École Nationale Supérieure de Techniques Avancées
Module : Commande des Systèmes

Examen du cours B7–1 "Filtrage bayésien optimal et approximation particulaire"
Lundi 19 octobre 2009, 8:30 à 10:00

Exercice 1 :

L'objectif de cet exercice est d'établir les équations du filtre bayésien, dans le cas plus général d'une chaîne de Markov observée dans un bruit additif défini par un modèle ARMA (auto–regressive moving average). Plus spécifiquement, on suppose que les états cachés sont à valeurs dans $\mathbb{R}^m$ et vérifient la relation récurrente

$$X_k = f_k(X_{k-1}, W_k) \ ,$$

où $\{W_k\}$ est une suite de vecteurs aléatoires indépendants, et que les observations sont à valeurs dans $\mathbb{R}^d$ et sont reliées aux états cachés par

$$Y_k = h_k(X_k) + V_k \ ,$$

où $\{V_k\}$ est une suite de vecteurs aléatoires *dépendants* à valeurs dans $\mathbb{R}^d$, définie par le modèle ARMA(p,q)

$$V_k + A_1\, V_{k-1} + \cdots + A_p\, V_{k-p} = B_0\, e_k + B_1\, e_{k-1} + \cdots + B_q\, e_{k-q} \ ,$$

paramétré par les matrices $A_1, \cdots, A_p$ et $B_0, B_1, \cdots, B_q$, et où $\{e_k\}$ est une suite de vecteurs aléatoires indépendants à valeurs dans $\mathbb{R}^d$. Par hypothèse, la matrice B_0 est inversible. On suppose que la condition initiale X_0 et les suites $\{W_k\}$ et $\{e_k\}$ sont mutuellement indépendantes.

On admet qu'il existe une représentation d'état pour le bruit d'observation, de la forme

$$U_k \ = \ A\, U_{k-1} + B\, e_{k-1} \ ,$$

$$V_k \ = \ C\, U_k + B_0\, e_k \ ,$$

où la suite $\{U_k\}$ est à valeurs dans $\mathbb{R}^{rd}$, et où A, B et C sont des matrices de dimension $rd \times rd$, $rd \times d$ et $d \times rd$, respectivement, avec $r = \max(p,q)$.

(i) **On définit la nouvelle variable d'état $Z_k = (X_k, U_k)$. Montrer que les nouveaux états cachés ainsi définis vérifient la relation récurrente**

$$Z_k = f_k^{A,B}(Z_{k-1}, e_{k-1}, W_k) \ , \tag{$\star$}$$

où on donnera l'expression de la fonction $f_k^{A,B}$ ainsi définie.

(ii) **Montrer que les observations sont reliées aux nouveaux états cachés par**

$$Y_k = h_k^C(Z_k) + B_0\, e_k \ , \tag{$\star\star$}$$

où on donnera l'expression de la fonction h_k^C ainsi définie.

On constate que les bruits qui apparaissent dans le modèle d'état $(\star)$ et les bruits qui apparaissent l'équation d'observation $(\star\star)$ ne sont pas indépendants. Le modèle obtenu n'appartient donc pas à la classe des modèles de Markov cachés.

(iii) **Montrer qu'on peut écrire le modèle $(\star)$, $(\star\star)$ sous la forme**

$$\begin{aligned} Z_k &= f_k^{A,B,C}(Z_{k-1}, Y_{k-1}, W_k) \ , \\ Y_k &= h_k^C(Z_k) + B_0\, e_k \ , \end{aligned}$$

où on donnera l'expression de la fonction $f_k^{A,B,C}$ ainsi définie. On pourra par exemple utiliser l'équation $(\star\star)$ exprimée à l'instant $(k-1)$ pour obtenir e_{k-1} en fonction de Y_{k-1} et Z_{k-1}, et reporter dans $(\star)$ l'expression obtenue.

On introduit l'hypothèse supplémentaire que le vecteur aléatoire e_k possède une densité $q_k(e)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$.

(iv) **À quelle classe de modèle vue en cours appartient le modèle obtenu à la question (iii) ? On pourra par exemple exprimer la distribution de probabilité conditionnelle jointe du couple (Z_k, Y_k) sachant (Z_{k-1}, Y_{k-1}).**

(v) **Donner l'équation vérifiée par le filtre bayésien $\mathbb{P}[Z_k \in dz \mid Y_{0:k}]$ pour le nouveau modèle. En déduire l'expression du filtre bayésien $\mathbb{P}[X_k \in dx \mid Y_{0:k}]$ pour le modèle de départ.**

Exercice 2 :

L'objectif de cet exercice est d'établir les équations du *lisseur* bayésien, dans le cas simple des modèles de Markov cachés. Ici, l'horizon temporel est fixé, c'est–à–dire que toutes les observations recueillies entre l'instant initial 0 et l'instant final n sont disponibles et peuvent être utilisées pour estimer l'état caché à un instant k intermédiaire entre 0 et n (le cas où $k = n$ correspond bien sûr au cas du *filtre* bayésien vu en cours). Pour fixer les notations, on suppose que les états cachés $\{X_k\}$ forment une chaîne de Markov à valeurs dans E, caractérisée par la distribution de probabilité initiale $\eta_0(dx)$, et les noyaux de probabilités de transition $Q_k(x, dx')$ pour tout $k = 1, \cdots, n$, et que *conditionnellement aux états cachés* les observations $\{Y_k\}$ forment une suite de variables aléatoires indépendantes à valeurs dans F, caractérisée par la densité de probabilité d'émission $g_k(x, y)$, d'où l'expression de la *fonction de vraisemblance* $g_k(x) = g_k(x, Y_k)$ avec l'abus de notation usuel, pour tout $k = 0, 1, \cdots, n$.

On se propose donc d'étudier la distribution de probabilité conditionnelle de l'état caché X_k à l'instant intermédiaire k sachant les observations $Y_{0:n} = (Y_0, \cdots, Y_n)$ entre l'instant initial 0 et l'instant final n, soit

$$\mu_{k|n}(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:n}] \; .$$

(i) **Montrer que le lisseur bayésien $\mu_{k|n}(dx)$ peut s'exprimer comme la distribution normalisée associée à une distribution non–normalisée $\gamma_{k|n}(dx)$ dont on donnera une représentation probabiliste, c'est–à–dire qu'on exprimera $\langle \gamma_{k|n}, \phi \rangle$ comme une espérance mathématique, pour toute fonction mesurable bornée ϕ. On pourra s'inspirer largement de la preuve vue en cours pour le filtre bayésien.**

(ii) **En utilisant la propriété de Markov, montrer qu'il est possible d'exprimer $\langle \gamma_{k|n}, \phi \rangle = \langle \gamma_k, v_{k|n}\, \phi \rangle$ en fonction de la distribution non–normalisée $\gamma_k(dx)$ associée au filtre bayésien et d'une fonction $v_{k|n}(x)$ indépendante de ϕ, qu'on exprimera comme une espérance mathématique.**

(iii) **Établir une relation de récurrence rétrograde, partant de $k = n$, pour la fonction $v_{k|n}$, c'est–à–dire permettant d'exprimer $v_{k-1|n}$ en fonction de $v_{k|n}$.**