PSC: Progettazione di sistemi di controllo	**a.a. 2010-2011**
Lezione 3 — Ottobre 19	
Docente: Luca Schenato	*Stesori: Cuccato Davide, De Franceschi Igor, Sartor Giorgio*

3.1 Filtro di Kalman in forma di informazione o di varianza inversa

Il filtro di Kalman in forma di informazione é una manipolazione delle consuete equazioni del filtro al fine di dar loro una forma particolarmente utile dal punto di vista implementativo, in quanto richiede l'inversione di una matrice $n \times n$ e non $m \times m$; in numerose applicazioni si ha $m > n$ e dunque la complessitá di calcolo risulta diminuita. Tale manipolazione si basa sul ben noto *lemma di inversione di matrice*, di seguito riportato, la cui dimostrazione puó essere trovata in un qualsiasi testo di algebra lineare:

Lemma 3.1. *Siano A e D due matrici quadrate invertibili; siano U e V altre due matrici di dimensioni opportune. Vale:*

$$(A + UDV)^{-1} = A^{-1} - A^{-1}U(D^{-1} + VA^{-1}U)^{-1}VA^{-1} \tag{3.1}$$

Si ricordino le equazioni relative al passo di aggiornamento del Filtro di Kalman:

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T \left(CP_{k+1|k}C^T + R\right)^{-1} CP_{k+1|k} \tag{3.2}$$

$$\hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + P_{k+1|k}C^T \left(CP_{k+1|k}C^T + R\right)^{-1} \left(y_{k+1} - C\hat{x}_{k+1|k}\right) \tag{3.3}$$

Per quel che riguarda la prima equazione relativa alla varianza di errore si ricava subito che la parte a destra dell' Equazione (3.1) coincide con l'Equazione (3.2) tramite la sostituzione $A^{-1} = P_{k+1|k}, U = C, D^{-1} = R, V = C^T$, da cui di ricava immediatamente:

$$P_{k+1|k+1} = (P_{k+1|k}^{-1} + C^T R^{-1} C)^{-1} \tag{3.4}$$

Per quel che rigaurda l'Equazione (3.3) di aggiornamento dello stato invece si trova:

$$
\begin{aligned}
\hat{x}_{k+1|k+1} &= \hat{x}_{k+1|k} + P_{k+1|k}C^T\left(CP_{k+1|k}C^T+R\right)^{-1}\left(y_{k+1}-C\hat{x}_{k+1|k}\right)\\
&= \left(I - P_{k+1|k}C^T\left(CP_{k+1|k}C^T+R\right)^{-1}C\right)\right)\hat{x}_{k+1|k} + P_{k+1|k}C^T\left(CP_{k+1|k}C^T+R\right)^{-1}y_{k+1}\\
&= \left(I - P_{k+1|k}C^T\left(CP_{k+1|k}C^T+R\right)^{-1}C\right)\right)P_{k+1|k}P^{-1}_{k+1|k}\hat{x}_{k+1|k}+\\
&\quad +P_{k+1|k}C^T\left(R^{-1}-R^{-1}C(P^{-1}_{k+1|k}+C^TR^{-1}C)^{-1}C^TR^{-1}\right)y_{k+1}\\
&= \left(P_{k+1|k} - P_{k+1|k}C^T\left(CP_{k+1|k}C^T+R\right)^{-1}CP_{k+1|k}\right)\right)P^{-1}_{k+1|k}\hat{x}_{k+1|k}+\\
&\quad +\left(P_{k+1|k}-P_{k+1|k}C^TR^{-1}C(P^{-1}_{k+1|k}+C^TR^{-1}C)^{-1}\right)C^TR^{-1}y_{k+1}\\
&= P_{k+1|k+1}P^{-1}_{k+1|k}\hat{x}_{k+1|k}+\\
&\quad +\left(P_{k+1|k}(P^{-1}_{k+1|k}+C^TR^{-1}C)-P_{k+1|k}C^TR^{-1}C\right)(P^{-1}_{k+1|k}+C^TR^{-1}C)^{-1}C^TR^{-1}y_{k+1}\\
&= P_{k+1|k+1}P^{-1}_{k+1|k}\hat{x}_{k+1|k}+P_{k+1|k+1}C^TR^{-1}y_{k+1}
\end{aligned}
$$

da cui di deduce che

$$\hat{x}_{k+1|k+1} = P_{k+1|k+1}(P^{-1}_{k+1|k}\hat{x}_{k+1|k}+C^TR^{-1}y_{k+1}) \tag{3.5}$$

OSSERVAZIONI:

- Nel Filtro di Kalman classico, le equazioni di aggiornamento (3.2)-(3.3) richiedono l'inversione **ad ogni passo** della matrice $\left(CP_{k+1|k}C^T+R\right)\in\mathbb{R}^{m\times m}$ che ha rango uguale alla dimensione del vettore di misura y, mentre le equazioni del Filtro di Kalman in forma di informazione (3.4)-(3.5) richiedono l'inversione di matrici $P_{k+1|k+1}\in\mathbb{R}^{n\times n}$ che anno la dimensione dello stato x. Quindi da un punto di vista computazionale e' possibile scegliere la forma piu' appropriata e quindi la complessitá del filtro di Kalman e' $O(\min(m^3,n^3))$, in quanto la complessità dell'inversione di una matrice è dell'ordine del cubo della sua dimensione.
- Si definisca $z_{k|h}\triangleq P^{-1}_{k|h}\hat{x}_{k|h}$ $\quad Z_{k|h}\triangleq P^{-1}_{k|h}$ $\quad i_k\triangleq C^TR^{-1}y_k$, dove i_k prende il nome di **informazione** relativa alla misura k-esima. Le equazioni del filtro di Kalman relative all'aggiornamento possono essere riscritte come

$$z_{k+1|k+1} = z_{k+1|k} + i_{k+1} \tag{3.6}$$
$$Z_{k+1|k+1} = Z_{k+1|k} + C^TR^{-1}C \tag{3.7}$$

  Si noti che $Z_{k|h}$ é l'inversa di una varianza, quindi piu' grande è tale matrice, minore è l'errore di stima. Inoltre le equazioni del predittore in questo nuovo spazio di stato sono date da

$$z_{k+1|k} = Z_{k+1|k}AZ^{-1}_{k|k}z_{k|k} \tag{3.8}$$
$$Z_{k+1|k} = (AZ^{-1}_{k|k}A^T+Q)^{-1} \tag{3.9}$$

  che si ottengono a partire da $\hat{x}_{k+1|k} = A\hat{x}_{k|k}$ e $P_{k+1|k} = AP_{k|k}A^T+Q$ e dalle definizioni di $z_{k|h}$ e $Z_{k|h}$.

Tali equazioni, ottenute da quelle canoniche di Kalman applicando le posizioni sopra definite, prendono il nome di **Filtro di Kalman in forma d'informazione**.

ESEMPIO 1 Consideriamo il caso:

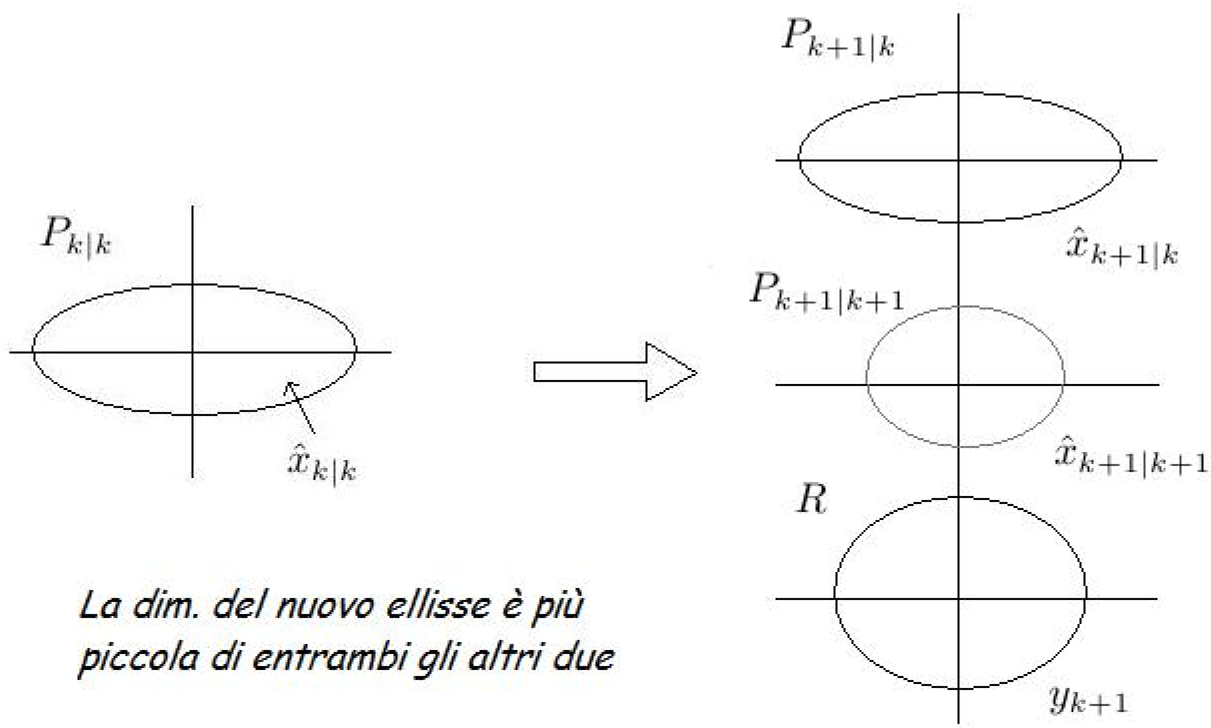
$$\begin{cases} x_{k+1} &= Ax_k + w_k \\ y_k &= x_k + v_k \end{cases}$$

dove si puó interpretare y_k come una misura rumorosa dello stato.

Si ha che:

$$\hat{x}_{k+1|k+1} = P_{k+1|k+1}(P_{k+1|k}^{-1}\hat{x}_{k+1|k} + R^{-1}y_{k+1})$$

dunque l'uscita del filtro risulta essere la somma pesata della stima precedente e delle misure; se il rumore $v(\cdot)$ ha varianza elevata, ovvero le misure sono molto rumorose, la matrice di peso R^{-1} risulterá composta da elementi piccoli rispetto a $P_{k+1|k}^{-1}$, quindi le misure conteranno poco rispetto alla stima. Se viceversa é l'errore di stima ad avere varianza elevata rispetto al rumore di misura, saranno le stime ad avere peso minore nella somma.



Consideriamo ora il caso di una matrice C generica di dimensioni $m \times n$ dove per $rank(C) = n < m$. Questa condizione serve a garantire che la matrice $C^T R^{-1} C$ sia invertibile. Allora:

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

É quindi possibile interpretare $\hat{x}^y_{k+1} = (C^T R^{-1} C)^{-1} C^T R^{-1} y_{k+1}$ come una stima dello stato basata sulla misura all'istante k, con $(P^y_{k+1})^{-1} = C^T R^{-1} C$ la relativa inversa della matrice di varianza d'errore, in corrispondenza a quanto fatto nel caso precedente.

Infine consideriamo il caso in cui $n > m$, e quindi $C^T R^{-1} C$ non é piú invertibile; sia, ad esempio:

$$C = \begin{bmatrix} 1 & 0 \end{bmatrix}, R = r,\ y_k \in \mathbb{R}$$

percui:

$$C^T R^{-1} C = \begin{bmatrix} r^{-1} & 0 \\ 0 & 0 \end{bmatrix} = (P^y_{k+1})^{-1}$$

Tale matrice non é invertibile ma é possibile fare un ragionamento intuitivo. Assumiano che anche la seconda componente dello stato sia direttamente osservabile, ma affetta da un rumore molto elevato, cioè:

$$\bar{y}_{k+1} = \begin{bmatrix} y_{k+1} \\ \xi_{k+1} \end{bmatrix},\ \bar{C} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \bar{R} = \begin{bmatrix} r & 0 \\ 0 & \sigma^2 \end{bmatrix},\quad \sigma^2 \gg r$$

dove ξ_{k+1} rappresenta la misura fittizia. Da un punto di vista pratico i due modelli dovrebbero dare risultati molto simili in quanto il secondo sensore non apporta informazione essendo affetto da rumore molto elevato. Infatti in questo caso:

$$\bar{C}^T \bar{R}^{-1} \bar{C} = \begin{bmatrix} r^{-1} & 0 \\ 0 & \sigma^{-2} \end{bmatrix} \approx \begin{bmatrix} r^{-1} & 0 \\ 0 & 0 \end{bmatrix} = (P^y_{k+1})^{-1}$$

Nel momento in cui viene fatto l'aggiornamento dello stato con questo modello si ha:

$$\bar{i}_{k+1} = \bar{C}^T \bar{R}^{-1} \bar{y}_{k+1} = \begin{bmatrix} r^{-1} y_{k+1} \\ \sigma^{-2} \xi_{k+1} \end{bmatrix} \approx \begin{bmatrix} r^{-1} y_{k+1} \\ 0 \end{bmatrix} = C^T R^{-1} y_{k+1} = i_{k+1}$$

quindi come era da aspettarsi la misura ξ_{k+1} viene "pesata" poco in quanto molto rumorosa.

È possibile quindi interpretare l'informazione $i_k = C^T R^{-1} y_{k+1}$ come la proiezione della stima dello stato ottenibile dalla sola misura y_{k+1} tenendo conto sia della varianza del rumore R sia dell' "osservabilità" diretta dello stato dato da C. Con ragionamenti simili, è possibile interpretare la matrice $C^T R^{-1} C$ come l'inversa dell'incertezza data dalla stima ottenibile con la sola misura y_{k+1}, quindi le direzioni che definiscono il sottospazio relativo al suo nucleo correspondono a direzioni in cui l'incertezza è totale.

ESEMPIO 2 Supponiamo di voler stimare l'evoluzione dello stato utilizzando le misure ottenute tramite il modello lineare:

$$\begin{cases} x_{i+1} = A x_i + w_i \\ y_i = C x_i + v_i \end{cases}$$

a cui aggiungiamo le seguenti ipotesi:

- $x_k, y_k \in \mathbb{R}$
- $w_i = 0 \quad \Longrightarrow \quad Q = q = 0$
- $v_i \sim \mathcal{N}(0, \sigma^2) \quad \Longrightarrow \quad R = r = \sigma^2$
- $A = 1, \quad C = 1$
- $P_0 = \infty, \quad \bar{x}_0 = 0$[1]

Da quete ipotesi si ricava che $x_{k+1} = x_k = \ldots = x_0 = \theta$, cioè il parametro da stimare non cambia nel tempo. Applichiamo a questo modello semplificato il filtro di Kalman in forma di informazione ottenendo le seguenti equazioni:

$$P_{k+1|k} = P_{k|k}$$

$$\hat{x}_{k+1|k} = \hat{x}_{k|k}$$

$$P_{k+1|k+1} = (P_{k|k}^{-1} + \tfrac{1}{\sigma^2})^{-1}$$

$$\hat{x}_{k+1|k+1} = P_{k+1|k+1}(P_{k|k}^{-1}\hat{x}_{k|k} + \tfrac{1}{\sigma^2}y_{k+1})$$

Calcoliamo adesso la stima del filtro di Kalman per ogni passo partendo dallo stato iniziale:

$$P_{0|0} = P_0 = \infty \quad \Longrightarrow \quad P_0^{-1} = 0$$
$$\hat{x}_{0|0} = \bar{x}_0 = 0$$

$$P_{1|1} = \sigma^2$$
$$\hat{x}_{1|1} = y_1$$

$$P_{2|2} = (\tfrac{1}{\sigma^2} + \tfrac{1}{\sigma^2})^{-1}$$
$$\hat{x}_{2|2} = \tfrac{\sigma^2}{2}(\tfrac{1}{\sigma^2}\hat{x}_{1|1} + \tfrac{1}{\sigma^2}y_2) = \tfrac{y_1+y_2}{2}$$

Si puó intuire facilmente come proseguiranno le iterazioni e si trova in generale:

$$P_{k|k} = \tfrac{\sigma^2}{k}$$

$$\hat{x}_{k|k} = \tfrac{1}{k}\textstyle\sum_{i=1}^{k} y_i = \tfrac{1}{k}((k-1)\hat{x}_{k-1|k-1} + y_k) = \tfrac{k-1}{k}\hat{x}_{k-1|k-1} + \tfrac{1}{k}y_k$$

Si puó notare come all'aumentare dei campioni la misura venga pesata sempre meno mentre la stima sempre di piú. Inoltre la varianza d'errore converge a zero, quindi si puó quindi affermare che la stima del filtro di Kalman converge al valore esatto del parametro incognito $x_k = x_0 = \theta$ all'aumentare dei campioni.
Se si volesse determinare quanti campioni sono sufficienti affinché la stima del filtro di Kalman

[1] Il fatto che la varianza iniziale P_0 sia infinita rende comunque ininfluente il valore iniziale $\bar{x}_0$.

sia sufficientemente accurata, é necessario determinare l'intervallo di confidenza utilizzando la formula:

$$\theta \in [\hat{x}_{k|k} - \alpha\frac{\sigma}{\sqrt{k}} \ , \ \hat{x}_{k|k} + \alpha\frac{\sigma}{\sqrt{k}}]$$

dove $\alpha = 1.96 \ , \ 2.57$ con confidenza rispettivamente del 95% e 99%, che si possono ottenere dall'ipotesi di gaussianità del rumore, cioè

$$\mathbb{P}\left[\theta \in [\hat{x}_{k|k} - \alpha\frac{\sigma}{\sqrt{k}} \ , \ \hat{x}_{k|k} + \alpha\frac{\sigma}{\sqrt{k}}]\right] \geq 1 - \delta$$

dove $\delta = 0.05$ per $\alpha = 1.96$ e $\delta = 0.01$ per $\alpha = 2.57$.