$

implicando que a soma de todos os elementos de uma mesma coluna de $\mathbf{P}_k$ deve ser 1.

■ Dependência do tempo k :

- Sendo $p_{ij,k}$ dependente do tempo: cadeia de Markov não-homogênea
- Sendo $p_{ij,k} = p_{ij}$ independente do tempo: cadeia de Markov homogênea.

■ Em alguns casos, $S = \infty$.

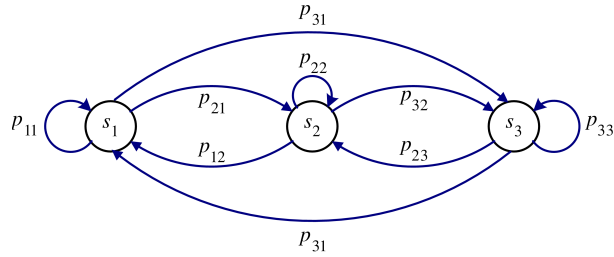
■ Em algumas situações, usa-se uma representação gráfica dita *diagrama de transição de estados*.

■ **Exemplo 3:** Diagrama de transição de estados de uma cadeia homogênea

Considere uma cadeia homogênea com $S = 3$. Sua matriz de transição de probabilidades é dada por

$$\mathbf{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{bmatrix}$$

e diagrama de transição de estados mostrado abaixo:



com

$$w_k = \begin{cases} +1 & \text{com probabilidade } \alpha \\ -1 & \text{com probabilidade } \beta = 1 - \alpha \end{cases} \quad (51)$$

Observa-se que

$$\Pr\{\theta_k = \mathbf{s}_i | \theta_{k-1} = \mathbf{s}_j\} = \begin{cases} \alpha & \text{se } s_i = s_j + 1 \\ \beta & \text{se } s_i = s_j - 1 \\ 0 & \text{qualquer outro caso.} \end{cases} \quad (52)$$

de modo que, considerando $s_n = n$, a partir das seguintes relações:

$$p_{ii} = \Pr\{\theta_k = i | \theta_{k-1} = i\} = 0 \quad (53)$$

$$p_{(i+1)i} = \Pr\{\theta_k = i + 1 | \theta_{k-1} = i\} = \alpha \quad (54)$$

$$p_{i(i+1)} = \Pr\{\theta_k = i | \theta_{k-1} = i + 1\} = \beta \quad (55)$$

Exemplo 4:

Considere o problema do Passeio Aleatório em uma única dimensão:

$$\theta_k = \begin{cases} \theta_{k-1} + 1 & \text{com probabilidade } \alpha \\ \theta_{k-1} - 1 & \text{com probabilidade } \beta = 1 - \alpha \end{cases} \quad (49)$$

tal que $\theta_0 = 0$ e $p + q = 1$. θ_k é uma cadeia de Markov em tempo discreto. No entanto, θ_k pode assumir um número infinito de valores no espaço discreto

$$\mathcal{S} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$$

Solução:

Este modelo pode ser escrito na forma

$$\theta_k = \theta_{k-1} + w_k \quad (50)$$

A matriz de transição de estados é dada por

$$\mathbf{P} = \begin{bmatrix} \ddots & & & & & & & & \\ & \ddots & & & & & & & \\ & & \ddots & & & & & & \\ \cdots & 0 & \alpha & 0 & \beta & 0 & & & \\ & \cdots & 0 & \alpha & 0 & \beta & 0 & \cdots & \\ & & \cdots & 0 & \alpha & 0 & \beta & 0 & \cdots \\ & & & & & \ddots & & & \\ & & & & & & \ddots & & \\ & & & & & & & \ddots & \end{bmatrix} \quad (56)$$

■ A dinâmica de π_k é obtida da seguinte forma: sendo

$$\pi_k = (\pi_{1,k} \ \pi_{2,k} \ \cdots \ \pi_{S,k})^T, \text{ com}$$

$$\pi_{i,k} = \Pr\{\theta_k = \mathbf{s}_i\} \quad (57)$$

que pode ser escrito como

$$\pi_{i,k} = \sum_{j=1}^S \Pr\{\theta_k = \mathbf{s}_i, \theta_{k-1} = \mathbf{s}_j\} \quad (58)$$

$$= \sum_{j=1}^S \Pr\{\theta_k = \mathbf{s}_i | \theta_{k-1} = \mathbf{s}_j\} \Pr\{\theta_{k-1} = \mathbf{s}_j\} \quad (59)$$

$$= \sum_{j=1}^S \Pr\{\theta_k = \mathbf{s}_i | \theta_{k-1} = \mathbf{s}_j\} \pi_{j,k-1} \quad (60)$$

$$= \sum_{j=1}^S p_{ij,k} \cdot \pi_{j,k-1} \quad (61)$$

$$= {}_i [\mathbf{P}_k] \boldsymbol{\pi}_{k-1} \quad (62)$$

em que ${}_i [\mathbf{P}_k]$ é o vetor dado pela i -ésima linha da matriz $\mathbf{P}_k$.

Exemplo 5: Um sistema estocástico linear com saltos Markovianos (simulação mjls)

Seja

$$x_{k+1} = a_{\theta_k} x_k + b_{\theta_k} u_k + \gamma_{\theta_k} w_{k+1} \quad (65)$$

$$y_k = x_k \quad (66)$$

com $\mathcal{S} = \{1, 2\}$, $w_{k+1} \sim \mathcal{N}(0, 1)$,

$$\begin{matrix} a_1 = 0,9 & b_1 = 0,1 & \gamma_1 = 0,1 \\ a_2 = 0,7 & b_2 = 0,6 & \gamma_2 = 0,05 \end{matrix}.$$

O processo θ_k é uma cadeia de Markov homogênea regida pelos parâmetros

$$\mathbf{P} = \begin{bmatrix} 1 - \varepsilon & \varepsilon \\ \varepsilon & 1 - \varepsilon \end{bmatrix}$$

com $\varepsilon = 0,005$ e $\boldsymbol{\pi}_0 = \begin{pmatrix} 1 & 0 \end{pmatrix}^T$.

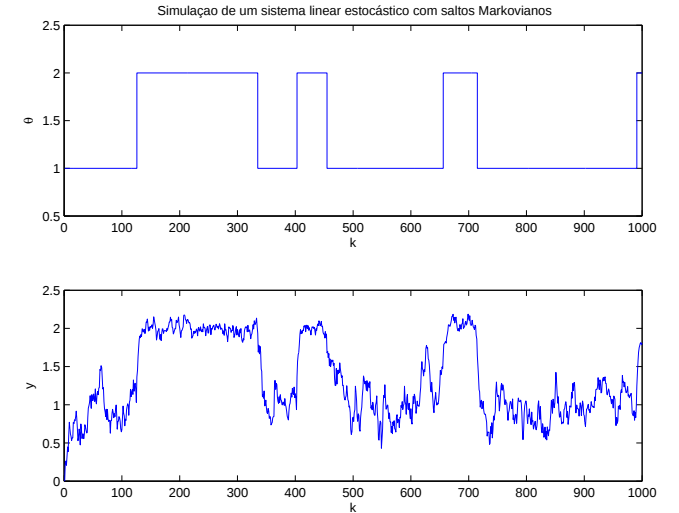
Ao aplicar o resultado acima $i = 1, \dots, S$, tem-se

$$\boldsymbol{\pi}_k = \begin{pmatrix} \pi_{1,k} \\ \pi_{2,k} \\ \vdots \\ \pi_{S,k} \end{pmatrix} = \begin{pmatrix} {}_1 [\mathbf{P}_k] \boldsymbol{\pi}_{k-1} \\ {}_2 [\mathbf{P}_k] \boldsymbol{\pi}_{k-1} \\ \vdots \\ {}_S [\mathbf{P}_k] \boldsymbol{\pi}_{k-1} \end{pmatrix} = \begin{pmatrix} {}_1 [\mathbf{P}_k] \\ {}_2 [\mathbf{P}_k] \\ \vdots \\ {}_S [\mathbf{P}_k] \end{pmatrix} \boldsymbol{\pi}_{k-1} \quad (63)$$

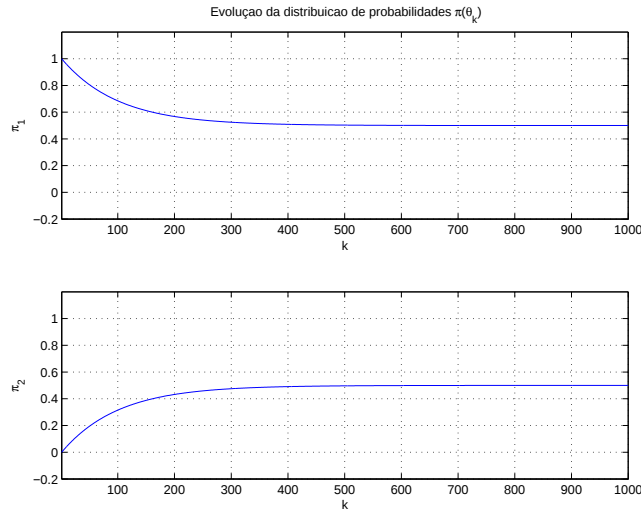
$$= \mathbf{P}_k \boldsymbol{\pi}_{k-1}. \quad (64)$$

■ A cadeia de Markov precisa de uma condição inicial, dada por $\boldsymbol{\pi}_0$. A cada instante discreto k , um valor de θ_k é escolhido dentre os S valores discretos de $\mathcal{S} = \{s_1, \dots, s_S\}$ segundo a distribuição $[\mathbf{P}_k]_{\theta_{k-1}}$, em que $[\cdot]_l$ é a l -ésima coluna da matriz argumento.

Simulação com 1000 amostras para $u_k = 1$, $k = 1, \dots, 1000$:



A evolução dos elementos de π_k é mostrada abaixo:



O problema de filtragem estocástica

- Seja $\mathbf{x}_k$ um processo em tempo discreto gerando medições $\mathbf{y}_k$, deseja-se determinar

$$p(\mathbf{x}_l | \mathbf{y}_1, \dots, \mathbf{y}_k)$$

de forma a obter então uma estimativa $\hat{\mathbf{x}}_k$. De acordo com l e k , o problema pode ser classificado como:

- Problema de suavização: $l < k$
- Problema de filtragem: $l = k$
- Problema de predição: $l > k$

- A resolução do problema de filtragem é primordial, pois tanto suavização como predição são obtidos a partir da solução do problema de filtragem.

- Pelo exemplo anterior, pode-se observar que com o aumento de k , os elementos da distribuição π_k da cadeia homogênea tendem a ficar constantes, ou seja

$$\pi_k \xrightarrow[k \rightarrow \infty]{} \bar{\pi} \quad (67)$$

Com isso, $\pi_{k+1} = \pi_k = \bar{\pi}$, o que leva a

$$\bar{\pi} = \mathbf{P} \bar{\pi} \quad (68)$$

ou ainda

$$\bar{\pi} (\mathbf{I} - \mathbf{P}) = \mathbf{0} \quad (69)$$

de forma que $\bar{\pi}$ pertence ao espaço nulo de $(\mathbf{I} - \mathbf{P})$, e também deve satisfazer a condição

$$\bar{\pi}_1 + \bar{\pi}_2 + \dots + \bar{\pi}_S = 1 \quad (70)$$

No caso do Exemplo 5, a eq. (69) implica em $\bar{\pi}_1 = \bar{\pi}_2$. Considerando $\bar{\pi}_1 + \bar{\pi}_2$, a única solução possível é $\bar{\pi}_1 = \bar{\pi}_2 = 0,5$, que foi verificada na simulação.

O problema de filtragem estocástica

- O restante desse material trata apenas do caso de modelo estocástico linear em tempo discreto.

- Assume-se que o estado evolui de acordo com o modelo de processo

$$\mathbf{x}_k = \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{B}_{k-1} \mathbf{u}_{k-1} + \mathbf{\Gamma}_k \mathbf{w}_k, \quad k = 1, \dots \quad (71)$$

e as medições são modeladas por

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k. \quad (72)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ descorrelacionados, ou seja

$$E\{\mathbf{w}_k \mathbf{v}_k^T\} = \mathbf{0}, \quad E\{\mathbf{w}_k \mathbf{x}_0^T\} = \mathbf{0}, \quad E\{\mathbf{v}_k \mathbf{x}_0^T\} = \mathbf{0}. \quad (73)$$

Filtragem estocástica linear

■ A partir de $p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k)$, estamos interessados em obter uma estimativa $\hat{\mathbf{x}}_k$ para $\mathbf{x}_k$.

- Várias escolhas possíveis: média, mediana, modo principal.
- Necessidade de definição de um critério a minimizar.
- Sendo o erro de estimação dado por $\tilde{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k$, a métrica seria

$$E\{L(\tilde{\mathbf{x}}_k)\}$$

com as possíveis escolhas [1]

- $L(\tilde{\mathbf{x}}_k) = \tilde{\mathbf{x}}_k^T \cdot \mathbf{S} \cdot \tilde{\mathbf{x}}_k$, $\mathbf{S} \geq \mathbf{0}$: critério quadrático L_2
- $L(\tilde{\mathbf{x}}_k) = \sum_i c_i (|\tilde{\mathbf{x}}_k|)_i$: critério L_1
- $L(\tilde{\mathbf{x}}_k) = \min(c, \tilde{\mathbf{x}}_k^T \cdot \mathbf{S} \cdot \tilde{\mathbf{x}}_k)$: critério robusto, que minimiza a influência de outliers.

Esse critério é conhecido como Erro Médio Quadrático, que é a soma dos elementos da diagonal de

$$\mathbf{P}_k = E\{\tilde{\mathbf{x}}_k \cdot \tilde{\mathbf{x}}_k^T | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (79)$$

$$= E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (80)$$

$$= E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (81)$$

ou seja,

$$E\{L(\tilde{\mathbf{x}}_k)\} = \text{tr} \mathbf{P}_k \quad (82)$$

Assim a escolha $\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k$ é a estimativa de mínima variância de $\mathbf{x}_k$, ou seja, aquela que minimiza o traço da matriz de covariâncias do erro de estimação.

Portanto, estamos apenas interessados em

$$\hat{\mathbf{x}}_k = E\{\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (83)$$

$$\mathbf{P}_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (84)$$

Filtragem estocástica linear

■ O critério quadrático é o mais interessante: sendo $\bar{\mathbf{x}}_k = E\{\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k\}$,

$$E\{L(\tilde{\mathbf{x}}_k)\} = E\{\tilde{\mathbf{x}}_k^T \cdot \mathbf{S} \cdot \tilde{\mathbf{x}}_k\} \quad (74)$$

$$= E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \cdot \mathbf{S} \cdot (\mathbf{x}_k - \bar{\mathbf{x}}_k)\} \quad (75)$$

$$+ E\{(\bar{\mathbf{x}}_k - \hat{\mathbf{x}}_k)^T \cdot \mathbf{S} \cdot (\bar{\mathbf{x}}_k - \hat{\mathbf{x}}_k)\} \quad (76)$$

que é minimizado com $\hat{\mathbf{x}}_k = \bar{\mathbf{x}}_k = E\{\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k\}$. Com esta escolha,

$$E\{L(\tilde{\mathbf{x}}_k)\} = E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \cdot \mathbf{S} \cdot (\mathbf{x}_k - \bar{\mathbf{x}}_k)\} \quad (77)$$

para qualquer $\mathbf{S}$. Se $\mathbf{S} = \mathbf{I}$,

$$E\{L(\tilde{\mathbf{x}}_k)\} = E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)^T \cdot (\mathbf{x}_k - \bar{\mathbf{x}}_k)\} = \sum_i E\{(\mathbf{x}_k - \bar{\mathbf{x}}_k)_i^2\} \quad (78)$$

Filtragem estocástica linear

■ Por simplicidade, considere o seguinte processo de Gauss-Markov:

$$\mathbf{x}_k = \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{\Gamma}_k \mathbf{w}_k \quad (85)$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \quad (86)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ e

$$E\{\mathbf{w}_k \mathbf{v}_k^T\} = \mathbf{0}. \quad (87)$$

■ Definição do problema: obter, no instante discreto k ,

$$\hat{\mathbf{x}}_k = E\{\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (88)$$

$$\mathbf{P}_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (89)$$

sabendo que $p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$ evolui no tempo de acordo com

$$p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_{k-1})p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})}{p(\mathbf{y}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})} \quad (90)$$

- Deve ser observado que em se tratando de distribuição Gaussiana para $p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k)$, então $\hat{\mathbf{x}}_k$ coincide com a estimativa de máximo a posteriori (MAP).

Filtragem estocástica linear

- Fase de predição: entre os instantes t_{k-1} e t_k , a estimativa predita é obtida por

$$\hat{\mathbf{x}}_{k|k-1} = E\{\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (95)$$

$$= E\{\mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (96)$$

$$= E\{\mathbf{A}_{k-1}\mathbf{x}_{k-1}|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} + E\{\mathbf{\Gamma}_k\mathbf{w}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (97)$$

Sendo $\mathbf{w}_k$ decorrelacionado das medições (87),

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}E\{\mathbf{x}_{k-1}|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (98)$$

$$= \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} \quad (99)$$

$\mathbf{P}_{k|k-1}$ é dada por

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{\Gamma}_k\mathbf{Q}_k\mathbf{\Gamma}_k^T \quad (100)$$

Filtragem estocástica linear

- O problema é tratado de forma recursiva, partindo de $\hat{\mathbf{x}}_0$ como estimativa ótima inicial, pois $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$.
- A cada instante $t_k = kT$, o procedimento de estimação de $\mathbf{x}_k$ é dividido em duas fases:
 - PREDIÇÃO: por meio do modelo de processo, determina-se $p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})$ cujos parâmetros são

$$\hat{\mathbf{x}}_{k|k-1} = E\{\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (91)$$

$$\mathbf{P}_{k|k-1} = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T|\mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (92)$$

- CORREÇÃO: por meio do modelo de medição, determina-se $p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k)$ cujos parâmetros são

$$\hat{\mathbf{x}}_k = E\{\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (93)$$

$$\mathbf{P}_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T|\mathbf{y}_1, \dots, \mathbf{y}_k\} \quad (94)$$

Filtragem estocástica linear

- Fase de correção: no instante t_k , se nenhuma medição estiver disponível, então $p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k) = p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})$, cujos parâmetros são

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} \quad (101)$$

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} \quad (102)$$

Se um vetor de medidas $\mathbf{y}_k$ estiver disponível, utiliza-se a regra de Bayes

$$p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_k) = \frac{p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_{k-1})p(\mathbf{x}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})}{p(\mathbf{y}_k|\mathbf{y}_1, \dots, \mathbf{y}_{k-1})} \quad (103)$$

com

$$p(\mathbf{y}_k|\mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = p(\mathbf{y}_k|\mathbf{x}_k) \quad (104)$$

$$= p_{\mathbf{v}_k}(\mathbf{y}_k - \mathbf{C}_k\mathbf{x}_k) \quad (105)$$

$$= \mathcal{N}(\mathbf{C}_k\mathbf{x}_k, \mathbf{R}_k) \quad (106)$$

devido a (86).

A predição constitui a distribuição *a priori*

$$p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1}) \quad (107)$$

Para determinar $p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1})$, parte-se de sua média

$$E\{\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} = E\{\mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (108)$$

$$= \mathbf{C}_k E\{\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (109)$$

$$= \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1} \quad (110)$$

Percebe-se que $p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1})$ é constante, e não depende de $\mathbf{x}_k$, como era de se esperar.

Agora, pode-se verificar que

$$p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k) = \frac{p(\mathbf{y}_k | \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1})}{p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1})}$$

com

$$p(\mathbf{y}_k | \mathbf{x}_k, \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \mathcal{N}(\mathbf{C}_k \mathbf{x}_k, \mathbf{R}_k)$$

$$p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \mathcal{N}(\hat{\mathbf{x}}_{k|k-1}, \mathbf{P}_{k|k-1})$$

$$p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \mathcal{N}(\mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}, \mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k)$$

se escreve da forma

$$p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k) = \beta_k \exp\left\{-\frac{1}{2} \vartheta_k\right\}$$

e da matriz de covariâncias

$$E\{(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (111)$$

$$= E\{(\mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})(\mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\}$$

$$= E\{(\mathbf{C}_k [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}] + \mathbf{v}_k)(\mathbf{C}_k [\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}] + \mathbf{v}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (112)$$

$$= \mathbf{C}_k E\{(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})(\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \mathbf{C}_k^T + E\{\mathbf{v}_k \mathbf{v}_k^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} \quad (113)$$

que, a partir de (100) se torna

$$E\{(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})^T | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}\} = \mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k. \quad (114)$$

Portanto,

$$p(\mathbf{y}_k | \mathbf{y}_1, \dots, \mathbf{y}_{k-1}) = \mathcal{N}(\mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}, \mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k) \quad (115)$$

com

$$\beta_k = \frac{\det(\mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{1/2}}{(2\pi)^{n/2} \det(\mathbf{R}_k)^{1/2} \det(\mathbf{P}_{k|k-1})^{1/2}}$$

e

$$\begin{aligned} \vartheta_k &= (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k)^T [\mathbf{R}_k]^{-1} (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k) \\ &\quad + (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{P}_{k|k-1}]^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \\ &\quad - (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k]^{-1} (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \end{aligned}$$

Como $p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k)$ é Gaussiano, a média e a moda são iguais, o que permite obter $\hat{\mathbf{x}}_k$ a partir de

$$\left. \frac{\partial p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k)}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k} = \left. \frac{\partial \ln [p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k)]}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k} = \mathbf{0}.$$

Sendo

$$\ln [p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k)] = \ln(\beta_k) - \frac{1}{2} \frac{\partial \vartheta_k}{\partial \mathbf{x}_k}$$

então

$$\frac{\partial \ln [p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k)]}{\partial \mathbf{x}_k} = -\frac{1}{2} \frac{\partial \vartheta_k}{\partial \mathbf{x}_k}$$

de forma que deseja-se resolver

$$\left. \frac{\partial \vartheta_k}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k} = \mathbf{0}$$

Mas

$$\begin{aligned} \frac{\partial \vartheta_k}{\partial \mathbf{x}_k} &= \frac{\partial \left((\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k)^T [\mathbf{R}_k]^{-1} (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k) \right)}{\partial \mathbf{x}_k} \\ &+ \frac{\partial \left((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{P}_{k|k-1}]^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \right)}{\partial \mathbf{x}_k} \end{aligned}$$

que resulta em

$$\hat{\mathbf{x}}_k = \left\{ \mathbf{C}_k^T [\mathbf{R}_k]^{-1} \mathbf{C}_k + [\mathbf{P}_{k|k-1}]^{-1} \right\}^{-1} \left[\mathbf{C}_k^T [\mathbf{R}_k]^{-1} \mathbf{y}_k + [\mathbf{P}_{k|k-1}]^{-1} \hat{\mathbf{x}}_{k|k-1} \right]$$

Sendo

$$\mathbf{P}_k = E\{(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T | \mathbf{y}_1, \dots, \mathbf{y}_k\}$$

pode-se mostrar que

$$\mathbf{P}_k^{-1} = \mathbf{P}_{k|k-1}^{-1} + \mathbf{C}_k^T [\mathbf{R}_k]^{-1} \mathbf{C}_k$$

Como

$$\begin{aligned} & \frac{\partial \left((\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k)^T [\mathbf{R}_k]^{-1} (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k) \right)}{\partial \mathbf{x}_k} \\ &= -2 (\mathbf{y}_k - \mathbf{C}_k \mathbf{x}_k)^T [\mathbf{R}_k]^{-1} \mathbf{C}_k \end{aligned}$$

e

$$\begin{aligned} & \frac{\partial \left((\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{P}_{k|k-1}]^{-1} (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1}) \right)}{\partial \mathbf{x}_k} \\ &= 2 (\mathbf{x}_k - \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{P}_{k|k-1}]^{-1} \end{aligned}$$

obtem-se

$$\left. \frac{\partial \vartheta_k}{\partial \mathbf{x}_k} \right|_{\mathbf{x}_k = \hat{\mathbf{x}}_k} = -2 (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k)^T [\mathbf{R}_k]^{-1} \mathbf{C}_k + 2 (\hat{\mathbf{x}}_k - \hat{\mathbf{x}}_{k|k-1})^T [\mathbf{P}_{k|k-1}]^{-1} = \mathbf{0}$$

■ Aplicando-se o lema da inversão matricial para esse resultado, pode-se mostrar que [1]

$$p(\mathbf{x}_k | \mathbf{y}_1, \dots, \mathbf{y}_k) = \mathcal{N}(\hat{\mathbf{x}}_k, \mathbf{P}_k) \quad (116)$$

com

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}) \quad (117)$$

$$\mathbf{P}_k = (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k|k-1} \quad (118)$$

$$= (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{G}_k \mathbf{C}_k)^T + \mathbf{G}_k \mathbf{R}_k \mathbf{G}_k^T \quad (119)$$

$\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_{k|k-1}$ é chamado de termo de inovação e $\mathbf{G}_k$ é o Ganho de Kalman, dado por

$$\mathbf{G}_k = \mathbf{P}_{k|k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (120)$$

■ Entrada exógena: se o modelo possuir uma entrada $\mathbf{u}_k$ de forma que (85) se torna

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k \quad (121)$$

o que altera as equações do filtro apenas na equação (99) da fase de predição, que passa a ser

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad (122)$$

e Eq. (100) não se altera uma vez que $\mathbf{u}_k$ não é uma variável aleatória.

(ii) Fase de correção (no instante t_k): dado a medição $\mathbf{y}_k$ com matriz de covariâncias $\mathbf{R}_k$

$$\mathbf{G}_k = \mathbf{P}_{k|k-1}\mathbf{C}_k^T (\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (128)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}) \quad (129)$$

$$\begin{aligned} \mathbf{P}_k &= (\mathbf{I} - \mathbf{G}_k\mathbf{C}_k) \mathbf{P}_{k|k-1} \\ &= (\mathbf{I} - \mathbf{G}_k\mathbf{C}_k) \mathbf{P}_{k|k-1} (\mathbf{I} - \mathbf{G}_k\mathbf{C}_k)^T + \mathbf{G}_k\mathbf{R}_k\mathbf{G}_k^T \end{aligned} \quad (130) \quad (131)$$

Filtro de Kalman (FK)

Algoritmo 1: *Filtro de Kalman Linear*

Seja o sistema dinâmico estocástico em tempo discreto

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k \quad (123)$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \quad (124)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ descorrelacionados, ou seja

$$E\{\mathbf{w}_k\mathbf{v}_k^T\} = \mathbf{0}, \quad E\{\mathbf{w}_k\mathbf{x}_0^T\} = \mathbf{0}, \quad E\{\mathbf{v}_k\mathbf{x}_0^T\} = \mathbf{0}. \quad (125)$$

O filtro de Kalman é o estimador de mínima variância de $\mathbf{x}_k$, composto de duas fases:

(i) Fase de predição (entre os instantes t_{k-1} e t_k):

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad (126)$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{\Gamma}_k\mathbf{Q}_k\mathbf{\Gamma}_k^T \quad (127)$$

Filtro de Kalman (FK)

Observação 1:

O FK é O ESTIMADOR Bayesiano para sistemas lineares variantes no tempo. Sistemas invariantes no tempo são apenas o caso especial, em que as matrizes do sistema são constantes.

Observação 2:

O conhecido filtro de Wiener é o estimador para sistemas lineares invariantes no tempo, sendo formulado na forma de resposta em frequência.

Observação 3:

A estimativa $\hat{\mathbf{x}}_k$ é de mínima variância, pois minimiza $tr\mathbf{P}_k$.

Observação 4:

A medição $\mathbf{y}_k$ não afeta a atualização de $\mathbf{P}_k$, apenas sua incerteza representada sob a forma de $\mathbf{R}_k$.

Filtro de Kalman (FK)

Observação 5:

Casos especiais:

(i) Medição não afeta estimativa:

$$\begin{aligned} \mathbf{R}_k &\rightarrow \infty \implies \mathbf{G}_k \rightarrow \mathbf{0}. \\ \mathbf{P}_{k|k-1} &\rightarrow \mathbf{0} \implies \mathbf{G}_k \rightarrow \mathbf{0}. \end{aligned}$$

(ii) Medição afeta ao máximo a estimativa:

$$\begin{aligned} \mathbf{R}_k &\rightarrow \mathbf{0} \implies \mathbf{G}_k \rightarrow \mathbf{P}_{k|k-1} \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_{k|k-1} \mathbf{C}_k)^{-1}. \\ \mathbf{P}_{k|k-1} &\rightarrow \infty \text{ em consequência de } \mathbf{Q}_k \rightarrow \infty \end{aligned}$$

filtro se tornam

$$\hat{x}_k = \hat{x}_{k|k-1} + \frac{\sigma_{k|k-1}^2}{\sigma_{k|k-1}^2 + \sigma_v^2} (y_k - \hat{x}_{k|k-1}) \quad (134)$$

$$= w \cdot \hat{x}_{k|k-1} + (1 - w) y_k \quad (135)$$

que é uma média ponderada com

$$w = \frac{\sigma_v^2}{\sigma_{k|k-1}^2 + \sigma_v^2}. \quad (136)$$

Portanto, $\hat{x}_k$ estará entre $\hat{x}_{k|k-1}$ e y_k , dependendo de w . No mais,

$$\sigma_k^2 = w \cdot \sigma_{k|k-1}^2 = (1 - w) \sigma_v^2 \quad (137)$$

implicando que $\sigma_k^2 \leq \sigma_{k|k-1}^2$ e $\sigma_k^2 \leq \sigma_v^2$ uma vez que $0 \leq w \leq 1$. Isto implica em $\sigma_k^2 \leq \min(\sigma_{k|k-1}^2, \sigma_v^2)$. Com este resultado, verifica-se a redução de incertezas após considerar-se a medição y_k .

Filtro de Kalman (FK)

Observação 6:

Sendo $\mathbf{P}_k$ e $\mathbf{P}_{k|k-1}$ positivas semi-definidas (i.e., são matrizes de covariâncias), pela equação (130) tem-se

$$(\mathbf{I} - \mathbf{G}_k \mathbf{C}_k) \succeq \mathbf{0} \implies \mathbf{I} \succeq \mathbf{G}_k \mathbf{C}_k \quad (132)$$

no sentido matricial. Isto implica em $\mathbf{P}_k \preceq \mathbf{P}_{k|k-1}$, resultando em redução de incerteza de predição após o processo de correção. Isto pode ser facilmente verificado para o caso escalar em que o estado é observado diretamente na saída

$$y_k = x_k + v_k \quad (133)$$

em que $x_{k|k-1} \sim \mathcal{N}(\hat{x}_{k|k-1}, \sigma_{k|k-1}^2)$ e $v_k \sim \mathcal{N}(0, \sigma_v^2)$. As equações de correção do

Aplicações do Filtro de Kalman

■ Estimador de estados: aplicação direta do filtro, na forma como foi apresentado, a um sistema com modelo linear.

Exemplo 6: FK para o modelo do exemplo 2 (simulação fk_foguete)

Considerando o modelo do foguete do Exemplo 2 subindo com empuxo constante $F_k = 10000 \text{ N}$, $M = 1000 \text{ kg}$, $g = 9,81 \text{ m/s}^2$ e $T = 1 \text{ s}$.

Aplicações do Filtro de Kalman

■ Fusão de dados sensoriais (caso de fusão simples):

Considere um sistema de medição de temperatura θ com dois sensores S_1 e S_2 . Cada sensor provê medições θ_1 e θ_2 com variâncias σ_1^2 e σ_2^2 . O modelo

$$\theta_k = \theta_{k-1} \quad (138)$$

$$y_k = \theta_k + v_k \quad (139)$$

em que k refere-se ao índice do sensor, pode ser usado para fusionar as medidas providas por esses sensores e obter uma estimativa de mínima variância para θ . $\theta_1 = \theta_1$, com variância σ_1^2 , uma estimativa de θ obtida apenas com a medida do sensor S_1 . Em $k = 2$, deverá ser considerada a medida de S_2 . Pode-se mostrar que

Aplicações do Filtro de Kalman

■ Fusão de dados sensoriais (generalização):

Em muitos casos, dado o modelo

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k \quad (140)$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \quad (141)$$

$\mathbf{u}_{k-1}$ é uma entrada medida por sensores, assim como o é $\mathbf{y}_k$. Ou seja, os sensores que proporcionam $\mathbf{u}_{k-1}$, apesar de introduzirem incerteza no modelo, contribuem para uma melhor predição. Nesse caso, considerando que $\mathbf{u}_{k-1}$ possui matriz de covariâncias $\mathbf{Q}_{\mathbf{u}_{k-1}}$, então as equações de predição do filtro se tornam

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad (142)$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{B}_{k-1}\mathbf{Q}_{\mathbf{u}_{k-1}}\mathbf{B}_{k-1}^T + \mathbf{\Gamma}_k\mathbf{Q}_k\mathbf{\Gamma}_k^T \quad (143)$$

e as equações de correção permanecem inalteradas

$\hat{\theta}_{2|1} = \theta_1$ e $\hat{\theta}_2 = \hat{\theta}$ com variância σ^2 em que

$$\hat{\theta} = \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2}\theta_1 + \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2}\theta_2$$

$$\sigma^2 = \frac{\sigma_1^2\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \leq \min(\sigma_1^2, \sigma_2^2)$$

E ainda, uma solução "ingênua" para o problema que seria

$$\hat{\theta}_i = \frac{\theta_1 + \theta_2}{2}$$

apresenta sempre variância

$$\sigma_i^2 = \frac{\sigma_1^2 + \sigma_2^2}{4}$$

mas $\sigma_i^2 \geq \sigma^2$, ou seja, essa solução apresenta incerteza no mínimo igual à da solução ótima.

Aplicações do Filtro de Kalman

■ Estimação sequencial da média de um sinal:

Supor que deseja-se estimar (rastrear) o valor médio x_k de um sinal discreto escalar z_k . O seguinte modelo de evolução da média é usado:

$$x_k = x_{k-1} + w_k \quad (144)$$

com $w_k \sim \mathcal{N}(0, q_k^2)$. A componente w_k se justifica por

- O sinal z_k possuir média variante no tempo (não estacionário);
- x_0 não ser bem conhecido *a priori*.

O modelo de medição é dado por

$$y_k = x_k + v_k \quad (145)$$

com $v_k \sim \mathcal{N}(0, r_k^2)$. A medição y_k é o próprio sinal, i.e., $y_k = z_k$. Pede-se:

- Determinar as equações do filtro de Kalman para este problema;
- Mostrar que o filtro de Kalman pode ser interpretado como a versão discreta (Euler) de um filtro passa-baixas com função de transferência

$$H(s) = \frac{\omega_c}{s + \omega_c}. \quad (146)$$

com ω_c variante no tempo.

Exemplo 7: FK para rastrear a média de um sinal (simulação *fk_media*)

O Filtro de Kalman pode ser usado para a estimação recursiva de θ com o uso do modelo

$$\begin{aligned} \theta_k &= \theta_{k-1} + \mathbf{w}_k \\ y_k &= \varphi_k^T \theta_k + v_k \end{aligned}$$

com condição inicial $\theta_0 \sim \mathcal{N}(\hat{\theta}_0, \mathbf{P}_0)$, $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$ e $v_k \sim \mathcal{N}(0, r_k^2)$. Assim, o algoritmo de estimação para θ_k pode incorporar predição e correção nas seguintes relações:

$$\mathbf{G}_k = \frac{(\mathbf{P}_{k-1} + \mathbf{Q}_k) \varphi_k}{r_k^2 + \varphi_k^T [\mathbf{P}_{k-1} + \mathbf{Q}_k] \varphi_k} \quad (147)$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \mathbf{G}_k (y_k - \varphi_k^T \hat{\theta}_{k-1}) \quad (148)$$

$$\mathbf{P}_k = (\mathbf{P}_{k-1} + \mathbf{Q}_k) - \frac{(\mathbf{P}_{k-1} + \mathbf{Q}_k) \varphi_k \varphi_k^T (\mathbf{P}_{k-1} + \mathbf{Q}_k)}{r_k^2 + \varphi_k^T [\mathbf{P}_{k-1} + \mathbf{Q}_k] \varphi_k} \quad (149)$$

Com $\mathbf{Q}_k = \mathbf{0}$, o algoritmo obtido é chamado de Estimador Recursivo de Ganho

Aplicações do Filtro de Kalman

■ Identificação recursiva de modelos lineares:

O modelo ARX

$$\begin{aligned} y_k A(q) &= u_k B(q) + v_k \\ A(q) &= 1 + a_1 q^{-1} + \dots + a_{n_a} q^{-n_a} \\ B(q) &= b_1 q^{-1} + \dots + b_{n_b} q^{-n_b} \end{aligned}$$

pode ser escrito na forma

$$\begin{aligned} y_k &= \varphi_k^T \theta \\ \varphi_k^T &= \begin{bmatrix} -y_{k-1} & \dots & -y_{k-n_a} & u_{k-1} & \dots & u_{k-n_b} \end{bmatrix} \\ \theta &= \begin{bmatrix} a_1 & \dots & a_{n_a} & b_1 & \dots & b_{n_b} \end{bmatrix} \end{aligned}$$

Normalizado [2]. Em que condições o algoritmo resultaria em

$$\mathbf{G}_k = \frac{\mathbf{P}_{k-1} \varphi_k}{\lambda_k + \varphi_k^T \mathbf{P}_{k-1} \varphi_k} \quad (150)$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} + \mathbf{G}_k (y_k - \varphi_k^T \hat{\theta}_{k-1}) \quad (151)$$

$$\mathbf{P}_k = \frac{1}{\lambda_k} \left[\mathbf{P}_{k-1} - \frac{\mathbf{P}_{k-1} \varphi_k \varphi_k^T \mathbf{P}_{k-1}}{\lambda_k + \varphi_k^T \mathbf{P}_{k-1} \varphi_k} \right] \quad (152)$$

que é o algoritmo dos mínimos quadrados recursivos com fator de esquecimento λ_k ?

Aplicações do Filtro de Kalman

Solução: $r_k^2 = 1$ e $\mathbf{Q}_k = \mathbf{P}_{k-1}(1 - \lambda_k)/\lambda_k$. Tem-se mínimos quadrados recursivos para $\lambda_k = 1$, o que leva a $\mathbf{Q}_k = \mathbf{0}$.

■ Detecção de eventos ou mudança: na equação de correção do filtro,

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1})$$

o termo de inovação $\nu_k = \mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}$ é uma medida da compatibilidade da medição $\mathbf{y}_k$ com sua predição $\mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}$. Grandes valores para esse termo indicam incompatibilidade, quer seja no modelo, quer seja na medição.

A compatibilidade deve ser medida por uma métrica estatística, tal como a distância de Mahalanobis:

$$d_k = \nu_k^T [\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k]^{-1} \nu_k$$

detecção menos susceptível a artefatos ou perturbações temporárias:

$$s_k = \sum_{i=0}^{q-1} d_{k-i}$$

de forma que $s_k \sim \chi_{qm}^2$.

■ A interpretação sobre a falha no teste depende do tipo de problema:

- Detecção de eventos: ocorrência de um evento que interferiu na medida $\mathbf{y}_k$;
- Detecção de mudança: importante alteração no modelo do sistema, que interferiu em $\mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}$.

Exemplo 8: FK para detecção de evento e estimação de parâmetros (simulação fk_evento)

em que $\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{R}_k$ é a matriz de covariâncias de ν_k . Pode-se mostrar que $d_k \sim \chi_m^2$ e que, se houver compatibilidade com a medida $\mathbf{y}_k$, o teste

$$d_k < \chi_m^2(1 - r)$$

se torna verdadeiro, com $\chi_m^2(1 - r)$ sendo um limiar constante que incorpora um percentual $1 - r$ da distribuição. Esses limiares podem ser obtidos da tabela do χ_m^2 [3]. Geralmente escolhe-se $1 - r = 0,95$:

$$(1 - r) = 0,95 \text{ com } m = 1 \implies \chi_1^2(1 - r) = 3,841$$

$$(1 - r) = 0,95 \text{ com } m = 2 \implies \chi_2^2(1 - r) = 5,991$$

$$(1 - r) = 0,95 \text{ com } m = 3 \implies \chi_3^2(1 - r) = 7,815$$

Em alguma situações, usa-se uma janela de q valores passados de ν_k para ter uma

Aplicações do Filtro de Kalman

■ Rastreamento de parâmetros

Considere o seguinte modelo de primeira ordem:

$$y_k = ay_{k-1} + bu_{k-1}$$

em que a e b são parâmetros a serem estimados a partir de medições y_k . Este problema pode ser atacado sob a ótica do *rastreamento de parâmetros*, em que são usados os seguintes modelos aproximados de evolução para estes parâmetros:

$$a_k \approx a_{k-1} + T\dot{a}_{k-1}$$

$$b_k \approx b_{k-1} + T\dot{b}_{k-1}$$

com $\dot{a}_{k-1}$ e $\dot{b}_{k-1}$ sendo taxas (ou velocidades) de variação dos parâmetros, a serem estimadas. T é o período de amostragem. Observa-se que com $\dot{a}_{k-1} = \dot{b}_{k-1} = 0$, o modelo fica similar ao de identificação de modelos lineares.

Pede-se para propor um estimador usando o Filtro de Kalman.

Exemplo 9: FK para rastrear parâmetros (simulação *fk_rastreamento_parametros*)

Nessa simulação, são verificadas as seguintes situações:

- $a_k = 0,8$ e $b_k = 0,2$;
- $a_k = 0,8$ e b_k sendo uma onda triangular.
- $a_k = 0,8$ e b_k sendo um passeio aleatório com $b_0 = 0,85$ e confinado a $0,2 \leq b_k < 2,9$.

Consistência do Filtro de Kalman

Teorema 1: Consistência de estimação diante de erro de modelagem ([1], pp. 248)

Seja o sistema

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k \quad (153)$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \quad (154)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \bar{\mathbf{Q}}_k)$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \bar{\mathbf{R}}_k)$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \bar{\mathbf{P}}_0)$ descorrelacionados, se o filtro de Kalman for empregado com parâmetros

$$\mathbf{P}_0 \geq \bar{\mathbf{P}}_0 \quad \mathbf{Q}_k \geq \bar{\mathbf{Q}}_k \quad \mathbf{R}_k \geq \bar{\mathbf{R}}_k$$

$\hat{\mathbf{x}}_k$ e $\hat{\mathbf{x}}_{k|k-1}$ são consistentes pois é garantido que

$$\mathbf{P}_k \geq \bar{\mathbf{P}}_k \quad \mathbf{P}_{k|k-1} \geq \bar{\mathbf{P}}_{k|k-1}$$

Aplicações do Filtro de Kalman

■ Outros exemplos:

- Filtragem adaptativa: caso em que $\mathbf{Q}_k$ e $\mathbf{R}_k$ são estimados on-line [4].
- Rastreamento de alvos [5].
- Estimação sob restrição (usando pseudo-medições $\mathbf{d} = \mathbf{C}\mathbf{x}_k$).
- Branqueamento: o resultado é o termo de inovação *a posteriori* $\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_k$.

Exemplo 10: Consistência do FK para o modelo do exemplo 2 (simulação *fk_foguete_consistencia*)

Considerando o modelo do foguete do Exemplo 2 subindo com empuxo constante $F_k = 10000 \text{ N}$, $M = 1000 \text{ kg}$, $g = 9,81 \text{ m/s}^2$ e $T = 1 \text{ s}$.

São tratados dois casos:

$$\text{caso 1 : } \mathbf{A}_k = \bar{\mathbf{A}}_k \quad \mathbf{B}_k = \bar{\mathbf{B}}_k \quad \mathbf{C}_k = \bar{\mathbf{C}}_k$$

$$\text{caso 2 : } \mathbf{A}_k = 0,99\bar{\mathbf{A}}_k \quad \mathbf{B}_k = \bar{\mathbf{B}}_k \quad \mathbf{C}_k = \bar{\mathbf{C}}_k$$

Para cada caso, considera-se também $\mathbf{P}_0 = \gamma\bar{\mathbf{P}}_0$, $\mathbf{Q}_k = \gamma\bar{\mathbf{Q}}_k$ e $\mathbf{R}_k = \gamma\bar{\mathbf{R}}_k$ para $\gamma = 0,1$, $\gamma = 1$ e $\gamma = 10$.

Variantes do Filtro de Kalman

■ Filtro de regime permanente

Considere o sistema

$$\mathbf{x}_k = \mathbf{A}\mathbf{x}_{k-1} + \mathbf{\Gamma}\mathbf{w}_k \quad (155)$$

$$\mathbf{y}_k = \mathbf{C}\mathbf{x}_k + \mathbf{v}_k \quad (156)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q})$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R})$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ descorrelacionados. Sendo $\mathbf{A}$ e $\mathbf{C}$ invariantes no tempo, os auto-valores de $\mathbf{A}$ estando dentro do círculo unitário e $\mathbf{w}_k$ e $\mathbf{v}_k$ estacionários, pode-se verificar que com $k \rightarrow \infty$, tem-se que

$$\mathbf{P}_{ss} = \mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-2} = \mathbf{P}_{k-2|k-3} = \dots \quad (157)$$

Isso implica que o ganho do filtro de Kalman alcançará um valor constante dado por

$$\mathbf{G}_{ss} = \mathbf{P}_{ss} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{ss} \mathbf{C}^T + \mathbf{R})^{-1}$$

Sendo $\mathbf{P}_{ss} = \mathbf{P}_{k|k-1} = \mathbf{P}_{k-1|k-2}$, então $\mathbf{P}_{ss}$ deve satisfazer

$$\mathbf{P}_{ss} = \mathbf{A} \mathbf{P}_{ss} \mathbf{A}^T - \mathbf{A} \mathbf{P}_{ss} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{ss} \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \mathbf{P}_{ss} \mathbf{A}^T + \mathbf{\Gamma} \mathbf{Q} \mathbf{\Gamma}^T$$

que é da forma da equação algébrica de Riccati para tempo discreto. Existem várias técnicas para resolução dessa equação. Uma delas é usada pela função `dare` do MATLAB.

que é constante, não precisando ser mais calculado a cada passo de amostragem.

Dado que

$$\mathbf{P}_{k|k-1} = \mathbf{A} \mathbf{P}_{k-1} \mathbf{A}^T + \mathbf{\Gamma} \mathbf{Q} \mathbf{\Gamma}^T \quad (158)$$

$$\mathbf{P}_{k-1} = (\mathbf{I} - \mathbf{G}_{k-1} \mathbf{C}) \mathbf{P}_{k-1|k-2} \quad (159)$$

com

$$\mathbf{G}_{k-1} = \mathbf{P}_{k-1|k-2} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{k-1|k-2} \mathbf{C}^T + \mathbf{R})^{-1}$$

então

$$\mathbf{P}_{k|k-1} = \mathbf{A} (\mathbf{I} - \mathbf{G}_{k-1} \mathbf{C}) \mathbf{P}_{k-1|k-2} \mathbf{A}^T + \mathbf{\Gamma} \mathbf{Q} \mathbf{\Gamma}^T \quad (160)$$

$$= \mathbf{A} \mathbf{P}_{k-1|k-2} \mathbf{A}^T \quad (161)$$

$$- \mathbf{A} \mathbf{P}_{k-1|k-2} \mathbf{C}^T (\mathbf{C} \mathbf{P}_{k-1|k-2} \mathbf{C}^T + \mathbf{R})^{-1} \mathbf{C} \mathbf{P}_{k-1|k-2} \mathbf{A}^T + \mathbf{\Gamma} \mathbf{Q} \mathbf{\Gamma}^T \quad (162)$$

$$(163)$$

Variantes do Filtro de Kalman

■ Filtro de informação

As equações do filtro de Kalman podem ser re-escritas empregando a matriz de informação

$$\mathcal{P}_k \triangleq \mathbf{P}_k^{-1}$$

Sendo o sistema

$$\mathbf{x}_k = \mathbf{A}_{k-1} \mathbf{x}_{k-1} + \mathbf{B}_{k-1} \mathbf{u}_{k-1} + \mathbf{w}_k \quad (164)$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k \quad (165)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathcal{Q}_k^{-1})$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathcal{R}_k^{-1})$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathcal{P}_0^{-1})$ descorrelacionados, então o filtro de informação emprega as seguintes equações:

(i) Fase de predição (entre os instantes t_{k-1} e t_k):

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad (166)$$

$$\mathcal{P}_{k|k-1} = \mathcal{Q}_k - \mathcal{Q}_k\mathbf{A}_{k-1}(\mathcal{P}_{k-1} + \mathbf{A}_{k-1}^T\mathcal{Q}_k\mathbf{A}_{k-1})^{-1}\mathbf{A}_{k-1}^T\mathcal{Q}_k \quad (167)$$

(ii) Fase de correção (no instante t_k): dado a medição $\mathbf{y}_k$ com matriz de informação $\mathcal{R}_k$

$$\mathcal{P}_k = \mathcal{P}_{k|k-1} + \mathbf{C}_k^T\mathcal{R}_k\mathbf{C}_k \quad (168)$$

$$\mathbf{G}_k = (\mathcal{P}_k)^{-1}\mathbf{C}_k^T\mathcal{R}_k\mathbf{C}_k \quad (169)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}) \quad (170)$$

Esse filtro apresenta vantagem computacional apenas de $m \gg n$.

O filtro ótimo para esse problema é dado pela formulação seguinte:

(i) Fase de predição (entre os instantes t_{k-1} e t_k):

$$\hat{\mathbf{x}}_{k|k-1} = \mathbf{A}_{k-1}\hat{\mathbf{x}}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} \quad (174)$$

$$\mathbf{P}_{k|k-1} = \mathbf{A}_{k-1}\mathbf{P}_{k-1}\mathbf{A}_{k-1}^T + \mathbf{\Gamma}_k\mathbf{Q}_k\mathbf{\Gamma}_k^T \quad (175)$$

(ii) Fase de correção (no instante t_k): dado a medição $\mathbf{y}_k$ com matriz de covariâncias $\mathbf{R}_k$

$$\mathbf{G}_k = (\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{S}_k)(\mathbf{C}_k\mathbf{P}_{k|k-1}\mathbf{C}_k^T + \mathbf{C}_k\mathbf{S}_k + \mathbf{S}_k^T\mathbf{C}_k^T + \mathbf{R}_k)^{-1} \quad (176)$$

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_{k|k-1} + \mathbf{G}_k(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_{k|k-1}) \quad (177)$$

$$\mathbf{P}_k = \mathbf{P}_{k|k-1} - \mathbf{G}_k(\mathbf{C}_k\mathbf{P}_{k|k-1} + \mathbf{S}_k^T) \quad (178)$$

Razões se ter $\mathbf{w}_k$ e $\mathbf{v}_k$ correlacionados:

- Ruído estrutural agindo nos modelos de processo e de medição
- Uso do mesmo sensor para se obter $\mathbf{u}_k$ e $\mathbf{y}_k$

Variantes do Filtro de Kalman

■ Filtro para processos com ruídos de processo e de medição correlacionados

Seja o sistema dinâmico estocástico em tempo discreto

$$\mathbf{x}_k = \mathbf{A}_{k-1}\mathbf{x}_{k-1} + \mathbf{B}_{k-1}\mathbf{u}_{k-1} + \mathbf{\Gamma}_k\mathbf{w}_k \quad (171)$$

$$\mathbf{y}_k = \mathbf{C}_k\mathbf{x}_k + \mathbf{v}_k \quad (172)$$

com $\mathbf{w}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{Q}_k)$, $\mathbf{v}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{R}_k)$, $\mathbf{x}_0 \sim \mathcal{N}(\hat{\mathbf{x}}_0, \mathbf{P}_0)$ tais que

$$E\{\mathbf{w}_k\mathbf{v}_k^T\} = \mathbf{S}_k, \quad E\{\mathbf{w}_k\mathbf{x}_0^T\} = \mathbf{0}, \quad E\{\mathbf{v}_k\mathbf{x}_0^T\} = \mathbf{0}. \quad (173)$$

ou seja, $\mathbf{w}_k$ e $\mathbf{v}_k$ são correlacionados tais que

$$E\left\{\begin{pmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{pmatrix} \begin{pmatrix} \mathbf{w}_k^T & \mathbf{v}_k^T \end{pmatrix}\right\} = \begin{pmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^T & \mathbf{R}_k \end{pmatrix}$$

Exemplo 11: Filtragem de uma versão correlata do modelo do exemplo 2 (simulação *fk_foguete_correlato*)

Considerando o modelo do foguete do Exemplo 2 subindo com empuxo constante $F_k = 10000 \text{ N}$, $M = 1000 \text{ kg}$, $g = 9,81 \text{ m/s}^2$ e $T = 1 \text{ s}$.

Tem-se que os ruídos de processo e de medição são relacionados por

$$\begin{pmatrix} \mathbf{w}_k \\ \mathbf{v}_k \end{pmatrix} = \mathbf{H}\boldsymbol{\lambda}_k$$

com $\boldsymbol{\lambda}_k \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$. Portanto

$$\begin{pmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^T & \mathbf{R}_k \end{pmatrix} = \mathbf{H}\mathbf{H}^T$$

Nessa simulação, é feita uma comparação entre o filtro de Kalman e o filtro de Kalman correlato.

Variantes do Filtro de Kalman

- Intersecção de covariâncias: trabalho...

Referências

- [1] JAZWINSKI, A. H. *Stochastic Processes and Filtering Theory*. [S.l.]: Academic Press, 1970.
- [2] LJUNG, L. *System Identification - Theory for the User*. Second edition. [S.l.]: Prentice Hall, 1999.
- [3] MEYER, P. *Probabilidade - aplicações à estatística*. Segunda edição. [S.l.]: Livros Técnicos e Científicos (LTC), 2003.
- [4] ANDERSON, B. D.; MOORE, J. B. *Optimal Filtering*. [S.l.]: Prentice Hall, 1979.
- [5] BAR-SHALOM, Y.; LI, X.-R. *Estimation and Tracking: Principles, Techniques and Software*. [S.l.]: Norwood: Artech House, 1987.