Université de Rennes 1
Master Recherche SISEA

Examen du cours UE S3-2
“Filtre de Kalman et modèles de Markov cachés”
Jeudi 23 janvier 2014, 10:15 à 12:15
— Corrigé —

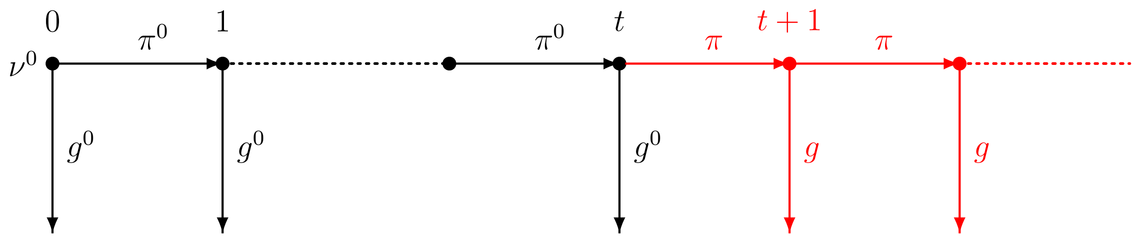
Exercice

Le but de cet exercice est de montrer comment une utilisation ingénieuse des variables forward et backward permet de détecter un éventuel changement de modèle, et le cas échéant d'estimer l'instant de changement.

On considère un modèle de Markov caché, avec des états dans un ensemble fini E et avec des observations *numériques*, à valeurs dans $\mathbb{R}^d$.

Au vu des observations $(Y_0, \cdots, Y_n)$, on veut détecter s'il existe ou non un instant t compris entre l'instant initial 0 et l'instant final n, tel que

- à l'instant 0, la loi initiale est $\nu^0 = (\nu^0_i)$,
- entre l'instant initial 0 et l'instant t, le modèle est nominal, c'est–à–dire que la matrice de transition est $\pi^0 = (\pi^0_{ij})$ et les densités d'émission sont $g^0(y) = (g^0_i(y))$,
- entre l'instant $(t + 1)$ et l'instant final n, le modèle est modifié, c'est–à–dire que la matrice de transition est $\pi = (\pi_{ij})$ et les densités d'émission sont $g(y) = (g_i(y))$.



On rappelle que la vraisemblance L^0_n du modèle nominal $\boldsymbol{M}^0 = (\nu^0, \pi^0, g^0)$ au vu des observations $(Y_0, \cdots, Y_n)$ peut s'exprimer comme

$$L^0_n = \sum_{i \in E} p^{0,i}_n = \sum_{i \in E} p^{0,i}_k \; v^{0,i}_k \;, \qquad \text{pour tout } k = 0, \cdots, n$$

au moyen des variables forward et backward associées au modèle nominal $\boldsymbol{M}^0$. De même, la vraisemblance L_n du modèle alternatif $\boldsymbol{M} = (\nu^0, \pi, g)$ (où le changement est déjà intervenu avant même l'instant initial 0) au vu des observations $(Y_0, \cdots, Y_n)$ peut s'exprimer comme

$$L_n = \sum_{i \in E} p_n^i = \sum_{i \in E} p_k^i \, v_k^i \; , \qquad \text{pour tout } k = 0, \cdots, n$$

au moyen des variables forward et backward associées au modèle alternatif $\boldsymbol{M}$.

(i) En adoptant la même démarche bayésienne et en suivant les mêmes étapes (distribution des états cachés + densité conditionnelle des observations sachant les états cachés $\rightarrow$ distribution conjointe des états cachés et des observations $\rightarrow$ densité des observations) que dans le cours, montrer que la vraisemblance du modèle avec changement à l'instant t au vu des observations $(Y_0, \cdots, Y_n)$ peut s'exprimer comme

$$L_n(t) = \sum_{i_0, \cdots, i_n \in E} \nu_{i_0}^0 \; \pi_{i_0,i_1}^0 \cdots \pi_{i_{t-1},i_t}^0 \; g_{i_0}^0(Y_0) \cdots g_{i_t}^0(Y_t)$$

$$\pi_{i_t,i_{t+1}} \; \cdots \pi_{i_{n-1},i_n} \; g_{i_{t+1}}(Y_{t+1}) \cdots g_{i_n}(Y_n) \; .$$

Solution

Pour le modèle avec changement à l'instant t, la distribution de probabilité des états cachés $X_{0:n}$ s'exprime à l'aide de la loi initiale ν^0 et des matrices de transition π^0 et π comme

$$\mathbb{P}[X_0 = i_0, \cdots, X_n = i_n] = \nu_{i_0}^0 \; \pi_{i_0,i_1}^0 \cdots \pi_{i_{t-1},i_t}^0 \; \pi_{i_t,i_{t+1}} \; \cdots \pi_{i_{n-1},i_n} \; ,$$

pour tout $i_0, \cdots, i_n \in E$, et la distribution de probabilité conditionnelle des observations $Y_{0:n}$ sachant les états cachés $X_{0:n}$ s'exprime à l'aide des densités d'émission $g^0(y)$ et $g(y)$ comme

$$\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n \mid X_0 = i_0, \cdots, X_n = i_n]$$

$$= g_{i_0}^0(y_0) \cdots g_{i_t}^0(y_t) \; g_{i_{t+1}}(y_{t+1}) \cdots g_{i_n}(y_n) \; dy_0 \cdots dy_n \; ,$$

pour tout $y_0, \cdots, y_n \in \mathbb{R}^d$ et pour tout $i_0, \cdots, i_n \in E$. On en déduit l'expressionn la distribution de probabilité conjointe des états cachés $X_{0:n}$ et des observations $Y_{0:n}$

$$\mathbb{P}[X_0 = i_0, \cdots, X_n = i_n, Y_0 \in dy_0, \cdots, Y_n \in dy_n]$$

$$= \nu_{i_0}^0 \; \pi_{i_0,i_1}^0 \cdots \pi_{i_{t-1},i_t}^0 \; \pi_{i_t,i_{t+1}} \; \cdots \pi_{i_{n-1},i_n}$$

$$g_{i_0}^0(y_0) \cdots g_{i_t}^0(y_t) \; g_{i_{t+1}}(y_{t+1}) \cdots g_{i_n}(y_n) \; dy_0 \cdots dy_n \; ,$$

pour tout $y_0, \cdots, y_n \in \mathbb{R}^d$ et pour tout $i_0, \cdots, i_n \in E$, et en sommant sur les états $i_0, \cdots, i_n \in E$ on obtient l'expression la distribution de probabilité des observations $Y_{0:n}$

$$\begin{aligned}\mathbb{P}[Y_0 \in dy_0, \cdots, Y_n \in dy_n] \;=\; & \sum_{i_0, \cdots, i_n \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; \pi_{i_t,i_{t+1}}\; \cdots \pi_{i_{n-1},i_n} \\ & \qquad g^0_{i_0}(y_0) \cdots g^0_{i_t}(y_t)\; g_{i_{t+1}}(y_{t+1}) \cdots g_{i_n}(y_n)\; dy_0 \cdots dy_n \;,\end{aligned}$$

pour tout $y_0, \cdots, y_n \in \mathbb{R}^d$. Clairement, cette distribution de probabilité admet une densité, ce qui définit la vraisemblance du modèle avec changement à l'instant t au vu des observations $(Y_0, \cdots, Y_n)$, comme

$$\begin{aligned}L_n(t) \;=\; & \sum_{i_0, \cdots, i_n \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; \pi_{i_t,i_{t+1}}\; \cdots \pi_{i_{n-1},i_n} \\ & \qquad g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t)\; g_{i_{t+1}}(Y_{t+1}) \cdots g_{i_n}(Y_n) \\ =\; & \sum_{i_0, \cdots, i_n \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t) \\ & \qquad \pi_{i_t,i_{t+1}}\; \cdots \pi_{i_{n-1},i_n}\; g_{i_{t+1}}(Y_{t+1}) \cdots g_{i_n}(Y_n) \;.\end{aligned}$$

□

(ii) En déduire que

$$L_n(t) = \sum_{i \in E} p^{0,i}_t\; v^i_t \;,$$

au moyen de la variable forward associée au modèle nominal M^0 et de la variable backward associée au modèle alternatif M.

Solution

Dans l'expression obtenue pour la vraisemblance du modèle avec changement à l'instant t, on somme d'abord par rapport aux états $i_{t+1}, \cdots, i_n \in E$, et on obtient

$$\begin{aligned}L_n(t) \;=\; & \sum_{i_0, \cdots, i_n \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t) \\ & \qquad \pi_{i_t,i_{t+1}}\; \cdots \pi_{i_{n-1},i_n}\; g_{i_{t+1}}(Y_{t+1}) \cdots g_{i_n}(Y_n) \\ =\; & \sum_{i_0, \cdots, i_t \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t) \\ & \qquad \sum_{i_{t+1}, \cdots, i_n \in E} \pi_{i_t,i_{t+1}}\; \cdots \pi_{i_{n-1},i_n}\; g_{i_{t+1}}(Y_{t+1}) \cdots g_{i_n}(Y_n) \\ =\; & \sum_{i_0, \cdots, i_t \in E} \nu^0_{i_0}\; \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\; g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t)\; v^{i_t}_t \;,\end{aligned}$$

en fonction de la variable backward v_t à l'instant t pour le modèle alternatif $\boldsymbol{M}$. Dans l'expression obtenue, on somme d'abord par rapport aux états $i_0, \cdots, i_{t-1} \in E$, et on obtient

$$
\begin{aligned}
L_n(t) &= \sum_{i_0,\cdots,i_t \in E} \nu^0_{i_0}\, \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\, g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t)\, v^{i_t}_t \\
&= \sum_{i_t \in E} [\sum_{i_0,\cdots,i_{t-1} \in E} \nu^0_{i_0}\, \pi^0_{i_0,i_1} \cdots \pi^0_{i_{t-1},i_t}\, g^0_{i_0}(Y_0) \cdots g^0_{i_t}(Y_t)]\, v^{i_t}_t \\
&= \sum_{i_t \in E} p^{0,i_t}_t\, v^{i_t}_t \ ,
\end{aligned}
$$

en fonction de la variable forward p^0_t à l'instant t pour le modèle nominal $\boldsymbol{M}^0$.

$\square$

Il suffit donc de résoudre deux équations, l'équation forward associée au modèle nominal $\boldsymbol{M}^0$ et l'équation backward associée au modèle alternatif $\boldsymbol{M}$, pour pouvoir calculer la fonction de vraisemblance $t \mapsto L_n(t)$ au vu des observations $(Y_0, \cdots, Y_n)$ pour l'estimation de l'instant de changement.

Problème

Le but de ce problème est d'établir d'une manière géométrique simple due à Ansley et Kohn (1982), les équations du lisseur de Kalman.

On considère une suite d'états cachés $\{X_k\}$ à valeurs dans $\mathbb{R}^m$, vérifiant

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

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

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

et on suppose que

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

On dispose de toutes les observations

$$Y_{0:n} = (Y_0, Y_1, \cdots, Y_n) \ ,$$

et l'objectif est d'estimer de façon optimale le vecteur aléatoire X_k à partir de $Y_{0:n}$, pour un instant k intermédiaire entre l'instant initial 0 et l'instant final n. Si on adopte le critère du minimum de variance, il s'agit donc de calculer la distribution de probabilité conditionnelle du vecteur aléatoire X_k sachant $Y_{0:n}$. Comme le cadre est gaussien, cette distribution de probabilité conditionnelle est gaussienne, et il suffit de calculer la moyenne

$$\widehat{X}_k^n = \mathbb{E}[X_k \mid Y_{0:n}] \ .$$

On suppose que les estimateurs (prédicteur et filtre, respectivement)

$$\widehat{X}_k^- = \mathbb{E}[X_k \mid Y_{0:k-1}] \qquad \text{et} \qquad \widehat{X}_k = \mathbb{E}[X_k \mid Y_{0:k}] \ ,$$

ont déjà été calculés dans une première phase et sont disponibles, pour tout instant k intermédiaire entre l'instant initial 0 et l'instant final n.

(i) Montrer que $\widehat{X}_n^n = \widehat{X}_n$ pour $k = n$. En d'autres termes, le lisseur et le filtre coïncident à l'instant final.

Solution

Clairement

$$\widehat{X}_n^n = \mathbb{E}[X_n \mid Y_{0:n}] = \widehat{X}_n \ .$$

□

On commence par établir, dans les deux questions ci–dessous, la propriété qualitative suivante (dont la démonstration ne nécessite aucun calcul, mais seulement des arguments qualitatifs) : le vecteur aléatoire $(Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n})$, où

$$Z_{k+1:n} = (W_{k+1}, \cdots, W_n, V_k, V_{k+1}, \cdots, V_n) \ ,$$

par définition, contient davantage d'information que le vecteur aléatoire $Y_{0:n}$.

(ii) Montrer que le vecteur Y_k peut s'exprimer comme transformation affine du vecteur (X_k, V_k), et que le vecteur $X_k = (X_k - \widehat{X}_k^-) + \widehat{X}_k^-$ peut s'exprimer comme transformation affine du vecteur $(Y_{0:k-1}, X_k - \widehat{X}_k^-)$.

En déduire que le vecteur Y_k peut s'exprimer comme transformation affine du vecteur $(Y_{0:k-1}, X_k - \widehat{X}_k^-, V_k)$.

Solution

Par définition

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

et

$$X_k = \widehat{X}_k^- + (X_k - \widehat{X}_k^-) \ ,$$

où l'état prédit $\widehat{X}_k^- = \mathbb{E}[X_k \mid Y_{0:k-1}]$ peut s'exprimer comme transformation affine du vecteur $Y_{0:k-1}$, d'après le conditionnement dans les vecteurs aléatoires gaussiens.

□

(iii) De même, montrer que le vecteur Y_{k+p} peut s'exprimer comme transformation affine du vecteur (X_{k+p}, V_{k+p}), et que le vecteur X_{k+p} peut s'exprimer comme transformation affine du vecteur $(X_k, W_{k+1}, \cdots, W_{k+p})$.

En déduire que le vecteur Y_{k+p} peut s'exprimer comme transformation affine du vecteur $(Y_{0:k-1}, X_k - \widehat{X}_k^-, W_{k+1}, \cdots, W_{k+p}, V_{k+p})$.

En déduire que le vecteur $Y_{0:n} = (Y_{0:k-1}, Y_k, \cdots, Y_n)$ peut s'exprimer comme transformation affine du vecteur $(Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n})$, où on rappelle que

$$Z_{k+1:n} = (W_{k+1}, \cdots, W_n, V_k, V_{k+1}, \cdots, V_n) \ ,$$

par définition.

Solution

Par définition
$$Y_{k+p} = H_{k+p}\, X_{k+p} + h_{k+p} + V_{k+p} \ ,$$
et on commence par montrer, par récurrence sur l'indice p, que le vecteur X_{k+p} peut s'exprimer comme transformation affine du vecteur $(X_k, W_{k+1}, \cdots, W_{k+p})$. Clairement, l'hypothèse de récurrence est vérifiée à l'étape $p = 0$. Par définition
$$X_{k+p} = F_{k+p}\, X_{k+p-1} + h_{k+p} + W_{k+p} \ ,$$
et si l'hypothèse de récurrence est vérifiée à l'étape $(p-1)$, alors le vecteur X_{k+p-1} peut s'exprimer comme transformation affine du vecteur $(X_k, W_{k+1}, \cdots, W_{k+p-1})$. On en déduit que l'hypothèse de récurrence est vérifiée à l'étape p. Finalement
$$X_k = \widehat{X}_k^- + (X_k - \widehat{X}_k^-) \ ,$$
où l'état prédit $\widehat{X}_k^- = \mathbb{E}[X_k \mid Y_{0:k-1}]$ peut s'exprimer comme transformation affine du vecteur $Y_{0:k-1}$, d'après le conditionnement dans les vecteurs aléatoires gaussiens.

□

(iv) Montrer que le vecteur aléatoire $(Y_{0:k-1}, X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ **et le vecteur aléatoire** $Z_{k+1:n} = (W_{k+1}, \cdots, W_n, V_k, V_{k+1}, \cdots, V_n)$ **sont indépendants.**

Montrer que le vecteur aléatoire $(X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ **et le vecteur aléatoire** $Y_{0:k-1}$ **sont indépendants.**

En déduire que le vecteur aléatoire U_{k-1}^n **défini par**
$$U_{k-1}^n = \mathbb{E}[X_{k-1} \mid Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n}] \ ,$$
vérifie
$$U_{k-1}^n = \widehat{X}_{k-1} + \mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid X_k - \widehat{X}_k^-] \ .$$

Solution

Le vecteur aléatoire $(Y_{0:k-1}, X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ dépend seulement du vecteur aléatoire $(Y_{0:k-1}, X_{k-1}, X_k)$, lequel est indépendant du vecteur aléatoire $Z_{k+1:n}$.

On a vu en cours que les vecteurs aléatoires $(X_k - \widehat{X}_k^-)$ et $Y_{0:k-1}$ sont indépendants. En procédant exactement de la même manière, on remarque que l'état estimé $\widehat{X}_{k-1} = \mathbb{E}[X_{k-1} \mid Y_{0:k-1}]$ peut s'exprimer comme transformation affine du vecteur $Y_{0:k-1}$, d'après le conditionnement dans les vecteurs aléatoires gaussiens. On en déduit que le vecteur aléatoire $(Y_{0:k-1}, X_{k-1} - \widehat{X}_{k-1})$ est gaussien comme transformation affine du vecteur aléatoire gaussien $(Y_{0:k-1}, X_{k-1})$, et compte tenu que
$$\mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid Y_{0:k-1}] = 0 \ ,$$

par définition, on en déduit que

$$\mathbb{E}[(X_{k-1} - \widehat{X}_{k-1})\, Y_{0:k-1}^*] = \mathbb{E}[\,\mathbb{E}[(X_{k-1} - \widehat{X}_{k-1}) \mid Y_{0:k-1}]\, Y_{0:k-1}^*\,] = 0\ ,$$

c'est–à–dire que les vecteurs aléatoires $(X_{k-1} - \widehat{X}_{k-1})$ et $Y_{0:k-1}$ sont indépendants.

On en déduit que

$$\begin{aligned} U_{k-1}^n &= \mathbb{E}[X_{k-1} \mid Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n}] \\ &= \widehat{X}_{k-1} + \mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n}] \\ &= \widehat{X}_{k-1} + \mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid Y_{0:k-1}, X_k - \widehat{X}_k^-] \\ &= \widehat{X}_{k-1} + \mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid X_k - \widehat{X}_k^-]\ , \end{aligned}$$

où on a utilisé dans l'avant–dernière égalité le fait que les vecteurs aléatoires $(Y_{0:k-1}, X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ et $Z_{k+1:n}$ sont indépendants, et où on a utilisé dans la dernière égalité le fait que les vecteurs aléatoires $(X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ et $Y_{0:k-1}$ sont indépendants.

□

Pour calculer la moyenne conditionnelle U_{k-1}^n du vecteur aléatoire X_{k-1} sachant $(Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n})$, il suffit donc de calculer la moyenne conditionnelle du vecteur aléatoire $(X_{k-1} - \widehat{X}_{k-1})$ sachant $(X_k - \widehat{X}_k^-)$.

(v) Montrer que le vecteur aléatoire $(X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ est gaussien, de moyenne nulle, et de matrice de covariance dont on donnera l'expression en fonction des matrices de covariance P_{k-1} et P_k^- (et de la matrice F_k).

Vérifier que la matrice de covariance $P_k^- = F_k\, P_{k-1}\, F_k^* + Q_k^W$ est inversible.

Solution

L'état estimé $\widehat{X}_{k-1} = \mathbb{E}[X_{k-1} \mid Y_{0:k-1}]$ et l'état prédit $\widehat{X}_k^- = \mathbb{E}[X_k \mid Y_{0:k-1}]$ peuvent s'exprimer comme transformation affine du vecteur $Y_{0:k-1}$, d'après le conditionnement dans les vecteurs aléatoires gaussiens. On en déduit que le vecteur aléatoire $(X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ est gaussien comme transformation affine du vecteur aléatoire gaussien $(Y_{0:k-1}, X_{k-1}, X_k)$. Clairement

$$\mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid Y_{0:k-1}] = 0 \quad \text{et} \quad \mathbb{E}[X_k - \widehat{X}_k^- \mid Y_{0:k-1}] = 0\ ,$$

et a fortiori

$$\mathbb{E}[X_{k-1} - \widehat{X}_{k-1}] = 0 \quad \text{et} \quad \mathbb{E}[X_k - \widehat{X}_k^-] = 0\ .$$

Par définition, les matrices de covariance sont données par

$$\mathbb{E}[\,(X_{k-1} - \widehat{X}_{k-1})\,(X_{k-1} - \widehat{X}_{k-1})^*\,] = P_{k-1} \ ,$$

et

$$\mathbb{E}[\,(X_k - \widehat{X}_k^-)\,(X_k - \widehat{X}_k^-)^*\,] = P_k^- \ ,$$

et il suffit de calculer la matrice de corrélation. Par différence

$$X_k - \widehat{X}_k^- = F_k\,X_{k-1} + f_k + W_k - (F_k\,\widehat{X}_{k-1} + f_k) = F_k\,(X_{k-1} - \widehat{X}_{k-1}) + W_k \ ,$$

de sorte que

$$\begin{aligned}
&\mathbb{E}[\,(X_{k-1} - \widehat{X}_{k-1})\,(X_k - \widehat{X}_k^-)^*\,] \\
&\qquad = \mathbb{E}[\,(X_{k-1} - \widehat{X}_{k-1})\,(F_k\,(X_{k-1} - \widehat{X}_{k-1}) + W_k)^*\,] \\
&\qquad = \mathbb{E}[\,(X_{k-1} - \widehat{X}_{k-1})\,(X_{k-1} - \widehat{X}_{k-1})^*\,]\ F_k^* + \mathbb{E}[\,(X_{k-1} - \widehat{X}_{k-1})\,W_k^*\,] \\
&\qquad = P_{k-1}\,F_k^* \ .
\end{aligned}$$

Dans cette dernière égalité, on a utilisé le fait que $(X_{k-1} - \widehat{X}_{k-1})$ et W_k sont indépendants, donc $\mathbb{E}[(X_{k-1} - \widehat{X}_{k-1})\,W_k^*] = 0$. On en déduit que le vecteur aléatoire $(X_{k-1} - \widehat{X}_{k-1}, X_k - \widehat{X}_k^-)$ est gaussien, de moyenne nulle et de matrice de covariance

$$\begin{pmatrix} P_{k-1} & P_{k-1}\,F_k^* \\ F_k\,P_{k-1} & P_k^- \end{pmatrix} \ .$$

Par hypothèse, la matrice Q_k^W est inversible, et a fortiori la matrice $P_k^- = F_k\,P_{k-1}\,F_k^* + Q_k^W$ est inversible.

$\square$

(vi) En déduire que

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

avec la matrice de gain définie par

$$L_k = P_{k-1}\,F_k^*\,[F_k\,P_{k-1}\,F_k^* + Q_k^W]^{-1} \ .$$

Solution

D'après le conditionnnement dans les vecteurs aléatoires gaussiens, on a immédiatement

$$\mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid X_k - \widehat{X}_k^-] = P_{k-1}\,F_k^*\ (P_k^-)^{-1}\ (X_k - \widehat{X}_k^-) = L_k\ (X_k - \widehat{X}_k^-) \ ,$$

avec la matrice de gain définie par

$$L_k = P_{k-1}\, F_k^* \,(P_k^-)^{-1} = P_{k-1}\, F_k^* \,[F_k\, P_{k-1}\, F_k^* + Q_k^W]^{-1} \ ,$$

et on en déduit que

$$U_{k-1}^n = \widehat{X}_{k-1} + \mathbb{E}[X_{k-1} - \widehat{X}_{k-1} \mid X_k - \widehat{X}_k^-] = \widehat{X}_{k-1} + L_k\, (X_k - \widehat{X}_k^-) \ .$$

$\square$

(vii) Montrer que

$$\widehat{X}_{k-1}^n = \mathbb{E}[X_{k-1} \mid Y_{0:n}] = \mathbb{E}[U_{k-1}^n \mid Y_{0:n}] \ ,$$

en utilisant le résultat obtenu à la question (iii).

En déduire que le lisseur de Kalman vérifie l'équation de récurrence rétrograde suivante

$$\widehat{X}_{k-1}^n = \widehat{X}_{k-1} + L_k\, (\widehat{X}_k^n - \widehat{X}_k^-) \qquad \textbf{avec} \qquad \widehat{X}_n^n = \widehat{X}_n \ ,$$

en utilisant le résultat obtenu à la question (vi).

Solution

Le vecteur aléatoire $(Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n})$ contient davantage d'information que le vecteur aléatoire $Y_{0:n}$, de sorte que

$$\widehat{X}_{k-1}^n = \mathbb{E}[X_{k-1} \mid Y_{0:n}] = \mathbb{E}[\, \mathbb{E}[X_{k-1} \mid Y_{0:k-1}, X_k - \widehat{X}_k^-, Z_{k+1:n}] \mid Y_{0:n}] = \mathbb{E}[U_{k-1}^n \mid Y_{0:n}] \ ,$$

et on vérifie que

$$\widehat{X}_{k-1}^n = \mathbb{E}[U_{k-1}^n \mid Y_{0:n}] = \widehat{X}_{k-1} + L_k\, (\mathbb{E}[X_k \mid Y_{0:n}] - \widehat{X}_k^-) = \widehat{X}_{k-1} + L_k\, (\widehat{X}_k^n - \widehat{X}_k^-) \ .$$

$\square$