Université de Rennes 1
Master EEEA (parcours SISEA)

Examen du cours
"Filtre de Kalman et modèles de Markov cachés"
Jeudi 22 février 2018, 10:15 à 12:15
— Corrigé —

Le but de ce problème est de proposer des formulations équivalentes pour le lissage de Kalman, et de comparer les mérites respectifs des différentes formulations, en terme de temps de calcul (nombre d'opérations mises en œuvre, matrices à inverser, etc.), de taille mémoire (nombre de variables calculées à conserver), d'utilisation/ré–utilisation des observations, d'hypothèses nécessaires.

Pour fixer les notations, on considère une suite d'états cachés $\{X_k\}$ à valeurs dans $\mathbb{R}^m$, vérifiant

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

où la suite $\{W_k\}$ prend aussi ses valeurs dans $\mathbb{R}^m$, et une suite d'observations $\{Y_k\}$ à valeurs dans $\mathbb{R}^d$, vérifiant

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

où la suite $\{V_k\}$ prend aussi ses valeurs dans $\mathbb{R}^d$, et on suppose que

- la condition initiale 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, de matrice de covariance Q_k^W,
- la suite $\{V_k\}$ est un bruit blanc gaussien, de matrice de covariance Q_k^V,
- les suites $\{W_k\}$ et $\{V_k\}$ et la condition initiale X_0 sont mutuellement indépendants.

Pour mémoire, sous l'hypothèse que

la matrice de covariance Q_k^V est inversible à chaque instant,

le meilleur estimateur (au sens du minimum de l'erreur quadratique moyenne) de l'état caché X_k sachant les observations passées $Y_{0:k} = (Y_0, \cdots, Y_k)$ jusqu'à l'instant courant seulement, est le *filtre* de Kalman $\widehat{X}_k$ donné par les équations suivantes : prédiction

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

$$P_k^- \ = \ F_k\, P_{k-1}\, F_k^* + Q_k^W \ ,$$

et correction

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

$$P_k \;=\; (I - K_k\,H_k)\;P_k^-\ \ ,$$

avec la matrice de gain de Kalman définie par

$$K_k = P_k^-\;H_k^*\,(Q_k^I)^{-1} \qquad \text{où} \qquad Q_k^I = H_k\,P_k^-\;H_k^* + Q_k^V\ ,$$

et avec les conditions initiales (formulées pour $k = 0$)

$$\widehat{X}_0^- = \bar{X}_0 \qquad \text{et} \qquad P_0^- = Q_0^X\ .$$

(i) Établir les identités (très simples) suivantes

$$(P_k^-)^{-1}\;P_k = (I - K_k\,H_k)^* \qquad \textbf{et} \qquad (P_k^-)^{-1}\;K_k = H_k^*\;(Q_k^I)^{-1}\ .$$

Solution

On rappelle que

$$P_k = (I - K_k\,H_k)\;P_k^-\ \ ,$$

de sorte que

$$P_k\;(P_k^-)^{-1} = I - K_k\,H_k\ \ ,$$

et par transposition on obtient

$$(P_k^-)^{-1}\;P_k = (I - K_k\,H_k)^*\ .$$

Par définition

$$K_k = P_k^-\;H_k^*\,(Q_k^I)^{-1}\ ,$$

de sorte que immédiatement

$$(P_k^-)^{-1}\;K_k = H_k^*\,(Q_k^I)^{-1}\ .$$

□

De même, sous l'hypothèse que

les matrices de covariance Q_k^W et Q_k^V sont inversibles à chaque instant, le meilleur estimateur (au sens de l'erreur quadratique moyenne) de l'état caché X_k sachant toutes les observations $Y_{0:n} = (Y_0, \cdots, Y_n)$, est le *lisseur* de Kalman $\widehat{X}_k^n$ donné par les équations rétrogrades suivantes

$$\widehat{X}_{k-1}^n = \widehat{X}_{k-1} + L_k \, (\widehat{X}_k^n - \widehat{X}_k^-) \; , \tag{1}$$

$$P_{k-1}^n = P_{k-1} + L_k \, (P_k^n - P_k^-) \, L_k^* \; ,$$

avec la matrice de gain

$$L_k = P_{k-1} \, F_k^* \, (P_k^-)^{-1} \; ,$$

et avec les conditions initiales (formulées pour $k = n$)

$$\widehat{X}_n^n = \widehat{X}_n \qquad \text{et} \qquad P_n^n = P_n \; .$$

Intuitivement, l'incertitude associée au lisseur de Kalman devrait être inférieure à l'incertitude associée au filtre de Kalman.

(ii) Montrer (par exemple par récurrence rétrograde) que les matrices de covariance d'erreur P_k et P_k^n, associées au filtre de Kalman et au lisseur de Kalman respectivement, vérifient la relation $P_k^n \leq P_k$ au sens des matrices symétriques.

Solution

Par définition $P_n^n = P_n$, c'est–à–dire que la relation est vraie au rang $k = n$.

Si la relation est vraie au rang k, c'est–à–dire si $P_k^n \leq P_k$, alors nécessairement $P_k^n \leq P_k^-$ puisque $P_k \leq P_k^-$. En d'autres termes, la différence $(P_k^n - P_k^-)$ est semi–définie négative, de sorte que la différence

$$P_{k-1}^n - P_{k-1} = L_k \, (P_k^n - P_k^-) \, L_k^* \; ,$$

aussi est semi–définie négative. En d'autres termes, $P_{k-1}^n \leq P_{k-1}$, c'est–à–dire que la relation est vraie au rang $(k-1)$.

□

Pour tout $k = 0, 1, \cdots, n$, on introduit les variables

$$r_k^n = F_k^* \, (P_k^-)^{-1} \, (\widehat{X}_k^n - \widehat{X}_k^-) \qquad \text{et} \qquad \Pi_k^n = F_k^* \, (P_k^-)^{-1} \, (P_k^n - P_k^-) \, (P_k^-)^{-1} \, F_k \; .$$

(iii) Vérifier que la matrice Π_k^n est symétrique et semi–définie négative.

Solution

Clairement, la matrice Π_k^n est symétrique. Et puisque la matrice $(P_k^n - P_k^-)$ est semi–définie négative, alors nécessairement la matrice Π_k^n aussi est semi–définie négative.

□

(iv) Montrer que nécessairement le lisseur de Kalman $\widehat{X}_k^n$ et la matrice de covariance d'erreur associée P_k^n vérifient les relations suivantes

$$\widehat{X}_k^n = \widehat{X}_k + P_k \, r_{k+1}^n \ , \tag{2}$$

$$P_k^n = P_k + P_k \, \Pi_{k+1}^n \, P_k \ .$$

SOLUTION

En écrivant l'équation (1) à l'instant k, on obtient

$$\begin{aligned}
\widehat{X}_k^n &= \widehat{X}_k + L_{k+1} \, (\widehat{X}_{k+1}^n - \widehat{X}_{k+1}^-) \\
&= \widehat{X}_k + P_k \, F_{k+1}^* \, (P_{k+1}^-)^{-1} \, (\widehat{X}_{k+1}^n - \widehat{X}_{k+1}^-) \\
&= \widehat{X}_k + P_k \, r_{k+1}^n \ .
\end{aligned}$$

De même

$$\begin{aligned}
P_k^n &= P_k + L_{k+1} \, (P_{k+1}^n - P_{k+1}^-) \, L_{k+1}^* \\
&= P_k + P_k \, F_{k+1}^* \, (P_{k+1}^-)^{-1} \, (P_{k+1}^n - P_{k+1}^-) \, (P_{k+1}^-)^{-1} \, F_{k+1} \, P_k \\
&= P_k + P_k \, \Pi_{k+1}^n \, P_k \ .
\end{aligned}$$

□

(v) Exprimer la différence $(\widehat{X}_k^n - \widehat{X}_k^-)$ en fonction de l'innovation $(Y_k - H_k \, \widehat{X}_k^-)$ et de la variable r_{k+1}^n introduite.

SOLUTION

D'après l'étape de correction du filtre de Kalman, on a

$$\widehat{X}_k = \widehat{X}_k^- + K_k \, (Y_k - H_k \, \widehat{X}_k^-) \ ,$$

et en reportant cette expression dans l'équation (2), on obtient

$$\widehat{X}_k^n = \widehat{X}_k + P_k \, r_{k+1}^n = \widehat{X}_k^- + K_k \, (Y_k - H_k \, \widehat{X}_k^-) + P_k \, r_{k+1}^n \ ,$$

ou de manière équivalente, par différence

$$\widehat{X}_k^n - \widehat{X}_k^- = K_k \, (Y_k - H_k \, \widehat{X}_k^-) + P_k \, r_{k+1}^n \ .$$

□

Pour tout $k = 0, 1, \cdots, n$, on définit la matrice

$$\Phi_k = (I - K_k \, H_k) \, F_k \ .$$

(vi) Montrer que la variable r_k^n vérifie l'équation rétrograde suivante

$$r_k^n = F_k^* \, H_k^* \, (Q_k^I)^{-1} \, (Y_k - H_k \, \widehat{X}_k^-) + \Phi_k^* \, r_{k+1}^n \;, \tag{3}$$

avec la condition initiale $r_{n+1}^n = 0$ par convention.

Solution

En reportant dans la définition de r_k^n l'expression obtenue en réponse à la question (v), on obtient

$$\begin{aligned} r_k^n &= F_k^* \, (P_k^-)^{-1} \, (\widehat{X}_k^n - \widehat{X}_k^-) \\ &= F_k^* \, (P_k^-)^{-1} \, K_k \, (Y_k - H_k \, \widehat{X}_k^-) + F_k^* \, (P_k^-)^{-1} \, P_k \, r_{k+1}^n \\ &= F_k^* \, H_k^* \, (Q_k^I)^{-1} \, (Y_k - H_k \, \widehat{X}_k^-) + F_k^* \, (I - K_k \, H_k)^* \, r_{k+1}^n \;, \end{aligned}$$

compte tenu des identités

$$(P_k^-)^{-1} \, K_k = H_k^* \, (Q_k^I)^{-1} \qquad \text{et} \qquad P_k \, (P_k^-)^{-1} = I - K_k \, H_k \;,$$

et on reconnait l'expression de la matrice $\Phi_k = (I - K_k \, H_k) \, F_k$.

□

(vii) En utilisant la même démarche, montrer que la variable Π_k^n vérifie l'équation rétrograde suivante

$$\Pi_k^n = -F_k^* \, H_k^* \, (Q_k^I)^{-1} \, H_k \, F_k + \Phi_k^* \, \Pi_{k+1}^n \, \Phi_k \;,$$

avec la condition initiale $\Pi_{n+1}^n = 0$ par convention.

Solution

En procédant comme dans la réponse à la question (v), d'après l'étape de correction du filtre de Kalman, on a

$$P_k = P_k^- - P_k^- \, H_k^* \, (Q_k^I)^{-1} \, H_k \, P_k^- \;,$$

et en reportant cette expression plus haut, on obtient

$$P_k^n = P_k + P_k \, \Pi_{k+1}^n \, P_k = P_k^- - P_k^- \, H_k^* \, (Q_k^I)^{-1} \, H_k \, P_k^- + P_k \, \Pi_{k+1}^n \, P_k \;,$$

ou de manière équivalente, par différence

$$P_k^n - P_k^- = -P_k^- \, H_k^* \, (Q_k^I)^{-1} \, H_k \, P_k^- + P_k \, \Pi_{k+1}^n \, P_k \;.$$

En reportant cette expression dans la définition de Π_k^n, on obtient

$$\begin{aligned}\Pi_k^n &= F_k^* \, (P_k^-)^{-1} \, (P_k^n - P_k^-) \, (P_k^-)^{-1} \, F_k \\ &= -F_k^* \, (P_k^-)^{-1} \, P_k^- \, H_k^* \, (Q_k^I)^{-1} \, H_k \, P_k^- \, (P_k^-)^{-1} \, F_k \\ &\quad + F_k^* \, (P_k^-)^{-1} \, P_k \, \Pi_{k+1}^n \, P_k \, (P_k^-)^{-1} \, F_k \\ &= -F_k^* \, H_k^* \, (Q_k^I)^{-1} \, H_k \, F_k + F_k^* \, (I - K_k \, H_k)^* \, \Pi_{k+1}^n \, (I - K_k \, H_k) \, F_k \ ,\end{aligned}$$

compte tenu de l'identité

$$(P_k^-)^{-1} \, P_k = (I - K_k \, H_k)^* \ ,$$

et on reconnait l'expression de la matrice $\Phi_k = (I - K_k \, H_k) \, F_k$.

$\square$

(viii) Comparer les deux approches équivalentes obtenues pour le lissage de Kalman, (1) vs. (2)–(3), selon les critères suivants : volume des calculs impliqués, volume de l'espace mémoire nécessaire, inversion de matrices, hypothèses nécessaires, etc.

Solution

Les deux approches partagent la même phase aller, qui comprend le calcul du filtre de Kalman $\widehat{X}_k$ et de la matrice de covariance d'erreur associée P_k. Une condition nécessaire pour cette phase aller est l'inversibilité de la matrice de covariance $Q_k^I = H_k \, P_k^- \, H_k^* + Q_k^V$ de dimension $d \times d$, et une condition suffisante est l'inversibilité de la matrice de covariance Q_k^V, une donnée du problème. Le calcul de la matrice inverse n'est pas nécessaire, mais la résolution de systèmes linéaires de dimension d de la forme $Q_k^I \, y = b$ est requise, et passe par exemple par la décomposition de Cholesky de la matrice Q_k^I.

Avec l'approche (1), une condition nécessaire pour la phase retour est l'inversibilité de la matrice de covariance $P_k^- = F_k \, P_{k-1} \, F_k^* + Q_k^W$ de dimension $m \times m$, et une condition suffisante est l'inversibilité de la matrice de covariance Q_k^W, une donnée du problème. Le calcul de la matrice inverse n'est pas nécessaire, mais la résolution de systèmes linéaires de dimension m de la forme $P_k^- \, x = b$ est requise, et passe par exemple par la décomposition de Cholesky de la matrice P_k^-. L'équation de récurrence rétrograde (1) pour le calcul du lisseur $\widehat{X}_{k-1}^n$ utilise les valeurs numériques du filtre $\widehat{X}_{k-1}$ et de la matrice de covariance d'erreur associée P_{k-1} (à partir desquelles il est facile de reconstruire les valeurs numériques du prédicteur $\widehat{X}_k^-$ et de la matrice de covariance d'erreur associée P_k^-). Ces valeurs numériques sont calculées dans la phase aller, et doivent donc être conservées en mémoire pour être utilisées dans la phase retour. En revanche, l'équation de récurrence rétrograde (1) pour le calcul du lisseur $\widehat{X}^n_{k-1}$ n'utilise ni la valeur numérique de l'observation Y_k ni celle de l'innovation $I_k = Y_k - H_k\, \widehat{X}^-_k$.

Avec l'approche (2)–(3), il n'y a pas de condition nécessaire d'inversibilité pour la phase retour qui ne soit pas déjà nécessaire pour la phase aller. L'expression (2) pour le calcul du lisseur $\widehat{X}^n_k$ utilise les valeurs numériques du filtre $\widehat{X}_k$ et de la matrice de covariance d'erreur associée P_k. Ces valeurs numériques sont calculées dans la phase aller, et doivent donc être conservées en mémoire pour être utilisées dans la phase retour. L'équation de récurrence rétrograde (3) pour le calcul de la variable r^n_k utilise la valeur numérique de l'observation Y_k ou de manière équivalente celle de l'innovation $I_k = Y_k - H_k\, \widehat{X}^-_k$. Ces valeurs numériques sont calculées dans la phase aller, et doivent donc être conservées en mémoire pour être utilisées dans la phase retour.

En conclusion :

- les deux approchent requièrent dans la phase aller une même condition d'inversibilité et l'inversion de systèmes linéaires de dimension d,
- l'approche (1) requiert dans la phase retour une condition d'inversibilité supplémentaire et l'inversion de systèmes linéaires de dimension m, tandis que l'approche (2)–(3) ne requiert aucune condition d'inversibilité supplémentaire,
- les deux approches utilisent dans la phase retour les valeurs numériques du filtre et de la matrice de covariance d'erreur associée — ces valeurs numériques sont calculées dans la phase aller, et doivent donc être conservées en mémoire pour être utilisées dans la phase retour,
- l'approche (2)–(3) utilise dans la phase retour la valeur numérique de l'observation ou de manière équivalente celle de l'innovation, tandis que l'approche (1) n'utilise aucune de ces valeurs numériques — ces valeurs numériques sont calculées dans la phase aller, et doivent donc être conservées en mémoire pour être utilisées dans la phase retour.

$\square$