

Une inégalité de concentration à gauche pour les processus empiriques

Thierry KLEIN
Université de versailles
45 avenue des Etats-Unis
78035 Versailles cedex
klein@math.uvsq.fr

Résumé

Nous donnons des constantes dans l'inégalité de concentration à gauche de Talagrand pour les processus empiriques indexés par des classes de fonctions, en partant de la méthode de Herbst. Le point nouveau est que la constante du facteur variance est exacte, ce qui répond à une conjecture de Massart.

Abstract

We give new constants in Talagrand's left concentration inequality for maxima of empirical processes. Our approach is based on the Herbst method. The improvement we get concerns the constant in the variance factor, which is the one conjectured by Massart.

1 Introduction

Soit $X_1, X_2, \dots$ une suite de variables aléatoires indépendantes de loi commune P à valeur dans $\mathcal{X}$ espace polonais muni de sa tribu borélienne. Soit $\mathcal{F}$ une famille dénombrable de fonctions mesurables de $\mathcal{X}$ dans $\mathbb{R}$, de carré intégrable sous P . On pose

$$S_n(f) = f(X_1) + \dots + f(X_n). \quad (1)$$

Nous regardons les propriétés de concentration de la variable aléatoire $Z = \sup\{S_n(f) : f \in \mathcal{F}\}$ autour de sa moyenne. Ce problème a été étudié dans une série de travaux successifs, en particulier par Talagrand [6] au moyen d'inégalités isopérimétriques pour les mesures produits. Il a obtenu une inégalité de type Bennett pour la déviation de Z par rapport à sa moyenne. Ensuite Ledoux [2] a montré que ces inégalités pouvaient aussi être obtenues à partir d'inégalités de type log-Sobolev pour les mesures produit. Cette technique permet de majorer la transformée de Laplace de Z . Massart [3] a clairement montré l'intérêt des méthodes entropiques (ou log-Sobolev) pour le calcul des constantes dans les inégalités de Talagrand, enfin Rio [5] a obtenu une inégalité de concentration à droite de type Bernstein avec un facteur variance asymptotiquement exact. Nous allons montrer dans cette note un analogue de l'inégalité de Rio [5] pour la concentration de Z à gauche de sa moyenne. Bousquet [1] a amélioré la fonction de taux, et a obtenu une inégalité de Bennett.

2 Résultat

Théorème 1 Soit $\mathcal{F}$ classe dénombrable de fonctions de $\mathcal{X}$ dans $] -\infty, 1]$, mesurables, de carré intégrable et d'espérance nulle sous P . Soit σ^2 telle que $\sigma^2 \geq P(f^2)$ pour toute f dans $\mathcal{F}$.

- Si les fonctions de $\mathcal{F}$ sont à valeurs dans $[-1, 1]$ alors pour tout $x > 0$,

$$P(Z \leq E(Z) - x) \leq \exp \left(-\frac{v_n}{16} h\left(\frac{4x}{v_n}\right) \right), \quad (2)$$

avec $v_n = n\sigma^2 + 2E(Z)$ et $h(x) = (1+x)\log(1+x) - x$.

- Si pour tout $f \in \mathcal{F}$, et tout $p \geq 2$, $|E(f^p(X_i))| \leq \frac{\sigma^2 p!}{2}$ alors

$$\forall t \in [0, 1[\quad L(t) \leq -te^{-t}E(Z) + \frac{n\sigma^2}{2} \frac{t^2(1+2t)}{1-t}, \quad (3)$$

et par conséquent pour tout x positif,

$$P(Z \leq E(Z) - x) \leq \exp \left(- (u-v)^2 \left(1 - 2\frac{v(u-v)}{u^2}\right) \right), \quad (4)$$

où $L(t)$ est la log-Laplace de $-Z$, $u = \sqrt{x + v_n/2}$, $v = \sqrt{v_n/2}$.

Remarque 1 La fonction de taux de $v_n t^2 (2(1-t))^{-1}$ est $(u-v)^2$, et nous avons en plus une correction multiplicative comprise entre $1/2$ et 1 tendant vers 1 quand xv_n^{-1} tend vers 0. Asymptotiquement on retrouve donc la fonction de taux des inégalités gaussiennes.

3 Preuve

3.1 Lemme préparatoire

Nous rappelons ici le Lemme 2.1 de Rio[5] (voir aussi Massart [3], lemme 8).

Lemme 1 Soient $(X_1, \dots, X_n)$ des variables aléatoires indépendantes à valeurs dans un espace polonais $\mathcal{X}$. Soit $\mathcal{F}_n^k$ la tribu engendrée par $(X_i)_{i \neq k}$. Posons $\phi(x) = x \log(x) - x + 1$. Soit h une variable aléatoire fonction mesurable de $(X_1, \dots, X_n)$, strictement positive et intégrable. alors, pour toute famille (h_k) de variables aléatoires réelles strictement positives et intégrables, respectivement $\mathcal{F}_n^k$ -mesurables,

$$E(h \log h) - E(h) \log E(h) \leq \sum_{k=1}^n E(h_k \phi(h/h_k)). \quad (5)$$

3.2 Preuve

Comme dans Rio [5] on peut supposer que $\mathcal{F}$ est finie et procéder ensuite par passage à la limite. On définit $\mathcal{F}_n^k$ la tribu engendrée par $(X_i)_{i \neq k}$, $Z_k = \sup\{S_n^k(f) : f \in \mathcal{F}\}$. Puisque $\mathcal{F}$ est finie, on peut opérer une sélection $\mathcal{F}_n^k$ -mesurable d'une fonction f_k telle que $Z_k = S_n^k(f_k)$, $Z'_k = S_n(f_k)$. On définit $\psi(x) = 1 - (1+x)e^{-x}$, $\eta_k = f_k(X_k)$. Soit $t > 0$, en appliquant (5) à $h = e^{-tZ}$ et $h_k = e^{-tZ_k}$, on obtient l'inégalité suivante pour la transformée de Laplace F de $-Z$.

$$tF'(t) - F(t) \log F(t) \leq \sum_{k=1}^n E(e^{-tZ_k} \psi(t(Z - Z_k))). \quad (6)$$

On note alors que

$$\psi(tx) = \frac{t^2}{1-t} x e^{-tx} + \left(1 - e^{-tx} \left(\frac{1-t+tx}{1-t}\right)\right) := q(t) x e^{-tx} + r(t, x), \quad (7)$$

si $q(t) = t^2/(1-t)$. Par conséquent (6) s'écrit encore

$$tF'(t) - F(t) \log F(t) \leq \sum_{k=1}^n E(e^{-tZ} \frac{t^2}{1-t} (Z - Z_k)) + \sum_{k=1}^n E(e^{-tZ_k} r(t, Z - Z_k)),$$

Comme $\sum_{k=1}^n (Z - Z_k) \leq Z$ et $q(t) > 0$, on en déduit que

$$tF'(t) - F(t) \log F(t) \leq -q(t)F'(t) + \sum_{k=1}^n E\left(e^{-tZ_k} r(t, Z - Z_k)\right). \quad (8)$$

Puisque

$$\frac{\partial r}{\partial x}(t, x) = \frac{t^2}{1-t} e^{-tx} (x - 1),$$

la fonction $r(t, x)$ est décroissante en x pour $0 \leq t < 1$ et $x \leq 1$.

Puisque $\eta_k = f_k(X_k) \leq (Z - Z_k) \leq 1$ et r est décroissante en x ,

$$tF' - F \log F \leq -q(t)F' + \frac{1}{1-t} \sum_{k=1}^n E\left(e^{-tZ_k} \left(1 - t - e^{-t\eta_k} (1 - t + t\eta_k)\right)\right). \quad (9)$$

Posons $S_k = 1 - t - e^{-t\eta_k} (1 - t + t\eta_k)$. L'équation (9) devient alors pour $0 \leq t < 1$,

$$tF'(t) - (1-t)F(t) \log F(t) \leq \sum_{k=1}^n E\left(e^{-tZ_k} E_n^k(S_k)\right), \quad (10)$$

où E_n^k est l'espérance conditionnelle par rapport à $\mathcal{F}_n^k$.

• **Sous la première hypothèse** : montrons à l'aide d'un développement en série entière de $\bar{S}_k$ que :

$$S_k \leq -t^2\eta_k + \eta_k^2(e^t - 1 - t + t^3). \quad (11)$$

Preuve de (11).

On pose pour la preuve $x = \eta_k$. On a alors, pour $t \in [0, 1]$,

$$\begin{aligned} S_k &= (1-t) - (1-t+tx)(1-tx+t^2x^2/2) - (1-t) \sum_{j=3}^{+\infty} (-1)^j x^j t^j / j! \\ &\quad - tx \sum_{j=3}^{+\infty} (-1)^j x^j t^j / j! , \\ &= -t^2x + t^2x^2/2 + t^3 \frac{x^2}{2} (1-x) + (1-t)x^2 \sum_{j=3}^{+\infty} (-1)^{j+1} x^{j-2} t^j / j! \\ &\quad + tx^2 \sum_{j=3}^{+\infty} (-1)^{j+1} x^{j-1} t^j / j! . \end{aligned}$$

Comme x est dans $[-1, 1]$, on a la majoration suivante :

$$\begin{aligned} S_k &\leq -t^2x + t^2x^2/2 + t^3x^2 + (1-t)x^2 \left(\frac{e^t - e^{-t}}{2} - t \right) + tx^2 \left(e^t - 1 - t - \frac{t^2}{2} \right), \\ &\leq -t^2x + x^2 \left(\frac{t^2}{2} + \frac{t^3}{2} - 2t + e^t \left(\frac{1+t}{2} \right) - e^{-t} \left(\frac{1-t}{2} \right) \right). \end{aligned}$$

Notons maintenant

$$\begin{aligned} \alpha(t, x) &= -t^2x + x^2 \left(\frac{t^2}{2} + \frac{t^3}{2} - 2t + e^t \left(\frac{1+t}{2} \right) - e^{-t} \left(\frac{1-t}{2} \right) \right), \\ \beta(t, x) &= -t^2x + x^2 \left(e^t - 1 - t + t^3 \right), \\ \gamma(t, x) &= \beta(t, x) - \alpha(t, x), \end{aligned}$$

et montrons que $\gamma \geq 0$. On remarque que le signe de γ est le même que celui de

$$\delta(t) = e^t(1-t) + e^{-t}(1-t) - 2 + 2t + t^3 - t^2.$$

comme $\delta(0) = 0$, $\delta'(t)e^t = -te^{-2t} + (2 + 3t^2 - 2t)e^t + (t-2)$, $\delta'(0) = 0$ et $\delta'(1) > 0$ et comme la fonction de $q(y) = -ty^2 + (2 - 2t + 3t^2)y + (t-2)$ est concave on en déduit que δ est positive, ce qui achève la preuve de (11).

Par (2.4) dans Rio [5] on sait que $E_n^k(\eta_k) = 0$ et $E_n^k(\eta_k^2) \leq \sigma^2$. On en déduit la majoration suivante :

$$E_n^k(S_k) \leq \sigma^2(e^t - 1 - t + t^3). \quad (12)$$

En remarquant que $e^{-tZ_k} \leq e^t E^{-tZ}$, on obtient l'inéquation différentielle suivante valable pour tout $0 \leq t < 1$:

$$tL'(t) - (1-t)L(t) \leq ne^t \sigma^2 (e^t - 1 - t + t^3). \quad (13)$$

Montrons maintenant que la fonction

$$G(t) = n\sigma^2 \frac{e^{4t} - 1 - 4t}{16} - te^{-t}E(Z) \quad (14)$$

est une sursolution de (13) sur $[0, 1]$. Soit

$$\begin{aligned} f(t) &= tG'(t) - (1-t)G(t) \\ &= \frac{n\sigma^2}{16} \left(4t(e^{4t} - 1) - (1-t)(e^{4t} - 1 - 4t) \right). \end{aligned}$$

Il s'agit donc de montrer que, pour tout t dans $[0, 1]$

$$f(t) \geq n\sigma^2 e^t (e^t - 1 - t + t^3). \quad (15)$$

Cela revient à montrer que , pour tout t dans $[0, 1]$

$$D(t) = e^{4t}(5t - 1) + 1 - t - 4t^2 - 16e^{2t} + 16e^t(1 + t - t^3) \geq 0.$$

Nous avons les formules suivantes pour les dérivées de D .

$$\begin{aligned} D' &= e^{4t}(2t + 1) - 1 - 8t - 32e^{2t} + 16e^t(2 + t - 3t^2 - t^3), \\ D^{(2)} &= 8e^{4t}(10t + 3) - 8 - 64e^{2t} + 16e^t(3 - 5t - 6t^2 - t^3), \\ D^{(3)} &= 16e^{4t}(5t + 11) - 128e^{2t} - 16e^t(2 + 17t - 9t^2 - t^3). \end{aligned}$$

Montrons que $D^{(3)} \geq 0$. Soit

$$\begin{aligned} R(t) &= \frac{D^{(3)}(t)}{8e^t}, \\ R(t) &= e^{3t}(11 + 5t) - 8e^t - (2 + 17t - 9t^2 - t^3), \\ R'(t) &= e^{3t}(38 + 15t) - 8e^t - (17 - 18t - 3t^2) > 0. \end{aligned}$$

Donc R est croissante sur $[0, 1]$. Or $R(0) = 1$ donc $R(t) > 0$ sur $[0, 1]$ et $D^{(3)} \geq 0$ sur $[0, 1]$. On en déduit alors facilement que $D(t) > 0$. Ce qui entraîne que $G(t)$ est une sursolution de (13).

Comme de plus $G(0) = L(0)$ et $G'(0) = L'(0)$, on en déduit que pour tout $t \in [0, 1]$ on a

$$L(t) \leq G(t). \quad (16)$$

Comme sur $[0, 1]$ on a $t(1 - e^{-t}) \leq 2\frac{e^{4t} - 1 - 4t}{16}$, on en déduit que pour $t \in [0, 1]$:

$$P(Z \leq E(Z) - x) \leq \exp \left(v_n \frac{e^{4t} - 1 - 4t}{16} - tx \right). \quad (17)$$

Preuve de (2) pour $x < v_n(e^4 - 1)/4$.

L'expression $\exp\left(v_n \frac{e^{4t} - 1 - 4t}{16} - tx\right)$ est optimisée pour t_0 vérifiant

$$4t_0 = \ln\left(1 + \frac{4x}{v_n}\right).$$

En réinjectant t_0 dans (17) on obtient (2).

Preuve de (2) pour $x \geq v_n(e^4 - 1)/4$.

Pour obtenir (2) quand $x \geq v_n(e^4 - 1)/4$ il suffit de remarquer que pour tout $f \in \mathcal{F}$ on a

$$P(Z \leq E(Z) - x) \leq P(S_n(f) \leq E(Z) - x), \quad (18)$$

et que le membre de droite de (18) peut être majoré d'après l'inégalité de Bennett classique, par

$$P(S_n(f) \leq E(Z) - x) \leq \exp\left(-n\sigma^2 h\left(\frac{x - E(Z)}{n\sigma^2}\right)\right). \quad (19)$$

On conclut en remarquant que pour $x \geq v_n(e^4 - 1)/4$, puisque $v_n/2 \geq E(Z)$,

$$x - E(Z) \geq x\left(1 - \frac{2}{e^4 - 1}\right) \geq \frac{24x}{25}.$$

Donc, comme h est croissante, (19) devient

$$P(S_n(f) \leq E(Z) - x) \leq \exp\left(-n\sigma^2 h\left(\frac{24x}{25n\sigma^2}\right)\right). \quad (20)$$

Or le majorant de (20) est à nouveau inférieur à la borne de (2).

• **Sous la deuxième hypothèse** : comme $\eta_k = f_k(X_k)$ avec f_k $\mathcal{F}_n^k$ -mesurable on a $|E_n^k(\eta_k^p)| \leq \sigma^2 p!/2$. Montrons alors, à l'aide d'un développement en série entière de S_k que :

$$E_n^k(S_k) \leq \sigma^2 \frac{t^2}{2(1-t)^2} (1 + 2t - 3t^2 + t^3). \quad (21)$$

Preuve de (21).

Posons $x = \eta_k$. Alors

$$S_k = -t^2 x + t^3 \frac{x^2}{2} (1 - x) + (1 - t) \sum_{i=3}^{+\infty} (-1)^{i+1} x^i \frac{t^i}{i!} + t \sum_{i=3}^{+\infty} (-1)^{i+1} x^{i+1} \frac{t^i}{i!},$$

et par conséquent

$$\begin{aligned} E_n^k(S_k) &\leq \frac{\sigma^2}{2} \left(3t^3 + t^2 + (1-t) \sum_{i=3}^{+\infty} t^i + t \sum_{i=3}^{+\infty} (i+1)t^i \right), \\ &\leq \frac{\sigma^2 t^2}{2(1-t)^2} (1 + 2t - 3t^2 + t^3). \end{aligned}$$

ce qui prouve (21).

On obtient alors l'inéquation différentielle suivante, valable pour tout $0 \leq t < 1$:

$$tL'(t) - (1-t)L(t) \leq \frac{n\sigma^2 t^2 e^t}{2(1-t)^2} (1 + 2t - 3t^2 + t^3). \quad (22)$$

Cette équation peut s'écrire

$$\left(\frac{L(t)}{te^{-t}} \right)' \leq \frac{n\sigma^2 e^{2t}}{2(1-t)^2} (1 + 2t - 3t^2 + t^3). \quad (23)$$

Définissons la fonction G par

$$G(t) = \frac{n\sigma^2 e^t (t + 2t^2)}{2(1-t)}.$$

Alors

$$G'(t) = \frac{n\sigma^2 e^t (1 + 5t - t^2 - 2t^3)}{2(1-t)^2}.$$

Nous allons montrer que pour tout t dans $[0, 1]$,

$$G'(t) \geq \frac{n\sigma^2 e^{2t}}{2(1-t)^2} (1 + 2t - 3t^2 + t^3).$$

Pour cela il suffit de montrer que

$$D(t) = 1 + 5t - t^2 - 2t^3 - e^t (1 + 2t - 3t^2 + t^3) \geq 0. \quad (\mathcal{H})$$

On a

$$\begin{aligned} D^{(3)}(t) &= -12 - e^t (-5 + 2t + 6t^2 + t^3) \leq 0, \\ D^{(2)}(t) &= -2 - 12t - e^t (-1 - 4t + 3t^2 + t^3), \quad D^{(2)}(0) = -1, \\ D'(t) &= 5 - 2t - 6t^2 - e^t (3 - 4t + t^3). \end{aligned}$$

On en déduit que D' est décroissante. $(\mathcal{H})$ est donc une conséquence immédiate de la positivité de la fonction concave D aux bords de l'intervalle. Après majoration et intégration de (23) on obtient :

$$L(t) \leq -te^{-t}E(Z) + \frac{n\sigma^2}{2} \frac{t^2(1+2t)}{(1-t)} \leq -tE(Z) + \frac{v_n}{2} \frac{t^2(1+2t)}{1-t}. \quad (24)$$

Par le calcul usuel de Cramer-Chernov on obtient

$$P(Z \leq E(Z) - x) \leq \exp \left(\frac{v_n}{2} \frac{t^2(1+2t)}{(1-t)} - tx \right). \quad (25)$$

Pour obtenir une borne on prend alors le réel t^* minimisant $\frac{1}{2}v_n t^2(1-t)^{-1} - tx$. Après calcul on trouve $t^* = 1 - w_n^{\frac{1}{2}}(x + w_n)^{-\frac{1}{2}}$, où $w_n = v_n/2$ que l'on réinjecte dans (25), pour obtenir (4) . $\square$

Références

- [1] Olivier Bousquet. A Bennett concentration inequality and its application to suprema of empirical processes. *C. R. Acad. Sci. Paris, Ser I* 334 2002. 495-500.
- [2] Michel Ledoux. On Talagrand's deviation inequalities for product measures. *ESAIM probability and statistics*, 1996. 1 :63-87.
- [3] Pascal Massart. About the constant in Talagrand's concentration inequalities for empirical processes. *Annals of Probab.* 2000. 28 :863-884.
- [4] Emmanuel Rio. Théorie asymptotique des processus aléatoires faiblement dépendants. J.M. Ghidaglia, X. Guyon (Ed.). *Mathématiques et Applications* 31. Springer. 2000.
- [5] Emmanuel Rio" Une inégalité de Bernstein pour les maxima de processus empiriques. *Preprint 57 de l'Université de Versailles-Saint-Quentin*. 2001.
- [6] Michel Talagrand, New concentration inequalities in product spaces. *Invent. Math.* 1996. 126 :503-563.