Available online at www.sciencedirect.com

ScienceDirect

C. R. Acad. Sci. Paris, Ser. I 345 (2007) 513–518

COMPTES RENDUS MATHEMATIQUE

http://france.elsevier.com/direct/CRASS1/

Probability Theory

Bounds on the concentration function in terms of the Diophantine approximation

Omer Friedland, Sasha Sodin

School of Mathematical Sciences, Tel Aviv University, Ramat Aviv, Tel Aviv 69978, Israel

Received 14 August 2007; accepted 28 September 2007

Presented by Jean Bourgain

Abstract

We demonstrate a simple analytic argument that may be used to bound the Lévy concentration function of a sum of independent random variables. The main application is a version of a recent inequality due to Rudelson and Vershynin, and its multidimensional generalization. ***To cite this article: O. Friedland, S. Sodin, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Résumé

Des bornes pour la fonction de concentration en matière d'approximation diophantienne. Nous montrons un simple raisonnement analytique qui peut être utile pour borner la fonction de concentration d'une somme des variables aléatoires indépendantes. L'application principale est une version de l'inégalité récente de Rudelson et Vershynin, et sa généralisation au cadre multidimensionel. ***Pour citer cet article : O. Friedland, S. Sodin, C. R. Acad. Sci. Paris, Ser. I 345 (2007).***

Version française abrégée

La *fonction de concentration de P. Lévy* d'une variable aléatoire S est définie par

$$\mathcal{Q}(S) = \sup_{x\in\mathbb{R}} \mathbb{P}\big\{|S-x| \leqslant 1\big\}.$$

Depuis Lévy, Littlewood–Offord, Esseen, Kolmogorov, Halász et autres, des nombreux travaux de théorie des probabilités étaient consacrés aux bornes pour la fonction de concentration d'une somme des variables aléatoires indépendantes.

Notre étude a était motivée par le récent travail de Rudelson et Vershynin [4], qui ont considéré le problème suivant. Soit X une variable aléatoire, et soient $X_1, \dots, X_n$ des échantillons indépendants de X ; soit aussi $\mathbf{a} \in \mathbb{R}^n$ un vecteur. Quelles conditions doit on imposer à $\mathbf{a}$ pour garantir que

E-mail addresses: omerfrie@post.tau.ac.il (O. Friedland), sodinale@post.tau.ac.il (S. Sodin).

1631-073X/$ – see front matter
doi:10.1016/j.crma.2007.10.006

$$\mathcal{Q}\left(\sum_{k=1}^{n} X_k a_k\right) \leqslant C/|\mathbf{a}|? \tag{1}$$

Si $X \sim N(0,1)$ est Gaussienne, (1) tient toujours (avec une constante absolue $C > 0$), et est precise quand la norme Euclidienne $|\mathbf{a}| \to \infty$. En revanche, si X a des atomes, (1) doit défaillir quand $|\mathbf{a}| \to \infty$.

En suivant l'approche de Rudelson et Vershynin, nous fournissons une condition suffisante en matière d'approximation Diophantine de $\mathbf{a}$. Le thèoréme prochain améliore légèrment le résultat [4, Théorème 1.3] :

Théorème 0.1. *Soient $X_1, \ldots, X_n$ des échantillons indépendants d'une variable aléatoire X qui vérifie $\mathcal{Q}(X) \leqslant 1 - p$, et soit $\mathbf{a} \in \mathbb{R}^n$. Soient $\alpha > 0$ et $0 < D < 1$ tels que*

$$|\eta \mathbf{a} - \mathbf{m}| \geqslant \alpha \quad \textit{pour } \mathbf{m} \in \mathbb{Z}^n,\ \eta \in \big[1/(2\|a\|_\infty), D\big].$$

Alors

$$\mathcal{Q}\left(\sum_{k=1}^{n} X_k a_k\right) \leqslant C\left\{\exp\left(-cp^2\alpha^2\right) + \frac{1}{pD}\frac{1}{|\mathbf{a}|}\right\}.$$

Ce résultat a un homologue multidimensionel. Si $\vec{S}$ est un vecteur aléatoire à $\mathbb{R}^d$, on note

$$\mathcal{Q}(\vec{S}) = \sup_{\vec{x} \in \mathbb{R}^d} \mathbb{P}\big\{|\vec{S} - \vec{x}| \leqslant 1\big\}.$$

Théorème 0.2. *Soient $X_1, \ldots, X_n$ des échantillons indépendants d'une variable aléatoire X qui vérifie $\mathcal{Q}(X) \leqslant 1 - p$, et soit $\vec{\mathbf{a}} = (\vec{a}_1, \ldots, \vec{a}_n) \in (\mathbb{R}^d)^n$. Nous supposons aussi que $\alpha > 0$ et $0 < D < d$ vérifient*

$$\sum_{k=1}^{n} (\vec{\eta} \cdot \vec{a}_k - m_k)^2 \geqslant \alpha^2 \quad \textit{pour } m_1, \ldots, m_n \in \mathbb{Z} \textit{ et } \vec{\eta} \in \mathbb{R}^d \textit{ tel que } \max_k |\vec{\eta} \cdot \vec{a}_k| \geqslant 1/2,\ |\vec{\eta}| \leqslant D.$$

Alors

$$\mathcal{Q}\left(\sum_{k=1}^{n} X_k \vec{a}_k\right) \leqslant C^d\left\{\exp\left(-cp^2\alpha^2\right) + \left(\frac{\sqrt{d}}{pD}\right)^d \det{}^{-1/2}\left[\sum_{k=1}^{n} \vec{a}_k \otimes \vec{a}_k\right]\right\}.$$

La démonstration s'appuie sur une méthode analytique simple et évite la lemma de Halász [2]. Nous montrons la preuve du Théorème 0.1 et indiquons les modifications que l'on doit pour prouver le Théorème 0.2.

1. Introduction

The *P. Lévy concentration function* of a random variable S is defined as $\mathcal{Q}(S) = \sup_{x \in \mathbb{R}} \mathbb{P}\{|S - x| \leqslant 1\}$. Since the work of Lévy, Littlewood–Offord, Erdős, Esseen, Kolmogorov and others, numerous results in probability theory concern upper bounds on the concentration function of the sum of independent random variables; a particularly powerful approach was introduced in the 1970s by Halász [2].

This Note was motivated by the recent work of Rudelson and Vershynin [4]. Let X be a random variable; let $X_1, \ldots, X_n$ be independent copies of X, and let $\mathbf{a} = (a_1, \ldots, a_n)$ be an n-tuple of real numbers.

In the Gaussian case $X \sim N(0,1)$, we have: $\sum_{k=1}^{n} a_k X_k \sim N(0, |\mathbf{a}|^2)$ (where $|\cdot|$ stands for Euclidean norm), and consequently

$$\mathcal{Q}\left(\sum_{k=1}^{n} a_k X_k\right) = \sqrt{\frac{2}{\pi|\mathbf{a}|}}\big(1 + \mathrm{o}(1)\big), \quad |\mathbf{a}| \to \infty. \tag{2}$$

On the other hand, if X has atoms, the left-hand side of (2) does not tend to 0 as $|\mathbf{a}| \to \infty$. Therefore one may ask, for which $\mathbf{a} \in \mathbb{R}^n$ is it true that

$$\mathcal{Q}\left(\sum_{k=1}^{n} a_k X_k\right) \leqslant C/|\mathbf{a}|? \tag{3}$$

Rudelson and Vershynin gave a bound in terms of Diophantine approximation of the vector **a**. Their approach makes use of a deep measure-theoretic lemma from [2]. Our goal is to show a simpler analytic method that may be of use in such problems. The following theorem is a (slightly improved) version of [4, Theorem 1.3]:

Theorem 1.1. *Let $X_1, \dots, X_n$ be independent copies of a random variable X such that $\mathcal{Q}(X) \leqslant 1 - p$, and let $\mathbf{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$. If, for some $0 < D < 1$ and $\alpha > 0$,*

$$|\eta \mathbf{a} - \mathbf{m}| \geqslant \alpha \quad \text{for } \mathbf{m} \in \mathbb{Z}^n, \eta \in \big[1/(2\|a\|_\infty), D\big], \tag{4}$$

then

$$\mathcal{Q}\left(\sum_{k=1}^n X_k a_k\right) \leqslant C\left\{\exp(-cp^2\alpha^2) + \frac{1}{pD}\frac{1}{|\mathbf{a}|}\right\}. \tag{5}$$

Here and further $C, c, C', c_1, \dots > 0$ denote numerical constants.

We also extend this result to the multidimensional case. The concentration function of an $\mathbb{R}^d$-valued random vector $\vec{S}$ is defined as

$$\mathcal{Q}(\vec{S}) = \sup_{x \in \mathbb{R}^d} \mathbb{P}\big\{|\vec{S} - \vec{x}| \leqslant 1\big\}.$$

Theorem 1.2. *Let $X_1, \dots, X_n$ be independent copies of a random variable X such that $\mathcal{Q}(X) \leqslant 1 - p$, and let $\vec{a}_1, \dots, \vec{a}_n \in (\mathbb{R}^d)^n$ be such that, for some $0 < D < d$ and $\alpha > 0$,*

$$\sum_{k=1}^n (\vec{\eta} \cdot \vec{a}_k - m_k)^2 \geqslant \alpha^2 \quad \text{for } m_1, \dots, m_n \in \mathbb{Z}, \vec{\eta} \in \mathbb{R}^d \text{ such that } \max_k |\vec{\eta} \cdot \vec{a}_k| \geqslant 1/2, |\vec{\eta}| \leqslant D. \tag{6}$$

Then

$$\mathcal{Q}\left(\sum_{k=1}^n X_k \vec{a}_k\right) \leqslant C^d \left\{\exp(-cp^2\alpha^2) + \left(\frac{\sqrt{d}}{pD}\right)^d \det{}^{-1/2}\left[\sum_{k=1}^n \vec{a}_k \otimes \vec{a}_k\right]\right\}, \tag{7}$$

where $C, c > 0$ are numerical constants.

Of course, Theorem 1.1 follows formally from Theorem 1.2. For simplicity of exposition we will prove Theorem 1.1 and indicate the adjustments that are necessary for $d > 1$.

2. Proof of Theorem 1.1

Step 1: By Chebyshev's inequality and the identity

$$\exp(-y^2) = \int_{-\infty}^{+\infty} \exp\{2\mathrm{i}y\eta - \eta^2\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}$$

it follows that

$$\mathbb{P}\left\{\left|\sum_{k=1}^n X_k a_k - x\right| \leqslant 1\right\} \leqslant e\mathbb{E}\exp\left\{-\left[\sum_{k=1}^n X_k a_k - x\right]^2\right\}$$
$$= e\mathbb{E}\int_{-\infty}^{+\infty} \exp\left\{2\mathrm{i}\left[\sum_{k=1}^n X_k a_k - x\right]\eta - \eta^2\right\}\frac{\mathrm{d}\eta}{\sqrt{\pi}}.$$

Now we can swap the expectation with the integral and take absolute value:

$$\mathbb{P}\left\{\left|\sum_{k=1}^{n} X_k a_k - x\right| \leqslant 1\right\} \leqslant e \int_{-\infty}^{+\infty} \prod_{k=1}^{n} \phi(2a_k\eta) \exp\{-2\mathrm{i}x\eta - \eta^2\} \frac{d\eta}{\sqrt{\pi}}$$
$$\leqslant e \int_{-\infty}^{+\infty} \prod_{k=1}^{n} \left|\phi(2a_k\eta)\right| \exp\{-\eta^2\} \frac{\mathrm{d}\eta}{\sqrt{\pi}},$$

where $\phi(\eta) = \mathbb{E}\exp(\mathrm{i}\eta X)$ is the characteristic function of every one of the X_k. Therefore

$$\mathcal{Q}\left(\sum_{k=1}^{n} X_k a_k\right) \leqslant e \int_{-\infty}^{+\infty} \prod_{k=1}^{n} \left|\phi(2a_k\eta)\right| \exp\{-\eta^2\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}. \tag{8}$$

Step 2 (this step is analogous to [2, §3] and [4, 4.2]): First,

$$\left|\phi(\eta)\right| \leqslant \exp\left(-\frac{1}{2}\left(1 - \left|\phi(\eta)\right|^2\right)\right).$$

Let X' be an independent copy of X, $X^\# = X - X'$. Observe that

$$q = \mathbb{P}\left\{|X^\#| \geqslant 2\right\} \geqslant p^2/2$$

and

$$\left|\phi(\eta)\right|^2 = \mathbb{E}\exp(\mathrm{i}\eta X^\#) = \mathbb{E}\cos(\eta X^\#) \leqslant (1-q) + q\mathbb{E}\left\{\cos(\eta X^\#)\,\middle|\,|X^\#| \geqslant 2\right\};$$

therefore

$$\int_{-\infty}^{+\infty} \prod_{k=1}^{n} \left|\phi(2a_k\eta)\right| \exp\{-\eta^2\} \frac{\mathrm{d}\eta}{\sqrt{\pi}} \leqslant \int_{-\infty}^{+\infty} \exp\left\{-\frac{q}{2}\mathbb{E}\left[\sum_{k=1}^{n}\left(1 - \cos\left(2a_k\eta X^\#\right)\right)\middle|\,|X^\#| \geqslant 2\right] - \eta^2\right\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}$$
$$\leqslant \mathbb{E}\left[\int_{-\infty}^{+\infty} \exp\left\{-\frac{q}{2}\sum_{k=1}^{n}\left(1 - \cos(2a_k\eta X^\#)\right) - \eta^2\right\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}\,\middle|\,|X^\#| \geqslant 2\right].$$

Replace the conditional expectation with supremum over the possible values $z \geqslant 2$ of $|X^\#|$ and recall that

$$1 - \cos\theta \geqslant c_1 \min_{m \in \mathbb{Z}} |\theta - 2\pi m|^2;$$

then

$$\int_{-\infty}^{+\infty} \prod_{k=1}^{n} \left|\phi(2\eta a_k)\right| \exp\{-\eta^2\} \frac{\mathrm{d}\eta}{\sqrt{\pi}} \leqslant \sup_{z \geqslant 2} \int_{-\infty}^{+\infty} \exp\left\{-\frac{q}{2}\sum_{k=1}^{n}\left(1 - \cos(2z\eta a_k)\right) - \eta^2\right\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}$$
$$\leqslant \sup_{z \geqslant 2} \int_{-\infty}^{+\infty} \exp\left\{-c_2 p^2 \sum_{k=1}^{n} \min_{m_k} |z\eta a_k - \pi m_k|^2 - \eta^2\right\} \frac{\mathrm{d}\eta}{\sqrt{\pi}}$$
$$= \sup_{z \geqslant 2/\pi} \int_{-\infty}^{+\infty} \exp\left\{-c_3 p^2 \sum_{k=1}^{n} \min_{m_k} |\eta a_k - m_k|^2 - (\eta/z)^2\right\} \frac{\mathrm{d}\eta}{z\sqrt{\pi}}. \tag{9}$$

Step 3: Denote

$$A = \left\{\eta \in \mathbb{R} \mid \forall \mathbf{m} \in \mathbb{Z}^n\colon\ |\eta\mathbf{a} - \mathbf{m}| \geqslant \alpha/2\right\}, \quad B = \mathbb{R} \setminus A.$$

Then the last integral in (9) can be split into

$$\int_{-\infty}^{+\infty} = \int_A + \int_B, \tag{10}$$

and

$$\int_A \leqslant \exp(-c_3 p^2 \alpha^2). \tag{11}$$

On the other hand, if $\eta', \eta'' \in B$, then $|\eta'\mathbf{a} - \pi\mathbf{m}'|, |\eta''\mathbf{a} - \pi\mathbf{m}''| < \alpha/2$ for some $\mathbf{m}', \mathbf{m}'' \in \mathbb{Z}^n$, and hence

$$\left|(\eta' - \eta'')\mathbf{a} - (\mathbf{m}' - \mathbf{m}'')\right| < \alpha.$$

Therefore by (4) either $|\eta' - \eta''| < 1/(2\|a\|_\infty)$ or $|\eta' - \eta''| > D$. In other words, $B \subset \bigcup_j B_j$, where B_j are intervals of length $\leqslant 1/\|a\|_\infty$ such that any two points belonging to different B_j are at least D-apart.

Step 4: For every j there exists $\eta_j \in B_j$ such that

$$\int_{B_j} = \exp(-\eta_j^2/z^2) \int_{B_j} \exp\left\{-c_3 p^2 \sum_{k=1}^n \min_{m_k} |\eta a_k - m_k|^2\right\} \frac{\mathrm{d}\eta}{z\sqrt{\pi}}.$$

By Hölder's inequality

$$\int_{B_j} \leqslant \exp(-\eta_j^2/z^2) \prod_{k=1}^n \left\{\int_{B_j} \exp\left\{-\frac{c_3 p^2 |\mathbf{a}|^2}{a_k^2} \min_{m_k} |\eta a_k - m_k|^2\right\} \frac{\mathrm{d}\eta}{z\sqrt{\pi}}\right\}^{a_k^2/|\mathbf{a}|^2}. \tag{12}$$

The length of the interval $a_k B_j$ is $\leqslant 1$; hence m_k (which is the closest integer to ηa_k) can obtain at most 2 values while $\eta \in B_j$. Therefore every one of the integrals on the right-hand side of (12) is bounded by

$$2\int_{-\infty}^{+\infty} \exp\{-c_3 p^2 |\mathbf{a}|^2 \eta^2\} \frac{\mathrm{d}\eta}{z\sqrt{\pi}} = \frac{C_1}{zp|\mathbf{a}|},$$

and therefore

$$\int_B \leqslant \sum_j \int_{B_j} \leqslant \frac{C_1}{zp|\mathbf{a}|} \sum_j \exp(-\eta_j^2/z^2).$$

Now, B_j (and hence η_j) are D-separated; therefore

$$\begin{aligned}\sum_j \exp(-\eta_j^2/z^2) &\leqslant 2\sum_{j=0}^\infty \exp\bigl(-(Dj/z)^2\bigr) \\ &\leqslant 2\left\{1 + \int_0^{+\infty} \exp\bigl(-(D\eta/z)^2\,\mathrm{d}\eta\bigr)\right\} \leqslant 2(1 + C_2 z/D) \leqslant C_3 z/D.\end{aligned}$$

Hence finally

$$\int_B \leqslant \frac{C_4}{pD|\mathbf{a}|};$$

combining this with (8)–(11) we deduce (5).

3. Remarks

(i) The results can be also used to estimate the formally more general form of the Lévy concentration function:

$$\mathcal{L}\left(\sum_{k=1}^n X_k \vec{a}_k; \varepsilon\right) = \sup_{\vec{x} \in \mathbb{R}^d} \mathbb{P}\left\{\left|\sum_{k=1}^n X_k \vec{a}_k - \vec{x}\right| \leqslant \varepsilon\right\}.$$

Indeed, $\mathcal{L}(\sum_{k=1}^n X_k \vec{a}_k; \varepsilon) = \mathcal{Q}(\sum_{k=1}^n X_k \vec{a}_k/\varepsilon)$, so one can just apply the theorems to $\vec{a}_k' = \vec{a}_k/\varepsilon$.

(ii) By similar reasoning, the assumption $\mathcal{Q}(X) \leqslant 1 - p$ can be replaced with $\mathcal{Q}(KX) \leqslant 1 - p$ (for an arbitrary $K > 0$); this will only influence the values of the constants in (5), (7).

(iii) The proof of Theorem 1.2 is parallel to that of Theorem 1.1. The main difference appears in Step 4, where instead of Hölder's inequality one should use the Brascamp–Lieb–Luttinger rearrangement inequality [1]. (Note that a different rearrangement inequality was applied to a similar problem by Howard [3].)

Acknowledgements

We are grateful to our supervisor Vitali Milman for his support and for encouraging to write this note. We thank Mark Rudelson and Roman Vershynin for stimulating discussions, and in particular for suggesting the current formulation of Theorem 1.2 with improved dependence on the dimension d.

References

[1] H.J. Brascamp, E.H. Lieb, J.M. Luttinger, A general rearrangement inequality for multiple integrals, J. Functional Anal. 17 (1974) 227–237.

[2] G. Halász, Estimates for the concentration function of combinatorial number theory and probability, Period. Math. Hungar. 8 (3–4) (1977) 197–211.

[3] R. Howard, Estimates on the concentration function in $\mathbb{R}^d$: Notes on Lectures of Oskolkov, http://www.math.sc.edu/~howard/Notes/concentration.pdf.

[4] M. Rudelson, R. Vershynin, The Littlewood–Offord problem and invertibility of random matrices, arxiv preprint, math/0703503.