# Université de Rennes 1
# Master recherche STI
# Examen du cours UE 14
# “Filtre de Kalman et modèles de Markov cachés”
# Jeudi 2 février 2006, 14:00 à 16:00
# — Corrigé —

## Exercice

On observe une suite $\{Y_k\}$ à valeurs réelles, définie par

$$Y_k = \theta_k + V_k \ ,$$

où la suite $\{V_k\}$ est un bruit blanc gaussien centré, de variance 1. La suite $\{\theta_k\}$ est une chaîne de Markov prenant deux valeurs réelles 0 ou $m$, indépendante de la suite $\{V_k\}$. La loi initiale est donc définie par un unique paramètre $0 \leq p_0 \leq 1$

$$p_0 = \mathbb{P}[\theta_0 = 0] = 1 - \mathbb{P}[\theta_0 = m] \ ,$$

et la matrice de transition est définie par deux paramètres $0 \leq a, b \leq 1$

$$a = \mathbb{P}[\theta_k = 0 \mid \theta_{k-1} = 0] = 1 - \mathbb{P}[\theta_k = m \mid \theta_{k-1} = 0] \ ,$$

et

$$b = \mathbb{P}[\theta_k = m \mid \theta_{k-1} = m] = 1 - \mathbb{P}[\theta_k = 0 \mid \theta_{k-1} = m] \ ,$$

qu'on supposera indépendants de l'indice $k$.

(i) **Donner l'expression vectorielle de la loi initiale en fonction du paramètre $p_0$, et donner l'expression matricielle de la matrice de transition en fonction des paramètres $a$ et $b$ (très facile).**

Solution

La loi initiale et la matrice de transition sont définies respectivement par

$$\nu = \begin{pmatrix} p_0 \\ 1 - p_0 \end{pmatrix} \qquad \text{et} \qquad \pi = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix} \ ,$$

d'où on tire

$$\pi^* = \begin{pmatrix} a & 1-b \\ 1-a & b \end{pmatrix} \ ,$$

pour un usage ultérieur.

□

(ii) **Donner l'expression des deux densités d'émission $\psi_0(y)$ et $\psi(y)$, indépendantes de l'indice $k$ et définies par**

$$\mathbb{P}[Y_k \in dy \mid \theta_k = 0] = \psi_0(y)\, dy \qquad \textbf{et} \qquad \mathbb{P}[Y_k \in dy \mid \theta_k = m] = \psi(y)\, dy \ ,$$

**respectivement.**

**Solution**

La distribution de probabilité de la variable $Y_k$ sachant $\theta_k$ est gaussienne, de moyenne $\theta_k$ et de variance 1. On en déduit immédiatement que

$$\psi_0(y) = \frac{1}{\sqrt{2\pi}}\ \exp\{-\tfrac{1}{2}\, y^2\} \qquad \text{et} \qquad \psi(y) = \frac{1}{\sqrt{2\pi}}\ \exp\{-\tfrac{1}{2}\, (y-m)^2\} \ ,$$

d'où on tire

$$\frac{\psi_0(y)}{\psi(y)} = \exp\{-\tfrac{1}{2}\, y^2 + \tfrac{1}{2}\, (y-m)^2\} = \exp\{-m\, y + \tfrac{1}{2}\, m^2\} = \exp\{-m\, (y-m) - \tfrac{1}{2}\, m^2\} \ ,$$

pour un usage ultérieur.

□

On définit le rapport de vraisemblance $L_k = \dfrac{\psi_0(Y_k)}{\psi(Y_k)}$, et la probabilité conditionnelle

$$p_k = \mathbb{P}[\theta_k = 0 \mid Y_{1:k}] = 1 - \mathbb{P}[\theta_k = m \mid Y_{1:k}] \ .$$

(iii) **Exprimer $p_k$ en fonction de $p_{k-1}$ et de $L_k$ (on pourra utiliser l'équation de Baum forward et exploiter le fait que $\mathbb{P}[\theta_k = m \mid Y_{1:k}] = 1 - p_k$).**

**Solution**

D'après l'équation de Baum forward, et en utilisant au besoin l'expression de la matrice $\pi^*$ donnée en réponse à la question (i), on a

$$p_k \ = \ c_k\ \psi_0(Y_k)\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1}))\ ,$$

$$1 - p_k \ = \ c_k\ \psi(Y_k)\ ((1-a)\, p_{k-1} + b\, (1-p_{k-1}))\ ,$$

d'où on tire la constante de normalisation

$$1/c_k = \psi_0(Y_k)\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1})) + \psi(Y_k)\ ((1-a)\, p_{k-1} + b\, (1-p_{k-1}))\ ,$$

en ajoutant membre à membre les deux égalités, et finalement

$$\begin{aligned} p_k \ &= \ \frac{\psi_0(Y_k)\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1}))}{\psi_0(Y_k)\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1})) + \psi(Y_k)\ ((1-a)\, p_{k-1} + b\, (1-p_{k-1}))} \\ &= \ \frac{L_k\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1}))}{L_k\ (a\, p_{k-1} + (1-b)\, (1-p_{k-1})) + ((1-a)\, p_{k-1} + b\, (1-p_{k-1}))}\ , \end{aligned}$$

en divisant numérateur et dénominateur par $\psi(Y_k)$.

□

Jusqu'à la fin de l'exercice, on considère le cas particulier où $a = b = 1$.

(iv) **Décrire la chaîne de Markov $\{\theta_k\}$ dans ce cas particulier (très facile).**

Solution

En faisant $a = b = 1$ dans la réponse à la question (i), il vient

$$\pi = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \; ,$$

c'est–à–dire que la matrice de transition est la matrice identité, ce qui signifie que la chaîne de Markov $\{\theta_k\}$ est constante, soit $\theta_k = \theta_0$.

□

(v) **En particularisant le résultat obtenu à la question (iii), exprimer $p_k$ en fonction de $p_{k-1}$ et de $L_k$, et en itérant, exprimer $p_k$ en fonction de $p_0$ et du produit $L_{k:1} = L_k \cdots L_1$.**

Solution

En faisant $a = b = 1$ dans la réponse à la question (iii), il vient

$$p_k = \frac{L_k \, p_{k-1}}{L_k \, p_{k-1} + (1 - p_{k-1})} \; ,$$

soit, en posant $u_k = 1/p_k$

$$u_k = 1 + \frac{1 - p_{k-1}}{L_k \, p_{k-1}} = 1 + \frac{1}{L_k} \, (u_{k-1} - 1) \; ,$$

et en itérant

$$u_k - 1 = \frac{1}{L_k \cdots L_1} \, (u_0 - 1) \qquad \text{ou encore} \qquad u_k = 1 + \frac{1}{L_k \cdots L_1} \, (u_0 - 1) \; ,$$

et finalement

$$p_k = \frac{1}{1 + \dfrac{1}{L_k \cdots L_1} \, (\dfrac{1}{p_0} - 1)} = \frac{L_k \cdots L_1 \, p_0}{L_k \cdots L_1 \, p_0 + (1 - p_0)} = \frac{L_{k:1} \, p_0}{L_{k:1} \, p_0 + (1 - p_0)} \; .$$

□

(vi) **Montrer comment retrouver simplement le résultat de la question (v).**

Solution

Compte tenu que la suite $\{\theta_k\}$ est constante, et d'après l'hypothèse de canal sans mémoire, il vient

$$\begin{aligned}\mathbb{P}[Y_{1:k} \in dy_{1:k} \mid \theta_k = 0] &= \mathbb{P}[Y_1 \in dy_1, \cdots, Y_k \in dy_k \mid \theta_1 = 0, \cdots, \theta_k = 0] \\ &= \psi_0(y_1) \cdots \psi_0(y_k)\, dy_1 \cdots dy_k \ ,\end{aligned}$$

et de même

$$\begin{aligned}\mathbb{P}[Y_{1:k} \in dy_{1:k} \mid \theta_k = m] &= \mathbb{P}[Y_1 \in dy_1, \cdots, Y_k \in dy_k \mid \theta_1 = m, \cdots, \theta_k = m] \\ &= \psi(y_1) \cdots \psi(y_k)\, dy_1 \cdots dy_k \ .\end{aligned}$$

D'après la formule de Bayes

$$\begin{aligned}p_k &= \mathbb{P}[\theta_k = 0 \mid Y_{1:k}] = c_k\, \psi_0(Y_1) \cdots \psi_0(Y_k)\, p_0 \ , \\ 1 - p_k &= \mathbb{P}[\theta_k = m \mid Y_{1:k}] = c_k\, \psi(Y_1) \cdots \psi(Y_k)\, (1 - p_0) \ ,\end{aligned}$$

d'où on tire la constante de normalisation

$$1/c_k = \psi_0(Y_1) \cdots \psi_0(Y_k)\, p_0 + \psi(Y_1) \cdots \psi(Y_k)\, (1 - p_0) \ ,$$

en ajoutant membre à membre les deux égalités, et finalement

$$p_k = \frac{\psi_0(Y_1) \cdots \psi_0(Y_k)\, p_0}{\psi_0(Y_1) \cdots \psi_0(Y_k)\, p_0 + \psi(Y_1) \cdots \psi(Y_k)\, (1 - p_0)} = \frac{L_{k:1}\, p_0}{L_{k:1}\, p_0 + (1 - p_0)} \ ,$$

en divisant numérateur et dénominateur par $\psi(Y_1) \cdots \psi(Y_k)$.

□

On définit l'estimateur du maximum a posteriori (MAP) de la façon suivante :

$$\theta_k^{\mathrm{MAP}} = 0 \qquad \text{si et seulement si} \qquad p_k \geq \tfrac{1}{2} \ .$$

(vii) **Montrer que**

$$\theta_k^{\mathrm{MAP}} = 0 \qquad \textbf{si et seulement si} \qquad L_{k:1} \geq \frac{1 - p_0}{p_0} \ ,$$

**et évaluer la probabilité de non–detection**

$$\mathbb{P}[\theta_k^{\mathrm{MAP}} = 0, \theta_0 = m] \ .$$

SOLUTION

D'après la réponse à la question (v), $p_k \geq \frac{1}{2}$ si et seulement si

$$\frac{L_{k:1}\; p_0}{L_{k:1}\; p_0 + (1-p_0)} \geq \tfrac{1}{2} \qquad \text{ou encore} \qquad 2\, L_{k:1}\; p_0 \geq L_{k:1}\; p_0 + (1-p_0)\ ,$$

c'est–à–dire finalement

$$L_{k:1} \geq \frac{1-p_0}{p_0}\ .$$

D'après la remarque faite dans la réponse à la question (v)

$$L_{k:1} = \frac{\psi_0(Y_k)}{\psi(Y_k)} \cdots \frac{\psi_0(Y_1)}{\psi(Y_1)} = \prod_{l=1}^{k} \exp\{-m\,(Y_l - m) - \tfrac{1}{2}\, m^2\}\ ,$$

et compte tenu que la suite $\{\theta_k\}$ est constante, sur l'ensemble $\{\theta_0 = m\}$ il vient

$$L_{k:1} = \exp\{-m\,\sum_{l=1}^{k}(Y_l - m) - \tfrac{1}{2}\, k\, m^2\} = \exp\{m\,\sqrt{k}\, S_k - \tfrac{1}{2}\,(m\,\sqrt{k})^2\}\ ,$$

où la variable aléatoire

$$S_k = -\frac{1}{\sqrt{k}}\, \sum_{l=1}^{k} (Y_l - \theta_l)\ ,$$

est gaussienne, centrée et de variance 1, indépendante de la suite $\{\theta_k\}$. On en déduit que

$$\begin{aligned}
\mathbb{P}[\theta_k^{\text{MAP}} = 0, \theta_0 = m] &= \mathbb{P}[L_{k:1} \geq \frac{1-p_0}{p_0}, \theta_0 = m] \\
&= \mathbb{P}[\exp\{m\,\sqrt{k}\, S_k - \tfrac{1}{2}\,(m\,\sqrt{k})^2\} \geq \frac{1-p_0}{p_0}, \theta_0 = m] \\
&= \mathbb{P}[S_k \geq \tfrac{1}{2}\, m\,\sqrt{k} + \frac{1}{m\,\sqrt{k}} \log \frac{1-p_0}{p_0}]\; \mathbb{P}[\theta_0 = m]\ ,
\end{aligned}$$

par indépendance, c'est–à–dire que

$$\mathbb{P}[\theta_k^{\text{MAP}} = 0, \theta_0 = m] = (1-p_0)\; \Phi(\tfrac{1}{2}\, m\,\sqrt{k} + \frac{1}{m\,\sqrt{k}} \log \frac{1-p_0}{p_0})\ ,$$

où

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_x^{\infty} e^{-\frac{1}{2}\, u^2}\, du\ ,$$

représente la probabilité qu'une variable aléatoire gaussienne, centrée et de variance 1, dépasse le niveau $x$.

$\square$

## Problème

**Partie A** L'objectif de cette première partie est de proposer un calcul efficace de la distribution de probabilité jointe des observations

$$\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] = p_n(y_0, \cdots, y_n)\, dy_0 \cdots dy_n \ ,$$

dans le modèle linéaire général

$$\left\{ \begin{aligned} X_{k+1} &= F_k\, X_k + f_k + G_k\, W_k \ , \\ Y_k &= H_k\, X_k + h_k + V_k \ , \end{aligned} \right.$$

sous les hypothèses habituelles :

- $X_0$ vecteur aléatoire gaussien, de moyenne $\bar{X}_0$ et de matrice de covariance $Q_0^X$,
- $\{W_k\}$ bruit blanc gaussien centré de matrice de covariance $Q_k$,
- $\{V_k\}$ bruit blanc gaussien centré de matrice de covariance $R_k$ inversible,
- $X_0$, $\{W_k\}$ et $\{V_k\}$ mutuellement indépendants.

(i) **Démontrer la formule suivante**

$$\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] = \prod_{k=0}^{n} \mathbb{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \cdots, Y_0 = y_0] \ ,$$

**valable en toute généralité, avec la convention** $\mathbb{P}[\cdot \mid \emptyset] = \mathbb{P}[\cdot]$**.**

Solution

En itérant la formule de Bayes, il vient

$$\begin{aligned} \mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] &= \mathbb{P}[Y_n \in dy_n \mid Y_0 = y_0, \cdots, Y_{n-1} = y_{n-1}] \\ &\qquad \mathbb{P}[Y_0 \in dy_0, \cdots, Y_{n-1} \in dy_{n-1}] \\ &= \prod_{k=1}^{n} \mathbb{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \cdots, Y_0 = y_0]\, \mathbb{P}[Y_0 \in dy_0] \\ &= \prod_{k=0}^{n} \mathbb{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \cdots, Y_0 = y_0] \ , \end{aligned}$$

avec la convention $\mathbb{P}[Y_0 \in dy_0 \mid \emptyset] = \mathbb{P}[Y_0 \in dy_0]$.

□

(ii) **Montrer que la distribution de probabilité conditionnelle de $Y_k$ sachant $(Y_{k-1} = y_{k-1}, \cdots, Y_0 = y_0)$ est gaussienne, de moyenne $\phi_k(y_0, \cdots, y_{k-1})$ et de matrice de covariance $\Xi_k$ dont on donnera l'expression.**

Solution

Pour $k = 0$, le vecteur aléatoire $Y_0$ est gaussien, de moyenne

$$\phi_0 = H_0 \, \bar{X}_0 + h_0 \ ,$$

et de matrice de covariance

$$\Xi_0 = H_0 \, Q_0^X \, H_0^* + R_0 \ .$$

Pour $k = 1, \cdots, n$, le vecteur aléatoire $Y_{0:k} = (Y_0, \cdots, Y_k)$ est gaussien, donc la distribution de probabilité conditionnelle du vecteur aléatoire $Y_k$ sachant $Y_{0:k-1} = y_{0:k-1}$ est gaussienne, de moyenne

$$\phi_k(y_0, \cdots, y_{k-1}) = \mathbb{E}[Y_k \mid Y_0 = y_0, \cdots, Y_{k-1} = y_{k-1}] \ ,$$

et de matrice de covariance déterministe

$$\begin{aligned}\Xi_k &= \mathbb{E}[\,(Y_k - \phi_k(y_0, \cdots, y_{k-1}))\,(Y_k - \phi_k(y_0, \cdots, y_{k-1}))^* \mid Y_0 = y_0, \cdots, Y_{k-1} = y_{k-1}] \\ &= \mathbb{E}[\,(Y_k - \phi_k(y_0, \cdots, y_{k-1}))\,(Y_k - \phi_k(y_0, \cdots, y_{k-1}))^*] \ ,\end{aligned}$$

indépendante de $y_{0:k-1} = (y_0, \cdots, y_{k-1})$. On sait même dire que

$$\Xi_k = H_k \, P_k^- \, H_k^* + R_k \ ,$$

où la matrice de covariance $P_k^-$ est donnée par les équations du filtre de Kalman.

□

(iii) **En déduire une expression de $p_n(y_0, \cdots, y_n)$.**

Solution

D'après la réponse à la question (ii)

$$\begin{aligned}&\mathbb{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \cdots, Y_0 = y_0] \\ &\qquad = c_k \, \exp\{-\tfrac{1}{2}\,(y_k - \phi_k(y_0, \cdots, y_{k-1}))^* \, \Xi_k^{-1} \, (y_k - \phi_k(y_0, \cdots, y_{k-1}))\} \ ,\end{aligned}$$

et d'autre part

$$\mathbb{P}[Y_0 \in dy_0] = c_0 \, \exp\{-\tfrac{1}{2}\,(y_0 - \phi_0)^* \, \Xi_0^{-1} \, (y_0 - \phi_0)\} \ ,$$

et en utilisant la réponse à la question (i), il vient

$$p_n(y_0, \cdots, y_n) = \prod_{k=1}^{n} c_k \, \exp\{-\tfrac{1}{2}\,(y_k - \phi_k(y_0, \cdots, y_{k-1}))^* \, \Xi_k^{-1}\,(y_k - \phi_k(y_0, \cdots, y_{k-1}))\}$$

$$c_0 \, \exp\{-\tfrac{1}{2}\,(y_0 - \phi_0)^* \, \Xi_0^{-1}\,(y_0 - \phi_0)\} \; .$$

□

La log–vraisemblance $\ell_n$ est définie par

$$\ell_n = \log \, p_n(Y_0, \cdots, Y_n) \; ,$$

où on a remplacé les variables muettes $(y_0, \cdots, y_n)$ par les observations $(Y_0, \cdots, Y_n)$.

(iv) **Donner un algorithme, basé sur le filtre de Kalman, permettant de calculer $\ell_n$.**

Solution

En utilisant d'abord l'expression obtenue à la question (ii), il vient

$$\phi_k(Y_0, \cdots, Y_{k-1}) = \mathbb{E}[Y_k \mid Y_0, \cdots, Y_{k-1}] = H_k \, \mathbb{E}[X_k \mid Y_0, \cdots, Y_{k-1}] + h_k = H_k \, \widehat{X}_k^- + h_k \; ,$$

et en utilisant ensuite l'expression obtenue à la question (iii), il vient

$$\begin{aligned}
\ell_n &= \log \, p_n(Y_0, \cdots, Y_n) \\
&= \sum_{k=1}^{n} \log c_k - \tfrac{1}{2} \sum_{k=1}^{n} (Y_k - \phi_k(Y_0, \cdots, Y_{k-1}))^* \, \Xi_k^{-1}\,(Y_k - \phi_k(Y_0, \cdots, Y_{k-1})) \\
&\qquad + \log c_0 - \tfrac{1}{2}\,(Y_0 - \phi_0)^* \, \Xi_0^{-1}\,(Y_0 - \phi_0) \\
&= c - \tfrac{1}{2} \sum_{k=0}^{n} (Y_k - H_k \, \widehat{X}_k^- - h_k)^* \, \Xi_k^{-1}\,(Y_k - H_k \, \widehat{X}_k^- - h_k) \; ,
\end{aligned}$$

compte tenu que $\widehat{X}_0^- = \bar{X}_0$ pour $k = 0$.

□

(v) **A quoi la log–vraisemblance $\ell_n$ peut–elle servir ?**

Solution

Comme la vraisemblance elle–même, la log–vraisemblance $\ell_n$ permet d'évaluer l'adéquation d'un modèle aux données, de comparer plusieurs modèles entre eux (au vu de leur adéquation aux données), d'identifier un modèle, etc.

□

**Partie B** On suppose dans cette seconde partie que le biais dans l'équation d'état est constant, c'est–à–dire que $f_k = f$ pour tout instant $k$. On se propose d'estimer le paramètre inconnu $f$ au vu des observations $(Y_0, \cdots, Y_n)$, dans le modèle suivant

$$\Sigma(f) \qquad \left\{ \begin{aligned} X_{k+1} &= F_k\, X_k + f + G_k\, W_k \ , \\[2ex] Y_k &= H_k\, X_k + h_k + V_k \ , \end{aligned} \right.$$

paramétré par $f$. On désigne respectivement par $\{\widehat{X}_k^-(f)\}$ et $\ell_n(f)$ le prédicteur de Kalman et la log–vraisemblance dans le modèle $\Sigma(f)$.

(vi) **Montrer par récurrence que la dépendance de $\widehat{X}_k^-(f)$ par rapport au paramètre $f$ est affine, c'est–à–dire que**

$$\widehat{X}_k^-(f) = \widehat{X}_k^-(0) + A_k^-\, f \ ,$$

**pour tout instant $k$, et donner l'expression de la matrice $A_k^-$.**

Solution

Les équations du filtre de Kalman pour le modèle $\Sigma(f)$ sont décrites par l'initialisation

$$\widehat{X}_0^-(f) = \bar{X}_0 \qquad \text{et} \qquad P_0^- = Q_0^X \ ,$$

l'étape de prédiction pour $k = 1, \cdots, n$

$$\begin{aligned} \widehat{X}_k^-(f) &= F_k\, \widehat{X}_{k-1}(f) + f \ , \\[2ex] P_k^- &= F_k\, P_{k-1}\, F_k^* + G_k\, Q_k^W\, G_k^* \ , \end{aligned}$$

et l'étape de correction pour $k = 0, 1, \cdots, n$

$$\begin{aligned} \widehat{X}_k(f) &= \widehat{X}_k^-(f) + K_k\, (Y_k - H_k\, \widehat{X}_k^-(f) - h_k) \ , \\[2ex] P_k &= (I - K_k\, H_k)\, P_k^- = P_k^- - P_k^-\, H_k^*\, \Xi_k^{-1}\, H_k\, P_k^- \ , \end{aligned}$$

avec le gain de Kalman

$$K_k = P_k^-\, H_k^*\, (H_k\, P_k^-\, H_k^* + R_k)^{-1} = P_k^-\, H_k^*\, \Xi_k^{-1} \ .$$

On remarque que les matrices de covariance d'erreur et le gain de Kalman ne dépendent pas de $f$. Clairement

$$\widehat{X}_0^-(f) = \widehat{X}_0^-(0) \ ,$$

c'est–à–dire que l'hypothèse de récurrence est vérifiée avec $A_0^- = 0$. On suppose ensuite que l'hypothèse de récurrence

$$\widehat{X}_k^-(f) = \widehat{X}_k^-(0) + A_k^-\, f \ ,$$

est vérifiée : alors

$$\begin{aligned}
\widehat{X}_k(f) &= \widehat{X}_k^-(f) + K_k\,(Y_k - H_k\,\widehat{X}_k^-(f) - h_k) \\
&= \widehat{X}_k^-(0) + K_k\,(Y_k - H_k\,\widehat{X}_k^-(0) - h_k) + (I - K_k\,H_k)\,A_k^-\, f \\
&= \widehat{X}_k(0) + (I - K_k\,H_k)\,A_k^-\, f \ ,
\end{aligned}$$

et

$$\begin{aligned}
\widehat{X}_{k+1}^-(f) &= F_{k+1}\,\widehat{X}_k(f) + f \\
&= F_{k+1}\,[\,\widehat{X}_k(0) + (I - K_k\,H_k)\,A_k^-\, f\,] + f \\
&= F_{k+1}\,\widehat{X}_k(0) + F_{k+1}\,(I - K_k\,H_k)\,A_k^-\, f + f \\
&= \widehat{X}_{k+1}^-(0) + F_{k+1}\,(I - K_k\,H_k)\,A_k^-\, f + f \ ,
\end{aligned}$$

c'est–à–dire que l'hypothèse de récurrence est vérifiée avec

$$A_{k+1}^- = F_{k+1}\,(I - K_k\,H_k)\,A_k^- + I \ .$$

□

(vii) **En utilisant la réponse à la question (iv), donner une expression de la log–vraisemblance $\ell_n(f)$ en fonction du paramètre $f$.**

Solution

En utilisant d'abord la réponse à la question (vi), il vient

$$Y_k - H_k\,\widehat{X}_k^-(f) - h_k = Y_k - H_k\,\widehat{X}_k^-(0) - h_k - H_k\,A_k^-\, f = I_k - H_k\,A_k^-\, f \ ,$$

où

$$I_k = Y_k - H_k\,\widehat{X}_k^-(0) - h_k \ ,$$

correspond à l'innovation dans le modèle $\Sigma(0)$, et en utilisant ensuite la réponse à la question (iv), il vient

$$\begin{aligned} -\ell_n(f) &= \tfrac{1}{2} \sum_{k=0}^{n} (Y_k - H_k\, \widehat{X}_k^-(f) - h_k)^*\, \Xi_k^{-1}\, (Y_k - H_k\, \widehat{X}_k^-(f) - h_k) \\ &= \tfrac{1}{2} \sum_{k=0}^{n} (I_k - H_k\, A_k^-\, f)^*\, \Xi_k^{-1}\, (I_k - H_k\, A_k^-\, f)\ , \end{aligned}$$

à une constante additive près, indépendante de $f$.

$\square$

(viii) **En déduire l'équation vérifiée par l'estimateur du maximum de vraisemblance $\widehat{f}_n$.**

Solution

L'estimateur du maximum de vraisemblance minimise la forme quadratique

$$\begin{aligned} -\ell_n(f) &= \tfrac{1}{2} \sum_{k=0}^{n} (I_k - H_k\, A_k^-\, f)^*\, \Xi_k^{-1}\, (I_k - H_k\, A_k^-\, f) \\ &= \tfrac{1}{2}\, f^*\, [\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, H_k\, A_k^-\, ]\, f - f^*\, [\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, I_k\, ] + \tfrac{1}{2} \sum_{k=0}^{n} I_k^*\, \Xi_k^{-1}\, I_k\ , \end{aligned}$$

et la condition d'optimalité du premier ordre donne l'équation suivante

$$[\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, H_k\, A_k^-\, ]\, f = \sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, I_k\ .$$

Si la matrice

$$\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, H_k\, A_k^-\ ,$$

est inversible, alors l'estimateur du maximum de vraisemblance est défini de manière unique comme

$$\widehat{f}_n = [\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, H_k\, A_k^-\, ]^{-1}\, [\sum_{k=0}^{n} (H_k\, A_k^-)^*\, \Xi_k^{-1}\, I_k\, ]\ .$$

$\square$