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

Examen du cours
“Filtre de Kalman et modèles de Markov cachés”
Lundi 8 janvier 2001, 14:00 à 15:30
— Corrigé —

Problème :

L'objectif de ce problème est de calculer la densité conditionnelle de l'état caché sachant les observations, pour une certaine classe de systèmes dynamiques non–linéaires.

Préliminaire

On commence par établir un certain nombre de résultats préliminaires, dans le cadre statique suivant.

Soit (X, W, V, Z) un vecteur aléatoire à valeurs dans $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^d \times \mathbb{R}^p$, de densité jointe $p_{X,W,V,Z}$. Soit f et h deux fonctions définies sur $\mathbb{R}^m$, à valeurs dans $\mathbb{R}^m$ et dans $\mathbb{R}^d$ respectivement. On suppose que les vecteurs aléatoires W et V sont indépendants du vecteur aléatoire (X, Z).

(i) **Montrer que la densité conditionnelle jointe du vecteur aléatoire** (X, W) **sachant** $Z = z$ **vérifie**

$$p_{X,W|Z=z}(x, w) = p_W(w)\ p_{X|Z=z}(x)\ .$$

Solution

En utilisant la formule de Bayes, et compte tenu que le vecteur W est indépendant du vecteur aléatoire (X, Z), on obtient

$$p_{X,W|Z=z}(x, w) = \frac{p_{X,W,Z}(x, w, z)}{p_Z(z)} = \frac{p_W(w)\ p_{X,Z}(x, z)}{p_Z(z)} = p_W(w)\ p_{X|Z=z}(x)\ .$$

□

(ii) **En déduire que pour toute fonction–test ϕ définie sur $\mathbb{R}^m$**

$$\mathbb{E}[\,\phi(f(X)+W) \mid Z=z] = \int_{\mathbb{R}^m} \phi(x') \; \Big[\int_{\mathbb{R}^m} p_W(x'-f(x))\, p_{X|Z=z}(x)\, dx \Big] \; dx' \; .$$

Solution

Par définition

$$\mathbb{E}[\,\phi(f(X)+W) \mid Z=z] = \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \phi(f(x)+w)\; p_{X,W|Z=z}(x,w)\, dx\, dw \; .$$

En utilisant l'expression obtenue en (i), et le changement de variable $x' = f(x)+w$, on obtient

$$\begin{aligned}\mathbb{E}[\,\phi(f(X)+W) \mid Z=z] &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^m} \phi(f(x)+w)\; p_W(w)\; p_{X|Z=z}(x)\, dx\, dw \\ &= \int_{\mathbb{R}^m} \phi(x') \; \Big[\int_{\mathbb{R}^m} p_W(x'-f(x))\, p_{X|Z=z}(x)\, dx \Big] \; dx' \; .\end{aligned}$$

□

(iii) **En déduire que la densité conditionnelle du vecteur aléatoire $X' = f(X) + W$ sachant $Z=z$ vérifie**

$$p_{X'|Z=z}(x') = \int_{\mathbb{R}^m} p_W(x'-f(x))\, p_{X|Z=z}(x)\, dx \; .$$

Solution

Par définition

$$\mathbb{E}[\,\phi(X') \mid Z=z] = \int_{\mathbb{R}^m} \phi(x')\; p_{X'|Z=z}(x')\, dx' \; .$$

D'autre part, en utilisant l'expression obtenue en (ii), on obtient

$$\begin{aligned}\mathbb{E}[\,\phi(X') \mid Z=z] &= \mathbb{E}[\,\phi(f(X)+W) \mid Z=z] \\ &= \int_{\mathbb{R}^m} \phi(x') \; \Big[\int_{\mathbb{R}^m} p_W(x'-f(x))\, p_{X|Z=z}(x)\, dx \Big] \; dx' \; .\end{aligned}$$

Comme l'égalité a lieu pour toute fonction–test ϕ définie sur $\mathbb{R}^m$, on en déduit que

$$p_{X'|Z=z}(x') = \int_{\mathbb{R}^m} p_W(x'-f(x))\, p_{X|Z=z}(x)\, dx \; .$$

□

(iv) **Montrer que pour toute fonction–test ϕ définie sur $\mathbb{R}^d$**

$$\mathbb{E}[\,\phi(h(X)+V) \mid X=x, Z=z] = \int_{\mathbb{R}^d} \phi(z')\, p_V(z'-h(x))\, dz' \ .$$

SOLUTION

Par définition

$$\mathbb{E}[\,\phi(h(X)+V) \mid X=x, Z=z] \;=\; \mathbb{E}[\,\phi(h(x)+V) \mid X=x, Z=z]$$

$$= \int_{\mathbb{R}^d} \phi(h(x)+v)\, p_{V|X=x,Z=z}(v)\, dv \ .$$

Compte tenu que le vecteur aléatoire V est indépendant du vecteur aléatoire (X,Z), on obtient

$$p_{V|X=x,Z=z}(v) = p_V(v) \ .$$

En utilisant le changement de variable $z' = h(x)+v$, on obtient

$$\mathbb{E}[\,\phi(h(X)+V) \mid X=x, Z=z] \;=\; \int_{\mathbb{R}^d} \phi(h(x)+v)\, p_V(v)\, dv$$

$$= \int_{\mathbb{R}^d} \phi(z')\, p_V(z'-h(x))\, dz' \ .$$

□

(v) **En déduire l'expression de la densité conditionnelle du vecteur aléatoire $Z' = h(X)+V$ sachant $X=x$ et $Z=z$. Cette densité conditionnelle dépend–elle de z ?**

SOLUTION

Par définition

$$\mathbb{E}[\,\phi(Z') \mid X=x, Z=z] = \int_{\mathbb{R}^d} \phi(z')\, p_{Z'|X=x,Z=z}(z')\, dz' \ .$$

D'autre part, en utilisant l'expression obtenue en (iv), on obtient

$$\mathbb{E}[\,\phi(Z') \mid X=x, Z=z] \;=\; \mathbb{E}[\,\phi(h(X)+V) \mid X=x, Z=z]$$

$$= \int_{\mathbb{R}^d} \phi(z')\, p_V(z'-h(x))\, dz' \ .$$

Comme l'égalité a lieu pour toute fonction–test ϕ définie sur $\mathbb{R}^d$, on en déduit que

$$p_{Z'|X=x,Z=z}(z') = p_V(z'-h(x)) \ .$$

On remarque que cette densité conditionnelle ne dépend pas de z.

□

(vi) **En déduire, en utilisant la formule de Bayes, que la densité conditionnelle du vecteur aléatoire X sachant $Z' = z'$ et $Z = z$ vérifie**

$$p_{X|Z'=z',Z=z}(x) = c \, p_V(z' - h(x)) \, p_{X|Z=z}(x) \ ,$$

où la constante de normalisation c ne dépend que de z et z'.

Solution

En toute généralité, en utilisant la formule de Bayes, on obtient

$$\begin{aligned} p_{X|Z'=z',Z=z}(x) &= \frac{p_{X,Z',Z}(x,z',z)}{p_{Z',Z}(z',z)} \\ &= \frac{p_{Z'|X=x,Z=z}(z') \, p_{X|Z=z}(x) \, p_Z(z)}{p_{Z'|Z=z}(z') \, p_Z(z)} \\ &= \frac{p_{Z'|X=x,Z=z}(z') \, p_{X|Z=z}(x)}{p_{Z'|Z=z}(z')} \ . \end{aligned}$$

En utilisant l'expression obtenue en (v), on obtient

$$p_{X|Z'=z',Z=z}(x) = c \, p_V(z' - h(x)) \, p_{X|Z=z}(x) \ ,$$

où la constante de normalisation

$$c = \frac{1}{p_{Z'|Z=z}(z')} \ ,$$

ne dépend que de z et z'.

□

Application au filtrage des systèmes non–linéaires

On considère le système dynamique non–linéaire suivant :

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

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

où les suites $\{X_k\}$, $\{W_k\}$, $\{V_k\}$ et $\{Y_k\}$ prennent leurs valeurs dans $\mathbb{R}^m$, $\mathbb{R}^m$, $\mathbb{R}^d$ et $\mathbb{R}^d$ respectivement, et où les fonctions f_k et h_k sont définies sur $\mathbb{R}^m$, à valeurs dans $\mathbb{R}^m$ et $\mathbb{R}^d$ respectivement. On suppose que

- le vecteur aléatoire X_0 a pour densité p_0^X,

- la suite aléatoire $\{W_k\}$ est un bruit blanc, et pour tout $k \geq 0$, le vecteur aléatoire W_k a pour densité p_k^W,
- la suite aléatoire $\{V_k\}$ est un bruit blanc, et pour tout $k \geq 0$, le vecteur aléatoire V_k a pour densité p_k^V,
- la condition initiale X_0 et les bruits $\{W_k\}$ et $\{V_k\}$ sont mutuellement indépendants.

Les densités p_0^X, p_k^W et p_k^V ne sont pas nécessairement supposées gaussiennes.

L'estimation optimale (au sens du minimum de la variance de l'erreur d'estimation) des états cachés X_k et X_{k+1} au vu des observations $Y_0, \cdots, Y_k$ repose sur le calcul des densités conditionnelles $p_k = p_{X_k | Y_0, \cdots, Y_k}$ et $p_{k+1}^- = p_{X_{k+1} | Y_0, \cdots, Y_k}$, définies respectivement par

$$\mathbb{P}[\, X_k \in dx \mid Y_0, \cdots, Y_k] \; = \; p_k(x) \, dx$$

$$\mathbb{P}[\, X_{k+1} \in dx' \mid Y_0, \cdots, Y_k] \; = \; p_{k+1}^-(x') \, dx' \; .$$

L'objectif des questions suivantes est d'obtenir des équations pour le calcul récursif des densités conditionnelles $\{p_k\}$, en utilisant les résultats obtenus dans la partie préliminaire.

(vii) **Comment peut–on définir la densité p_0^- ?**

SOLUTION

Nécessairement $p_0^- = p_0^X$.

□

(viii) **Donner l'expression de la densité p_{k+1}^- en fonction de la densité p_k (étape de prédiction).** [Indication : utiliser le résultat obtenu en (iii), avec $X = X_k$, $W = W_k$ et $Z = (Y_0, \cdots, Y_k)$]

SOLUTION

Par hypothèse, $X = X_k$ ne dépend que de $X_0, W_0, \cdots, W_{k-1}$, et $Z = (Y_0, \cdots, Y_k)$ ne dépend que de $X_0, W_0, \cdots, W_{k-1}, V_0, \cdots, V_k$, donc le vecteur aléatoire $W = W_k$ est indépendant du vecteur aléatoire $(X, Z) = (X_k, Y_0, \cdots, Y_k)$. D'autre part, on remarque que $X' = f_k(X) + W = X_{k+1}$.

On peut donc appliquer le résultat obtenu en (iii), et on obtient

$$\begin{aligned} p_{k+1}^-(x') &= p_{X_{k+1} | Y_0, \cdots, Y_k}(x') \\ &= \int_{\mathbb{R}^m} p_{W_k}(x' - f_k(x)) \, p_{X_k | Y_0, \cdots, Y_k}(x) \, dx \\ &= \int_{\mathbb{R}^m} p_k^W(x' - f_k(x)) \, p_k(x) \, dx \; . \end{aligned}$$

□

(ix) **Donner l'expression de la densité p_k en fonction de la densité p_k^- (étape de correction).** [Indication : utiliser le résultat obtenu en (vi), avec $X = X_k$, $V = V_k$ et $Z = (Y_0, \cdots, Y_{k-1})$]

Solution

Par hypothèse, $X = X_k$ ne dépend que de $X_0, W_0, \cdots, W_{k-1}$, et $Z = (Y_0, \cdots, Y_{k-1})$ ne dépend que de $X_0, W_0, \cdots, W_{k-2}, V_0, \cdots, V_{k-1}$, donc le vecteur aléatoire $V = V_k$ est indépendant du vecteur aléatoire $(X, Z) = (X_k, Y_0, \cdots, Y_{k-1})$. D'autre part, on remarque que $Z' = h_k(X) + V = Y_k$.

On peut donc appliquer le résultat obtenu en (vi), et on obtient

$$\begin{aligned} p_k(x) &= p_{X_k|Y_0,\cdots,Y_k}(x') \\ &= c_k \, p_{V_k}(Y_k - h_k(x)) \, p_{X_k|Y_0,\cdots,Y_{k-1}}(x) \\ &= c_k \, p_k^V(Y_k - h_k(x)) \, p_k^-(x) \ , \end{aligned}$$

où la constante de normalisation c_k ne dépend que de $(Y_0, \cdots, Y_k)$.

□

(x) **Décrire l'analogie avec l'équation de Baum forward pour les modèles de Markov cachés.**

Solution

La suite $\{X_k\}$ est une chaîne de Markov dans $\mathbb{R}^m$, de densité de probabilité de transition définie par

$$\mathbb{P}[X_{k+1} \in dx' \mid X_k = x] = q_k(x, x') \, dx' \ .$$

Par définition, et compte tenu que le vecteur aléatoire X_k est indépendant du vecteur aléatoire W_k

$$\begin{aligned} \int_{\mathbb{R}^m} \phi(x') \, q_k(x, x') \, dx' &= \mathbb{E}[\phi(X_{k+1}) \mid X_k = x] \\ &= \mathbb{E}[\phi(f_k(X_k) + W_k) \mid X_k = x] \\ &= \mathbb{E}[\phi(f_k(x) + W_k) \mid X_k = x] \\ &= \int_{\mathbb{R}^m} \phi(f_k(x) + w) \, p_k^W(w) \, dw \ . \end{aligned}$$

En utilisant le changement de variable $x' = f_f(x) + w$, on obtient

$$\int_{\mathbb{R}^m} \phi(x')\, q_k(x,x')\, dx' = \int_{\mathbb{R}^m} \phi(x')\, p_k^W(x' - f_k(x))\, dx' \ .$$

Comme l'égalité a lieu pour toute fonction–test ϕ définie sur $\mathbb{R}^m$, on obtient

$$q_k(x,x') = p_k^W(x' - f_k(x)) \ .$$

L'équation obtenue en (viii) s'écrit

$$p_{k+1}^-(x') = \int_{\mathbb{R}^m} q_k(x,x')\, p_k(x)\, dx = Q_k^*\, p_k(x) \ ,$$

où l'opérateur linéaire Q_k^* est défini par

$$Q_k^*\, p(x) = \int_{\mathbb{R}^m} q_k(x,x')\, p(x)\, dx \ ,$$

pour toute densité de probabilité p définie sur $\mathbb{R}^m$.

D'autre part, l'équation obtenue en (ix) s'écrit

$$p_k(x) = c_k\, p_k^V(Y_k - h_k(x))\, p_k^-(x) = c_k\, \Psi_k(x)\, p_k^-(x) \ ,$$

où la fonction Ψ_k est définie sur $\mathbb{R}^m$ par

$$\Psi_k(x) = p_k^V(Y_k - h_k(x)) \ .$$

□

Cas particulier des systèmes linéaires gaussiens

On se propose de retrouver les équations du filtre de Kalman, dans le cas particulier d'un modèle linéaire gaussien. On suppose dorénavant que les fonctions f_k et h_k sont de la forme

$$f_k(x) = A_k\, x + a_k \qquad \text{et} \qquad h_k(x) = C_k\, x + c_k \ ,$$

et on suppose que

- la densité p_0^X est gaussienne, de moyenne $\bar{X}_0$ et de matrice de covariance Q_0^X,
- pour tout $k \geq 0$, la densité p_k^W est gaussienne, de moyenne nulle et de matrice de covariance Q_k^W,
- pour tout $k \geq 0$, la densité p_k^V est gaussienne, de moyenne nulle et de matrice de covariance Q_k^V.

On suppose également que les matrices Q_0^X, Q_k^W et Q_k^V sont inversibles.

On montre le résultat souhaité par récurrence.

(xi) **En remplacant dans l'équation obtenue en (viii), la densité conditionnelle p_k par la densité gaussienne de moyenne $\widehat{X}_k$ et de matrice de covariance P_k, montrer que la densité conditionnelle p_{k+1}^- est une densité gaussienne de moyenne $\widehat{X}_{k+1}^-$ et de matrice de covariance P_{k+1}^- dont on donnera l'expression en fonction de $\widehat{X}_k$ et P_k.**

Solution

Par définition

$$p_k^W(w) \propto \exp\{ -\tfrac{1}{2}\, w^* \,(Q_k^W)^{-1}\, w \} \ .$$

En supposant que la densité de probabilité conditionnelle p_k est gaussienne, de moyenne $\widehat{X}_k$ et de matrice de covariance P_k, c'est–à–dire

$$p_k(x) \propto \exp\{ -\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) \} \ ,$$

et en remplaçant dans l'équation obtenue en (viii), on obtient

$$\begin{aligned} p_{k+1}^-(x') = \int_{\mathbb{R}^m} & \exp\{ -\tfrac{1}{2}\,(x' - A_k\, x - a_k)^*\,(Q_k^W)^{-1}\,(x' - A_k\, x - a_k) \} \\ & \exp\{ -\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) \}\ dx \ . \end{aligned}$$

On en déduit que la fonction caractéristique Φ_{k+1}^- de la densité de probabilité conditionnelle p_{k+1}^- vérifie

$$\begin{aligned} \Phi_{k+1}^-(u) &= \int_{\mathbb{R}^m} \exp\{ i\, u^*\, x' \}\ p_{k+1}^-(x')\, dx' \\ &\propto \int_{\mathbb{R}^m} \Big[\int_{\mathbb{R}^m} \exp\{ i\, u^*\, x' \} \\ &\qquad\qquad \exp\{ -\tfrac{1}{2}\,(x' - A_k\, x - a_k)^*\,(Q_k^W)^{-1}\,(x' - A_k\, x - a_k) \}\ dx' \Big] \\ &\qquad \exp\{ -\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) \}\ dx \\ &\propto \int_{\mathbb{R}^m} \exp\{ i\, u^*\,(A_k\, x + a_k) - \tfrac{1}{2}\, u^*\, Q_k^W\, u \} \\ &\qquad \exp\{ -\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) \}\ dx \\ &= \exp\{ i\, u^*\, a_k - \tfrac{1}{2}\, u^*\, Q_k^W\, u \}\ \Phi_k(A_k^*\, u) \ , \end{aligned}$$

où Φ_k est la fonction caractéristique de la densité de probabilité conditionnelle p_k, c'est–à–dire

$$\Phi_k(v) = \int_{\mathbb{R}^m} \exp\{\, i\, v^*\, x \,\}\; p_k(x)\, dx = \exp\{\, i\, v^*\, \widehat{X}_k - \tfrac{1}{2}\, v^*\, P_k\, v \,\} \;\; .$$

On en déduit que

$$\Phi^-_{k+1}(u) = \exp\{\, i\, u^*\, a_k - \tfrac{1}{2}\, u^*\, Q^W_k\, u \,\}\; \exp\{\, i\, u^*\, A_k\, \widehat{X}_k - \tfrac{1}{2}\, u^*\, A_k\, P_k\, A^*_k\, u \,\} \;\; ,$$

c'est–à–dire que Φ^-_{k+1} est la fonction caractéristique de la densité de probabilité gaussienne, de moyenne $\widehat{X}^-_{k+1} = A_k\, \widehat{X}_k + a_k$ et de matrice de covariance $P^-_{k+1} = A_k\, P_k\, A^*_k + Q^W_k$.

□

(xii) **En remplacant dans l'équation obtenue en (ix), la densité conditionnelle p^-_k par la densité gaussienne de moyenne $\widehat{X}^-_k$ et de matrice de covariance P^-_k, montrer que la densité conditionnelle p_k est une densité gaussienne de moyenne $\widehat{X}_k$ et de matrice de covariance P_k dont on donnera l'expression en fonction de $\widehat{X}^-_k$ et P^-_k.**

Solution

Par définition

$$p^V_k(v) \propto \exp\{\, -\tfrac{1}{2}\, v^*\, (Q^V_k)^{-1}\, v \,\} \;\; .$$

En supposant que la densité de probabilité conditionnelle p^-_k est gaussienne, de moyenne $\widehat{X}^-_k$ et de matrice de covariance P^-_k, c'est–à–dire

$$p^-_k(x) \propto \exp\{\, -\tfrac{1}{2}\, (x - \widehat{X}^-_k)^*\, (P^-_k)^{-1}\, (x - \widehat{X}^-_k) \,\} \;\; ,$$

et en remplaçant dans l'équation obtenue en (ix), on obtient

$$\begin{aligned}
p_k(x) \;&\propto\; \exp\{\, -\tfrac{1}{2}\, (Y_k - C_k\, x - c_k)^*\, (Q^V_k)^{-1}\, (Y_k - C_k\, x - c_k) \,\} \\
&\qquad \exp\{\, -\tfrac{1}{2}\, (x - \widehat{X}^-_k)^*\, (P^-_k)^{-1}\, (x - \widehat{X}^-_k) \,\} \\
&\propto\; \exp\{\, -\tfrac{1}{2}\, x^*\, C^*_k\, (Q^V_k)^{-1}\, C_k\, x - \tfrac{1}{2}\, x^*\, (P^-_k)^{-1}\, x \\
&\qquad\qquad + x^*\, C^*\, (Q^V_k)^{-1}\, (Y_k - c_k) + x^*\, (P^-_k)^{-1}\, \widehat{X}^-_k \,\} \\
&\propto\; \exp\{\, -\tfrac{1}{2}\, (x - \widehat{X}_k)^*\, P^{-1}_k\, (x - \widehat{X}_k) \,\} \;\; ,
\end{aligned}$$

en posant

$$P^{-1}_k = C^*\, (Q^V_k)^{-1}\, C_k + (P^-_k)^{-1} \;\; ,$$

et

$$C^* \, (Q_k^V)^{-1} \, (Y_k - c_k) + (P_k^-)^{-1} \, \widehat{X}_k^- = P_k^{-1} \, \widehat{X}_k \ .$$

En utilisant le lemme d'inversion matriciel, on obtient

$$P_k = P_k^- - P_k^- \, C_k^* \, [C_k \, P_k^- \, C_k^* + Q_k^V]^{-1} \, C_k \, P_k^- = [I - K_k \, C_k] \, P_k^- \ ,$$

et

$$\widehat{X}_k = \widehat{X}_k^- + K_k \, [Y_k - C_k \, \widehat{X}_k^- - c_k] \ .$$

$\square$