

Université de Rennes 1
DEA STIR
Signal — Télécommunications — Images — Radar
Option Signal

Examen du cours
“Filtre de Kalman et modèles de Markov cachés”
Jeudi 7 janvier 1999, 15:30 à 17:00
— Corrigé —

L’objectif de ce problème est d’étudier les systèmes linéaires à paramètre markovien, qui sont souvent utilisés pour modéliser des systèmes linéaires avec coefficients inconnus, ou avec coefficients susceptibles de changer brusquement.

Partie A On considère le modèle d’état suivant :

- L’état $\{X_k, k \geq 0\}$ est défini par le système linéaire suivant, à valeurs dans \mathbf{R}^m

$$X_{k+1} = F(s_k) X_k + W_k ,$$

où la suite $\{W_k, k \geq 0\}$ est un bruit blanc gaussien centré, de matrice de covariance Q inversible, et où la suite $\{s_k, k \geq 0\}$ est une chaîne de Markov à espace d’état fini $E = \{1, \dots, N\}$, de loi initiale

$$\mathbf{P}[s_0 = i] = \nu_i ,$$

et de matrice de transition

$$\mathbf{P}[s_{k+1} = j \mid s_k = i] = \pi_{i,j} .$$

La condition initiale X_0 est une variable aléatoire gaussienne, de moyenne μ_0 et de matrice de covariance Σ_0 inversible.

On suppose que la chaîne de Markov $\{s_k, k \geq 0\}$, la condition initiale X_0 , et le bruit blanc gaussien $\{W_k, k \geq 0\}$ sont mutuellement indépendants.

On considère d’abord le cas où l’état $\{X_k, k \geq 0\}$ est directement observé, et on se propose de calculer la distribution de probabilité conditionnelle de s_n sachant (X_0, \dots, X_n) .

- (i) Donner l'expression de la densité conditionnelle de X_{k+1} sachant (X_k, s_k) , définie par

$$\mathbf{P}[X_{k+1} \in dx' \mid X_k = x, s_k = i] = \Gamma_i(x, x') dx' ,$$

pour tout $i \in E$.

SOLUTION

Conditionnellement à $(X_k = x, s_k = i)$, la v.a. X_{k+1} est gaussienne, de moyenne $F_i x$ (avec la notation $F_i = F(i)$), et de matrice de covariance Q inversible. On a donc

$$\Gamma_i(x, x') = \frac{1}{(\sqrt{2\pi})^m \sqrt{\det Q}} \exp \left\{ -\frac{1}{2} (x' - F_i x)^* Q^{-1} (x' - F_i x) \right\} .$$

□

- (ii) La propriété de canal sans mémoire est-elle vérifiée ? Donner l'expression de la distribution de probabilité conditionnelle

$$\mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] .$$

SOLUTION

On vérifie d'abord que

$$\mathbf{P}[X_0 \in dx \mid s_n = i_n, \dots, s_0 = i_0] = \mathbf{P}[X_0 \in dx] = p(x) dx ,$$

avec

$$p(x) = \frac{1}{(\sqrt{2\pi})^m \sqrt{\det \Sigma_0}} \exp \left\{ -\frac{1}{2} (x - \mu_0)^* \Sigma_0^{-1} (x - \mu_0) \right\} .$$

D'autre part, on a la décomposition suivante

$$\begin{aligned} & \mathbf{P}[X_{k+1} \in dx_{k+1}, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[X_{k+1} \in dx_{k+1} \mid X_k = x_k, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \end{aligned}$$

$$\mathbf{P}[X_k \in dx_k, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] .$$

Conditionnellement à $(s_n = i_n, \dots, s_0 = i_0)$, on a

$$X_{k+1} = F_{i_k} X_k + W_k ,$$

c'est-à-dire que la suite $\{X_k, k = 0, \dots, n\}$ est un processus gaussien, mais n'est pas une suite de v.a. indépendantes : La propriété de *canal sans mémoire* n'est donc *pas* vérifiée.

Quand on connaît déjà (X_k, s_k) , connaître $(X_{k-1}, \dots, X_0, s_n, \dots, s_{k+1}, s_{k-1}, \dots, s_0)$ n'apporte aucune information supplémentaire sur X_{k+1} . Par conséquent

$$\begin{aligned} & \mathbf{P}[X_{k+1} \in dx_{k+1} \mid X_k = x_k, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[X_{k+1} \in dx_{k+1} \mid X_k = x_k, s_k = i_k] = \Gamma_{i_k}(x_k, x_{k+1}) dx_{k+1} . \end{aligned}$$

On en déduit que

$$\begin{aligned} & \mathbf{P}[X_{k+1} \in dx_{k+1}, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[X_k \in dx_k, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] \Gamma_{i_k}(x_k, x_{k+1}) dx_{k+1} , \end{aligned}$$

et en itérant cette relation, on obtient

$$\begin{aligned} & \mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] \\ &= p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{n-1}}(x_{n-1}, x_n) dx_0 dx_1 \cdots dx_n , \end{aligned}$$

pour tout $i_0, \dots, i_n \in E$.

□

(iii) **En déduire la distribution de probabilité jointe**

$$\mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] .$$

SOLUTION

Compte tenu que

$$\mathbf{P}[s_n = i_n, \dots, s_0 = i_0] = \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{n-1}, i_n} ,$$

on a immédiatement

$$\begin{aligned} & \mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0 \mid s_n = i_n, \dots, s_0 = i_0] \mathbf{P}[s_n = i_n, \dots, s_0 = i_0] \\ &= \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{n-1}, i_n} p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{n-1}}(x_{n-1}, x_n) dx_0 dx_1 \cdots dx_n , \end{aligned}$$

pour tout $i_0, \dots, i_n \in E$.

□

(iv) **Établir l'équation de Baum forward permettant de calculer la distribution de probabilité conditionnelle de s_n sachant (X_0, \dots, X_n) .**

Comme dans la Section 6.1 du cours, on définit la distribution de probabilité jointe des observations passées (X_0, \dots, X_k) et de l'état présent s_k de la façon suivante

$$\mathbf{P}[X_0 \in dx_0, \dots, X_k \in dx_k, s_k = i] = \alpha_i^k[x_0, \dots, x_k] dx_0 \cdots dx_k ,$$

et on définit la variable forward $p^k = (p_i^k)$ par

$$p_i^k = \alpha_i^k[X_0, \dots, X_k] ,$$

pour tout $i \in E$.

Comme dans la démonstration du Théorème 6.4, par définition

$$\begin{aligned} & \alpha_i^k[x_0, \dots, x_k] dx_0 \cdots dx_k = \mathbf{P}[X_0 \in dx_0, \dots, X_k \in dx_k, s_k = i] \\ &= \sum_{i_0, \dots, i_{k-1} \in E} \mathbf{P}[X_0 \in dx_0, \dots, X_k \in dx_k, s_0 = i_0, \dots, s_{k-1} = i_{k-1}, s_k = i] \\ &= \sum_{i_0, \dots, i_{k-1} \in E} \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{k-1}, i} p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{k-1}}(x_{k-1}, x_k) dx_0 dx_1 \cdots dx_k . \end{aligned}$$

De même

$$\begin{aligned} & \mathbf{P}[X_0 \in dx_0, \dots, X_k \in dx_k, X_{k+1} \in dx_{k+1}, s_k = i, s_{k+1} = j] \\ &= \sum_{i_0, \dots, i_{k-1} \in E} \mathbf{P}[X_0 \in dx_0, \dots, X_k \in dx_k, X_{k+1} \in dx_{k+1}, \\ & \quad s_0 = i_0, \dots, s_{k-1} = i_{k-1}, s_k = i, s_{k+1} = j] \\ &= \sum_{i_0, \dots, i_{k-1} \in E} \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{k-1}, i} \pi_{i, j} \\ & \quad p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{k-1}}(x_{k-1}, x_k) \Gamma_i(x_k, x_{k+1}) dx_0 dx_1 \cdots dx_k dx_{k+1} \\ &= \pi_{i, j} \Gamma_i(x_k, x_{k+1}) \alpha_i^k[x_0, \dots, x_k] dx_0 \cdots dx_k dx_{k+1} , \end{aligned}$$

pour tout $i \in E$. En sommant pour tout $i \in E$, on obtient

$$\alpha_j^{k+1}[x_0, \dots, x_{k+1}] = \sum_{i \in E} \pi_{i, j} \Gamma_i(x_k, x_{k+1}) \alpha_i^k[x_0, \dots, x_k] .$$

En substituant les véritables observations (X_0, \dots, X_{k+1}) à la place des variables (muettes) (x_0, \dots, x_{k+1}) , on obtient

$$p_j^{k+1} = \sum_{i \in E} \pi_{i, j} \Gamma_i(X_k, X_{k+1}) p_i^k ,$$

pour tout $j \in E$, avec la condition initiale

$$p_i^0 = \nu_i p(X_0) ,$$

pour tout $i \in E$. On en déduit

$$\mathbf{P}[s_n = i \mid X_0, \dots, X_n] = \frac{p_i^n}{\sum_{j \in E} p_j^n} ,$$

pour tout $i \in E$.

□

Partie B On considère à présent le cas où l'état $\{X_k, k \geq 0\}$ n'est pas directement observé, c'est-à-dire qu'on considère le modèle suivant :

- L'état $\{X_k, k \geq 0\}$ est défini par le système linéaire suivant, à valeurs dans \mathbf{R}^m

$$X_{k+1} = F(s_k) X_k + W_k ,$$

où la suite $\{W_k, k \geq 0\}$ est un bruit blanc gaussien centré, de matrice de covariance Q inversible, et où la suite $\{s_k, k \geq 0\}$ est une chaîne de Markov à espace d'état fini $E = \{1, \dots, N\}$, de loi initiale

$$\mathbf{P}[s_0 = i] = \nu_i ,$$

et de matrice de transition

$$\mathbf{P}[s_{k+1} = j \mid s_k = i] = \pi_{i,j} .$$

La condition initiale X_0 est une variable aléatoire gaussienne, de moyenne μ_0 et de matrice de covariance Σ_0 inversible.

- La suite des observations $\{Y_k, k \geq 0\}$ à valeurs dans \mathbf{R}^d , est définie par

$$Y_k = H X_k + V_k ,$$

où la suite $\{V_k, k \geq 0\}$ est un bruit blanc gaussien centré, de matrice de covariance R inversible.

On suppose que la chaîne de Markov $\{s_k, k \geq 0\}$, la condition initiale X_0 , et les bruits blancs gaussiens $\{W_k, k \geq 0\}$ et $\{V_k, k \geq 0\}$ sont mutuellement indépendants.

Dans cette partie, seule la suite $\{Y_k, k \geq 0\}$ est observée, et on se propose de calculer la distribution de probabilité conditionnelle jointe de (X_n, s_n) sachant (Y_0, \dots, Y_n) .

- (v) **Donner l'expression de la densité conditionnelle de Y_k sachant (X_k, s_k) , définie par**

$$\mathbf{P}[Y_k \in dy \mid X_k = x, s_k = i] = \psi(x, y) dy ,$$

et vérifier que cette densité ne dépend pas de $i \in E$.

SOLUTION

Conditionnellement à $(X_k = x, s_k = i)$, la v.a. Y_k est gaussienne, de moyenne Hx , et de matrice de covariance R inversible. On remarque que les deux premiers moments ne dépendent pas de $i \in E$, et

$$\psi(x, y) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det R}} \exp \left\{ -\frac{1}{2} (y - Hx)^* R^{-1} (y - Hx) \right\} .$$

□

(vi) **Montrer que la propriété de canal sans mémoire est vérifiée, et en déduire l'expression de la distribution de probabilité conditionnelle**

$$\mathbf{P}[Y_n \in dy_n, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] .$$

SOLUTION

Comme à la question (ii), on a la décomposition suivante

$$\begin{aligned} & \mathbf{P}[Y_k \in dy_k, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \dots, Y_0 = y_0, X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ & \quad \mathbf{P}[Y_{k-1} \in dy_{k-1}, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] . \end{aligned}$$

Quand on connaît déjà X_k , connaître $(Y_{k-1}, \dots, Y_0, X_n, \dots, X_{k+1}, X_{k-1}, \dots, X_0, s_n, \dots, s_0)$ n'apporte aucune information supplémentaire sur Y_k . Par conséquent

$$\begin{aligned} & \mathbf{P}[Y_k \in dy_k \mid Y_{k-1} = y_{k-1}, \dots, Y_0 = y_0, X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[Y_k \in dy_k \mid X_k = x_k] = \psi(x_k, y_k) dy_k . \end{aligned}$$

On en déduit que

$$\begin{aligned} & \mathbf{P}[Y_k \in dy_k, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[Y_{k-1} \in dy_{k-1}, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ & \quad \psi(x_k, y_k) dy_k , \end{aligned}$$

et en itérant cette relation, on obtient

$$\begin{aligned} & \mathbf{P}[Y_n \in dy_n, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \psi(x_0, y_0) \cdots \psi(x_n, y_n) dy_0 \cdots dy_n , \end{aligned}$$

pour tout $i_0, \dots, i_n \in E$. La propriété de *canal sans mémoire* est donc vérifiée.

□

(vii) **En déduire la distribution de probabilité jointe**

$$\mathbf{P}[Y_n \in dy_n, \dots, Y_0 \in dy_0, X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] .$$

SOLUTION

Compte tenu que, d'après le résultat de la question (iii)

$$\begin{aligned} & \mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{n-1}, i_n} p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{n-1}}(x_{n-1}, x_n) dx_0 dx_1 \cdots dx_n, \end{aligned}$$

on a immédiatement

$$\begin{aligned} & \mathbf{P}[Y_n \in dy_n, \dots, Y_0 \in dy_0, X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \mathbf{P}[Y_n \in dy_n, \dots, Y_0 \in dy_0 \mid X_n = x_n, \dots, X_0 = x_0, s_n = i_n, \dots, s_0 = i_0] \\ & \quad \mathbf{P}[X_n \in dx_n, \dots, X_0 \in dx_0, s_n = i_n, \dots, s_0 = i_0] \\ &= \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{n-1}, i_n} p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{n-1}}(x_{n-1}, x_n) dx_0 dx_1 \cdots dx_n \\ & \quad \psi(x_0, y_0) \cdots \psi(x_n, y_n) dy_0 \cdots dy_n, \end{aligned}$$

pour tout $i_0, \dots, i_n \in E$.

□

(viii) **Établir l'équation de Baum forward permettant de calculer la distribution de probabilité conditionnelle jointe de (X_n, s_n) sachant (Y_0, \dots, Y_n) , puis la distribution de probabilité conditionnelle marginale de s_n sachant (Y_0, \dots, Y_n) .**

SOLUTION

Comme à la question (iv), et comme dans la Section 6.1 du cours, on définit la distribution de probabilité jointe des observations passées (Y_0, \dots, Y_k) et de l'état présent (X_k, s_k) de la façon suivante

$$\mathbf{P}[Y_0 \in dy_0, \dots, Y_k \in dy_k, X_k \in dx, s_k = i] = \alpha_i^k[x, y_0, \dots, y_k] dx dy_0 \cdots dy_k,$$

et on définit la variable forward $p^k = (p_i^k(\cdot))$ par

$$p_i^k(x) = \alpha_i^k[x, Y_0, \dots, Y_k],$$

pour tout $i \in E$ et tout $x \in \mathbf{R}^m$.

Comme dans la démonstration du Théorème 6.4, par définition

$$\begin{aligned}
& \alpha_i^k[x, y_0, \dots, y_k] dx dy_0 \cdots dy_k = \mathbf{P}[Y_0 \in dy_0, \dots, Y_k \in dy_k, X_k \in dx, s_k = i] \\
&= \sum_{i_0, \dots, i_{k-1} \in E} \int_{\mathbf{R}^m} \cdots \int_{\mathbf{R}^m} \mathbf{P}[Y_0 \in dy_0, \dots, Y_k \in dy_k, \\
&\quad X_0 \in dx_0, \dots, X_{k-1} \in dx_{k-1}, X_k \in dx, s_0 = i_0, \dots, s_{k-1} = i_{k-1}, s_k = i] \\
&= \sum_{i_0, \dots, i_{k-1} \in E} \int_{\mathbf{R}^m} \cdots \int_{\mathbf{R}^m} \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{k-1}, i} \\
&\quad p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{k-1}}(x_{k-1}, x) dx_0 \cdots dx_{k-1} dx \\
&\quad \psi(x_0, y_0) \cdots \psi(x_{k-1}, y_{k-1}) \psi(x, y_k) dy_0 \cdots dy_k .
\end{aligned}$$

De même

$$\begin{aligned}
& \mathbf{P}[Y_0 \in dy_0, \dots, Y_k \in dy_k, Y_{k+1} \in dy_{k+1}, X_k \in dx, X_{k+1} \in dx', s_k = i, s_{k+1} = j] \\
&= \sum_{i_0, \dots, i_{k-1} \in E} \int_{\mathbf{R}^m} \cdots \int_{\mathbf{R}^m} \mathbf{P}[Y_0 \in dy_0, \dots, Y_k \in dy_k, Y_{k+1} \in dy_{k+1}, \\
&\quad X_0 \in dx_0, \dots, X_{k-1} \in dx_{k-1}, X_k \in dx, X_{k+1} \in dx', \\
&\quad s_0 = i_0, \dots, s_{k-1} = i_{k-1}, s_k = i, s_{k+1} = j] \\
&= \sum_{i_0, \dots, i_{k-1} \in E} \int_{\mathbf{R}^m} \cdots \int_{\mathbf{R}^m} \nu_{i_0} \pi_{i_0, i_1} \cdots \pi_{i_{k-1}, i} \pi_{i, j} \\
&\quad p(x_0) \Gamma_{i_0}(x_0, x_1) \cdots \Gamma_{i_{k-1}}(x_{k-1}, x) \Gamma_i(x, x') dx_0 \cdots dx_{k-1} dx dx' \\
&\quad \psi(x_0, y_0) \cdots \psi(x_{k-1}, y_{k-1}) \psi(x, y_k) \psi(x', y_{k+1}) dy_0 \cdots dy_k dy_{k+1} \\
&= \pi_{i, j} \Gamma_i(x, x') \psi(x', y_{k+1}) \alpha_i^k[x, y_0, \dots, y_k] dx dx' dy_0 \cdots dy_k dy_{k+1} ,
\end{aligned}$$

pour tout $i \in E$. En sommant pour tout $i \in E$, et en intégrant par rapport à la variable $x \in \mathbf{R}^m$, on obtient

$$\alpha_j^{k+1}[x', y_0, \dots, y_{k+1}] = \psi(x', y_{k+1}) \sum_{i \in E} \int_{\mathbf{R}^m} \pi_{i, j} \Gamma_i(x, x') \alpha_i^k[x, y_0, \dots, y_k] dx .$$

En substituant les véritables observations (Y_0, \dots, Y_{k+1}) à la place des variables (muettes) (y_0, \dots, y_{k+1}) , on obtient

$$p_j^{k+1}(x') = \psi(x', Y_{k+1}) \sum_{i \in E} \int_{\mathbf{R}^m} \pi_{i,j} \Gamma_i(x, x') p_i^k(x) dx ,$$

pour tout $j \in E$ et tout $x' \in \mathbf{R}^m$, avec la condition initiale

$$p_i^0(x) = \nu_i p(x) ,$$

pour tout $i \in E$ et tout $x \in \mathbf{R}^m$. On en déduit la distribution de probabilité conditionnelle jointe

$$\mathbf{P}[X_n \in dx, s_n = i \mid Y_0, \dots, Y_n] = \frac{p_i^n(x) dx}{\sum_{j \in E} \int_{\mathbf{R}^m} p_j^n(x') dx'} ,$$

et les distributions de probabilité conditionnelles marginales

$$\mathbf{P}[X_n \in dx \mid Y_0, \dots, Y_n] = \frac{\sum_{i \in E} p_i^n(x) dx}{\sum_{j \in E} \int_{\mathbf{R}^m} p_j^n(x') dx'} ,$$

et

$$\mathbf{P}[s_n = i \mid Y_0, \dots, Y_n] = \frac{\int_{\mathbf{R}^m} p_i^n(x) dx}{\sum_{j \in E} \int_{\mathbf{R}^m} p_j^n(x') dx'} ,$$

pour tout $i \in E$ et tout $x \in \mathbf{R}^m$.

□