École Nationale Supérieure de Techniques Avancées
examen du cours SOD333
“Filtrage bayésien et approximation particulaire”

vendredi 23 octobre 2020, 13:30 à 16:00

—— addendum ——

[...] Pour fixer les idées, on considère un modèle de Markov caché, où les états cachés $\{X_k\}$ forment une chaîne de Markov à valeurs dans E, et où *conditionnellement aux états cachés* les observations $\{Y_k\}$ forment une suite de variables aléatoires indépendantes (hypothèse de canal sans mémoire). Ce modèle est caractérisé par

- la distribution de probabilité initiale $\eta_0(dx)$ à l'instant 0,
- le noyau de probabilités de transition $Q_k(x, dx')$, c'est–à–dire que
$$Q_k(x, dx') = \mathbb{P}[X_k \in dx' \mid X_{k-1} = x] \;,$$
pour tout $k = 1, \cdots, n$,
- la fonction de vraisemblance $g_k(x')$ pour tout $k = 0, 1, \cdots, n$.

[...] Plus généralement, on introduit la fonction
$$f_k(x) = \mathbb{E}[\, \phi(X_n) \prod_{q=k+1}^{n} g_q(X_q) \mid X_k = x\,] \;, \tag{3}$$
pour tout $k = 0, 1, \cdots, n-1$, avec la convention $f_n = \phi$ pour $k = n$. Pour $x \in E$ fixé, cette expression peut s'interpréter comme une espérance pour le modèle de Markov caché caractérisé par

- la distribution de probabilité initiale δ_x à l'instant k,
- le noyau de probabilités de transition $Q_p(x, dx')$, pour tout $p = k+1, \cdots, n$,
- la fonction de vraisemblance $g_p(x')$ pour tout $p = k+1, \cdots, n$.

(iii) **Montrer que la suite de fonctions $\{f_k\}$ définie en (3) vérifie l'équation récurrente dans le sens rétrograde**
$$f_{k-1} = Q_k(g_k\, f_k) \;, \tag{4}$$
pour tout $k = n, n-1, \cdots, 1$, avec la condition initiale $f_n = \phi$ pour $k = n$.

La définition (3) exprimée pour $k = n - 1$ donne

$$f_{n-1}(x) = \mathbb{E}[\, \phi(X_n)\, g_n(X_n)) \mid X_{n-1} = x\,] = Q_n(g_n\, \phi)(x) = Q_n(g_n\, f_n)(x) \ ,$$

c'est–à–dire que l'équation (4) est vérifiée pour $k = n$.

Solution proposée dans le corrigé

On rappelle que, d'après la propriété de Markov, si la variable aléatoire Z est mesurable par rapport aux états futurs $X_{k:n}$, alors

$$\mathbb{E}[Z \mid X_{k-1}] = \mathbb{E}[\, \mathbb{E}[Z \mid X_{0:k}] \mid X_{k-1}] = \mathbb{E}[\, \mathbb{E}[Z \mid X_k] \mid X_{k-1}]$$

d'où on déduit que

$$\mathbb{E}[Z \mid X_{k-1} = x] = \mathbb{E}[\, \mathbb{E}[Z \mid X_k] \mid X_{k-1} = x] \qquad \kappa\text{–presque partout,}$$

où $\kappa(dx)$ désigne la distribution de probabilité de la variable aléatoire X_{k-1}.

Par ailleurs, on remarque que

$$g_k(x)\, f_k(x) = \mathbb{E}[\phi(X_n)\, \prod_{q=k}^{n} g_q(X_q) \mid X_k = x] \ ,$$

de sorte que

$$g_k(X_k)\, f_k(X_k) = \mathbb{E}[\phi(X_n)\, \prod_{q=k}^{n} g_q(X_q) \mid X_k] \ ,$$

et par définition

$$\begin{aligned} Q_k(g_k\, f_k)(x) &= \mathbb{E}[g_k(X_k)\, f_k(X_k) \mid X_{k-1} = x] \\ &= \mathbb{E}[\, \mathbb{E}[\phi(X_n)\, \prod_{q=k}^{n} g_q(X_q) \mid X_k] \mid X_{k-1} = x] \\ &= \mathbb{E}[\phi(X_n)\, \prod_{q=k}^{n} g_q(X_q) \mid X_{k-1} = x] \\ &= f_{k-1}(x) \qquad \kappa\text{–presque partout.} \end{aligned}$$

$\square$

On rappelle que $\mathbb{E}[Z \mid X' = x', X = x]$ est toute fonction mesurable définie sur $E \times E$ telle que

$$\begin{aligned}
\mathbb{E}[Z\ 1_{((X', X) \in A)}] &= \int\int_A \mathbb{E}[Z \mid X' = x', X = x]\ \mathbb{P}[X' \in dx', X \in dx] \\
&= \int\int_A \mathbb{E}[Z \mid X' = x', X = x]\ \mathbb{P}[X' \in dx' \mid X = x]\ \mathbb{P}[X \in dx]\ ,
\end{aligned}$$

pour tout borélien $A \subset E \times E$, et en particulier pour $A = E \times B$ on obtient

$$\begin{aligned}
\mathbb{E}[Z\ 1_{(X \in B)}] &= \int\int_{E\times B} \mathbb{E}[Z \mid X' = x', X = x]\ \mathbb{P}[X' \in dx' \mid X = x]\ \mathbb{P}[X \in dx] \\
&= \int_B [\int_E \mathbb{E}[Z \mid X' = x', X = x]\ \mathbb{P}[X' \in dx' \mid X = x]\,]\ \mathbb{P}[X \in dx]\ .
\end{aligned}$$

de sorte que

$$\mathbb{E}[Z \mid X = x] = \int_E \mathbb{E}[Z \mid X' = x', X = x]\ \mathbb{P}[X' \in dx' \mid X = x] \qquad \kappa\text{–presque partout,}$$

où $\kappa(dx)$ désigne la distribution de probabilité de la variable aléatoire X_{k-1}. On en déduit que

$$\begin{aligned}
f_{k-1}(x) &= \mathbb{E}[\phi(X_n)\ \prod_{p=k}^{n} g_p(X_p) \mid X_{k-1} = x] \\
&= \int_E \mathbb{E}[\phi(X_n)\ \prod_{p=k}^{n} g_p(X_p) \mid X_k = x', X_{k-1} = x]\ \mathbb{P}[X_k \in dx' \mid X_{k-1} = x] \\
&= \int_E g_k(x')\ \mathbb{E}[\phi(X_n)\ \prod_{p=k+1}^{n} g_p(X_p) \mid X_k = x']\ Q_k(x, dx') \\
&= \int_E g_k(x')\ f_k(x')\ Q_k(x, dx') \\
&= Q_k(g_k\, f_k)(x) \qquad \kappa\text{–presque partout,}
\end{aligned}$$

en utilisant la propriété de Markov, c'est–à–dire que l'équation (4) est vérifiée pour tout $k = 1, \cdots, n$.

$\square$

Solution proposée par Jérôme Bonnin et Claude Hamou

Par définition

$$
\begin{aligned}
f_k(x_k) &= \mathbb{E}[\,\phi(X_n)\ \prod_{q=k+1}^{n} g_q(X_q)) \mid X_k = x_k\,] \\
&= \int_E \cdots \int_E \phi(x_n)\ \prod_{q=k+1}^{n} g_q(x_q)\ \mathbb{P}[X_{k+1:n} \in dx_{k+1:n} \mid X_k = x_k\,] \\
&= \int_E \cdots \int_E \phi(x_n)\ \prod_{q=k+1}^{n} g_q(x_q)\ \prod_{q=k+1}^{n} Q_q(x_{q-1}, dx_q)\ ,
\end{aligned}
$$

pour tout k. Il en résulte que

$$
\begin{aligned}
f_{k-1}(x_{k-1}) &= \int_E \cdots \int_E \phi(x_n) \prod_{q=k}^{n} g_q(x_q) \prod_{q=k}^{n} Q_q(x_{q-1}, dx_q) \\
&= \int_E \cdots \int_E \phi(x_n)\ \prod_{q=k+1}^{n} g_q(x_q)\ \prod_{q=k+1}^{n} Q_q(x_{q-1}, dx_q)\ g_k(x_k)\, Q_k(x_{k-1}, dx_k) \\
&= \int_E [\int_E \cdots \int_E \phi(x_n)\ \prod_{q=k+1}^{n} g_q(x_q)\ \prod_{q=k+1}^{n} Q_q(x_{q-1}, dx_q)\,]\ g_k(x_k)\, Q_k(x_{k-1}, dx_k) \\
&= \int_E f_k(x_k)\ g_k(x_k)\, Q_k(x_{k-1}, dx_k) \\
&= Q_k(g_k\, f_k)(x_{k-1})\ ,
\end{aligned}
$$

c'est–à–dire que l'équation (4) est vérifiée pour tout $k = 1, \cdots, n$.

□

Soit x et k fixés. Pour tout entier l compris entre k et n, on considère la distribution de Feynman–Kac non–normalisée $\gamma_l^{x,k}$ définie par

$$\langle \gamma_l^{x,k}, \phi \rangle = \mathbb{E}[\phi(X_l^{x,k}) \prod_{p=k}^{l} g_p(X_p^{x,k})] = g_k(x)\ \mathbb{E}[\phi(X_l) \prod_{p=k+1}^{l} g_p(X_p) \mid X_k = x] \ ,$$

pour toute fonction mesurable bornée ϕ. Ici $\{X_p^{x,k}\}$ désigne la chaîne de Markov partant de l'état x à l'instant k, et possédant le même de noyau de probabilité de transition que la chaîne de Markov $\{X_p\}$. En particulier pour $l = n$, on a

$$\langle \gamma_n^{x,k}, \phi \rangle = g_k(x)\ \mathbb{E}[\phi(X_n) \prod_{p=k+1}^{n} g_p(X_p) \mid X_k = x] = g_k(x)\ f_k(x) \ ,$$

et pour $l = k$, on a

$$\langle \gamma_k^{x,k}, \phi \rangle = g_k(x)\ \phi(x) \ .$$

Clairement, la suite $\{\gamma_l^{x,k}\}$ vérifie la relation récurrente

$$\gamma_l^{x,k} = \gamma_{l-1}^{x,k}\, R_l \ ,$$

dans le sens *direct*, et en itérant cette relation on obtient

$$\gamma_n^{x,k} = \gamma_k^{x,k}\, R_{k+1} \cdots R_n \ ,$$

de sorte que

$$\langle \gamma_n^{x,k}, \phi \rangle = \langle \gamma_k^{x,k}\, R_{k+1} \cdots R_n, \phi \rangle = \langle \gamma_k^{x,k}, R_{k+1} \cdots R_n\, \phi \rangle = g_k(x)\ R_{k+1} \cdots R_n\, \phi(x) \ .$$

En identifiant les deux expressions pour $\langle \gamma_n^{x,k}, \phi \rangle$, on obtient

$$g_k(x)\ f_k(x) = g_k(x)\ R_{k+1} \cdots R_n\, \phi(x) \ ,$$

de sorte que

$$f_k(x) = R_{k+1} \cdots R_n\, \phi(x) \ ,$$

pour tout $x \in E$, c'est–à–dire que

$$f_k = R_{k+1} \cdots R_n\, \phi \ .$$

En particulier

$$f_{k-1} = R_k \cdots R_n\, \phi = R_k\, R_{k+1} \cdots R_n\, \phi = R_k\, f_k = Q_k(g_k\, f_k) \ ,$$

c'est–à–dire que l'équation (4) est vérifiée pour tout $k = 1, \cdots, n$.

$\square$