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

Examen du cours
"Filtre de Kalman et modèles de Markov cachés"
Mardi 17 décembre 2002, 10:30 à 12:00
— Corrigé —

Problème :

L'objectif de ce problème est d'établir par une autre méthode, et de généraliser, les équations du filtre non–linéaire optimal. Soit $\{X_n\}$ une chaîne de Markov à valeurs dans $E = \mathbb{R}^m$, de noyau de probabilités de transition $Q(x, dx')$, et pour tout $n \geq 0$, soit g_n une fonction strictement positive définie sur E.

Pour tout $n \geq 0$, on considère les mesures positives $\gamma_n(dx)$ et $\gamma_n^-(dx)$ sur E, définies par les formules de Feynman–Kac

$$\langle \gamma_n, f \rangle = \int_E f(x)\, \gamma_n(dx) = \mathbb{E}[f(X_n) \prod_{k=0}^{n} g_k(X_k)] \ ,$$

et

$$\langle \gamma_n^-, f \rangle = \int_E f(x)\, \gamma_n^-(dx) = \mathbb{E}[f(X_n) \prod_{k=0}^{n-1} g_k(X_k)] \ ,$$

respectivement, pour toute fonction test f définie sur E. A partir de $\gamma_n(dx)$ et $\gamma_n^-(dx)$, on peut obtenir par normalisation les distributions de probabilité $\mu_n(dx)$ et $\mu_n^-(dx)$ sur E, définies par

$$\mu_n(dx) = \frac{\gamma_n(dx)}{\langle \gamma_n, 1 \rangle} \qquad \text{et} \qquad \mu_n^-(dx) = \frac{\gamma_n^-(dx)}{\langle \gamma_n^-, 1 \rangle} \ .$$

(i) **En considérant la fonction $f \equiv 1$, montrer que**

$$\langle \gamma_n, 1 \rangle = \langle \gamma_n^-, g_n \rangle \qquad \textbf{et} \qquad \langle \gamma_n^-, 1 \rangle = \langle \gamma_{n-1}, 1 \rangle \ .$$

Solution

Par définition

$$\langle \gamma_n, 1\rangle = \mathbb{E}[\prod_{k=0}^{n} g_k(X_k)\,] = \mathbb{E}[g_n(X_n)\ \prod_{k=0}^{n-1} g_k(X_k)\,] = \langle \gamma_n^-, g_n\rangle\ ,$$

et

$$\langle \gamma_n^-, 1\rangle = \mathbb{E}[\prod_{k=0}^{n-1} g_k(X_k)\,] = \langle \gamma_{n-1}, 1\rangle\ .$$

□

(ii) **En déduire que**

$$\langle \gamma_n, 1\rangle = \prod_{k=0}^{n} \langle \mu_k^-, g_k\rangle\ .$$

Solution

D'après la réponse à la question (i), on a

$$\langle \gamma_n, 1\rangle = \langle \gamma_n^-, g_n\rangle = \frac{\langle \gamma_n^-, g_n\rangle}{\langle \gamma_n^-, 1\rangle}\ \langle \gamma_n^-, 1\rangle = \langle \mu_n^-, g_n\rangle\ \langle \gamma_{n-1}, 1\rangle\ ,$$

et en itérant cette identité, on obtient

$$\langle \gamma_n, 1\rangle = \langle \mu_n^-, g_n\rangle \cdots \langle \mu_1^-, g_1\rangle\ \langle \gamma_0, 1\rangle\ .$$

Avec la convention

$$\langle \gamma_0^-, f\rangle = \mathbb{E}[f(X_0)\,]\ ,$$

pour toute fonction test f définie sur E, il vient

$$\langle \gamma_0, 1\rangle = \langle \gamma_0^-, g_0\rangle = \frac{\langle \gamma_0^-, g_0\rangle}{\langle \gamma_0^-, 1\rangle}\ \langle \gamma_0^-, 1\rangle = \langle \mu_0^-, g_0\rangle\ .$$

On en déduit que

$$\langle \gamma_n, 1\rangle = \langle \mu_n^-, g_n\rangle \cdots \langle \mu_1^-, g_1\rangle\ \langle \mu_0^-, g_0\rangle = \prod_{k=0}^{n} \langle \mu_k^-, g_k\rangle\ .$$

□

(iii) **En déduire qu'il est possible de définir les mesures positives $\gamma_n(dx)$ et $\gamma_n^-(dx)$ à partir des distributions de probabilité $\mu_n(dx)$ et $\mu_n^-(dx)$.**

SOLUTION

Par définition, les distributions de probabilité $\mu_n(dx)$ et $\mu_n^-(dx)$ sont obtenues par normalisation des mesures positives $\gamma_n(dx)$ et $\gamma_n^-(dx)$. Inversement,

$$\gamma_n(dx) = \mu_n(dx)\ \langle \gamma_n, 1 \rangle = \mu_n(dx)\ \prod_{k=0}^{n} \langle \mu_k^-, g_k \rangle\ ,$$

et

$$\gamma_n^-(dx) = \mu_n^-(dx)\ \langle \gamma_n^-, 1 \rangle = \mu_n^-(dx)\ \langle \gamma_{n-1}, 1 \rangle = \mu_n^-(dx)\ \prod_{k=0}^{n-1} \langle \mu_k^-, g_k \rangle\ ,$$

c'est–à–dire que les mesures positives $\gamma_n(dx)$ et $\gamma_n^-(dx)$ peuvent être définies à partir des distributions de probabilités $\mu_n(dx)$ et $\mu_n^-(dx)$.

□

(iv) **Montrer que**

$$\langle \gamma_n, f \rangle = \langle \gamma_n^-, g_n\, f \rangle \qquad \textbf{et} \qquad \langle \gamma_n^-, f \rangle = \langle \gamma_{n-1}, Q\, f \rangle\ ,$$

pour toute fonction test f définie sur E.

SOLUTION

Pour toute fonction test f définie sur E

$$\langle \gamma_n, f \rangle = \mathbb{E}[f(X_n)\ \prod_{k=0}^{n} g_k(X_k)\,] = \mathbb{E}[f(X_n)\ g_n(X_n)\ \prod_{k=0}^{n-1} g_k(X_k)\,] = \langle \gamma_n^-, g_n\, f \rangle\ ,$$

et

$$\begin{aligned} \langle \gamma_n^-, f \rangle = \mathbb{E}[f(X_n)\ \prod_{k=0}^{n-1} g_k(X_k)\,] &= \mathbb{E}[\,\mathbb{E}[f(X_n) \mid X_0, \cdots, X_{n-1}\,]\ \prod_{k=0}^{n-1} g_k(X_k)\,] \\ &= \mathbb{E}[Q\, f(X_{n-1})\ \prod_{k=0}^{n-1} g_k(X_k)\,] = \langle \gamma_{n-1}, Q\, f \rangle\ , \end{aligned}$$

compte tenu que

$$\mathbb{E}[f(X_n) \mid X_0, \cdots, X_{n-1}\,] = \mathbb{E}[f(X_n) \mid X_{n-1}\,] = Q\, f(X_{n-1})\ ,$$

d'après la propriété de Markov.

□

(v) **En déduire que**

$$\mu_n(dx') = \frac{g_n(x')\, \mu_n^-(dx')}{\langle \mu_n^-, g_n \rangle} \qquad \textbf{et} \qquad \mu_n^-(dx') = \int_E \mu_{n-1}(dx)\, Q(x, dx') \; .$$

SOLUTION

On déduit de la réponse à la question (iv) que

$$\langle \mu_n, f \rangle = \frac{\langle \gamma_n, f \rangle}{\langle \gamma_n, 1 \rangle} = \frac{\langle \gamma_n^-, g_n\, f \rangle}{\langle \gamma_n^-, g_n \rangle} = \frac{\langle \mu_n^-, g_n\, f \rangle}{\langle \mu_n^-, g_n \rangle} \; ,$$

pour toute fonction test f définie sur E, c'est–à–dire que

$$\mu_n(dx') = \frac{g_n(x')\; \mu_n^-(dx')}{\langle \mu_n^-, g_n \rangle} \; .$$

De même

$$\langle \mu_n^-, f \rangle = \frac{\langle \gamma_n^-, f \rangle}{\langle \gamma_n^-, 1 \rangle} = \frac{\langle \gamma_{n-1}, Q\, f \rangle}{\langle \gamma_{n-1}, 1 \rangle} = \langle \mu_{n-1}, Q\, f \rangle = \int_E \int_E \mu_{n-1}(dx)\; Q(x, dx')\; f(x') \; ,$$

pour toute fonction test f définie sur E, c'est–à–dire que

$$\mu_n^-(dx') = \int_E \mu_{n-1}(dx)\; Q(x, dx') \; .$$

□

Plus généralement, on considère les mesures $\sigma_n(dx)$ et $\sigma_n^-(dx)$ sur E, définies par

$$\langle \sigma_n, f \rangle = \int_E f(x)\, \sigma_n(dx) = \mathbb{E}[f(X_n)\, [\sum_{k=0}^{n} c_k(X_k)\,]\, \prod_{k=0}^{n} g_k(X_k)\,] \; ,$$

et

$$\langle \sigma_n^-, f \rangle = \int_E f(x)\, \sigma_n^-(dx) = \mathbb{E}[f(X_n)\, [\sum_{k=0}^{n-1} c_k(X_k)\,]\, \prod_{k=0}^{n-1} g_k(X_k)\,] \; ,$$

respectivement, pour toute fonction test f définie sur E, et les versions normalisées $\rho_n(dx)$ et $\rho_n^-(dx)$, définies par

$$\rho_n(dx) = \frac{\sigma_n(dx)}{\langle \gamma_n, 1 \rangle} \qquad \text{et} \qquad \rho_n^-(dx) = \frac{\sigma_n^-(dx)}{\langle \gamma_n^-, 1 \rangle} \; .$$

(vi) **Montrer que**

$$\langle \sigma_n, f \rangle = \langle \sigma_n^-, g_n\, f \rangle + \langle \gamma_n^-, c_n\, g_n\, f \rangle \qquad \textbf{et} \qquad \langle \sigma_n^-, f \rangle = \langle \sigma_{n-1}, Q\, f \rangle \; ,$$

pour toute fonction test f définie sur E.

Solution

Pour toute fonction test f définie sur E

$$
\begin{aligned}
\langle \sigma_n, f \rangle &= \mathbb{E}[f(X_n)\, [\sum_{k=0}^{n} c_k(X_k)\,]\, \prod_{k=0}^{n} g_k(X_k)\,] \\
&= \mathbb{E}[f(X_n)\, [\sum_{k=0}^{n-1} c_k(X_k) + c_n(X_n)\,]\, g_n(X_n)\, \prod_{k=0}^{n-1} g_k(X_k)\,] \\
&= \mathbb{E}[f(X_n)\, g_n(X_n)\, [\sum_{k=0}^{n-1} c_k(X_k)\,]\, \prod_{k=0}^{n-1} g_k(X_k)\,] \\
&\qquad + \mathbb{E}[f(X_n)\, c_n(X_n)\, g_n(X_n)\, \prod_{k=0}^{n-1} g_k(X_k)\,] \\
&= \langle \sigma_n^-, g_n\, f \rangle + \langle \gamma_n^-, c_n\, g_n\, f \rangle \ ,
\end{aligned}
$$

et de même, en raisonnant comme au point (iv)

$$
\begin{aligned}
\langle \sigma_n^-, f \rangle &= \mathbb{E}[f(X_n)\, [\sum_{k=0}^{n-1} c_k(X_k)\,]\, \prod_{k=0}^{n-1} g_k(X_k)\,] \\
&= \mathbb{E}[\mathbb{E}[f(X_n) \mid X_0, \cdots, X_{n-1}\,]\, [\sum_{k=0}^{n-1} c_k(X_k)\,]\, \prod_{k=0}^{n-1} g_k(X_k)\,] \\
&= \mathbb{E}[Q\, f(X_{n-1})\, [\sum_{k=0}^{n-1} c_k(X_k)\,]\, \prod_{k=0}^{n-1} g_k(X_k)\,] = \langle \sigma_{n-1}, Q\, f \rangle \ .
\end{aligned}
$$

$\square$

(vii) **En déduire que**

$$\rho_n(dx') = \frac{g_n(x')\, \rho_n^-(dx')}{\langle \mu_n^-, g_n \rangle} + \frac{c_n(x')\, g_n(x')\, \mu_n^-(dx')}{\langle \mu_n^-, g_n \rangle} \ ,$$

et

$$\rho_n^-(dx') = \int_E \rho_{n-1}(dx)\, Q(x, dx') \ .$$

Solution

On déduit de la réponse à la question (vi) que

$$\langle \rho_n, f \rangle = \frac{\langle \sigma_n, f \rangle}{\langle \gamma_n, 1 \rangle} = \frac{\langle \sigma_n^-, g_n\, f \rangle}{\langle \gamma_n^-, g_n \rangle} + \frac{\langle \gamma_n^-, c_n\, g_n\, f \rangle}{\langle \gamma_n^-, g_n \rangle} = \frac{\langle \rho_n^-, g_n\, f \rangle}{\langle \mu_n^-, g_n \rangle} + \frac{\langle \mu_n^-, c_n\, g_n\, f \rangle}{\langle \mu_n^-, g_n \rangle} \ ,$$

pour toute fonction test f définie sur E, c'est–à–dire que

$$\rho_n(dx') = \frac{g_n(x')\ \rho_n^-(dx')}{\langle \mu_n^-, g_n \rangle} + \frac{c_n(x')\, g_n(x')\ \mu_n^-(dx')}{\langle \mu_n^-, g_n \rangle} \ .$$

De même

$$\langle \rho_n^-, f \rangle = \frac{\langle \sigma_n^-, f \rangle}{\langle \gamma_n^-, 1 \rangle} = \frac{\langle \sigma_{n-1}, Q\, f \rangle}{\langle \gamma_{n-1}, 1 \rangle} = \langle \rho_{n-1}, Q\, f \rangle = \int_E \int_E \rho_{n-1}(dx)\ Q(x, dx')\ f(x') \ ,$$

pour toute fonction test f définie sur E, c'est–à–dire que

$$\rho_n^-(dx') = \int_E \rho_{n-1}(dx)\ Q(x, dx') \ .$$

□

Application au filtrage On suppose que la chaîne de Markov $\{X_n\}$ n'est pas observée directement, mais qu'on dispose d'une suite d'observations $\{Y_n\}$ à valeurs dans $F = \mathbb{R}^d$, vérifiant la propriété de *canal sans mémoire*, avec la probabilité d'émission

$$\mathbb{P}[Y_n \in dy \mid X_n = x] = \bar{g}_n(x, y)\, dy \ .$$

(viii) **Montrer que la loi jointe des états cachés** $(X_0, \cdots, X_n)$ **et des observations** $(Y_0, \cdots, Y_n)$ **vérifie**

$$\mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n, Y_0 \in dy_0, \cdots, Y_n \in dy_n]$$

$$= \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]\ \prod_{k=0}^{n} \bar{g}_k(x_k, y_k)\, dy_0 \cdots dy_n \ ,$$

et que

$$\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] = \mathbb{E}[\prod_{k=0}^{n} \bar{g}_k(X_k, y_k)\,]\ dy_0 \cdots dy_n \ .$$

Solution

D'après la formule de Bayes, et d'après la propriété de canal sans mémoire

$$\mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n, Y_0 \in dy_0, \cdots, Y_n \in dy_n]$$

$$= \mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n \mid X_0 = x_0, \cdots, X_n = x_n]\ \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]$$

$$= \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]\ \prod_{k=0}^{n} \bar{g}_k(x_k, y_k)\, dy_0 \cdots dy_n \ .$$

En intégrant par rapport aux variables $x_0, \cdots, x_n$, on obtient

$$\begin{aligned}
&\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] \\
&\quad = \int_E \cdots \int_E \prod_{k=0}^n \bar{g}_k(x_k, y_k)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]\, dy_0 \cdots dy_n \\
&\quad = \mathbb{E}[\prod_{k=0}^n \bar{g}_k(X_k, y_k)\,]\, dy_0 \cdots dy_n \;.
\end{aligned}$$

$\square$

(ix) **En déduire que la loi conditionelle jointe des états cachés** $(X_0, \cdots, X_n)$ **sachant** $(Y_0 = y_0, \cdots, Y_n = y_n)$ **vérifie**

$$\mathbb{E}[f_n(X_0, \cdots, X_n) \mid Y_0 = y_0, \cdots, Y_n = y_n] = \frac{\mathbb{E}[f_n(X_0, \cdots, X_n)\ \prod_{k=0}^n \bar{g}_k(X_k, y_k)\,]}{\mathbb{E}[\prod_{k=0}^n \bar{g}_k(X_k, y_k)\,]} \;,$$

pour toute fonction test f_n **définie sur l'espace produit** $E^{n+1} = \underbrace{E \times \cdots \times E}_{(n+1)\ \textbf{fois}}$**.**

SOLUTION

D'après la réponse à la question (viii) et d'après la formule de Bayes, il vient

$$\begin{aligned}
&f_n(x_0, \cdots, x_n)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n, Y_0 \in dy_0, \cdots, Y_n \in dy_n] \\
&\quad = f_n(x_0, \cdots, x_n)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]\ \prod_{k=0}^n \bar{g}_k(x_k, y_k)\, dy_0 \cdots dy_n \\
&\quad = f_n(x_0, \cdots, x_n)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n \mid Y_0 = y_0, \cdots, Y_n = y_n] \\
&\qquad\qquad \mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] \;,
\end{aligned}$$

pour toute fonction test f_n définie sur l'espace produit E^{n+1}. En intégrant par rapport aux variables $x_0, \cdots, x_n$, on obtient

$$\begin{aligned}
&\int_E \cdots \int_E f_n(x_0, \cdots, x_n)\ \prod_{k=0}^n \bar{g}_k(x_k, y_k)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n]\, dy_0 \cdots dy_n \\
&\quad = \int_E \cdots \int_E f_n(x_0, \cdots, x_n)\, \mathbb{P}[X_0 \in dx_0, \cdots, X_n \in dx_n \mid Y_0 = y_0, \cdots, Y_n = y_n] \\
&\qquad\qquad \mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] \;,
\end{aligned}$$

c'est–à–dire que

$$\mathbb{E}[f_n(X_0,\cdots,X_n)\ \prod_{k=0}^{n} \bar{g}_k(X_k,y_k)\,]\,dy_0\cdots dy_n$$

$$= \mathbb{E}[f_n(X_0,\cdots,X_n) \mid Y_0=y_0,\cdots,Y_n=y_n]\ \mathbb{E}[\prod_{k=0}^{n} \bar{g}_k(X_k,y_k)\,]\,dy_0\cdots dy_n\ .$$

On en déduit que

$$\mathbb{E}[f_n(X_0,\cdots,X_n) \mid Y_0=y_0,\cdots,Y_n=y_n] = \frac{\mathbb{E}[f_n(X_0,\cdots,X_n)\ \prod_{k=0}^{n} \bar{g}_k(X_k,y_k)\,]}{\mathbb{E}[\prod_{k=0}^{n} \bar{g}_k(X_k,y_k)\,]}\ .$$

$\square$

Pour tout $k \geq 0$, on introduit la fonction strictement positive g_k, définie par la relation $g_k(x) = \bar{g}_k(x,y_k)$ pour tout $x \in E$.

(x) **Déduire de la question (ix) que la loi conditionelle de l'état caché** X_n **sachant** $(Y_0=y_0,\cdots,Y_n=y_n)$ **vérifie**

$$\mathbb{E}[f(X_n) \mid Y_0=y_0,\cdots,Y_n=y_n] = \frac{\mathbb{E}[f(X_n)\ \prod_{k=0}^{n} g_k(X_k)\,]}{\mathbb{E}[\prod_{k=0}^{n} g_k(X_k)\,]}\ ,$$

pour toute fonction test f **définie sur** E**.**

Solution

En particulier pour la fonction test f_n définie sur l'espace produit E^{n+1} par la relation

$$f_n(x_0,\cdots,x_n) = f(x_n)\ ,$$

où f est une fonction test définie sur E, il vient

$$\mathbb{E}[f(X_n) \mid Y_0=y_0,\cdots,Y_n=y_n] = \frac{\mathbb{E}[f(X_n)\ \prod_{k=0}^{n} g_k(X_k)\,]}{\mathbb{E}[\prod_{k=0}^{n} g_k(X_k)\,]}\ ,$$

où pour tout $k \geq 0$, la fonction g_k est définie par la relation

$$g_k(x) = \bar{g}_k(x,y_k)\ .$$

$\square$

(xi) **En déduire que la réponse à la question (v) permet de retrouver les équations du filtre non–linéaire optimal .**

Solution

On remarque que

$$\mathbb{E}[f(X_n) \mid Y_0 = y_0, \cdots, Y_n = y_n] = \frac{\mathbb{E}[f(X_n) \displaystyle\prod_{k=0}^{n} g_k(X_k)\,]}{\mathbb{E}[\displaystyle\prod_{k=0}^{n} g_k(X_k)\,]} = \frac{\langle \gamma_n, f \rangle}{\langle \gamma_n, 1 \rangle} = \langle \mu_n, f \rangle \ ,$$

et on déduit de la réponse à la question (v) que

$$\mu_n(dx') = \frac{g_n(x')\ \mu_n^-(dx')}{\langle \mu_n^-, g_n \rangle} \ ,$$

et

$$\mu_n^-(dx') = \int_E \mu_{n-1}(dx)\ Q(x, dx') \ .$$

□