9-19-2016

On a multidimensional spherically invariant extension of the Rademacher-Gaussian comparison

Iosif Pinelis
Michigan Technological University

Follow this and additional works at: https://digitalcommons.mtu.edu/math-fp

Part of the Mathematics Commons

Electron. Commun. Probab. **21** (2016), no. 67, 1–5.
DOI: 10.1214/16-ECP23
ISSN: 1083-589X

ELECTRONIC COMMUNICATIONS in PROBABILITY

On a multidimensional spherically invariant extension of the Rademacher–Gaussian comparison

Iosif Pinelis*

Abstract

It is shown that

$$\mathsf{P}(\|a_1U_1+\cdots+a_nU_n\|>u)\leqslant c\,\mathsf{P}(a\|Z_d\|>u)$$

for all real u, where $U_1,\dots,U_n$ are independent random vectors uniformly distributed on the unit sphere in $\mathbb{R}^d$, $a_1,\dots,a_n$ are any real numbers, $a:=\sqrt{(a_1^2+\cdots+a_n^2)/d}$, Z_d is a standard normal random vector in $\mathbb{R}^d$, and $c=2e^3/9=4.46\dots$. This constant factor is about 89 times as small as the one in a recent result by Nayar and Tkocz, who proved, by a different method, a corresponding conjecture by Oleszkiewicz. As an immediate application, a corresponding upper bound on the tail probabilities for the norm of the sum of arbitrary independent spherically invariant random vectors is given.

Keywords: probability inequalities; generalized moment comparison; tail comparison; sums of independent random vectors; Gaussian random vectors; uniform distribution on spheres.
AMS MSC 2010: Primary 60E15, Secondary 60G15; 60G50.
Submitted to ECP on April 7, 2016, final version accepted on September 7, 2016.

In this note, we shall present an upper bound on the tail probability for the Euclidean norm of the weighted sum of independent random vectors distributed uniformly on the unit sphere in $\mathbb{R}^d$ that is about 89 times as small as the corresponding bound recently obtained in [8].

To provide a relevant context, let us begin by introducing the class C^2_{conv} of all even twice differentiable functions $h\colon\mathbb{R}\to\mathbb{R}$ whose second derivative h'' is convex. Let $\varepsilon,\varepsilon_1,\dots,\varepsilon_n$ be independent Rademacher random variables (r.v.'s), and let $\xi_1,\dots,\xi_n$ be any independent symmetric r.v.'s with $\mathsf{E}\,\xi_i^2=1$ for all i.

Take any natural d. For any vectors x and y in $\mathbb{R}^d$, let, as usual, $x\cdot y$ denote the standard inner product of x and y, and then let $\|x\|:=\sqrt{x\cdot x}$.

Theorem 2.3 in [9] states that $\mathsf{E}\,h\big(\sqrt{\boldsymbol{\varepsilon}A\boldsymbol{\varepsilon}^T}\big)\leqslant\mathsf{E}\,h\big(\sqrt{\boldsymbol{\xi}A\boldsymbol{\xi}^T}\big)$ for any $h\in C^2_{\mathsf{conv}}$ and any nonnegative definite $n\times n$ matrix $A\in\mathbb{R}^{n\times n}$, where $\boldsymbol{\varepsilon}:=[\varepsilon_1,\dots,\varepsilon_n]$ and $\boldsymbol{\xi}:=[\xi_1,\dots,\xi_n]$. This can be restated as the following generalized moment comparison:

$$\mathsf{E}\,h(\|\varepsilon_1x_1+\cdots+\varepsilon_nx_n\|)\leqslant\mathsf{E}\,h(\|\xi_1x_1+\cdots+\xi_nx_n\|)\tag{1}$$

for any $h\in C^2_{\mathsf{conv}}$ and any (nonrandom) vectors $x_1,\dots,x_n$ in $\mathbb{R}^d$; indeed, any nonnegative definite matrix $A\in\mathbb{R}^{n\times n}$ is the Gram matrix of some vectors $x_1,\dots,x_n$ in $\mathbb{R}^d$ for some natural d, and then $\|\alpha_1x_1+\cdots+\alpha_nx_n\|=\sqrt{\boldsymbol{\alpha}A\boldsymbol{\alpha}^T}$ for any $\boldsymbol{\alpha}:=[\alpha_1,\dots,\alpha_n]\in\mathbb{R}^{1\times n}$. From the comparison (1) of generalized moments of the r.v.'s $\|\varepsilon_1x_1+\cdots+\varepsilon_nx_n\|$ and

*Michigan Technological University, United States of America. E-mail: ipinelis@mtu.edu

$\|\xi_1 x_1 + \cdots + \xi_n x_n\|$, a tail comparison was extracted ([9, Theorem 2.4]), an equivalent form of which is the inequality

$$\mathsf{P}(\|\varepsilon_1 x_1 + \cdots + \varepsilon_n x_n\| > u) < c\,\mathsf{P}(\|Z_r\| > u) \tag{2}$$

for all real u, where $x_1, \ldots, x_n$ are any (nonrandom) vectors in $\mathbb{R}^d$ whose Gram matrix is an orthoprojector of rank r, Z_r is a standard normal random vector in $\mathbb{R}^r$, and

$$c = c_3 := 2e^3/9 = 4.46\ldots. \tag{3}$$

A special case of (2) is the inequality

$$\mathsf{P}(|\varepsilon_1 a_1 + \cdots + \varepsilon_n a_n| > u) \leqslant c\,\mathsf{P}(|Z_1| > u) \tag{4}$$

for all real u, where $a_1, \ldots, a_n$ are any real numbers such that

$$a_1^2 + \cdots + a_n^2 = 1.$$

The quoted results generalize and refine results of [4, 5]. In turn, they were further developed in [10, 11].

A simple inductive argument, which was direct rather than based on a generalized moment comparison, was offered in [3], where (4) was proved with $c \approx 12.01$. Based in part on that inductive argument in [3], the constant c in (4) was improved to $\approx 1.01c_*$ in [13] and then to c_* in [2], where $c_* := \mathsf{P}(|\varepsilon_1 + \varepsilon_2| \geqslant 2)/\,\mathsf{P}(|Z_1| \geqslant \sqrt{2}) = 3.17\ldots$, so that c_* is the best possible value of c in (4).

In [1], another kind of multidimensional generalized moment comparison was obtained. A continuous function $f\colon \mathbb{R}^d \to \mathbb{R}$ is called bisubharmonic if the (Sobolev–Schwartz) distribution $\Delta^2 f$ is a nonnegative Radon measure on $\mathbb{R}^d$, where Δ is the Laplace operator on $\mathbb{R}^d$. By [1, Theorem 3], for any continuous function $f\colon \mathbb{R}^d \to \mathbb{R}$ one has

$$f \text{ is bisubharmonic if and only if } \mathsf{E}\, f(y + U\sqrt{t}) \text{ is convex in } t \in (0, \infty) \text{ for each } y \in \mathbb{R}^d, \tag{5}$$

where U is a random vector uniformly distributed on the unit sphere S^{d-1} in $\mathbb{R}^d$.

Let $U_1, \ldots, U_n$ be independent copies of U. Theorem 1 in [1] states that

$$\mathsf{E}\, f(a_1 U_1 + \cdots + a_n U_n) \leqslant \mathsf{E}\, f(b_1 U_1 + \cdots + b_n U_n), \tag{6}$$

where f is a bisubharmonic function and $a_1, \ldots, a_n, b_1, \ldots, b_n$ are real numbers such that the n-tuple $(b_1^2, \ldots, b_n^2)$ is majorized by $(a_1^2, \ldots, a_n^2)$ in the sense of the Schur majorization (see e.g. [7]).

One may note that, whereas in (1) each of the random summands $\varepsilon_1 x_1, \ldots, \varepsilon_n x_n$, $\xi_1 x_1, \ldots, \xi_n x_n$ is distributed on a straight line through the origin, each of the random summands $a_1 U_1, \ldots, a_n U_n, b_1 U_1, \ldots, b_n U_n$ in (6) is uniformly distributed on a sphere centered at the origin.

Since the distributions of the random vectors $a_1 U_1 + \cdots + a_n U_n$ and $b_1 U_1 + \cdots + b_n U_n$ are clearly spherically invariant, without loss of generality one may assume that the function f in (6) is spherically invariant as well, that is, $f(x)$ depends on $x \in \mathbb{R}^d$ only through $\|x\|$. If f is indeed a spherically invariant bisubharmonic function, it then follows from (6) and [1, formulas (1.2), (1.3)] that

$$\mathsf{E}\, f(a_1 U_1 + \cdots + a_n U_n) \leqslant \mathsf{E}\, f(a Z_d), \tag{7}$$

where

$$a := \sqrt{(a_1^2 + \cdots + a_n^2)/d}; \tag{8}$$

cf. [1, Corollary 1].

Let $C^2_{\mathsf{conv}}(H)$ denote the class of all spherically invariant twice differentiable functions f from a Hilbert space H to $\mathbb{R}$ whose second derivative f'' is convex in the sense that the function $H \ni x \mapsto f''(x; y, y)$ is convex for each $y \in H$, where $f''(x; y, y)$ is the value of the second derivative of the function $\mathbb{R} \ni t \mapsto f(x + ty)$ at $t = 0$. The class $C^2_{\mathsf{conv}}(H)$ was characterized in [12], with some applications. Clearly, $C^2_{\mathsf{conv}}(\mathbb{R})$ coincides with the class C^2_{conv} defined in the beginning of this note.

K. Oleszkiewicz conjectured [8] that

$$\mathsf{P}(\|a_1 U_1 + \cdots + a_n U_n\| > u) \leqslant c\, \mathsf{P}(a\|Z_d\| > u) \tag{9}$$

for some universal constant c and all real u, where $a_1, \ldots, a_n, a, U_1, \ldots, U_n, Z_d$ are as before; clearly, (9) is a generalization of (4). This conjecture was proved in [8] with $c = 397$ based, in part, on the idea from [3].

Using inequality (2.6) in [9], one can improve the lower bound $1/397$ in [8, Lemma 1] to $1/e^2$ and thus improve the constant c in (9) from 397 to $e^2 = 7.38\ldots$. Indeed, let, as usual, Φ denote the standard normal distribution function. Then, by inequality (2.6) in [9], $g(d) := \mathsf{P}(\|Z_d\| \geqslant \sqrt{d+2}\,) > 1 - \Phi\big((\sqrt{d+2} - \sqrt{d-1}\,)\sqrt{2}\,\big) =: q(d)$, which latter is clearly increasing in d, with $q(4) > 1/e^2$, whence $g(d) > 1/e^2$ for $d = 4, 5, \ldots$, whereas $g(2) = 1/e^2 < g(3)$. So, $\mathsf{P}(\|Z_d\| \geqslant \sqrt{d+2}\,) = g(d) \geqslant 1/e^2$ for $d = 2, 3, \ldots$. Similarly, $\mathsf{P}(\|Z_d\| \geqslant \sqrt{d}\,) \geqslant 1/e$ for $d = 2, 3, \ldots$ (but a lower bound on $\mathsf{P}(\|Z_d\| \geqslant \sqrt{d}\,)$ is not really needed in the proof of the main result in [8]).

The aim of this note is to point out that, based on the generalized moment comparison (7) and results in [9, 10], one can further improve the constant c in (9):

Theorem 1. Inequality (9) holds (for all real u) with c as in (3). The strict version of (9), again with c as in (3), also holds.

Our method is quite different from that of [8]. In view of (7), Theorem 1 is an immediate corollary of the following two lemmas.

Lemma 1. For any function $h \in C^2_{\mathsf{conv}}$, the function $f\colon \mathbb{R}^d \to \mathbb{R}$ defined by the formula $f(x) := h(\|x\|)$ for $x \in \mathbb{R}^d$ is a spherically invariant bisubharmonic function.

Lemma 2. Let ξ be any nonnegative r.v. such that

$$\mathsf{E}\, h(\xi) \leqslant \mathsf{E}\, h(\|Z_d\|) \quad \text{for all} \quad h \in C^2_{\mathsf{conv}}. \tag{10}$$

Then

$$\mathsf{P}(\xi > u) < c_3\, \mathsf{P}(\|Z_d\| > u) \tag{11}$$

for all real u, with c_3 defined in (3).

Proof of Lemma 1. Let U be as in (5) and then let ε be a Rademacher r.v. independent of U. For all $t \in (0, \infty)$ and $y \in \mathbb{R}^d$

$$\mathsf{E}\, f(y + U\sqrt{t}) = \mathsf{E}\, f(y + \varepsilon U\sqrt{t}) = \mathsf{E}\, h(\|y + \varepsilon U\sqrt{t}\|) = \mathsf{E}\,\mathsf{E}_U\, g_{b_U,h}(\beta_U + \varepsilon\sqrt{t}), \tag{12}$$

where E_U denotes the conditional expectation given U, $g_{b,h}(u) := h\big(\sqrt{u^2 + b}\,\big)$ for $b \in [0, \infty)$ and $u \in \mathbb{R}$, $\beta_U := y \cdot U$, and $b_U := \|y\|^2 - (y \cdot U)^2 \geqslant 0$, so that the r.v. ε is independent of the pair (b_U, β_U), which latter is a function of U. By [9, Lemma 3.1], $g_{b,h} \in C^2_{\mathsf{conv}}$ for each $b \in [0, \infty)$. Hence, by [14, Lemma 3.1] or [9, Proposition A.1], $\mathsf{E}_U\, g_{b_U,h}(\beta_U + \varepsilon\sqrt{t})$ is convex in $t \in (0, \infty)$. So, in view of (12), $\mathsf{E}\, f(y + U\sqrt{t})$ is convex in $t \in (0, \infty)$. Now it follows by (5) that the function f is indeed bisubharmonic. That f is spherically invariant is trivial. □

Proof of Lemma 2. Taken almost verbatim, the proof of Theorem 2.4 in [9] (based on Theorem 2.3 in [9]) can also serve as a proof of Lemma 2. Indeed, no properties of the r.v. $\varepsilon\Pi\varepsilon^T$ were used in the proof of [9, Theorem 2.4] except that this nonnegative r.v. satisfies the inequality in [9, Theorem 2.3] with $A = \Pi$ and $\xi = Z_n$, which can then be written as (10) with $\xi = \sqrt{\varepsilon\Pi\varepsilon^T}$ and d equal the rank of Π. (Note here a typo in [9]: in place of "Theorem 2.3" in line 7- on page 363 there, it should be "Theorem 2.4".)

Instead of following the entire proof of [9, Theorem 2.4], one can alternatively reason as follows. Let ξ be any nonnegative r.v. such that (10) holds. Then [9, Lemma 3.5] holds with ξ^2 in place of $\boldsymbol{\varepsilon}\Pi\boldsymbol{\varepsilon}^T$. So, in view of [9, formula (3.11)] and [10, formula (22) in Theorem 3.11], inequality (11) holds for $u \geqslant \mu_r$, with $r := d$ and μ_r defined on page 362 in [9]. The cases $r^{1/2} \leqslant u \leqslant \mu_r$ and $0 \leqslant u \leqslant r^{1/2}$ are considered as was done at the end of the proof of [9, Lemma 3.6], starting at the middle of page 365 in [9]. The case $u < 0$ is trivial. $\square$

An immediate application of Theorem 1 is

Corollary 1. Let $X_1, \dots, X_n$ be any independent spherically invariant random vectors in $\mathbb{R}^d$, which are also independent of the Gaussian random vector Z_d. Then

$$\mathsf{P}(\|X_1 + \dots + X_n\| > u) < \frac{2e^3}{9}\, \mathsf{P}\left(\sqrt{\|X_1\|^2 + \dots + \|X_n\|^2}\, \|Z_d\| > u\right) \tag{13}$$

for all real u.

This corollary follows from Theorem 1 by the conditioning on $\|X_1\|, \dots, \|X_n\|$, because for each $i = 1, \dots, n$ the conditional distribution of the spherically invariant random vector X_i given $\|X_i\| = a_i$ is the distribution of $a_i U_i$.

In the case when the independent spherically invariant random vectors $X_1, \dots, X_n$ are bounded almost surely by positive real numbers $b_1, \dots, b_n$, respectively, one can obviously replace $\sqrt{\|X_1\|^2 + \dots + \|X_n\|^2}$ in the bound in (13) by $\sqrt{b_1^2 + \dots + b_n^2}$. The resulting bound, but with the constant factor 397 in place of $\frac{2e^3}{9} = 4.46\dots$, was obtained in [8].

Similarly to the extension (13) of inequality (9), one can extend (7) as follows:

$$\mathsf{E}\, f(X_1 + \dots + X_n) \leqslant \mathsf{E}\, f\left(\sqrt{\|X_1\|^2 + \dots + \|X_n\|^2}\; Z_d\right) \tag{14}$$

for any spherically invariant bisubharmonic function f, where $X_1, \dots, X_n$ are as in Corollary 1.

A related result was obtained in [6]: if $X_1, \dots, X_n$ are independent identically distributed spherically invariant random vectors in $\mathbb{R}^d$ such that $\mathsf{E}\, h(\|X_1\|^2) \leqslant \mathsf{E}\, h(\|Z_d\|^2)$ for all nonnegative convex functions $h\colon \mathbb{R} \to \mathbb{R}$, then

$$\mathsf{E}\, \|a_1 X_1 + \dots + a_n X_n\|^p \leqslant \mathsf{E}\, \|a Z_d \sqrt{d}\|^p \tag{15}$$

for real $p \geqslant 3$, where $a_1, \dots, a_n, a$ are as in (7)–(8).

References

[1] A. Baernstein, II and R. C. Culverhouse. Majorization of sequences, sharp vector Khinchin inequalities, and bisubharmonic functions. *Studia Math.*, 152(3):231–248, 2002. MR-1916226

[2] V. K. Bentkus and D. Dzindzalieta. A tight Gaussian bound for weighted sums of Rademacher random variables. *Bernoulli*, 21(2):1231–1237, 2015. MR-3338662

[3] S. G. Bobkov, F. Götze, and C. Houdré. On Gaussian and Bernoulli covariance representations. *Bernoulli*, 7(3):439–451, 2001. MR-1836739

[4] M. L. Eaton. A note on symmetric Bernoulli random variables. *Ann. Math. Statist.*, 41:1223–1226, 1970. MR-0268930

[5] M. L. Eaton. A probability inequality for linear combinations of bounded random variables. *Ann. Statist.*, 2:609–613, 1974.

[6] H. König and S. Kwapień. Best Khintchine type inequalities for sums of independent, rotationally invariant random vectors. *Positivity*, 5(2):115–152, 2001. MR-1825172

[7] A. W. Marshall and I. Olkin. *Inequalities: theory of majorization and its applications*, volume 143 of *Mathematics in Science and Engineering*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1979. MR-0552278

[8] P. Nayar and T. Tkocz. A multidimensional analogue of the Rademacher-Gaussian tail comparison. 2016. arXiv:1602.07995 [math.PR].

[9] I. Pinelis. Extremal probabilistic problems and Hotelling's T^2 test under a symmetry condition. *Ann. Statist.*, 22(1):357–368, 1994. MR-1272088

[10] I. Pinelis. Optimal tail comparison based on comparison of moments. In *High dimensional probability (Oberwolfach, 1996)*, volume 43 of *Progr. Probab.*, pages 297–314. Birkhäuser, Basel, 1998. MR-1652335

[11] I. Pinelis. Fractional sums and integrals of r-concave tails and applications to comparison probability inequalities. In *Advances in stochastic inequalities (Atlanta, GA, 1997)*, volume 234 of *Contemp. Math.*, pages 149–168. Amer. Math. Soc., Providence, RI, 1999. MR-1694770

[12] I. Pinelis. Spherically symmetric functions with a convex second derivative and applications to extremal probabilistic problems. *Math. Inequal. Appl.*, 5(1):7–26, 2002. MR-1880267

[13] I. Pinelis. Toward the best constant factor for the Rademacher-Gaussian tail comparison. *ESAIM Probab. Stat.*, 11:412–426, 2007. MR-2339301

[14] S. A. Utev. Extremal problems in moment inequalities. In *Limit theorems of probability theory*, volume 5 of *Trudy Inst. Mat.*, pages 56–75, 175. "Nauka" Sibirsk. Otdel., Novosibirsk, 1985. MR-0821753