Записки научных
семинаров ПОМИ
Том 294, 2002 г.

**В. Н. Судаков**

# РАВНОМЕРНЫЕ ПО РАЗМЕРНОСТИ ОЦЕНКИ РАЗБРОСА РАСПРЕДЕЛЕНИЙ ГАУССОВСКИХ СЛУЧАЙНЫХ ФУНКЦИОНАЛОВ ДЛЯ НЕКОТОРЫХ КЛАССОВ МНОГОМЕРНЫХ ВЕРОЯТНОСТНЫХ МЕР

**0.** Введем используемые в дальнейшем термины и напомним некоторые определения и известные результаты.

Пространство всех вероятностных мер на пространстве $S$ будем обозначать $\mathcal{P}(S)$.

Центрированная гауссовская мера $\gamma$ в конечномерном евклидовом пространстве $\Omega$ называется в дальнейшем субстандартной, если определяемая этой мерой евклидова норма (возможно, псевдонорма, но мы будем и в этом случае употреблять термин "норма") в сопряженном с $\Omega$ пространстве $F$, индуцированная вложением $F \subset L^2(\gamma)$), мажорируется стандартной евклидовой нормой сопряженного с евклидовым пространства. Иными словами, если мера $\gamma$ есть образ стандартной гауссовской меры $\gamma_\Omega$ при линейном отображении $\Omega \to \Omega$ с нормой, не превосходящей единицы.

Вероятностная мера $\mu$ на $\mathbb{R}$ называется сублапласовской, если она является образом $N(0,1)\circ\varphi^{-1}$ стандартной одномерной гауссовской меры $N(0,1)$ при некотором нерастягивающем, т.е. липшицевом с константой единица отображении $\varphi\colon \mathbb{R} \to \mathbb{R}$.

Расстоянием Канторовича между мерами $\mu$ и $\nu$ в пространстве вероятностных мер на метрическом пространстве с метрикой $\rho$ называется величина

$$\varkappa(\mu,\nu) = \inf \iint \rho(x,y)\, m(dx\, dy),$$

где инфимум берется по всем вероятностным мерам $m$ на квадрате этого пространства с маргинальными распределениями $\mu$

Работа поддержана грантами РФФИ 99-15-96019, 02-01-00263 и 99-01-04027.

и $\nu$. Для метрического пространства $\mathbb{R}$ расстояние Канторовича равно $\int |F_\mu(t) - F_\nu(t)|dt$. Если меры $\mu$ и $\nu$ на $\mathbb{R}$ — дискретные, представляющие собой равномерные распределения на конечных подмножествах, соответственно, $\{s_1, \ldots, s_n\}$ и $\{t_1, \ldots, t_n\}$, где

$$s_1 \leqslant \cdots \leqslant s_n, \quad t_1 \leqslant \cdots \leqslant t_n$$

(кратность учитывается), то

$$\varkappa(\mu, \nu) = n^{-1} \sum_1^n |t_k - s_k|.$$

Если $\varphi$ — функция, определенная на некотором конечномерном векторном пространстве $\Omega$ с заданной на этом пространстве субстандартной гауссовской мерой $\gamma$, удовлетворяющая условию Липшица с константой единица, то ее распределение $\gamma \circ \varphi^{-1}$ относительно меры $\gamma$ является сублапласовским [1, 2].

**1.** В работе для двух классов борелевских вероятностных мер $\mathbf{P}$ (гауссовских и равномерно распределенных на конечном подмножестве линейно независимых векторов) на конечномерном векторном пространстве $F$ произвольной размерности исследуется следующая задача. Пусть на сопряженном с $F$ векторном пространстве $\Omega$ рассматривается евклидова норма $\|\cdot\|_{\mathbf{P}}$, которая индуцируется на $\Omega$ естественным вложением $\Omega \subset L^2(F, \mathbf{P})$, и задана стандартная по этой норме гауссовская мера $\gamma_{\mathbf{P}}$. Существует ли такая абсолютная (не зависящая от размерности пространства $F$) константа $C$, что для некоторого распределения $\overline{\mathbf{P}}$ на прямой $\mathbb{R}$ выполняется неравенство

$$\mathbf{E}\varkappa(\overline{\mathbf{P}}, \mathbf{P} \circ \omega^{-1}) \leqslant C?$$

(Такое распределение можно назвать “псевдотипичным” распределением для выбираемых случайным образом в соответствии с “естественной” для евклидового пространства $\Omega$ стандартной гауссовской мерой элементов этого пространства, которые относительно вероятностной меры $\mathbf{P}$ являются случайными величинами. В отличие от ранее установленных [3] теорем о существовании для любого $\varepsilon > 0$ $\varepsilon$-типичных распределений, когда рассматривавшиеся сферически инвариантные гауссовские меры $\gamma$ на $\Omega$ не были стандартными, но их ковариационные квадратичные формы стягивались с увеличением размерности к нулевой таким образом, что типичные для $\Omega$ нормы выбираемых из распределения $\gamma$ элементов $\omega \in \Omega$ оставались приблизительно единичными, иначе говоря, распределения относительно $\mathbf{P}$ типичных по такой гауссовской мере элементов не вырождались, теперь ковариационная квадратичная форма меры $\gamma$ остается единичной, вследствие чего нормы извлекаемых по мере $\gamma$ элементов пространства $\Omega$ неограниченно растут, т.е. растут их дисперсии относительно меры $\mathbf{P}$, и обсуждаемый феномен состоит не в том, что распределения возникающих семейств случайных элементов пространства $\mathcal{P}(\mathbb{R})$ с ростом размерности все теснее концентрируются, но в том, что несмотря на "расплывание" всех распределений вида $\mathbf{P} \circ \omega^{-1}$, они все в пространстве $\mathcal{P}(\mathbb{R})$ постоянно держатся вместе, хотя и не сближаются, а не рассеиваются, как можно было бы ожидать.)

**2.** Рассмотрим как первый пример в качестве меры $\mathbf{P}$ дискретную вероятностную меру на пространстве $F$, сосредоточенную на линейно независимых элементах $f_1, f_2, \dots, f_d$ векторного пространства $F$, такую, что для $k = 1, \dots, d$ $\mathbf{P}(\{f_k\}) = d^{-1}$. Не умаляя общности, мы можем считать, что векторы $f_k$ образуют базис в пространстве $F$, переходя в противном случае к подпространству, натянутому на них. На сопряженном с $F$ векторном пространстве $\Omega := F'$, $\dim \Omega = d$, рассмотрим евклидову норму $\|\cdot\|_\Omega$, определяемую как норму в пространстве $L^2(\mathbf{P})$. Можно говорить, что норма $|\cdot|_\Omega$ в пространстве $\Omega$ — это евклидова норма, индуцированная каноническим отображением $F \to L^2(F, \mathbf{P})/\Omega^\circ$, и так это и будет делаться в дальнейшем, при рассмотрении мер $\mathbf{P}$ более сложной природы.

Пусть теперь $\gamma_0^d$ — стандартная гауссовская мера в евклидовом пространстве $\Omega$ (ее ковариационная квадратичная форма $q$ на сопряженном пространстве $F$ совпадает с квадратом нормы сопряженного пространства).

**Теорема 1.** *Существует такая абсолютная (не зависящая от размерности) константа $C$, что для пары независимо по мере $\gamma_0^d$ выбранных элементов $\omega', \omega'' \in \Omega$ распределения (случайного) расстояния в пространстве $(\mathcal{P}, \varkappa)$ распределений этих линейных функционалов*

$$\varkappa(\mathbf{P} \circ \omega'^{-1}, \mathbf{P} \circ \omega''^{-1})$$

*является сублапласовским со средним, не превосходящим $C$.*

**Доказательство.** Рассмотрим евклидово пространство $\Omega \oplus \Omega$ со стандартной гауссовской мерой $\gamma_0^d \otimes \gamma_0^d$ на нем, так что выбор элемента $(\omega', \omega'') \in \Omega \oplus \Omega$ из распределения $\gamma_0^d \otimes \gamma_0^d$ означает повторную независимую объема 2 выборку из распределения $\gamma_0^d$. В пространстве $\Omega$ введем ортонормированную систему координат, взяв в качестве координатных функционалов этементы $d_k^{-\frac{1}{2}}$, $k = 1, \dots, d$. Теперь элементы $\omega' = (\omega'_1, \dots, \omega'_d)$ и $\omega'' = (\omega''_1, \dots, \omega''_d)$ можно рассматривать как две независимо произведенные повторные независимые выборки из нормального распределения $N(0, d)$, причем выборочные распределения для этих выборок совпадают с распределениями относительно меры $\mathbf{P}$ случайных относительно меры $\gamma_0^d$ величин $\omega'$ и $\omega''$. Интересующее нас расстояние Канторовича между распределениями $\mathbf{P}\omega'^{-1}$ и $\mathbf{P}\omega''^{-1}$ — это расстояние Канторовича между описанными выборочными распределениями.

Как отмечалось в п. 0, расстояние Канторовича $\varkappa$ между двумя выборочными дискретными распределениями с равными нагрузками по $d^{-\frac{1}{2}}$ в точках $d^{\frac{1}{2}}\omega'_{(k)}$, $k = 1, \dots, d$, $\omega'_1 \leqslant \omega'_{(2)} \leqslant \dots \leqslant \omega'_{(d)}$ и $d^{\frac{1}{2}}\omega''_{(k)}$, $k = 1, \dots, d$, $\omega''_1 \leqslant \omega''_{(2)} \leqslant \dots \leqslant \omega''_{(d)}$ равно

$$\varkappa(\mathbf{P} \circ \omega'^{-1}, \mathbf{P} \circ \omega''^{-1}) = \sum_{k=1}^{d} d^{-\frac{1}{2}} |\omega'_{(k)} - \omega''_{(k)}|.$$

Непосредственно проверяется, что эта величина не превосходит

$$(\sum_{k=1}^{d} (\omega'_{(k)} - \omega''_{(k)})^2)^{\frac{1}{2}} \leqslant (\sum_{k=1}^{d} (\omega'_k - \omega''_k)^2)^{\frac{1}{2}} = \|(\omega', \omega'')\|_{\Omega \oplus \Omega}.$$

Таким образом, функция $\kappa$

$$\kappa \colon \Omega \oplus \Omega \to \mathbb{R}, \quad \kappa(\omega', \omega'') = \varkappa(\mathbf{P} \circ \omega'^{-1}, \mathbf{P} \circ \omega''^{-1})$$

удовлетворояет условию Липшица с константой единица, и поэтому образ при указанном отображении $\kappa$ стандартной гауссовской меры $\gamma_0^d \otimes \gamma_0^d$ в пространстве $\Omega \oplus \Omega$ является сублапласовским.

Остается доказать, что среднее этого распределения (всегда конечное ввиду сублапласовости) ограничено некоторой константой $C$, не зависящей от размерности. Поскольку расстояние $\varkappa$ между двумя эмпирическими распределениями с одинаковым объемом выборки равно среднему расстоянию между одноименными членами вариационных рядов этих выборок, которые можно рассматривать как одноименные квантили этих эмпирических выборок, мы воспользуемся хорошо изученным предельным поведением эмпирических выборочных функций и выборочных квантилей (см., например, [4, 5]). Распределение выборочной квантили порядка $p$ выборки объема $n$ из распределения $N(0,1)$ асимптотически нормально со средним $\Phi^{-1}(p)$ (теоретическая квантиль) и дисперсией

$$\sigma^2 = \frac{p(1-p)}{n(\phi(\Phi^{-1})^2}.$$

(Здесь $\phi$ — плотность и $\Phi$ — функция распределения стандартного нормального закона.) Учитывая величину этой дисперсии и тот факт, что распределение $\mathcal{L}(\|X-Y\|)$, где $X$ и $Y$ — независимые случайные величины с распределениями $N(0,\sigma^2)$, имеет среднее, пропорциональное значеению $\sigma$, мы в конце концов приходим р требуемому заключению, что среднее значение величины $d^{-1}\sum_{k=1}^{d}\|X_{(k)}-Y_{(k)}\|$, где $X_{(k)}$ и $Y_{(k)}$ — члены вариационного ряда для двух независимо полученных повторных независимых выборок объема $d$ из распределения $N(0,d)$ ограничено некоторой константой, не зависящей от числа $d$, чем и завершается доказательство теоремы 1 — рассмотрение класса дискретных распределений в конечномерном векторном пространстве $F$, равномерных на элементах некоторого алгебраического базиса.

**3.** Рассмотрим теперь другой класс вероятностных распределений на конечномерном векторном пространстве $F$: класс невырожденных гауссовских мер.

**Теорема 2.** *Пусть* $\mathbf{P}$ *— центрированная гауссовская мера в векторном пространстве* $F$ *конечной размерности* $d$*, и пусть* $\gamma$ *— центрированная гауссовская мера в сопряженном с* $F$ *пространстве* $\Omega$*, субстандартная относительно евклидовой нормы в пространстве* $F$*, индуцированной вложением* $\Omega \subset L^2(\mathbf{P})$*. Тогда существует такая абсолютная (не зависящая от размерности) константа* $C$*, что для пары независимо по мере* $\gamma$ *выбранных элементов* $\omega', \omega'' \in \Omega$ *распределения (случайного) расстояния в пространстве* $(\mathcal{P}, \varkappa)$ *распределений этих линейных функционалов*

$$\varkappa(\mathbf{P}\circ\omega'^{-1}, \mathbf{P}\circ\omega''^{-1})$$

*является сублапласовским со средним, не превосходящим $C$.*

**Доказательство.** Определенная как и выше функция $\kappa$ на евклидовом пространстве $\Omega \oplus \Omega, \gamma \otimes \gamma$ является, как и выше, липшицевой с константой 1, поэтому ее распределение относительно субстандартной меры $\gamma \otimes \gamma$ является сублапласовской мерой на $\mathbb{R}$. Остается оценить ее среднее. Величина $\kappa(\omega', \omega'')$ равна $\varkappa(N(0, |\omega'|^2, N(0, |\omega''|^2))$. Поскольку распределение евклидовой нормы как липшицевой функции относительно субстандартной гауссовской меры является сублапласовским, его дисперсия не превосходит единицы, и разность $\|\omega'\| - \|\omega''\|$ имеет нулевое среднее и дисперсию, не превосходящую двух. С другой стороны, величина $\varkappa(N(0, \|\omega'\|^2, N(0, \|\omega''\|^2))$ пропорциональна модулю разности стандартных отклонений $|\|\omega'\| - \|\omega''\||$, следовательно, ее среднее не превосходит некоторой абсолютной константы.

## Литература


1. В. Н. Судаков, Б. С. Цирельсон, *Экстремальные свойства полупространств для сферически инвариантных мер.* — Записки научн. семин. ЛОМИ **41** (1974), 14–24.
2. B. S. Cirel'son, I. A. Ibragimov, V. N. Sudakov, *Norms of Gaussian sample functions.* — Lecture Notes in Mathematics **550** (1976), 20–41.
3. В. Н. Судаков, *Типичные распределения линейных функционалов на пространствах высокой размерности.* — ДАН СССР **243**, No. 6 (1978), 1402–1405.
4. П. Биллингсли, *Сходимость вероятностных мер.* М., Наука (1977).
5. Д. Кокс, Д. Хинкли, *Теоретическая статистика.* М., Мир (1978).



Sudakov V. N. Uniform in dimension estimations of the scatter of distributions of Gaussian random functionals for some classes of multivariate measures.

Uniform in dimension estimations of the scatter of dstributions of linear functionals chosen at random with respect to Gaussian distribution are established for uniform discrete distributions and for Gaussian distributions on a vector space.



С.-Петербургское отделение
Математического института
им. В. А. Стеклова РАН


Поступило 20 августа 2002 г.