

Lezione 5 — Ottobre 22

Docente: Luca Schenato

Stesori: Alberton R., Ausserer M., Barazzuol A.

5.1 Equazioni filtro di Kalman

Nei prossimi esempi si vedranno delle ulteriori applicazioni delle equazioni del filtro di Kalman in forma di informazione che per comodità vengono qui di seguito riassunte:

$$\hat{x}_{k+1|k+1} = A\hat{x}_{k|k} + K_{k+1} (y_{k+1} - CA\hat{x}_{k|k}) \quad (5.1)$$

$$P_{k+1|k} = AP_{k|k}A^T + Q \quad (5.2)$$

$$P_{k+1|k+1} = P_{k+1|k} - P_{k+1|k}C^T (CP_{k+1|k}C^T + R)^{-1} CP_{k+1|k} \quad (5.3)$$

$$K_{k+1} = P_{k+1|k}C^T (CP_{k+1|k}C^T + R)^{-1} \quad (5.4)$$

con condizioni iniziali:

$$P_{0|-1} = P_0 \quad \hat{x}_0 = \bar{x}_0 \quad (5.5)$$

Si continua l'esempio 4 della lezione precedente e si prende in considerazione il vettore delle misure:

$$y_i = [x_i \quad 1] \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\theta} + v_i, \quad i = 1, \dots, N$$

$$v_i \sim \mathcal{N}(0, \sigma^2)$$

Ponendoci nel caso di stima di un vettore θ costante ($A = I, Q = 0, P_0 = \infty$) e ponendo $y = [y_1 \cdots y_N]^T$, con $R = \sigma^2 I$ la stima diventa:

$$\hat{\theta} = \hat{x}_{0|0} = (C^T C)^{-1} C^T y$$

Nel caso di rumori con varianze diverse e possibilmente correlati ovvero, $R \neq \sigma^2 I$ si ottiene:

$$\hat{\theta} = \hat{x}_{0|0} = (C^T R^{-1} C)^{-1} C^T R^{-1} y$$

chiamato 'stimatore ai minimi quadrati pesati'.

Esempio 5 (Stima distribuita): Si considerino le matrici

$$C = \begin{bmatrix} C_1 \\ \vdots \\ C_N \end{bmatrix}, \quad R = \begin{bmatrix} R_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & R_N \end{bmatrix}$$

Si vuole vedere come si scrive lo stimatore ai minimi quadrati pesati in questa situazione. Supponendo che la matrice C sia composta dai vettori riga C_i si ottiene:

$$C^T R^{-1} C = \begin{bmatrix} C_1^T & \cdots & C_N^T \end{bmatrix} \begin{bmatrix} R_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & R_N \end{bmatrix} \begin{bmatrix} C_1 \\ \vdots \\ C_N \end{bmatrix} = \sum_{i=1}^N C_i^T R_i^{-1} C_i$$

e dunque ciò che si ottiene è che l'inverso dell'incertezza complessiva è data dalla somma dell'inverso delle incertezze di ogni misura.

Vale poi:

$$C^T R^{-1} y = \sum_{i=1}^N C_i R_i^{-1} y_i$$

e dunque lo stimatore finale è:

$$\hat{\theta} = \left(\sum_{i=1}^N C_i^T R_i^{-1} C_i \right)^{-1} \left(\sum_{i=1}^N C_i R_i^{-1} y_i \right)$$

Moltiplicando e dividendo per $1/N$ si ha:

$$\hat{\theta} = \left(1/N \sum_{i=1}^N C_i^T R_i^{-1} C_i \right)^{-1} \left(1/N \sum_{i=1}^N C_i R_i^{-1} y_i \right)$$

Calcolando le due medie precedenti si ottiene la stima cercata che verrà chiamata '*stimatore ai minimi quadrati distribuita*' e vale:

$$\hat{\theta} \sim \mathcal{N}(\theta, \left(\sum_{i=1}^N C_i^T R_i^{-1} C_i \right)^{-1})$$

per cui è ancora uno stimatore corretto.

Esempio 6 (Stima ai minimi quadrati iterativa):

Si prende ora in considerazione il modello precedente (con lo stesso significato delle variabili)

$$y_k = C_k \theta + v_k, \quad v_k \sim \mathcal{N}(0, R_k), \quad k = 1, \dots$$

ma in questo caso il pedice k sta a significare che le misure non arrivano più tutte assieme ma in maniera sequenziale. Anche in questo caso il parametro non noto resta costante nel tempo, e dunque si ha: $A = I, Q = 0, P_0 = \infty$ per cui anche $P_{1|0} = P_{0|0} = \infty$. Con questa inizializzazione l'algoritmo del filtro di Kalman in forma di informazione diventa:

$$\begin{aligned} \hat{x}_{k|k} &= (\sum_{i=1}^k C_i^T R_i^{-1} C_i)^{-1} (\sum_{i=1}^k C_i^T R_i^{-1} y_i) \\ P_{k|k} &= (\sum_{i=1}^k C_i^T R_i^{-1} C_i)^{-1} = (P_{k-1|k-1}^{-1} + C_k^T R_k^{-1} C_k)^{-1} \end{aligned}$$

si osservi come la varianza dell'errore di stima diminuisca al crescere di k , e in particolare diminuisce in maniera pesata lungo alcune direzioni. Questa può essere riscritta come:

$$P_{k|k}^{-1} = P_{k-1|k-1}^{-1} + C_k^T R_k^{-1} C_k$$

e dunque lo stimatore diventa:

$$\hat{x}_{k|k} = (P_{k-1|k-1}^{-1} + C_k^T R_k^{-1} C_k)^{-1} (P_{k-1|k-1}^{-1} \hat{x}_{k-1|k-1} + C_k^T R_k^{-1} y_k)$$

ad ogni passo è necessario quindi invertire la matrice P , questo è il prezzo che viene pagato nel fare il conto iterativamente. Nel caso in cui l'inversa non esistesse si possono adottare due soluzioni. La prima consiste nel prendere $P_0 = aI$ con $a \gg 1$ in modo tale che l'inversa esista già dal primo passo. Nel secondo caso si aspetta qualche passo fino a che il termine in sommatoria diventi definito positivo e dal quel punto si inizia ad usare l'algoritmo di stima. La differenza fra le due soluzioni è che la seconda soluzione non dà delle stime nei primi passi a differenza della prima. D'altra parte, nel momento in cui la matrice della seconda soluzione diventa invertibile, la stima fornita è esatta, mentre la stima della prima soluzione è approssimata. Tuttavia, questa differenza diventa sempre più piccola man mano che si accumulano misure e quindi i due approcci sono molto simili da un punto di vista pratico.

Esempio 7 (Stima ai minimi quadrati dinamica o adattiva):

Può succedere che il parametro θ cambi nel tempo. Per risolvere questo problema, si può far ripartire l'algoritmo con il nuovo valore di θ oppure usare un'algoritmo che tiene conto dei possibili cambiamenti di θ . Un modello molto semplice che tiene conto di questi possibili cambiamenti è il seguente:

$$\theta_{k+1} = \theta_k + w_k, \quad w_k \sim \mathcal{N}(0, qI).$$

Si tratta quindi di usare $A = I$ e $Q = qI$ nelle equazioni del filtro di Kalman. Il nuovo θ è quindi uguale a quello precedente a cui viene sommato un rumore bianco. Se il parametro q

è piccolo si ha che θ cambia di poco. Il filtro userá quindi molte misure in modo da filtrare il rumore. Si hanno però dei transitori di stima lenti. Se q è invece grande il filtro ha un transitorio di stima veloce, ma le stime saranno però soggette a piú rumore.

Il fatto che q regoli i pesi dati alle misure e che quindi renda piú o meno “passa-basso” il filtro di Kalman, lo si può facilmente vedere andando a riscrivere le equazioni del filtro di Kalman tenendo conto che in questo specifico caso $P_{k+1|k} = P_{k|k} + qI$ e $\hat{x}_{k+1|k} = \hat{x}_{k|k}$. Infatti si ha che:

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

quindi se q cresce il peso dato alla stima precedente (e di conseguenza alle misure precedenti) diminuisce a favore della misura corrente y_{k+1} e viceversa se q diminuisce.