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

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).**

(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.

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

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).**

(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 $L_{k:1} = L_k \cdots L_1$.**

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

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

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

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 $\mathbb{P}[\cdot \mid \emptyset] = \mathbb{P}[\cdot]$.

(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 Ξ_k dont on donnera l'expression.**

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

La log–vraisemblance ℓ_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 ℓ_n.**

(v) **A quoi la log–vraisemblance ℓ_n peut–elle servir ?**

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

$$\Sigma(f) \qquad \begin{cases} X_{k+1} = F_k\, X_k + f + G_k\, W_k \ , \\ \\ Y_k = H_k\, X_k + h_k + V_k \ , \end{cases}$$

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^-.

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

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