Note di Matematica Vol. 18 - n. 1, 113-117 (1998)

EXPONENTIAL BOUNDS FOR THE DISTRIBUTION OF THE NORM OF SUBGAUSSIAN RANDOM VECTORS[1]

RITA GIULIANO ANTONINI

0. INTRODUCTION

Let X be a real random variable, subgaussian in the sense of [1]. It is shown in [1] that

$$P(|X| > t) \leq 2\exp\left(-\frac{t^2}{2\tau^2(X)}\right) \tag{0.1}$$

where $\tau(X)$ is the gaussian standard of X.

Two classes of subgaussian Banach space-valued random vectors are defined in the paper [2]; for both of them it is proved that $E[e^{\varepsilon\|X\|^2}] < \infty$ for some ε. This yields

$$P(\|X\| > t) = P(e^{\varepsilon\|X\|^2} > e^{\varepsilon t^2}) \leq E[e^{\varepsilon\|X\|^2}]\exp(-\varepsilon t^2) = k\exp(-\varepsilon t^2). \tag{0.2}$$

Bounds of the type of (0.1) or (0.2) are what we call exponential bounds. In this paper we derive an exponential bound for the distribution of $\|X\|$, where X is a subgaussian $\mathbb{R}^n$-valued random vector, and we identify the numbers k and ε of (0.2).

Our bound will appear as a generalization of (0.1); it can be used for estimating the tail distribution of $\|X\|$ in various contexts (e.g. in the study of the asymptotic behaviour of subgaussian processes).

1. THE MAIN RESULT

Let X be a random vector taking its values in $\mathbb{R}^n$, subgaussian in the sense of [3], i.e. we assume that there exists a symmetric positive definite $n \times n$ matrix R such that

$$E[e^{<x,X>}] \leq \exp\left(\frac{1}{2} < Rx, x >\right) \tag{1.1}$$

for all $x \in \mathbb{R}^n$ (we shall say also that X is subgaussian with respect to R).

In what follows, the term "vector" will always mean "column vector". We shall adopt the following conventions.

Let k be an integer, with $0 \leq k \leq n$, and denote by I the set

$$I = \begin{cases} \{i_1, \dots, i_k\} & \text{for } k \geq 1 \\ \emptyset & \text{for } k = 0 \end{cases}, \tag{1.2}$$

[1]This paper is partially supported by GNAFA, CNR and MURST.

where $i_1, \dots, i_k$ are integers such that $1 \leq i_1 < i_2 < \dots < i_k \leq n$.

Consider the $n \times n$ matrix $M_I = \{m_{ij}^{(I)}\}$ with

$$m_{ij}^{(I)} = \begin{cases} 1 & \text{for } i = j \notin I \\ -1 & \text{for } i = j \in I \\ 0 & \text{elsewhere.} \end{cases}$$

Clearly $M_\emptyset = Id$, while in the other cases the action of M_I on any vector $\underline{x} = (x_1, \dots, x_n)^T$ is to change the sign of $x_{i_1}, \dots, x_{i_k}$. When there will be no risk of confusion (i.e. when I is fixed) we shall denote by $\tilde{\underline{x}} = (\tilde{x}_1, \dots, \tilde{x}_n)^T$ the vector $M_I \underline{x}$.

Let now $R = \{r_{ij}\}$ be any $n \times n$ matrix. We shall denote by R_I the matrix $M_I R M_I$. Let $\xi^{(n)}$ be the vector in $\mathbb{R}^n$ whose components are all 1. We shall say that R has property P iff for every I, the vector $R_I^{-1}\xi^{(n)}$ has all strictly positive components and the same happens for each of the diagonal submatrices of R (obviously, for each submatrix, I is a subset of the set of its indexes; moreover, the vector with all components 1 and M_I have the suitable order, i.e. the same as the submatrix).

The main result of this paper is the following

Proposition 1.3. *Let $X = (X_1, \dots, X_n)^T$ be subgaussian with respect to R and assume that R has property P; then, for every $t > 0$, we have*

$$P(\|X\| > t) \leq \frac{3^n + (-1)^{n-1}}{2} \exp\left(-\frac{t^2}{2n}\alpha\right) \tag{1.4}$$

where α is a number, depending only on R, that will be identified in the course of the proof.

We shall use the following

Lemma 1.5. *Assume that X is subgaussian with respect to R. Let $\xi^{(n)}$ be the vector in $\mathbb{R}^n$ whose components are all 1, and assume that, for every I, the vector $R_I^{-1}\xi^{(n)}$ has all strictly positive components. Then, for every $t > 0$, we have*

$$P(|X_1| > t, \dots, |X_n| > t) \leq 2^n \exp\left(-\frac{t^2}{2}\beta\right) \tag{1.6}$$

where

$$\beta = \min_I < R_I^{-1}\xi^{(n)}, \xi^{(n)} >.$$

For the proof of (1.5), we need another

Lemma 1.7. *Assume that X and R are as in (1.5), and put $\tilde{X} = (\tilde{X}_1, \dots, \tilde{X}_n)^T = M_I X$.*

Fix $u = (u_1, \dots, u_n)^T$ with $u_i \geq 0$ for every $i = 1, \dots, n$, and assume that $R_I^{-1}u$ is a vector with all positive components. Then

$$P(\tilde{X}_1 > u_1, \dots, \tilde{X}_n > u_n) \leq \exp\left(-\frac{1}{2} < R_I^{-1}u, u >\right).$$

Proof of (1.7). For every $\lambda = (\lambda_1, \dots, \lambda_n)$ with $\lambda_i > 0$ for every $i = 1, \dots, n$ we have

$$P(\tilde{X}_1 > u_1, \dots, \tilde{X}_n > u_n) = P(e^{\lambda_1 \tilde{X}_1} > e^{\lambda_1 u_1}, \dots, e^{\lambda_{n\bar{n}} \tilde{X}_n} > e^{\lambda_n u_n}) \leq$$
$$\leq E[e^{<\lambda . \tilde{X}>}] \exp(- <\lambda, u>) = E[e^{<\lambda . M_I X>}] \exp(- <\lambda, u>) =$$
$$= E[e^{<M_I \lambda . X>}] \exp(- <\lambda, u>) \leq \exp\left(\frac{1}{2} <RM_I\lambda, M_I\lambda> - <\lambda, u>\right) =$$
$$= \exp\left(\frac{1}{2} <M_I R M_I \lambda, \lambda> - <\lambda, u>\right) = \exp\left(\frac{1}{2} <R_I \lambda, \lambda> - <\lambda, u>\right).$$

By minimizing in λ, we find that the minimum of the last quantity is attained in $\lambda = R_I^{-1} u$, and is equal to

$$\exp\left(\frac{1}{2} <R_I^{-1} u, u> - <R_I^{-1} u, u>\right) = \exp\left(-\frac{1}{2} <R_I^{-1} u, u>\right).$$

□

Remark 1.8. The above lemma is proved in [4] in the particular case $I = \emptyset$.

Proof of (1.5). By writing

$$\{|X_i| > t\} = \{X_i > t\} \cup \{-X_i > t\},$$

it is easy to see that the probability in (1.6) can be split into sum of 2^n terms; each of them is of the form

$$P(M_I X \in (t, +\infty)^n), \tag{1.9}$$

where I is a suitable set of indexes, of the type considered at the beginning of this section.

Suppose now I fixed, so that (1.9) can be written in the more understandable form

$$P(\tilde{X}_1 > t, \dots, \tilde{X}_n > t) \tag{1.10}$$

and let $\underline{t}$ be the vector in $\mathbb{R}^n$ with all components equal to t.

Then $\underline{t} = t\xi^{(n)}$, and, by lemma (1.7), (1.10) is not greater than

$$\exp\left(-\frac{1}{2} <R_I^{-1} \underline{t}, \underline{t}>\right) = \exp\left(-\frac{t^2}{2} <R_I^{-1} \xi^{(n)}, \xi^{(n)}>\right) \leq \exp\left(-\frac{t^2}{2} \beta\right).$$

□

We are now in a position to prove (1.3).

Let C_t be the closed ball in $\mathbb{R}^n$ centered at the origin and having radius t, and Q_t the cube

$$Q_t = \left\{(x_1, \dots, x_n) : |x_i| \leq \frac{t}{\sqrt{n}} \text{ for every } i = 1, \dots, n\right\}.$$

Then $Q_t \subset C_t$, so that

$$P(\|X\| > t) = P(X \in C_t^C) \leq P(X \in Q_t^C) \leq P\left(\bigcup_{i=1}^{n}\left\{|X_i| > \frac{t}{\sqrt{n}}\right\}\right). \tag{1.11}$$

By the inclusion-exclusion formula, the last probability in (1.11) is not greater than

$$\sum_{i=1}^{n} P\left(|X_i| > \frac{t}{\sqrt{n}}\right) + \sum_{1 \leq i < j < k \leq n} P\left(|X_i| > \frac{t}{\sqrt{n}}, |X_j| > \frac{t}{\sqrt{n}}, |X_k| > \frac{t}{\sqrt{n}}\right) + \ldots \tag{1.12}$$

In order to make our reasoning as easy as possible, we focus our attention, for a moment, on the term

$$P\left(|X_1| > \frac{t}{\sqrt{n}}, |X_2| > \frac{t}{\sqrt{n}}, |X_3| > \frac{t}{\sqrt{n}}\right).$$

It is easy to see that the 3-dimensional vector (X_1, X_2, X_3) is subgaussian with respect to the 3×3 submatrix of R obtained by cancelling all rows and columns in it, except the ones having indexes $1, 2, 3$. Hence we can apply lemma (1.5) to the vector (X_1, X_2, X_3), and we get

$$P\left(|X_1| > \frac{t}{\sqrt{n}}, |X_2| > \frac{t}{\sqrt{n}}, |X_3| > \frac{t}{\sqrt{n}}\right) \leq 2^3 \exp\left(-\frac{t^2}{2n}\beta_{1,2,3}\right),$$

where $\beta_{1,2,3}$ is defined as β of lemma (1.5).

The above argument applies to every vector of the form (X_i, X_j, X_k), so that we can write

$$\sum_{1 \leq i < j < k \leq n} P\left(|X_i| > \frac{t}{\sqrt{n}}, |X_j| > \frac{t}{\sqrt{n}}, |X_k| > \frac{t}{\sqrt{n}}\right) \leq$$

$$\leq \sum_{1 \leq i < j < k \leq n} 2^3 \exp\left(-\frac{t^2}{2n}\beta_{i,j,k}\right) \leq \binom{n}{3} 2^3 \exp\left(-\frac{t^2}{2n}\beta_3\right),$$

where

$$\beta_3 = \min_{i,j,k} \beta_{i,j,k}.$$

One can reason the same way for each sum appearing in (1.12), and obtains that (1.12) is not greater than

$$\sum_{k=0}^{\left[\frac{n-1}{2}\right]} \binom{n}{2k+1} 2^{2k+1} \exp\left(-\frac{t^2}{2n}\beta_{2k+1}\right).$$

Put

$$\alpha = \min_k \beta_{2k+1}.$$

Then the above sum is majorized by

$$\left(\sum_{k=0}^{\left[\frac{n-1}{n}\right]} \binom{n}{2k+1} 2^{2k+1}\right) \exp\left(-\frac{t^2}{2n}\alpha\right) = \frac{3^n + (-1)^{n-1}}{2} \exp\left(-\frac{t^2}{2n}\alpha\right),$$

where the last equality is easily proved by induction on n. □

REFERENCES

[1] V. V. Buldygin, Yu. V. Kozachenko, *Sub-Gaussian random variables*, Ukrainian Math. J., **32** (1980), 483-489.
[2] R. Fukuda, *Exponential integrability of sub-Gaussian vectors*, Probab. Theory Related Fields, **85** (1990), 505-521.
[3] E. I. Ostrovskii, *Exponential Estimates for the Distribution of the Maximum of a non - Gaussian Random Field*, Theory Probab. Appl., **35** (1980), 487-499.
[4] V. V. Buldygin, Yu. V. Kozachenko, *Sub-Gaussian random vectors and processes*, Theory Probab. Math. Statist., **36** (1988), 9-20.

R. Giuliano Antonini
Dipartimento di Matematica
Universitá di Pisa
Via F. Buonarroti 2
56100 Pisa
ITALY