Université de Rennes 1
Master Recherche SISEA

Examen du cours UE S3-2
“Filtre de Kalman et modèles de Markov cachés”
Vendredi 25 janvier 2013, 10:15 à 12:15
— Corrigé —

Problème

Le but de ce problème est d'établir les équations du lisseur de Kalman pour le système linéaire gaussien suivant :

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

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

sous les hypothèses habituelles :

- *X_0 variable aléatoire gaussienne.*
- *$\{W_k\}$ bruit blanc gaussien centré de matrice de covariance Q^W_k inversible.*
- *$\{V_k\}$ bruit blanc gaussien centré de matrice de covariance Q^V_k inversible.*
- *X_0, $\{W_k\}$ et $\{V_k\}$ mutuellement indépendants.*

Par rapport aux hypothèses du cours, on suppose donc que le bruit de modèle aussi possède une matrice de covariance *inversible*.

Soit n un instant final *fixé* : on souhaite calculer la loi conditionnelle de la variable aléatoire X_k sachant $Y_{0:n} = (Y_0, \cdots, Y_n)$, pour tout instant k compris entre 0 et n. Il existe plusieurs approches pour traiter ce problème, qui donnent lieu à des formulations différentes pour le même résultat. L'approche proposée ici reproduit celle vue en cours pour les modèles de Markov cachés.

On denote par $\widehat{X}^n_k$ et par P^n_k la moyenne et la matrice de covariance de cette loi conditionnelle, i.e.

$$\widehat{X}^n_k = \mathbb{E}[X_k \mid Y_{0:n}] \qquad \text{et} \qquad P^n_k = \mathbb{E}[\,(X_k - \widehat{X}^n_k)\,(X_k - \widehat{X}^n_k)^*\,] \;,$$

respectivement, et on dénote par $\widehat{X}_k$ et P_k la moyenne et la matrice de covariance du filtre de Kalman, pour tout $k = 0, \cdots, n$.

Généralités

Dans cette première partie, on établit des propriétés qui seront utiles pour la suite, et qui sont valables en toute généralité, c'est–à–dire qui ne dépendent pas du caractère gaussien. On rappelle la définition du *filtre* bayésien vu en cours

$$\mu_k(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:k}] \ ,$$

et on introduit aussi le *lisseur* bayésien, défini par

$$\mu_k^n(dx) = \mathbb{P}[X_k \in dx \mid Y_{0:n}] \ .$$

On admet que la distribution de probabilité conjointe des états cachés $X_{0:n} = (X_0, \cdots, X_n)$ et des observations $Y_{0:n} = (Y_0, \cdots, Y_n)$ peut s'écrire

$$\mathbb{P}[X_{0:n} \in dx_{0:n}, Y_{0:n} \in dy_{0:n}\,] = \eta_0(dx_0) \ \prod_{k=1}^{n} Q_k(x_{k-1}, dx_k) \ \prod_{k=0}^{n} g_k(x_k, y_k) \ dy_{0:k} \ ,$$

en fonction de la distribution de probabilité initiale

$$\mathbb{P}[X_0 \in dx] = \eta_0(dx) \ ,$$

des noyaux de probabilité de transition

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

pour tout $k = 1, \cdots, n$, et des densités d'émission

$$\mathbb{P}[Y_k \in dy \mid X_k = x] = g_k(x, y) \ dy \ ,$$

pour tout $k = 0, 1, \cdots, n$ (on ne demande pas de démontrer ce résultat).

(i) **Montrer que**

$$\mathbb{P}[X_{0:n} \in dx_{0:n}, Y_{0:n} \in dy_{0:n}\,]$$
$$= \mathbb{P}[X_{0:k} \in dx_{0:k}, Y_{0:k} \in dy_{0:k}\,] \ \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p) \ dy_{k+1:n} \ .$$

Solution

En regroupant les facteurs de 0 à k d'une part, et les facteurs de $k+1$ à n d'autre part, on décompose la distribution de probabilité conjointe des états cachés $X_{0:n}$ et des observations $Y_{0:n}$ en produit de deux facteurs

$$\mathbb{P}[X_{0:n} \in dx_{0:n}, Y_{0:n} \in dy_{0:n}\,] \ = \ [\, \eta_0(dx_0) \ \prod_{p=1}^{k} Q_p(x_{p-1}, dx_p) \ \prod_{p=0}^{k} g_p(x_p, y_p) \ dy_{0:k}\,]$$
$$[\, \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p)\,] \ dy_{k+1:n} \ ,$$

et on remarque que le premier facteur est juste la distribution de probabilité conjointe des états cachés $X_{0:k}$ et des observations $Y_{0:k}$.

$\square$

(ii) **En intégrant par rapport à** $x_{0:k-1} = (x_0, \cdots, x_{k-1})$ **et par rapport à** $x_{k+1:n} = (x_{k+1}, \cdots, x_n)$**, montrer que**

$$\mathbb{P}[X_k \in dx_k, Y_{0:n} \in dy_{0:n}\,] = \mathbb{P}[X_k \in dx_k, Y_{0:k} \in dy_{0:k}\,]\; v_k(x_k, y_{k+1:n})\; dy_{k+1:n} \ ,$$

où

$$v_k(x_k, y_{k+1:n}) = \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p) \ ,$$

par définition.

Solution

Il s'agit d'intégrer par rapport à $x_{0:k-1}$ et par rapport à $x_{k+1:n}$, les deux membres de l'identité

$$\begin{aligned} &\mathbb{P}[X_{0:n} \in dx_{0:n}, Y_{0:n} \in dy_{0:n}\,] \\ &\qquad = \mathbb{P}[X_{0:k} \in dx_{0:k}, Y_{0:k} \in dy_{0:k}\,] \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p)\; dy_{k+1:n} \ . \end{aligned}$$

En intégrant par rapport à $x_{0:k-1}$ et par rapport à $x_{k+1:n}$ la distribution de probabilité conjointe des états cachés $X_{0:n}$ et des observations $Y_{0:n}$, c'est–à–dire la quantité figurant au membre de gauche, on obtient par marginalisation la distribution de probabilité conjointe de l'état caché X_k et des observations $Y_{0:n}$. Par ailleurs, on remarque que la quantité

$$\mathbb{P}[X_{0:k} \in dx_{0:k}, Y_{0:k} \in dy_{0:k}\,] \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p)\; dy_{k+1:n} \ ,$$

figurant au membre de droite, s'exprime comme le produit d'une quantité

$$\mathbb{P}[X_{0:k} \in dx_{0:k}, Y_{0:k} \in dy_{0:k}\,] \ ,$$

qui ne dépend pas des variables $x_{k+1:n}$, et d'une quantité

$$\prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p) \ \prod_{p=k+1}^{n} g_p(x_p, y_p)\, dy_{k+1:n} \ ,$$

qui ne dépend pas des variables $x_{0:k-1}$. Le résultat de l'intégration par rapport à $x_{0:k-1}$ et par rapport à $x_{k+1:n}$ s'exprime donc comme le produit du résultat de l'intégration par rapport à $x_{0:k-1}$ du premier facteur et du résultat de l'intégration par rapport à $x_{k+1:n}$

du deuxième facteur. En intégrant par rapport à $x_{0:k-1}$ la distribution de probabilité conjointe des états cachés $X_{0:k}$ et des observations $Y_{0:k}$, c'est–à–dire le premier facteur, on obtient par marginalisation la distribution de probabilité conjointe de l'état caché X_k et des observations $Y_{0:k}$. En intégrant par rapport à $x_{k+1:n}$ le deuxième facteur, on obtient $v_k(x_k, y_{k+1:n})\, dy_{k+1:n}$ par définition.

$\square$

(iii) **Montrer que**

$$v_{k-1}(x_{k-1}, y_{k:n}) = \int_{\mathbb{R}^m} Q_k(x_{k-1}, dx_k)\ g_k(x_k, y_k)\ v_k(x_k, y_{k+1:n})\ .$$

Cette identité est valable pour toute suite $y_{k:n} = (y_k, \cdots, y_n)$ et en remplaçant les variables muettes $y_{k:n} = (y_k, \cdots, y_n)$ par les observations $Y_{k:n} = (Y_k, \cdots, Y_n)$, montrer la relation de récurrence rétrograde suivante

$$v_{k-1}(x) = \int_{\mathbb{R}^m} Q_k(x_{k-1}, dx_k)\ g_k(x_k)\ v_k(x_k)\ ,$$

où $v_{k-1}(x) = v_{k-1}(x, Y_{k:n})$, $v_k(x) = v_k(x, Y_{k+1:n})$ et $g_k(x) = g_k(x, Y_k)$ par définition.

Solution

En intégrant d'abord par rapport à $x_{k+1:n}$ puis par rapport à x_k, on obtient

$$\begin{aligned}
v_{k-1}(x_{k-1}, y_{k:n}) &= \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \prod_{p=k}^{n} Q_p(x_{p-1}, dx_p)\ \prod_{p=k}^{n} g_p(x_p, y_p) \\
&= \int_{\mathbb{R}^m} Q_k(x_{k-1}, dx_k)\ g_k(x_k, y_k) \\
&\qquad \int_{\mathbb{R}^m} \cdots \int_{\mathbb{R}^m} \prod_{p=k+1}^{n} Q_p(x_{p-1}, dx_p)\ \prod_{p=k+1}^{n} g_p(x_p, y_p) \\
&= \int_{\mathbb{R}^m} Q_k(x_{k-1}, dx_k)\ g_k(x_k, y_k)\ v_k(x_k, y_{k+1:n})\ ,
\end{aligned}$$

en remarquant que le résultat de l'intégrale intérieure est juste $v_k(x_k, y_{k+1:n})$.

En remplaçant les variables muettes $y_{k:n}$ par les observations $Y_{k:n}$, et en posant

$v_{k-1}(x) = v_{k-1}(x, Y_{k:n})$, $v_k(x) = v_k(x, Y_{k+1:n})$ et $g_k(x) = g_k(x, Y_k)$, on obtient directement

$$v_{k-1}(x) = \int_{\mathbb{R}^m} Q_k(x_{k-1}, dx_k)\; g_k(x_k)\; v_k(x_k)\ ,$$

ce qui s'interprète comme une relation de récurrence rétrograde, qui permet d'exprimer la fonction v_{k-1} à partir de la fonction v_k.

---------- □

(iv) **À partir du résultat obtenu à la question (ii), montrer que**

$$\mathbb{P}[X_k \in dx_k \mid Y_{0:n} = y_{0:n}\,] = \frac{\mathbb{P}[X_k \in dx_k \mid Y_{0:k} = y_{0:k}\,]\; v_k(x_k, y_{k+1:n})}{\displaystyle\int_{\mathbb{R}^m} \mathbb{P}[X_k \in dz \mid Y_{0:k} = y_{0:k}\,]\; v_k(z, y_{k+1:n})}\ .$$

Cette identité est valable pour toute suite $y_{0:n} = (y_0, \cdots, y_n)$ **et en remplaçant les variables muettes** $y_{0:n} = (y_0, \cdots, y_n)$ **par les observations** $Y_{0:n} = (Y_0, \cdots, Y_n)$**, montrer que le lisseur bayésien** $\mu_k^n(dx)$ **s'exprime comme**

$$\mu_k^n(dx) \ = \ \frac{v_k(x)\; \mu_k(dx)}{\displaystyle\int_{\mathbb{R}^m} v_k(z)\; \mu_k(dz)}\ ,$$

en fonction du filtre bayésien $\mu_k(dx)$ **et de la fonction** $v_k(x)$ **introduite à la question (iii) et définie à une constante multiplicative près par** $v_k(x) = v_k(x, Y_{k+1:n})$**.**

---------- Solution ----------

D'après la formule de Bayes, l'identité obtenue à la question (ii) s'écrit

$$\begin{aligned}
&\mathbb{P}[X_k \in dx_k \mid Y_{0:n} = y_{0:n}\,]\; \mathbb{P}[Y_{0:n} \in dy_{0:n}\,] \\
&\qquad = \mathbb{P}[X_k \in dx_k \mid Y_{0:k} = y_{0:k}\,]\; v_k(x_k, y_{k+1:n})\; \mathbb{P}[Y_{0:k} \in dy_{0:k}\,]\; dy_{k+1:n}\ .
\end{aligned}$$

En intégrant par rapport à x_k, on obtient

$$\begin{aligned}
&\mathbb{P}[Y_{0:n} \in dy_{0:n}\,] \\
&\qquad = [\int_{\mathbb{R}^m} \mathbb{P}[X_k \in dx_k \mid Y_{0:k} = y_{0:k}\,]\; v_k(x_k, y_{k+1:n})\,]\; \mathbb{P}[Y_{0:k} \in dy_{0:k}\,]\; dy_{k+1:n}\ ,
\end{aligned}$$

et en divisant membre–à–membre, on obtient

$$\mathbb{P}[X_k \in dx_k \mid Y_{0:n} = y_{0:n}\,] = \frac{\mathbb{P}[X_k \in dx_k \mid Y_{0:k} = y_{0:k}\,]\; v_k(x_k, y_{k+1:n})}{\displaystyle\int_{\mathbb{R}^m} \mathbb{P}[X_k \in dz \mid Y_{0:k} = y_{0:k}\,]\; v_k(z, y_{k+1:n})}\ .$$

En remplaçant les variables muettes $y_{0:n}$ par les observations $Y_{0:n}$, et en rappelant que $v_k(x) = v_k(x, Y_{k+1:n})$, on obtient directement

$$\mathbb{P}[\, X_k \in dx \mid Y_{0:n}\,] = \frac{\mathbb{P}[\, X_k \in dx \mid Y_{0:k}\,]\; v_k(x)}{\displaystyle\int_{\mathbb{R}^m} \mathbb{P}[\, X_k \in dz \mid Y_{0:k}\,]\; v_k(z)} \; .$$

□

Cas particulier gaussien

À partir de maintenant, on exploite le caractère gaussien du problème, qui permet d'obtenir des expressions explicites.

(v) **Montrer que**

$$\mathbb{P}[\, X_k \in dx' \mid X_{k-1} = x] = \exp\{-\tfrac{1}{2}\,(x' - F_k\, x)^*\,(Q_k^W)^{-1}\,(x' - F_k\, x)\}\; \frac{dx'}{\sqrt{\det(2\pi Q_k^W)}} \; ,$$

avec une densité de transition gaussienne, de moyenne $F_k\, x$ et de matrice de covariance Q_k^W (supposée inversible par hypothèse).

Solution

D'après le modèle a priori

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

où les vecteurs aléatoires X_{k-1} et W_k sont indépendants. L'information $X_{k-1} = x$ n'apprend donc rien de plus sur le vecteur aléatoire W_k, de sorte que la distribution de probabilité conditionnelle du vecteur aléatoire X_k sachant $X_{k-1} = x$ est gaussienne, de moyenne $F_k\, x$ et de matrice de covariance Q_k^W supposée inversible.

□

(vi) **Montrer que**

$$\mathbb{P}[\, Y_k \in dy \mid X_k = x] = \exp\{-\tfrac{1}{2}\,(y - H_k\, x)^*\,(Q_k^V)^{-1}\,(y - H_k\, x)\}\; \frac{dy}{\sqrt{\det(2\pi Q_k^V)}} \; ,$$

avec une densité d'émission gaussienne, de moyenne $H_k\, x$ et de matrice de covariance Q_k^V (supposée inversible par hypothèse).

Solution

D'après l'équation d'observation (modèle de capteur)

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

où les vecteurs aléatoires X_k et V_k sont indépendants. L'information $X_k = x$ n'apprend donc rien de plus sur le vecteur aléatoire V_k, de sorte que la distribution de probabilité conditionnelle du vecteur aléatoire Y_k sachant $X_k = x$ est gaussienne, de moyenne $H_k\, x$ et de matrice de covariance Q^V_k supposée inversible.

□

La fin du problème consiste à montrer (par récurrence) que

$$v_k(x) = \exp\{-\tfrac{1}{2}\, x^*\, M_k\, x + m_k^*\, x\} \ ,$$

à une constante multiplicative près pour tout $k = (n-1), \cdots, 1, 0$, et à établir une relation de récurrence rétrograde pour le vecteur m_k et pour la matrice symétrique semi–définie positive M_k.

(vii) **On suppose l'hypothèse de récurrence vraie au rang k. Montrer que**

$$g_k(x')\; v_k(x') = \exp\{-\tfrac{1}{2}\, x'^*\, M_{k-1/2}\, x' + m_{k-1/2}^*\, x'\} \ ,$$

à une constante multiplicative près, avec

$$M_{k-1/2} = H_k^*\, (Q_k^V)^{-1}\, H_k + M_k \ ,$$

et

$$m_{k-1/2} = H_k^*\, (Q_k^V)^{-1}\, Y_k + m_k \ .$$

Solution

En développant, on obtient

$$\begin{aligned}
&\tfrac{1}{2}\, (Y_k - H_k\, x')^*\, (Q_k^V)^{-1}\, (Y_k - H_k\, x') + \tfrac{1}{2}\, x'^*\, M_k\, x' - m_k^*\, x' \\
&\qquad = \tfrac{1}{2}\, x'^*\, (H_k^*\, (Q_k^V)^{-1}\, H_k + M_k)\, x' - (Y_k^*\, (Q_k^V)^{-1}\, H_k + m_k^*)\, x' + \tfrac{1}{2}\, Y_k^*\, (Q_k^V)^{-1}\, Y_k \\
&\qquad = \tfrac{1}{2}\, x'^*\, M_{k-1/2}\, x' - m_{k-1/2}^*\, x' + \text{cste}
\end{aligned}$$

à une constante additive près, avec

$$M_{k-1/2} = H_k^*\, (Q_k^V)^{-1}\, H_k + M_k \qquad \text{et} \qquad m_{k-1/2} = H_k^*\, (Q_k^V)^{-1}\, Y_k + m_k \ ,$$

de sorte que

$$\begin{aligned}
g_k(x')\; v_k(x') &= \exp\{-\tfrac{1}{2}\, (Y_k - H_k\, x')^*\, (Q_k^V)^{-1}\, (Y_k - H_k\, x')\}\; \exp\{-\tfrac{1}{2}\, x'^*\, M_k\, x' + m_k^*\, x'\} \\
&= \exp\{-\tfrac{1}{2}\, x'^*\, M_{k-1/2}\, x' + m_{k-1/2}^*\, x'\} \ .
\end{aligned}$$

à une constante multiplicative près.

□

(viii) **En complétant le carré et en intégrant par rapport à la variable x', montrer que**

$$v_{k-1}(x) = \int_{\mathbb{R}^m} Q_k(x, dx')\; g_k(x')\; v_k(x') = \exp\{-\tfrac{1}{2}\, x^*\, M_{k-1}\, x + m_{k-1}^*\, x\}\ ,$$

à une constante multiplicative près, c'est–à–dire que l'hypothèse de récurrence est vraie au rang $(k-1)$, avec

$$\begin{aligned} M_{k-1} &= F_k^*\, [\, (Q_k^W)^{-1} - (Q_k^W)^{-1}\, ((Q_k^W)^{-1} + M_{k-1/2})^{-1}\, (Q_k^W)^{-1}\,]\, F_k \\ &= F_k^*\, (Q_k^W)^{-1}\, [\, Q_k^W - ((Q_k^W)^{-1} + M_{k-1/2})^{-1}\,]\, (Q_k^W)^{-1}\, F_k\ , \end{aligned}$$

et

$$m_{k-1} = F_k^*\, (Q_k^W)^{-1}\, [\, (Q_k^W)^{-1} + M_{k-1/2}\,]^{-1}\, m_{k-1/2}\ .$$

Solution

En développant et en complètant le carré, on obtient

$$\begin{aligned} &\tfrac{1}{2}\, (x' - F_k\, x)^*\, (Q_k^W)^{-1}\, (x' - F_k\, x) + \tfrac{1}{2}\, x'^*\, M_{k-1/2}\, x' - m_{k-1/2}^*\, x' \\ &\quad = \tfrac{1}{2}\, x'^*\, ((Q_k^W)^{-1} + M_{k-1/2})\, x' - ((F_k\, x)^*\, (Q_k^W)^{-1} + m_{k-1/2}^*)\, x' \\ &\qquad\quad + \tfrac{1}{2}\, (F_k\, x)^*\, (Q_k^W)^{-1}\, F_k\, x \\ &\quad = \tfrac{1}{2}\, x'^*\, \Sigma^{-1}\, x' - b^*\, x' + \tfrac{1}{2}\, c \\ &\quad = \tfrac{1}{2}\, (x' - \Sigma\, b)^*\, \Sigma^{-1}\, (x' - \Sigma\, b) - \tfrac{1}{2}\, b^*\, \Sigma\, b + \tfrac{1}{2}\, c\ , \end{aligned}$$

avec

$$\Sigma^{-1} = (Q_k^W)^{-1} + M_{k-1/2}\ ,$$

$$b^* = (F_k\, x)^*\, (Q_k^W)^{-1} + m_{k-1/2}^* \qquad \text{soit} \qquad b = (Q_k^W)^{-1}\, F_k\, x + m_{k-1/2}\ ,$$

$$c = (F_k\, x)^*\, (Q_k^W)^{-1}\, F_k\, x\ .$$

On en déduit que

$$\begin{aligned} v_{k-1}(x) &= \int_{\mathbb{R}^m} Q_k(x, dx')\; g_k(x')\; v_k(x') \\ &= \int_{\mathbb{R}^m} \exp\{-\tfrac{1}{2}\, (x' - F_k\, x)^*\, (Q_k^W)^{-1}\, (x' - F_k\, x)\} \\ &\qquad\qquad \exp\{-\tfrac{1}{2}\, x'^*\, M_{k-1/2}\, x' + m_{k-1/2}^*\, x'\}\; dx' \\ &= \int_{\mathbb{R}^m} \exp\{-\tfrac{1}{2}\, (x' - \Sigma\, b)^*\, \Sigma^{-1}\, (x' - \Sigma\, b) + \tfrac{1}{2}\, b^*\, \Sigma\, b - \tfrac{1}{2}\, c\}\; dx' \\ &= \exp\{\tfrac{1}{2}\, b^*\, \Sigma\, b - \tfrac{1}{2}\, c\}\ , \end{aligned}$$

à une constante multiplicative près, et on vérifie que

$$\begin{aligned}
-\tfrac{1}{2}\, c + \tfrac{1}{2}\, b^* \, \Sigma \, b = -\tfrac{1}{2}\, (F_k \, x)^* \, (Q^W_k)^{-1} \, F_k \, x \\
+ \tfrac{1}{2}\, ((F_k \, x)^* \, (Q^W_k)^{-1} + m^*_{k-1/2}) \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \\
((Q^W_k)^{-1} \, F_k \, x + m_{k-1/2}) \\
= -\tfrac{1}{2}\, (F_k \, x)^* \, [\, (Q^W_k)^{-1} - (Q^W_k)^{-1} \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \, (Q^W_k)^{-1} \,] \, F_k \, x \\
+ [\, m^*_{k-1/2} \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \, (Q^W_k)^{-1} \,] \, F_k \, x \\
+ \tfrac{1}{2}\, m^*_{k-1/2} \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \, m_{k-1/2} \\
= -\tfrac{1}{2}\, x^* \, M_{k-1} \, x + m^*_{k-1} \, x + \text{cste} \; ,
\end{aligned}$$

à une constante additive près, avec

$$\begin{aligned}
M_{k-1} &= F^*_k \, [\, (Q^W_k)^{-1} - (Q^W_k)^{-1} \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \, (Q^W_k)^{-1} \,] \, F_k \; , \\
m_{k-1} &= F^*_k \, (Q^W_k)^{-1} \, ((Q^W_k)^{-1} + M_{k-1/2})^{-1} \, m_{k-1/2} \; .
\end{aligned}$$

□

(ix) **En déduire que le lisseur bayésien est (associé à) une distribution de probabilité gaussienne, de moyenne et de matrice de covariance définies par**

$$\widehat{X}^n_k = [\, P^{-1}_k + M_k \,]^{-1} \, [\, P^{-1}_k \, \widehat{X}_k + m_k \,] \qquad \textbf{et} \qquad P^n_k = [\, P^{-1}_k + M_k \,]^{-1} \; ,$$

respectivement.

Solution

D'après les équations du filtre de Kalman

$$P_k = P^-_k - P^-_k \, H^*_k \, (H_k \, P^-_k \, H^*_k + Q^V_k)^{-1} \, H_k \, P^-_k \; ,$$

et

$$P^-_k = F_k \, P_{k-1} \, F^*_k + Q^W_k \; ,$$

où la matrice de covariance Q^W_k est supposée inversible par hypothèse, de sorte que a fortiori la matrice de covariance P^-_k est inversible. A fortiori la matrice

$$H^*_k \, (Q^V_k)^{-1} \, H_k + (P^-_k)^{-1} \; ,$$

est inversible, et a pour inverse

$$(H^*_k \, (Q^V_k)^{-1} \, H_k + (P^-_k)^{-1})^{-1} = P^-_k - P^-_k \, H^*_k \, (H_k \, P^-_k \, H^*_k + Q^V_k)^{-1} \, H_k \, P^-_k = P_k \; ,$$

d'après le lemme d'inversion matricielle. On en déduit que la matrice de covariance P_k est inversible.

En développant et en complétant le carré, on obtient

$$\begin{aligned}
&\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) + \tfrac{1}{2}\, x^*\, M_k\, x - m_k^*\, x \\
&= \tfrac{1}{2}\, x^*\,(P_k^{-1} + M_k)\, x - (\widehat{X}_k^*\, P_k^{-1} + m_k^*)\, x + \tfrac{1}{2}\, \widehat{X}_k^*\, P_k^{-1}\, \widehat{X}_k \\
&= \tfrac{1}{2}\, x^*\, \Sigma^{-1}\, x - b^*\, x + \tfrac{1}{2}\, c \\
&= \tfrac{1}{2}\,(x - \Sigma\, b)^*\, \Sigma^{-1}\,(x - \Sigma\, b) + \text{cste}\ ,
\end{aligned}$$

à une constante additive près, avec

$$\Sigma^{-1} = P_k^{-1} + M_k\ ,$$

$$b^* = \widehat{X}_k^*\, P_k^{-1} + m_k^* \qquad \text{soit} \qquad b = P_k^{-1}\, \widehat{X}_k + m_k\ .$$

On en déduit que

$$\begin{aligned}
\mathbb{P}[X_k \in dx \mid Y_{0:n}] &= \mathbb{P}[X_k \in dx \mid Y_{0:k}]\ v_k(x) \\
&= \exp\{-\tfrac{1}{2}\,(x - \widehat{X}_k)^*\, P_k^{-1}\,(x - \widehat{X}_k) - \tfrac{1}{2}\, x^*\, M_k\, x + m_k^*\, x\} \\
&= \exp\{-\tfrac{1}{2}\,(x - \Sigma\, b)^*\, \Sigma^{-1}\,(x - \Sigma\, b)\}\ ,
\end{aligned}$$

à une constante multiplicative près, c'est–à–dire que le lisseur bayésien est (associé à) une distribution de probabilité gaussienne, de moyenne

$$\widehat{X}_{k|n} = \Sigma\, b = (P_k^{-1} + M_k)^{-1}\ (P_k^{-1}\, \widehat{X}_k + m_k)\ ,$$

et de matrice de covariance

$$P_{k|n} = \Sigma = (P_k^{-1} + M_k)^{-1}\ .$$

$\square$