École Nationale Supérieure
de Techniques Avancées
Module : Commande des Systèmes

Examen du cours B7–1
"Filtrage bayésien optimal
et approximation particulaire"
Lundi 19 octobre 2009, 8:30 à 10:00

Exercice 1 :

L'objectif de cet exercice est d'établir les équations du filtre bayésien, dans le cas plus général d'une chaîne de Markov observée dans un bruit additif défini par un modèle ARMA (auto–regressive moving average). Plus spécifiquement, on suppose que les états cachés sont à valeurs dans $\mathbb{R}^m$ et vérifient la relation récurrente

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

où $\{W_k\}$ est une suite de vecteurs aléatoires indépendants, et que les observations sont à valeurs dans $\mathbb{R}^d$ et sont reliées aux états cachés par

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

où $\{V_k\}$ est une suite de vecteurs aléatoires *dépendants* à valeurs dans $\mathbb{R}^d$, définie par le modèle ARMA(p,q)

$$V_k + A_1\, V_{k-1} + \cdots + A_p\, V_{k-p} = B_0\, e_k + B_1\, e_{k-1} + \cdots + B_q\, e_{k-q} \ ,$$

paramétré par les matrices $A_1, \cdots, A_p$ et $B_0, B_1, \cdots, B_q$, et où $\{e_k\}$ est une suite de vecteurs aléatoires indépendants à valeurs dans $\mathbb{R}^d$. Par hypothèse, la matrice B_0 est inversible. On suppose que la condition initiale X_0 et les suites $\{W_k\}$ et $\{e_k\}$ sont mutuellement indépendantes.

On admet qu'il existe une représentation d'état pour le bruit d'observation, de la forme

$$U_k \ = \ A\, U_{k-1} + B\, e_{k-1} \ ,$$

$$V_k \ = \ C\, U_k + B_0\, e_k \ ,$$

où la suite $\{U_k\}$ est à valeurs dans $\mathbb{R}^{rd}$, et où A, B et C sont des matrices de dimension $rd \times rd$, $rd \times d$ et $d \times rd$, respectivement, avec $r = \max(p,q)$.

(i) **On définit la nouvelle variable d'état $Z_k = (X_k, U_k)$. Montrer que les nouveaux états cachés ainsi définis vérifient la relation récurrente**

$$Z_k = f_k^{A,B}(Z_{k-1}, e_{k-1}, W_k) \ , \tag{$\star$}$$

où on donnera l'expression de la fonction $f_k^{A,B}$ ainsi définie.

Solution

Clairement

$$Z_k = \begin{pmatrix} X_k \\ \\ U_k \end{pmatrix} = \begin{pmatrix} f_k(X_{k-1}, W_k) \\ \\ A\, U_{k-1} + B\, e_{k-1} \end{pmatrix} = f_k^{A,B}(Z_{k-1}, e_{k-1}, W_k) \ ,$$

et la fonction $f_k^{A,B}$ est définie par

$$f_k^{A,B}(z, e, w) = \begin{pmatrix} f_k(x, w) \\ \\ A\, u + B\, e \end{pmatrix} \ ,$$

pour tout $z = (x, u)$, pour tout e et pour tout w.

□

(ii) **Montrer que les observations sont reliées aux nouveaux états cachés par**

$$Y_k = h_k^C(Z_k) + B_0\, e_k \ , \tag{$\star\star$}$$

où on donnera l'expression de la fonction h_k^C ainsi définie.

Solution

Clairement

$$Y_k = h_k(X_k) + V_k = h_k(X_k) + C\, U_k + B_0\, e_k = h_k^C(Z_k) + B_0\, e_k \ ,$$

et la fonction h_k^C est définie par

$$h_k^C(z) = h_k(x) + C\, u \ ,$$

pour tout $z = (x, u)$.

□

On constate que les bruits qui apparaissent dans le modèle d'état ($\star$) et les bruits qui apparaissent l'équation d'observation ($\star\star$) ne sont pas indépendants. Le modèle obtenu n'appartient donc pas à la classe des modèles de Markov cachés.

(iii) **Montrer qu'on peut écrire le modèle ($\star$), ($\star\star$) sous la forme**

$$Z_k \;=\; f_k^{A,B,C}(Z_{k-1}, Y_{k-1}, W_k) \;,$$

$$Y_k \;=\; h_k^C(Z_k) + B_0\, e_k \;,$$

où on donnera l'expression de la fonction $f_k^{A,B,C}$ ainsi définie. On pourra par exemple utiliser l'équation ($\star\star$) exprimée à l'instant $(k-1)$ pour obtenir e_{k-1} en fonction de Y_{k-1} et Z_{k-1}, et reporter dans ($\star$) l'expression obtenue.

Solution

Compte tenu que la matrice B_0 est inversible, par hypothèse, l'équation ($\star\star$) exprimée à l'instant $(k-1)$ donne

$$e_{k-1} = B_0^{-1}\,(Y_{k-1} - h_{k-1}(X_{k-1}) - C\,U_{k-1}) \;,$$

et en reportant dans ($\star$), on obtient

$$U_k = A\,U_{k-1} + B\,e_{k-1} = A\,U_{k-1} + B\,B_0^{-1}\,(Y_{k-1} - h_{k-1}(X_{k-1}) - C\,U_{k-1}) \;,$$

de sorte que

$$Z_k = \begin{pmatrix} X_k \\ \\ U_k \end{pmatrix} = \begin{pmatrix} f_k(X_{k-1}, W_k) \\ \\ A\,U_{k-1} + B\,B_0^{-1}\,(Y_{k-1} - h_{k-1}(X_{k-1}) - C\,U_{k-1}) \end{pmatrix}$$

$$= f_k^{A,B,C}(Z_{k-1}, Y_{k-1}, W_k) \;,$$

où la fonction $f_k^{A,B,C}$ est définie par

$$f_k^{A,B,C}(z, y, w) = \begin{pmatrix} f_k(x, w) \\ \\ A\,u + B\,B_0^{-1}\,(y - h_{k-1}(x) - C\,u) \end{pmatrix} \;,$$

pour tout $z = (x, u)$, pour tout y et pour tout w.

$\square$

On introduit l'hypothèse supplémentaire que le vecteur aléatoire e_k possède une densité $q_k(e)$ par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$.

(iv) **À quelle classe de modèle vue en cours appartient le modèle obtenu à la question (iii) ? On pourra par exemple exprimer la distribution de probabilité conditionnelle jointe du couple (Z_k, Y_k) sachant (Z_{k-1}, Y_{k-1}).**

Compte tenu que $\{W_k\}$ et $\{e_k\}$ sont deux suites indépendantes de vecteurs aléatoires indépendants, les nouveaux états cachés et les observations forment conjointement une chaîne de Markov à valeurs dans $\mathbb{R}^{m+rd} \times \mathbb{R}^d$. Le noyau de probabilités de transition se factorise comme

$$\mathbb{P}[Z_k \in dz', Y_k \in dy' \mid Z_{k-1} = z, Y_{k-1} = y]$$

$$= \mathbb{P}[Y_k \in dy' \mid Z_k = z', Z_{k-1} = z, Y_{k-1} = y]\, \mathbb{P}[Z_k \in dz' \mid Z_{k-1} = z, Y_{k-1} = y]$$

$$= \mathbb{P}[Y_k \in dy' \mid Z_k = z']\, \mathbb{P}[Z_k \in dz' \mid Z_{k-1} = z, Y_{k-1} = y]$$

et compte tenu que les vecteurs aléatoires Z_k et e_k sont indépendants, on vérifie que

$$\begin{aligned}\mathbb{E}[\Psi(Y_k) \mid Z_k = z'] &= \mathbb{E}[\Psi(h_k^C(Z_k) + B_0\, e_k) \mid Z_k = z'] \\ &= \int_{\mathbb{R}^d} \Psi(h_k^C(z') + B_0\, e)\, q_k(e)\, de \\ &= \int_{\mathbb{R}^d} \Psi(y')\, q_k(B_0^{-1}\, (y' - h_k^C(z')))\, \frac{dy'}{|\det B_0|}\ ,\end{aligned}$$

pour toute fonction test Ψ définie sur $\mathbb{R}^d$, avec le changement de variable $y' = h_k^C(z') + B_0\, e$ et le jacobien associé, de sorte que la distribution de probabilité conditionnelle

$$\mathbb{P}[Y_k \in dy' \mid Z_k = z'] = q_k(B_0^{-1}\, (y' - h_k^C(z')))\, \frac{dy'}{|\det B_0|}\ ,$$

est absolument continue par rapport à la mesure de Lebesgue sur $\mathbb{R}^d$. On en déduit que le noyau de probabilités de transition peut s'écrire comme

$$\mathbb{P}[Z_k \in dz', Y_k \in dy' \mid Z_{k-1} = z, Y_{k-1} = y]$$

$$= \mathbb{P}[Y_k \in dy' \mid Z_k = z']\, \mathbb{P}[Z_k \in dz' \mid Z_{k-1} = z, Y_{k-1} = y]$$

$$= \mathbb{P}[Z_k \in dz' \mid Z_{k-1} = z, Y_{k-1} = y]\, g_k(z', y')\, dy'$$

$$= R_k(z, y, y', dz')\, dy'\ ,$$

avec l'expression suivante pour la densité

$$g_k(z', y') = q_k(B_0^{-1}\, (y' - h_k^C(z')))\ ,$$

à une constante multiplicative près. Le modèle obtenu à la question (iii) appartient donc à la classe des chaînes de Markov partiellement observées. Pour être complet

$$\begin{aligned}
&\mathbb{P}[Z_k \in dz' \mid Z_{k-1} = z, Y_{k-1} = y] \\
&\quad = \mathbb{P}[X_k \in dx', U_k \in du' \mid X_{k-1} = x, U_{k-1} = u, Y_{k-1} = y] \\
&\quad = \mathbb{P}[X_k \in dx' \mid U_k = u', X_{k-1} = x, U_{k-1} = u, Y_{k-1} = y] \\
&\qquad\qquad \mathbb{P}[U_k \in du' \mid X_{k-1} = x, U_{k-1} = u, Y_{k-1} = y] \\
&\quad = \mathbb{P}[X_k \in dx' \mid X_{k-1} = x]\; \delta_{A\,u + B\,B_0^{-1}\,(y - h_{k-1}(x) - C\,u)}(du') \; .
\end{aligned}$$

□

(v) **Donner l'équation vérifiée par le filtre bayésien $\mathbb{P}[Z_k \in dz \mid Y_{0:k}]$ pour le nouveau modèle. En déduire l'expression du filtre bayésien $\mathbb{P}[X_k \in dx \mid Y_{0:k}]$ pour le modèle de départ.**

Solution

On introduit le noyau positif non–normalisé

$$R_k(z, dz') = R_k(z, Y_{k-1}, Y_k, dz') \; ,$$

avec l'abus de notation usuel, pour tout $k = 1, \cdots, n.$. Le filtre bayésien pour le nouveau modèle, défini par

$$\mu_k(dz) = \mathbb{P}[Z_k \in dz \mid Y_{0:k}] \; ,$$

ou de manière équivalente par

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

vérifie donc l'équation récurrente

$$\mu_k = \frac{\mu_{k-1}\; R_k}{\langle \mu_{k-1}\; R_k, 1 \rangle} \; ,$$

et le filtre bayésien pour le modèle de départ est obtenu par marginalisation comme

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

□

Complément : modèle ARMA et représentation d'état

Quitte à poser $A_{p+1} = \cdots = A_q = 0$ au cas où $q \geq p+1$, ou bien $B_{q+1} = \cdots = B_p = 0$ au cas où $p \geq q+1$, on peut toujours se ramener, sans perte de généralité, au cas où $p = q = r$, avec

$$V_k + A_1\, V_{k-1} + \cdots + A_r\, V_{k-r} = B_0\, e_k + B_1\, e_{k-1} + \cdots + B_r\, e_{k-r} \; . \qquad (\bullet)$$

On va montrer que la suite $\{V_k\}$ définie par la relation

$$V_k = (\, I \quad 0 \quad \cdots \quad 0\,)\, U_k + B_0\; e_k \; ,$$

à l'aide d'une suite auxiliaire $\{U_k\}$ définie par

$$U_k = \begin{pmatrix} -A_1 & I & 0 & \cdots & \cdots & \cdots & 0 \\ -A_2 & 0 & I & 0 & & & \vdots \\ \vdots & & \ddots & \ddots & \ddots & & \vdots \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 0 & I & 0 \\ \vdots & & & & & 0 & I \\ -A_r & 0 & \cdots & \cdots & \cdots & \cdots & 0 \end{pmatrix} U_{k-1} + \begin{pmatrix} B_1 - A_1\; B_0 \\ B_2 - A_2\; B_0 \\ \vdots \\ \vdots \\ \vdots \\ \vdots \\ B_r - A_r\; B_0 \end{pmatrix} e_{k-1} \; ,$$

vérifie le modèle ARMA(r,r) introduit en $(\bullet)$. Ici, U_k est un vecteur de dimension rd et le vecteur U_k^i dénote le i–ème sous–bloc de dimension d, pour tout $i = 1, \cdots, r$. On a d'abord

$$\begin{aligned} V_k &= U_k^1 + B_0\; e_k \\ &= (-A_1\; U_{k-1}^1 + U_{k-1}^2 + (B_1 - A_1\; B_0)\; e_{k-1}) + B_0\; e_k \\ &= U_{k-1}^2 - A_1\; (U_{k-1}^1 + B_0\; e_{k-1}) + B_1\; e_{k-1} + B_0\; e_k \\ &= U_{k-1}^2 - A_1\; V_{k-1} + B_0\; e_k + B_1\; e_{k-1} \; , \end{aligned}$$

et on formule donc l'hypothèse de récurrence

$$V_k = U_{k-i}^{i+1} - (A_1\; V_{k-1} + \cdots + A_i\; V_{k-i}) + B_0\; e_k + \cdots + B_i\; e_{k-i} \; ,$$

au rang i. L'hypothèse est vérifiée pour $i = 0$ et pour $i = 1$, et on suppose qu'elle est vérifiée au rang i, soit

$$\begin{aligned}
V_k &= U^{i+1}_{k-i} - (A_1\, V_{k-1} + \cdots + A_i\, V_{k-i}) + B_0\, e_k + \cdots + B_i\, e_{k-i} \\
&= (-A_{i+1}\, U^1_{k-i-1} + U^{i+2}_{k-i-1} + (B_{i+1} - A_{i+1}\, B_0)\, e_{k-i-1}) \\
&\qquad\qquad - (A_1\, V_{k-1} + \cdots + A_i\, V_{k-i}) + B_0\, e_k + \cdots + B_i\, e_{k-i} \\
&= U^{i+2}_{k-i-1} - A_{i+1}\, (U^1_{k-i-1} + B_0\, e_{k-i-1}) + B_{i+1}\, e_{k-i-1} \\
&\qquad\qquad - (A_1\, V_{k-1} + \cdots + A_i\, V_{k-i}) + B_0\, e_k + \cdots + B_i\, e_{k-i} \\
&= U^{i+2}_{k-i-1} - (A_1\, V_{k-1} + \cdots + A_{i+1}\, V_{k-i-1}) + B_0\, e_k + \cdots + B_{i+1}\, e_{k-i-1} \ ,
\end{aligned}$$

c'est–à–dire que l'hypothèse est vérifiée au rang $(i+1)$. En particulier au rang $i = r - 1$, on obtient

$$\begin{aligned}
V_k &= U^r_{k-r+1} - (A_1\, V_{k-1} + \cdots + A_{r-1}\, V_{k-r+1}) + B_0\, e_k + \cdots + B_{r-1}\, e_{k-r+1} \\
&= (-A_r\, U^1_{k-r} + (B_r - A_r\, B_0)\, e_{k-r}) \\
&\qquad\qquad - (A_1\, V_{k-1} + \cdots + A_{r-1}\, V_{k-r+1}) + B_0\, e_k + \cdots + B_{r-1}\, e_{k-r+1} \\
&= -A_r\, (U^1_{k-r} + B_0\, e_{k-r}) + B_r\, e_{k-r} \\
&\qquad\qquad - (A_1\, V_{k-1} + \cdots + A_{r-1}\, V_{k-r+1}) + B_0\, e_k + \cdots + B_{r-1}\, e_{k-r+1} \\
&= -(A_1\, V_{k-1} + \cdots + A_r\, V_{k-r}) + B_0\, e_k + \cdots + B_r\, e_{k-r} \ ,
\end{aligned}$$

c'est–à–dire que

$$V_k + A_1\, V_{k-1} + \cdots + A_r\, V_{k-r} = B_0\, e_k + \cdots + B_r\, e_{k-r} \ ,$$

ou en d'autres termes, le modèle $(\bullet)$ est vérifié.

Exercice 2 :

L'objectif de cet exercice est d'établir les équations du *lisseur* bayésien, dans le cas simple des modèles de Markov cachés. Ici, l'horizon temporel est fixé, c'est–à–dire que toutes les observations recueillies entre l'instant initial 0 et l'instant final n sont disponibles et peuvent être utilisées pour estimer l'état caché à un instant k intermédiaire entre 0 et n (le cas où $k = n$ correspond bien sûr au cas du *filtre* bayésien vu en cours). Pour fixer les notations, on suppose que les états cachés $\{X_k\}$ forment une chaîne de Markov à valeurs dans E, caractérisée par la distribution de probabilité initiale $\eta_0(dx)$, et les noyaux de probabilités de transition $Q_k(x, dx')$ pour tout $k = 1, \cdots, n$, et que *conditionnellement aux états cachés* les observations $\{Y_k\}$ forment une suite de variables aléatoires indépendantes à valeurs dans F, caractérisée par la densité de probabilité d'émission $g_k(x, y)$, d'où l'expression de la *fonction de vraisemblance* $g_k(x) = g_k(x, Y_k)$ avec l'abus de notation usuel, pour tout $k = 0, 1, \cdots, n$.

On se propose donc d'étudier la distribution de probabilité conditionnelle de l'état caché X_k à l'instant intermédiaire k sachant les observations $Y_{0:n} = (Y_0, \cdots, Y_n)$ entre l'instant initial 0 et l'instant final n, soit

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

(i) **Montrer que le lisseur bayésien $\mu_{k|n}(dx)$ peut s'exprimer comme la distribution normalisée associée à une distribution non–normalisée $\gamma_{k|n}(dx)$ dont on donnera une représentation probabiliste, c'est–à–dire qu'on exprimera $\langle \gamma_{k|n}, \phi \rangle$ comme une espérance mathématique, pour toute fonction mesurable bornée ϕ. On pourra s'inspirer largement de la preuve vue en cours pour le filtre bayésien.**

Solution

La distribution de probabilité conditionnelle jointe des états cachés $X_{0:n}$ sachant les observations $Y_{0:n}$ est donnée par

$$\mathbb{P}[X_{0:n} \in dx_{0:n} \mid Y_{0:n}] = \frac{\displaystyle\prod_{p=0}^{n} g_p(x_p) \; \mathbb{P}[X_{0:n} \in dx_{0:n}]}{\displaystyle\mathbb{E}[\prod_{p=0}^{n} g_p(X_p)\,]} \ ,$$

où les fonctions de vraisemblance $g_0(x), \cdots, g_n(x)$ sont définies par abus de notation comme

$$g_p(x) = g_p(x, Y_p) \ ,$$

pour tout $p = 0, 1, \cdots, n$, et dépendent implicitement des observations $Y_{0:n}$, mais celles–ci sont considérées comme fixées dans toutes les expressions. En particulier, pour toute fonction test ϕ définie sur E

$$\mathbb{E}[\phi(X_k) \mid Y_{0:n}] = \frac{\displaystyle\int_E \cdots \int_E \phi(x_k) \prod_{p=0}^{n} g_p(x_p)\ \mathbb{P}[X_{0:n} \in dx_{0:n}]}{\displaystyle\mathbb{E}[\prod_{p=0}^{n} g_p(X_p)\,]} = \frac{\displaystyle\mathbb{E}[\,\phi(X_k) \prod_{p=0}^{n} g_p(X_p)\,]}{\displaystyle\mathbb{E}[\prod_{p=0}^{n} g_p(X_p)\,]}\ ,$$

de sorte que la distribution de probabilité conditionnelle de l'état caché X_k sachant $Y_{0:n}$ est finalement donné par

$$\langle \mu_{k|n}, \phi \rangle = \mathbb{E}[\phi(X_k) \mid Y_{0:n}] = \frac{\langle \gamma_{k|n}, \phi \rangle}{\langle \gamma_{k|n}, 1 \rangle}\ ,$$

où la distribution non–normalisée $\gamma_{k|n}(dx)$ est définie par

$$\langle \gamma_{k|n}, \phi \rangle = \mathbb{E}[\phi(X_k) \prod_{p=0}^{n} g_p(X_p)\,]\ .$$

$\square$

(ii) **En utilisant la propriété de Markov, montrer qu'il est possible d'exprimer $\langle \gamma_{k|n}, \phi \rangle = \langle \gamma_k, v_{k|n}\, \phi \rangle$ en fonction de la distribution non–normalisée $\gamma_k(dx)$ associée au filtre bayésien et d'une fonction $v_{k|n}(x)$ indépendante de ϕ, qu'on exprimera comme une espérance mathématique.**

Solution

Clairement, $\gamma_{n|n} = \gamma_n$ de sorte que la propriété est vérifiée pour $k = n$, avec la fonction $v_{n|n}(x) \equiv 1$. On considère ensuite le cas où $k \leq n-1$. En conditionnant par rapport à $X_{0:k} = (X_0, \cdots, X_k)$, puis en utilisant la propriété de Markov, on obtient

$$\begin{aligned}
\langle \gamma_{k|n}, \phi \rangle &= \mathbb{E}[\,\phi(X_k) \prod_{p=0}^{n} g_p(X_p)\,] \\
&= \mathbb{E}[\,\phi(X_k) \prod_{q=0}^{k} g_q(X_q) \prod_{p=k+1}^{n} g_p(X_p)\,] \\
&= \mathbb{E}[\,\phi(X_k) \prod_{q=0}^{k} g_q(X_q)\ \mathbb{E}[\prod_{p=k+1}^{n} g_p(X_p) \mid X_{0:k}\,]\,] \\
&= \mathbb{E}[\,\phi(X_k) \prod_{q=0}^{k} g_q(X_q)\ \mathbb{E}[\prod_{p=k+1}^{n} g_p(X_p) \mid X_k\,]\,] \\
&= \mathbb{E}[\,\phi(X_k)\ v_{k|n}(X_k) \prod_{q=0}^{k} g_q(X_q)\,] = \langle \gamma_k, v_{k|n}\, \phi \rangle\ ,
\end{aligned}$$

où la fonction $v_{k|n}$ est définie par

$$v_{k|n}(x) = \mathbb{E}[\prod_{p=k+1}^{n} g_p(X_p) \mid X_k = x] .$$

$\square$

(iii) **Établir une relation de récurrence rétrograde, partant de $k = n$, pour la fonction $v_{k|n}$, c'est–à–dire permettant d'exprimer $v_{k-1|n}$ en fonction de $v_{k|n}$.**

Solution

On remarque que $v_{k|n}(a)$ peut s'exprimer comme

$$g_k(a)\ v_{k|n}(a) = \mathbb{E}[\prod_{p=k}^{n} g_p(X_p) \mid X_k = a] = \mathbb{E}[\prod_{p=k}^{n} g_p(X_p^a)]$$

pour la chaîne de Markov $\{X_p^a\}$ définie entre les instants k et n, caractérisée par la distribution de probabilité initiale $\delta_a(dx)$ à l'instant k et par les noyaux de probabilités de transition $Q_p(x, dx')$ pour tout $p = k+1, \cdots, n$. On reconnait la constante de normalisation $\langle \gamma_n^a, 1 \rangle$ associée à la distribution non–normalisée $\gamma_n^a(dx)$ donnée par

$$\langle \gamma_\ell^a, \phi \rangle = \mathbb{E}[\phi(X_n^a) \prod_{p=k}^{\ell} g_p(X_p)] ,$$

pour toute fonction test ϕ définie sur E et pour tout $\ell = k, \cdots, n$. La relation de récurrence $\gamma_p^a = \gamma_{p-1}^a\ R_p$ pour les distributions non–normalisées, définie en terme du noyau positif non–normalisé $R_p(x, dx') = Q_p(x, dx')\ g_p(x')$ pour tout $p = k+1, \cdots, n$, donne $\gamma_n^a = \gamma_k^a\ R_{k+1:n}$ par itération, et clairement $\gamma_k^a = g_k(a)\ \delta_a$, de sorte que

$$g_k(a)\ v_{k|n}(a) = \langle \gamma_n^a, 1 \rangle = \langle \gamma_k^a\ R_{k+1:n}, 1 \rangle = \langle \gamma_k^a, R_{k+1:n}\ 1 \rangle = g_k(a)\ R_{k+1:n}\ 1(a) ,$$

et on en déduit que $v_{k|n}(a) = R_{k+1:n}\ 1(a)$. On obtient alors facilement la relation de récurrence souhaitée

$$\begin{aligned} v_{k-1|n}(x) = R_{k:n}\ 1(x) &= \int_E R_k(x, dx')\ R_{k+1:n}\ 1(x') \\ &= \int_E R_k(x, dx')\ v_{k|n}(x') \\ &= \int_E Q_k(x, dx')\ g_k(x')\ v_{k|n}(x') , \end{aligned}$$

pour tout $k = 1, \cdots, n-1$, avec la condition initiale

$$v_{n-1|n}(x) = \mathbb{E}[\, g_n(X_n) \mid X_{n-1} = x\,] = \int_E Q_n(x, dx')\; g_n(x')\; v_{n|n}(x')\ ,$$

pour $k = n-1$, compte tenu que $v_{n|n}(x) \equiv 1$, c'est–à–dire que finalement

$$v_{k-1|n} = Q_k\ (g_k\, v_{k|n})\ ,$$

pour tout $k = 1, \cdots, n$.

$\square$