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

Examen du cours
“Filtre de Kalman et modèles de Markov cachés”
Jeudi 8 janvier 1998, 14:00 à 15:30
— Corrigé —

Exercice :

Dans le système GPS (global positioning system) chaque satellite du réseau transmet à intervalle régulier un message indiquant en particulier sa position, et la date. En comparant la date d'émission et la date de réception du message, le récepteur obtient une mesure de sa distance par rapport au satellite.

Les horloges des satellites du réseau sont parfaitement synchrones, mais l'horloge du récepteur peut différer d'une quantité Δ supposée constante au cours du temps, mais inconnue, appelée *offset*.

Pour chaque $i = 1, \cdots, I$, la mesure de pseudo–distance Y_k^i dont dispose le récepteur à l'instant t_k de la part du i–ème satellite, est

$$Y_k^i = \|r^i(t_k^i) - r_k\| + c\,\Delta + V_k^i \ ,$$

où c est la vitesse de la lumière, $r^i(t_k^i)$ est la position du i–ème satellite à l'instant t_k^i où le message a été émis, et r_k est la position du récepteur à l'instant t_k.

On suppose que la suite $\{V_k^i \, , \, k \geq 0\}$ est un bruit blanc gaussien de variance R^i. On note $Y_k = (Y_k^i \, , \, i = 1, \cdots, I\}$ la collection des mesures de pseudo–distance dont dispose le récepteur à l'instant t_k de la part de l'ensemble des satellites du réseau.

(i) **Commenter la terminologie pseudo–distance.**

Solution

Dans l'horloge du satellite, la date de réception du message est $(t_k + \Delta)$, donc l'intervalle de temps nécessaire pour parcourir la véritable distance $\|r^i(t_k^i) - r_k\|$ à la vitesse c est $t_k^i - (t_k + \Delta)$.

Comme l'offset Δ est inconnu, l'intervalle de temps pris en compte par le récepteur est $t_k^i - t_k$, à partir de quoi il construit la mesure de pseudo–distance

$$c\,(t_k^i - t_k) = c\,(t_k^i - (t_k + \Delta)) + c\,\Delta = \|r^i(t_k^i) - r_k\| + c\,\Delta\ .$$

□

On suppose que le mobile portant le récepteur se déplace à une vitesse constante, mais inconnue. On dénote par $X = (r, v, \Delta)$ la variable d'état (position et vitesse du mobile, et offset du récepteur). Il s'agit d'un vecteur de dimension 7.

(ii) **On pose $X_k = X(t_k)$. Ecrire l'équation d'etat, exprimant X_{k+1} en fonction de X_k, sous la forme**

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

c'est–à–dire donner l'expression de la matrice F_k.

Solution

On a

$$\begin{cases} r_{k+1} = r_k + (t_{k+1} - t_k)\,v_k \\ v_{k+1} = v_k \\ \Delta_{k+1} = \Delta_k \end{cases}$$

soit

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

avec

$$F_k = \begin{pmatrix} I & (t_{k+1} - t_k) \cdot I & 0 \\ 0 & I & 0 \\ 0 & 0 & 1 \end{pmatrix}\ ,$$

où I désigne la matrice identité 3×3.

□

Compte tenu que la vitesse du mobile n'est pas exactement constante, on utilisera plutôt l'équation d'etat suivante

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

où la suite $\{W_k\,,\ k \geq 0\}$ est un bruit blanc gaussien de matrice de covariance Q.

On suppose que l'état initial X_0 est un vecteur aléatoire gaussien, de moyenne $\bar{X}_0$, et de matrice de covariance Q_0^X. On suppose que les bruits $\{V_k^i\,,\ k \geq 0\}$ pour $i = 1, \cdots, I$, et $\{W_k\,,\ k \geq 0\}$, et l'état initial X_0 sont mutuellement indépendants.

(iii) **Donner les équations du filtre de Kalman étendu pour l'estimation de l'état X_k basé sur les observations $\mathcal{Y}_k = (Y_0, \cdots, Y_k)$.**

——————————————— SOLUTION ———————————————

Le calcul se fait en deux étapes.

□ **Etape de prédiction** On a

$$\widehat{X}^-_{k+1} = F_k \, \widehat{X}_k$$

$$P^-_{k+1} = F_k \, P_k \, F^*_k + Q \; .$$

□ **Etape de correction** La fonction d'observation h^i_k correspondant à la pseudo–distance du i–ème satellite s'écrit

$$h^i_k(X) = h^i_k(r, v, \Delta) = \| r^i(t^i_k) - r \| + c \, \Delta \; .$$

Le gradient par rapport à la position, à la vitesse, et à l'offset s'écrit respectivement

$$\frac{\partial h^i_k}{\partial r}(r, v, \Delta) = \frac{-[\, r^i(t^i_k) - r \,]^*}{\| r^i(t^i_k) - r \|} \qquad \text{(vecteur–ligne de dimension 3)},$$

$$\frac{\partial h^i_k}{\partial v}(r, v, \Delta) = 0 \qquad \text{(vecteur–ligne de dimension 3)},$$

$$\frac{\partial h^i_k}{\partial \Delta}(r, v, \Delta) = c \qquad \text{(scalaire)}.$$

Il en résulte que le gradient de la fonction d'observation h^i_k, évalué à l'estimateur courant $\widehat{X}^-_k = (\widehat{r}^-_k, \widehat{v}^-_k, \widehat{\Delta}^-_k)$, s'écrit

$$H^i_k = \nabla h^i_k(\widehat{X}^-_k) = \Big(\frac{-[\, r^i(t^i_k) - \widehat{r}^-_k \,]^*}{\| r^i(t^i_k) - \widehat{r}^-_k \|}, 0, c \Big) \qquad \text{(vecteur–ligne de dimension 7)}.$$

On pose

$$h_k = \begin{pmatrix} h^1_k \\ \vdots \\ h^I_k \end{pmatrix} \qquad \text{(vecteur de dimension } I\text{)},$$

$$H_k = \begin{pmatrix} H^1_k \\ \vdots \\ H^I_k \end{pmatrix} \qquad \text{(matrice } I \times 7\text{)},$$

$$R = \begin{pmatrix} R^1 & & \\ & \ddots & \\ & & R^I \end{pmatrix} \qquad \text{(matrice } I \times I \text{ diagonale),}$$

et on a

$$\widehat{X}_k \;=\; \widehat{X}_k^- + K_k\,[Y_k - h_k(\widehat{X}_k^-)]$$

$$P_k \;=\; [I - K_k\,H_k]\,P_k^-$$

où le gain de Kalman K_k est défini par

$$K_k = P_k^-\,H_k^*\,[\,H_k\,P_k^-\,H_k^* + R\,]^{-1}\;.$$

$\square$

Problème :

L'objectif de ce problème est d'étudier la détection d'un changement brusque dans les probabilités de transition d'une chaîne de Markov cachée, et l'estimation de l'instant r de changement (en supposant qu'il ait été décidé qu'un changement a eu lieu). On s'intéressera ici à proposer une méthode *efficace* de calcul de la fonction de vraisemblance associée au problème d'estimation.

On observe la suite $(Y_0, \cdots, Y_n)$ sur un horizon n fixé, et pour chaque valeur $r = 0, \cdots, n$, on considère le modèle $\mathbf{M}(r)$ suivant :

- $\{X_k\,,\, k = 0, \cdots, n\}$ est une chaîne de Markov à espace d'état fini $E = \{1, \cdots, N\}$, de loi initiale
$$\mathbf{P}^{(r)}[X_0 = i] = \nu_i \ ,$$
et de matrice de transition
$$\mathbf{P}^{(r)}[X_{k+1} = j \mid X_k = i] = \pi^k_{i,j}(r) = \begin{cases} \pi_{i,j} & \text{si } k = 0, \cdots, r-1 \\ \\ \bar{\pi}_{i,j} & \text{si } k = r, \cdots, n-1 \end{cases}$$
dépendant du temps.

- La suite des observations $\{Y_k\,,\, k = 0, \cdots, n\}$ est à valeurs dans $\mathbf{R}^d$, la propriété de canal sans mémoire est vérifiée, avec les densités d'observation
$$\mathbf{P}^{(r)}[Y_k \in dy \mid X_k = i] = \psi_i(y)\, dy \ .$$

Pour chaque valeur $r = 0, \cdots, n$, on se propose de calculer la vraisemblance $\ell(r)$ du modèle $\mathbf{M}(r)$, c'est–à–dire de calculer la distribution de probabilité jointe des observations $(Y_0, \cdots, Y_n)$ sous la probabilité $\mathbf{P}^{(r)}$
$$\mathbf{P}^{(r)}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] = \ell(r)\ dy_0 \cdots dy_n \ .$$

Première Méthode

(i) **Pour chaque valeur $r = 0, \cdots, n$, montrer que**
$$\ell(r) = \sum_{i \in E} p^n_i(r) \ ,$$
où la suite $\{p^k(r)\,,\, k = 0, \cdots, n\}$ vérifie une équation forward de Baum dont on donnera l'expression.

Solution

D'après le Théorème 6.4 et la Remarque 6.5 du cours

$$\ell(r) = \sum_{i \in E} p_i^n(r) \ ,$$

où la suite $\{p^k(r) \, , \, k = 0, \cdots, n\}$ vérifie l'équation de Baum forward

$$p^{k+1}(r) = \Psi(Y_{k+1}) \, [\pi^k(r)]^* \, p^k(r) \ ,$$

pour tout $k = 0, \cdots, n-1$, avec la condition initiale

$$p^0(r) = \Psi(Y_0) \, \nu \ .$$

Remarque. On peut décomposer cette équation en deux parties : Avant et après l'instant r de changement. On obtient ainsi

$$p^{k+1}(r) = \Psi(Y_{k+1}) \, \pi^* \, p^k(r) \ ,$$

pour tout $k = 0, \cdots, r-1$, et

$$p^{k+1}(r) = \Psi(Y_{k+1}) \, \bar{\pi}^* \, p^k(r) \ ,$$

pour tout $k = r, \cdots, n-1$.

□

(ii) **Combien d'équations faut–il résoudre avec cette méthode pour calculer $\ell(r)$ pour toutes les valeurs $r = 0, \cdots, n$?**

Solution

Il faut résoudre $(n + 1)$ équations, une pour chaque valeur $r = 0, \cdots, n$.

□

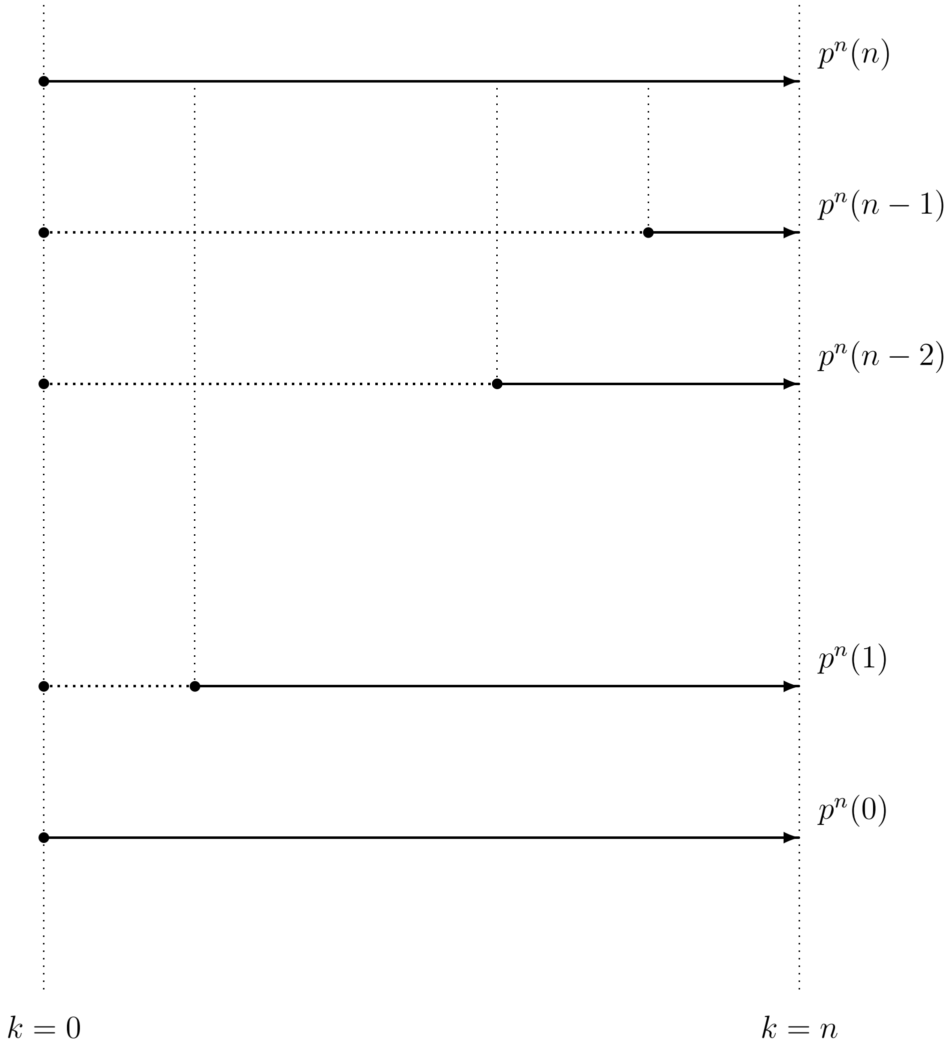
(iii) **Montrer qu'on peut réduire de moitié le volume des calculs nécessaires.**

Solution

On commence par calculer la suite *nominale* $\{p^k(n) \, , \, k = 0, \cdots, n\}$ pour $r = n$, qui correspond au cas où il n'y a pas eu de changement sur l'horizon n.

On remarque que la suite $\{p^k(n-1) \, , \, k = 0, \cdots, n\}$ pour $r = n-1$ coïncide avec la suite nominale jusqu'à l'instant $k = n-1$, et on peut économiser les calculs correspondants. Plus généralement, la suite $\{p^k(r) \, , \, k = 0, \cdots, n\}$ coïncide avec la suite nominale jusqu'à l'instant $k = r$, et on peut économiser les calculs correspondants.

Finalement, on doit calculer la suite $\{p^k(0) \, , \, k = 0, \cdots, n\}$ pour $r = 0$, qui correspond au cas où le changement a déja eu lieu avant l'instant initial.



Dans la méthode telle qu'elle est décrite en (i), on doit résoudre $(n+1)$ équations sur l'horizon n, soit un volume de calcul proportionnel à $n\,(n+1)$.

Dans la variante proposée ici, on doit résoudre l'equation pour la suite *nominale*, et la moitié des calculs correspondant aux n équations restantes, soit un volume de calcul proportionnel à $n + \frac{1}{2}\,n^2 = \frac{1}{2}\,n\,(n+2)$.

$\square$

Pour réduire encore le volume des calculs, on considère la méthode suivante.

Deuxième Méthode

(iv) **Pour chaque valeur $r = 0, \cdots, n$, donner l'expression de l'équation backward de Baum vérifiée par la suite backward $\{v^k(r)\, ,\, k = n, \cdots, 0\}$ associée à $\{p^k(r)\, ,\, k = 0, \cdots, n\}$.**

Solution

D'après le Théorème 6.9 du cours, la suite backward $\{v^k(r)\ ,\ k = n, \cdots, 0\}$ vérifie l'équation backward de Baum

$$v^k(r) = \pi^k(r)\,[\,\Psi(Y_{k+1})\, v^{k+1}(r)\,]\ ,$$

pour tout $k = n-1, \cdots, 0$, avec la condition initiale

$$v^n(r) \equiv 1\ .$$

Remarque. On peut aussi décomposer cette équation en deux parties : Après et avant l'instant r de changement. On obtient ainsi

$$v^k(r) = \bar{\pi}\,[\,\Psi(Y_{k+1})\, v^{k+1}(r)\,]\ ,$$

pour tout $k = n-1, \cdots, r$, et

$$v^k(r) = \pi\,[\,\Psi(Y_{k+1})\, v^{k+1}(r)\,]\ ,$$

pour tout $k = r-1, \cdots, 0$.

□

(v) **Pour chaque valeur $r = 0, \cdots, n$, montrer que la quantité**

$$\sum_{i \in E} p_i^k(r)\ v_i^k(r)$$

ne dépend pas de l'instant k. En déduire que

$$\ell(r) = \sum_{i \in E} p_i^r(r)\ v_i^r(r)\ .$$

Solution

D'après la Proposition 6.10 du cours, les équations forward et backward sont duales l'une de l'autre, et on a

$$\ell(r) = \sum_{i \in E} p_i^n(r) = \sum_{i \in E} p_i^n(r)\ v_i^n(r) = \sum_{i \in E} p_i^k(r)\ v_i^k(r)\ ,$$

pour tout $k = 0, \cdots, n$. En particulier pour $k = r$

$$\ell(r) = \sum_{i \in E} p_i^r(r)\ v_i^r(r)\ .$$

□

(vi) **Pour chaque valeur $r = 0, \cdots, n$, montrer que $p^r(r) = p^r$, où la suite $\{p^k \, , \, k = 0, \cdots, n\}$ vérifie une équation forward de Baum dont on donnera l'expression.**

SOLUTION

On introduit l'équation de Baum forward

$$p^{k+1} = \Psi(Y_{k+1}) \, \pi^* \, p^k \ , \tag{$\star$}$$

pour tout $k = 0, \cdots, n-1$, avec la condition initiale

$$p^0 = \Psi(Y_0) \, \nu \ .$$

Il s'agit de l'équation vérifiée par la suite *nominale* évoquée dans la solution de (iii), et il résulte de la remarque faite à la fin de la solution de (i) que

$$p^k = p^k(r) \ ,$$

pour tout $k = 0, \cdots, r$ et en particulier pour $k = r$, d'où le résultat.

□

(vii) **Pour chaque valeur $r = 0, \cdots, n$, montrer que $v^r(r) = \bar{v}^r$, où la suite $\{\bar{v}^k \, , \, k = n, \cdots, 0\}$ vérifie une équation backward de Baum dont on donnera l'expression.**

SOLUTION

On introduit l'équation de Baum backward

$$\bar{v}^k = \bar{\pi} \, [\, \Psi(Y_{k+1}) \, \bar{v}^{k+1} \,] \ , \tag{$\star\star$}$$

pour tout $k = n-1, \cdots, 0$, avec la condition initiale

$$\bar{v}^n \equiv 1 \ .$$

Il résulte de la remarque faite à la fin de la solution de (iv) que

$$\bar{v}^k = v^k(r) \ ,$$

pour tout $k = n, \cdots, r$ et en particulier pour $k = r$, d'où le résultat.

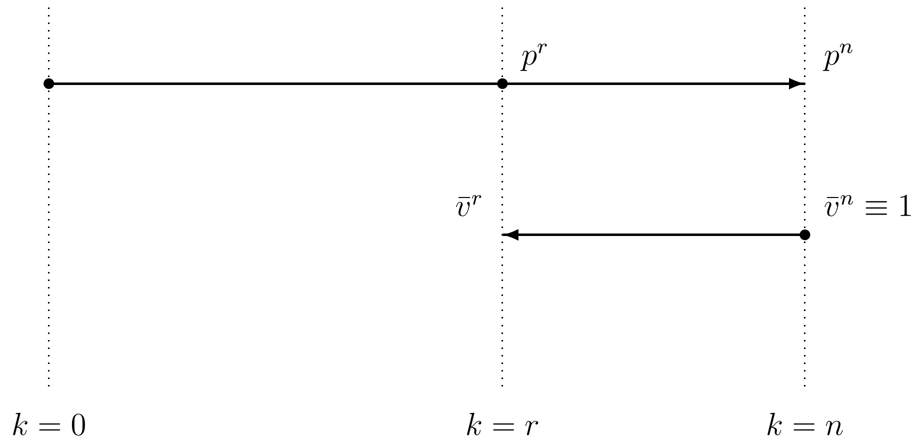
□

(viii) **Combien d'équations suffit–il de résoudre avec cette méthode pour calculer $\ell(r)$ pour toutes les valeurs $r = 0, \cdots, n$?**

Solution

Il suffit de résoudre les deux équations ($\star$) et ($\star\star$). En effet, pour tout $r = 0, \cdots, n$

$$\ell(r) = \sum_{i \in E} p_i^r(r)\; v_i^r(r) = \sum_{i \in E} p_i^r\; \bar{v}_i^r\ .$$



Le volume de calcul est proportionnel à $2\,n$.

□

(ix) **Comparer les deux méthodes du point de vue de la quantité de données à garder en mémoire.**

Solution

Dans la première méthode telle qu'elle est décrite en (i), on résoud séparément $(n+1)$ équations, et il suffit de garder en mémoire à chaque instant la valeur courante de la solution de l'équation considérée, soit un vecteur de dimension N.

Dans la variante proposée en (iii), on doit résoudre l'équation ($\star$) pour la suite *nominale*, et garder en mémoire la suite entière, soit $(n+1)$ vecteurs de dimension N. Il faut ensuite résoudre séparément les n équations restantes, et il suffit de garder en mémoire à chaque instant la valeur courante de la solution de l'équation considérée, soit un vecteur de dimension N. Au total on doit garder en mémoire $(n+2)$ vecteurs de dimension N.

Dans la seconde méthode, on doit résoudre l'équation ($\star$) pour la suite *nominale*, et garder en mémoire la suite entière, soit $(n+1)$ vecteurs de dimension N. On peut ensuite résoudre séparément l'équation ($\star\star$), et il suffit de garder en mémoire à chaque instant la valeur courante de la solution, soit un vecteur de dimension N. Au total on doit garder en mémoire $(n+2)$ vecteurs de dimension N.

□