PROBABILITÉS ET STATISTIQUES

L'objet du problème est l'étude d'un algorithme stochastique de minimisation d'une fonction H sur un ensemble E^S où S est de cardinal fini n (En pratique n est très grand, par exemple en traitement d'images $n \approx 10^6$). Cet algorithme dépend d'une suite de paramètres réels positifs (T_k).

Dans la première partie $E = \mathbb{R}$ et H est une forme quadratique définie positive ; elle est donc convexe et a un minimum unique. Les vecteurs aléatoires qui interviennent dans l'algorithme sont Gaussiens. L'algorithme converge pourvu que la suite (T_k) tende vers 0.

Dans la seconde partie E est fini et H quelconque. L'algorithme converge pourvu que la suite (T_k) tende vers 0 en décroissant et en restant minorée par une suite $\left(\frac{\gamma}{\ln k}\right)$.

Mise à part cette problématique commune, les deux parties sont indépendantes et peuvent être abordées dans l'ordre qui conviendra le mieux à chaque candidat. Celui-ci pourra admettre le résultat de certaines questions pour traiter les suivantes à condition de l'indiquer clairement.

PREMIERE PARTIE

Toutes les variables aléatoires intervenant dans cette partie sont définies sur un même espace probabilisé. On note $\mathcal{F}$ la tribu des parties probabilisables et E l'espérance mathématique, ou moyenne, ou intégrale par rapport à la probabilité.

Q.1. Soit $S = \{1, 2, \dots, n\}$ et soit $\{e_s\}_{s \in S}$ la base canonique de $\mathbb{R}^n$; on note de la même lettre une application linéaire de $\mathbb{R}^n$ dans $\mathbb{R}^n$ et la matrice (n,n) qui la représente sur cette base, un vecteur de $\mathbb{R}^n$ et la matrice colonne $(n,1)$ qui le représente ; I_n désigne l'application identique de $\mathbb{R}^n$ sur lui-même et ${}^tA$ la matrice transposée de la matrice A.

Un vecteur aléatoire à valeurs dans $\mathbb{R}^n$ et suivant une loi de Laplace-Gauss sera dit *Gaussien*. Il sera dit Gaussien *strict* si sa matrice des variances-covariances est régulière (pour abréger, on appellera celle-ci *covariance* et on la notera cov).

Soit $X = \sum_{s \in S} X_s e_s$ un vecteur Gaussien strict de $\mathbb{R}^n$ de moyenne $E(X) = m = \sum_{s \in S} m_s e_s$ et de covariance $\operatorname{cov}(X) = E((X-m)\,{}^t(X-m)) = \Gamma$.

On note γ l'inverse de Γ, γ_{sk} les éléments de γ, f la densité de X par rapport à la mesure de Lebesgue de support $\mathbb{R}^n$.

Pour tout s de S, on note P_s l'application de $\mathbb{R}^n$ dans $\mathbb{R}^n$ qui à x associe $x_s e_s$: $P_s x = x_s e_s$; et on pose :

$$x_{(s)} = (I_n - P_s)\, x,$$
$$X_{(s)} = (I_n - P_s)\, X.$$

1.1. Déterminer la loi de $X_{(s)}$; on précisera : sa moyenne,
sa covariance,
sa fonction caractéristique,

et sa densité notée g_s par rapport à la mesure de Lebesgue ayant pour support $(I_n - P_s)\,\mathbb{R}^n$.

1.2. Montrer que l'on a, pour tout x de $\mathbb{R}^n$:

$$f(x) = g_s(x_{(s)})\, h_s(x)$$

avec $h_s(x) = \sqrt{\frac{\gamma_{ss}}{2\pi}} \exp\left(-\frac{1}{2}\,{}^t(x-m)\,\gamma\, Q_s\, \gamma\,(x-m)\right)$ où $Q_s = \frac{1}{\gamma_{ss}} P_s$.

1.3. On dit qu'une fonction φ de $\mathbb{R}$ (respectivement $\mathbb{R}^n$) dans $\mathbb{C}$ est à *croissance lente* s'il existe un k de $\mathbb{N}$ tel que $\lambda \to \frac{|\varphi(\lambda)|}{(1+|\lambda|)^k}$ soit bornée sur $\mathbb{R}$ (respectivement $\lambda \to \frac{|\varphi(\lambda)|}{(1+\|\lambda\|)^k}$ bornée sur $\mathbb{R}^n$).

Tournez la page S.V.P.

A toute fonction φ de $\mathbb{R}$ dans $\mathbb{C}$ mesurable et à croissance lente, on associe la fonction Ψ_s de $(I_n - P_s)\,\mathbb{R}^n$ dans $\mathbb{C}$ définie par :

$$\Psi_s(\alpha) = \int_{\mathbb{R}} \varphi(\lambda)\, h_s(\lambda e_s + \alpha)\, d\lambda .$$

On note $\mathcal{F}_{(s)}$ la plus petite sous-tribu de $\mathcal{F}$ rendant mesurable $X_{(s)}$. Montrer que l'espérance conditionnelle de $\varphi(X_s)$ par rapport à cette tribu : $E(\varphi(X_s) \mid \mathcal{F}_{(s)})$ est définie et que $\Psi_s(X_{(s)})$ est un représentant de $E(\varphi(X_s) \mid \mathcal{F}_{(s)})$.

Q.2.

2.1. Pour α fixé dans $(I_n - P_s)\,\mathbb{R}^n$, montrer que la fonction de $\mathbb{R}$ dans $\mathbb{R}$:

$$\lambda \longrightarrow h_s(\lambda e_s + \alpha)$$

est une densité de probabilité Gaussienne par rapport à la mesure de Lebesgue de support $\mathbb{R}$.

Soit $Z_s(\alpha)$ une variable aléatoire réelle ayant cette densité.

2.2. Montrer que $E(Z_s(\alpha)\, e_s) = Q_s\, \gamma\, (m-\alpha)$.

2.3. Montrer que $\operatorname{cov}(Z_s(\alpha)\, e_s) = Q_s$.

2.4. Déterminer la fonction caractéristique de $Z_s(\alpha)$.

Q.3. Pour tout k entier strictement positif, on note s_k l'élément de S tel que : $s_k \equiv k$ (modulo n).

On considère une suite $(Y^k)_{k\geq 1}$ de vecteurs aléatoires à valeurs dans $\mathbb{R}^n$.

On note $\mathcal{F}^k$ la plus petite sous-tribu de $\mathcal{F}$ rendant mesurables $Y^1, Y^2, \ldots, Y^k$.

On pose $\alpha^k = (I_n - P_{s_k})\, Y^k$.

On suppose que

i) Y^1 est Gaussien de moyenne $m + \mu^1$, de covariance Λ^1 ;

ii) pour toute application φ mesurable et à croissance lente de $\mathbb{R}^n$ dans $\mathbb{C}$,

$$\int_{\mathbb{R}} \varphi(\lambda e_{s_k} + \alpha^k)\, h_{s_k}(\lambda e_{s_k} + \alpha^k)\, d\lambda$$

est un représentant de $E(\varphi(Y^{k+1}) / \mathcal{F}^k)$.

3.1. u étant un vecteur de $\mathbb{R}^n$, exprimer $E(\exp(i\, {}^t u\, Y^{k+1}) / \mathcal{F}^k)$ en fonction de Y^k.

3.2. En déduire que, pour tout k, Y^k est un vecteur gaussien à valeurs dans $\mathbb{R}^n$.

Calculer $\mu^k = E(Y^k - m)$ et $E((Y^k - m)\, {}^t(Y^k - m))$.

3.3. On pose, pour tout x de $\mathbb{R}^n$:

$$\|x\|_\gamma = \sqrt{{}^t x\, \gamma x} .$$

Montrer que $\|\mu^k\|_\gamma$ est une fonction décroissante de k.

3.4. Montrer que le produit :

$(I_n - Q_n \gamma) \ldots (I_n - Q_1 \gamma)$ est une application linéaire strictement contractante pour $\|\ \|_\gamma$.

En déduire que $\lim\limits_{k\to\infty} E(Y^k) = m$, puis que la suite $(Y^k)_{k\geq 1}$ converge en loi vers X.

3.5. Quelle est la loi de Y^k dans le cas particulier :

$\mu^1 = 0$, $\Lambda^1 = \Gamma$?

Q.4. On se donne une suite $(T_k)_{k\geq 1}$ de réels strictement positifs telle que :

$$\lim_{k\ \infty} T_k = 0 .$$

On reprend l'algorithme de la question 3 en y remplaçant la fonction h_{s_k} par la fonction :

$$x \longrightarrow h_k^1(x) = \sqrt{\frac{\gamma_{s_k s_k}}{\pi\, T_k}}\ \exp\left(-\frac{1}{T_k}\ {}^t(x-m)\, \gamma\, Q_{s_k}\, \gamma\, (x-m)\right)$$

Tournez la page S.V.P.

et on suppose maintenant que, pour tout $k \geqslant 1$,

$$\int_{\mathbb{R}} \varphi(\lambda e_{s_k} + \alpha^k)\, h_k^1(\lambda e_{s_k} + \alpha^k)\, d\lambda$$

est un représentant de $E\left(\varphi(Y^{k+1}) / \mathcal{F}_k\right)$.

Montrer que la suite $(Y_k)_{k \geqslant 1}$ converge en probabilité vers m.

DEUXIEME PARTIE

Dans cette partie, Ω est un ensemble fini et toutes les probabilités sont définies sur $\mathcal{P}(\Omega)$.

Etant données deux probabilités μ et ν, on pose

$$\|\mu - \nu\| = \operatorname*{Max}_{A \subset \Omega} |\mu(A) - \nu(A)| \;;$$

on rappelle que

$$\|\mu - \nu\| = \sum_{\omega \in \Omega} (\mu(\{\omega\}) - \nu(\{\omega\}))^+$$

$$= \frac{1}{2} \sum_{\omega \in \Omega} |\mu(\{\omega\}) - \nu(\{\omega\})|$$

où $x^+ = \operatorname{Max}(x, 0)$.

Q.5. On appelle *transition* sur Ω une fonction P de $\Omega \times \mathcal{P}(\Omega)$ dans $[0,1]$ telle que, pour tout ω de Ω, la fonction $A \to P(\omega, A)$ soit une probabilité ; on note P_ω cette probabilité.

On pose alors $\Delta(P) = \operatorname*{Max}_{\omega, \omega'} \|P_\omega - P_{\omega'}\|$.

Etant données une probabilité μ et une transition P, on note μP la probabilité définie par

$$\mu P = \sum_{\omega} \mu(\{\omega\})\, P_\omega$$

et on dit que μ est *invariante par* P si $\mu P = \mu$.

Etant données deux transitions P et Q, on note PQ la transition définie par :

$$PQ(\omega, A) = \sum_{\omega' \in \Omega} P(\omega, \{\omega'\})\, Q(\omega', A).$$

5.1. Montrer que, pour toute transition P, il existe une fonction non identiquement nulle, α, de Ω dans $\mathbb{R}$ telle que, pour tout ω de Ω :

$$\alpha(\omega) = \sum_{\omega' \in \Omega} \alpha(\omega')\, P(\omega', \{\omega\}).$$

Montrer que $|\alpha|$ satisfait la même équation et en déduire qu'il existe une probabilité μ invariante par P.

5.2. Montrer que quelles que soient les deux probabilités μ et ν et les deux transitions P et Q, on a :

$$\|\mu P - \nu P\| \leqslant \|\mu - \nu\|\, \Delta(P)$$

et $\quad \Delta(PQ) \leqslant \Delta(P)\, \Delta(Q)$.

5.3. Montrer que :

$$\Delta(P) = 1 - \min_{\omega, \omega'} \sum_{\omega''} \min(P(\omega, \{\omega''\}), P(\omega', \{\omega''\})).$$

Q.6. On considère une suite $(P^k)_{k \in \mathbb{N}}$ de transitions sur Ω et on définit la suite $(P^{k,\ell})_{\ell \geqslant k+1}$ par :

$P^{k,k+1} = P^{k+1}$ et, pour $\ell \geqslant k+2$,

$P^{k,\ell} = P^{k,\ell-1}\, P^\ell$.

On dit que la suite $(P^k)_{k \in \mathbb{N}}$ est *faiblement ergodique* si, pour tout k de $\mathbb{N}$,

$$\lim_{\ell \infty} \Delta(P^{k,\ell}) = 0.$$

On dit que la suite $(P^k)_{k \in \mathbb{N}}$ est *fortement ergodique* s'il existe une probabilité ν, telle que, pour toute probabilité μ et pour tout k dans $\mathbb{N}$, on ait :

$$\lim_{\ell \infty} \|\mu P^{k,\ell} - \nu\| = 0.$$

6.1. Montrer que $(P^k)_{k \in \mathbb{N}}$ est faiblement ergodique si et seulement si, quelles que soient les deux probabilités μ et ν et l'entier k, on a

$$\lim_{\ell \infty} ||\mu P^{k,\ell} - \nu P^{k,\ell}|| = 0 .$$

6.2. Soit, pour tout k dans $\mathbb{N}$, ν_k une probabilité invariante par P^k. Montrer que, si $\sum_k ||\nu_{k+1} - \nu_k||$ converge, et si $(P^k)_{k \in \mathbb{N}}$ est faiblement ergodique, alors $(P^k)_{k \in \mathbb{N}}$ est fortement ergodique.

Q.7. Dans cette question, Ω est l'ensemble des applications d'un ensemble S de cardinal fini n dans un ensemble fini E : $\Omega = E^S$; on note $(\omega_s)_{s \in S}$ un élément de Ω.

A toute fonction H de Ω dans $\mathbb{R}$, et à tout réel strictement positif T, on associe la probabilité $\mu_{H,T}$ définie par

$$\mu_{H,T}(\{\omega\}) = C_{H,T} \exp\left(-\frac{H(\omega)}{T}\right)$$

avec
$$C_{H,T} = \frac{1}{\sum_{\omega \in \Omega} \exp\left(-\frac{H(\omega)}{T}\right)} .$$

Pour chaque s de S, on associe à tout ω de Ω sa restriction $\omega_{(s)}$ à $S \setminus \{s\}$ (complémentaire de $\{s\}$ dans S) :

$$\omega_{(s)} = (\omega_j)_{j \in S \setminus \{s\}} ,$$

et la fonction H_s définie sur $E \times E^{S \setminus \{s\}}$ par :

$$H_s(\omega_s, \omega_{(s)}) = H(\omega).$$

On associe à H_s la transition π_s^T sur Ω définie par :

$$\pi_s^T(\omega, \{\omega'\}) = \begin{cases} C_{s,T} \exp\left(-\dfrac{H_s(\omega'_s, \omega_{(s)})}{T}\right) & \text{si } \omega_{(s)} = \omega'_{(s)} \\ 0 & \text{si } \omega_{(s)} \neq \omega'_{(s)} \end{cases}$$

où $C_{s,T}$ ne dépend pas de ω'_s.

Pour une énumération $\{s_1, s_2, \ldots, s_n\}$ de S, on pose

$$P^T = \pi_{s_1}^T \pi_{s_2}^T \ldots \pi_{s_n}^T$$

7.1. Montrer que $\mu_{H,T}$ est invariante par P^T.

7.2. On note Ω_o le sous-ensemble de Ω sur lequel H est minimum : ω_o appartient à Ω_o si et seulement si $H(\omega_o) = \min_{\omega \in \Omega} H(\omega)$.

Soit μ_o la probabilité uniforme sur Ω_o ; montrer que :

$$\lim_{T \downarrow 0} ||\mu_{H,T} - \mu_o|| = 0 .$$

7.3. Montrer qu'il existe un nombre réel $\delta \geqslant 0$, ne dépendant que de H, tel que $\Delta(P^T) \leqslant 1 - \exp\left(-\frac{n\delta}{T}\right)$.

7.4. En déduire l'existence d'un nombre $\gamma > 0$ tel que, pour toute suite décroissante $(T_k)_{k \in \mathbb{N}}$ de réels > 0 satisfaisant

$\lim_{k \infty} T_k = 0$ et, pour tout $k \geqslant 2$, $T_k \geqslant \frac{\gamma}{\ln k}$,

la suite $(P^{T_k})_{k \in \mathbb{N}}$ est fortement ergodique.

- FIN -