Теорія Ймовір. та Матем. Статист.
Вип. 80, 2009

Theor. Probability and Math. Statist.
No. 80, 2010, Pages 25–36
S 0094-9000(2010)00792-7
Article electronically published on August 18, 2010

INEQUALITIES FOR THE DISTRIBUTIONS OF FUNCTIONALS OF SUB-GAUSSIAN VECTORS

UDC 519.21

V. V. BULDYGIN AND E. D. PECHUK

ABSTRACT. Exponential inequalities for moment generating functions and for distributions of sub-Gaussian random vectors are studied in the paper.

1. Introduction

There is a number of papers devoted to exponential bounds for moment generating functions and for distributions of random variables $\|X\|$ and $\|X\| - \mathsf{E}\,\|X\|$, where X is either a *sub-Gaussian random vector* or a *sub-Gaussian random sequence* and where $\|\cdot\|$ is a norm in the corresponding space (see, for example, [1]–[6]).

Some known inequalities for the norms are improved in this paper. We also obtain their generalizations for a wider class of functionals. The main Theorem 3.1 and its corollaries are stated in Section 3. The proof of Theorem 3.1 is given in Section 5. Necessary definitions and related results are considered in Sections 2 and 4. A "contracting" transformation in the class of sub-Gaussian random variables is studied in Section 6. This transformation helps to improve some results under additional conditions.

The results obtained in the paper can be generalized for the case of *φ-sub-Gaussian random vectors*, as well. Some related results are obtained in [5]. We state the corresponding generalizations in Section 3; however, the proofs are provided for sub-Gaussian random vectors only.

2. Definitions

Norms and G-functions. Let $\mathbf{R} = (-\infty, \infty)$ be the set of real numbers.

For elements $\vec{x} = (x_1, \dots, x_m)$ and $\vec{y} = (y_1, \dots, y_m)$ of the space $\mathbf{R}^m$, we consider the *scalar product* $\langle \vec{x}, \vec{y} \rangle = \sum_{k=1}^m x_k y_k$, the *p-norms*

$$\|\vec{x}\|_p = \left(\sum_{k=1}^m |x_k|^p \right)^{1/p}, \qquad \vec{x} \in \mathbf{R}^m,\ p \in [1, \infty),$$

and the max-*norm*

$$\|\vec{x}\|_\infty = \max_{1 \le k \le m} |x_k|, \qquad \vec{x} \in \mathbf{R}^m.$$

In the *space of all real sequences* $\mathbf{R}^\infty$, we consider the *p-norms*

$$\|\mathbf{x}\|_p = \left(\sum_{k=1}^\infty |x_k|^p \right)^{1/p}, \qquad \mathbf{x} = (x_k, k \ge 1) \in \mathbf{R}^\infty,\ p \in [1, \infty),$$

2000 *Mathematics Subject Classification.* Primary 60G60, 60G17.

Key words and phrases. Sub-Gaussian random vectors, distributions of functionals, exponential inequalities.

and the sup-*norm*

$$\|\mathbf{x}\|_\infty = \sup_{k\ge1}|x_k|, \qquad \mathbf{x} = (x_k, k\ge 1)\in \mathbf{R}^\infty.$$

Along with the norms $\|\cdot\|_p$, $p\in[1,\infty)$, we study more general nonlinear functionals called G-functionals.

Let $G(\vec{x}) = G(x_1,\dots,x_m)$ be a positive function that is continuous in $\mathbf{R}^m$ and such that $G(\vec{0}) = G(0,\dots,0) = 0$. Moreover we assume that the partial derivatives

$$\frac{\partial G(\vec{x})}{\partial x_k}, \qquad k = 1,\dots,m,$$

exist and are continuous in $\mathbf{R}^m\setminus S_0$, where $S_0 = \bigcup_{k=1}^m\{\vec{x}\in\mathbf{R}^m\colon x_k\in S_k\}$ and S_k, $k=1,\dots,m$, are finite sets of numbers.

Let

$$\nabla G(\vec{x}) = \left(\frac{\partial G(\vec{x})}{\partial x_1},\dots,\frac{\partial G(\vec{x})}{\partial x_m}\right), \qquad \vec{x} = (x_1,\dots,x_m)\in\mathbf{R}^m\setminus S_0.$$

Consider the following condition:

$$(1)\qquad \left|\frac{\partial G(\vec{x})}{\partial x_k}\right| \le c_k, \qquad \vec{x}\in\mathbf{R}^m\setminus S_0,\ k=1,\dots,m,$$

where $c_k\in[0,\infty)$, $k=1,\dots,m$, and $\vec{c} = (c_1,\dots,c_m)$.

We denote by $\mathfrak{G}_0(\vec{c})$ the set of functions G satisfying condition (1).

If $G(\vec{x}) = \|\vec{x}\|_p$, $p\in[1,\infty)$, then

$$(2)\qquad \nabla\|\vec{x}\|_p = \left(\frac{\partial\|\vec{x}\|_p}{\partial x_1},\dots,\frac{\partial\|\vec{x}\|_p}{\partial x_m}\right), \qquad \vec{x} = (x_1,\dots,x_m)\in\mathbf{R}^m\setminus S_0,$$

where $S_0 = \bigcup_{k=1}^m\{\vec{x}\in\mathbf{R}^m\colon x_k = 0\}$,

$$\frac{\partial\|\vec{x}\|_p}{\partial x_k} = \frac{s(x_k)|x_k|^{p-1}}{\|\vec{x}\|_p^{p-1}}, \qquad k=1,\dots,m,$$

and where $s(x) = 1$ for $x\ge 0$ and $s(x) = -1$ for $x<0$.

It follows from (2) that, for all $p\in[1,\infty)$,

$$(3)\qquad \left|\frac{\partial\|\vec{x}\|_p}{\partial x_k}\right| \le 1, \qquad \vec{x}\in\mathbf{R}^m\setminus S_0,\ k=1,\dots,m.$$

Thus we have $\vec{c} = \vec{1} = (1,\dots,1)$ for $G = \|\cdot\|_p$, $p\in[1,\infty)$.

Random variables, vectors, and sequences. We assume that the random variables, vectors, and sequences considered below are defined on a common complete probability space $\{\Omega,\mathfrak{F},\mathsf{P}\}$.

Denote by $L_\infty(\Omega)$ the space of almost surely bounded random variables Z equipped with the norm

$$|Z|_\infty = \operatorname*{ess\,sup}_{\omega\in\Omega}|Z(\omega)|;$$

here $\mathbf{1}_A$ is the indicator of the set A.

We recall some definitions related to sub-Gaussian random variables, vectors, and sequences (see [2]).

A random variable ξ is called *sub-Gaussian* if there exists a number $a\in[0,\infty)$ such that

$$\mathsf{E}\exp\{\lambda\xi\} \le \exp\left\{\frac{a^2\lambda^2}{2}\right\}$$

for all $\lambda\in\mathbf{R}$.

The number

$$\tau(\xi) = \inf\left\{a \geq 0\colon\ \mathsf{E}\exp\{\lambda\xi\} \leq \exp\left\{\frac{a^2\lambda^2}{2}\right\}, \lambda \in \mathbf{R}\right\}$$

is called the *standard of a sub-Gaussian random variable* ξ. The functional $\tau(\xi)$ is a *norm* in the *Banach space* $\mathrm{Sub}(\Omega)$ *of all sub-Gaussian random variables*. The sub-Gaussian random variable ξ is assumed to be centered ($\mathsf{E}\,\xi = 0$) and such that

$$\mathsf{E}\exp\{\lambda\xi\} \leq \exp\left\{\frac{\tau^2(\xi)\lambda^2}{2}\right\}, \qquad \lambda \in \mathbf{R}.$$

The so-called *φ-sub-Gaussian random variables* are a generalization of sub-Gaussian random variables (see [2]). The φ-sub-Gaussian random variables form a Banach space denoted by $\mathrm{Sub}_\varphi(\Omega)$ and equipped with the norm τ_φ.

A random vector $\vec{X} = (X_1, \dots, X_m)$ is called *sub-Gaussian* (*φ-sub-Gaussian*) if all its coordinates are sub-Gaussian (*φ-sub-Gaussian*) random variables.

A sequence of sub-Gaussian (φ-sub-Gaussian) random variables

$$(X_k, k \geq 1)$$

is called a *sub-Gaussian* (*φ-sub-Gaussian*) *sequence*.

We use the following notation for a sub-Gaussian (*φ-sub-Gaussian*) random vector $\vec{X} = (X_1, \dots, X_m)$ and for a sub-Gaussian (*φ-sub-Gaussian*) sequence $\mathbf{X} = (X_k, k \geq 1)$:

$$B(\vec{c}; \vec{X}) = \sum_{k=1}^m c_k\tau(X_k), \qquad B(\vec{X}) = \sum_{k=1}^m \tau(X_k), \qquad B(\mathbf{X}) = \sum_{k=1}^\infty \tau(X_k),$$

$$B_\varphi(\vec{c}; \vec{X}) = \sum_{k=1}^m c_k\tau_\varphi(X_k), \qquad B_\varphi(\vec{X}) = \sum_{k=1}^m \tau_\varphi(X_k), \qquad B_\varphi(\mathbf{X}) = \sum_{k=1}^\infty \tau_\varphi(X_k).$$

3. Inequalities for functionals of sub-Gaussian vectors

We consider some general inequalities for functions of the class $\mathfrak{G}_0(\vec{c})$. The main Theorem 3.1 is proved in Section 5.

Theorem 3.1. *Let $G \in \mathfrak{G}_0(\vec{c})$. If $\vec{X}$ is a sub-Gaussian random vector in $\mathbf{R}^m$, then*

$$\mathsf{E}\exp\left\{\lambda G(\vec{X})\right\} \leq 2\exp\left\{\frac{\lambda^2 B^2(\vec{c}; \vec{X})}{2}\right\}, \qquad \lambda \in \mathbf{R}, \tag{4}$$

and

$$\mathsf{P}\left\{G(\vec{X}) \geq x\right\} \leq 2\exp\left\{-\frac{x^2}{2B^2(\vec{c}; \vec{X})}\right\}, \qquad x > 0. \tag{5}$$

Remark 3.1. If $\vec{X}$ is a φ-sub-Gaussian random vector in $\mathbf{R}^m$, then

$$\mathsf{E}\exp\left\{\lambda G(\vec{X})\right\} \leq 2\exp\left\{\varphi(\lambda B_\varphi(\vec{c}; \vec{X}))\right\}, \qquad \lambda \in \mathbf{R}, \tag{6}$$

and

$$\mathsf{P}\left\{G(\vec{X}) \geq x\right\} \leq 2\exp\left\{-\varphi^*\big(x/B_\varphi(\vec{c}; \vec{X})\big)\right\}, \qquad x > 0, \tag{7}$$

where φ^* is the conjugate function to φ; that is, φ^* is the Young–Fenchel transform of the function φ (see [2, p. 60]). Recall that $\varphi^*(x) = x^2/2$, $x \in \mathbf{R}$, if and only if $\varphi(x) = x^2/2$, $x \in \mathbf{R}$.

Remark 3.2. Since

$$\max\left\{ \mathsf{E}\exp\left\{\lambda\big(G(\vec{X})-\mathsf{E}\,G(\vec{X})\big)\right\}, \mathsf{E}\exp\left\{\lambda\big(\mathsf{E}\,G(\vec{X})-G(\vec{X})\big)\right\}\right\} \le \mathsf{E}\exp\left\{|\lambda|G(\vec{X})\right\}$$

for all $\lambda\in\mathbf{R}$, inequalities (4)–(7) hold if we use there

$$G(\vec{X})-\mathsf{E}\,G(\vec{X}) \quad \text{or} \quad \mathsf{E}\,G(\vec{X})-G(\vec{X})$$

instead of $G(\vec{X})$. Moreover,

$$\mathsf{P}\left\{\big|G(\vec{X})-\mathsf{E}\,G(\vec{X})\big|\ge x\right\}\le \kappa(x)\,\mathsf{P}\left\{G(\vec{X})\ge x\right\},$$

where $\kappa(x)=2$ for $0<x\le \mathsf{E}\,G(\vec{X})$ and $\kappa(x)=1$ for $x>\mathsf{E}\,G(\vec{X})$.

Consider some examples.

Example 3.1. Let

$$G(x_1,\dots,x_m)=\sum_{k=1}^m \ln(1+|x_k|^{a_k}),$$

where $a_k\in[1,\infty)$, $k=1,\dots,m$. It is clear that

$$\left|\frac{\partial G(\vec{x})}{\partial x_k}\right|\le c(a_k), \qquad \vec{x}\in\mathbf{R}^m\setminus S_0,$$

for all $k=1,\dots,m$, where $c(1)=1$, $c(a)=(a-1)^{(a-1)/a}$, $a>1$, and

$$S_0=\bigcup_{k=1}^m\{\vec{x}\in\mathbf{R}^m\colon x_k=0\}.$$

Therefore $G\in\mathfrak{G}_0(\vec{c})$, where $\vec{c}=(c(a_1),\dots,c(a_m))$. Also, if $\vec{X}$ is a sub-Gaussian random vector in $\mathbf{R}^m$, then

$$\mathsf{P}\left\{\prod_{k=1}^m(1+|X_k|^{a_k})\ge e^x\right\}\le 2\exp\left\{-\frac{x^2}{2B^2(\vec{c};\vec{X})}\right\}, \qquad x>0.$$

Example 3.2. Since $\|\vec{\cdot}\|_p\in\mathfrak{G}_0(\vec{1})$, $p\in[1,\infty)$, we have

$$\mathsf{E}\exp\left\{\lambda\|\vec{X}\|_p\right\}\le 2\exp\left\{\frac{\lambda^2B^2(\vec{X})}{2}\right\}, \qquad \lambda\in\mathbf{R}, \tag{8}$$

and

$$\mathsf{P}\left\{\|\vec{X}\|_p\ge x\right\}\le 2\exp\left\{-\frac{x^2}{2B^2(\vec{X})}\right\}, \qquad x>0, \tag{9}$$

for all sub-Gaussian random vectors $\vec{X}$ in $\mathbf{R}^m$ and for all $p\in[1,\infty]$.

Further, if $\vec{X}$ is a φ-sub-Gaussian random vector in $\mathbf{R}^m$, then

$$\mathsf{E}\exp\left\{\lambda\|\vec{X}\|_p\right\}\le 2\exp\left\{\varphi\big(\lambda B_\varphi(\vec{X})\big)\right\}, \qquad \lambda\in\mathbf{R}, \tag{10}$$

and

$$\mathsf{P}\left\{\|\vec{X}\|_p\ge x\right\}\le 2\exp\left\{-\varphi^*\big(x/B_\varphi(\vec{X})\big)\right\}, \qquad x>0, \tag{11}$$

for all $p\in[1,\infty]$.

One can derive inequalities (8)–(11) from the results of the paper [5], since $\|\vec{X}\|_p\le\|\vec{X}\|_1$ for all $p\in[1,\infty]$. Note also that these inequalities hold if we use

$$\|\vec{X}\|_p-\mathsf{E}\,\|\vec{X}\|_p \quad \text{or} \quad \mathsf{E}\,\|\vec{X}\|_p-\|\vec{X}\|_p$$

instead of $\|\vec{X}\|_p$. Moreover,

$$\mathsf{P}\left\{\left|\|\vec{X}\|_p - \mathsf{E}\,\|\vec{X}\|_p\right| \geq x\right\} \leq \kappa(x)\,\mathsf{P}\left\{\|\vec{X}\|_p \geq x\right\} \leq \kappa(x)\,\mathsf{P}\left\{\|\vec{X}\|_1 \geq x\right\}$$

for all $p \in [1, \infty]$, where $\kappa(x) = 2$ for $0 < x \leq \mathsf{E}\,\|\vec{X}\|_p$ and $\kappa(x) = 1$ for $x > \mathsf{E}\,\|\vec{X}\|_p$; see Remark 3.2.

The inequalities obtained in Example 3.2 can easily be generalized to the case of sub-Gaussian (φ-sub-Gaussian) sequences.

Corollary 3.1. *Let $\mathbf{X} = (X_k, k \geq 1)$ be a sub-Gaussian sequence. If*

$$\sum_{k=1}^{\infty} \tau(X_k) < \infty, \tag{12}$$

then

$$\mathsf{E}\exp\{\lambda\|\mathbf{X}\|_p\} \leq 2\exp\left\{\frac{\lambda^2 B^2(\mathbf{X})}{2}\right\}, \qquad \lambda \in \mathbf{R}, \tag{13}$$

and

$$\mathsf{P}\{\|\mathbf{X}\|_p \geq x\} \leq 2\exp\left\{-\frac{x^2}{2B^2(\mathbf{X})}\right\}, \qquad x > 0, \tag{14}$$

for all $p \in [1, \infty]$.

Further, if $\mathbf{X} = (X_k, k \geq 1)$ is a φ-sub-Gaussian sequence and

$$\sum_{k=1}^{\infty} \tau_\varphi(X_k) < \infty, \tag{15}$$

then

$$\mathsf{E}\exp\{\lambda\|\mathbf{X}\|_p\} \leq 2\exp\{\varphi(\lambda B_\varphi(\mathbf{X}))\}, \qquad \lambda \in \mathbf{R}, \tag{16}$$

and

$$\mathsf{P}\{\|\mathbf{X}\|_p \geq x\} \leq 2\exp\{-\varphi^*(x/B_\varphi(\mathbf{X}))\}, \qquad x > 0, \tag{17}$$

for all $p \in [1, \infty]$.

Remark 3.3. If either assumption (12) or assumption (15) holds, then (see [2, pp. 17 and 73]) $\mathsf{E}\,\|\mathbf{X}\|_p < \infty$ for all $p \in [1, \infty]$. This together with Corollary 3.1 implies that $\|\mathbf{X}\|_p$ can be substituted by

$$\|\mathbf{X}\|_p - \mathsf{E}\,\|\mathbf{X}\|_p \quad \text{or} \quad \mathsf{E}\,\|\mathbf{X}\|_p - \|\mathbf{X}\|_p$$

in inequalities (13), (14), and in (16), (17). Moreover,

$$\mathsf{P}\left\{\left|\|\mathbf{X}\|_p - \mathsf{E}\,\|\mathbf{X}\|_p\right| \geq x\right\} \leq \kappa(x)\,\mathsf{P}\{\|\mathbf{X}\|_p \geq x\} \leq \kappa(x)\,\mathsf{P}\{\|\mathbf{X}\|_1 \geq x\}$$

for all $p \in [1, \infty]$, where $\kappa(x) = 2$ for $0 < x \leq \mathsf{E}\,\|\mathbf{X}\|_p$ and $\kappa(x) = 1$ for $x > \mathsf{E}\,\|\mathbf{X}\|_p$.

4. Auxiliary results

To prove Theorem 3.1, we need some auxiliary results considered in this section. Note that some similar results are known in the literature (see [1]–[6]).

Lemma 4.1. *If $a_1, \dots, a_m$ are nonzero real numbers, then the minimum of the function*

$$f(\vec{x}) = a_1^2 x_1 + \cdots + a_m^2 x_m$$

in the set

$$S = \left\{ \vec{x} = (x_1, \dots, x_m) \colon x_j > 1, j = 1, \dots, m, \frac{1}{x_1} + \cdots + \frac{1}{x_m} = 1 \right\}$$

is equal to

$$\left(\sum_{j=1}^{m} |a_j| \right)^2 .$$

Proof. Another proof of this result can be found in [3]. However the method used in [3] is rather complicated and we propose a simpler proof below.

Since $x_j > 0$, $j = 1, \dots, m$, in the set S, we get

$$|a_1| + \cdots + |a_m| = \frac{|a_1|\sqrt{x_1}}{\sqrt{x_1}} + \cdots + \frac{|a_m|\sqrt{x_m}}{\sqrt{x_m}}.$$

The Cauchy–Bunyakovskiĭ inequality implies that

$$\frac{|a_1|\sqrt{x_1}}{\sqrt{x_1}} + \cdots + \frac{|a_m|\sqrt{x_m}}{\sqrt{x_m}} \le \left(\sum_{j=1}^{m} |a_j|^2 x_j \right)^{\frac{1}{2}} \left(\frac{1}{x_1} + \cdots + \frac{1}{x_m} \right)^{\frac{1}{2}} = \left(\sum_{j=1}^{m} |a_j|^2 x_j \right)^{\frac{1}{2}}$$

for all $(x_1, \dots, x_m) \in S$. Hence

$$\left(|a_1| + \cdots + |a_m| \right)^2 \le \sum_{j=1}^{m} |a_j|^2 x_j$$

for all $(x_1, \dots, x_m) \in S$. Note that the inequality becomes an equality if

$$x_j = \frac{\sum_{k=1}^{m} |a_k|}{|a_j|} > 1, \qquad j = 1, \dots, m. \qquad \square$$

Lemma 4.2. *If V is a sub-Gaussian random variable and $U \in L_\infty(\Omega, \mathfrak{F}, \mathsf{P})$, then*

$$\mathsf{E} \exp\{\lambda U V\} \le \mathsf{E} \exp\{|\lambda| \cdot |U| \cdot |V|\} \le 2 \exp\left\{ \frac{\lambda^2}{2} \tau^2(V) |U|_\infty^2 \right\}, \qquad \lambda \in \mathbf{R}.$$

Proof. Consider

$$\cosh(x) = \frac{1}{2} \left(\exp\{x\} + \exp\{-x\} \right), \qquad x \in \mathbf{R}.$$

Since V is a sub-Gaussian random variable, we obtain

$$\begin{aligned} \mathsf{E} \exp\{|\lambda| \cdot |U| \cdot |V|\} &\le \mathsf{E} \exp\{|\lambda| \cdot |U|_\infty |V|\} \le 2\, \mathsf{E} \cosh(|\lambda| \cdot |U|_\infty |V|) \\ &= 2\, \mathsf{E} \cosh(|\lambda| \cdot |U|_\infty V) \le 2 \exp\left\{ \frac{\lambda^2}{2} \tau^2(V) |U|_\infty^2 \right\} \end{aligned}$$

for all $\lambda \in \mathbf{R}$. $\square$

Remark 4.1. If the random variables U and V in Lemma 4.2 are independent, then

$$\mathsf{E} \exp\{\lambda U V\} \le \exp\left\{ \frac{\lambda^2}{2} \tau^2(V) |U|_\infty^2 \right\}, \qquad \lambda \in \mathbf{R}.$$

Lemma 4.3. *If $V_1, \dots, V_m$ are sub-Gaussian random variables and*

$$U_k \in L_\infty(\Omega, \mathfrak{F}, \mathsf{P}), \qquad k = 1, \dots, m,$$

then

$$\mathsf{E}\exp\left\{\lambda \sum_{k=1}^m U_k V_k\right\} \le \mathsf{E}\exp\left\{|\lambda| \sum_{k=1}^m |U_k| \cdot |V_k|\right\} \le 2\exp\left\{\frac{\lambda^2}{2}\left(\sum_{k=1}^m \tau(V_k)|U_k|_\infty\right)^2\right\},$$
$$\lambda \in \mathbf{R}.$$

Proof. Without loss of generality one can assume that $\tau(V_k)|U_k|_\infty > 0$, $k = 1, \dots, m$. The Hölder inequality and Lemma 4.2 imply that

$$\begin{aligned}\mathsf{E}\exp\left\{|\lambda| \sum_{k=1}^m |U_k| \cdot |V_k|\right\} &\le \prod_{k=1}^m \left(\mathsf{E}\exp\{p_k|\lambda| \cdot |U_k| \cdot |V_k|\}\right)^{1/p_k} \\ &\le \prod_{k=1}^m \left(2\exp\left\{\frac{\lambda^2}{2} p_k^2 \tau^2(V_k)|U_k|_\infty^2\right\}\right)^{1/p_k} = 2\exp\left\{\frac{\lambda^2}{2}\sum_{k=1}^m p_k \tau^2(V_k)|U_k|_\infty^2\right\}\end{aligned}$$

for all $\lambda \in \mathbf{R}$ and $p_1 > 1, \dots, p_m > 1$ such that $(1/p_1) + \cdots + (1/p_m) = 1$. Putting $p_k = (\sum_{k=1}^m \tau(V_k)|U_k|_\infty)/\tau(V_k)|U_k|_\infty$, $k = 1, \dots, m$, we prove the desired inequality.

Note that Lemma 4.1 is not used in the proof of Lemma 4.3. On the other hand, this shows that the result of Lemma 4.3 cannot be improved in general. □

Remark 4.2. Let $V_1, \dots, V_m$ be sub-Gaussian random variables and

$$U_k \in L_\infty(\Omega, \mathfrak{F}, \mathsf{P}), \qquad k = 1, \dots, m.$$

If random variables $U_1V_1, \dots, U_mV_m$ are independent and U_k and V_k are mutually independent for all $k = 1, \dots, m$, then

$$\mathsf{E}\exp\left\{\lambda \sum_{k=1}^m U_k V_k\right\} \le \exp\left\{\frac{\lambda^2}{2}\sum_{k=1}^m \tau^2(V_k)|U_k|_\infty^2\right\}, \qquad \lambda \in \mathbf{R}.$$

Lemma 4.3 and relation (1) imply the following result.

Lemma 4.4. *Let $G \in \mathfrak{G}_0(\vec{c})$ and let $\vec{V} = (V_1, \dots, V_m)$ be a sub-Gaussian random vector. If $\vec{Z} = (Z_1, \dots, Z_m)$ is a random vector such that*

$$\sum_{k=1}^m \mathsf{P}\{Z_k \in S_k\} = 0,$$

then

$$\mathsf{E}\exp\left\{\lambda\left\langle \nabla G(\vec{Z}), \vec{V}\right\rangle\right\} \le 2\exp\left\{\frac{\lambda^2}{2}\left(\sum_{k=1}^m c_k \tau(V_k)\right)^2\right\}, \qquad \lambda \in \mathbf{R}.$$

5. Proof of Theorem 3.1

We use the gradient method in the proof of Theorem 3.1 (see, for example, [3, Theorem 1.1] and [7, Theorem 4.3.4]).

Proof. Consider the following vector-valued random process:

$$\vec{Z}(\theta) = \theta\vec{X}, \qquad \theta \in [0, 1],$$

and note that $\vec{Z}(1) = \vec{X}$ and $\vec{Z}(0) = \vec{0}$.

For all $\omega \in \Omega$, the corresponding path of this process is differentiable with respect to θ and

$$\frac{d}{d\theta}\vec{Z}(\theta) = \vec{X}, \qquad \theta \in [0, 1].$$

Assume that

$$\mathsf{P}\{X_k = 0\} = 0, \qquad k = 1, \dots, m. \tag{18}$$

Then there exists a random event $\Omega_1 \subset \Omega$ such that $\mathsf{P}(\Omega_1) = 1$ and, for all $\omega \in \Omega_1$, the corresponding paths of the random process $G(\vec{Z}(\theta))$, $\theta \in [0,1]$, are continuously differentiable with respect to θ except for a finite set of points $\theta \in [0,1]$ such that $\min_{1\le k\le m} \rho(Z_k(\theta), S_k) = 0$, where $\rho(a, S)$ is the distance between a number a and a set S. For all θ, except for the points mentioned above, we have

$$\frac{d}{d\theta} G(\vec{Z}(\theta)) = \sum_{k=1}^{m} \left(\frac{\partial G(\vec{x})}{\partial x_k}\right)_{\vec{x}=\vec{Z}(\theta)} \cdot \frac{d}{d\theta} Z_k(\theta) = \left\langle \nabla G(\vec{Z}(\theta)), \vec{X} \right\rangle .$$

Thus

$$G(\vec{X}) = \int_0^1 \frac{d}{d\theta} G(\vec{Z}(\theta))\, d\theta = \int_0^1 \left\langle \nabla G(\vec{Z}(\theta)), \vec{X} \right\rangle d\theta$$

almost surely.

This together with the Jensen inequality implies that

$$\begin{aligned} \mathsf{E}\exp\{\lambda G(X)\} &= \mathsf{E}\exp\left\{\lambda \int_0^1 \left\langle \nabla G(\vec{Z}(\theta)), \vec{X} \right\rangle d\theta\right\} \\ &\le \int_0^1 \mathsf{E}\exp\left\{\lambda \left\langle \nabla G(\vec{Z}(\theta)), \vec{X} \right\rangle\right\} d\theta. \end{aligned}$$

The latter result and Lemma 4.4 imply inequality (4). Therefore inequality (4) is proved under the extra assumption (18).

To avoid assumption (18) we consider a sub-Gaussian vector

$$\vec{X} = (X_1, \dots, X_m)$$

and another random vector $\vec{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_m)$ that is independent of $\vec{X}$, where $\varepsilon_1, \dots, \varepsilon_m$ are independent identically distributed random variables assuming values ± 1 each with the probability $\frac{1}{2}$. Using the random vectors $\vec{X}$ and $\vec{\varepsilon}$, we construct another random vector

$$\vec{X}^{(\alpha)} = \left(X_1^{(\alpha)}, \dots, X_m^{(\alpha)}\right), \qquad \alpha > 0,$$

where

$$X_k^{(\alpha)} = X_k + \alpha \varepsilon_k \mathbf{1}_{\{X_k = 0\}}, \qquad k = 1, \dots, m.$$

For all $\alpha > 0$, $\vec{X}^{(\alpha)}$ is a sub-Gaussian random vector satisfying condition (18). Thus inequality (4) holds for all $\vec{X}^{(\alpha)}$, $\alpha > 0$.

Moreover,

$$\left|\tau\left(X_k^{(\alpha)}\right) - \tau(X_k)\right| \le \tau\left(\alpha \varepsilon_k \mathbf{1}_{\{X_k=0\}}\right) \le \alpha \quad \text{and} \quad \left|X_k^{(\alpha)} - X_k\right| \le \alpha$$

for all $k = 1, \dots, m$. Hence

$$\lim_{\alpha\downarrow 0} B\left(\vec{c}; \vec{X}^{(\alpha)}\right) = B(\vec{c}; \vec{X}), \qquad \lim_{\alpha\downarrow 0} G\left(X^{(\alpha)}\right) = G(X), \qquad \lim_{\alpha\downarrow 0} \mathsf{E}\, G\left(\vec{X}^{(\alpha)}\right) = \mathsf{E}\, G(\vec{X}).$$

Inequality (4) for a general sub-Gaussian random vector $\vec{X}$ follows from the latter equalities and the Fatou theorem.

In its turn, inequality (5) follows from (4); see, for example, [2, p. 14]. Therefore the proof of Theorem 3.1 is complete. □

6. A "contracting" transformation of random variables

Let ξ be a centered random variable (that is, $\mathsf{E}\,\xi = 0$). Consider another random variable $T\xi$ defined by

$$T\xi = |\xi| - \mathsf{E}\,|\xi|.$$

The transformation $\xi \mapsto T\xi$ is called the "*contraction*" of the random variable ξ. This name is explained by the inequality $\mathsf{E}\,|T\xi|^2 \le \mathsf{E}\,|\xi|^2$, which means that the nonlinear transformation T is a contraction in the space $L_2(\Omega)$ and its contraction coefficient is given by

$$\varkappa_2(\xi) = \left(\frac{\mathsf{E}\,|T\xi|^2}{\mathsf{E}\,|\xi|^2}\right)^{\frac{1}{2}} = \left(1 - \frac{(\mathsf{E}\,|\xi|)^2}{\mathsf{E}\,|\xi|^2}\right)^{\frac{1}{2}}.$$

For nondegenerate random variables $\xi \in L_2(\Omega)$, one always has $\varkappa_2(\xi) < 1$. For example,

$$\varkappa_2(\xi) = \left(1 - \frac{2}{\pi}\right)^{\frac{1}{2}}$$

for a Gaussian random variable ξ with parameters $(0, \sigma^2)$. For a Rademacher random variable ξ assuming values ± 1 with probability $\frac{1}{2}$, we have $T\xi = 0$ and $\varkappa_2(\xi) = 0$.

For other Banach spaces of random variables, the evaluation of the contraction coefficient can be a complicated problem. The following example provides a nondegenerate random variable whose contraction coefficient is the same in any Banach space.

Example 6.1. Let $A \in \mathfrak{F}$ and $\mathsf{P}(A) = p \in [0,1]$. Then

$$\mathbf{1}_A = (\mathbf{1}_A(\omega), \omega \in \Omega)$$

is a Bernoulli random variable assuming values 1 and 0 with probabilities p and $q = 1-p$, respectively. Thus $\xi_p = \mathbf{1}_A - p$ is a centered Bernoulli random variable assuming values q and $-p$ with probabilities p and q, respectively. Hence

$$T\xi_p = |\mathbf{1}_A - p| - \mathsf{E}\,|\mathbf{1}_A - p| = q\mathbf{1}_A + p(1 - \mathbf{1}_A) - 2pq = (q-p)\,(\mathbf{1}_A - p) = (q-p)\xi_p,$$

that is, $T\xi_p = (q-p)\xi_p$. Note that $\xi_p = 0$ for $p = 1$ or $p = 0$, while $T\xi_p = 0$ for $p = 1$, $p = 0$, or $p = \frac{1}{2}$. This means that

$$\mathfrak{n}(T\xi_p) = |q-p|\mathfrak{n}(\xi_p)$$

for an arbitrary norm $\mathfrak{n}$ in every Banach space of random variables $\mathfrak{N}(\Omega)$ such that $L_\infty(\Omega) \subset \mathfrak{N}(\Omega)$.

The transformation $\xi \mapsto T\xi$ is well defined in the Banach space of sub-Gaussian random variables $(\mathrm{Sub}(\Omega), \tau)$. Let

$$\varkappa_{\mathrm{sub}}(\xi) = \frac{\tau(T\xi)}{\tau(\xi)}, \qquad \xi \in \mathrm{Sub}(\Omega).$$

We agree that $\varkappa_{\mathrm{sub}}(\xi) = 0$ if $\tau(\xi) = 0$.

An upper bound for $\varkappa_{\mathrm{sub}}(\xi)$ over the whole space of sub-Gaussian random variables is found in the following assertion.

Lemma 6.1. *In the class of sub-Gaussian random variables we have*:

$$K_{\mathrm{sub}} = \sup_{\xi \in \mathrm{Sub}(\Omega)} \varkappa_{\mathrm{sub}}(\xi) \le 2\,(2)^{\frac{1}{2}}\,(2.05)^{\frac{1}{4}}\,(e)^{\frac{1}{16}} < 4. \tag{19}$$

Proof. Let $\xi \in \mathrm{Sub}(\Omega)$. Then, for all $n \geq 1$,

$$\mathsf{E}(T\xi)^{2n} = \mathsf{E}(|\xi| - \mathsf{E}\,|\xi|)^{2n} \leq \sum_{k=0}^{n} C_{2n}^{2k}\,\mathsf{E}\,|\xi|^{2(n-k)}(\mathsf{E}\,|\xi|)^{2k}.$$

Since

$$\mathsf{E}\,|\xi|^a \leq 2\left(\frac{a}{e}\right)^{a/2}\tau^a(\xi)$$

for all $a > 0$ (see [2, p. 17]), we get

$$\mathsf{E}(T\xi)^{2n} \leq \left(\frac{16n\tau^2(\xi)}{e}\right)^n, \qquad n \geq 1,$$

and

$$\theta_1(T\xi) = \sup_{n\geq 1}\left(\frac{2^n n!}{(2n)!}\,\mathsf{E}(T\xi)^{2n}\right)^{\frac{1}{2n}} \leq 2\,(2)^{\frac{1}{2}}\,(e)^{\frac{1}{16}}\,\tau(\xi).$$

This implies inequality (19), since

$$\tau(T\xi) \leq (2.05)^{\frac{1}{4}}\,\theta_1(T\xi)$$

(see Corollary 2.2 in [6]). □

Below we evaluate the coefficient $\varkappa_{\mathrm{sub}}(\xi)$ for some examples. Note that both cases are possible, namely $\varkappa_{\mathrm{sub}}(\xi) < 1$ and $\varkappa_{\mathrm{sub}}(\xi) > 1$.

Example 6.2. Let ξ_p, $p \in (0,1)$, be a nondegenerate centered Bernoulli random variable (see Example 6.1). Then

$$\varkappa_{\mathrm{sub}}(\xi_p) = |1-2p| < 1$$

and

$$\tau(T\xi_p) \leq |1-2p|\max\{p, 1-p\},$$

since $\tau(\xi_p) \leq \max\{p, 1-p\}$.

Example 6.3. Let ξ be a Gaussian random variable with parameters $(0,\sigma^2)$. Then $\tau(T\xi) = \tau(\xi) \leq \sigma$ and $\varkappa_{\mathrm{sub}}(\xi) \leq 1$.

Example 6.4. Let η_p, $p \in [0,1]$, be a random variable assuming three values, namely 0 with probability $q = 1-p$ and ± 1 with probabilities $p/2$. It is known ([2, p. 25]) that $\tau(\eta_p) = \sqrt{p}$ for $p \in [\frac{1}{3}, 1]$. Since $T\eta_p = \xi_p$,

$$\tau(T\eta_p) = \tau(\xi_p) \leq \max\{p, 1-p\}, \qquad p \in (0,1)$$

(see Example 6.2). Thus

$$\varkappa_{\mathrm{sub}}(\eta_p) \leq \max\left\{\sqrt{p}, \frac{1}{\sqrt{p}} - \sqrt{p}\right\}, \qquad p \in \left[\frac{1}{3}, 1\right).$$

Note that

$$\varkappa_{\mathrm{sub}}(\eta_p) > 1, \qquad \frac{1}{3} \leq p < \left(\frac{\sqrt{5}-1}{2}\right)^2,$$

and

$$\frac{1}{\sqrt{2}} \leq \varkappa_{\mathrm{sub}}(\eta_p) \leq \frac{2}{\sqrt{3}}, \qquad \frac{1}{3} \leq p < 1.$$

The following assertion exhibits a possible application of the "contracting" transformation introduced above.

Proposition 6.1. *If $\xi_1,\dots,\xi_m$ are sub-Gaussian random variables, then*

$$\mathsf{E}\exp\left\{\lambda\left(\sum_{k=1}^m|\xi_k|-\mathsf{E}\sum_{k=1}^m|\xi_k|\right)\right\}\le\exp\left\{\frac{\lambda^2}{2}\tau^2\left(\sum_{k=1}^m(|\xi_k|-\mathsf{E}\,|\xi_k|)\right)\right\}\tag{20}$$

$$\le\exp\left\{\frac{\lambda^2}{2}\left(\sum_{k=1}^m\varkappa_{\mathrm{sub}}(\xi_k)\tau(\xi_k)\right)^2\right\}\le\exp\left\{8\lambda^2\left(\sum_{k=1}^m\tau(\xi_k)\right)^2\right\},\tag{21}$$

$$\lambda\in\mathbf{R}.$$

Proof. Inequality (20) follows, since the random variable $\sum_{k=1}^m|\xi_k|-\mathsf{E}\sum_{k=1}^m|\xi_k|$ is sub-Gaussian by Lemma 6.1. In its turn, relation (20) implies (21), since τ is a norm in the space $\mathrm{Sub}(\Omega)$, whence we derive that

$$\tau\left(\sum_{k=1}^m(|\xi_k|-\mathsf{E}\,|\xi_k|)\right)\le\sum_{k=1}^m\tau(|\xi_k|-\mathsf{E}\,|\xi_k|)=\sum_{k=1}^m\varkappa_{\mathrm{sub}}(\xi_k)\tau(\xi_k).$$

This together with Lemma 6.1 implies (21). □

If $\varkappa_{\mathrm{sub}}(\xi_k)\le 1$, $k=1,\dots,m$, then inequality (21) implies the result of Example 3.2 for $\|\vec{X}\|_1-\mathsf{E}\,\|\vec{X}\|_1$.

For independent random variables $\xi_1,\dots,\xi_m$, we prove the following assertion.

Proposition 6.2. *If $\xi_1,\dots,\xi_m$ are independent sub-Gaussian random variables, then*

$$\mathsf{E}\exp\left\{\lambda\left(\sum_{k=1}^m|\xi_k|-\mathsf{E}\sum_{k=1}^m|\xi_k|\right)\right\}\le\exp\left\{\frac{\lambda^2}{2}\sum_{k=1}^m\varkappa_{\mathrm{sub}}^2(\xi_k)\tau^2(\xi_k)\right\}\tag{22}$$

$$\le\exp\left\{8\lambda^2\sum_{k=1}^m\tau^2(\xi_k)\right\},\qquad\lambda\in\mathbf{R}.\tag{23}$$

Proof. Since the random variables $\xi_1,\dots,\xi_m$ are independent, the sub-Gaussian random variables $\xi_1-\mathsf{E}\,\xi_1,\dots,\xi_m-\mathsf{E}\,\xi_m$ are also independent. Thus

$$\mathsf{E}\exp\left\{\lambda\left(\sum_{k=1}^m|\xi_k|-\mathsf{E}\sum_{k=1}^m|\xi_k|\right)\right\}=\prod_{k=1}^m\mathsf{E}\exp\left\{\lambda\left(|\xi_k|-\mathsf{E}\,|\xi_k|\right)\right\}$$
$$\le\prod_{k=1}^m\mathsf{E}\exp\left\{\frac{\lambda^2}{2}\varkappa_{\mathrm{sub}}^2(\xi_k)\tau^2(\xi_k)\right\}=\exp\left\{\frac{\lambda^2}{2}\sum_{k=1}^m\varkappa_{\mathrm{sub}}^2(\xi_k)\tau^2(\xi_k)\right\},\qquad\lambda\in\mathbf{R}.$$

Therefore inequality (22) is proved.

Now relation (22) and Lemma 6.1 imply inequality (23). □

7. Concluding remarks

Exponential bounds for moment generating functions and the corresponding exponential inequalities for the distributions of nonlinear functionals of sub-Gaussian and φ-sub-Gaussian random vectors are obtained in the paper. The proofs of the results of the paper are based on the gradient and "contracting" methods. Our general inequalities imply the corresponding inequalities for the p-norms.

Acknowledgement

The authors are indebted to Prof. Yu. V. Kozachenko for valuable comments, which improved the presentation.

Bibliography

1. R. Fukuda, *Exponential integrability of sub-Gaussian vectors*, Probab. Theory Relat. Fields **85** (1990), no. 4, 505–521. MR1061942 (92e:60005)
2. V. V. Buldygin and Yu. V. Kozachenko, *Metric Characterization of Random Variables and Random Processes*, TViMS, Kiev, 1998; English transl., American Mathematical Society, Providence, RI, 2000. MR1743716 (2001g:60089)
3. R. Giuliano Antonini, *Exponential bounds for the distribution for the norm of scalarly subgaussian random vectors*, Sezione di analisi matematica e probabilita, vol. 2.240.957, Universita di Pisa, Maggio, 1996.
4. R. Giuliano Antonini, Yu. V. Kozachenko, and A. M. Tegza, *Inequalities for the norms of sub-Gaussian vectors and the accuracy of the modeling of random processes*, Teor. Imovirnost. Matem. Statist. **66** (2002), 58–66; English transl in. Theory Probab. Math. Stat. **66** (2003), 63–72. MR1931677 (2003h:60053)
5. R. Giuliano Antonini and Yu. V. Kozachenko, *A note on the asymptotic behavior of sequences of generalized sub-Gaussian random vectors*, Rand. Oper. Stoch. Equ. **13** (2005), no. 1, 39–52. MR2130246 (2006b:60020)
6. N. N. Vakhaniya, V. V. Kvaratskheliya, and V. I. Tarieladze, *Weakly sub-Gaussian random elements in Banach spaces*, Ukr. Mat. Zh. **57** (2005), no. 9, 1187–1208; English transl. in Ukr. Math. J. **57** (2005), no. 9, 1387–1412. MR2216040 (2007c:60006)
7. V. I. Bogachev, *Gaussian Measures*, Nauka, Moscow, 1997; English transl. in American Mathematical Society, Providence, RI, 1998. MR1663929 (2000b:60097); MR1642391 (2000a:60004)

Department of Mathematical Analysis and Probability Theory, Faculty for Physics and Mathematics, National Technical University of Ukraine "KPI", Peremogy Avenue, 37, Kiev 03056, Ukraine

E-mail address: matan@ntu-kpi.kiev.ua

Department of Mathematical Analysis and Probability Theory, Faculty for Physics and Mathematics, National Technical University of Ukraine "KPI", Peremogy Avenue, 37, Kiev 03056, Ukraine

E-mail address: matan@ntu-kpi.kiev.ua

Received 17/NOV/2008

Translated by OLEG KLESOV