PSC: Progettazione di sistemi di controllo	a.a. 2010-2011
Lezione 2 — 15 Ottobre 2010	
Docente: Luca Schenato	*Stesori: Schenato*

2.1 Il filtro di Kalman: derivazione delle equazioni

Si consideri il modello stocastico lineare tempo invariante:

$$\begin{cases} x_{k+1} &= Ax_k + w_k \\ y_k &= Cx_k + v_k \end{cases} \tag{2.1}$$

dove:

$$\begin{cases} v_k &\sim \mathcal{N}(0,R), \quad \mathbb{E}[v_k v_h] = R\delta(k-h) \\ w_k &\sim \mathcal{N}(0,Q), \quad \mathbb{E}[w_k w_h] = Q\delta(k-h) \\ x_0 &\sim \mathcal{N}(\bar{x}_0, P_0) \end{cases} \tag{2.2}$$

e v_k, w_k, x_0 sono v.a. gaussiane[1] a media nulla, incorrelate fra di loro.

Il filtro di Kalman è definito come:

$$\hat{x}_{k+1|k+1} = \mathbb{E}\left[x_{k+1}|y_0,\ldots,y_{k+1}\right] = \mathbb{E}\left[x_{k+1}|y_{k+1},Y^k\right]$$

dove $Y^k = (y_k,\ldots,y_1,y_0)$. Vogliamo trovarne un'espressione ricorsiva che ne faciliti il calcolo ad ogni passo.

L'espressione esplicita di $\mathbb{E}[X|Y]$ è facilmente calcolabile se X e Y sono due variabili aleatorie congiuntamente gaussiane e si conoscono le espressioni di μ_X, μ_Y e $\begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$.

Nel nostro caso consideriamo $X = x_{k+1}|Y^k$ e $Y = y_{k+1}|Y^k$; è, quindi, necessario calcolare le espressioni esplicite di:

$$\mu_X = \mathbb{E}\left[x_{k+1}|Y^k\right] \tag{2.3}$$

$$\mu_Y = \mathbb{E}\left[y_{k+1}|Y^k\right] \tag{2.4}$$

$$P_{k+1|k} = \Sigma_{XX} = Var\left[x_{k+1}|Y^k\right] \tag{2.5}$$

$$\Sigma_{YY} = Var\left[y_{k+1}|Y^k\right] \tag{2.6}$$

$$\Sigma_{XY} = \Sigma_{YX}^T = Cov\left[x_{k+1}, y_{k+1}|Y^k\right]. \tag{2.7}$$

da esse dedurremo l'espressione dello stimatore ottimo come:

$$\mathbb{E}[X|Y] = \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \Sigma_{XY}\Sigma_{YY}^{-1}\left(y_{k+1} - \hat{y}_{k+1|k}\right) \tag{2.8}$$

[1] Si ricordi che il simbolo $\mathcal{N}$ indica una v.a. gaussiana

e la varianza dell'errore di stima come:

$$\Sigma_{X|Y} = P_{k+1|k+1} = \Sigma_{XX} - \Sigma_{XY}\Sigma_{YY}^{-1}\Sigma_{YX} \tag{2.9}$$

Calcoliamo in forma esplicita i termini 2.3, 2.4, 2.5, 2.6 e 2.7 ricordando che le medie di prodotti di termini incorrelati sono nulle.

$$\mathbb{E}\left[x_{k+1}|Y^k\right] = \mathbb{E}\left[Ax_k + w_k|Y^k\right] = A\mathbb{E}\left[x_k|Y^k\right] + \mathbb{E}\left[w_k|Y^k\right] = A\hat{x}_{k|k} = \hat{x}_{k+1|k}$$

$$\mathbb{E}\left[y_{k+1}|Y^k\right] = \mathbb{E}\left[Cx_{k+1} + v_{k+1}|Y^k\right] = C\mathbb{E}\left[x_{k+1}|Y^k\right] + \mathbb{E}\left[v_{k+1}|Y^k\right] = C\hat{x}_{k+1|k} = CA\hat{x}_{k|k}$$

$$\begin{aligned}
Var\left[x_{k+1}|Y^k\right] &= \mathbb{E}\left[\left(x_{k+1} - \hat{x}_{k+1|k}\right)\left(x_{k+1} - \hat{x}_{k+1|k}\right)^T|Y^k\right] = P_{k+1|k} \\
&= \mathbb{E}\left[\left(Ax_k + w_k - A\hat{x}_{k|k}\right)\left(Ax_k + w_k - A\hat{x}_{k|k}\right)^T|Y^k\right] \\
&= A\mathbb{E}\left[\left(x_k - \hat{x}_{k|k}\right)\left(x_k - \hat{x}_{k|k}\right)^T|Y^k\right]A^T + A\mathbb{E}\left[\left(x_k - \hat{x}_{k|k}\right)w_k^T|Y^k\right] \\
&\quad + \mathbb{E}\left[w_k\left(x_k - \hat{x}_{k|k}\right)^T|Y^k\right]A^T + \mathbb{E}\left[w_k w_k^T|Y^k\right] \\
&= AP_{k|k}A^T + Q
\end{aligned}$$

$$\begin{aligned}
Var\left[y_{k+1}|Y^k\right] &= \mathbb{E}\left[\left(y_{k+1} - \hat{y}_{k+1|k}\right)\left(y_{k+1} - \hat{y}_{k+1|k}\right)^T|Y^k\right] \\
&= \mathbb{E}\left[\left(Cx_{k+1} + v_{k+1} - C\hat{x}_{k+1|k}\right)\left(Cx_{k+1} + v_{k+1} - C\hat{x}_{k+1|k}\right)^T|Y^k\right] \\
&= C\mathbb{E}\left[\left(x_{k+1} - \hat{x}_{k+1|k}\right)\left(x_{k+1} - \hat{x}_{k+1|k}\right)^T|Y^k\right]C^T + C\mathbb{E}\left[\left(x_{k+1} - \hat{x}_{k+1|k}\right)v_{k+1}^T|Y^k\right] \\
&\quad + \mathbb{E}\left[v_{k+1}\left(x_{k+1} - \hat{x}_{k+1|k}\right)^T|Y^k\right]C^T + \mathbb{E}\left[v_{k+1}v_{k+1}^T|Y^k\right] \\
&= CP_{k+1|k}C^T + R
\end{aligned}$$

$$\begin{aligned}
Cov\left[x_{k+1}, y_{k+1}|Y^k\right] &= \mathbb{E}\left[\left(x_{k+1} - \hat{x}_{k+1|k}\right)\left(y_{k+1} - \hat{y}_{k+1|k}\right)^T|Y^k\right] \\
&= \mathbb{E}\left[\left(Ax_k - A\hat{x}_{k|k} + w_k\right)\left(CAx_k - CA\hat{x}_{k|k} + v_{k+1} + Cw_k\right)^T|Y^k\right] \\
&= A\mathbb{E}\left[\left(x_k - \hat{x}_{k|k}\right)\left(x_k - \hat{x}_{k|k}\right)^T|Y^k\right]A^TC^T + \mathbb{E}\left[w_k w_k^T|Y^k\right]C^T \\
&= AP_{k|k}A^TC^T + QC^T \\
&= P_{k+1|k}C^T
\end{aligned}$$

Si deduce che, definito $z = \begin{bmatrix} x_{k+1} \\ y_{k+1} \end{bmatrix}$, si ha:

$$p\left(z|Y^k\right) \sim \mathcal{N}\left(\begin{bmatrix} \hat{x}_{k+1|k} \\ C\hat{x}_{k+1|k} \end{bmatrix}, \begin{bmatrix} P_{k+1|k} & P_{k+1|k}C^T \\ CP_{k+1|k} & CP_{k+1|k}C^T + R \end{bmatrix}\right)$$

con:

$$\hat{x}_{k+1|k} = A\hat{x}_{k|k}$$
$$P_{k+1|k} = AP_{k|k}A^T + Q$$

Calcoliamo quindi:

$$p\left(x_{k+1}|Y^{k+1}\right) \sim \mathcal{N}\left(\hat{x}_{k+1|k+1}, P_{k+1|k+1}\right)$$

Sostituendo i termini calcolati nelle espressioni 2.8 e 2.9 si ottiene per lo stimatore ottimo:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1}\left(y_{k+1} - C\hat{x}_{k+1|k}\right)$$

che posto:

$$K_{k+1|k+1} = P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1}$$

diviene:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1|k+1}\left(y_{k+1} - C\hat{x}_{k+1|k}\right)$$

dove $K_{k+1|k+1}$ è detto guadagno di Kalman. E per quanto riguarda l'errore di stima:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1}CP_{k+1|k}$$

Riassumendo le equazioni ricorsive del filtro di Kalman risultano essere:

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1|k+1}\left(y_{k+1} - C\hat{x}_{k+1|k}\right) \quad (2.10)$$
$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1}CP_{k+1|k} \quad (2.11)$$
$$K_{k+1|k+1} = P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{-1} \quad (2.12)$$

con le condizioni iniziali:

$$\hat{x}_{0|-1} = \mathbb{E}\left[x_0\right] \qquad P_{0|-1} = Var\left[x_0\right]. \quad (2.13)$$

Nel caso in cui la varianza dell'errore sia semplicemente semidefinita positiva, cioe' $R \geq 0$, le equazioni del filtro di Kalman possono essere generalizzate semplicemente sostituendo l'operazione di inversa con quella di pseudoinversa:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{\dagger}CP_{k+1|k} \quad (2.14)$$
$$K_{k+1|k+1} = P_{k+1|k}C^T\left(CP_{k+1|k}C^T + R\right)^{\dagger} \quad (2.15)$$

2.2 Il filtro di Kalman: ricapitolazione

Dalle definizioni considerate precedentemente e riscritte qui per convenienza:

$$\begin{aligned} \hat{x}_{k|h} &\triangleq \mathbb{E}[x_k | y_h, \ldots, y_0, x_0] \\ P_{k|h} &\triangleq \mathbb{E}[(x_k - \hat{x}_{k|h})(x_k - \hat{x}_{k|h})^T | y_h, \ldots, y_0, x_0] \end{aligned} \tag{2.16}$$

e considerando che $\hat{x}_{k|h}$ è anch'essa una variabile aleatoria Gaussiana, si possono ricavare le seguenti equazioni del filtro di Kalman, che rappresentano un metodo ricorsivo per il calcolo della stima ottima:

$$\begin{aligned}
&\left.\begin{aligned} \hat{x}_{k+1|k} &= A\hat{x}_{k|k} \\ P_{k+1|k} &= AP_{k|k}A^T + Q \end{aligned}\right\} \text{predizione} \\
&\left.\begin{aligned} \hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + K_{k+1}(y_{k+1} - C\hat{x}_{k+1|k}) \\ P_{k+1|k+1} &= P_{k+1|k} - P_{k+1|k}C^T(CP_{k+1|k}C^T + R)^{-1}CP_{k+1|k} \end{aligned}\right\} \text{aggiornamento} \\
&\left. K_{k+1} \triangleq P_{k+1|k}C^T(CP_{k+1|k}C^T + R)^{-1} \right\} \text{guadagno di Kalman} \\
&\left.\begin{aligned} \hat{x}_{0|-1} &= \bar{x}_0 \\ P_{0|-1} &= P_0 \end{aligned}\right\} \text{inizializzazione}
\end{aligned} \tag{2.17}$$

Riassumiamo le ipotesi sotto cui sono valide le equazioni del filtro (2.17):

- gaussianità di w_k e v_k
- bianchezza del rumore: $\mathbb{E}[w_k w_h^T] = Q\delta(k-h), \mathbb{E}[v_k v_h^T] = R\delta(k-h)$ (da cui si ricava che $\mathbb{E}[(x_k - \hat{x}_{k|k})w_k^T | y_h, \ldots, y_0] = 0$ e $\mathbb{E}[(x_k - \hat{x}_{k|k-1})v_k^T | y_h, \ldots, y_0] = 0$)
- w_k e v_k a media nulla
- incorrelazione tra v_k e w_k, cioe' $\mathbb{E}[w_k v_h^T] = 0, \forall h, k$
- $R > 0$, $Q \geq 0$

Osservazioni:

Si ritiene utile a questo riassumere alcune importanti osservazioni sulle equazioni del filtro di Kalman scritte.

1. Il processo non deve necessariamente essere stazionario e puo' essere generalizzato per un sistema tempo variante semplicemente apportando la seguente sostituzione: $A \leftarrow A_k$; $C \leftarrow C_{k+1}$; $R \leftarrow R_{k+1}$; $Q \leftarrow Q_k$.

2. L'ipotesi di incorrelazione tra v_k e w_k non è necessaria, e le equazioni possono essere modificate in maniera opportuna per tenere conto del fatto che $\mathbb{E}[v_k w_k^T] = S \neq 0$ (vedi equazioni in [1], pag. 290).

3. La condizione di positivita' di R puo' essere rilassata a semidefinita-positiva ($R \geq 0$) semplicemente sostituendo nelle equazioni la pseudoinversa al posto dell'inversa.

4. L'ipotesi di media nulla non e' fondamentale. Infatti, se consideriamo $w_k = \mathcal{N}(\mu_Q, Q)$, $v_k = \mathcal{N}(\mu_R, R)$ e' sufficiente considerare modificare le stesse equazioni del filtro di Kalman dove il sistema dinamico diventa $x_{k+1} = Ax_k + \mu_Q + \tilde{w}_k, \quad y_k = Cx_k + \mu_R + \tilde{v}_k$, dove $\tilde{w}_k = \mathcal{N}(0, Q), \tilde{v}_k = \mathcal{N}(0, R)$

5. La stima è lineare nelle misure. Infatti possiamo scrivere l'equazione di aggiornamento del filtro come $\hat{x}_{k+1|k+1} = (I - K_{k+1}C)\hat{x}_{k+1|k} + K_{k+1}y_{k+1}$

6. $\mathbb{E}[x_k|y_k, \ldots, y_0] = \mathbb{E}[x_k|\hat{x}_{k-1|k-1}, y_k]$ cioè tutta l'informazione dall'istante 0 all'istante $(k-1)$ è contenuta nella stima all'istante $(k-1)$.

7. Il guadagno ottimo, $K = K_k$, è tempo variante, anche nel caso in cui il sistema sia tempo invariante.

8. una ipotesi che non puo' essere rilassata e' la bianchezza del segnale. Nel caso in cui w_k, v_k siano *rumori colorati*, cioe' $[v_k v_h^T] = R_{h-k}, [w_k w_h^T] = Q_{h-k}$, le precedenti equazioni non possono essere generalizzate. In genere, si cerca di ottenere w_k, v_k come l'uscita di un sistema lineare stabile (filtro passa-basso) che ha come ingresso del rumore bianco $\tilde{w}_k, \tilde{v}_k$. Si tratta quindi di identificare una rappresentazione di tale rumore con un sistema del tipo $w_{k+1} = Fw_k + \tilde{w}_k$, cioe' matrice F e varianza di $\tilde{w}_k$ a partire dallo spettro R_{h-k}, Q_{h-k}.

9. Un'altra ipotesi che non puo' essere rilassata e' la gaussianita' dei rumori. Nel caso questi non siano gaussiani le precedenti equazioni non sono corrette. Riguardo a questa ipotesi si vedano anche le considerazioni riguardo il filtro con guadagno costante nella prossima sezione.

2.3 Filtro a guadagno costante

Osservando le equazioni (2.17), si può pensare di sostituire a K_k un guadagno costante K ovvero di introdurre un nuovo stimatore non più ottimo ma simile al precedente:

$$\hat{g}(y) \triangleq \hat{x}_{k+1|k+1} = A\hat{x}_{k|k} + K_{k+1}(y_{k+1} - CA\hat{x}_{k|k}) \tag{2.18}$$

Si definisce allora:

$$\tilde{g}(y) \triangleq \tilde{x}_{k+1|k+1} = A\tilde{x}_{k|k} + K(y_{k+1} - CA\tilde{x}_{k|k}) \tag{2.19}$$

Come nel caso del filtro di Kalman si avrà:

$$\begin{aligned}
\tilde{x}_{k+1|k} &= A\tilde{x}_{k|k} && (2.20)\\
\tilde{P}_{k|k} &= \mathbb{E}[(x_k - \tilde{x}_{k|k})(x_k - \tilde{x}_{k|k})^T] && (2.21)\\
\tilde{P}_{k+1|k} &= \mathbb{E}[(x_{k+1} - \tilde{x}_{k+1|k})(x_{k+1} - \tilde{x}_{k+1|k})^T] && (2.22)
\end{aligned}$$

Da (2.19) e (2.20) è facile ricavare:

$$\begin{aligned}\tilde{x}_{k+1|k} &= A(\underbrace{A\tilde{x}_{k-1|k-1}}_{\tilde{x}_{k|k-1}} + K(y_k - CA\tilde{x}_{k-1|k-1})) \\ &= A\tilde{x}_{k|k-1} + AK(y_k - C\tilde{x}_{k|k-1})\end{aligned} \tag{2.23}$$

da cui si ottiene la seguente espressione per l'errore di stima del filtro appena introdotto:

$$\begin{aligned}\tilde{e}_{k+1|k} &\triangleq x_{k+1} - \tilde{x}_{k+1|k} \\ &= Ax_k + w_k - (A\tilde{x}_{k|k-1} + AK(Cx_k + v_k - C\tilde{x}_{k|k-1})) \\ &= A(I - KC)\tilde{e}_{k|k-1} + w_k - AKv_k\end{aligned} \tag{2.24}$$

Per ipotesi dall'incorrelazione dei rumore w_k, v_k, si deduce che anche $\tilde{e}_{k|k-1}$ è incorrelato sia da w_k che da v_k, in quanto $\tilde{e}_{k|k-1} = f(x_0, w_0, \ldots, w_{k-1}, v_0, \ldots, v_{k-1})$, cioe' dipende solo dai rumori passati.

È evidente che la scelta del guadagno K sia un compromesso tra stabilità del sistema (l'errore non deve divergere) e riduzione del rumore v_k, cioe' si vuole che $A(I - KC)$ sia piccolo, ma allo stesso tempo non si vuole scegliere K troppo grande per non amplificare troppo il rumore che entra nella dinamica dell'errore con il termine AKv_k.

Dalle (2.22) e (2.24), ricordando l'ipotesi di incorrelazione tra v_k, w_k e $\tilde{e}_{k|k-1}$, si ricava:

$$\begin{aligned}\tilde{P}_{k+1|k} &= \mathbb{E}[\tilde{e}_{k+1|k}\tilde{e}^T_{k+1|k}] \\ &= \mathbb{E}[A(I - KC)\tilde{e}_{k|k-1}\tilde{e}^T_{k|k-1}(I - KC)^T A^T] + \mathbb{E}[w_k w_k^T] + \mathbb{E}[AKv_k v_k^T K^T A^T] \\ &= A(I - KC)\tilde{P}_{k|k-1}(I - KC)^T A^T + Q + AKRK^T A^T\end{aligned} \tag{2.25}$$

con, al solito:

$$\tilde{P}_{0|-1} = P_0 \tag{2.26}$$

Il filtro così introdotto non richiede inversioni di matrici e la memorizzazione di $P_{k|k-1}$, tuttavia ciò si paga con un peggioramento delle prestazioni. Infatti non essendo ottimo il filtro con guadagno statico necessariamente dobbiamo avere:

$$\tilde{P}_{k|k-1} \geq P_{k|k-1} \ \forall k \tag{2.27}$$

Dimostreremo comunque che esiste un opportuno valore del guadagno $K = K^{opt}$ tale che:

$$\begin{cases} \tilde{P}_{k|k-1} & \rightarrow & \tilde{P}_\infty \\ P_{k|k-1} & \rightarrow & P_\infty \\ \tilde{P}_\infty & = & P_\infty \end{cases} \tag{2.28}$$

cioè a regime le covarianze degli errori di stima dei due filtri coincidono. Ciò significa che, terminato il transitorio iniziale, si deve avere:

$$\tilde{x}_{k+1|k} \rightarrow \hat{x}_{k+1|k} \tag{2.29}$$

e quindi i due filtri hanno comportamento identico.

Bibliografia

[1] Giorgio Picci. *Fitraggio Statistico (Wiener, Levinson, Kalman) e Applicazioni*. Libreria Progetto, 2006.