

Sur l'intégrabilité des vecteurs gaussiens

Michel Talagrand

Equipe d'Analyse Université Paris VI, Tour 46, 4, Place Jussieu, F-75230 Paris Cedex 05,
France

Summary. Let T be any set, and let $(X_t)_{t \in T}$ be a separable gaussian process with mean zero on T . Assume that X is almost surely bounded, and let $N = \sup_{t \in T} |X_t|$.

Let $\sigma = \sup_{t \in T} (EX_t^2)^{1/2}$, and let

$$\tau = \sup \{ \lambda \geq 0; P\{N < \lambda\} = 0 \}.$$

Let $\beta > \tau/\sigma^2$. We prove that

$$E \left(\exp \left(\frac{1}{2\sigma^2} N^2 - \beta N \right) \right) < +\infty.$$

Examples are given to show that this result cannot be readily improved. When $\tau = 0$, we also show that the distribution of N has a continuous density with respect to Lebesgue measure.

1. Introduction

Soient E un espace vectoriel, et (f_n) une suite de formes linéaires sur E . Soit \mathcal{B} la tribu sur E engendrée par les f_n . Soit μ une mesure gaussienne sur (E, \mathcal{B}) , c'est-à-dire une probabilité sur (E, \mathcal{B}) telle que la loi de toute famille finie de f_n soit gaussienne centrée. Posons $N(x) = \sup_n |f_n(x)|$. On suppose $N(x) < +\infty$ μ .p.s. Soient

$$\tau = \sup \{ \lambda; \mu(\{x; N(x) < \lambda\}) = 0 \}, \quad \sigma = \sup_n (E f_n^2)^{1/2}.$$

Un théorème célèbre de Fernique [3] montre que pour $\alpha < \frac{1}{2\sigma^2}$, on a $E(\exp \alpha N^2) < +\infty$. Nous allons préciser ce résultat.

Théorème 1. Pour $\beta > \tau/\sigma^2$, on a

$$E \left(\exp \left(\frac{N^2}{2\sigma^2} - \beta N \right) \right) < +\infty.$$

Un cas particulier de ce résultat a été démontré par M. Weber [5] (c'est ce qui nous a suggéré le résultat général).

Théorème 2. a) *Il existe un exemple de la situation précédente où*

$$E \left(\exp \left(\frac{1}{2\sigma^2} N^2 - \frac{\tau}{\sigma^2} N \right) \right) = +\infty.$$

b) *Il existe un exemple de la situation précédente où $\tau > 0$ mais où*
 $E \left(\exp \left(\frac{1}{2\sigma^2} N^2 - \varepsilon N \right) \right) < +\infty$ *pour $\varepsilon > 0$.*

Posons

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-u^2/2} du.$$

Un résultat remarquable de A. Ehrhard [2] affirme que la fonction

$$h(t) = \Phi^{-1}(\mu(\{x; N(x) \leq t\}))$$

est concave sur $]\tau, \infty[$. Notre résultat peut s'interpréter par la formule

$$l = \lim_t h(t) - t\sigma \geq -\tau/\sigma.$$

On a bien sur $l \leq 0$.

La concavité de h montre que $h(t) \leq t\sigma + l$. Il est naturel de se demander à quelle vitesse h s'approche de son asymptote. La réponse est que cela peut être arbitrairement lentement.

Théorème 3. *Soit $\varphi: [0, \infty[\rightarrow \mathbb{R}$ décroissante avec $\varphi(t) \rightarrow 0$. Il existe alors un exemple de la situation précédente où $\tau = l = 0$, $\sigma = 1$ et $h(t) \leq t - \varphi(t)$ pour tout t assez grand.*

Remarque. Dans le cas où la condition d'entropie utilisée par Weber est vérifiée, les calculs de celui-ci donnent des résultats plus précis que la méthode générale exposée ici, et donnent en particulier une très bonne estimation de $h(t) - t\sigma$ pour $t \rightarrow \infty$.

Théorème 4. *Si $\tau = 0$, la distribution de N possède une densité continue par rapport à la mesure de Lebesgue.*

2. Preuve du théorème 1

Réduction. On applique le procédé d'orthonormalisation de Graham-Schmidt à la suite (f_n) , considérée comme suite de $L^2(\mu)$. On peut écrire $f_n = \sum_{i \leq n} a_{i,n} e_n$, où e_n est une forme linéaire mesurable sur (E, \mathcal{B}) , et où la suite (e_n) est orthonormale dans $L^2(\mu)$. Considérons l'application $\theta: x \rightarrow (e_n(x))$ de E dans $\mathbb{R}^{\mathbb{N}}$. La mesure image est la gaussienne canonique sur $\mathbb{R}^{\mathbb{N}}$, c'est-à-dire la mesure produit quand chaque facteur est muni de la loi normale. On est donc ramené à ce cas, avec de plus chaque f_n dans $\mathbb{R}^{(\mathbb{N})}$. On peut donc prendre pour \mathcal{B} la tribu borélienne.

Lemme 1. Pour $x \in \mathbb{R}^{\mathbb{N}}$, soit $R_n(x)$ l'élément de l^∞ obtenu à partir de x en remplaçant les n premières coordonnées par des zéros. Alors on a

$$\mu \cdot \text{p.s.} \quad \lim N(R_n(x)) = \tau.$$

Preuve. Première étape. On montre que pour tout n , on a $N(R_n(x)) \geq \tau$ p.s. Il en résulte que $\liminf N(R_n(x)) \geq \tau$ $\mu \cdot \text{p.s.}$ Soit $\tau' < \tau$ et soit $A = \{x \in \mathbb{R}^{\mathbb{N}}; N(R_n(x)) < \tau'\}$. Soit

$$B = \{x \in \mathbb{R}^{\mathbb{N}}; \forall i \leq n, |x_i| \leq (\tau - \tau')/n\sigma\}.$$

Alors $\mu(B) > 0$. De plus puisque chaque f_m est de norme $\leq \sigma$ dans $L^2(\mu)$, on a $|a_{i,m}| \leq \sigma$ pour chaque i, m . Il en résulte que pour $x \in B$, on a $N(x - R_n(x)) \leq \tau - \tau'$. Ainsi pour $x \in A \cap B$, on a $N(x) < \tau$. Par définition de τ , ceci montre que $\mu(A \cap B) = 0$. D'autre part, μ étant une mesure produit, on a $\mu(A \cap B) = \mu(A)\mu(B)$, et $\mu(B) > 0$, d'où $\mu(A) = 0$. Ceci prouve le résultat.

Deuxième étape. On montre que si $\tau' > \tau$ et si $\varepsilon > 0$, il existe un n tel que $\mu(L) \geq 1 - \varepsilon$, où

$$L = \{x \in \mathbb{R}^{\mathbb{N}}; N(R_n(x)) \leq \tau'\}.$$

Soit

$$D = \{x; N(x) \leq \tau'\}.$$

Par hypothèse on a $\mu(D) > 0$. Puisque $N(x) = \text{Sup} |f_n(x)|$, il existe un entier m tel que si

$$D' = \{x; \forall i \leq m, |f_i(x)| \leq \tau'\}$$

on a $\mu(D')(1 - \varepsilon/2) \leq \mu(D)$. Puisque chaque f_i ne dépend que d'un nombre fini de coordonnées on peut écrire $D' = M \times F_n$ où $M \subset \mathbb{R}^n$ et $F_n = \mathbb{R}^{l_n, \infty}$. On désigne par λ_n, ν_n les gaussiennes canoniques de \mathbb{R}^n, F_n . On a $\mu = \lambda_n \times \nu_n$. Le théorème de Fubini montre qu'il existe $x \in \mathbb{R}^n$ tel que si on pose

$$G = \{y \in F_n; N(x, y) \leq \tau'\},$$

on a $\nu_n(G) \geq 1 - \varepsilon/2$.

Par symétrie on a $\nu_n(-G) \geq 1 - \varepsilon/2$, et ainsi $\nu_n(G \cap -G) \geq 1 - \varepsilon$. Or, pour $y \in G \cap -G$, on a $N(x, y) \leq \tau'$, et $N(x, y) = N(-x, y) \leq \tau'$, d'où $N(0, y) \leq \tau'$. Ceci montre que $\mathbb{R}^n \times (G \cap -G) \subset L$, d'où $\mu(L) \geq 1 - \varepsilon$.

Troisième étape. Le théorème de Fernique montre que $EN(x) < +\infty$. Si \mathcal{B}_n est la σ -algèbre engendrée par les coordonnées de rang $\geq n+1$ la suite $(N(R_n(x)))_n$ est une sous-martingale par rapport à la suite décroissante (\mathcal{B}_n) . Elle converge donc, et les deux étapes précédentes montrent que sa limite est p.s. égale à τ .

Remarque. Le lemme 1 possède un intérêt en soi. Toutefois, pour la suite, nous n'utiliserons que la deuxième étape. En particulier, notre preuve n'utilisera pas le théorème de Fernique.

Lemme 2. Soient B un convexe borélien symétrique de $\mathbb{R}^{\mathbb{N}}$ et K la boule unité de l^2 . Supposons que $\nu_n(H) \geq 1/2$, où $H = \{y \in F_n; (0, y) \in B\}$. Alors pour tout $t \geq 0$ on a :

$$\mu(\{z, z \notin B + tK\}) \leq P(\chi^2 > t^2)$$

où χ^2 a une loi chi 2 à $n+1$ degrés de liberté.

Preuve. Soit K' la boule unité de $l^2([n, \infty[)$. Le point crucial est l'inégalité de C. Borell [1], qui affirme que $v_n(H + uK') \geq \Phi(u)$ pour $u \geq 0$. Pour $x \in \mathbb{R}^n$, et $y \in F_n$, $y \in H + uK'$, on a $(x, y) \in B - tK$, si $u^2 + \|x\|_2^2 \leq t^2$. Ainsi, pour $t \geq \|x\|_2$,

$$\begin{aligned} v_n(\{y \in F_n; (x, y) \notin B + tK\}) \\ \leq v_n(\{y \in F_n; y \notin B' + (t^2 - \|x\|_2^2)^{1/2} K'\}) \leq 1 - \Phi((t^2 - \|x\|_2^2)^{1/2}). \end{aligned}$$

En appliquant le théorème de Fubini on a donc

$$\mu(\{z; z \notin B + tK\}) \leq \lambda_n(\{x; \|x\|_2 \geq t\}) + \frac{1}{\sqrt{2\pi}} \int_A e^{-s^2/2} ds d\lambda_n(x)$$

où

$$A = \{(x, s) \in \mathbb{R}^{n+1}; \|x\|_2 \leq t, \|x\|_2^2 + s^2 \geq t^2\}$$

et ainsi

$$\mu(\{z; z \notin B + tK\}) \leq \frac{1}{\sqrt{2\pi}} \int_{s^2 + \|x\|_2^2 \geq t^2} e^{-s^2/2} ds d\lambda_n(x)$$

ce qui est le résultat cherché.

Lemme 3. Pour tout $\tau' > \tau$, il existe un entier n tel que

$$\forall t \geq 0, \quad \mu(\{z; N(z) > t + \tau'\}) \leq P(\sigma^2 \chi^2 > t^2)$$

où χ^2 est une loi chi 2 à n degrés de liberté.

Preuve. Avec les notations de la deuxième étape du lemme 1, on peut choisir n tel que $v_n(H) \geq 1/2$, où

$$H = \{y \in F_2; N(0, y) \leq \tau'\}.$$

Soit $B = \{z; N(z) \leq 1\}$. Puisque chaque f_n est de norme au plus σ dans l^2 pour $z \in K$, on a $N(z) = \sup_n |f_n(z)| \leq \sigma$, d'où $K \subset \sigma B$. Il résulte alors du lemme 2 que:

$$\mu(\{z; z \notin \tau' B + \sigma u B\}) \leq P(\chi^2 < u^2)$$

ce qui est le résultat cherché.

Le théorème 1 est maintenant une simple conséquence du lemme 2, en choisissant $\tau < \tau' < \beta\sigma^2$ et en écrivant

$$\begin{aligned} & \int_{N(x) > \tau'} \exp(N^2(x)/2\sigma^2 - \beta N(x)) d\mu(x) \\ &= \int_0^\infty \exp((t + \tau')^2/2\sigma^2 - \beta(t + \tau')) \mu(\{x; N(x) > t + \tau'\}) dt \\ &\leq K \int_0^\infty \exp(t^2/2\sigma^2 - (\beta - \tau'/\sigma^2)t) P(\chi^2 > t^2/\sigma^2) dt < +\infty. \end{aligned}$$

Remarque. Une fois effectuée la réduction à \mathbb{R}^N , muni de la tribu borélienne \mathcal{B} , on n'utilise plus la forme explicite de N , mais seulement le fait que N est une

quasi-semi-norme mesurable, et que $N(x) \leq \sigma$ pour x dans B , (c'est-à-dire dans la boule unité de l'espace auto-reproduisant). On peut formuler des énoncés très généraux du type de ceux donnés dans [1], dont la preuve se ramène par des procédés, maintenant classiques, au cas mentionné. Nous ne l'avons pas fait pour ne pas obscurcir la simplicité du résultat. La même remarque s'applique aux trois lemmes.

3. Preuve des théorèmes 2 et 3

Nous utiliserons l'observation classique suivante:

Le théorème central limite montre que la mesure gaussienne canonique μ_n sur \mathbb{R}^n , la loi de $\|x\|^2$ est approximativement $N(n, 3\sqrt{n})$. En particulier μ_n est concentrée près de la sphère de rayon \sqrt{n} . On désigne par (n_k) une suite telle que

$$\mu_{n_k} \{x \in \mathbb{R}^{n_k}; \|\|x\| - \sqrt{n_k}\| \leq n_k^{1/4}\} \leq 2^{-k}. \tag{1}$$

On pose $I_p =]\sum_{i < p} n_{k_i}, \sum_{i \leq p} n_{k_i}]$. On pose $F_1 = \mathbb{R}$, $F_2 = \mathbb{R}^{11, \infty}$, on désigne par μ_1 (resp. μ_2) la gaussienne canonique sur F_1 (resp. F_2), et on pose $E = F_1 \times F_2$, $\mu = \mu_1 \times \mu_2$.

La condition (1) montre que pour μ_2 presque tout $y \in F_2$, on a

$$\lim_{k \rightarrow \infty} n_k^{-1} \sum_{i \in I_k} y_i^2 = 1. \tag{2}$$

Deux des exemples vont être obtenus en posant $N(x) = \sup_{f \in D} |f(x)|$, où $D \subset \mathbb{R}^{(\mathbb{N})}$ est dénombrable.

Preuve du théorème 2, (a). On prend pour D l'ensemble des suites finies de rationnels (z_n) qui vérifient la condition:

$$\sup \left(\sum z_n^2, \sup_k n_k \sum_{i \in I_k} z_i^2 \right) \leq 1. \tag{3}$$

On a $\sigma = 1$, et (2) montre que $\tau = 1$ et que $N < \infty$ μ .p.s. Pour $x \in F_1$, pour μ_2 presque tout $y \in F_2$, tout $\varepsilon > 0$ et tout k assez grand, il existe des $(z_i)_{i \in I_k}$ rationnels tels que $\sum_{i \in I_k} z_i y_i \geq 1 - \varepsilon$. Si on définit

$$z' \text{ par } z'_i = \left(1 - \frac{1}{n_k}\right) \text{ signe}(x), \quad z'_i = z_i \text{ pour } i \in I_k \text{ et } z'_i = 0 \text{ sinon, alors } z' \in D \text{ et}$$

$$z'(x, y) \geq |x| (1 - n_k^{-1}) + 1 - \varepsilon.$$

Il en résulte que $N(x, y) \geq |x| + 1$. On a donc

$$E \exp \left(\frac{N^2}{2} - N \right) \geq E \left[\exp \left(\frac{t^2}{2} + 1 \right) \exp(-N(y)) \right]$$

$$\geq a \int \exp \left(\frac{t^2}{2} + 1 \right) \exp \left(-\frac{t^2}{2} \right) dt = + \infty$$

pour un $a > 0$, ce qui montre le résultat.

Preuve du théorème 3. On se donne une suite de réels $b_i > 0$, avec $b_i \rightarrow 0$. On considère maintenant l'ensemble D des suites finies de rationnels (z_n) telles que

$$\sup \left(\sum z_n^2; \sup_k \frac{n_k}{b_k} \sum_{i \in I_k} z_i^2 \right) \leq 1. \quad (4)$$

On a toujours $\sigma = 1$, et on voit sans peine que $\tau = 0$. Soit

$$A = \{y \in F_2; \text{Inf}(\sum_k y_i^2)^{1/2} \geq 1/2\}.$$

La condition (2) montre que $a = \mu_2(A) > 0$. Pour $y \in A$, on a, comme dans la preuve précédente,

$$N(x, y) \geq \sup_k \{|x| (1 - n_k^{-1}) + b_k/2\} = u(x) \geq x.$$

Il en résulte que, puisque $N(x, y) \geq |x|$, on a

$$\begin{aligned} \mu(\{(x, y); N(x, y) < t\}) &\leq (1-a) \mu_1([-t, t]) + a \mu_1([-v(t), v(t)]) \\ &\leq (1-a) \mu_1([-\infty, t]) + a \mu_1([-\infty, v(t)]) \end{aligned}$$

où $v(t)$ est le sup des s tels que $u(s) < t$. Le choix des b_k étant arbitraire, on peut imposer à la fonction $u(x) - u$ de tendre vers zéro aussi lentement que spécifié à l'avance. Il en est de même de la fonction $t - v(t)$. Or, la fonction Φ^{-1} étant convexe sur $]0, 1[$, on a

$$\Phi^{-1} \circ \mu(\{z; N(z) \leq t\}) \leq (1-a)t + av(t) = t - a(t - v(t)).$$

Le résultat en découle.

Preuve du théorème 2(b). Toujours sur \mathbb{R}^N , on pose $N(x) = \sup_n |x_n| (2 \log n)^{-1/2}$.

Il est classique que $\sigma = \tau = 1$, et des estimations élémentaires montrent que $\Phi^{-1} \circ \mu(\{z; N(z) \leq t\}) - t \rightarrow 0$.

4. Preuve du théorème 4

On opère la même réduction que dans le cas du théorème 1.

Lemme 4. Avec les notations du lemme 1, on a $\mu.p.s.$

$$\lim_n N(x - R_n(x)) = N(x).$$

Preuve. Puisque N est semi-continue inférieurement, on a $\liminf_n N(x - R_n(x)) \geq N(x)$ pour tout x . D'autre part, si \mathcal{B}'_n désigne la tribu engendrée par les coordonnées de rang $\leq n$, la suite $(N(x - R_n(x)))$ est une surmartingale pour la suite \mathcal{B}'_n . Puisque $\int N(x - R_n(x)) d\mu(x) \leq \int N(x) d\mu(x)$, cette sous-martingale converge dans L^1 vers $N(x)$, donc aussi p.s.

On pose

$$B = \{x \in \mathbb{R}^N; N(x) \leq 1\}.$$

Le théorème d'Ehrhard [2] montre que la fonction

$$h(t) = \Phi^{-1}(\mu(tB))$$

est concave. En chaque point, elle est donc dérivable à droite et à gauche. Il en est donc de même de la fonction $t \rightarrow \mu(tB)$. Il suffit de montrer que $a=0$, où

$$a = \limsup_{h \rightarrow 0} (2h)^{-1} (\mu((1+h)B) - \mu((1-h)B)).$$

Soit $\varepsilon > 0$, fixé, et soit $\eta > 0$, que l'on précisera ultérieurement. Il existe $0 < \beta < 1/10$ tel que

$$\mu((1-3\beta)B) \geq \mu(B) - \eta.$$

L'hypothèse que $\tau=0$, et les lemmes 1 et 4 impliquent qu'il existe n tel que $v_n(H) \geq 1 - \eta$ et $\lambda_n(A) \geq \mu((1-3\beta)B) - \eta$, où

$$H = \{y \in F_n; N(0, y) \leq \beta/2\}$$

$$A = \{x \in \mathbb{R}^n; N(x, 0) \leq 1 - 3\beta\}.$$

Etant donné $t \in \mathbb{R}^+$ on considère la transformation P_t (resp. Q_t) de \mathbb{R}^N qui consiste à multiplier les coordonnées de rang $\leq n$ (resp. $> n$) par t .

Le résultat d'Ehrhard montre que la fonction $h(t) = \Phi^{-1} \circ \mu(P_t(B))$ est concave. En particulier il existe $1 < t_0 < 1 + \beta$ tel que pour $t \leq t_0$ on ait $\mu(P_t(B)) \leq \mu(B) + \eta$.

Soit maintenant $t \in [1 - \beta, t_0]$, et posons $C = P_t(B)$. La fonction $g(u) = \Phi^{-1} \circ \mu(Q_u(C))$ est concave. Pour $x \in A$, on a

$$N(tx, 0) \leq (1 + \beta)(1 - 3\beta) \leq 1 - \beta$$

d'où pour $y \in H$ on a $N(tx, 2y) \leq 1$; c'est-à-dire que $Q_2 \circ P_t(x, y) \in B$. Ainsi $A \times H \subset Q_{1/2}(C)$. Il en résulte que

$$\mu(Q_{1/2}(C)) \geq \lambda_n(A) v_n(H) \geq (1 - \eta)(\mu(B) - \eta)$$

$$\mu(Q_1(C)) \leq \mu(B) + \eta.$$

Si η est assez petit, cela implique que $g(1/2) \geq g(1) - \varepsilon$. Puisque g est concave croissante, pour $u \geq 3/4$, on a $g(u) - g(1) \leq 4\varepsilon|u - 1|$. Puisque Φ est lipschitzienne de rapport ≤ 1 , on a

$$|\mu(Q_t(P_t(B))) - \mu(P_t(B))| \leq 4\varepsilon|t - 1|.$$

Pour $0 \leq h \leq \inf(\beta, t_0 - 1)$ on a, puisque $tB = Q_t(P_t(B))$:

$$(2h)^{-1} \mu((1+h)B) - \mu((1-h)B) \leq 8\varepsilon + 2h^{-1} (\mu(P_{1+h}(B)) - \mu(P_{1-h}(B)))$$

et ainsi $a \leq 8\varepsilon + b$ où

$$b = \limsup_h (2h)^{-1} (\mu(P_{1+h}(B)) - \mu(P_{1-h}(B))).$$

Pour $y \in F_n$, $D \subset \mathbb{R}^N$, soit $D_y = \{x \in \mathbb{R}^N; N(x, y) \in D\}$. On a $tB_y = (P_t(B))_y$. Ainsi on a

$$\mu(P_{1+h}(B)) - \mu(P_{1-h}(B)) = \int_{F_n} \lambda_n((1+h)B_y) - \lambda_n((1-h)B_y) d\nu_n(y).$$

Pour chaque y , B_y est convexe, et l'ensemble des B_y est borné. Comme λ_n a une densité bornée par rapport à la mesure de Lebesgue, on voit sans peine (par exemple en utilisant les coordonnées polaires) que l'on est dans un cas de convergence dominée, et donc que

$$b \leq \int_{F_n} \limsup (2h)^{-1} (\lambda_n((1+h)B_y) - \lambda_n((1-h)B_y)) d\nu_n(y).$$

Pour un ensemble D , on a $\lambda_n(tD) = \int_D h_t(u) du$ où h_t est la densité de la mesure ν donnée par $\nu(A) = \lambda_n(tA)$. On peut dériver sous l'intégrale; et $t \rightarrow \mu_1(tD)$ est infiniment dérivable. Il s'ensuit que $b=0$, et termine la preuve.

Remarque. Un exemple de [4] montre que le théorème 4 est faux sans l'hypothèse $\tau=0$.

Bibliographie

1. Borell, C.: The Brunn-Minkowski Inequality in Gauss Space. *Inventiones Math.* **30**, 205-216 (1975)
2. Ehrhard, A.: Symétrisation dans l'espace de Gauss
3. Fernique, X.: Régularité des trajectoires des fonctions aléatoires Gaussiennes. *Lecture notes in Math.* N° **480**. Berlin-Heidelberg-New York: Springer 1974
4. Tsirelson, V.S.: The density of the distribution of the maximum of a Gaussian process. *Theory Probability Appl.* 847-856 (1975)
5. Weber, M.: Analyse asymptotique des processus gaussiens stationnaires. *Ann. Inst. Henri Poincaré*, Vol. **XVI**, n° 2, 117-176 (1980)

Reçu le 20 Juin 1983