

E. A. Gorbunov, E. Vorontsova, A. V. Gasnikov, On the Upper Bound for the Expectation of the Norm of a Vector Uniformly Distributed on the Sphere and the Phenomenon of Concentration of Uniform Measure on the Sphere, *Mat. Zametki*, 2019, Volume 106, Issue 1, 13–23

DOI: 10.4213/mzm12041

Use of the all-Russian mathematical portal Math-Net.Ru implies that you have read and agreed to these terms of use

<http://www.mathnet.ru/eng/agreement>

Download details:

IP: 176.57.72.123

August 22, 2023, 20:36:48





УДК 519.21

О верхней оценке математического ожидания нормы равномерно распределенного на сфере вектора и явлении концентрации равномерной меры на сфере

Э. А. Горбунов, Е. А. Воронцова, А. В. Гасников

Рассматривается задача построения верхних оценок математического ожидания нормы равномерно распределенного на единичной евклидовой сфере вектора.

Библиография: 15 названий.

Ключевые слова: концентрация меры, равномерно распределенный на сфере вектор.

DOI: <https://doi.org/10.4213/mzm12041>

1. Введение. Пусть $e \in RS_2^n(1)$ – случайный вектор, имеющий равномерное распределение на n -мерной единичной евклидовой сфере.

В настоящей работе рассматривается задача построения уточненных верхних оценок математического ожидания нормы вектора e . От точности построения таких оценок зависят, например, оценки скорости сходимости ускоренного метода (Accelerated by Coupling Directional Search – ACDS), построенного на базе специального каплинга спусков по направлению в форме градиентного спуска и метода зеркального спуска [1].

Первые такие оценки были получены в 2014 г. в [2], статья опубликована в 2016 году в [3]. Кроме того, независимо от работы [3] в 2015 г. похожие оценки были сделаны в [4], опубликовано в 2017 г. в [5]; и, также независимо, в 2015 г. одна из возможных оценок математического ожидания нормы равномерно распределенного на единичной евклидовой сфере вектора была получена в [6].

Основным результатом данной работы является теорема 1.

2. Постановка задачи и формулировка результата. Пусть задан некоторый (неслучайный) вектор s с единичной евклидовой сферы. Не умаляя общности, будем считать, что вектор s направлен вдоль первой координатной оси (если это не так, то мы можем перейти к нужному базису). Тогда с вероятностью хотя бы $1 - (2/c)e^{-c^2/2}$ будет выполнено неравенство $|\langle s, e \rangle| \leq c/\sqrt{n-1}$ (см. теорему 2.7

Работа А. В. Гасникова поддержана программой Президента РФ (грант № МД-1320.2018.1). Работа Э. А. Горбунова поддержана Российским фондом фундаментальных исследований (грант № 18-31-20005 мол_a_вед). Работа Е. А. Воронцовой поддержана Российским фондом фундаментальных исследований (грант № 18-29-03071).

и рисунок 2.2 из [7] и [8]). То есть, если взять $c = 10$, то получим, что с большой вероятностью выполнено неравенство $\langle s, e \rangle^2 \leq 100/n$ (множество, на котором $\langle s, e \rangle^2 \leq 100/n$, обозначим через A_s ; как мы видим, при достаточно больших n вероятностная мера множества A_s велика). Кроме того, можно показать, что $\mathbb{E}[\langle s, e \rangle^2] = 1/n$ (см., например, лемму В.10 из [9]).

Рассмотрим ∞ -норму, которая для произвольного вектора $x \in \mathbb{R}^n$ задается формулой

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad \text{где } x = (x_1, x_2, \dots, x_n)^\top.$$

Заметим, что функция $f(e) = \|e\|_\infty$ является липшицевой с константой 1 в евклидовой норме. Рассмотрим константу M_f такую, что

$$\mathbb{P}_e\{f(e) \geq M_f\} \geq \frac{1}{2}, \quad \mathbb{P}_e\{f(e) \leq M_f\} \geq \frac{1}{2}.$$

Тогда верно неравенство (см. [10], [11])

$$\mathbb{P}_e\{|f(e) - M_f| > t\} \leq 4e^{-t^2/4}, \quad t > 0.$$

Это означает, что случайная величина $\|e\|_\infty$ принимает очень близкие к $\mathbb{E}[\|e\|_\infty]$ (M_f и $\mathbb{E}[f(e)]$ асимптотически близки, см. [12]) значения на множестве достаточно большой меры. Кроме того, можно показать, что максимальная по модулю компонента вектора e с вероятностью не меньше $1 - 1/(n\sqrt{n})$ принимает значения по модулю меньше $2\sqrt{\ln n}/\sqrt{n-1}$ (множество, на котором $\|e\|_\infty \leq 2\sqrt{\ln n}/\sqrt{n-1}$, обозначим через B_∞). Тогда $\mathbb{E}[\langle s, e \rangle^2 \|e\|_\infty^2]$ близко к среднему значению случайной величины $\langle s, e \rangle^2 \|e\|_\infty^2$ на множестве $A_e \cap B_\infty$ (чья вероятностная мера по-прежнему велика), на котором она не превосходит $400 \ln n/n^2$. Константа в этой оценке сильно завышена и она уточняется далее в теореме 1 (причем не только для ∞ -нормы). Однако такого рода рассуждения, вытекающие из явления концентрации равномерной меры на сфере, поясняют причины возникновения такой оценки, а также ее целесообразность в терминах вхождения размерности пространства n .

Сформулируем и докажем достаточно известный факт, заключающийся в том, что векторная q -норма является невозрастающей функцией от q для любого фиксированного вектора.

ЛЕММА 1. *Для любого вектора $x \in \mathbb{R}^n$ (и для любого $n \in \mathbb{N}$) выполнено неравенство:*

$$\|x\|_{p_1} \leq \|x\|_{p_2}, \tag{2.1}$$

где $p_1 \geq p_2$ и под знаком $\|\cdot\|_q$ понимается векторная q -норма (норма Гёльдера с показателем q).

ДОКАЗАТЕЛЬСТВО. Не умаляя общности, будем считать, что все компоненты вектора x являются ненулевыми (если вектор x нулевой, то его норма равна норме его подвектора меньшей размерности, полученного удалением нулевых компонент, и, соответственно, можно рассматривать этот подвектор; если же вектор $x = 0$, то неравенство (2.1) тоже верно). Пусть

$$g_x(p) \stackrel{\text{def}}{=} \ln \|x\|_p = \ln \left(\sum_{k=1}^n |x_k|^p \right)^{1/p} = \frac{1}{p} \ln \left(\sum_{k=1}^n |x_k|^p \right).$$

Тогда

$$\frac{dg_x(p)}{dp} = -\frac{1}{p^2} \ln\left(\sum_{k=1}^n |x_k|^p\right) + \frac{1}{p} \cdot \frac{\sum_{k=1}^n \ln(|x_k|) \cdot |x_k|^p}{\sum_{k=1}^n |x_k|^p}.$$

Так как $\ln y$ – вогнутая по y функция, по неравенству Йенсена получаем

$$\begin{aligned} \frac{dg_x(p)}{dp} &\leq \frac{1}{p} \ln\left(\sum_{k=1}^n |x_k|^p\right)^{-1/p} + \frac{1}{p} \ln\left(\sum_{k=1}^n |x_k| \cdot \frac{|x_k|^p}{\sum_{k=1}^n |x_k|^p}\right) \\ &= \frac{1}{p} \ln\left(\frac{\sum_{k=1}^n |x_k|^{p+1}}{(\sum_{k=1}^n |x_k|^p)^{(p+1)/p}}\right) \leq \frac{1}{p} \ln\left(\frac{\sum_{k=1}^n |x_k|^{p+1}}{\sum_{k=1}^n (|x_k|^p)^{(p+1)/p}}\right) = 0, \end{aligned}$$

то есть функция $g_x(p)$ – невозрастающая функция на $[1, +\infty)$. Лемма доказана.

Имеет место следующая теорема, являющаяся следствием явления концентрации равномерной меры на сфере вокруг экватора (см. также [13]; северный полюс задается градиентом $\nabla f(x)$).

ТЕОРЕМА 1. Пусть $e \in RS_2^n(1)$, $n \geq 8$, $s \in \mathbb{R}^n$. Тогда

$$\mathbb{E}[\|e\|_q^2] \leq \min\{q - 1, 16 \ln n - 8\} n^{2/q-1}, \quad 2 \leq q \leq \infty, \quad (2.2)$$

$$\mathbb{E}[\langle s, e \rangle^2 \|e\|_q^2] \leq \sqrt{3} \|s\|_2^2 \min\{2q - 1, 32 \ln n - 8\} n^{2/q-2}, \quad 2 \leq q \leq \infty, \quad (2.3)$$

где под знаком $\|\cdot\|_q$ понимается векторная q -норма (норма Гёльдера с показателем q).

ДОКАЗАТЕЛЬСТВО. Докажем вспомогательное неравенство:

$$\mathbb{E}[\|e\|_q^2] \leq (q - 1) n^{2/q-1}, \quad 2 \leq q < \infty. \quad (2.4)$$

Во-первых,

$$\mathbb{E}[\|e\|_q^2] = \mathbb{E}\left[\left(\sum_{k=1}^n |e_k|^q\right)^{2/q}\right] \stackrel{(1)}{\leq} \left(\mathbb{E}\left[\sum_{k=1}^n |e_k|^q\right]\right)^{2/q} \stackrel{(2)}{=} (n \mathbb{E}[|e_2|^q])^{2/q}, \quad (2.5)$$

где (1) выполнено в силу вероятностного неравенства Йенсена (функция $\varphi(x) = x^{2/q}$ является вогнутой, так как $q \geq 2$), а переход (2) корректен в силу линейности математического ожидания и одинаковой распределенности компонент вектора e .

Во-вторых, воспользуемся тем (лемма Пуанкаре, см., например, [14; п. 6.3]), что

$$e \stackrel{d}{=} \frac{\xi}{\sqrt{\xi_1^2 + \dots + \xi_n^2}}, \quad (2.6)$$

где $\xi = (\xi_1, \xi_2, \dots, \xi_n)^\top$ – n -мерный гауссовский случайный вектор с нулевым математическим ожиданием и единичной ковариационной матрицей, а “ $\stackrel{d}{=}$ ” обозначает равенство по распределению. Тогда

$$\begin{aligned} \mathbb{E}[|e_2|^q] &= \mathbb{E}\left[\frac{|\xi_2|^q}{(\xi_1^2 + \dots + \xi_n^2)^{q/2}}\right] \\ &= \int \dots \int_{\mathbb{R}^n} |x_2|^q \left(\sum_{k=1}^n x_k^2\right)^{-q/2} \cdot \frac{1}{(2\pi)^{n/2}} \cdot \exp\left(-\frac{1}{2} \sum_{k=1}^n x_k^2\right) dx_1 \dots dx_n. \end{aligned}$$

Перейдем к сферическим координатам:

$$\begin{aligned}
 x_1 &= r \cos \varphi \sin \theta_1 \cdots \sin \theta_{n-2}, \\
 x_2 &= r \sin \varphi \sin \theta_1 \cdots \sin \theta_{n-2}, \\
 x_3 &= r \cos \theta_1 \sin \theta_2 \cdots \sin \theta_{n-2}, \\
 x_4 &= r \cos \theta_2 \sin \theta_3 \cdots \sin \theta_{n-2}, \\
 &\dots\dots\dots \\
 x_n &= r \cos \theta_{n-2}, \\
 &r > 0, \quad \varphi \in [0, 2\pi), \quad \theta_i \in [0, \pi], \quad i = 1, \dots, n-2;
 \end{aligned}$$

якобиан преобразования координат равен

$$\det \left(\frac{\partial(x_1, \dots, x_n)}{\partial(r, \varphi, \theta_1, \theta_2, \dots, \theta_{n-2})} \right) = r^{n-1} \sin \theta_1 (\sin \theta_2)^2 \cdots (\sin \theta_{n-2})^{n-2}.$$

Тогда математическое ожидание $\mathbb{E}[|e_2|^q]$ можно записать в виде

$$\begin{aligned}
 \mathbb{E}[|e_2|^q] &= \int \cdots \int_{\substack{r>0, \varphi \in [0, 2\pi), \\ \theta_i \in [0, \pi], i=1, \dots, n-2}} r^{n-1} |\sin \varphi|^q |\sin \theta_1|^{q+1} |\sin \theta_2|^{q+2} \cdots |\sin \theta_{n-2}|^{q+n-2} \\
 &\quad \times \frac{e^{-r^2/2}}{(2\pi)^{n/2}} dr \cdots d\theta_{n-2} \\
 &= \frac{1}{(2\pi)^{n/2}} I_r \cdot I_\varphi \cdot I_{\theta_1} \cdot I_{\theta_2} \cdots I_{\theta_{n-2}},
 \end{aligned}$$

где

$$\begin{aligned}
 I_r &= \int_0^{+\infty} r^{n-1} e^{-r^2/2} dr, & I_\varphi &= \int_0^{2\pi} |\sin \varphi|^q d\varphi = 2 \int_0^\pi |\sin \varphi|^q d\varphi, \\
 I_{\theta_i} &= \int_0^\pi |\sin \theta_i|^{q+i} d\theta_i, & i &= 1, \dots, n-2.
 \end{aligned}$$

Вычислим эти интегралы. Начнем с I_r :

$$I_r = \int_0^{+\infty} r^{n-1} e^{-r^2/2} d \stackrel{(r=\sqrt{2t})}{=} \int_0^{+\infty} (2t)^{n/2-1} e^{-t} dt = 2^{n/2-1} \Gamma\left(\frac{n}{2}\right).$$

Чтобы вычислить остальные интегралы, будет полезно рассмотреть следующий интеграл ($\alpha > 0$):

$$\begin{aligned}
 \int_0^\pi |\sin \varphi|^\alpha d\varphi &= 2 \int_0^{\pi/2} |\sin \varphi|^\alpha d\varphi \\
 &= 2 \int_0^{\pi/2} (\sin^2 \varphi)^{\alpha/2} d\varphi \stackrel{(t=\sin^2 \varphi)}{=} \int_0^1 t^{(\alpha-1)/2} (1-t)^{-1/2} dt \\
 &= B\left(\frac{\alpha+1}{2}, \frac{1}{2}\right) = \frac{\Gamma(\frac{\alpha+1}{2})\Gamma(\frac{1}{2})}{\Gamma(\frac{\alpha+2}{2})} = \sqrt{\pi} \frac{\Gamma(\frac{\alpha+1}{2})}{\Gamma(\frac{\alpha+2}{2})}.
 \end{aligned}$$

Отсюда получаем, что

$$\begin{aligned}
 \mathbb{E}[|e_2|^q] &= \frac{1}{(2\pi)^{n/2}} I_r \cdot I_\varphi \cdot I_{\theta_1} \cdot I_{\theta_2} \cdot \dots \cdot I_{\theta_{n-2}} \\
 &= \frac{2^{n/2-1}}{(2\pi)^{n/2}} \cdot \Gamma\left(\frac{n}{2}\right) \cdot 2 \frac{\sqrt{\pi} \Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+2}{2})} \cdot \frac{\sqrt{\pi} \Gamma(\frac{q+2}{2})}{\Gamma(\frac{q+3}{2})} \cdot \frac{\sqrt{\pi} \Gamma(\frac{q+3}{2})}{\Gamma(\frac{q+4}{2})} \cdot \dots \cdot \frac{\sqrt{\pi} \Gamma(\frac{q+n-1}{2})}{\Gamma(\frac{q+n}{2})} \\
 &= \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})}.
 \end{aligned} \tag{2.7}$$

Покажем, что для всех $q \geq 2$

$$\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})} \leq \left(\frac{q-1}{n}\right)^{q/2}. \tag{2.8}$$

Сначала убедимся, что неравенство (2.8) выполнено для $q = 2$ (и произвольного n):

$$\frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{2+1}{2})}{\Gamma(\frac{2+n}{2})} - \frac{1}{n} = \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2}) \cdot \frac{1}{2}\Gamma(\frac{1}{2})}{\frac{n}{2}\Gamma(\frac{n}{2})} - \frac{1}{n} = \frac{1}{n} - \frac{1}{n} = 0 \leq 0.$$

Рассмотрим функцию (вообще говоря, двух аргументов)

$$f_n(q) = \frac{1}{\sqrt{\pi}} \cdot \frac{\Gamma(\frac{n}{2})\Gamma(\frac{q+1}{2})}{\Gamma(\frac{q+n}{2})} - \left(\frac{q-1}{n}\right)^{q/2}$$

при $q \geq 2$. Также введем в рассмотрение функцию $\psi(x) = d(\ln(\Gamma(x)))/dx$ при $x > 0$ (*дигамма-функция*). Для гамма-функции выполняется тождество

$$\Gamma(x+1) = x\Gamma(x), \quad x > 0.$$

Возьмем от этого тождества логарифм и продифференцируем по x :

$$\begin{aligned}
 \ln \Gamma(x+1) &= \ln \Gamma(x) + \ln x, \\
 \frac{d(\ln(\Gamma(x+1)))}{dx} &= \frac{d(\ln(\Gamma(x)))}{dx} + \frac{1}{x},
 \end{aligned}$$

что можно записать через дигамма-функцию:

$$\psi(x+1) = \psi(x) + \frac{1}{x}. \tag{2.9}$$

Покажем, что дигамма-функция возрастает при $x > 0$. Для этого докажем неравенство:

$$(\Gamma'(x))^2 < \Gamma(x)\Gamma''(x). \tag{2.10}$$

Действительно,

$$\begin{aligned}
 (\Gamma'(x))^2 &= \left(\int_0^{+\infty} e^{-t} \ln t \cdot t^{x-1} dt \right)^2 \\
 &\stackrel{(1)}{<} \int_0^{+\infty} (e^{-t/2} t^{(x-1)/2})^2 dt \cdot \int_0^{+\infty} (e^{-t/2} t^{(x-1)/2} \ln t)^2 dt \\
 &= \underbrace{\int_0^{+\infty} e^{-t} t^{x-1} dt}_{\Gamma(x)} \cdot \underbrace{\int_0^{+\infty} e^{-t} t^{x-1} \ln^2 t dt}_{\Gamma''(x)},
 \end{aligned}$$

где (1) следует из неравенства Коши–Буняковского (причем неравенство строгое, ибо функции $e^{-t/2}t^{(x-1)/2}$ и $e^{-t/2}t^{(x-1)/2} \ln t$ линейно независимы). Из неравенства (2.10) следует, что

$$\frac{d^2(\ln \Gamma(x))}{dx^2} = \left(\frac{\Gamma'(x)}{\Gamma(x)} \right)' = \frac{\Gamma''(x)}{\Gamma(x)} - \frac{(\Gamma'(x))^2}{(\Gamma(x))^2} \stackrel{(2.10)}{>} 0,$$

т.е. дигамма-функция возрастает.

Теперь покажем, что $f_n(q)$ убывает на отрезке $[2, +\infty)$. Для этого достаточно рассмотреть $\ln(f_n(q))$:

$$\begin{aligned} \ln(f_n(q)) &= \ln\left(\frac{\Gamma\left(\frac{n}{2}\right)}{\sqrt{\pi}}\right) + \ln\left(\Gamma\left(\frac{q+1}{2}\right)\right) - \ln\left(\Gamma\left(\frac{q+n}{2}\right)\right) - \frac{q}{2}(\ln(q-1) - \ln n), \\ \frac{d(\ln(f_n(q)))}{dq} &= \frac{1}{2}\psi\left(\frac{q+1}{2}\right) - \frac{1}{2}\psi\left(\frac{q+n}{2}\right) - \frac{1}{2}\ln(q-1) - \frac{q}{2(q-1)} + \frac{1}{2}\ln n. \end{aligned}$$

Покажем, что $d(\ln(f_n(q)))/dq < 0$ при $q \geq 2$. Пусть $k = \lfloor n/2 \rfloor$ (ближайшее целое число, не превосходящее $n/2$). Тогда $\psi((q+n)/2) > \psi(k-1+(q+1)/2)$ и $\ln n \leq \ln(2k+1)$, откуда следует, что

$$\begin{aligned} \frac{d(\ln(f_n(q)))}{dq} &< \frac{1}{2}\left(\psi\left(\frac{q+1}{2}\right) - \psi\left(k-1+\frac{q+1}{2}\right)\right) \\ &\quad - \frac{1}{2}\ln(q-1) - \frac{q}{2(q-1)} + \frac{1}{2}\ln(2k+1) \\ &\stackrel{(2.9)}{=} \frac{1}{2}\left(\psi\left(\frac{q+1}{2}\right) - \sum_{i=1}^{k-1} \frac{1}{\frac{q+1}{2} + k - i - 1} - \psi\left(\frac{q+1}{2}\right)\right) \\ &\quad - \frac{q}{2(q-1)} + \frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) \\ &\stackrel{(1)}{\leq} -\frac{1}{2}\sum_{i=1}^{k-1} \frac{2}{q-1+2k-2i} - \frac{1}{q-1} + \frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) \\ &= -\frac{1}{2}\left(\frac{2}{q-1} + \frac{2}{q+1} + \frac{2}{q+3} + \dots + \frac{2}{q+2k-3}\right) + \frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) \\ &\stackrel{(2)}{<} -\frac{1}{2}\ln\left(\frac{q+2k-1}{q-1}\right) + \frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) \\ &\stackrel{(3)}{\leq} -\frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) + \frac{1}{2}\ln\left(\frac{2k+1}{q-1}\right) = 0, \end{aligned}$$

где (1) и (3) выполнены в силу неравенства $q \geq 2$, (2) следует из оценки сверху интеграла от функции $1/x$ интегралом от ее верхней ступенчатой мажоранты $g(x) = 1/(q-1+2i)$, $x \in [q-1+2i, q-1+2i+2]$, $i = 0, \dots, 2k-1$:

$$\frac{2}{q-1} + \frac{2}{q+1} + \frac{2}{q+3} + \dots + \frac{2}{q+2k-3} > \int_{q-1}^{q+2k-1} \frac{1}{x} dx = \ln\left(\frac{q+2k-1}{q-1}\right).$$

Итак, мы показали, что $d(\ln(f_n(q)))/dq < 0$ для $q \geq 2$ и произвольного натурального n . Следовательно, для любого фиксированного n функция $f_n(q)$ убывает по q , а значит, $f_n(q) \leq f_n(2) = 0$, т.е. справедливо неравенство (2.8). Отсюда

и из (2.5), (2.7) получаем, что для любого $q \geq 2$

$$\mathbb{E}[\|e\|_q^2] \stackrel{(2.5)}{\leq} (n\mathbb{E}[|e_2|^q])^{2/q} \stackrel{(2.7),(2.8)}{\leq} (q-1)n^{2/q-1}. \quad (2.11)$$

Неравенство (2.11) нет смысла использовать при больших q (относительно n). Рассмотрим правую часть неравенства (2.11) как функцию q и найдем ее минимум при $q \geq 2$. Рассмотрим $h_n(q) = \ln(q-1) + (2/q-1)\ln n$ (логарифм правой части (2.11)). Производная $h(q)$:

$$\begin{aligned} \frac{dh(q)}{dq} &= \frac{1}{q-1} - \frac{2\ln n}{q^2}, \\ \frac{1}{q-1} - \frac{2\ln n}{q^2} &= 0, \\ q^2 - 2q\ln n + 2\ln n &= 0. \end{aligned}$$

Если $n \geq 8$, то точка минимума на множестве $[2, +\infty)$ есть

$$q_0 = \ln n \left(1 + \sqrt{1 - \frac{2}{\ln n}} \right)$$

(в случае $n \leq 7$ оказывается, что $q_0 = 2$; везде далее считаем, что $n \geq 8$). Поэтому для всех $q > q_0$ более точная оценка будет следующей:

$$\begin{aligned} \mathbb{E}[\|e\|_q^2] &\stackrel{(1)}{\leq} \mathbb{E}[\|e\|_{q_0}^2] \stackrel{(2.11)}{\leq} (q_0-1)n^{2/q_0-1} \stackrel{(2)}{\leq} (2\ln n - 1)n^{2/\ln n - 1} \\ &= (2\ln n - 1)e^2 \frac{1}{n} \leq (16\ln n - 8) \frac{1}{n} \leq (16\ln n - 8)n^{2/q-1}, \end{aligned} \quad (2.12)$$

где (1) верно в силу леммы 1, (2) следует из неравенств $q_0 \leq 2\ln n$, $q_0 \geq \ln n$. Объединяя оценки (2.11) и (2.12), получаем неравенство (2.2).

Теперь перейдем к доказательству неравенства (2.3). Во-первых, получим оценку для $\sqrt{\mathbb{E}[\|e\|_q^4]}$. В силу вероятностного неравенства Йенсена ($q \geq 2$)

$$\begin{aligned} \mathbb{E}[\|e\|_q^4] &= \mathbb{E} \left[\left(\left(\sum_{k=1}^n |e_k|^q \right)^2 \right)^{2/q} \right] \leq \left(\mathbb{E} \left[\left(\sum_{k=1}^n |e_k|^q \right)^2 \right] \right)^{2/q} \\ &\stackrel{(1)}{\leq} \left(\mathbb{E} \left[\left(n \sum_{k=1}^n |e_k|^{2q} \right) \right] \right)^{2/q} \stackrel{(2)}{=} (n^2 \mathbb{E}[|e_2|^{2q}])^{2/q} \\ &\stackrel{(2.7),(2.8)}{\leq} n^{4/q} \left(\left(\frac{2q-1}{n} \right)^{2q/2} \right)^{2/q} = (2q-1)^2 n^{4/q-2}, \end{aligned}$$

где (1) следует из неравенства

$$\left(\sum_{k=1}^n nx_k \right)^2 \leq n \sum_{k=1}^n nx_k^2 \quad \text{для } x_1, x_2, \dots, x_n \in \mathbb{R},$$

а (2) есть следствие линейности математического ожидания и одинаковой распределенности компонент вектора e . Отсюда получаем оценку

$$\sqrt{\mathbb{E}[\|e\|_q^4]} \leq (2q-1)n^{2/q-1}. \quad (2.13)$$

Рассмотрим правую часть неравенства (2.13) как функцию q и найдем ее минимум при $q \geq 2$. Рассмотрим $h_n(q) = \ln(2q - 1) + (2/q - 1) \ln n$ (логарифм правой части (2.13)). Производная $h(q)$:

$$\begin{aligned} \frac{dh(q)}{dq} &= \frac{2}{2q-1} - \frac{2 \ln n}{q^2}, \\ \frac{2}{2q-1} - \frac{2 \ln n}{q^2} &= 0, \\ q^2 - 2q \ln n + \ln n &= 0. \end{aligned}$$

Если $n \geq 3$, то точка минимума на множестве $[2, +\infty)$ есть

$$q_0 = \ln n \left(1 + \sqrt{1 - \frac{1}{\ln n}} \right)$$

(в случае $n \leq 2$ оказывается, что $q_0 = 2$; везде далее считаем, что $n \geq 3$). Поэтому для всех $q > q_0$ более точная оценка будет следующей:

$$\begin{aligned} \sqrt{\mathbb{E}[\|e\|_q^4]} &\stackrel{(1)}{<} \sqrt{\mathbb{E}[\|e\|_{q_0}^4]} \stackrel{(2.13)}{\leq} (2q_0 - 1)n^{2/q_0-1} \stackrel{(2)}{\leq} (4 \ln n - 1)n^{2/\ln n-1} \\ &= (4 \ln n - 1)e^2 \frac{1}{n} \leq (32 \ln n - 8) \frac{1}{n} \leq (32 \ln n - 8)n^{2/q-1}, \end{aligned} \quad (2.14)$$

где (1) верно в силу неравенства $\|e\|_q < \|e\|_{q_0}$ для $q > q_0$, (2) следует из неравенств $q_0 \leq 2 \ln n$, $q_0 \geq \ln n$. Объединяя оценки (2.13) и (2.14), получаем неравенство

$$\sqrt{\mathbb{E}[\|e\|_q^4]} \leq \min\{2q - 1, 32 \ln n - 8\} n^{2/q-1}. \quad (2.15)$$

Теперь найдем $\mathbb{E}[\langle s, e \rangle^4]$, где $s \in \mathbb{R}^n$ – некоторый вектор. Пусть $S_n(r)$ – площадь поверхности n -мерной евклидовой сферы радиуса n , $d\sigma(e)$ – ненормированная равномерная мера на n -мерной евклидовой сфере. В данных обозначениях

$$S_n(r) = S_n(1)r^{n-1}, \quad \frac{S_{n-1}(1)}{S_n(1)} = \frac{n-1}{n\sqrt{\pi}} \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})}.$$

Кроме того, пусть φ – угол между s и e . Тогда

$$\begin{aligned} \mathbb{E}[\langle s, e \rangle^4] &= \frac{1}{S_n(1)} \int_S \langle s, e \rangle^4 d\sigma(\varphi) = \frac{1}{S_n(1)} \int_0^\pi \|s\|_2^4 \cos^3 \varphi S_{n-1}(\sin \varphi) d\varphi \\ &= \|s\|_2^4 \frac{S_{n-1}(1)}{S_n(1)} \int_0^\pi \cos^4 \varphi \sin^{n-2} \varphi d\varphi \\ &= \|s\|_2^4 \cdot \frac{n-1}{n\sqrt{\pi}} \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})} \int_0^\pi \cos^4 \varphi \sin^{n-2} \varphi d\varphi. \end{aligned} \quad (2.16)$$

Отдельно вычислим интеграл:

$$\begin{aligned} &\int_0^\pi \cos^4 \varphi \sin^{n-2} \varphi d\varphi \\ &= 2 \int_0^{\pi/2} \cos^4 \varphi \sin^{n-2} \varphi d\varphi \stackrel{(t=\sin^2 \varphi)}{=} \int_0^{\pi/2} t^{(n-3)/2} (1-t)^{3/2} dt = B\left(\frac{n-1}{2}, \frac{5}{2}\right) \\ &= \frac{\Gamma(\frac{5}{2})\Gamma(\frac{n-1}{2})}{\Gamma(\frac{n+4}{2})} = \frac{\frac{3}{2} \cdot \frac{1}{2} \Gamma(\frac{1}{2})\Gamma(\frac{n-1}{2})}{\frac{n+2}{2} \cdot \Gamma(\frac{n+2}{2})} = \frac{3}{n+2} \cdot \frac{\sqrt{\pi} \Gamma(\frac{n-1}{2})}{2\Gamma(\frac{n+2}{2})}. \end{aligned}$$

Отсюда и из (2.16) получаем, что

$$\begin{aligned}\mathbb{E}[\langle s, e \rangle^4] &= \|s\|_2^4 \cdot \frac{n-1}{n\sqrt{\pi}} \frac{\Gamma(\frac{n+2}{2})}{\Gamma(\frac{n+1}{2})} \cdot \frac{3}{n+2} \cdot \frac{\sqrt{\pi}\Gamma(\frac{n-1}{2})}{2\Gamma(\frac{n+2}{2})} \\ &= \|s\|_2^4 \cdot \frac{3(n-1)}{2n(n+2)} \cdot \frac{\Gamma(\frac{n-1}{2})}{\frac{n-1}{2}\Gamma(\frac{n-1}{2})} = \frac{3\|s\|_2^4}{n(n+2)} \leq \frac{3\|s\|_2^4}{n^2}.\end{aligned}\quad (2.17)$$

Чтобы доказать неравенство (2.3), осталось воспользоваться (2.15), (2.17) и неравенством Коши–Буняковского ($(\mathbb{E}[XY])^2 \leq \mathbb{E}[X^2] \cdot \mathbb{E}[Y^2]$):

$$\mathbb{E}[\langle s, e \rangle^2 \|e\|_q^2] \leq \sqrt{\mathbb{E}[\langle s, e \rangle^4] \cdot \mathbb{E}[\|e\|_q^4]} \leq \sqrt{3}\|s\|_2^2 \min\{2q-1, 32\ln n-8\}n^{2/q-2}.$$

Теорема доказана.

3. Вычислительные эксперименты. Для уточнения констант в верхних оценках теоремы 1 были проведены вычислительные эксперименты. При генерации случайных векторов, равномерно распределенных на поверхности n -мерной евклидовой сферы, использовалась лемма Пуанкаре (см. (2.6)) о том, что компоненты любого вектора e , имеющего такое распределение, можно представлять как отношения $e_k/\sqrt{e_1^2 + \dots + e_n^2}$, где все e_k , $k = 1, 2, \dots$ – независимые одинаково распределенные случайные величины, имеющие стандартное нормальное распределение $N(0, 1)$.

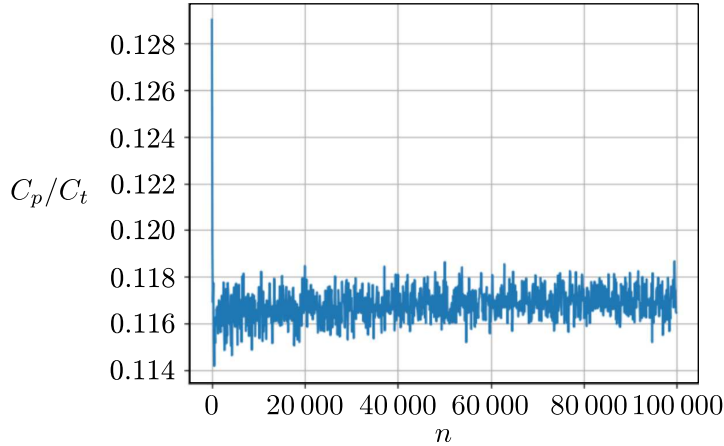


Рис. 1. Уточнение константы в оценке (2.2), n – размерность пространства, $C_t = 16 \ln n - 8$, $q = \infty$.

На рис. 1 приведены результаты эксперимента по оценке математического ожидания ∞ -нормы векторов $e \in RS_2^n(1)$. По теореме 1 при $q = \infty$ неравенство (2.2) имеет вид

$$\mathbb{E}[\|e\|_\infty^2] \leq C_t n^{-1},$$

где $C_t = 16 \ln n - 8$. Эти же константы (назовем их в этом случае C_p) можно оценить практически путем вычисления $\mathbb{E}[\|e\|_\infty^2]$ методом Монте-Карло. Это было сделано для n от 10 до 10^5 , и на рис. 1 приведено отношение C_p/C_t для разных n . Оказалось, что отношение с ростом n не меняется, что значит, что теоретическая оценка с точностью до константы верна.

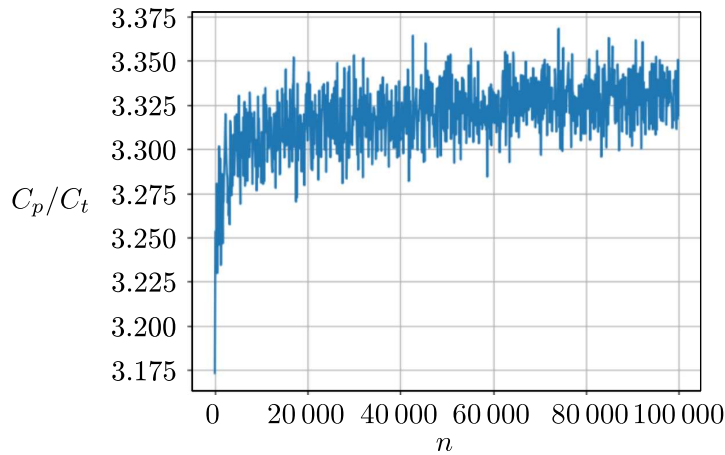


Рис. 2. Уточнение константы в оценке (2.3), n – размерность пространства, $C_t = \sqrt{3}\|s\|_2^2 \min\{2q - 1, 32 \ln n - 8\}$, $q = \infty$.

Такие же эксперименты были проведены и для оценки константы

$$C_t = \sqrt{3}\|s\|_2^2 \min\{2q - 1, 32 \ln n - 8\}$$

в неравенстве (2.3) при $q = \infty$:

$$\mathbb{E}[\langle s, e \rangle^2 \|e\|_q^2] \leq C_t n^{-2}.$$

Результаты экспериментов (см. рис. 2) также подтверждают, что теоретическая оценка C_t с точностью до константы верна.

Код на языке Python всех вычислительных экспериментов выложен на Github [15].

Авторы выражают благодарность Павлу Евгеньевичу Двуреченскому за помощь в работе.

СПИСОК ЦИТИРОВАННОЙ ЛИТЕРАТУРЫ

- [1] E. Vorontsova, A. Gasnikov, E. Gorbunov, *Accelerated Directional Search with Non-Euclidean Prox-Structure*, 2019, [arXiv: 1710.00162](#).
- [2] A. Gasnikov, A. Lagunovskaya, I. Usmanova, F. Fedorenko, *Gradient-Free Prox-Methods with Inexact Oracle for Stochastic Convex Optimization Problems on a Simplex*, 2014, [arXiv: 1412.3890](#).
- [3] А. В. Гасников, А. А. Лагуновская, И. Н. Усманова, Ф. А. Федоренко, “Безградиентные прокс-методы с неточным оракулом для негладких задач выпуклой стохастической оптимизации на симплексе”, *Автомат. и телемех.*, 2016, № 10, 57–77.
- [4] O. Shamir, *An Optimal Algorithm for Bandit and Zero-Order Convex Optimization with Two-Point Feedback*, 2015, [arXiv: 1507.08752](#).
- [5] O. Shamir, “An optimal algorithm for bandit and zero-order convex optimization with two-point feedback”, *J. Mach. Learn. Res.*, **18** (2017), Paper No. 52.
- [6] J. C. Duchi, M. I. Jordan, M. J. Wainwright, A. Wibisono, “Optimal rates for zero-order convex optimization: the power of two function evaluations”, *IEEE Trans. Inform. Theory*, **61:5** (2015), 2788–2806.
- [7] A. Blum, J. Hopcroft, R. Kannan, *Foundations of Data Science*, <https://www.cs.cornell.edu/jeh/book.pdf>, 2018.

- [8] K. Ball, “An elementary introduction to modern convex geometry”, *Flavors of Geometry*, Math. Sci. Res. Inst. Publ., **31**, Cambridge Univ. Press, Cambridge, 1997.
- [9] L. Bogolubsky, P. Dvurechensky, A. Gasnikov, G. Gusev, Yu. Nesterov, A. Raigorodskii, A. Tikhonov, M. Zhukovskii, *Learning Supervised PageRank with Gradient-Based and Gradient-Free Optimization Methods*, <https://papers.nips.cc/paper/6565-learning-supervised-pagerank-with-gradient-based-and-gradient-free-optimization-methods.pdf>, 2016.
- [10] S. Boucheron, G. Lugosi, P. Massart, *Concentration Inequalities. A Nonasymptotic Theory of Independence*, Oxford Univ. Press, Oxford, 2013.
- [11] V. Milman, G. Schechtman, *Asymptotic Theory of Finite Dimensional Normed Spaces*, Lecture Notes in Math., **1200**, Springer-Verlag, Berlin, 1986.
- [12] В. А. Зорич, *Математический анализ в задачах естествознания*, МЦНМО, М., 2008.
- [13] A. Bayandina, A. Gasnikov, F. Guliev, A. Lagunovskaya, *Gradient-Free Two-Points Optimal Method for Non Smooth Stochastic Convex Optimization Problem with Additional Small Noise*, **1701.03821**, 2017.
- [14] M. Lifshits, *Lectures on Gaussian Processes*, Springer, Heidelberg, 2012.
- [15] Э. А. Горбунов, Е. А. Воронцова, *Вычислительные эксперименты, иллюстрирующие явление концентрации равномерной меры на поверхности евклидовой сферы в малой окрестности экватора*, <https://github.com/evorontsova/Concentration-of-Measure>, 2015.

Э. А. Горбунов

Московский физико-технический институт
(государственный университет),
г. Долгопрудный, Московская обл.
E-mail: ed-gorbunov@yandex.ru

Поступило

13.04.2018

Принята к публикации

12.09.2018

Е. А. Воронцова

Дальневосточный федеральный университет,
г. Владивосток;
Université Grenoble Alpes, Франция
E-mail: vorontsovaea@gmail.com

А. В. Гасников

Московский физико-технический институт
(государственный университет),
г. Долгопрудный, Московская обл.;
Кавказский математический центр,
Адыгейский государственный университет,
г. Майкоп
E-mail: gasnikov@yandex.ru