École Nationale de la Statistique
et de l'Analyse de l'Information
examen du cours
"Filtrage linéaire et non–linéaire"

vendredi 24 janvier 2020, 14:00 à 16:00

Borne de Cramér–Rao a posteriori et filtre de Kalman

On considère le système linéaire gaussien

$$\begin{aligned} X_k &= F_k\, X_{k-1} + W_k \\ Y_k &= H_k\, X_k + V_k \end{aligned} \qquad (\star)$$

avec les conditions habituelles

- l'état initial 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 (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^W à l'instant k,
- la suite $\{V_k\}$ est un bruit blanc gaussien (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^V à l'instant k,
- l'état initial X_0 et les suites $\{W_k\}$ et $\{V_k\}$ sont mutuellement indépendants.

On suppose en outre que

- la matrice de covariance Q_0^X est inversible et les matrices de covariance Q_k^W et Q_k^V sont inversibles, à tout instant k.

On se propose de montrer la propriété suivante

> la matrice de covariance P_k^- de l'erreur de prédiction et la matrice de covariance P_k de l'erreur d'estimation (données par les équations du filtre de Kalman) sont inversibles, à tout instant k,

et d'établir une équation récurrente pour les matrices inverses, notées I_k^- et I_k respectivement (attention à ne pas confondre avec la même notation utilisée pour le processus d'innovation). On pourra utiliser en particulier le lemme d'inversion matricielle vu en cours.

(i) **Montrer que la matrice P_0^- est inversible, et donner l'expression de la matrice inverse I_0^-.**

SOLUTION

Par hypothèse, la matrice de covariance Q_0^X est inversible, c'est–à–dire que la matrice $P_0^- = Q_0^X$ est inversible et $I_0^- = (Q_0^X)^{-1}$.

$\square$

(ii) **Pour tout instant $k \geq 0$, montrer que si la matrice P_k^- est inversible, alors la matrice P_k est inversible, et donner l'expression de la matrice inverse I_k en fonction de la matrice I_k^-.**

SOLUTION

En guise de préliminaire, on rappelle que si une matrice symétrique M est définie positive, donc inversible, alors son inverse M^{-1} est elle–même symétrique et définie positive.

Par hypothèse, la matrice de covariance Q_k^V est inversible, donc définie positive, et a fortiori la matrice de covariance

$$H\, P_k^-\, H_k^* + Q_k^V \ ,$$

est définie positive, donc inversible. Si en outre la matrice de covariance P_k^- est inversible, donc définie positive, alors la matrice inverse $I_k^- = (P_k^-)^{-1}$ est elle–même symétrique et définie positive, et a fortiori la matrice symétrique

$$H_k^*\, (Q_k^V)^{-1}\, H_k + I_k^- \ ,$$

est définie positive, donc inversible. Il résulte alors du lemme d'inversion matricielle que

$$(H_k^*\, (Q_k^V)^{-1}\, H_k + I_k^-)^{-1} = P_k^- - P_k^-\, H_k^*\, (H\, P_k^-\, H_k^* + Q_k^V)^{-1}\, H_k\, P_k^- = P_k \ ,$$

c'est–à–dire que la matrice de covariance P_k est inversible et

$$I_k = I_k^- + H_k^*\, (Q_k^V)^{-1}\, H_k \ .$$

$\square$

(iii) **Pour tout instant $k \geq 1$, montrer que si la matrice P_{k-1} est inversible, alors la matrice P_k^- est inversible, et donner l'expression de la matrice inverse I_k^- en fonction de la matrice I_{k-1}.**

Solution

Par hypothèse, la matrice de covariance Q_k^W est inversible, donc définie positive, et a fortiori la matrice de covariance

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

est définie positive, donc inversible. Si en outre la matrice de covariance P_{k-1} est inversible, donc définie positive, alors la matrice inverse $I_{k-1} = P_{k-1}^{-1}$ est elle–même symétrique et définie positive, et a fortiori la matrice symétrique

$$I_{k-1} + F_k^* \, (Q_k^W)^{-1} \, F_k \ ,$$

est définie positive, donc inversible. Il résulte alors du lemme d'inversion matricielle que

$$\begin{aligned} I_k^- &= (F_k \, I_{k-1}^{-1} \, F_k^* + Q_k^W)^{-1} \\ &= (Q_k^W)^{-1} - (Q_k^W)^{-1} \, F_k \, (I_{k-1} + F_k^* \, (Q_k^W)^{-1} \, F_k)^{-1} \, F_k^* \, (Q_k^W)^{-1} \ . \end{aligned}$$

□

On considère ensuite le système non–linéaire (mais avec des bruits gaussiens)

$$X_k = b_k(X_{k-1}) + W_k$$

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

avec les conditions habituelles

- l'état initial 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 (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^W à l'instant k,
- la suite $\{V_k\}$ est un bruit blanc gaussien (une suite de vecteurs aléatoires gaussiens indépendants et centrés) de matrice de covariance Q_k^V à l'instant k,
- l'état initial X_0 et les suites $\{W_k\}$ et $\{V_k\}$ sont mutuellement indépendants,

- les fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$ sont dérivables.

On suppose en outre (comme dans le cas du système linéaire gaussien décrit en ($\star$)) que

> la matrice de covariance Q_0^X est inversible et les matrices de covariance Q_k^W et Q_k^V sont inversibles, à tout instant k.

(iv) **Rappeler l'équation récurrente vue en cours vérifiée par la matrice d'information de Fisher J_k, avec l'expression des matrices D_k^{11}, D_k^{12}, D_k^{22} et E_k. On ne demande pas l'expression générale, mais l'expression explicite à l'aide des matrices jacobiennes des fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$.**

Solution

On a vu en cours que la matrice d'information de Fisher vérifie l'équation récurrente

$$J_k^- = D_k^{22} - D_k^{21}\,(J_{k-1} + D_k^{11})^{-1}\,D_k^{12} \qquad \text{et} \qquad J_k = J_k^- + E_k\ ,$$

avec

$$\begin{aligned}
D_k^{11} &= \mathbb{E}[\,(b'_k(X_{k-1}))^*\,(Q_k^W)^{-1}\,b'_k(X_{k-1})\,]\ ,\\
D_k^{12} &= -\mathbb{E}[\,(b'_k(X_{k-1}))^*\,]\,(Q_k^W)^{-1}\ ,\\
D_k^{22} &= (Q_k^W)^{-1}\ ,\\
E_k &= \mathbb{E}[\,(h'_k(X_k))^*\,(Q_k^V)^{-1}\,h'_k(X_k)\,]\ ,
\end{aligned}$$

et avec l'initialisation

$$J_0^- = -\mathbb{E}[\,\frac{\partial^2}{\partial x_0^2}\log p_0(X_0)\,] \qquad \text{et} \qquad J_0 = J_0^- + E_0\ .$$

□

(v) **Montrer que l'initialisation est donnée par $J_0^- = (Q_0^X)^{-1}$.**

Solution

Par hypothèse, la matrice de covariance Q_0^X est inversible, et il existe donc une densité pour la loi initiale, définie par

$$p_0(x_0) = \frac{1}{\sqrt{\det(2\pi\,Q_0^X)}}\ \exp\{-\tfrac{1}{2}\,(x_0 - \bar{X}_0)^*\,(Q_0^X)^{-1}\,(x_0 - \bar{X}_0)\}\ .$$

L'opposé de la log–densité s'écrit

$$- \log p_0(x_0) = \frac{1}{2} \, (x_0 - \bar{X}_0)^* \, (Q_0^X)^{-1} \, (x_0 - \bar{X}_0) + \mathrm{cste} \; ,$$

et la matrice hessienne

$$- \frac{\partial^2}{\partial x_0^2} \log p_0(x_0) = (Q_0^X)^{-1} \; ,$$

ne dépend pas de la variable x_0. On en déduit que l'initialisation est donnée par

$$J_0^- = - \mathbb{E}[\, \frac{\partial^2}{\partial x_0^2} \log p_0(X_0) \,] = (Q_0^X)^{-1} \; .$$

(vi) **Donner l'expression des matrices D_k^{11}, D_k^{12}, D_k^{22} et E_k dans le cas particulier d'un système linéaire gaussien comme celui décrit en $(\star)$, où les fonctions $x \mapsto b_k(x)$ et $x \mapsto h_k(x)$ sont linéaires.**

Ré–écrire l'équation récurrente vérifiée par la matrice d'information de Fisher J_k dans ce cas particulier.

Solution

Dans le cas particulier d'un système linéaire gaussien avec $b_k(x) = F_k \, x$ et $h_k(x) = H_k \, x$, les matrices jacobiennes $b'_k(x) = F_k$ et $h'_k(x) = H_k$ ne dépendent pas de la variable x, de sorte que

$$\begin{aligned}
D_k^{11} &= \mathbb{E}[\, (b'_k(X_{k-1}))^* \, (Q_k^W)^{-1} \, b'_k(X_{k-1}) \,] = F_k^* \, (Q_k^W)^{-1} \, F_k \; , \\
D_k^{12} &= - \mathbb{E}[\, (b'_k(X_{k-1}))^* \,] \, (Q_k^W)^{-1} = - F_k^* \, (Q_k^W)^{-1} \; , \\
D_k^{22} &= (Q_k^W)^{-1} \; , \\
E_k &= \mathbb{E}[\, (h'_k(X_k))^* \, (Q_k^V)^{-1} \, h'_k(X_k) \,] = H_k^* \, (Q_k^V)^{-1} \, H_k \; ,
\end{aligned}$$

et l'équation récurrente vérifiée par la matrice d'information de Fisher devient

$$J_k^- = (Q_k^W)^{-1} - (Q_k^W)^{-1} \, F_k \, (J_{k-1} + F_k^* \, (Q_k^W)^{-1} \, F_k)^{-1} \, F_k^* \, (Q_k^W)^{-1} \; ,$$

et

$$J_k = J_k^- + H_k^* \, (Q_k^V)^{-1} \, H_k \; .$$

(vii) **En déduire que, dans le cas particulier d'un système linéaire gaussien, la matrice d'information de Fisher J_k coïncide avec la matrice I_k, étudiée aux questions (i) à (iii) et définie comme l'inverse de la matrice de covariance P_k de l'erreur d'estimation (donnée par les équations du filtre de Kalman).**

Interpréter la borne de Cramér–Rao a posteriori ainsi obtenue.

Solution

On vérifie que l'initialisation J_0^- établie en réponse à la question (v) coïncide avec l'initialisation I_0^- établie en réponse à la question (i).

On constate que l'équation récurrente

$$J_k^- = (Q_k^W)^{-1} - (Q_k^W)^{-1} F_k \, (J_{k-1} + F_k^* \, (Q_k^W)^{-1} F_k)^{-1} F_k^* \, (Q_k^W)^{-1} \ ,$$

et

$$J_k = J_k^- + H_k^* \, (Q_k^V)^{-1} H_k \ ,$$

établie en réponse à la question (vi) et vérifiée par la matrice d'information de Fisher, coïncide avec l'équation récurrente

$$I_k^- = (Q_k^W)^{-1} - (Q_k^W)^{-1} F_k \, (I_{k-1} + F_k^* \, (Q_k^W)^{-1} F_k)^{-1} F_k^* \, (Q_k^W)^{-1} \ ,$$

et

$$I_k = I_k^- + H_k^* \, (Q_k^V)^{-1} H_k \ ,$$

établie en réponse aux questions (iii) et (ii) et vérifiée par l'inverse de la matrice de covariance de l'erreur d'estimation.

Il en résulte que la matrice d'information de Fisher J_n coïncide avec l'inverse de la matrice de covariance P_n de l'erreur d'estimation, ou de manière équivalente, l'inverse J_n^{-1} de la matrice d'information de Fisher coïncide avec la matrice de covariance P_n de l'erreur d'estimation. La borne de Cramér–Rao a posteriori devient donc

$$\mathbb{E}[\, (\psi(Y_{0:n}) - X_n) \, (\psi(Y_{0:n}) - X_n)^* \,] \geq P_n \ ,$$

et cette borne, valable pour un estimateur quelconque, est atteinte pour le filtre de Kalman comme estimateur particulier.

$\square$