Université de Rennes 1
Master EEA (parcours SISEA)

Examen du cours
“Filtrage de Kalman et modèles de Markov cachés”
Jeudi 13 février 2020, 8:00 à 10:00

Exercice 1

Lisseur dans un modèle de Markov caché à espace d'état fini

On considère un modèle de Markov caché $\{(X_k, Y_k)\}$ caractérisé

- par la loi initiale $\nu = (\nu_i)$ définie par
$$\nu_i = \mathbb{P}[X_0 = i] \qquad \text{pour tout } i \in E,$$
- par la matrice de transition $\pi = (\pi_{i,j})$ définie par
$$\pi_{i,j} = \mathbb{P}[X_k = j \mid X_{k-1} = i] \qquad \text{pour tout } i, j \in E,$$
- et dans le cas *symbolique*, par les probabilités d'émission $b = (b_i^\ell)$ définies par
$$b_i^\ell = \mathbb{P}[Y_k = \ell \mid X_k = i] \qquad \text{pour tout } i \in E \text{ et tout } \ell \in O,$$
ou dans le cas *numérique*, par les densités d'émission $g = (g_i)$ définies par
$$g_i(y)\, dy = \mathbb{P}[Y_k \in dy \mid X_k = i] \qquad \text{pour tout } i \in E \text{ et tout } y \in \mathbb{R}^d.$$

Pour gérer à la fois le cas *symbolique* et le cas *numérique*, on introduit la notation
$$g_k^j = \begin{cases} b_j^{Y_k} \;, & \text{dans le cas symbolique} \\[2ex] g_j(Y_k) \;, & \text{dans le cas numérique} \end{cases}$$
pour tout $j \in E$.

On rappelle que le lisseur non–normalisé $q_k = (q_k^i)$ défini par
$$q_k^i = \mathbb{P}[X_k = i \mid Y_0, \cdots, Y_n]\; L_n \qquad \text{pour tout } i \in E, \text{ avec} \qquad L_n = \sum_{i \in E} q_k^i \;,$$
peut s'obtenir comme le produit composante–par–composante
$$q_k^i = p_k^i \; v_k^i \qquad \text{pour tout } i \in E, \tag{$\star$}$$
de la variable forward $p_k = (p_k^i)$ et de la variable backward $v_k = (v_k^i)$.

(i) Rappeler l'équation de récurrence vérifiée par la variable forward p_k et l'équation de récurrence rétrograde vérifiée par la variable backward v_k.

Le but de cet exercice est de proposer la procédure alternative suivante

> une équation de récurrence pour la variable forward p_k et une équation de récurrence rétrograde pour la variable q_k directement, sans utiliser la variable backward v_k,

au lieu de la procédure

> une équation de récurrence pour la variable forward p_k, une équation de récurrence rétrograde pour la variable backward v_k séparément, et le produit composante–par–composante $(\star)$ pour obtenir la variable q_k,

vue en cours.

(ii) Montrer que le lisseur non–normalisé vérifie une équation de récurrence rétrograde de la forme

$$q_{k-1}^i = \sum_{j \in E} t_k^{i,j} \, q_k^j \qquad \textbf{pour tout } i \in E, \tag{$\star\star$}$$

où les coefficients de la matrice $t_k = (t_k^{i,j})$ dépendent des coefficients de la matrice de transition $\pi = (\pi_{i,j})$ et des composantes de la variable forward à l'instant $(k-1)$, mais ne dépendent pas des composantes de la variable backward.

Donner l'expression de la condition initiale, exprimée à l'instant final n.

[Indication : Pour obtenir l'équation $(\star\star)$, on mettra en œuvre les étapes suivantes

- écrire l'équation $(\star)$ à l'instant $(k-1)$,
- utiliser l'équation backward, pour exprimer les composantes de la variable backward à l'instant $(k-1)$ en fonction des composantes de la variable backward à l'instant k,
- utiliser l'équation $(\star)$ à l'instant k, pour éliminer les composantes de la variable backward à l'instant k,
- utiliser l'équation forward, pour exprimer les composantes de la variable forward à l'instant k en fonction des composantes de la variable forward à l'instant $(k-1)$.]

(iii) **Exprimer les coefficients de la matrice $t_k = (t_k^{i,j})$ à l'aide des composantes de la variable forward *normalisée* à l'instant $(k-1)$.**

(iv) **Montrer qu'à tout instant k, la somme des coefficients de chaque *colonne* est égale à 1, c'est–à–dire que la matrice $t_k = (t_k^{i,j})$ est une matrice stochastique *rétrograde* (et peut donc s'interpréter comme la matrice de transition d'une chaîne de Markov *rétrograde*).**

En déduire que toute suite vérifiant l'équation ($\star\star$) possède la propriété que la somme des coefficients est constante (ne dépend pas de l'instant considéré). Cette propriété est–elle surprenante, ou était–elle prévisible, et pourquoi ?

(v) **Donner l'équation de récurrence rétrograde vérifiée par le lisseur *normalisé* $\bar{q}_k = (\bar{q}_k^i)$, défini par**

$$\bar{q}_k^i = \mathbb{P}[X_k = i \mid Y_0, \cdots, Y_n] \qquad \textbf{pour tout } i \in E.$$

(vi) **Comparer les deux procédures en termes de temps de calcul, d'encombrement mémoire, etc. On pourra moduler la réponse selon que l'objectif est d'obtenir le lisseur**

- **à un instant intermédiaire k unique,**
- **à tous les instants compris entre l'instant initial 0 et l'instant final n.**

Exercice 2

Borne de Cramér–Rao a posteriori et filtre de Kalman

On considère un système non–linéaire (mais avec des bruits gaussiens)

$$X_k \;=\; b_k(X_{k-1}) + W_k \;,$$

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

avec les conditions habituelles

- l'état initial X_0 est un vecteur aléatoire gaussien de moyenne $\bar{X}_0$ et de matrice de covariance Q_0^X,
- la suite $\{W_k\}$ est un bruit blanc gaussien (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^W à l'instant k,
- la suite $\{V_k\}$ est un bruit blanc gaussien (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^V à l'instant k,
- l'état initial X_0 et les suites $\{W_k\}$ et $\{V_k\}$ sont mutuellement indépendants,
- les fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$ sont dérivables.

On suppose en outre que

- la matrice de covariance Q_0^X est inversible,
- à tout instant k, les matrices de covariance Q_k^W et Q_k^V sont inversibles.

(i) Rappeler l'équation de récurrence vue en cours vérifiée par la matrice d'information de Fisher J_k, avec l'expression des matrices D_k^-, D_k et D_k^+. On ne demande pas l'expression générale, mais l'expression explicite à l'aide des matrices jacobiennes des fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$.

(ii) En utilisant le résultat vu en cours pour l'expression de la matrice d'information de Fisher dans le cas statique, donner l'expression explicite à l'aide de la matrice jacobienne de la fonction $x \mapsto h_0(x)$, de la condition initiale J_0 définie par

$$J_0 = -\mathbb{E}[\, \frac{\partial^2}{\partial x_0^2} \log p_0(X_0, Y_0) \,] \;.$$

A partir de maintenant, on considère le cas particulier d'un système linéaire gaussien

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

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

(iii) Donner l'expression des matrices D_k^-, D_k et D_k^+ dans ce cas particulier d'un système linéaire gaussien, où les fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$ sont linéaires.

Ré–écrire l'équation de récurrence vérifiée par la matrice d'information de Fisher J_k, et donner l'expression de la condition initiale J_0, dans ce cas particulier.

Il pourra être utile de décomposer l'équation de récurrence vérifiée par la matrice d'information de Fisher J_k en deux étapes, en introduisant par exemple une matrice notée J_k^- à l'étape intermédiaire.

(iv) En procédant par récurrence et en utilisant le lemme d'inversion matricielle vu en cours, montrer que dans le cas particulier d'un système linéaire gaussien, l'inverse J_k^{-1} de la matrice d'information de Fisher J_k coïncide avec la matrice de covariance P_k de l'erreur d'estimation (donnée par les équations du filtre de Kalman).

Interpréter la borne de Cramér–Rao a posteriori ainsi obtenue.