Université de Rennes 1
Master Recherche SISEA

Examen du cours UE S3-2
“Filtre de Kalman et modèles de Markov cachés”
Jeudi 29 janvier 2009, 14:00 à 16:00
— Corrigé —

Exercice

Le but de cet exercice est de comprendre d'un point de vue qualitatif et quantitatif, l'impact d'une coopération entre algorithmes d'estimation. Pour être concret, on considère deux entités mobiles (par exemple deux robots ou deux individus) qui tentent séparément d'estimer leur position (par exemple à l'aide d'un modèle de déplacement, ou de mesures inertielles, et de mesures fournies par des capteurs) et on suppose qu'à un certain instant on dispose d'une mesure supplémentaire de leur position relative, c'est–à–dire de la différence entre les deux positions (par exemple quand les deux entités établissent un contact direct, par exemple un contact visuel dans le cas de deux individus). Pour simplifier encore, on définit le modèle statique suivant :

- la position inconnue de l'entité numéro i est estimée par le vecteur μ_i, et l'erreur additive d'estimation est supposée gaussienne centrée et de matrice de covariance P_i, pour $i = 1, 2$,
- on dispose d'une mesure Y de la différence $X_1 - X_2$ entre les deux positions, et l'erreur additive de mesure est supposée gaussienne centrée et de matrice de covariance Σ,
- on suppose que les différentes sources d'erreur ou d'incertitude sont statistiquement indépendantes.

On pose $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$ et $Z = \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} X_1 \\ X_2 \\ Y \end{pmatrix}$.

(i) **Montrer que le vecteur aléatoire Z est gaussien, et donner l'expression du vecteur moyenne $\bar{Z}$ en fonction de μ_1 et μ_2, et de la matrice de covariance Q_Z en fonction de P_1, P_2 et Σ.**

Solution

On a
$$X_i = \mu_i + W_i \qquad \text{avec} \qquad W_i \sim \mathcal{N}(0, P_i)$$
pour $i = 1, 2$ et
$$Y = (X_1 - X_2) + V = (\mu_1 - \mu_2) + W_1 - W_2 + V \qquad \text{avec} \qquad V \sim \mathcal{N}(0, \Sigma)$$
et compte tenu que les variables aléatoires W_1, W_2 et V sont indépendantes, on en déduit que le vecteur aléatoire
$$Z = \begin{pmatrix} X_1 \\ X_2 \\ Y \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_1 - \mu_2 \end{pmatrix} + \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ I & -I & I \end{pmatrix} \begin{pmatrix} W_1 \\ W_2 \\ V \end{pmatrix} ,$$
est gaussien, comme transformation affine d'un vecteur aléatoire gaussien, de moyenne
$$\bar{Z} = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \mu_1 - \mu_2 \end{pmatrix} ,$$
et de matrice de covariance
$$Q_Z = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ I & -I & I \end{pmatrix} \begin{pmatrix} P_1 & 0 & 0 \\ 0 & P_2 & 0 \\ 0 & 0 & \Sigma \end{pmatrix} \begin{pmatrix} I & 0 & I \\ 0 & I & -I \\ 0 & 0 & I \end{pmatrix} = \begin{pmatrix} P_1 & 0 & P_1 \\ 0 & P_2 & -P_2 \\ P_1 & -P_2 & P_1 + P_2 + \Sigma \end{pmatrix} .$$
□

(ii) **En utilisant un résultat donné en cours (mais sans le redémontrer), montrer que la distribution conditionnelle du vecteur aléatoire X sachant Y est gaussienne, de vecteur moyenne et de matrice de covariance**
$$\widehat{X} = \begin{pmatrix} \widehat{X}_1 \\ \widehat{X}_2 \end{pmatrix} \qquad \textbf{et} \qquad R = \begin{pmatrix} R_1 & R_{12} \\ R_{21} & R_2 \end{pmatrix}$$
respectivement, dont on donnera les expressions en fonction de μ_1, μ_2, P_1, P_2 et Σ.

Solution

On utilise le résultat fondamental sur le conditionnement dans les vecteurs aléatoires gaussiens, et on obtient
$$\widehat{X} = \begin{pmatrix} \widehat{X}_1 \\ \widehat{X}_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix} + \begin{pmatrix} P_1 \\ -P_2 \end{pmatrix} \; (P_1 + P_2 + \Sigma)^{-1} \; (Y - (\mu_1 - \mu_2)) ,$$

et

$$R = \begin{pmatrix} R_1 & R_{12} \\ R_{21} & R_2 \end{pmatrix} = \begin{pmatrix} P_1 & 0 \\ 0 & P_2 \end{pmatrix} - \begin{pmatrix} P_1 \\ -P_2 \end{pmatrix} (P_1 + P_2 + \Sigma)^{-1} \begin{pmatrix} P_1 & -P_2 \end{pmatrix} \; .$$

On en déduit que

$$\widehat{X}_1 = \mu_1 + P_1 \, (P_1 + P_2 + \Sigma)^{-1} \, (Y - (\mu_1 - \mu_2)) \; ,$$

et

$$\widehat{X}_2 = \mu_2 - P_2 \, (P_1 + P_2 + \Sigma)^{-1} \, (Y - (\mu_1 - \mu_2)) \; ,$$

et

$$R_1 = P_1 - P_1 \, (P_1 + P_2 + \Sigma)^{-1} \, P_1 \qquad \text{et} \qquad R_2 = P_2 - P_2 \, (P_1 + P_2 + \Sigma)^{-1} \, P_2 \; ,$$

et

$$R_{12} = P_1 \, (P_1 + P_2 + \Sigma)^{-1} \, P_2 \; .$$

□

(iii) **On suppose que les matrices de covariance $P_1 = \sigma_1^2 \, I$ et $P_2 = \sigma_2^2 \, I$ sont diagonales (et même multiples scalaires de la matrice identité), et par exemple $\sigma_1 \leq \sigma_2$. Lequel des deux estimateurs μ_1 ou μ_2 subit la correction la plus forte après prise en compte de la mesure de la position relative ? Finalement, comment se comparent les matrices de covariance R_1 et R_2, c'est–à–dire lequel des deux estimateurs $\widehat{X}_1$ ou $\widehat{X}_2$ possède la plus petite variance, après correction ?**

Au besoin, on pourra commencer par considérer le cas plus simple où la matrice de covariance $\Sigma = s^2 \, I$ elle aussi est multiple scalaire de la matrice identité.

Solution

On remarque que

$$\widehat{X}_1 - \mu_1 = \sigma_1^2 \, u \qquad \text{et} \qquad \widehat{X}_2 - \mu_2 = -\sigma_2^2 \, u \; ,$$

avec

$$u = (P_1 + P_2 + \Sigma)^{-1} \, (Y - (\mu_1 - \mu_2)) \; .$$

On en déduit que les corrections $(\widehat{X}_1 - \mu_1)$ et $(\widehat{X}_2 - \mu_2)$ sont colinéaires, de sens opposés puisque $(\widehat{X}_1 - \mu_1) \, (\widehat{X}_2 - \mu_2)^* \leq 0$, et que l'estimateur le moins précis subit la correction la plus forte, puisque $|\widehat{X}_1 - \mu_1| \leq |\widehat{X}_2 - \mu_2|$, compte tenu que $\sigma_1 \leq \sigma_2$.

Dans le cas simple où $\Sigma = s^2 \, I$, on obtient

$$R_1 = \frac{\sigma_1^2 \, (\sigma_2^2 + s^2)}{\sigma_1^2 + \sigma_2^2 + s^2} \, I \qquad \text{et} \qquad R_2 = \frac{\sigma_2^2 \, (\sigma_1^2 + s^2)}{\sigma_1^2 + \sigma_2^2 + s^2} \, I \; ,$$

de sorte que $R_1 \leq R_2$, compte tenu que $\sigma_1 \leq \sigma_2$. Dans le cas général, on remarque que

$$R_i = \sigma_i^2 \, I - \sigma_i^4 \, (P_1 + P_2 + \Sigma)^{-1} \ ,$$

pour $i = 1, 2$, de sorte que par différence

$$\begin{aligned} R_1 - R_2 &= (\sigma_1^2 - \sigma_2^2) \, I - (\sigma_1^4 - \sigma_2^4) \, (P_1 + P_2 + \Sigma)^{-1} \\ &= (\sigma_1^2 - \sigma_2^2) \, (I - (\sigma_1^2 + \sigma_2^2) \, (P_1 + P_2 + \Sigma)^{-1}) \\ &= (\sigma_1^2 - \sigma_2^2) \, (P_1 + P_2 + \Sigma - (\sigma_1^2 + \sigma_2^2) \, I) \, (P_1 + P_2 + \Sigma)^{-1} \\ &= (\sigma_1^2 - \sigma_2^2) \, \Sigma \, (P_1 + P_2 + \Sigma)^{-1} \ . \end{aligned}$$

La matrice de covariance Σ admet la décomposition $\Sigma = O \, D \, O^*$ où la matrice O est orthogonale et où les composantes de la matrice diagonale D sont positives. Elle admet donc une racine carrée symétrique $\Sigma^{1/2} = O \, D^{1/2} \, O^*$ qui commute avec Σ (et avec $P_1 + P_2 = (\sigma_1^2 + \sigma_2^2) \, I$ aussi, bien sûr). On vérifie que

$$\Sigma^{1/2} \, (P_1 + P_2 + \Sigma)^{-1} \, \Sigma^{1/2} \, (P_1 + P_2 + \Sigma) = \Sigma \ ,$$

de sorte que

$$\Sigma^{1/2} \, (P_1 + P_2 + \Sigma)^{-1} \, \Sigma^{1/2} = \Sigma \, (P_1 + P_2 + \Sigma)^{-1} \ ,$$

et en reportant ci–dessus

$$R_1 - R_2 = (\sigma_1^2 - \sigma_2^2) \, \Sigma^{1/2} \, (P_1 + P_2 + \Sigma)^{-1} \, \Sigma^{1/2} \ ,$$

de sorte que $R_1 \leq R_2$, compte tenu que $\sigma_1 \leq \sigma_2$ et que la matrice symétrique

$$\Sigma^{1/2} \, (P_1 + P_2 + \Sigma)^{-1} \, \Sigma^{1/2} \ ,$$

est positive. On en déduit que l'estimateur initialement le plus précis, reste le plus précis après correction.

$\square$

Problème

Le but de ce problème est d'établir les équations de filtres gaussiens, comme alternatives au filtre de Kalman étendu, pour des systèmes non–linéaires avec bruits additifs gaussiens. Au lieu de s'appuyer sur une linéarisation des fonctions autour de l'estimateur courant, l'idée qui sera développée dans la suite consiste

- à approcher les différentes densités conditionnelles par des densités gaussiennes ayant la même moyenne et la même matrice de covariance,
- à utiliser des formules de quadrature, développées initialement pour le calcul numérique d'intégrales, pour approcher ces moyennes et ces matrices de covariance conditionnelles.

Pour être plus spécifique, on considère une suite d'états cachés $\{X_k\}$ à valeurs dans $\mathbb{R}^m$, vérifiant

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

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

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

et on suppose que

- la condition initiale X_0 est un vecteur aléatoire gaussien, de moyenne m_0 et de matrice de covariance Σ_0,
- la suite $\{W_k\}$ est un bruit blanc gaussien, de matrice de covariance Q_k,
- la suite $\{V_k\}$ est un bruit blanc gaussien, de matrice de covariance R_k,
- les suites $\{W_k\}$ et $\{V_k\}$ et la condition initiale X_0 sont mutuellement indépendants,

et pour simplifier encore, on suppose que les matrices de covariance Σ_0, Q_k et R_k sont inversibles. On rappelle que les distributions de probabilité conditionnelles

$$\mathbb{P}[X_k \in dx \mid Y_{0:k}] = p_k(x)\, dx \qquad \text{et} \qquad \mathbb{P}[X_k \in dx \mid Y_{0:k-1}] = p_k^-(x)\, dx$$

vérifient

$$p_k^-(x') = \frac{1}{(\sqrt{2\pi})^m \sqrt{\det Q_k}} \int_{\mathbb{R}^m} \exp\{-\tfrac{1}{2}\,(x' - f_k(x))^*\, Q_k^{-1}\,(x' - f_k(x))\}\; p_{k-1}(x)\, dx \ ,$$

et

$$p_k(x) \propto \exp\{-\tfrac{1}{2}\,(Y_k - h_k(x))^*\, R_k^{-1}\,(Y_k - h_k(x))\}\; p_k^-(x) \ ,$$

à une constante multiplicative près. Le calcul des deux premiers moments (moyenne et matrice de covariance) de la densité $p_k^-(x')$ est assez facile, et fait l'objet de la question (i). En revanche, le calcul des deux premiers moments de la densité $p_k(x)$ n'est pas immédiat, et on utilise l'approche suivante. On rappelle que la distribution de probabilité conditionnelle jointe

$$\mathbb{P}[X_k \in dx, Y_k \in dy \mid Y_{0:k-1}] = q_k^-(x,y)\, dx\, dy$$

vérifie

$$q_k^-(x,y) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det R_k}} \; \exp\{-\tfrac{1}{2}\,(y - h_k(x))^*\, R_k^{-1}\,(y - h_k(x))\}\; p_k^-(x) \; .$$

Le calcul des deux premiers moments de la densité jointe $q_k^-(x,y)$ est assez facile, et fait l'objet de la question (iii). En vertu du principe énoncé en introduction, on approche la densité jointe $q_k^-(x,y)$ par la densité gaussienne ayant les mêmes deux premiers moments, et on utilise le résultat fondamental sur le conditionnement dans les vecteurs aléatoires gaussiens pour calculer les deux premiers moments de la densité $p_k(x)$.

(i) **Montrer que la moyenne conditionnelle du vecteur aléatoire X_k sachant $Y_{0:k-1}$, c'est–à–dire la moyenne de la densité conditionnelle $p_k^-(x')$, est donnée par**

$$\widehat{X}_k^- = \int_{\mathbb{R}^m} x'\, p_k^-(x')\, dx' = \int_{\mathbb{R}^m} f_k(x)\, p_{k-1}(x)\, dx \; ,$$

et de même que la matrice de covariance conditionnelle du vecteur aléatoire X_k sachant $Y_{0:k-1}$, c'est–à–dire la matrice de covariance de la densité conditionnelle $p_k^-(x')$, est donnée par

$$\begin{aligned} P_k^- &= \int_{\mathbb{R}^m} (x' - \widehat{X}_k^-)\,(x' - \widehat{X}_k^-)^*\, p_k^-(x')\, dx' \\ &= Q_k + \int_{\mathbb{R}^m} (f_k(x) - \widehat{X}_k^-)\,(f_k(x) - \widehat{X}_k^-)^*\, p_{k-1}(x)\, dx \; , \end{aligned}$$

Solution

On introduit la notation

$$m(x,x') = \frac{1}{(\sqrt{2\pi})^m \sqrt{\det Q_k}} \; \exp\{-\tfrac{1}{2}\,(x' - f_k(x))^*\, Q_k^{-1}\,(x' - f_k(x))\} \; ,$$

pour représenter la densité gaussienne m–dimensionnelle, paramétrée par x, de moyenne $f_k(x)$ et de matrice de covariance Q_k, de sorte que

$$\int_{\mathbb{R}^m} x'\; m(x,x')\, dx' = f_k(x) \qquad \text{et} \qquad \int_{\mathbb{R}^m} (x' - f_k(x))\; (x' - f_k(x))^*\; m(x,x')\, dx' = Q_k \; ,$$

et

$$
\begin{aligned}
\int_{\mathbb{R}^m} (x'-u)\,(x'-u)^*\, m(x,x')\,dx' \;=\; & \int_{\mathbb{R}^m} (x'-f_k(x))\,(x'-f_k(x))^*\, m(x,x')\,dx' \\
& + \int_{\mathbb{R}^m} (x'-f_k(x))\,(f_k(x)-u)^*\, m(x,x')\,dx' \\
& + \int_{\mathbb{R}^m} (f_k(x)-u)\,(x'-f_k(x))^*\, m(x,x')\,dx' \\
& + \int_{\mathbb{R}^m} (f_k(x)-u)\,(f_k(x)-u)^*\, m(x,x')\,dx' \\
= \; & Q_k + (f_k(x)-u)\,(f_k(x)-u)^* \;,
\end{aligned}
$$

pour tout $u \in \mathbb{R}^m$, compte tenu que le terme croisé

$$
\int_{\mathbb{R}^m} (x'-f_k(x))\,(f_k(x)-u)^*\, m(x,x')\,dx' = \Big\{ \int_{\mathbb{R}^m} (x'-f_k(x))\, m(x,x')\,dx' \Big\}\,(f_k(x)-u)^* \;,
$$

est nul. On vérifie que

$$
\begin{aligned}
\widehat{X}_k^- &= \int_{\mathbb{R}^m} x'\, p_k^-(x')\,dx' \\
&= \int_{\mathbb{R}^m} x' \,\Big\{ \int_{\mathbb{R}^m} m(x,x')\, p_{k-1}(x)\,dx \Big\}\,dx' \\
&= \int_{\mathbb{R}^m} \Big\{ \int_{\mathbb{R}^m} x'\, m(x,x')\,dx' \Big\}\, p_{k-1}(x)\,dx \\
&= \int_{\mathbb{R}^m} f_k(x)\, p_{k-1}(x)\,dx \;,
\end{aligned}
$$

et

$$
\begin{aligned}
P_k^- &= \int_{\mathbb{R}^m} (x'-\widehat{X}_k^-)\,(x'-\widehat{X}_k^-)^*\, p_k^-(x')\,dx' \\
&= \int_{\mathbb{R}^m} (x'-\widehat{X}_k^-)\,(x'-\widehat{X}_k^-)^* \Big\{ \int_{\mathbb{R}^m} m(x,x')\, p_{k-1}(x)\,dx \Big\}\, p_k^-(x')\,dx' \\
&= \int_{\mathbb{R}^m} \Big\{ \int_{\mathbb{R}^m} (x'-\widehat{X}_k^-)\,(x'-\widehat{X}_k^-)^*\, m(x,x')\,dx' \Big\}\, p_{k-1}(x)\,dx \\
&= \int_{\mathbb{R}^m} \Big\{ Q_k + (f_k(x)-\widehat{X}_k^-)\,(f_k(x)-\widehat{X}_k^-)^* \Big\}\, p_{k-1}(x)\,dx \\
&= Q_k + \int_{\mathbb{R}^m} (f_k(x)-\widehat{X}_k^-)\,(f_k(x)-\widehat{X}_k^-)^*\, p_{k-1}(x)\,dx \;.
\end{aligned}
$$

$\square$

En vertu du principe énoncé en introduction, on approche la densité $p_{k-1}(x)$ par la densité gaussienne de moyenne $\widehat{X}_{k-1}$ et de matrice de covariance P_{k-1}.

(ii) **En effectuant le changement de variable $x \to \widehat{X}_{k-1} + S_{k-1}\, u$, où la matrice S_{k-1} est définie par $P_{k-1} = S_{k-1}\ S_{k-1}^*$, montrer qu'on obtient les approximations**

$$\widehat{X}_k^- = \int_{\mathbb{R}^m} \widehat{f}_k(u)\ \exp\{-\tfrac{1}{2}\,|u|^2\}\ \frac{du}{(2\pi)^{m/2}}\ ,$$

et

$$P_k^- = Q_k + \int_{\mathbb{R}^m} (\widehat{f}_k(u) - \widehat{X}_k^-)\,(\widehat{f}_k(u) - \widehat{X}_k^-)^*\ \exp\{-\tfrac{1}{2}\,|u|^2\}\ \frac{du}{(2\pi)^{m/2}}\ ,$$

où $\widehat{f}_k(u) = f_k(\widehat{X}_{k-1} + S_{k-1}\, u)$ par définition.

Solution

D'après l'expression obtenue en réponse à la question (i), on a

$$\begin{aligned}
\widehat{X}_k^- &= \int_{\mathbb{R}^m} f_k(x)\, p_{k-1}(x)\, dx \\
&\approx \int_{\mathbb{R}^m} f_k(x)\ \exp\{-\tfrac{1}{2}\,(x-\widehat{X}_{k-1})^*\, P_{k-1}^{-1}\,(x-\widehat{X}_{k-1})\}\ \frac{dx}{(2\pi)^{m/2}\sqrt{\det P_{k-1}}} \\
&= \int_{\mathbb{R}^m} f_k(\widehat{X}_{k-1} + S_{k-1}\, u)\ \exp\{-\tfrac{1}{2}\, u^*\, S_{k-1}^*\, P_{k-1}^{-1}\, S_{k-1}\, u\}\ \frac{|\det S_{k-1}|\ dx}{(2\pi)^{m/2}\sqrt{\det P_{k-1}}} \\
&= \int_{\mathbb{R}^m} \widehat{f}_k(u)\ \exp\{-\tfrac{1}{2}\,|u|^2\}\ \frac{du}{(2\pi)^{m/2}}\ ,
\end{aligned}$$

et

$$P_k^- = Q_k + \int_{\mathbb{R}^m} (\widehat{f}_k(u) - \widehat{X}_k^-)\,(\widehat{f}_k(u) - \widehat{X}_k^-)^*\ \exp\{-\tfrac{1}{2}\,|u|^2\}\ \frac{du}{(2\pi)^{m/2}}\ ,$$

où $\widehat{f}_k(u) = f_k(\widehat{X}_{k-1} + S_{k-1}\, u)$ par définition.

$\square$

(iii) **Montrer que la moyenne conditionnelle du vecteur aléatoire $Z_k = (X_k, Y_k)$ sachant $Y_{0:k-1}$, c'est–à–dire la moyenne de la densité conditionnelle $q_k^-(x,y)$, est donnée par**

$$\widehat{X}_k^- = \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} x\, q_k^-(x,y)\, dx\, dy = \int_{\mathbb{R}^m} x\, p_k^-(x)\, dx\ ,$$

c'est–à–dire qu'on retrouve le résultat obtenu à la question (i), et

$$\widehat{Y}_k^- = \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} y\, q_k^-(x,y)\, dx\, dy = \int_{\mathbb{R}^m} h_k(x)\, p_k^-(x)\, dx \ ,$$

et de même que la matrice de coavariance conditionnelle du vecteur aléatoire $Z_k = (X_k, Y_k)$ sachant $Y_{0:k-1}$, c'est–à–dire la matrice de covariance de la densité conditionnelle $q_k^-(x,y)$, est donnée par

$$\begin{aligned} P_k^- &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (x - \widehat{X}_k^-)\, (x - \widehat{X}_k^-)^*\, q_k^-(x,y)\, dx\, dy \\ &= \int_{\mathbb{R}^m} (x - \widehat{X}_k^-)\, (x - \widehat{X}_k^-)^*\, p_k^-(x)\, dx \ , \end{aligned}$$

c'est–à–dire qu'on retrouve le résultat obtenu à la question (i)

$$\begin{aligned} C_k &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (x - \widehat{X}_k^-)\, (y - \widehat{Y}_k^-)^*\, q_k^-(x,y)\, dx\, dy \\ &= \int_{\mathbb{R}^m} (x - \widehat{X}_k^-)\, (h_k(x) - \widehat{Y}_k^-)^*\, p_k^-(x) \ , \end{aligned}$$

et

$$\begin{aligned} \Xi_k &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (y - \widehat{Y}_k^-)\, (y - \widehat{Y}_k^-)^*\, q_k^-(x,y)\, dx\, dy \\ &= R_k + \int_{\mathbb{R}^m} (h_k(x) - \widehat{Y}_k^-)\, (h_k(x) - \widehat{Y}_k^-)^*\, p_k^-(x)\, dx \ , \end{aligned}$$

SOLUTION

En intégrant par rapport à la variable y, on remarque que

$$\int_{\mathbb{R}^d} q_k^-(x,y)\; dy = p_k^-(x) \ ,$$

c'est–à–dire que la densité $p_k^-(x)$ est la densité marginale de la densité jointe $q_k^-(x,y)$. On en déduit que

$$\begin{aligned} \widehat{X}_k^- &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} x\; q_k^-(x,y)\, dx\, dy \\ &= \int_{\mathbb{R}^m} x \,\{ \int_{\mathbb{R}^d} q_k^-(x,y)\, dy \,\}\, dx \\ &= \int_{\mathbb{R}^m} x\, p_k^-(x)\, dx \ , \end{aligned}$$

et

$$\begin{aligned} P_k^- &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (x - \widehat{X}^-)\ (x - \widehat{X}^-)^*\ q_k^-(x,y)\, dx\, dy \\ &= \int_{\mathbb{R}^m} (x - \widehat{X}^-)\ (x - \widehat{X}^-)^*\ \{\ \int_{\mathbb{R}^d} q_k^-(x,y)\, dy\ \}\, dx \\ &= \int_{\mathbb{R}^m} (x - \widehat{X}^-)\ (x - \widehat{X}^-)^*\ p_k^-(x)\, dx\ , \end{aligned}$$

c'est–à–dire qu'on retrouve les résultats obtenus à la question (i),

On introduit la notation

$$m(x,y) = \frac{1}{(\sqrt{2\pi})^d \sqrt{\det R_k}}\ \exp\{-\tfrac{1}{2}\,(y - h_k(x))^*\, R_k^{-1}\,(y - h_k(x))\}\ ,$$

pour représenter la densité gaussienne d–dimensionnelle, paramétrée par x, de moyenne $h_k(x)$ et de matrice de covariance R_k, de sorte que

$$\int_{\mathbb{R}^d} y\ m(x,y)\, dy = h_k(x) \qquad \text{et} \qquad \int_{\mathbb{R}^d} (y - h_k(x))\ (y - h_k(x))^*\ m(x,y)\, dy = R_k\ ,$$

et

$$\begin{aligned} \int_{\mathbb{R}^d} (y - v)\ (y - v)^*\ m(x,y)\, dy\ &=\ \int_{\mathbb{R}^d} (y - h_k(x))\ (y - h_k(x))^*\ m(x,y)\, dy \\ &\quad + \int_{\mathbb{R}^d} (y - h_k(x))\ (h_k(x) - v)^*\ m(x,y)\, dy \\ &\quad + \int_{\mathbb{R}^d} (h_k(x) - v)\ (y - h_k(x))^*\ m(x,y)\, dy \\ &\quad + \int_{\mathbb{R}^d} (h_k(x) - v)\ (h_k(x) - v)^*\ m(x,y)\, dy \\ &=\ R_k + (h_k(x) - v)\ (h_k(x) - v)^*\ , \end{aligned}$$

pour tout $v \in \mathbb{R}^d$, compte tenu que le terme croisé

$$\int_{\mathbb{R}^d} (y - h_k(x))\ (h_k(x) - v)^*\ m(x,y)\, dy = \{\ \int_{\mathbb{R}^d} (y - h_k(x))\ m(x,y)\, dy\ \}\, (h_k(x) - v)^*\ ,$$

est nul. On vérifie que

$$
\begin{aligned}
\widehat{Y}_k^- &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} y \; q_k^-(x,y)\, dx\, dy \\
&= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} y \; m(x,y)\, p_k^-(x)\, dx\, dy \\
&= \int_{\mathbb{R}^m} \{ \int_{\mathbb{R}^d} y \; m(x,y)\, dy \,\}\, p_k^-(x)\, dx \\
&= \int_{\mathbb{R}^m} h_k(x)\, p_k^-(x)\, dx \ ,
\end{aligned}
$$

$$
\begin{aligned}
C_k &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (x - \widehat{X}_k^-)\; (y - \widehat{Y}_k^-)^*\; q_k^-(x,y)\, dx\, dy \\
&= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (x - \widehat{X}_k^-)\; (y - \widehat{Y}_k^-)^*\; m(x,y)\; p_k^-(x)\, dx\, dy \\
&= \int_{\mathbb{R}^m} (x - \widehat{X}_k^-)\; \{ \int_{\mathbb{R}^d} (y - \widehat{Y}_k^-)^*\; m(x,y)\; dy \,\}\, p_k^-(x)\, dx \\
&= \int_{\mathbb{R}^m} (x - \widehat{X}_k^-)\; (h_k(x) - \widehat{Y}_k^-)^*\; p_k^-(x)\, dx \ ,
\end{aligned}
$$

et

$$
\begin{aligned}
\Xi_k &= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (y - \widehat{Y}_k^-)\; (y - \widehat{Y}_k^-)^*\; q_k^-(x,y)\, dx\, dy \\
&= \int_{\mathbb{R}^m} \int_{\mathbb{R}^d} (y - \widehat{Y}_k^-)\; (y - \widehat{Y}_k^-)^*\; m(x,y)\; p_k^-(x)\, dx\, dy \\
&= \int_{\mathbb{R}^m} \{ \int_{\mathbb{R}^d} (y - \widehat{Y}_k^-)\; (y - \widehat{Y}_k^-)^*\; m(x,y)\; dy \,\}\, p_k^-(x)\, dx \\
&= \int_{\mathbb{R}^m} \{ R_k + (h_k(x) - \widehat{Y}_k^-)\; (h_k(x) - \widehat{Y}_k^-)^* \,\}\, p_k^-(x)\, dx \\
&= R_k + \int_{\mathbb{R}^m} (h_k(x) - \widehat{Y}_k^-)\; (h_k(x) - \widehat{Y}_k^-)^*\; p_k^-(x)\, dx \ .
\end{aligned}
$$

□

Si on approche la densité jointe $q_k^-(x,y)$ par la densité gaussienne de moyenne et de matrice de covariance

$$
\begin{pmatrix} \widehat{X}_k^- \\ \widehat{Y}_k^- \end{pmatrix} \qquad \text{et} \qquad \begin{pmatrix} P_k^- & C_k \\ C_k^* & \Xi_k \end{pmatrix} \ ,
$$

respectivement, alors on obtient par conditionnement les approximations suivantes

$$\widehat{X}_k = \widehat{X}_k^- + C_k \, \Xi_k^{-1} \, (Y_k - \widehat{Y}_k^-) \ ,$$

et

$$P_k = P_k^- - C_k \, \Xi_k^{-1} \, C_k^* \ ,$$

pour les deux premiers moments de la densité $p_k(x)$.

En vertu du principe énoncé en introduction, on approche la densité $p_k^-(x)$ par la densité gaussienne de moyenne $\widehat{X}_k^-$ et de matrice de covariance P_k^-.

(iv) **En effectuant le changement de variable $x \to \widehat{X}_k^- + S_k^- \, u$, où la matrice S_k^- est définie par $P_k^- = S_k^- \, (S_k^-)^*$, montrer qu'on obtient les approximations**

$$\widehat{Y}_k^- = \int_{\mathbb{R}^m} \widehat{h}_k(u) \ \exp\{-\tfrac{1}{2} \, |u|^2\} \ \frac{du}{(2\pi)^{m/2}} \ ,$$

$$C_k = S_k^- \ \int_{\mathbb{R}^m} u \, (\widehat{h}_k(u) - \widehat{Y}_k^-)^* \ \exp\{-\tfrac{1}{2} \, |u|^2\} \ \frac{du}{(2\pi)^{m/2}} \ ,$$

et

$$\Xi_k = R_k + \int_{\mathbb{R}^m} (\widehat{h}_k(u) - \widehat{Y}_k^-) \, (\widehat{h}_k(u) - \widehat{Y}_k^-)^* \ \exp\{-\tfrac{1}{2} \, |u|^2\} \ \frac{du}{(2\pi)^{m/2}} \ ,$$

où $\widehat{h}_k(u) = h_k(\widehat{X}_k^- + S_k^- \, u)$ par définition.

Il reste donc à calculer les intégrales des fonctions non–linéaires $\widehat{f}_k(u)$, $\widehat{f}_k(u) \, \widehat{f}_k^*(u)$, $\widehat{h}_k(u)$, $u \, \widehat{h}_k^*(u)$ et $\widehat{h}_k(u) \, \widehat{h}_k^*(u)$ par rapport à la densité gaussienne réduite centrée, ce qui peut être réalisé de manière approchée, en utilisant des formules de quadrature de type Gauss–Hermite. En dimension $m = 1$, ces formules de quadrature sont de la forme

$$\int_{\mathbb{R}} F(u) \ \exp\{-\tfrac{1}{2} \, |u|^2\} \ \frac{du}{\sqrt{2\pi}} \approx \sum_{i=1}^p w_i \, F(u^i) \ ,$$

où les poids positifs $w_1, \ldots, w_p$ et les points $u^1, \ldots, u^p$ sont choisis de sorte que l'approximation soit exacte pour tout polynôme de degré inférieur ou égal à p.

(v) **Montrer que les formules de quadrature en dimension m peuvent s'écrire**

$$\int_{\mathbb{R}^m} F(u) \ \exp\{-\tfrac{1}{2} \, |u|^2\} \ \frac{du}{(2\pi)^{m/2}} \approx \sum_{i_1=1}^p \cdots \sum_{i_m=1}^p w_{i_1} \cdots w_{i_m} \ F(u_1^{i_1}, \ldots, u_m^{i_m}) \ .$$

Combien d'évaluations de la fonction F sont–elles nécessaires ?