Inégalités de concentration
Sorbonne Université
2020-2021

Examen

1. Une amélioration de l'inégalité d'Azuma–Hoeffding

Soit $(Z_i)_{i=0}^n$ une martingale adaptée à la filtration $(\mathcal{F}_i)_{i=0}^n$, telle que pour tout $i \in [\![1, n]\!]$, on a $a_i \leq Z_i - Z_{i-1} \leq b_i$. On sait par l'inégalité d'Azuma–Hoeffding que pour tout $t \geq 0$,

$$\mathbf{P}\left(Z_n - Z_0 \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

Le but de cet exercice est de montrer que l'on a même

$$\mathbf{P}\left(\max_{0\leq k\leq n}(Z_k - Z_0) \geq t\right) \leq \exp\left(-\frac{2t^2}{\sum_{i=1}^n (b_i - a_i)^2}\right).$$

(1) (Inégalité de Doob–Kolmogorov) Montrer que si $(Y_k)_{k=0}^n$ est une sous-martingale positive, alors pour tout $a > 0$,

$$\mathbf{P}\left(\max_{0\leq k\leq n} Y_k \geq a\right) \leq \frac{\mathbf{E}[Y_n]}{a}.$$

On pourra introduire le temps d'arrêt $\tau = \inf\{s \geq 0,\, Y_s \geq a\} \wedge n$ (avec la convention $\inf \emptyset = +\infty$) et utiliser le fait que $\mathbf{E}[Y_\tau] \leq \mathbf{E}[Y_n]$, par le théorème d'arrêt de Doob.

(2) Montrer que, pour tout $\lambda \in \mathbb{R}$, la suite $\left(e^{\lambda(Z_k - Z_0)}\right)_{k=0}^n$ est une sous-martingale adaptée à la filtration $(\mathcal{F}_k)_{k=0}^n$.

(3) Conclure.

2. Percolation de premier passage

Soit $G = (V, E)$ un graphe fini connexe. On munit chaque arête $e \in E$ d'un poids aléatoire X_e. Les variables $X_e, e \in E$ sont indépendantes et possède le même second moment $\mathbf{E}X_e^2 = \sigma^2$. Soient $u, v \in V$ deux sommets fixés. On s'intéresse au poids minimal d'un chemin allant de u à v (le poids d'un chemin correspondant à la somme des poids des arêtes le long de ce chemin), soit

$$Z = \min_{P\in\mathcal{P}_{u,v}} \sum_{e\in P} X_e\,,$$

où $\mathcal{P}_{u,v}$ est l'ensemble des chemins allant de u à v.

(1) Montrer que

$$\mathbf{Var}(Z) \leq \sigma^2 \mathbf{E}[L_\star]\,,$$

où $L_\star$ est la longueur (i.e. le nombre d'arêtes) d'un chemin optimal (i.e. d'un chemin qui réalise le minimum dans la définition de Z).

(2) On suppose que pour tout $e \in E$, on a $0 \leq X_e \leq 1$. Soit $f : [0,1]^{|E|} \to \mathbb{R}_+$ définie par

$$\forall x \in [0,1]^{|E|}\,,\ f(x) = \min_{P\in\mathcal{P}_{u,v}} \sum_{e\in P} x_e = \min_{t\in S} \sum_{e\in E} t_e x_e\,,$$

où S est l'ensemble des vecteurs $t = (t_e)_{e\in E}$ avec $t_e = \mathbb{1}_{e\in P}$, pour P parcourant $\mathcal{P}_{u,v}$. Montrer que pour tout $t \geq 0$, on a

$$\mathbf{P}\left(Z - \mathbf{E}Z \geq t\right) \leq \exp\left(-\frac{t^2}{2\mathbf{E}[L_\star]}\right),$$

et

$$\mathbf{P}\left(Z - \mathbf{E}Z \leq -t\right) \leq \exp\left(-\frac{t^2}{2d_{u,v}}\right),$$

où $d_{u,v}$ correspond à la longueur du plus long chemin entre u et v.

3. Matrice de covariance empirique

Soient $X_1, \dots, X_n$ des vecteurs de $\mathbb{R}^d$ ($d \geq 1$), i.i.d. de même loi que X, avec X d'espérance nulle et de matrice de covariance $\Sigma = \mathbf{E}[X\,{}^tX]$. Un estimateur naturel de Σ est donné par la matrice de covariance empirique :

$$\widehat{\Sigma} = \frac{1}{n}\sum_{i=1}^{n} X_i\,{}^tX_i\,.$$

Dans ce qui suit, $\|\cdot\|$ désigne la norme euclidienne, $\|\cdot\|_{\mathsf{op}}$ désigne la norme d'opérateur, et $\mathbb{S}^{d-1}$ désigne la sphère de $\mathbb{R}^d$, c'est-à-dire l'ensemble des vecteurs $x \in \mathbb{R}^d$ tels que $\|x\| = 1$. On rappelle que pour une matrice $A \in \mathcal{M}_{d,d}(\mathbb{R})$ symétrique, $\|A\|_{\mathsf{op}} = \max_{z\in\mathbb{S}^{d-1}} |\,{}^tzAz|$.

(1) On suppose dans cette question qu'il existe $c > 0$ tel que $\|X\| \leq c$.

(a) Montrer que $\|X\,{}^tX - \Sigma\|_{\mathsf{op}} \leq 2c^2$ et que $\left\|\mathbf{E}\left[(X\,{}^tX - \Sigma)^2\right]\right\|_{\mathsf{op}} \leq c^2\|\Sigma\|_{\mathsf{op}}$.

(b) En déduire que pour tout $t \geq 0$,

$$\mathbf{P}\left(\|\widehat{\Sigma} - \Sigma\|_{\mathsf{op}} \geq t\right) \leq 2d\exp\left(-\frac{nt^2}{2c^2\left(\|\Sigma\|_{\mathsf{op}} + \frac{2t}{3}\right)}\right).$$

(2) On relâche maintenant l'hypothèse sur X en supposant seulement que X est sous-gaussienne avec facteur de variance σ^2, au sens où pour tout $u \in \mathbb{S}^{d-1}$, la variable ${}^tuX$ est sous-gaussienne avec facteur de variance σ^2, i.e. pour tout $\lambda \in \mathbb{R}$, $\log \mathbf{E}e^{\lambda\,{}^tuX} \leq \frac{\sigma^2\lambda^2}{2}$.

(a) Soit $v \in \mathbb{S}^{d-1}$ et soit $Y = {}^tvX$. Montrer que pour tout entier $k \geq 1$, on a $\mathbf{E}[Y^{2k}] \leq 2^{k+1}\sigma^{2k}k!$, et en déduire que pour tout $\lambda \in \mathbb{R}$,

$$\log \mathbf{E}e^{\lambda(Y^2 - \mathbf{E}Y^2)} \leq \begin{cases} 8\sigma^4\lambda^2 & \text{si } \lambda \leq 0, \\ \frac{8\sigma^4\lambda^2}{1-2\sigma^2\lambda} & \text{si } 0 \leq \lambda < \frac{1}{2\sigma^2}. \end{cases}$$

(b) Montrer que pour tout $v \in \mathbb{S}^{d-1}$ et pour tout $t \geq 0$, on a

$$\mathbf{P}\left(\left|\,{}^tv(\widehat{\Sigma} - \Sigma)v\right| \geq t\right) \leq 2\exp\left(-\frac{nt^2}{4(8\sigma^4 + \sigma^2 t)}\right).$$

(c) Pour $\delta > 0$, un δ-recouvrement de $\mathbb{S}^{d-1}$ est un sous-ensemble fini $S_\delta \subset \mathbb{S}^{d-1}$ tel que pour tout $z \in \mathbb{S}^{d-1}$, il existe $u \in S_\delta$ avec $\|z - u\| \leq \delta$. Montrer que pour tout $\delta \in]0, 1/2[$, on a

$$\|\widehat{\Sigma} - \Sigma\|_{\mathsf{op}} \leq \frac{1}{1-2\delta}\max_{u\in S_\delta} |\,{}^tu(\widehat{\Sigma} - \Sigma)u|\,.$$

(d) On choisit maintenant $\delta = 1/4$ et l'on admet qu'il existe un $1/4$-recouvrement S de $\mathbb{S}^{d-1}$ tel que $|S| \leq 9^d$. Montrer que pour tout $t \geq 0$,

$$\mathbf{P}\left(\|\widehat{\Sigma} - \Sigma\|_{\mathsf{op}} \geq t\right) \leq 2\cdot 9^d\exp\left(-\frac{nt^2}{8(16\sigma^4 + \sigma^2 t)}\right).$$

4. Méthode entropique et inégalité de Harris

(A) Inégalité de Harris. Soient $f, g : \mathbb{R}^n \to \mathbb{R}$ deux fonctions croissantes (en chacune de leurs coordonnées) et soit $X = (X_1, \dots, X_n)$ une suite de variables aléatoires réelles indépendantes. Montrer que

$$\mathbf{E}\left[f(X)g(X)\right] \geq \mathbf{E}\left[f(X)\right]\mathbf{E}\left[g(X)\right]\,.$$

(On pourra raisonner par récurrence.)

(B) Soit $f : \mathbb{R}^n \to \mathbb{R}$ et $Z = f(X_1, \dots, X_n)$ avec $X = (X_1, \dots, X_n)$ une suite de variables aléatoires réelles indépendantes.

(1) En utilisant la sous-additivité de l'entropie, montrer que pour tout $\lambda \in \mathbb{R}$,

$$\mathbf{Ent}\left[e^{\lambda Z}\right] \leq \sum_{i=1}^{n}\mathbf{E}\left[e^{\lambda Z}\phi\left(\lambda(Z_i' - Z)\right)\right],$$

avec $Z_i' = f(X_1, \dots, X_{i-1}, X_i', X_{i+1}, \dots, X_n)$ où $(X_i')_{i=1}^n$ est une copie indépendante de $(X_i)_{i=1}^n$, et $\phi(u) = e^u - u - 1$.

(2) En écrivant $Z_i' - Z = (Z_i' - Z)_+ - (Z - Z_i')_+$ et le fait que Z et Z_i' sont échangeables, montrer que pour tout $\lambda \in \mathbb{R}$,
$$\mathbf{E}\left[e^{\lambda Z}\phi\left(\lambda(Z_i' - Z)\right)\right] = \mathbf{E}\left[e^{\lambda Z}\tau\left(\lambda(Z_i' - Z)_+\right)\right] ,$$
où $\tau(u) = u(e^u - 1)$.

(3) En déduire que pour tout $\lambda \le 0$, $\mathbf{Ent}\left[e^{\lambda Z}\right] \le \sum_{i=1}^n \mathbf{E}\left[\lambda^2 e^{\lambda Z}(Z_i' - Z)_+^2\right]$.

(4) On suppose que f est croissante en chacune de ses coordonnées, et qu'il existe une fonction $g : \mathbb{R}^n \to \mathbb{R}$, elle aussi croissante en chacune de ses coordonnées, telle que
$$\mathbf{E}\left[\sum_{i=1}^n (Z_i' - Z)_+^2 \mid X\right] \le g(X)\,.$$
Montrer que pour tout $t \ge 0$, on a
$$\mathbf{P}\left(Z - \mathbf{E}Z \le -t\right) \le \exp\left(-\frac{t^2}{4\mathbf{E}[g(X)]}\right) .$$