Université de Rennes 1
Master Recherche SISEA

Examen du cours UE S3-2
“Filtre de Kalman et modèles de Markov cachés”
Jeudi 15 janvier 2015, 14:00 à 16:00

Problème

Le but de ce problème est d'établir les équations du lisseur de Kalman, vu comme estimateur du maximum a posteriori (MAP), et en utilisant un principe de programmation dynamique *similaire* à l'algorithme de Viterbi vu en cours et adapté ici au cas des systèmes linéaires gaussiens.

On considère une suite d'états cachés $\{X_k\}$ à valeurs dans $\mathbb{R}^m$, vérifiant

$$X_k = F_k\, X_{k-1} + W_k \ ,$$

où $\{W_k\}$ est à valeurs dans $\mathbb{R}^m$, et une suite d'observations $\{Y_k\}$ à valeurs dans $\mathbb{R}^d$, vérifiant

$$Y_k = H_k\, X_k + V_k \ ,$$

et on suppose que

- la condition initiale X_0 est un vecteur aléatoire gaussien, de moyenne $\bar{X}_0$ et de matrice de covariance *inversible* Q_0^X,
- la suite $\{W_k\}$ est un bruit blanc gaussien, de matrice de covariance *inversible* Q_k^W,
- la suite $\{V_k\}$ est un bruit blanc gaussien, de matrice de covariance *inversible* Q_k^V,
- la condition initiale X_0 et les suites $\{W_k\}$ et $\{V_k\}$ sont mutuellement indépendants.

On dispose de toutes les observations

$$Y_{0:n} = (Y_0, Y_1, \cdots, Y_n) \ ,$$

et l'objectif est d'estimer de façon optimale le vecteur aléatoire X_k à partir de $Y_{0:n}$, pour un instant k intermédiaire entre l'instant initial 0 et l'instant final n. Si on adopte le critère du minimum de variance, il s'agit donc de calculer la distribution de probabilité conditionnelle du vecteur aléatoire X_k sachant $Y_{0:n}$. Comme le cadre est gaussien, cette distribution de probabilité conditionnelle est gaussienne, et il suffit de calculer la moyenne

$$\widehat{X}_k^n = \mathbb{E}[X_k \mid Y_{0:n}] \ .$$

On suppose que les estimateurs (prédicteur et filtre, respectivement)

$$\widehat{X}_k^- = \mathbb{E}[X_k \mid Y_{0:k-1}] \qquad \text{et} \qquad \widehat{X}_k = \mathbb{E}[X_k \mid Y_{0:k}] \ ,$$

et les matrices de covariance d'erreur associées

$$P_k^- = \mathbb{E}[(X_k - \widehat{X}_k^-)\,(X_k - \widehat{X}_k^-)^*] \qquad \text{et} \qquad P_k = \mathbb{E}[(X_k - \widehat{X}_k)\,(X_k - \widehat{X}_k)^*] \ ,$$

ont déjà été calculés dans une première phase et sont disponibles, pour tout instant k intermédiaire entre l'instant initial 0 et l'instant final n.

(i) Montrer que $\widehat{X}_n^n = \widehat{X}_n$ pour $k = n$. En d'autres termes, le lisseur et le filtre coïncident à l'instant final.

On rappelle que la distribution de probabilité conditionnelle

$$\mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n \mid Y_0, \cdots, Y_n] \ ,$$

est gausienne, et on suppose que cette distribution de probabilité possède une densité (par rapport à la mesure de Lebesgue), notée $p(x_0, \cdots, x_n \mid Y_0, \cdots, Y_n)$.

(ii) Montrer que la moyenne conditionnelle $(\widehat{X}_0^n, \cdots, \widehat{X}_n^n)$ coïncide avec le mode conditionnel $(X_0^{\mathrm{MAP}}, \cdots, X_n^{\mathrm{MAP}})$, également appelé estimateur du maximum a posteriori (MAP), défini comme le maximum de la fonction

$$(x_0, \cdots, x_n) \longmapsto p(x_0, \cdots, x_n \mid Y_0, \cdots, Y_n) \ ,$$

c'est–à–dire

$$(X_0^{\mathrm{MAP}}, \cdots, X_n^{\mathrm{MAP}}) = \operatorname*{argmax}_{x_0, \cdots, x_n} \ p(x_0, \cdots, x_n \mid Y_0, \cdots, Y_n) \ .$$

On se propose donc dans la suite du problème d'établir les équations pour calculer l'estimateur MAP, défini de manière équivalente comme le minimum de la fonction

$$(x_0, \cdots, x_n) \longmapsto -\log p(x_0, \cdots, x_n \mid Y_0, \cdots, Y_n) \ ,$$

c'est–à–dire

$$(X_0^{\mathrm{MAP}}, \cdots, X_n^{\mathrm{MAP}}) = \operatorname*{argmin}_{x_0, \cdots, x_n} \{-\log p(x_0, \cdots, x_n \mid Y_0, \cdots, Y_n)\} \ .$$

On rappelle que si la suite $(x_0^{\text{opt}}, \cdots, x_k^{\text{opt}})$ atteint le minimum de la fonction

$$(x_0, \cdots, x_k) \longmapsto -\log p(x_0, \cdots, x_k \mid Y_0, \cdots, Y_k) \ ,$$

alors nécessairement x_k^{opt} atteint le minimum de la fonction

$$x_k \longmapsto \min_{x_0, \cdots, x_{k-1}} \{-\log p(x_0, \cdots, x_{k-1}, x_k \mid Y_0, \cdots, Y_k)\} \ ,$$

ce qui justifie d'introduire la fonction *valeur*

$$V_k(x_k) = \min_{x_0, \cdots, x_{k-1}} \{-\log p(x_0, \cdots, x_{k-1}, x_k \mid Y_0, \cdots, Y_k)\} \ ,$$

pour tout $x_k \in \mathbb{R}^m$. On admet l'expression suivante (à une constante multiplicative près) pour la densité conditionnelle

$$\begin{aligned} & p(x_0, \cdots, x_{k-1}, x_k \mid Y_0, \cdots, Y_k) \\ & \quad = \exp\{-\tfrac{1}{2}\,(x_0 - \bar{X}_0)^*\,(Q_0^X)^{-1}\,(x_0 - \bar{X}_0)\} \\ & \qquad\qquad \prod_{l=1}^{k} \exp\{-\tfrac{1}{2}\,(x_l - F_l\,x_{l-1})^*\,(Q_l^W)^{-1}\,(x_l - F_l\,x_{l-1})\} \\ & \qquad\qquad \prod_{l=0}^{k} \exp\{-\tfrac{1}{2}\,(Y_l - H_l\,x_l)^*\,(Q_l^V)^{-1}\,(Y_l - H_l\,x_l)\} \ , \end{aligned}$$

valide pour toute suite $x_0, \cdots, x_k$.

(iii) Montrer que la suite $\{V_k\}$ vérifie la relation de récurrence

$$\begin{aligned} V_k(x_k) \ = \ & \min_{x_{k-1}} \{V_{k-1}(x_{k-1}) + \tfrac{1}{2}\,(x_k - F_k\,x_{k-1})^*\,(Q_k^W)^{-1}\,(x_k - F_k\,x_{k-1})\} \\ & + \tfrac{1}{2}\,(Y_k - H_k\,x_k)^*\,(Q_k^V)^{-1}\,(Y_k - H_k\,x_k) \ , \end{aligned}$$

pour tout $x_k \in \mathbb{R}^m$, avec la condition initiale

$$\begin{aligned} V_0(x_0) \ = \ & \tfrac{1}{2}\,(x_0 - \bar{X}_0)^*\,(Q_0^X)^{-1}\,(x_0 - \bar{X}_0) \\ & + \tfrac{1}{2}\,(Y_0 - H_0\,x_0)^*\,(Q_0^V)^{-1}\,(Y_0 - H_0\,x_0) \ , \end{aligned}$$

pour tout $x_0 \in \mathbb{R}^m$.

La suite $\{V_k\}$ est instrumentale et permet de définir à chaque instant k la fonction

$$I_{k-1}(x_k) = \operatorname*{argmin}_{x_{k-1}} \{V_{k-1}(x_{k-1}) + \tfrac{1}{2}\,(x_k - F_k\,x_{k-1})^*\,(Q_k^W)^{-1}\,(x_k - F_k\,x_{k-1})\} \ ,$$

qui s'interprète comme un pointeur vers un état à l'instant précédent $(k-1)$.

(iv) Montrer que la suite $\{X_k^{\rm MAP}\}$ vérifie la relation de récurrente rétrograde

$$X_{k-1}^{\rm MAP} = I_{k-1}(X_k^{\rm MAP}) \ ,$$

avec la condition initiale

$$X_n^{\rm MAP} = \mathop{\rm argmin}_{x_n} \ V_n(x_n) \ .$$

On propose de montrer par récurrence que la fonction valeur vérifie

$$V_k(x_k) = \tfrac{1}{2}\, (x_k - \widehat{X}_k)^* \, P_k^{-1} \, (x_k - \widehat{X}_k) \ , \qquad (\star)$$

pour tout $x_k \in \mathbb{R}^m$, à une constante additive près, et que

$$I_{k-1}(x_k) = \widehat{X}_{k-1} + L_k \, (x_k - \widehat{X}_k^-) \qquad \text{avec} \qquad L_k = P_{k-1} \, F_k^* \, (P_k^-)^{-1} \ ,$$

pour tout $x_k \in \mathbb{R}^m$.

(v) On suppose que l'hypothèse de récurrence ($\star$) est vérifiée à l'instant $(k-1)$. Montrer que

$$\min_{x_{k-1}} \{V_{k-1}(x_{k-1}) + \tfrac{1}{2}\, (x_k - F_k \, x_{k-1})^* \, (Q_k^W)^{-1} \, (x_k - F_k \, x_{k-1})\}$$

$$= \tfrac{1}{2}\, (x_k - \widehat{X}_k^-)^* \, (P_k^-)^{-1} \, (x_k - \widehat{X}_k^-) \ ,$$

et que

$$\begin{aligned} I_{k-1}(x_k) \ &= \ \mathop{\rm argmin}_{x_{k-1}} \{V_{k-1}(x_{k-1}) + \tfrac{1}{2}\, (x_k - F_k \, x_{k-1})^* \, (Q_k^W)^{-1} \, (x_k - F_k \, x_{k-1})\} \\ &= \ \widehat{X}_{k-1} + L_k \ (x_k - \widehat{X}_k^-) \ , \end{aligned}$$

pour tout $x_k \in \mathbb{R}^m$.

(vi) En utilisant la relation de récurrence établie à la question (iii), montrer que

$$\begin{aligned} V_k(x_k) \ &= \ \min_{x_{k-1}} \{V_{k-1}(x_{k-1}) + \tfrac{1}{2}\, (x_k - F_k \, x_{k-1})^* \, (Q_k^W)^{-1} \, (x_k - F_k \, x_{k-1})\} \\ &\qquad + \tfrac{1}{2}\, (Y_k - H_k \, x_k)^* \, (Q_k^V)^{-1} \, (Y_k - H_k \, x_k) \\ &= \ \tfrac{1}{2}\, (x_k - \widehat{X}_k)^* \, P_k^{-1} \, (x_k - \widehat{X}_k) \ , \end{aligned}$$

pour tout $x_k \in \mathbb{R}^m$, c'est–à–dire que l'hypothèse de récurrence ($\star$) est vérifiée à l'instant k.

(vii) Montrer que la relation de récurrence rétrograde établie à la question (iv) pour la suite $\{X_k^{\rm MAP}\}$ peut être reformulée comme

$$X_{k-1}^{\rm MAP} = \widehat{X}_{k-1} + L_k \ (X_k^{\rm MAP} - \widehat{X}_k^-) \ ,$$

avec la condition initiale $X_n^{\rm MAP} = \widehat{X}_n$.