ПРОБЛЕМЫ МАТЕМАТИЧЕСКОГО АНАЛИЗА

Январь 2010 **Выпуск 44. С. 55–64**

С. Г. Бобков
Университет Миннесоты, США
bobko001@umn.edu

О КОНЦЕНТРАЦИИ МЕРЫ НА КУБЕ

Изучается свойство концентрации равномерного распределения на кубе в классе множеств, инвариантных при перестановках координат. Библиография: 10 назв.

1. Введение

Обозначим через $\mathbf{P}$ равномерное распределение вероятности (т.е. сужение меры Лебега) на n-мерный куб $[0,1]^n$. Ввиду известного свойства безразмерной концентрации этой меры для любого измеримого множества $A \subset \mathbf{R}^n$ такого, что $\mathbf{P}(A) \geqslant 1/2$, имеем

$$\mathbf{P}(A + rB_2) \geqslant 1 - e^{-cr^2}, \qquad r > 0. \tag{1.1}$$

Здесь B_2 обозначает евклидов шар в $\mathbf{R}^n$ единичного радиуса и с центом в начале координат, так что сумма Минковского $A + rB_2$ является открытой r-окрестностью A относительно евклидова расстояния, и $c > 0$ — константа. Рассуждая, например, так же, как в [**1**], можно получить неравенство (1.1). Рассмотрим отображение

$$T(x_1, \ldots, x_n) = (\Phi(x_1), \ldots, \Phi(x_n))$$

из $\mathbf{R}^n$ на $(0,1)^n$, где Φ — функция стандартного нормального распределения на прямой. Оно порождает каноническую гауссову меру γ_n на $\mathbf{R}^n$ в $\mathbf{P}$, т.е. $\mathbf{P}(A) = \gamma_n(T^{-1}(A))$, и имеет константу Липшица

$$\|T\|_{\mathrm{Lip}} = \|\Phi\|_{\mathrm{Lip}} = \frac{1}{\sqrt{2\pi}}.$$

Поэтому ввиду изопериметрического неравенства для гауссовой меры

$$\gamma_n(A + rB_2) \geqslant \Phi\left(\Phi^{-1}(\gamma_n(A)) + r\right),$$

справедливого для всех измеримых множеств $A \subset \mathbf{R}^n$ и приводящего в неравенству

$$\gamma_n(A + rB_2) \geqslant 1 - e^{-r^2/2}$$

в случае $\gamma_n(A) \geqslant 1/2$, получаем (1.1), где $c = \pi$. Имеется много других подходов к изучению концентрации (изложение современного состояния дел по изучению концентрации меры можно найти в [**2**]).

Работа выполнена при частичной финансовой поддержке NSF.

В терминах функций (1.1) эквивалентно следующему: для любой функции $f : \mathbf{R}^n \to \mathbf{R}$ такой, что

$$\|f\|_{\mathrm{Lip}} \leqslant 1,$$

и $\mathbf{P}$-средним

$$\mathbf{E}f = \int\limits_{[0,1]^n} f(x)\,dx$$

имеем

$$\mathbf{P}\{|f - \mathbf{E}f| \geqslant r\} \leqslant 2\,e^{-cr^2}, \qquad h > 0 \tag{1.2}$$

(при некоторой константе $c > 0$). Правая часть (1.2) асимптотически неулучшаема в классе всех липшицевых функций. Однако во многих интересных случаях оценка (1.2) является грубой. Например, значения липшицевой функции $f(x) = \max\{x_1, \ldots, x_n\}$ концентрируются вокруг точки 1 и, более того, концентрация растет с ростом n. В более общем случае, когда требуется улучшить (1.2), рассматривают задачу для уклонений U-статистик

$$f(x) = \sum U(x_{i_1}, \ldots, x_{i_k}),$$

где U — "хорошая" функция от $k \leqslant n$ переменных и суммирование выполняется по всем наборам целых чисел $1 \leqslant i_1 < \cdots < i_k \leqslant n$.

Эти примеры побуждают к исследованию свойства концентрации равномерного распределения, а также других мер произведения, в классе подмножеств, инвариантных при перестановках. Будем говорить, что множество $A \subset \mathbf{R}^n$ *симметрично при перестановках координат* (или просто *инвариантно при перестановках*), если одновременно с каждой точкой $x = (x_1, \ldots, x_n)$ оно содержит все точки вида $x = (x_{\pi(1)}, \ldots, x_{\pi(n)})$, где π — произвольная перестановка множества $\{1, \ldots, n\}$. Для таких множеств можно значительно улучшить (1.1).

Теорема 1.1. *Для любого измеримого множества $A \subset \mathbf{R}^n$, симметричного при заменах координат, и $\mathbf{P}(A) \geqslant 1/2$*

$$\mathbf{P}(A + rB_\infty) \geqslant 1 - e^{-cnr^2}, \qquad r > 0, \tag{1.3}$$

где $c > 0$ — абсолютная константа.

Далее B_p обозначает единичный ℓ^p-шар относительно нормы

$$\|x\|_p = (|x_1|^p + \cdots + |x_n|^p)^{1/p}$$

(где $x = (x_1, \ldots, x_n) \in \mathbf{R}^n$, $1 \leqslant p \leqslant \infty$). В частности, $B_\infty = (-1,1)^n$ — открытый n-мерный куб в $\mathbf{R}^n$ со стороной $(-1,1)$.

На языке функций теорему 1.1 можно сформулировать следующим образом.

Теорема 1.2. *Пусть f — функция на $[0,1]^n$, симметричная относительно перестановок координат и такая, что для всех $x, y \in [0,1]^n$*

$$|f(x) - f(y)| \leqslant \|x - y\|_\infty. \tag{1.4}$$

Тогда для всех $r > 0$ с некоторой абсолютной константой $c > 0$

$$\mathbf{P}\{|f - \mathbf{E}f| > r\} \leqslant 2\,e^{-cnr^2}. \tag{1.5}$$

Таким образом, в классе подмножеств из $[0,1]^n$, инвариантных при перестановках, имеется более сильная концентрация гауссова типа. Чтобы сравнить (1.1) и (1.3), перепишем последнее неравенство в виде

$$\mathbf{P}\Big(A + \frac{r}{\sqrt{n}}\,B_\infty\Big) \geqslant 1 - e^{-cr^2}$$

и заметим, что

$$A + \frac{r}{\sqrt{n}}\,B_\infty \subset A + rB_2.$$

Аналогично, для функций условие (1.4) означает, что f дифференцируема почти всюду (п.в.) и ее градиент удовлетворяет соотношению

$$\|\nabla f\|_1 = \sum_{i=1}^n \left|\frac{\partial f(x)}{\partial x_i}\right| \leqslant 1 \quad \text{п.в.}$$

Поскольку

$$\|\nabla f\|_1 \leqslant \sqrt{n}\,\|\nabla f\|_2,$$

из (1.5) вытекает гауссово неравенство отклонения (1.2).

Теоремы 1.1–1.2 можно свести к задаче о концентрации равномерного распределения μ_n на симплексе

$$\Delta_n = \{y \in \mathbf{R}_+^n : y_1 + \cdots + y_n \leqslant 1\}$$

относительно ℓ^1-нормы. Для доказательства справедливости (1.3) нам потребуется следующее утверждение.

Теорема 1.3. *Для любого измеримого множества $A \subset \Delta_n$ такого, что $\mu_n(A) \geqslant 1/2$,*

$$\mu_n(A + rB_1) \geqslant 1 - e^{-cnr^2}, \qquad r > 0, \tag{1.6}$$

где $c > 0$ — абсолютная константа.

Это неравенство концентрации известно, однако оно было установлено для ℓ^1-шара B_1 вместо Δ_n (что, фактически, эквивалентно случаю симплекса). Более точно, если ν_n — нормированная мера Лебега на B_1 и множество $A \subset B_1$ имеет ν_n-меру не меньше $1/2$, то, как было впервые показано в [**3**],

$$\nu_n(A + rB_1) \geqslant 1 - n\,e^{-cnr^2}.$$

(На самом деле, в [**3**] было установлено это неравенство для всех ℓ_p-шаров при $1 \leqslant p \leqslant 2$.) Впоследствии удалось избавиться [**4**] от множителя n, который не являлся необходимым. В обоих случаях доказательство существенно основано на изопериметрической теореме Талаграна для экспоненциальной меры произведения [**5**]. Несмотря на то, что эта глубокая теорема к настоящему времени стала стандартным фактом и имеется несколько различных подходов к ее доказательству (см. [**6, 7**]), мы дадим прямое доказательство по индукции неравенства концентрации (1.6).

2. Сведение к симплексу

Для данной перестановки π множества $\{1, \ldots, n\}$ определим

$$\Omega(\pi) = \left\{x \in [0,1]^n : 0 \leqslant x_{\pi(1)} \leqslant \ldots \leqslant x_{\pi(n)} \leqslant 1\right\}.$$

Пусть Ω_n соответствует тождественной перестановке $\pi(i) = i$. Тогда по модулю 0 получаем разбиение куба на $n!$ подмножеств. В частности, для любого измеримого множества A в $[0,1]^n$ получаем

$$\mathbf{P}(A) = \frac{1}{n!}\sum_\pi \mathbf{P}_\pi(A) \tag{2.1}$$

и

$$\mathbf{P}(A + rB_\infty) = \frac{1}{n!}\sum_\pi \mathbf{P}_\pi(A + rB_\infty),$$

где $\mathbf{P}_\pi$ обозначает равномерное распределение на Ω_π. Если A симметрично при перестановке координат, члены в каждой сумме совпадают, откуда

$$\mathbf{P}(A) = \mathbf{P}_n(A), \qquad \mathbf{P}(A + rB_\infty) = \mathbf{P}_n(A + rB_\infty),$$

где через $\mathbf{P}_n$ мы временно обозначили равномерное распределение на Ω_n.

По аналогии с (2.1), при заданной (измеримой, ограниченной) функции f на $[0,1]^n$ имеем

$$\int f\,d\mathbf{P} = \frac{1}{n!}\sum_\pi \int f\,d\mathbf{P}_\pi.$$

Если f инвариантна при перестановке координат, $\mathbf{P}$-распределение функции f совпадает с ее $\mathbf{P}_n$-распределением, и то же верно для $\|\nabla f\|$ относительно любой нормы на $\mathbf{R}^n$, инвариантной при перестановках.

Теперь рассмотрим меру $\mathbf{P}_n$. Ее носитель Ω_n может быть идентифицирован с симплексом

$$\Delta_n = \{y \in \mathbf{R}^n_+ : y_1 + \cdots + y_n \leqslant 1\}$$

с помощью отображения

$$T : x \to y = (x_1, x_2 - x_1, \ldots, x_n - x_{n-1}).$$

Это отображение преобразует $\mathbf{P}_n$ в равномерную меру μ_n на Δ_n. Поэтому для любого измеримого множества $A \subset \mathbf{R}^n$ имеем

$$\mathbf{P}_n(A) = \mu_n(T(A))$$

и ввиду линейности T

$$\mathbf{P}_n(A + rB_\infty) = \mu_n(T(A) + rT(B_\infty)).$$

Заметим, что $T(B_\infty)$ содержит ℓ^1-шар B_1: начиная с точки $y = (y_1, \ldots, y_n)$ такой, что $|y_1| + \cdots + |y_n| \leqslant 1$, точка $T^{-1}(y)$ имеет координаты $(y_1, y_1 + y_2, \ldots, y_1 + \cdots + y_n)$ и поэтому принадлежит B_∞. Следовательно,

$$\mathbf{P}_n(A + rB_\infty) \geqslant \mu_n(T(A) + rB_1).$$

По этим же причинам, начиная с локально липшицевой функции f на Ω_n, функция $y \to f(T^{-1}(y)) = f(y_1, y_1 + y_2, \ldots, y_1 + \cdots + y_n)$ определена на Δ_n и удовлетворяет соотношению

$$\|\nabla f(T^{-1}(y))\|_\infty \leqslant \|(\nabla f)(T^{-1})\|_1.$$

Эти факты можно резюмировать следующим образом.

Лемма 2.1. *Для данных чисел $p \in (0,1)$, $r > 0$ и функции R на $[p,1]$ любое свойство концентрации на симплексе вида*

$$\mu_n(A + rB_1) \geqslant R(\mu_n(A)),$$

справедливое для всех измеримых множеств $A \subset \Delta_n$ таких, что $\mu_n(A) \geqslant p$, влечет аналогичное свойство на кубе

$$\mathbf{P}(A + rB_\infty) \geqslant R(\mathbf{P}(A))$$

в классе всех измеримых множеств $A \subset [0,1]^n$, симметричных при перестановках координат и таких, что $\mathbf{P}(A) \geqslant p$.

Сформулируем лемму 2.1 на языке функций.

Лемма 2.2. *Пусть $\alpha = \alpha(r)$ определено на $(0, +\infty)$. Предположим, что для каждой функции f на Δ_n такой, что $|f(x) - f(y) \leqslant \|x - y\|_1$ при $x, y \in \Delta_n$,*

$$\mu_n\left\{\left|f - \int f\, d\mu_n\right| \geqslant r\right\} \leqslant \alpha(r), \quad r > 0.$$

Тогда для каждой функции f на $[0,1]^n$, инвариантной при перестановке координат и такой, что $|f(x) - f(y) \leqslant \|x - y\|_\infty$ при $x, y \in [0,1]^n$, справедливо неравенство

$$\mathbf{P}\left\{\left|f - \int f\, d\mathbf{P}\right| \geqslant r\right\} \leqslant \alpha(r), \quad h > 0.$$

Замечание 2.3. Ввиду вышесказанного вместо симметрии при перестановках в заключении леммы 2.2 и тем самым теоремы 1.2 достаточно потребовать равенства $\mathbf{P}_\pi$-средних f на Ω_π. Аналогично в лемме 2.1 и теореме 1.1 можно ограничиться требованием равенства объемов пересечений $A \cap \Omega_\pi$.

3. Распределение последней координаты

Таким образом, благодаря леммам 2.1 и 2.2 изучение распределений функционалов, инвариантных при перестановках, на кубе сводится к изучению распределений общих функционалов на вероятностном пространстве (Δ_n, μ_n). На первом шаге полезно изучить, как последняя координата

$$\eta_n(x) = x_n, \qquad x = (x_1, \ldots, x_n),$$

концентрируется вокруг своего μ_n-среднего. Обозначим через F_n функцию распределения случайной величины η_n. Ясно, что

$$F_n(t) = 1 - (1-t)^n, \quad 0 \leqslant t \leqslant 1,$$

и прямыми вычислениями получаем

$$\mathbf{E}\eta_n = \frac{1}{n+1}, \qquad \mathrm{Var}(\eta_n) = \frac{n}{(n+1)^2(n+2)}.$$

Дисперсия имеет порядок $\frac{1}{n^2}$. Мы уточним это свойство в терминах хвостов липшицевых функций от η_n.

Лемма 3.1. *Для любой функции* $g : (0,1) \to \mathbf{R}$ *такой, что* $\|g\|_{\mathrm{Lip}} \leqslant \sigma$ *и* $\mathbf{E}g(\eta_n) = 0$,

$$\mathbf{E}e^{\lambda g(\eta_n)} \leqslant e^{4\lambda^2\sigma^2/n^2}, \quad \textit{для} \quad |\lambda|\,\sigma \leqslant \frac{n}{2}. \tag{3.1}$$

Доказательство. Ввиду однородности (3.1) по g можно считать, что $\sigma = 1$.

Сначала полезно представить меру F_n как липшицево преобразование канонической двусторонней экспоненциальной меры ν на вещественной прямой с плотностью

$$p(t) = \frac{1}{2}\,e^{-|t|}, \quad t \in \mathbf{R}.$$

Пусть

$$F(t) = \int_{-\infty}^{t} p(s)\,ds$$

обозначает соответствующую функцию распределения, которая имеет обратную функцию F^{-1} : $(0,1) \to \mathbf{R}$. Легко проверить, что

$$I(t) \equiv p(F^{-1}(t)) = \min\{t, 1-t\}, \qquad 0 < t < 1.$$

Аналогично определим

$$I_n(t) \equiv p_n(F_n^{-1}(t)) = n(1-t)^{(n-1)/n}, \qquad 0 < t < 1,$$

где $p_n = F_n'$ — плотность η_n. Рассмотрим возрастающее отображение $T_n : \mathbf{R} \to (0,1)$, которое преобразует ν в меру F_n, т.е. $T_n(t) = F_n^{-1}(F(t))$. Тогда $T_n'(F^{-1}(t)) = I(t)/I_n(t)$, так что липшицева полунорма T_n записывается так:

$$\|T_n\|_{\mathrm{Lip}} = \sup_{0<t<1} \frac{I(t)}{I_n(t)} = \frac{I(\frac{1}{2})}{I_n(\frac{1}{2})} = \frac{2^{1/n}}{n}.$$

Следовательно, липшицева полунорма суперпозиции $h = g(T_n)$ будет не более, чем $\frac{2^{1/n}}{n}$.

Теперь мы используем следующее свойство меры ν (см. [**8**]), которое является эквивалентной переформулировкой одномерного изопериметрического неравенства

$$\nu(A + (-r,r)) \geqslant F(F^{-1}(\nu(A)) + r), \quad r > 0.$$

Именно, для любой функции $h : \mathbf{R} \to \mathbf{R}$ такой, что $\|h\|_{\mathrm{Lip}} \leqslant b$, существует неубывающая функция $\widetilde{h} : \mathbf{R} \to \mathbf{R}$, которая имеет такое же распределение относительно ν, как h, и такое, что

$$\|\widetilde{h}\|_{\mathrm{Lip}} \leqslant b.$$

С другой стороны (см. [**9**, лемма 2]), для заданных $\lambda \in \mathbf{R}$ и случайной величины η с конечным первым моментом величина

$$\mathbf{E}\, e^{\lambda(h(\eta)-\mathbf{E}\, h(\eta))}$$

корректно определена для всех неубывающих $h : \mathbf{R} \to \mathbf{R}$ таких, что $\|h\|_{\mathrm{Lip}} \leqslant b$ и принимает максимум в этом классе на линейной функции $h(t) = bt$. Предположим, что η распределено согласно ν, так что можно положить

$$\eta_n = T_n(\eta).$$

Применяя это свойство к функции $\widetilde{h}$, которая соответствует $h = g(T_n)$ и имеет липшицеву полунорму не более, чем $\dfrac{2^{1/n}}{n}$, заключаем, что при $|\lambda| < n/2^{1/n}$

$$\mathbf{E}\, e^{\lambda g(\eta_n)} = \mathbf{E}\, e^{\lambda \widetilde{h}(\eta)} \leqslant \mathbf{E}\, e^{2^{1/n}\lambda\eta/n} = \frac{1}{1-(\frac{2^{1/n}\lambda}{n})^2}.$$

Можно теперь воспользоваться простой оценкой

$$\frac{1}{1-s} \leqslant e^{2s}$$

при $0 \leqslant s \leqslant \frac{1}{2}$, из которой получаем

$$\mathbf{E}\, e^{\lambda g(\eta_n)} \leqslant e^{2^{1+2/n}\,\lambda^2/n^2},$$

если

$$\Big(\frac{2^{1/n}\lambda}{n}\Big)^2 \leqslant \frac{1}{2},$$

т.е.

$$|\lambda| \leqslant n\, 2^{-(\frac{1}{2}+\frac{1}{n})}.$$

Если $n \geqslant 2$, эта область значений покрывает интервал $|\lambda| \leqslant n/2$, и мы получаем (3.1).

Замечание 3.2. Утверждение леммы 3.1 можно усилить при $n = 1$. Тогда случайная величина η_1 равномерно распределена на (0,1), и, по аналогичным причинам, для всех $\lambda \in \mathbf{R}$

$$\mathbf{E}e^{\lambda g(\eta_1)} \leqslant \mathbf{E}e^{\lambda\sigma(\eta_1-\frac{1}{2})} = \frac{\mathrm{sh}(\lambda\sigma/2)}{\lambda\sigma/2} \leqslant e^{\lambda^2\sigma^2/8}.$$

Следующее наблюдение, полученное прямым приложением теоремы Фубини, показывает, как мера F_n соотносится с равномерным распределением μ_n на симплексе Δ_n.

Лемма 3.3. *Пусть ξ_n — случайный вектор в $\mathbf{R}^n$, распределенный согласно μ_n. Тогда для всех $n \geqslant 2$ случайные векторы*

$$((1-\eta_n)\xi_{n-1}, \eta_n), \quad \xi_n$$

равнораспределены, если ξ_{n-1} и η_n независимы.

4. Отклонения липшицевых функций на симплексе

Теперь мы в состоянии перейти к изучению задач об отклонении на симплексе и доказать функциональный вариант теоремы 1.3. Напомним, что $\Delta_n = \{x \in \mathbf{R}^n_+ : x_1 + \cdots + x_n \leqslant 1\}$ снабжено равномерной мерой μ_n.

Теорема 4.1. *Для каждой функции f на Δ_n такой, что*

$$|f(x) - f(y)| \leqslant \|x-y\|_1,$$

если $x, y \in \Delta_n$, имеем

$$\mu_n\Big\{\Big|f - \int f\, d\mu_n\Big| \geqslant r\Big\} \leqslant 2\, e^{-cnr^2}, \quad r > 0, \tag{4.1}$$

где $c > 0$ — константа.

Можно взять, например, $c = 1/144$.

Это неравенство можно переписать в эквивалентном виде через норму Орлича $\|\cdot\|_{\psi_2}$, порожденную функцией Юнга $\psi_2(t) = e^{t^2} - 1$, следующим образом:

$$\left\| f - \int f\,d\mu_n \right\|_{\psi_2} \leqslant \frac{C}{\sqrt{n}},$$

где C — константа. В частности, μ_n-дисперсия таких функций ограничена C^2/n (с некоторой другой константой).

Доказательство теоремы 4.1. Можно считать, что f гладкая и

$$\max_{1\leqslant i\leqslant n} \left| \frac{\partial f(x)}{\partial x_i} \right| \leqslant \sigma \tag{4.2}$$

в $\mathrm{int}\,(\Delta_n)$ симплекса (где σ — положительный параметр). Используя индукцию по n, мы проверим, что при условии (4.2) справедливо неравенство

$$\mathbf{E}\, e^{\lambda f(\xi_n)} \leqslant e^{\lambda \mathbf{E} f(\xi_n) + c_n \lambda^2 \sigma^2} \quad \text{при} \quad |\lambda|\,\sigma \leqslant \lambda_n, \tag{4.3}$$

для соответствующим образом выбранных последовательностей c_n и λ_n, где ξ_n — случайный вектор в $\mathbf{R}^n$ с распределением μ_n.

Согласно замечанию 3.2 можно выбрать $c_1 = 1/8$, что подходит при любом $\lambda_1 > 0$.

Шаг индукции. Пусть $n \geqslant 2$. Поставим в соответствие $f(x)$, $x = (x_1, \ldots, x_{n-1}, x_n) \in \Delta_n$, функции

$$g(u,t) = f((1-t)u, t), \qquad h(u) = \int_0^1 g(u,t)\,dF_n(t) = \mathbf{E} g(u, \eta_n),$$

где $u = (u_1, \ldots, u_{n-1}) \in \Delta_{n-1}$, $0 < t < 1$, и где случайный вектор η_n имеет распределение F_n, как в предыдущем параграфе. Тогда для каждого $i = 1, \ldots, n-1$

$$\frac{\partial g(u,t)}{\partial u_i} = (1-t)\,\frac{\partial f((1-t)u,t)}{\partial x_i}$$

что в силу (4.2) не превышает $\sigma(1-t)$ по абсолютному значению. Следовательно,

$$\left| \frac{\partial h(u)}{\partial u_i} \right| \leqslant \sigma \int_0^1 (1-t)\,dF_n(t) = \sigma\,\frac{n}{n+1}. \tag{4.4}$$

Кроме того,

$$\frac{\partial g(u,t)}{\partial t} = -\sum_{i=1}^{n-1} \frac{\partial f((1-t)u,t)}{\partial x_i}\,u_i + \frac{\partial f((1-t)u,t)}{\partial x_n},$$

откуда опять в силу (4.2) имеем

$$\left| \frac{\partial g(u,t)}{\partial t} \right| \leqslant \sigma\left(\sum_{i=1}^{n-1} u_i + 1 \right) \leqslant 2\sigma.$$

Поэтому в силу леммы 3.1, примененной к функции $t \to g(u,t)$ с 2σ вместо липшецевой полунормы для всех $u \in \Delta_{n-1}$

$$\mathbf{E}\, e^{\lambda g(u,\eta_n)} \leqslant e^{\lambda \mathbf{E} g(u,\eta_n) + 16\lambda^2\sigma^2/n^2} = e^{\lambda h(u)}\, e^{16\lambda^2\sigma^2/n^2}, \qquad |\lambda|\,\sigma \leqslant \frac{n}{4}. \tag{4.5}$$

Согласно (4.4) относительно ℓ^1-расстояния на Δ_{n-1} функция h имеет липшицеву константу, ограниченную величиной $\sigma\,\frac{n}{n+1}$. Поэтому согласно предположению индукции (4.3) при размерности $n-1$ имеем

$$\mathbf{E}\, e^{\lambda h(\xi_{n-1})} \leqslant e^{\lambda \mathbf{E} h(\xi_{n-1}) + c_{n-1} (\frac{n}{n+1})^2 \lambda^2 \sigma^2}, \quad \text{для} \quad |\lambda|\,\sigma \leqslant \frac{n+1}{n}\,\lambda_{n-1}, \tag{4.6}$$

где ξ_{n-1} имеет распределение μ_{n-1} (предполагается, что ξ_{n-1} не зависит от η_n). Подставив $u = \xi_{n-1}$ в (4.5), интегрируя по μ_{n-1} и применяя (4.6), получаем неравенство

$$\mathbf{E}\, e^{\lambda g(\xi_{n-1},\eta_n)} \leqslant e^{\lambda \mathbf{E} h(\xi_{n-1})}\, e^{\left[c_{n-1}(\frac{n}{n+1})^2+\frac{16}{n^2}\right]\lambda^2\sigma^2}, \tag{4.7}$$

справедливое для всех λ таких, что

$$|\lambda|\,\sigma \leqslant \min\left\{\frac{n}{4}, \frac{n+1}{n}\,\lambda_{n-1}\right\}. \tag{4.8}$$

Однако по лемме 3.3 случайные величины $f(\xi_n)$ и $g(\xi_{n-1},\eta_n)$ равнораспределенные. Кроме того,

$$\mathbf{E}h(\xi_{n-1}) = \mathbf{E}g(\xi_{n-1},\eta_n) = \mathbf{E}f(\xi_n).$$

Поэтому из (4.7) и (4.8) вытекает требуемое утверждение (4.3) при размерности n, если

$$c_n \geqslant c_{n-1}\left(\frac{n}{n+1}\right)^2 + \frac{16}{n^2}, \qquad \lambda_n \leqslant \min\left\{\frac{n}{4}, \frac{n+1}{n}\,\lambda_{n-1}\right\}.$$

Для второго неравенства можно положить $\lambda_n = \alpha(n+1)$, где

$$\alpha(n+1) \leqslant \frac{n}{4}.$$

Поскольку $n \geqslant 2$, оптимальным будет выбор $\alpha = 1/6$.

Первое неравенство можно переписать в терминах $d_n = (n+1)\,c_n$ следующим образом:

$$d_n \geqslant \left(1 - \frac{1}{n+1}\right) d_{n-1} + \frac{16\,(n+1)}{n^2}.$$

Следовательно, чтобы доказать, что оптимальное значение d_n ограничено константой d (что можно выполнить индукцией по n), достаточно оценить через d величину $\dfrac{16\,(n+1)^2}{n^2}$. Наихудшая ситуация соответствует $n = 2$, когда $d = 36$. Поэтому можно положить $c_n = 36/(n+1)$.

Ограничиваясь случаем $\sigma = 1$ и $\mathbf{E}f(\xi_n) = 0$, мы тем самым установили, что если

$$\max_{1\leqslant i\leqslant n}\left|\frac{\partial f(x)}{\partial x_i}\right| \leqslant 1 \tag{4.9}$$

в $\mathrm{int}\,(\Delta_n)$, то

$$\mathbf{E}e^{\lambda f(\xi_n)} \leqslant e^{36\,\lambda^2/(n+1)}, \tag{4.10}$$

если

$$|\lambda| \leqslant \frac{n+1}{6}.$$

Однако область значений λ достаточно велика, так чтобы можно было продолжить (4.10) на все λ вещественной прямой. Действительно, для некоторой точки $x_0 \in \Delta_n$ необходимо $f(x_0) = 0$ (так как f имеет нулевое среднее). Поскольку $|f(x) - f(y)| \leqslant \|x-y\|_1$, функция f должна принимать значения в $[-1,1]$, поэтому всегда имеем

$$\int e^{\lambda f}\,d\mu_n \leqslant e^{|\lambda|} \leqslant e^{36\,\lambda^2/(n+1)},$$

где второе неравенство справедливо при $|\lambda| \geqslant (n+1)/36$. Ввиду этого можно также покрыть остальные значения λ в (4.10).

Наконец, начиная с (4.10) и применяя неравенство Чебышева, получаем для любого $r > 0$

$$\mathrm{Prob}\,\{f(\xi_n) \geqslant r\} \leqslant e^{-(n+1)\,r^2/144}, \tag{4.11}$$

и, комбинируя с аналогичным неравенством для функции $-f$,

$$\mathrm{Prob}\,\{|f(\xi_n)| \geqslant r\} \leqslant 2\,e^{-(n+1)\,r^2/144}.$$

Последнее неравенство приводит к (4.1) при $c = 1/144$.

На этом шаге условие (4.9) можно слегка ослабить до условия Липшица

$$|f(x) - f(y)| \leqslant \|x-y\|_1 \quad (x, y \in \Delta_n).$$

Теорема 4.1 доказана. $\square$

5. Гауссова концентрация на симплексе

Переход от функциональных неравенств (как (4.1) в теореме 4.1) к неравенствам концентрации (как (1.6) в теореме 1.3) стандартен. Напомним ход рассуждений. Предпочтительно начать с односторонней оценки (4.11), которую перепишем в виде

$$\mu_n\left\{f - \int f\,d\mu_n \geqslant r\right\} \leqslant e^{-cnr^2}, \quad r > 0. \tag{5.1}$$

Эта оценка справедлива при $c = 1/144$ для любой функции f на Δ_n такой, что $|f(x) - f(y)| \leqslant \|x - y\|_1$, если $x, y \in \Delta_n$.

Прежде всего напомним следующий факт. Для заданной случайной величины $\xi \geqslant 0$ с медианой в нуле, т.е. такой, что

$$\mathrm{Prob}\,\{\xi > 0\} \leqslant \frac{1}{2},$$

имеем

$$\mathbf{E}\xi \leqslant \sqrt{\mathrm{Var}(\xi)}. \tag{5.2}$$

Действительно, полагая $F = \{\xi > 0\}$ и используя неравенство Коши, получаем соотношение

$$(\mathbf{E}\xi)^2 = (\mathbf{E}\,\xi\,1_F)^2 \leqslant \mathbf{E}\xi^2\,\mathrm{Prob}\,(F) \leqslant \frac{1}{2}\,\mathbf{E}\xi^2,$$

которое в точности такое же, как (5.2).

Как уже упоминалось, из неравенства (4.1) получаем с точностью до константы $C > 0$ следующее неравенство:

$$\mathrm{Var}_{\mu_n}(f) \leqslant \frac{C^2}{n}, \tag{5.3}$$

где Var_{μ_n} обозначает дисперсию относительно меры μ_n. В частности, его можно применить к функции расстояния

$$f_A(x) \,=\, \mathrm{dist}(A, x) \,=\, \inf_{y \in A} \|x - y\|_1, \qquad x \in \Delta_n,$$

для непустых подмножеств A из Δ_n. Если, кроме того,

$$\mu_n(A) \geqslant \frac{1}{2},$$

то, комбинируя (5.2) и (5.3), получаем

$$\mathbf{E}_{\mu_n} f_A = \int f_A(x)\,d\mu_n(x) \leqslant \frac{C}{\sqrt{n}}.$$

Наконец, для любого

$$r \geqslant \frac{2C}{\sqrt{n}}$$

имеем

$$\mu_n\{f_A \geqslant r\} \leqslant \mu_n\left\{f_A - \mathbf{E}_{\mu_n} f_A \geqslant r - \frac{C}{\sqrt{n}}\right\} \leqslant \mu_n\left\{f_A - \mathbf{E}_{\mu_n} f_A \geqslant \frac{r}{2}\right\} \leqslant \; e^{-cnr^2/4},$$

где на последнем шаге мы применили (5.1). Полученное неравенство остается верным при

$$0 < r < \frac{2C}{\sqrt{n}}$$

с возможно меньшей константой c, так как

$$e^{-cnr^2/4} \geqslant \frac{1}{2} \geqslant \mu_n\{f_A \geqslant r\}.$$

Остается заметить, что $\{f_A \geqslant r\}$ является дополнением множества $A + rB_1$ в Δ_n, т.е. мы получили неравенство концентрации вида

$$1 - \mu_n(A + rB_1) \leqslant e^{-cnr^2}, \quad r > 0. \tag{5.4}$$

Замечание 5.1. Отметим также изопериметрическое неравенство Содина [**10**], которое доставляет аналогичное неравенство концентрации относительно евклидова расстояния:

$$1 - \mu_n(A + rB_2) \leqslant e^{-cnr}, \quad r > 0. \tag{5.5}$$

Воспользуясь $B_2 \subset \sqrt{n}\, B_1$ и выполняя замену переменных, на основе (5.5) получаем экспоненциальную оценку

$$1 - \mu_n(A + rB_1) \leqslant \ e^{-c\sqrt{n}r},$$

которая, однако, слабее оценки (5.4).

Литература

1. B. S. Cirel'son, I. A. Ibragimov, V. N. Sudakov, "Norms of Gaussian sample functions", In: *Proceedings of the Third Japan-USSR Symposium on Probability Theory (Tashkent, 1975)*, pp. 20–41. *Lect. Notes Math.* **550**, Springer, Berlin (1976).
2. M. Ledoux, *The Concentration of Measure Phenomenon* Am. Math. Soc., Providence, RI (2001).
3. J. Arias-de-Reyna, R. Villa, "The uniform concentration of measure phenomenon in ℓ_p^n $(1 \leqslant p \leqslant 2)$", In: *Geometric Aspects of Functional Analysis*, pp. 13–18. *Lect. Notes Math.* **1745**, Springer, Berlin (2000).
4. G. Schechtman, "An editorial comment on the preceding paper: "The uniform concentration of measure phenomenon in ℓ_p^n $(1 \leqslant p \leqslant 2)$" by J. Arias-de-Reyna and R. Villa", In: *Geometric Aspects of Functional Analysis*, pp. 19–20. *Lect. Notes Math.* **1745**, Springer, Berlin (2000).
5. M. Talagrand, "A new isoperimetric inequality and the concentration of measure phenomenon" In: *Geometric Aspects of Functional Analysis (1989–90)*, pp. 94–124. *Lect. Notes Math.* **1469**, Springer, Berlin (1991).
6. B. Maurey, "Some deviation inequalities", *Geom. Funct. Anal.* **1**, No. 2, 188–197 (1991).
7. S. Bobkov, M. Ledoux, "Poincaré's inequalities and Talagrand's concentration phenomenon for the exponential distribution", *Probab. Theory Related Fields* **107**, No. 3, 383–400 (1997).
8. S. G. Bobkov, "Extremal properties of half-spaces for log-concave distributions", *Ann. Probab.* **24**, No. 1, 35–48 (1996).
9. S. G. Bobkov, "Локализованное доказательство изопериметрического неравенства Бэкри — Леду и некоторые приложения", *Теор. вероятн. примен.* **47**, No. 2, 340–346 (2002).
10. S. Sodin, "An isoperimetric inequality on the l_p balls", *Ann. Inst. Henri Poincaré Probab. Stat.* **44**, No. 2, 362–373 (2008).

Статья поступила в редакцию 7 января 2010 г.