COMPITO DI
IDENTIFICAZIONE DEI MODELLI ED ANALISI DEI DATI
28 gennaio 2010

Domanda 1. Un'azienda ha due impianti di produzione, diciamo A e B, che producono lo stesso prodotto; i prodotti dei due impianti sono indistinguibili alla vista. La linea di produzione dell'impianto A è più moderna e, di conseguenza, più veloce ed affidabile di quella dell'impianto B. Di conseguenza l'impianto A produce, nello stesso tempo, il doppio dell'impianto B. I due impianti producono in maniera ininterrotta. Inoltre, il prodotto dell'impianto A risulta di qualità superiore. La superiore qualità del prodotto della linea A si riflette sul tempo di vita (=tempo prima della prima rottura).

Il tempo di vita T si può modellare con una variabile aleatoria esponenziale, la cui densità di probabilità è data da:

$$f_T(t) = \begin{cases} 0 & t < 0 \\ \lambda e^{-\lambda t} & t \geq 0 \end{cases}$$

Questo significa che la probabilità che l'istante di prima rottura sia compreso in un intervallo $[a, b]$, $b \geq a \geq 0$ è data da:

$$P[T \in [a, b]] = \int_a^b \lambda e^{-\lambda t}\, dt$$

Il parametro λ caratterizza la densità di probabilità; è facile mostrare che $E[T] = \frac{1}{\lambda}$. Si assuma che il tempo di vita T_A di un pezzo prodotto dalla linea A soddisfi $E[T_A] = \frac{1}{\lambda_A} = 2$ mentre il tempo di vita T_B di un pezzo prodotto dalla linea B soddisfi $E[T_B] = \frac{1}{\lambda_B} = 1$.

1. Sia T il tempo di vita di un pezzo preso a caso dal mercato. Si ricavi la distribuzione di probabilità e la densità di probabilità di T e se ne calcoli la media $E[T]$.

2. Si assuma che il pezzo comprato non si sia ancora rotto fino all'istante $T = 4$. Qual'è la probabilità che il pezzo in questione sia stato prodotto dall'impianto B?

Domanda 2. Si assuma che dei dati ingresso-uscita $\{u(t), y(t)\}$ siano esattamente generati dal modello[1]

$$y(t) = \frac{u(t) + u(t-1) + u(t-2)}{3} + e_0(t)$$

dove $e_0(t)$ è un rumore bianco Gaussiano, a media nulla e varianza $\sigma^2 = 1$. Si assuma che anche $u(t)$ sia un rumore bianco Gaussiano a media nulla e varianza $\sigma_u^2 = 2$, indipendente da $e_0(t)$, .

Si utilizzi un modello del tipo

$$y(t) = b_0 u(t) + b_1 u(t-1) + e(t)$$

per fare identificazione utilizzando il metodo PEM.

1. Supponendo siano disponibili dati $y(t)$, $t = 1, .., N$ e $u(t)$, $t = 0, .., N$, si scriva l'espressione dello stimatore $\hat{\theta}_N$ PEM del parametro vettoriale $\theta := [b_0\ b_1]^\top$ (in questo caso esiste un'espressione in forma chiusa).

2. Si dica dove converge lo stimatore PEM $\hat{\theta}_N$ per $N \to \infty$.

[1] L'uscita $y(t)$ è una versione rumorosa della media dei tre campioni passati di $u(t)$.

3. Facendo l'approssimazione[2]

$$\frac{1}{N}\sum_{i=1}^{N}\psi(t)\psi^\top(t) \simeq E[\psi(t)\psi^\top(t)] = R_1 = \sigma_u^2 I$$

dove $\psi(t) := [u(t)\ u(t-1)]^\top$, si ricavi l'espressione della varianza dello stimatore $\hat{\theta}_N$.

Domanda 3. Dei dati misurati $y(t)$ $t = 1, .., N$ si possono modellare come la somma di una componente "a bassa frequenza" $d(t)$ e di una componente sinusoidale $p(t)$ a frequenza nota ω_0, ma ampiezza e fase incognite; si assuma che $p(t)$ sia indipendente da $d(t)$ e che il segnale $d(t)$ si possa modellare come un processo stazionario, $ARMA(n, n)$, dove n e i parametri del modello sono incogniti. Si assuma inoltre che siano disponibili, da un esperimento preliminare, misure di $d(t)$, $t = t_0, ..., t_0 + M$ dove $t_0 + M << 0$.

Si descriva un procedimento per:

1. Identificare un modello per il processo $d(t)$.

2. Utilizzando il filtro di Kalman, costruire un algoritmo per la predizione della componente sinusoidale $p(t)$ a partire dalle misure $y(t)$, $t = 1, .., N$.

3. Come usereste il filtro di Kalman per costruire un predittore[3] a 2 passi $\hat{p}(t + 2|t) := \hat{E}[p(t + 2)|y(s), s = 1, .., t]$?
Suggerimento: il filtro di Kalman è un algoritmo per la predizione/stima lineare a minima varianza dello stato. Se si vuole predire $p(t + 2)$ basta fare in modo che $p(t + 1)$ sia nello stato.

Domanda 4. Sia θ una grandezza incognita che si può modellare come una variabile aleatoria Gaussiana a media m e varianza σ_θ^2. Un sensore fornisce, ad intervalli regolari, misure rumorose di θ della forma

$$y(k) = \theta + w(k), \quad k = 1, .., N$$

dove $w(k)$ sono variabili aleatorie Gaussiane, a media nulla e varianza σ_k^2, indipendenti e indipendenti da θ (si noti che la varianza dipende dall'istante temporale k).

1. Si scriva l'espressione dello stimatore a Minima Varianza $\hat{\theta}_N := \hat{E}[\theta|y(1), .., y(N)]$

2. Supponendo arrivi una nuova misura $y(N + 1)$, descrivere come si può calcolare $\hat{\theta}_{N+1} = \hat{E}[\theta|y(1), .., y(N), y(N + 1)]$ senza ri-processare tutte le misure.

[2] Questa apprrossimazione è giustificata purchè N sia "grande"
[3] Si ricordi che $\hat{E}[\cdot|\cdot]$ denota il miglior stimatore lineare.

SOLUZIONI

Domanda 1.[7 Punti]

1. La distribuzione di probabilità di una variabile esponenziale si ricava da:

$$P[T \leq t] = \int_0^T \lambda e^{-\lambda t}\, dt = 1 - e^{-\lambda t}$$

Di conseguenza, usando il Teorema della Probabilità Totale

$$\begin{aligned} F_T(t) &= P[T \leq t] = P[T \leq t|A]P[A] + P[T \leq t|B]P[B] \\ &= F_{T_A}(t)\tfrac{2}{3} + F_{T_B}(t)\tfrac{1}{3} \\ &= (1 - e^{-\lambda_A t})\tfrac{2}{3} + (1 - e^{-\lambda_B t})\tfrac{1}{3} \\ &= 1 - \tfrac{2}{3}e^{-\frac{t}{2}} - \tfrac{1}{3}e^{-t} \quad t \geq 0 \end{aligned}$$

e zero per $t < 0$. La densità di probabilità è nulla per $t < 0$ mentre per $t \geq 0$ si ottiene derivando:

$$f_T(t) = \frac{dF_T(t)}{dt} = \frac{2}{3}\left(\frac{1}{2}e^{-\frac{t}{2}}\right) + \frac{1}{3}\left(e^{-t}\right) \quad t \geq 0$$

Per quanto riguarda la media:

$$E[T] = E[T|A]P[A] + E[T|B]P[B] = \frac{1}{\lambda_A}\frac{2}{3} + \frac{1}{\lambda_B}\frac{1}{3} = \frac{5}{3}$$

2. La probabilità richiesta si ottiene usando la regola di Bayes

$$\begin{aligned} P[B|T > 4] &= \tfrac{P[T>4|B]P[B]}{P[T>4]} = \tfrac{\left(1-F_{T_B}(4)\right)\frac{1}{3}}{1-F_T(4)} \\ &= \tfrac{e^{-4}\frac{1}{3}}{\frac{2}{3}e^{-2}+\frac{1}{3}e^{-4}} \\ &= \tfrac{1}{2e^2+1} \simeq 0.0634 \end{aligned}$$

Domanda 2.[9 Punti]

1. Lo stimatore PEM si ottiene risolvendo

$$\begin{aligned} \hat{\theta}_N &:= \arg\min_\theta \tfrac{1}{N}\textstyle\sum_{t=1}^N (y(t) - \hat{y}_\theta(t|t-1))^2 \\ &= \arg\min_\theta \tfrac{1}{N}\textstyle\sum_{t=1}^N (y(t) - \psi^\top(t)\theta)^2 \end{aligned}$$

dove $\theta = [b_0\ b_1]^\top$ e $\psi(t) := [u(t)\ u(t-1)]^\top$. Svolgendo i calcoli si ricava che l'unico punto di minimo è data da:

$$\hat{\theta}_N = \left[\frac{1}{N}\sum_{t=1}^N \psi(t)\psi^\top(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^N \psi(t)y(t)$$

2. Come noto dalla Teoria, lo stimatore PEM converge, per $N \to \infty$, all'insieme dei punti di minimo del funzionale

$$J(\theta) := E[(y(t) - \hat{y}_\theta(t|t-1))^2]$$

Usando il fatto che i dati $y(t)$ sono generati dal modello

$$y(t) = \frac{u(t) + u(t-1) + u(t-2)}{3} + e_0(t)$$

il funzionale $J(\theta)$ risulta:

$$\begin{array}{rcl} J(\theta) & = & E\left[\left(\frac{u(t)+u(t-1)+u(t-2)}{3} + e_0(t) - b_0 u(t) - b_1 u(t-1)\right)^2\right] \\ & = & E\left[\left(\left(\frac{1}{3} - b_0\right) u(t) + \left(\frac{1}{3} - b_1\right) u(t-1) + \frac{1}{3} u(t-2) + e_0\right)^2\right] \\ & = & \left(\frac{1}{3} - b_0\right)^2 \sigma_u^2 + \left(\frac{1}{3} - b_1\right)^2 \sigma_u^2 + \left(\frac{1}{3}\right)^2 \sigma_u^2 + \sigma^2 \end{array}$$

dove si è usato il fatto che $u(t)$ è un rumore bianco e che $\{e_0(t)\}$ è indipendente (e quindi scorrelato) da $\{u(t)\}$. Chiaramente, l'unico punto di minimo di $J(\theta)$ è

$$\left\{ \begin{array}{l} b_0 = \frac{1}{3} \\ b_1 = \frac{1}{3} \end{array} \right.$$

e quindi $\hat{\theta}_N \to \left[\frac{1}{3}\ \frac{1}{3}\right]^\top$.

3. Si osservi preliminarmente che, definendo $\theta_0 := \left[\frac{1}{3}\ \frac{1}{3}\right]^\top$, si può scrivere:

$$y(t) = \frac{u(t) + u(t-1) + u(t-2)}{3} + e_0(t) = \psi^\top(t)\theta_0 + \frac{1}{3} u(t-2) + e_0(t)$$

Per calcolare la varianza, esprimiamo lo stimatore nella seguente forma:

$$\begin{array}{rcl} \hat{\theta}_N & = & \left[\frac{1}{N}\sum_{t=1}^N \psi(t)\psi^\top(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^N \psi(t) y(t) \\ & = & \left[\frac{1}{N}\sum_{t=1}^N \psi(t)\psi^\top(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^N \psi(t)\left(\psi^\top(t)\theta_0 + \frac{1}{3}u(t-2) + e_0(t)\right) \\ & = & \theta_0 + \left[\frac{1}{N}\sum_{t=1}^N \psi(t)\psi^\top(t)\right]^{-1} \frac{1}{N}\sum_{t=1}^N \psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right) \\ & \simeq & \theta_0 + \frac{1}{\sigma_u^2}\frac{1}{N}\sum_{t=1}^N \psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right) \end{array}$$

dove si è usata l'approssimazione suggerita

$$\frac{1}{N}\sum_{t=1}^N \psi(t)\psi^\top(t) \simeq E[\psi(t)\psi^\top(t)] = \sigma_u^2 I \tag{1}$$

Di conseguenza si ottiene:

$$\tilde{\theta}_N := \hat{\theta}_N - \theta_0 \simeq \frac{1}{\sigma_u^2}\frac{1}{N}\sum_{t=1}^N \psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right)$$

Usando il fatto che $\{u(t)\}$ è bianco e scorrelato da $\{e_0(t)\}$, si verifica facilmente che la variabile aleatoria a secondo membro ha media zero e quindi $E[\tilde{\theta}_N] \simeq 0$.
Ci rimane da calcolare la varianza di $\tilde{\theta}_N$ che risulta[4]:

$$\begin{array}{rcl} Var\{\tilde{\theta}_N\} & \simeq & \frac{1}{\sigma_u^4}\frac{1}{N^2}\sum_{t=1}^N\sum_{s=1}^N E\left[\left(\psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right)\right)\left(\psi(s)\left(\frac{1}{3}u(s-2) + e_0(s)\right)\right)^\top\right] \\ & = & \frac{1}{\sigma_u^4}\sum_{t=1}^N E\left[\psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right)\left(\psi(s)\left(\frac{1}{3}u(s-2) + e_0(s)\right)\right)^\top\right] \end{array}$$

Usando il fatto che $\{u(t)\}$ e $\{e_0(t)\}$ sono Gaussiani, bianchi (e quindi a componenti indipendenti) e indipendenti si verifica facilmente che:

$$E\left[\psi(t)\left(\frac{1}{3}u(t-2) + e_0(t)\right)\left(\psi(s)\left(\frac{1}{3}u(s-2) + e_0(s)\right)\right)^\top\right] = \left\{ \begin{array}{cc} \sigma_u^2 I\left(\frac{1}{9}\sigma_u^2 + \sigma^2\right) & t = s \\ 0 & t \neq s \end{array} \right.$$

[4]Il "circa" uguale è dovuto all'approssimazione suggerita (1)

Di conseguenza

$$Var\{\tilde{\theta}_N\} \simeq \frac{1}{\sigma_u^4}\frac{1}{N^2}\sum_{t=1}^{N}\sigma_u^2 I\left(\frac{1}{9}\sigma_u^2+\sigma^2\right) = \frac{1}{N}\left(\frac{1}{9}+\frac{\sigma^2}{\sigma_u^2}\right)I$$

Domanda 3.[9 Punti]

1. Per prima cosa è necessario trovare un modello $ARMA(n,n)$ per il segnale $d(t)$, usando i dati $d(t)$, $t = t_0, .., t_0 + M$.

   Si può utilizzare il metodo PEM applicato alla classe di modelli

   $$d(t) = H(z)e(t)$$

   dove $e(t)$ è un rumore bianco a varianza σ^2 (da stimare) e $H(z)$ è la funzione di trasferimento di un modello $ARMA(n,n)$,

   $$H(z) = \frac{1 + c_1 z^{-1} + .. + c_n z^{-n}}{1 + a_1 z^{-1} + ... + a_n z^{-n}}$$

   dove $H(z)$ deve essere causale e causalmente invertibile (BIBO con inversa BIBO).

   Per ogni ordine fissato si trovano gli stimatori $\hat{c}_1, .., \hat{c}_n$ $\hat{a}_1, .., \hat{a}_n$ e della varianza dell'innovazione $\hat{\sigma}_n^2$. L'ordine n si può stimare utilizzando uno dei criteri visti a lezione, ad esempio AIC o BIC.

2. Per costruire un filtro di Kalman è necessario procurarsi un modello di stato per il processo $y(t) = d(t) + p(t)$.

   Il modello per $d(t)$ è stato ottenuto, in forma di funzione di trasferimento, al punto precedente. Si tratta di trovare una realizzazione minima (A_H, B_H, C_H, D_H) di $H(z)$, i.e.

   $$\begin{cases} x_d(t+1) &= A_H x_d(t) + B_H e(t) \\ d(t) &= C_H x_d(t) + D_H e(t) \end{cases}$$

   in modo che

   $$H(z) = C_H(zI - A_H)^{-1}B_H + D_H.$$

   Si osservi che la varianza dell'innovazione $e(t)$ è stata stimata al punto 1.

   Il modello per il segnale sinusoidale si può scrivere, come visto a lezione, nella forma

   $$\begin{cases} x_p(t+1) &= A_p x_p(t) \\ p(t) &= C_p x_p(t) \end{cases}$$

   dove

   $$A_p = \begin{bmatrix} \cos(\omega_0) & -\sin(\omega_0) \\ \sin(\omega_0) & \cos(\omega_0) \end{bmatrix} \quad C_p = \begin{bmatrix} 1 & 0 \end{bmatrix}$$

   Ne consegue che il segnale $y(t)$ ammette un modello in forma di stato

   $$\begin{cases} x(t+1) &= Ax(t) + Be(t) \\ y(t) &= Cx(t) + De(t) \end{cases}$$

   dove

   $$x(t) := \begin{bmatrix} x_d(t) \\ x_p(t) \end{bmatrix} \quad A = \begin{bmatrix} A_H & 0 \\ 0 & A_p \end{bmatrix}$$

   $$B := \begin{bmatrix} B_H \\ 0 \end{bmatrix} \quad C = \begin{bmatrix} C_H & C_p \end{bmatrix}$$

ed il segnale $p(t)$ si ottiene da

$$p(t) = C_p x_p(t) = [0 \; C_p] x(t).$$

Utilizzando il filtro di Kalman si ottiene il predittore (lineare a minima varianza) $\hat{x}(t+1|t) = [\hat{x}_d(t+1|t)^\top \quad \hat{x}_p(t+1|t)^\top]^\top$. Di conseguenza il predittore per $p(t)$ si ottiene dall'equazione:

$$\hat{p}(t+1|t) = C_p \hat{x}_p(t+1|t) = [0 \;\; C_p] \hat{x}(t+1|t)$$

3. Come suggerito nel testo, per costruire il predittore a due passi è sufficiente "aumentare" lo stato $x(t)$ nel modo seguente:

$$x_a(t) := \begin{bmatrix} x_d(t) \\ x_p(t) \\ x_p(t+1) \end{bmatrix}$$

Il modello di stato diventa ora:

$$\begin{cases} x_a(t+1) &= A_a x_a(t) + B_a e(t) \\ y(t) &= C_a x_a(t) + De(t) \end{cases}$$

dove

$$A_a = \begin{bmatrix} A_H & 0 & 0 \\ 0 & 0 & I \\ 0 & 0 & A_p \end{bmatrix}$$

$$B_a := \begin{bmatrix} B_H \\ 0 \\ 0 \end{bmatrix} \quad C_a = \begin{bmatrix} C_H & C_p & 0 \end{bmatrix}$$

e il segnale $p(t+1)$ si ottiene da

$$p(t+1) = C_p x_p(t+1) = [0 \; 0 \; C_p] x_a(t).$$

Utilizzando il filtro di Kalman si ottiene il predittore (lineare a minima varianza) $\hat{x}_a(t+1|t) = [\hat{x}_d(t+1|t)^\top \;\; \hat{x}_p(t+1|t)^\top \;\; \hat{x}_p(t+2|t)^\top]^\top$. Di conseguenza il predittore a 2 passi per $p(t)$ si ottiene dall'equazione:

$$\hat{p}(t+2|t) = C_p \hat{x}_p(t+2|t) = [0 \;\; 0 \;\; C_p] \hat{x}_a(t+1|t)$$

Domanda 4.[8 Punti]

1. Il modello $y(k) = \theta + w(k)$, $k = 1, .., N$ si può scrivere in forma compatta [5]

$$Y = S\theta + W$$

dove $Y := [y(1) \;\; y(2) \;\; y(3) ... y(N)]^\top$ $S := [1 \;\; 1 \;\; 1 ... 1]^\top$ (N righe e 1 colonna), $W := [w(1) \; w(2) \; w(3) ... w(N)]^\top$.

Lo stimatore a minima varianza di θ basato su Y si scrive nella forma:

$$\hat{\theta}_N := \hat{E}[\theta|Y] = m + cov(\theta, Y) \left[Var\{Y\}\right]^{-1} (Y - Sm)$$

dove

$$cov(\theta, Y) = cov(\theta, S\Theta) + cov(\theta, W) = \sigma_\theta^2 S^\top$$

[5]Questo non è altro che il modello lineare che abbiamo studiato nella stima Bayesiana statica.

e, usando il fatto che θ e W sono scorrelate

$$Var\{Y\} = Var\{S\theta\} + Var\{W\} = \sigma_\theta^2 SS^\top + R$$

dove

$$R := \mathrm{diag}\{\sigma_1^2, .., \sigma_N^2\}.$$

Ne consegue che

$$\hat{\theta}_N = m + \sigma_\theta^2 S^\top \left(\sigma_\theta^2 SS^\top + R\right)^{-1} (Y - Sm)$$

2. Per ottenere una versione "ricorsiva" basta osservare che

$$\begin{array}{rcl} \hat{\theta}_{N+1} := \hat{E}[\theta|Y, y(N+1)] & = & m + \hat{E}[(\theta - m)|Y, \tilde{y}(N+1)] \\ & = & m + \hat{E}[\theta - m|Y] + \hat{E}[\theta - m|\tilde{y}(N+1)] \\ & = & \hat{\theta}_N + \hat{E}[\theta - m|\tilde{y}(N+1)] \end{array} \tag{2}$$

dove

$$\tilde{y}(N+1) := y(N+1) - \hat{E}[y(N+1)|Y]$$

rappresenta "l'innovazione" di $y(N+1)$ rispetto a Y. La seconda equaglianza in (2) deriva dal fatto che Y e $\tilde{y}(N+1)$ sono, per costruzione, scorrelati.

L'innovazione $\tilde{y}(N+1)$ si calcola da:

$$\begin{array}{rcl} \tilde{y}(N+1) & = & y(N+1) - \hat{E}[y(N+1)|Y] \\ & = & y(N+1) - \hat{E}[\theta + w(N+1)|Y] \\ & = & y(N+1) - \hat{\theta}_N \\ & = & \theta + w(N+1) - \hat{\theta}_N \end{array} \tag{3}$$

Di conseguenza, usando (2),

$$\begin{array}{rcl} \hat{\theta}_{N+1} & = & \hat{\theta}_N + \hat{E}[\theta - m|\tilde{y}(N+1)] \\ & = & \hat{\theta}_N + cov(\theta - m, \tilde{y}(N+1)) \left[Var\{\tilde{y}(N+1)\}\right]^{-1} \tilde{y}(N+1) \end{array}$$

Mancano ora da calcolare solo la varianza e covarianza nell'ultima equazione:

$$\begin{array}{rcl} cov(\theta - m, \tilde{y}(N+1)) & = & cov(\theta - \hat{\theta}_N + \hat{\theta}_N - m, \theta - \hat{\theta}_N + w(N+1)) \\ & = & cov(\theta - \hat{\theta}_N, \theta - \hat{\theta}_N + w(N+1)) + cov(\hat{\theta}_N - m, \theta - \hat{\theta}_N + w(N+1)) \\ & = & Var\{\hat{\theta}_N - \theta\} \\ & = & \sigma_\theta^2 - \sigma_\theta^2 S^\top \left(\sigma_\theta^2 SS^\top + R\right)^{-1} S\sigma_\theta^2 \end{array}$$

dove si è usato il fatto che $\hat{\theta}_N - \theta$ è scorrelato da $\hat{\theta}_N - m$ (dimostrato a lezione) e da $w(N+1)$ (questo "nuovo" rumore e' scrorrelato da θ e dalle missure precedenti). Infine abbiamo:

$$\begin{array}{rcl} Var\{\tilde{y}(N+1)\} & = & Var\{\theta - \hat{\theta}_N + w(N+1)\} \\ & = & Var\{\theta - \hat{\theta}_N\} + Var\{w(N+1)\} \\ & = & \sigma_\theta^2 - \sigma_\theta^2 S^\top \left(\sigma_\theta^2 SS^\top + R\right)^{-1} S\sigma_\theta^2 + \sigma_{N+1}^2 \end{array}$$

Di conseguenza, definendo

$$\sigma_{\tilde{\theta}}^2 := Var\{\hat{\theta}_N - \theta\} = \sigma_\theta^2 - \sigma_\theta^2 S^\top \left(\sigma_\theta^2 SS^\top + R\right)^{-1} S\sigma_\theta^2$$

lo stimatore $\hat{\theta}_{N+1}$ si può calcolare da:

$$\begin{array}{rcl} \hat{\theta}_{N+1} & = & \hat{\theta}_N + \sigma_{\tilde{\theta}}^2 \left(\sigma_{\tilde{\theta}}^2 + \sigma_{N+1}^2\right)^{-1} \left(y(N+1) - \hat{\theta}_N\right) \\ & = & \frac{\sigma_{N+1}^2}{\sigma_{\tilde{\theta}}^2 + \sigma_{N+1}^2} \hat{\theta}_N + \frac{\sigma_{\tilde{\theta}}^2}{\sigma_{\tilde{\theta}}^2 + \sigma_{N+1}^2} y(N+1) \end{array}$$