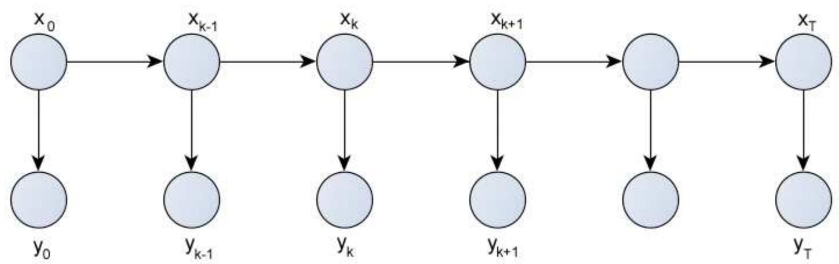
PSC: Progettazione di sistemi di controllo	a.a. 2010-2011
Lezione 10 — Novembre 5, 2010	
Docente: Luca Schenato	*Stesori: G. Pattarello, D. Volpato*

10.1 Kalman Smoother

Si vuole calcolare la densità di probabilità

$$p(x_k|Y^T) \qquad \text{dove} \qquad x_k|Y^T \sim \mathcal{N}(\hat{x}_{k|T}, P_{k|T}) \qquad \text{e} \qquad Y^T = \{y_0, \cdots, y_T\}$$

e per fare questo si cerca di spezzarla in parti condizionatamente indipendenti.



Nella Fig. 10.1 è possibile osservare come, in un grafo, si possano vedere le componenti disgiunte rispetto al condizionamento ad un nodo. Si procede per via grafica ad eliminare il nodo rispetto a cui stiamo condizionando e tutti i suoi archi; le componenti disgiunte sono ora immediatamente individuabili.

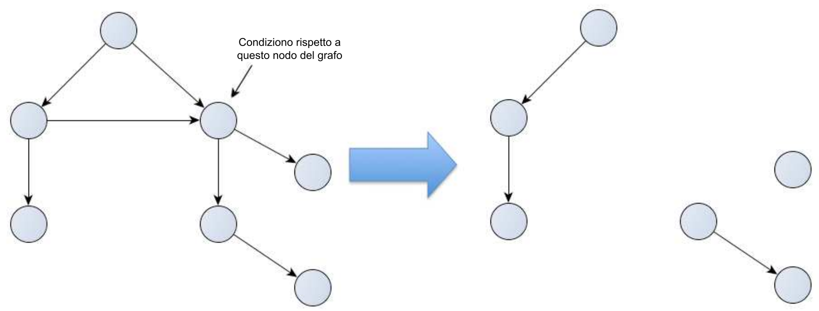


Figura 10.1. Individuazione componenti indipendenti rispetto al condizionamento ad un nodo di un grafo

Nel nostro caso, applicando Bayes, si condizionerà rispetto a x_k e quindi avremo quanto mostrato in Fig. 10.2.

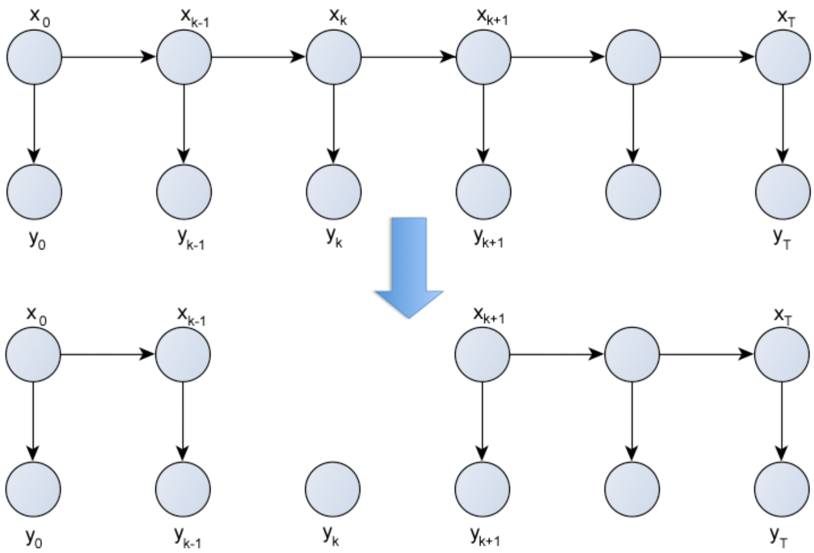


Figura 10.2. Componenti indipendenti rispetto al condizionamento a x_k

Il problema quindi si spezza in due parti:

- stima di x_k con le misure passate
- stima di x_k con le misure future

Le due stime sono indipendenti. Si otterrà $p(x_k|Y^T)$ come stima connessa tra le due stime ottenute.

10.1.1 Equazioni dell'interpolatore

Per la proprietà di Bayes si ha che:

$$p(x_k|Y^T) = \frac{p(y_0, \cdots, y_k, y_{k+1}, \cdots, y_T|x_k)p(x_k)}{p(Y^T)} \propto p(y_0, \cdots, y_k, y_{k+1}, \cdots, y_T|x_k)p(x_k)$$

Ora si spezza la catena di misure alla k-esima misura, supponendo di includere la k-esima misura tra le misure future ottenendo:

$$p(x_k|Y^T) \propto \overbrace{p(y_0, \cdots, y_{k-1}|x_k)}^{predittore-passato} \; \overbrace{p(y_k|x_k)}^{pred.-presente} \; \overbrace{p(y_{k+1}, \cdots, y_T|x_k)}^{predittore-futuro} p(x_k)$$

Definendo

$$\widetilde{Y}^k \triangleq \{y_k, y_{k+1}, \cdots, y_T\}$$

si può scrivere tutto nel più compatto

$$p(x_k|Y^T) \propto p(Y^{k-1}|x_k)p(\widetilde{Y}^k|x_k)p(x_k)$$

Applicando Bayes si ha che

$$p(x_k|Y^T) \propto \frac{p(x_k|Y^{k-1})p(Y^{k-1})}{p(x_k)}\frac{p(x_k|\widetilde{Y}^k)p(\widetilde{Y}^k)}{p(x_k)}p(x_k)$$

cosiderando che $p(Y^{k-1})$ e $p(\widetilde{Y}^k)$ sono numeri noti, allora possono essere pensati conglobati nella costante di normalizzazione; quindi operando le dovute semplificazioni risulta che:

$$p(x_k|Y^T) \propto p(x_k|Y^{k-1})\frac{p(x_k|\widetilde{Y}^k)}{p(x_k)}$$

dove:

$$p(x_k|Y^{k-1}) \sim \mathcal{N}_{x_k}(\hat{x}_{k|k-1}, P_{k|k-1}) \qquad \text{filtro di Kalman classico}$$

e

$$\frac{p(x_k|\widetilde{Y}^k)}{p(x_k)} \qquad \text{sembrerebbe il filtro di Kalman all'indietro.}$$

Considerando le due funzioni di distribuzione, lo si può vedere come rapporto tra due funzioni esponenziali in x_k, quindi

$$\frac{p(x_k|\widetilde{Y}^k)}{p(x_k)} = \frac{\mathcal{N}_{x_k}(\mu_1, \Sigma_1)}{\mathcal{N}_{x_k}(\mu_2, \Sigma_2)} \propto \mathcal{N}_{x_k}(\mu, \Sigma)$$

Infatti si ha considera il prodotto (o il rapporto) di due variabili aleatorie gaussiane, questo a sua volta deve essere proporzionale ad un'altra variabili aleatorie gaussiana con un'opportuna scelta di media e varianza. In particolare per quel che riguarda il prodotto di v.a. gaussiane si ha:

$$\mathcal{N}_x(\mu_1, \Sigma_1)\mathcal{N}_x(\mu_2, \Sigma_2) \propto \mathcal{N}_x(\mu, \Sigma)$$

Considerando i termini quadratici delle due variabili aleatorie si ha:

$$\begin{aligned}
&(x-\mu_1)^T\Sigma_1^{-1}(x-\mu_1) + (x-\mu_2)^T\Sigma_2^{-1}(x-\mu_2)\\
&x^T\Sigma_1^{-1}x - 2x^T\Sigma_1^{-1}\mu_1 + \mu_1^T\Sigma_1^{-1}\mu_1 + x^T\Sigma_2^{-1}x - 2x^T\Sigma_2^{-1}\mu_2 + \mu_2^T\Sigma_2^{-1}\mu_2
\end{aligned}$$

e queste somme devono essere uguali a:

$$\begin{aligned}
&(x-\mu)^T\Sigma^{-1}(x-\mu)\\
&x^T\Sigma^{-1}x - 2x^T\Sigma^{-1}\mu + \mu^T\Sigma^{-1}\mu
\end{aligned}$$

Imponendo l'uguaglianza tra i termini corrispondenti risulta:

$$\begin{aligned}
\Sigma^{-1} &= \Sigma_1^{-1} + \Sigma_2^{-1}\\
\mu &= \Sigma\left(\Sigma_1^{-1}\mu_1 + \Sigma_2^{-1}\mu_2\right)
\end{aligned}$$

Utilizzando lo stesso ragionamento possiamo quindi scrivere che sicuramente

$$\frac{p(x_k|\widetilde{Y}^k)}{p(x_k)} \propto \mathcal{N}_{x_k}(\widetilde{x}_{k|k}, \widetilde{P}_{k|k})$$

per un opportuna scelta di $\widetilde{x}_{k|k}, \widetilde{P}_{k|k}$. Definiamo quindi in maniera arbitraria tale quantità come

$$p(x_k|\widetilde{Y}^k, Inv) \triangleq \mathcal{N}_{x_k}(\widetilde{x}_{k|k}, \widetilde{P}_{k|k})$$

Per ricavare $p(x_k|\widetilde{Y}^k, Inv)$ si suppone che siano note:

- $p(x_{k+1}|\widetilde{Y}^{k+1}, Inv)$
- $p(x_{k+1}|x_k)$, parte della dinamica
- $p(y_k|x_k)$, parte della misura

$$p(x_k|\widetilde{Y}^k, Inv) \propto \frac{p(x_k|\widetilde{Y}^k)}{p(x_k)} = \frac{p(x_k|y_k, \widetilde{Y}^{k+1})}{p(x_k)}$$

Applicando la regola di Bayes si ha che:

$$\begin{aligned}
&= \frac{p(y_k|x_k, \widetilde{Y}^{k+1})p(x_k|\widetilde{Y}^{k+1})}{p(y_k|\widetilde{Y}^{k+1})} \frac{1}{p(x_k)} \\
&= p(y_k|x_k)\frac{p(x_k|\widetilde{Y}^{k+1})}{p(x_k)}
\end{aligned}$$

si applica ora la regola di marginalizzazione su $p(x_k|\widetilde{Y}^{k+1})$ e si ha:

$$= \frac{p(y_k|x_k)}{p(x_k)} \int_{x_{k+1}} p(x_k, x_{k+1}, |\widetilde{Y}^{k+1})dx_{k+1}$$

si applica Bayes su $p(x_k, x_{k+1}, |\widetilde{Y}^{k+1})$

$$\begin{aligned}
&= \frac{p(y_k|x_k)}{p(x_k)} \int_{x_{k+1}} p(x_k|x_{k+1}, \widetilde{Y}^{k+1})p(x_{k+1}|\widetilde{Y}^{k+1})dx_{k+1} \\
&= \frac{p(y_k|x_k)}{p(x_k)} \int_{x_{k+1}} p(x_k|x_{k+1})p(x_{k+1}|\widetilde{Y}^{k+1})dx_{k+1}
\end{aligned}$$

quindi si applica Bayes su $p(x_k|x_{k+1})$

$$\begin{aligned}
&= \frac{p(y_k|x_k)}{p(x_k)} \int_{x_{k+1}} \frac{p(x_{k+1}|x_k)p(x_k)}{p(x_{k+1})}p(x_{k+1}|\widetilde{Y}^{k+1})dx_{k+1} \\
&= p(y_k|x_k) \int_{x_{k+1}} p(x_{k+1}|x_k)p(x_{k+1}|\widetilde{Y}^{k+1}, Inv)dx_{k+1}
\end{aligned}$$

È stata dunque ottenuta l'equazione di aggiornamento di $p(x_k|\widetilde{Y}^k, Inv)$ a partire dal suo valore successivo $p(x_{k+1}|\widetilde{Y}^{k+1}, Inv)$. Questa risulta essere speculare a quella ottenuta per il filtro di Kalman in avanti, e quindi possiamo applicare gli stessi ragionamenti:

$$p(x_{k+1}|x_k)p(x_{k+1}|\widetilde{Y}^{k+1}, Inv) \sim \mathcal{N}_{x_k}(\widetilde{x}_{k|k+1}, \widetilde{P}_{k|k+1})$$

Possiamo ora calcolare in maniera esplicita le equazioni di aggiornamento della media e varianza andando a sostituire i valori delle varie densità di probabilità coinvolte. Si ottiene:

$$\mathcal{N}_{x_k}(\widetilde{x}_{k|k+1}, \widetilde{P}_{k|k+1}) \propto \int_{x_{k+1}} \mathcal{N}_{x_{k+1}}(Ax_k, Q)\mathcal{N}_{x_k}(\widetilde{x}_{k+1|k+1}, \widetilde{P}_{k+1|k+1})dx_{k+1} \tag{10.1}$$

$$\propto \int_{x_{k+1}} \mathcal{N}_{\begin{bmatrix} x_k \\ x_{k+1} \end{bmatrix}}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}, \begin{bmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{bmatrix}\right)dx_{k+1} \tag{10.2}$$

$$\propto \mathcal{N}_{x_k}(\mu_1, \Sigma_{11}) \tag{10.3}$$

Si è dunque ottenuto:

$$\mathcal{N}_{x_k}(\widetilde{x}_{k|k+1}, \widetilde{P}_{k|k+1}) = \mathcal{N}_{x_k}(\mu_1, \Sigma_{11})$$

Imponendo ora l'uguaglianza tra gli esponenti (termini quadratici) si ha che:

$$\left(x_k - \widetilde{x}_{k|k+1}\right)^T \widetilde{P}_{k|k+1}^{-1} \left(x_k - \widetilde{x}_{k|k+1}\right)$$

deve essere uguale alla marginalizzazione rispetto a x_{k+1} della v.a. gaussiana definita dai seguenti termini quadratici:

$$\begin{aligned} &(x_{k+1} - Ax_k)^T Q^{-1} (x_{k+1} - Ax_k) + \left(x_{k+1} - \widetilde{x}_{k+1|k+1}\right)^T \widetilde{P}_{k+1|k+1}^{-1} \left(x_{k+1} - \widetilde{x}_{k+1|k+1}\right) \\ &= \begin{bmatrix} x_k - \mu_1 \\ x_{k+1} - \mu_2 \end{bmatrix}^T \begin{bmatrix} \Gamma_{11} & \Gamma_{12} \\ \Gamma_{21} & \Gamma_{22} \end{bmatrix} \begin{bmatrix} x_k - \mu_1 \\ x_{k+1} - \mu_2 \end{bmatrix} + cost \end{aligned}$$

con

$$\Gamma = \Sigma^{-1}$$

si ha quindi:

$$\begin{aligned} &x_{k+1}^T Q^{-1} x_{k+1} + x_k^T A^T Q^{-1} A x_k - 2x_k^T A^T Q^{-1} x_{k+1} + x_{k+1}^T \widetilde{P}_{k+1|k+1}^{-1} x_{k+1} + \widetilde{x}_{k+1|k+1}^T \widetilde{P}_{k+1|k+1}^{-1} \widetilde{x}_{k+1|k+1} - \\ &- 2x_{k+1}^T P_{k+1|k+1}^{-1} \widetilde{x}_{k+1|k+1} = \\ &x_k^T \Gamma_{11} x_k + \mu_1^T \Gamma_{11} \mu_1 - 2x_k^T \Gamma_{11} \mu_1 + 2x_k^T \Gamma_{12} x_{k+1} - 2\mu_1^T \Gamma_{12} x_{k+1} - 2x_k^T \Gamma_{12} \mu_2 + 2\mu_1^T \Gamma_{12} \mu_2 + \\ &+ x_{k+1}^T \Gamma_{22} x_{k+1} + \mu_2^T \Gamma_{22} \mu_2 - 2x_{k+1}^T \Gamma_{22} \mu_2 \end{aligned}$$

risulta che:

$$\begin{aligned} \Gamma_{11} &= A^T Q^{-1} A \\ \Gamma_{22} &= Q^{-1} + \widetilde{P}_{k+1|k+1}^{-1} \\ \Gamma_{12} &= \Gamma_{21}^T = A^T Q^{-1} \end{aligned}$$

ed:

$$\begin{aligned}
\Gamma_{11}\mu_1 + \Gamma_{12}\mu_2 = 0 \qquad &\Rightarrow \qquad \mu_1 = -\Gamma_{11}^{-1}\Gamma_{12}\mu_2 \\
\Gamma_{22}\mu_2 + \Gamma_{21}\mu_1 = \widetilde{P}_{k+1|k+1}^{-1}\widetilde{x}_{k+1|k+1} \qquad &\Rightarrow \qquad (\Gamma_{22} - \Gamma_{21}\Gamma_{11}^{-1}\Gamma_{12})\mu_2 = \widetilde{P}_{k+1|k+1}^{-1}\widetilde{x}_{k+1|k+1}
\end{aligned}$$

quindi supponendo inoltre che A sia invertibile:

$$(Q^{-1} + \widetilde{P}_{k+1|k+1}^{-1} - Q^{-1}A(A^TQ^{-1}A)^{-1}A^TQ^{-1})\mu_2 = \widetilde{P}_{k+1|k+1}^{-1}\widetilde{x}_{k+1|k+1} \qquad \Rightarrow \qquad \mu_2 = \widetilde{x}_{k+1|k+1}$$

Infine:

$$\begin{aligned}
\mu_1 &= (A^TQ^{-1}A)^{-1}(A^TQ^{-1})\mu_2 \\
&= (A^TQ^{-1}A)^{-1}(A^TQ^{-1})\widetilde{x}_{k+1|k+1} \\
&= A^{-1}\widetilde{x}_{k+1|k+1}
\end{aligned}$$

allora

$$\Sigma = \begin{bmatrix} A^{-1}\widetilde{P}_{k+1|k+1}(A^T)^{-1} + A^{-1}Q(A^T)^{-1} & A^{-1}\widetilde{P}_{k+1|k+1} \\ \widetilde{P}_{k+1|k+1}(A^T)^{-1} & \widetilde{P}_{k+1|k+1} \end{bmatrix}$$

quindi risulta in fine che:

$$\mathcal{N}_{x_k}(\mu_1, \Sigma_{11}) = \mathcal{N}_{x_k}(A^{-1}\widetilde{x}_{k+1|k+1}, A^{-1}\widetilde{P}_{k+1|k+1}(A^T)^{-1} + A^{-1}Q(A^T)^{-1})$$

Analizzando il sistema, si può pensare di invertirne la dinamica predicendo in questo modo il passato dal futuro

$$\begin{aligned}
x_{k+1} &= Ax_k + w_k \\
x_k &= A^{-1}(x_{k+1} - w_k) = A^{-1}x_{k+1} + A^{-1}w_k = A^{-1}x_{k+1} + \widetilde{w}_k \qquad \widetilde{w}_k \sim \mathcal{N}(0, A^{-1}Q(A^{-1})^T)
\end{aligned}$$

Invertendo la dinamica praticamente si dovrebbe andare a sostituire ad A e Q rispettivamente A^{-1} e $A^{-1}Q(A^{-1})^T$). Rimane ora solo da calcolare l'inizializzazione delle equazione. Partendo dalla definizione si ha che

$$\mathcal{N}_{x_T}(\widetilde{x}_{T|T}, \widetilde{P}_{T|T}) = p(x_T|\widetilde{Y}^T, Inv) = p(x_T|y_T, Inv) \propto \frac{p(x_T|y_T)}{p(x_T)} = \frac{p(y_T|x_T)p(x_T)}{p(y_T)p(x_T)} \propto p(y_T|x_T)$$

Da questo si ricava con semplici manipolazioni che

$$p(y_T|x_T) = \mathcal{N}_{y_T}(Cx_T, R) \propto \mathcal{N}_{x_T}(C^TR^{-1}y_T, (C^TR^{-1}C)^{-1})$$

Ricapitolando si hanno dunque le seguenti formule:

$$\begin{aligned}
\hat{x}_{k+1|k} &= A\hat{x}_{k|k} \\
P_{k+1|k} &= AP_{k|k}A^T + Q \\
P_{k+1|k+1}^{-1} &= P_{k+1|k}^{-1} + C^TR^{-1}C \\
\hat{x}_{k+1|k+1} &= P_{k+1|k+1}(P_{k+1|k}^{-1}\hat{x}_{k+1|k} + C^TR^{-1}y_{k+1}) \\
P_{0|-1} &= P_0 \qquad \hat{x}_{0|-1} = \bar{x}_0
\end{aligned}$$

$$
\begin{aligned}
\widetilde{x}_{k|k+1} &= A^{-1}\widetilde{x}_{k+1|k+1} \\
\widetilde{P}_{k|k+1} &= A^{-1}\widetilde{P}_{k+1|k+1}(A^T)^{-1} + A^{-1}Q(A^T)^{-1} \\
\widetilde{P}^{-1}_{k|k} &= P^{-1}_{k|k+1} + C^T R^{-1} C \\
\widetilde{x}_{k|k} &= \widetilde{P}_{k|k}(\widetilde{P}^{-1}_{k|k+1}\widetilde{x}_{k|k+1} + C^T R^{-1} y_k) \\
\widetilde{x}_{T|T} &= C^T R^{-1} y_T \\
\widetilde{P}^{-1}_{T|T} &= C^T R^{-1} C
\end{aligned}
$$

allora, unendo i due pezzi (fusione convessa delle due parti indipendenti di cui si è discusso all'inizio) si ha

$$
\begin{aligned}
P^{-1}_{k|T} &= P^{-1}_{k|k-1} + \widetilde{P}^{-1}_{k|k} \\
&= P^{-1}_{k|k} + \widetilde{P}^{-1}_{k|k+1} \\
&= P^{-1}_{k|k+1} + P^{-1}_{k|k-1} + C^T R^{-1} C
\end{aligned}
$$

e

$$
\begin{aligned}
\hat{x}_{k|T} &= P_{k|T}(P^{-1}_{k|k-1}\hat{x}_{k|k-1} + \widetilde{P}^{-1}_{k|k}\widetilde{x}_{k|k}) \\
&= P_{k|T}(P^{-1}_{k|k}\hat{x}_{k|k} + \widetilde{P}^{-1}_{k|k+1}\widetilde{x}_{k|k+1}) \\
&= P_{k|T}(P^{-1}_{k|k-1}\hat{x}_{k|k-1} + \widetilde{P}^{-1}_{k|k+1}\widetilde{x}_{k|k+1} + C^T R^{-1} y_k)
\end{aligned}
$$