Filtraggio alla Kalman

Risultati utili per la dimostrazione

Risultato 1: stimatore ottimo di una variabile non osservabile

Ipotesi:

- x variabile casuale non accessibile;
- y variabile casuale accessibile;
- $\hat{x}=h(y)$ stimatore di x noto y ;
- $J[h(\cdot)]=E[\|x-\hat{x}\|^2]=E[\|x-h(y)\|^2]$ cifra di merito.

Tesi:

- $\hat{x}=E[x|y]$ è lo stimatore ottimo, ossia $argmin\, J[h(\cdot)]=E[x|y]$.

Risultato 2: stimatore ottimo lineare

Ipotesi:

- x variabile casuale non accessibile;
- y variabile casuale accessibile;
- $\hat{x}=h(y)$ stimatore di x noto y , $h=\alpha y+\beta$ funzione lineare;
- $J[h(\cdot)]=E[\|x-\hat{x}\|^2]=E[\|x-h(y)\|^2]$ cifra di merito.

Ipotesi non restrittive:

- $E[x]=0$;
- $E[y]=0$.

Tesi:

- $\hat{y}=\frac{\Lambda_{xy}}{\Lambda_{yy}}y$ è lo stimatore ottimo. Per definizione, $\Lambda_{xy}=E[xy]$.

Dimostrazione:

$$J[h(\cdot)]=E[\|x-\hat{x}\|^2]=E[\|x-h(y)\|^2]=E[\|x-\alpha y-\beta\|^2]$$

$$J[h(\cdot)]=E[x^2]+\alpha^2 E[y^2]+\beta^2-2\alpha E[xy]+2\beta E[x]+\alpha\beta E[y]$$

$$J[h(\cdot)]=E[x^2]+\alpha^2 E[y^2]+\beta^2-2\alpha E[xy]$$

Dal momento che β^2 è sempre positivo, sicuramente sarà $\beta=0$, quindi

$$J[h(\cdot)]=E[x^2]+\alpha^2 E[y^2]-2\alpha E[xy]$$

$$\frac{\partial J[h(\cdot)]}{\partial\alpha}=2\alpha E[y^2]-2E[xy]$$ quindi

$$\hat{\alpha}=\frac{E[xy]}{E[y^2]}$$ c.v.d.

Risultato 3: interpretazione geometrica delle variabili casuali

Si definisce V lo spazio vettoriale di elementi v definito come segue:

- V è l'insieme di vettori v_i che corrispondono alle variabili casuali x_i ;
- la somma di vettori è definita come la somma fra le variabili casuali: $v_1+v_2=x_1+x_2$;
- il prodotto scalare è definito come il valore atteso del prodotto fra le variabili casuali: $v_1 \cdot v_2=E[x_1 x_2]$;

Si dimostra che:

1. V rispetta tutte le proprietà di uno spazio vettoriale;
2. v_1 e v_2 sono ortogonali se e solo se x_1 e x_2 sono indipendenti: $v_1 \perp v_2 \Leftrightarrow E[x_1 x_2]=0$;
3. la lunghezza di un vettore v è calcolabile come: $\|v\|=V[x]$;
4. si possono definire sottospazi vettoriali di V , indicati con V' , cui corrispondono opportuni insiemi X' di variabili casuali;
5. la lunghezza della proiezione ortogonale v_P di un vettore v su un sottospazio V' si può calcolare come $\|v_P\|=\|proiezione(v, V')\|=E[x|X']$, dove X' corrisponde a V' , se si suppone che lo stimatore E sia lineare. Questo discende dalla definizione di proiezione ortogonale che di fatto è identica allo stimatore bayesiano lineare ottimo;
6. v_P è il vettore in V' a minima distanza da v ; la variabile casuale corrispondente x_P è il predittore in X' a minima varianza di x ;
7. se la proiezione ortogonale di v su V' è nulla allora v è ortogonale a tutti i vettori in V' (vale anche l'implicazione inversa);
8. se $E[x|X']=0$ allora x è indipendente da tutte le variabili casuali in X' (vale anche l'implicazione inversa);
9. qualunque vettore v può essere scomposto nella somma di due vettori: la sua proiezione su un sottospazio V' e un vettore ortogonale a V' , ossia: $v=v_P+w, v_P=proiezione(v, V') \wedge w \perp V'$;
10. qualunque variabile casuale x può essere scomposta nella somma di due variabili x_P e k , ossia $x=x_P+k$, dove x_P corrisponde a v_p e $k \perp X'$. In particolare vale $E[x]=E[x_P]+E[k]=E[x|X']+E[k]$.

Risultato 4: equazione di Riccati

Ipotesi:

- $P(t)=V[v(t)]$ varianza dell'errore di predizione dello stato;
- $K(t)=(F P(t) H^T+V_{12})(H P(t) H^T-V_2)^{-1}$;

Tesi:

- $P(t+1)=F P(t) F^T-K(t)(F P(t) H^T+V_{12})^T$.

Ipotesi

Restrittive:

- sistema SISO o MIMO;
- uscita del sistema “sporcata” da un vettore di rumore bianco $v_2(t)$ a media nulla e matrice delle varianze V_2 definita positiva;
- stato del sistema “sporcato” da un vettore di rumore bianco $v_1(t)$ a media nulla e matrice delle varianze V_1 semidefinita positiva;
- sistema strettamente proprio ($D=0$);
- uscita del sistema accessibile, stato non accessibile;
- Y_t insieme dei dati disponibili al tempo t .

Non restrittive:

- sistema lineare;
- sistema tempo-invariante;
- $G=0$;
- sistema in rappresentazione di stato $\begin{cases} x(t+1)=F\,x(t)+G\,u(t)+v_1(t) \\ y(t)=H\,x(t)+v_2(t) \end{cases}$;
- matrice delle covarianze tra $v_1(t)$ e $v_2(t)$ nulla: $V_{12}=0$;

Tesi

$\hat{x}(t+1|t)=F\,\hat{x}(t|t-1)+K(t)\,e(t)$, predittore lineare ottimo dello stato

$\hat{y}(t|t-1)=H\,\hat{x}(t|t-1)$, predittore lineare ottimo dell'uscita

$e(t)=y(t)-\hat{y}(t|t-1)$

$K(t)=(F\,P(t)\,H^T)(H\,P(t)\,H^T+V_2)^{-1}$

$P(t+1)=F\,P(t)\,F^T+V_1-K(t)(F\,P(t)\,H^T)^T$

$\hat{x}(0)=E[x(0)]$

$P(0)=V[x(0)]$

V_1 , V_2 parametri del filtro.

Dimostrazione

$\hat{y}(t+1|t)=?$

Per il risultato 1:

$\hat{y}(t+1|t)=E[y(t+1)|Y_t]=E[H\,x(t+1)+v_2(t+1)|Y_t]$

$\hat{y}(t+1|t)=H\,E[x(t+1)|Y_t]+E[v_2(t+1)|Y_t]$

$v_2(t+1)$ è sicuramente indipendente da tutte le variabili in Y_t , quindi il vettore corrispondente sarà ortogonale al sottospazio (risultato 3). Posso concludere $E[v_2(t+1)|Y_t]=0$.

$\hat{y}(t+1|t)=H\,E[x(t+1)|Y_t]$

$\hat{y}(t+1|t)=H\,\hat{x}(t+1|t)$

$\hat{x}(t+1|t)=?$

$\hat{x}(t+1|t)=E[x(t+1)|Y_t]$

Per il risultato 3.10 posso anche scrivere:

$$\hat{x}(t+1|t)=E[x(t+1)|Y_{t-1}]+E[x(t+1)|e(t)]$$

$$\hat{x}(t+1|t)=E[Fx(t)+v_1(t)|Y_{t-1}]+E[x(t+1)|e(t)]$$

$$\hat{x}(t+1|t)=F\,E[x(t)|Y_{t-1}]+E[v_1(t)|Y_{t-1}]+E[x(t+1)|e(t)]$$

Come al passaggio precedente, $E[v_1(t)|Y_{t-1}]=0$. Inoltre, per il risultato 1, $E[x(t)|Y_{t-1}]=\hat{x}(t|t-1)$ e per il risultato 2 $E[x(t+1)|e(t)]=\frac{\Lambda_{x(t+1)e(t)}}{\Lambda_{e(t)e(t)}}e(t)$.

$$\Lambda_{x(t+1)e(t)}=?$$

$$\Lambda_{x(t+1)e(t)}=E[x(t+1)e(t)]=E[x(t+1)(y(t)-\hat{y}(t|t-1))^T]$$

$$\Lambda_{x(t+1)e(t)}=E[x(t+1)(H\,x(t)+v_2(t)-\hat{y}(t|t-1))^T]$$

$$\Lambda_{x(t+1)e(t)}=E[(F\,x(t)+v_1(t))(H\,x(t)+v_2(t)-\hat{y}(t|t-1))^T]$$

Si può osservare che i prodotti $x(t)v_2(t)$, $v_1(t)x(t)$, $v_1(t)v_2(t)$ e $v_1(t)\hat{y}(t|t-1)$ sono tutti nulli, quindi la scrittura si riduce a:

$$\Lambda_{x(t+1)e(t)}=E[F\,x(t)(H\,x(t)-\hat{y}(t|t-1))^T]$$

$$\Lambda_{x(t+1)e(t)}=E[F\,x(t)(H\,x(t)-H\,\hat{x}(t|t-1))^T]$$

$$\Lambda_{x(t+1)e(t)}=E[F\,x(t)(H\,\nu(t))^T]$$

$$\Lambda_{x(t+1)e(t)}=E[F\,x(t)\nu(t)^T\,H^T]$$

$$\Lambda_{x(t+1)e(t)}=F\,E[x(t)\nu(t)^T]H^T=F\,E[(\hat{x}(t|t-1)+\nu(t))\,\nu(t)^T]H^T$$

$$\Lambda_{x(t+1)e(t)}=F\,E[\nu(t)\,\nu(t)^T]H^T=F\,V[\nu(t)]H^T$$

Si pone per definizione $P(t)=V[\nu(t)]$, quindi:

$$\Lambda_{x(t+1)e(t)}=F\,P(t)H^T \quad .$$

$$\Lambda_{e(t)e(t)}=?$$

$$\Lambda_{e(t)e(t)}=E[e(t)e(t)^T]=E[(y(t)-\hat{y}(t|t-1))(y(t)-\hat{y}(t|t-1))^T]$$

$$\Lambda_{e(t)e(t)}=E[(Hx(t)+v_2(t)-H\,\hat{x}(t|t-1))(Hx(t)+v_2(t)-H\,\hat{x}(t|t-1))^T]$$

$$\Lambda_{e(t)e(t)}=E[(H\,\nu(t)+v_2(t))(H\,\nu(t)+v_2(t))^T]$$

$$\Lambda_{e(t)e(t)}=E[H\,\nu(t)\nu(t)^T\,H^T+v_2(t)v_2(t)^T]$$

$$\Lambda_{e(t)e(t)}=HP(t)H^T+V_2$$

Riassumendo:

$$\hat{x}(t+1|t)=F\,\hat{x}(t|t+1)+F\,P(t)H^T+(HP(t)H^T+V_2)^{-1}e(t)$$

Si definisce:

$$K(t)=F\,P(t)H^T+(HP(t)H^T+V_2)^{-1}$$

A questo punto, con il risultato 4, si arriva alla formulazione della tesi.

Generalizzazioni del filtraggio alla Kalman

Matrice $V_{12} \neq 0$

$\hat{x}(t+1|t) = F\,\hat{x}(t|t-1) + K(t)\,e(t)$, predittore lineare ottimo dello stato

$\hat{y}(t|t-1) = H\,\hat{x}(t|t-1)$, predittore lineare ottimo dell'uscita

$e(t) = y(t) - \hat{y}(t|t-1)$

$K(t) = (F\,P(t)H^T + V_{12})(H\,P(t)\,H^T + V_2)^{-1}$

$P(t+1) = F\,P(t)\,F^T + V_1 - K(t)(F\,P(t)\,H^T + V_{12})^T$

$\hat{x}(0) = E[x(0)]$

$P(0) = V[x(0)]$

V_{12} , V_1 e V_2 parametri del filtro.

Filtraggio dello stato

Se F è non singolare, vale la formula:

$\hat{x}(t|t) = F^{-1}\hat{x}(t+1|t)$ viceversa, se $V_{12} = 0$:

$\hat{x}(t|t) = \hat{x}(t|t-1) + K_0(t)\,e(t)$ con

$K_0(t) = P(t)\,H^T(HP(t)\,H^T + V_2)$.

La varianza dell'errore di filtraggio dello stato si calcola come:

$$V[x(t) - \hat{x}(t|t)] = \frac{P(t)\,V_2}{HP(t)\,H^T + V_2}$$

Predittore a r passi

Vale la formula:

$\hat{x}(t+r|t) = F^{r-1}\,\hat{x}(t+1|t)$

La varianza dell'errore di predizione vale:

$$V[x(t+r) - \hat{x}(t+r|t)] = F^{2(r-1)}\,P(t) + V_1\frac{1-F^{2(r-1)}}{1-F^2}$$

Sistemi tempo-varianti con ingresso esogeno

$$\begin{cases} x(t+1) = F(t)\,x(t) + G(t)\,u(t) + v_1(t) \\ y(t) = H(t)\,x(t) + v_2(t) \end{cases}$$

$\hat{x}(t+1|t) = F(t)\,\hat{x}(t|t-1) + G(t)\,u(t) + K(t)\,e(t)$, predittore lineare ottimo dello stato

$\hat{y}(t|t-1) = H(t)\,\hat{x}(t|t-1)$, predittore lineare ottimo dell'uscita

$e(t) = y(t) - \hat{y}(t|t-1)$

$K(t) = (F(t)\,P(t)\,H(t)^T + V_{12})(H(t)\,P(t)\,H(t)^T - V_2)^{-1}$

$P(t+1) = F(t)\,P(t)\,F(t)^T + V_1 - K(t)(F(t)\,P(t)\,H(t)^T + V_{12})^T$

$\hat{x}(0) = E[x(0)]$

$P(0)=V[x(0)]$

V_{12} , V_1 e V_2 parametri del filtro.

Sistemi non lineari (filtro di Kalman esteso)

$$\begin{cases} x(t+1)=f(x(t))+v_1(t) \\ y(t)=h(x(t))+v_2(t) \end{cases} \quad , \quad f \text{ ed } h \text{ funzioni non lineari.}$$

Ci sono due alternative, supporre $K(t)$ lineare e tempo-variante oppure f ed h parametriche di parametro θ .

Nel primo caso, si possono ottenere le matrici $F(t)$ ed $H(t)$ per linearizzazione intorno al valore assunto dal predittore dello stato $\hat{x}$ ed utilizzare il filtro di Kalman per sistemi tempo-varianti.

$$F(t)=\frac{\partial f}{\partial x}\Big|_{x(t)=\hat{x}(t|t-1)}$$

$$H(t)=\frac{\partial h}{\partial x}\Big|_{x(t)=\hat{x}(t|t-1)}$$

Alternativamente, il sistema può essere descritto come:

$$\begin{cases} x(t+1)=f(x(t),\theta(t))+v_1(t) \\ \theta(t+1)=\theta(t)+v_3(t) \\ y(t)=h(x(t))+v_2(t) \end{cases}$$

$x_E=[x_1 \ \ x_2 \ \ \dots \ \ x_n \ \ \theta]$

$$\begin{cases} x_E(t+1)=f(x_E(t))+v_1(t) \\ y(t)=h(x_E(t))+v_2(t) \end{cases} .$$