École Nationale Supérieure de Techniques Avancées
module : Commande des Systèmes

examen du cours OROC–SC–FP
"Filtrage bayésien et approximation particulaire"

vendredi 27 octobre 2017, 13:30 à 15:30

Problème

L'objectif de ce problème est d'étudier une stratégie alternative d'approximation par échantillonnage pondéré d'une distribution de Gibbs–Boltzmann

$$\mu = g \cdot \eta = \frac{g\, \eta}{\langle \eta, g \rangle} = \frac{\gamma}{\langle \gamma, 1 \rangle} \qquad \text{soit} \qquad \langle \mu, \phi \rangle = \frac{\langle \eta, g\, \phi \rangle}{\langle \eta, g \rangle} = \frac{\langle \gamma, \phi \rangle}{\langle \gamma, 1 \rangle}$$

lorsque la fonction de pondération / sélection g est bornée supérieurement, positive (mais pas strictement positive), et d'intégrale non–nulle par rapport à la distribution de probabilité η, c'est–à–dire lorsque

$$g(x) \geq 0 \text{ pour tout } x \in E \qquad \text{et} \qquad \sup_{x \in E} g(x) = M < \infty \qquad \text{et} \qquad \langle \eta, g \rangle > 0 \ ,$$

et on souhaite en particulier se prémunir contre le risque d'une extinction du système de particules.

Pour tout entier N, on définit les approximations

$$\eta_N = S^N(\eta) = \frac{1}{N} \sum_{i=1}^N \delta_{\xi_i} \qquad \text{et} \qquad \mu_N = g \cdot \eta_N = \frac{\displaystyle\sum_{i=1}^N g(\xi_i)\; \delta_{\xi_i}}{\displaystyle\sum_{i=1}^N g(\xi_i)}$$

où $(\xi_1, \cdots, \xi_i, \cdots)$ est une suite de variables aléatoires i.i.d. générées selon la distribution de probabilité η. L'approximation μ_N n'est bien définie que si la constante de normalisation

$$\langle \eta_N, g \rangle = \frac{1}{N} \sum_{i=1}^N g(\xi_i) > 0$$

est non nulle (strictement positive).

Pour toute fonction ϕ mesurable bornée (et donc intégrable par rapport à la distribution de probabilité μ), on définit la fonction recentrée $\bar{\phi} = \phi - \langle \mu, \phi \rangle$, et pour tout entier N, on pose

$$A_N = \sum_{i=1}^{N} [g(\xi_i)\, \bar{\phi}(\xi_i) - \langle \eta, g\, \bar{\phi} \rangle] = \sum_{i=1}^{N} g(\xi_i)\, \bar{\phi}(\xi_i) \qquad \text{et} \qquad D_N = \sum_{i=1}^{N} g(\xi_i) \ .$$

(i) **Sous réserve que $\langle \eta_N, g \rangle \neq 0$, établir l'expression suivante**

$$\langle \mu_N - \mu, \phi \rangle = \frac{\langle \eta_N - \eta, g\, \bar{\phi} \rangle}{\langle \eta_N, g \rangle} = \frac{A_N}{D_N} \ ,$$

pour l'erreur d'approximation.

La difficulté bien sûr consiste à controler la constante de normalisation $\langle \eta_N, g \rangle$ et à garantir qu'elle soit bien strictement positive. Pour tout $H > 0$ (destiné à tendre vers l'infini), on introduit à cet effet la variable aléatoire

$$T = T_H = \inf\{N \in \mathbb{N}^* \ : \ \sum_{i=1}^{N} g(\xi_i) \geq H\, M\} \ ,$$

à valeur entière (on rappelle que M dénote le supremum (ou juste une borne) de la fonction g supposée bornée supérieurement).

(ii) **Montrer que la variable aléatoire $T = T_H$ à valeur entière est un temps d'arrêt pour la tribu engendrée par les variables** $(\xi_1, \cdots, \xi_i, \cdots)$

- **p.s. fini,**
- **et intégrable (de moyenne finie).**

On introduit les approximations η_T et μ_T et les variables A_T et D_T, où les sommes portent sur un nombre aléatoire de termes (il suffit de remplacer N par T dans les formules).

(iii) **Montrer que p.s. la constante de normalisation $\langle \eta_T, g \rangle$ est non nulle (strictement positive), et en déduire que l'approximation μ_T est bien définie.**

(iv) **Montrer l'encadrement suivant**

$$H\, M \leq D_T \leq (H + 1)\, M \ .$$

On rappelle les deux identités de Wald

Soit $(X_1, \cdots, X_i, \cdots)$ *des variables aléatoires i.i.d. et soit*

$$S_N = \sum_{i=1}^{N} X_i$$

la somme partielle définie pour tout entier N*. Soit* τ *un temps d'arrêt (pour la filtration engendrée par les variables aléatoires introduites ci–dessus) et intégrable, c'est–à–dire tel que* $\mathbb{E}[\tau] < \infty$*.*

(i) Si $m = \mathbb{E}[X_1] < \infty$*, alors*

$$\mathbb{E}[S_\tau] = m\, \mathbb{E}[\tau] \ .$$

(ii) Si en outre $\sigma^2 = \mathbb{E}|X_1 - m|^2 < \infty$*, alors*

$$\mathbb{E}|S_\tau - m\,\tau|^2 = \sigma^2\, \mathbb{E}[\tau] \ .$$

(v) **En utilisant la première identité de Wald, montrer que**

$$\mathbb{E}[T] \le (H+1)\, \frac{M}{\langle \eta, g \rangle} \ .$$

(vi) **En utilisant la deuxième identité de Wald, montrer que**

$$\mathbb{E}|A_T|^2 \le (H+1)\, \frac{M}{\langle \eta, g \rangle}\, \langle \eta, g^2\, \bar{\phi}^2 \rangle \ .$$

(vii) **En déduire que l'erreur d'approximation vérifie**

$$\mathbb{E}|\langle \mu_T - \mu, \phi \rangle|^2 = \mathbb{E}|\frac{A_T}{D_T}|^2 \le \frac{1}{H}\, \frac{H+1}{H}\, \frac{\langle \eta, g \rangle}{M}\, \frac{\langle \eta, g^2\, \bar{\phi}^2 \rangle}{\langle \eta, g \rangle^2} \ .$$

Sélection binaire

On s'intéresse à présent au cas particulier où la fonction de pondération / sélection g est binaire, c'est–à–dire à valeurs 0 ou 1, et on va montrer qu'il est possible d'obtenir des résultats plus précis voire explicites, quitte à modifier légèrement l'algorithme.

(viii) **En appliquant directement les résultats précédents, montrer que l'erreur d'approximation vérifie**

$$\mathbb{E}|\langle \mu_T - \mu, \phi \rangle|^2 = \mathbb{E}|\frac{A_T}{D_T}|^2 \leq \frac{1}{H} \, \frac{H+1}{H} \, \mathrm{var}(\phi, \mu) \ .$$

(ix) **Si H est un entier, montrer que le temps d'arrêt peut alors s'écrire**

$$T = T_H = \inf\{N \in \mathbb{N}^* \ : \ \sum_{i=1}^{N} g(\xi_i) = H\} \ ,$$

c'est–à–dire comme le nombre d'expériences nécessaires pour obtenir exactement H succès dans un tirage de Bernoulli où la probabilité de succès est $0 < p = \langle \eta, g \rangle \leq 1$.

(x) **Montrer que $D_T = H$ et $D_{T-1} = H - 1$.**

(xi) **En reprenant les étapes suivies aux questions (v) et (vi), montrer que l'erreur d'approximation vérifie**

$$\mathbb{E}|\langle \mu_T - \mu, \phi \rangle|^2 = \mathbb{E}|\frac{A_T}{D_T}|^2 = \frac{1}{H} \, \mathrm{var}(\phi, \mu) \ ,$$

exactement.

On rappelle que le nombre T d'expériences nécessaires pour obtenir exactement H succès dans un tirage de Bernoulli où la probabilité de succès est p, suit une loi binomiale inverse (ou loi de Pascal), caractérisée par

$$\mathbb{P}[T = N] = \binom{N-1}{H-1} \, p^H \, (1-p)^{N-H} \qquad \text{pour tout entier } N \geq H.$$

(xii) **Commenter la formule ci–dessus. Montrer que**

$$\mathbb{E}[\frac{1}{T-1}] = \frac{p}{H-1} \ .$$

On introduit l'approximation $\eta_T^\bullet = \eta_{T-1}$, où il suffit de remplacer N par $T-1$ (au lieu de T) dans les formules.

(xiii) **Montrer que la constante de normalisation vérifie**

$$\langle \eta_T^\bullet, g \rangle = \frac{H-1}{T-1} \ ,$$

et en déduire qu'il s'agit d'une approximation non–biaisée de la constante de normalisation $\langle \eta, g \rangle$.