Вестник Сыктывкарского университета.
Сер.1.Вып.1.1995

УДК 517.2

О НЕРАВЕНСТВАХ ГРОССА И ТАЛАГРАНА НА ДИСКРЕТНОМ КУБЕ[1]

С.Г.Бобков

В работе используется подход Дейвиса-Саймона для получения оценок больших уклонений для выпуклых липшицевых функционалов от бернуллиевских независимых случайных величин. Изучаются соотношения между неравенствами Гросса и Талаграна на дискретном кубе.

Обозначим через P_n равномерное распределение на $\Omega_n = \{-1,1\}^n$, приписывающее каждой точке $x \in \Omega_n$ массу $P_n(\{x\}) = 2^{-n}$. Пусть $f_A(x)$ означает кратчайшее евклидово расстояние от точки x до непустого множества $A \subset \mathbb{R}^n$. Если A выпукло, то справедливо следующее неравенство М.Талаграна [8]:

$$\mathrm{E}\,\mathrm{e}^{f_A^2/8} \leq \frac{1}{P_n(A)}, \tag{1}$$

где математическое ожидание (интеграл) понимается в смысле меры P_n. Так как $\{f_A < h\} = A^h$ $(h > 0)$ есть h-окрестность A и по неравенству Чебышева $P_n\{f_A \geq h\} \leq (\mathrm{E}\,\mathrm{e}^{f_A^2/8})\mathrm{e}^{-h^2/8}$, получаем неравенство изопериметрического типа:

$$P_n(A^h) \geq 1 - \frac{1}{P_n(A)}\mathrm{e}^{-h^2/8}. \tag{2}$$

Неравенство (2) может быть записано также на языке функций. Если $f : \mathbb{R}^n \to \mathbb{R}$ – выпуклая липшицева функция, $||f||_{Lip} \leq 1$,

[1]Работа выполнена при частичной поддержке Международного Научного Фонда, грант №NXZ000, Российского Фонда фундаментальных исследований, грант №93-011-1454, Конкурсного Центра фундаментального естествознания при Санкт-Петербургском госуниверситете и Грантового Центра по исследованиям в области математики при Новосибирском госуниверситете.

то применяя (2) к множествам вида $A = \{f \leq m_p\}$, где m_p – какая-нибудь) квантиль f порядка $p \in (0,1)$, т.е., такое число, что $P_n\{f \leq m_p\} \geq p$, $P_n\{f \geq m_p\} \geq 1-p$, и замечая, что $A^h \subset \{f < m_p + h\}$, приходим к неравенствам для уклонений f:

$$P_n\{f - m_p \geq h\} \leq \frac{1}{p}\mathrm{e}^{-h^2/8}. \tag{3}$$

Аналогично рассматривая $A = \{f \leq m_p - h\}$, имеем

$$P_n\{f - m_p \leq -h\} \leq \frac{1}{p}\mathrm{e}^{-h^2/8}. \tag{4}$$

Обратно, (3) превращается в (2) на функциях вида $f = f_A$. Неравенства (1)–(4) были улучшены М.Талаграном в [9]. В частности, при $p = \frac{1}{2}$, когда $m = m_p$ – медиана f, и при $h \geq \sqrt{\log 2}$ (2) и (3) уточнены следующим образом [9, с.127]:

$$P_n(A^h) \ \geq \ 1 - \exp\left\{-\frac{1}{4}\left(h - \sqrt{\log 2}\right)^2\right\}, \tag{5}$$

$$P_n(f - m \geq h) \ \leq \ \exp\left\{-\frac{1}{4}\left(h - \sqrt{\log 2}\right)^2\right\}. \tag{6}$$

Предположение о выпуклости A и f существенно во всех написанных неравенствах: если минимизировать левую часть (2) в классе всех множеств $A \subset \mathbb{R}^n$ с фиксированной мерой $p = P_n(A)$ (эта задача решена Харпером в [6]), то окажется, что с ростом n наименьшее значение $P_n(A^h)$ стремится к p, в то время, как правая часть (2) – почти 1 при больших h вне зависимости от n. Наименьшее значение $P_n(A^h)$ на классе выпуклых A с заданной мерой до сих пор неизвестно, так же, как и неизвестно, можно ли заменить правую часть (5) и (6) оценкой $K\mathrm{e}^{-h^2/4}$ с некоторой универсальной постоянной K. Но известно, что коэффициент $-1/4$ при h^2 не может быть улучшен [8, сс.60–61]. Тем не менее, если m в (6) заменить математическим ожиданием $\mathrm{E}\, f$, то $\sqrt{\log 2}$ можно опустить:

$$P_n(f - \mathrm{E}\, f \geq h) \leq \mathrm{e}^{-h^2/4}. \tag{7}$$

В данной заметке мы покажем, как (7) и другие неравенства, близкие к (1)–(6), можно получить в качестве следствий логарифмического неравенства Л.Гросса, причем условие выпуклости f в (3)–(4)

и (6)–(7) может быть немного ослаблено. Обозначим через $s_i(x)$, где $x \in \Omega_n$, $1 \leq i \leq n$, точку с координатами x_j при $j \neq i$, и $-x_i$ при $j = i$ (точка $s_i(x)$ – сосед x по i-ой координате). Для каждой функции $f : \Omega_n \to \mathbb{R}$ определен ее дискретный градиент ∇f, так что

$$|\nabla f(x)|^2 = \sum_{i=1}^{n} \left| \frac{f(x) - f(s_i(x))}{2} \right|^2.$$

Имеет место следующее неравенство, справедливое для всех функций f на Ω_n.

$$\mathrm{E}\, f^2 \log f^2 - \mathrm{E}\, f^2 \log \mathrm{E}\, f^2 \leq 2\mathrm{E}\, |\nabla f|^2. \tag{8}$$

(8) доказано Л.Гроссом в [3]. В случае $n = 1$, полагая $a = f(-1)$, $b = f(1)$, (8) превращается в элементарное неравенство

$$\frac{a^2 \log a^2 + b^2 \log b^2}{2} - \frac{a^2 + b^2}{2} \log \frac{a^2 + b^2}{2} \leq \frac{(a-b)^2}{2}.$$

При $n > 1$, доказательство (8) легко провести по индукции. Если его применить к функциям вида $f_n = g(\frac{x_1 + \ldots + x_n}{\sqrt{n}})$, где g – гладкая функция на прямой с ограниченной второй производной, заметить, что $|\nabla f_n(x)|^2 = g'(\frac{x_1 + \ldots + x_n}{\sqrt{n}})^2 + O(\frac{1}{n})$ и, устремляя $n \to \infty$, применить центральную предельную теорему, то в пределе (8) запишется в виде

$$\mathrm{E}_\gamma g^2 \log g^2 - \mathrm{E}_\gamma g^2 \log \mathrm{E}_\gamma g^2 \leq 2\mathrm{E}_\gamma (g')^2, \tag{9}$$

где E_γ означает математическое ожидание по отношению к стандартной гауссовской мере γ на $\mathbb{R}$. Обычно, неравенством Гросса называют (9) и его многомерный аналог. Вышеупомянутое рассуждение приведено им в [4] и представляет собой упрощение первоначального доказательства (9), также основанного на (8). Им также отмечено [4, с.21], что (8) при $n = 1$ было ранее открыто в одной работе Бонами [2], долго остававшейся незамеченной многими исследователями, включая и самого Гросса. Публикация же Гросса [3] вызвала живой интерес, о чем можно судить по обширной библиографии (перечень работ, мотивированных (9), займет несколько страниц). Многие исследуемые вопросы концентрируются вокруг следующего: что можно сказать о вероятностном распределении μ в $\mathbb{R}^n$, удовлетворяющем неравенству

$$\mathrm{E}_\mu f^2 \log f^2 - \mathrm{E}_\mu f^2 \log \mathrm{E}_\mu f^2 \leq 2\sigma^2 \mathrm{E}_\mu |Df|^2, \tag{10}$$

где f – произвольная гладкая функция на $\mathbb{R}^n$ и Df – ее обычный градиент, $|Df(x)|^2 = \sum_1^n |\partial f(x)/\partial x_i|^2$, а математической ожидание понимается в смысле меры μ. Применим (10) к функции вида $f = e^{tg/2}$, где g – ограниченная гладкая липшицева функция, $|Dg(x)| \le 1$ для всех $x \in \mathbb{R}^n$ (что означает, что $||g||_{Lip} \le 1$). Введем функцию u равенством $e^{tu(t)} = \mathrm{E}_\mu e^{tg} = \mathrm{E}_\mu f^2$, $t \in \mathbb{R}$. Замечая, что

$$\mathrm{E}_\mu f^2 \log f^2 = \frac{d}{dt}\mathrm{E}_\mu f^2, \qquad |Df|^2 = \frac{t^2}{4}|Dg|^2 f^2,$$

(10) примет вид $t^2 u'(t)\mathrm{E}_\mu f^2 \le \frac{\sigma^2 t^2}{2}\mathrm{E}_\mu |Dg|^2 f^2$, и так как $|Dg| \le 1$, получаем $u'(t) \le \frac{\sigma^2}{2}$. Разложения Тейлора $\mathrm{E}_\mu e^{tg} = 1 + t\mathrm{E}_\mu g + O(t^2)$ и $e^{tu(t)} = 1 + tu(t) + O(t^2)$ дают $u(0) = \mathrm{E}_\mu g$. Следовательно, для любого $t \ge 0$ имеем

$$u(t) - \mathrm{E}_\mu g \le \frac{\sigma^2 t}{2}.$$

Окончательно получаем

$$\mathrm{E}_\mu e^{tg} \le \exp\left\{t\mathrm{E}_\mu g + \frac{\sigma^2 t^2}{2}\right\}, \quad t \ge 0. \tag{11}$$

Впервые такое рассуждение приведено М.Леду [7, с.101], но идея получения дифференциальных неравенств приписывается Хербсту(I.Herbst, не опубликовано) и Дейвису и Саймону [5] (допустившим небольшую ошибку при решении одного дифференциального неравенства).Исправляя неточности, Аида, Масуда и Шигекава [1] провели рассуждение для функций вида $f = e^{tg^2}$ и получили следующее неравенство (теорема 3.3):

$$\mathrm{E}_\mu e^{tg^2} \le \exp\left\{\frac{t}{1 - 2\sigma^2 t}\mathrm{E}_\mu g^2\right\}, \quad 0 \le t < 1/(2\sigma^2). \tag{12}$$

Покажем теперь, как точно такие же рассуждения можно провести в случае дискретного куба применительно к выпуклым функциям, причем свойство выпуклости можно понимать в некотором более широком смысле. Обозначим через $\overline{\Omega}_n$ объединение всех ребер куба $[-1,1]^n$. Будем говорить, что функция $f : \overline{\Omega}_n \to \mathbb{R}$ квазивыпукла, если она непрерывна и выпукла на каждом ребре. В каждой вершине $x \in \Omega_n$ определены частные производные $\partial f/\partial x_i$: на ребре Δ_i, соединяющим x и $s_i(x)$, f имеет неубывающую производную g_i.

Полагаем

$$\frac{\partial f(x)}{\partial x_i} = \lim_{\substack{y \in \Delta_i \\ y \neq x \\ y \to x}} g_i(y), \qquad |Df(x)|^2 = \sum_{i=1}^{n} \left|\frac{\partial f(x)}{\partial x_i}\right|^2.$$

Заметим, что если h – непрерывная выпуклая функция на $[-1,1]$, то

$$\left|\frac{h(1)-h(-1)}{2}\right|^2 \le h'(-1+0)^2 + h'(1-0)^2.$$

Следовательно, для квазивыпуклой функции f на $\overline{\Omega}_n$ имеем

$$\left|\frac{f(x)-f(s_i(x))}{2}\right|^2 \le \left|\frac{\partial f(x)}{\partial x_i}\right|^2 + \left|\frac{\partial f(s_i(x))}{\partial x_i}\right|^2. \tag{13}$$

Так как мера P_n инвариантна относительно отображения $x \to s_i(x)$, $\mathrm{E}\left|\frac{\partial f(s_i(x))}{\partial x_i}\right|^2 = \mathrm{E}\left|\frac{\partial f(x)}{\partial x_i}\right|^2$, поэтому, дискретное логарифмическое неравенство (8) и (13) влекут

$$\mathrm{E}\, f^2 \log f^2 - \mathrm{E}\, f^2 \log \mathrm{E}\, f^2 \le 4\mathrm{E}\,|Df|^2.$$

Следовательно, мера $\mu = P_n$ удовлетворяет (10) с $\sigma^2 = 2$ на классе квазивыпуклых функций. Пусть g – произвольная квазивыпуклая функция на $\overline{\Omega}_n$ с $|Dg(x)| \le 1$ для всех $x \in \Omega_n$. Тогда функция $f = \mathrm{e}^{tg/2}$ тоже квазивыпукла при всех $t \ge 0$, причем $|Df(x)|^2 = \frac{t^2}{4}|Dg(x)|^2 f(x)^2$. Следовательно, (10) можно применить к f при всех $t \ge 0$, и мы приходим к (11):

Теорема 1. *Для любой квазивыпуклой функции f на $\overline{\Omega}_n$ с $|Df| \le 1$ на Ω_n и для любых $t \ge 0$*

$$\mathrm{E}\,\mathrm{e}^{tf} \le \mathrm{e}^{t\mathrm{E}\, f + t^2}. \tag{14}$$

В частности, для всех $h \ge 0$

$$P_n\{f - \mathrm{E}\, f \ge h\} \le \mathrm{e}^{-h^2/4}. \tag{15}$$

Покажем, каким образом (15) следует из (14). По неравенству Чебышева $P\{f - \mathrm{E}\, f \ge h\} \le \mathrm{E}\,\mathrm{e}^{t(f-\mathrm{E}\, f)}\mathrm{e}^{-th} \le \mathrm{e}^{t^2-th}$. Минимизируя по

≥ 0 ($t = \frac{h}{2}$), получаем (15). Таким образом, (7) выполняется для всех гладких выпуклых $f : \mathbb{R}^n \to \mathbb{R}$ с $||f||_{Lip} \leq 1$, так как для гладких f $||f||_{Lip} = \sup_{x\in\mathbb{R}^n} |Df(x)|$. Поскольку гладкими выпуклыми функциями можно аппроксимировать по распределению любые выпуклые функции, (7) сохраняется для всех выпуклых f с $||f||_{Lip} \leq 1$.

Покажем теперь, как можно перейти от (15) к неравенствам изопериметрического типа, аналогичным (2)–(4). Применим (15) к функциям вида $f = f_A$, где $A \subset \mathbb{R}^n$ – выпуклое множество меры $p = P_n(A)$, $0 < p < 1$. Так как f выпукла и $||f||_{Lip} \leq 1$, имеем

$$1 - P_n(A^h) = P_n(f \geq h) \leq \mathrm{e}^{-\frac{1}{4}(h - \mathrm{E}\, f)^2}, \qquad h \geq \mathrm{E}\, f. \tag{16}$$

Чтобы оценить сверху $\mathrm{E}\, f$, воспользуемся следующим элементарным предложением: если ξ – случайная величина, такая, что $\xi \geq 0$, $P\{\xi = 0\} = p$, $\mathrm{D}\,\xi \leq \sigma^2$, то $\mathrm{E}\,\xi \leq \sqrt{\frac{1-p}{p}}\sigma$ ($\mathrm{D}\,\xi$ означает дисперсию $\mathrm{E}\,\xi^2 - (\mathrm{E}\,\xi)^2$). Функция f, как случайная величина, обладает перечисленными свойствами, причем $\sigma^2 = 2$. Действительно, если неравенство (10) применить к функциям вида $1 + \varepsilon f$ и устремить $\varepsilon \to 0$, то получится неравенство типа Пуанкаре (см. также [4], теорема 2.5):

$$\mathrm{E}_\mu f^2 - (\mathrm{E}_\mu f)^2 \leq \sigma^2 \mathrm{E}_\mu |Df|^2. \tag{17}$$

В частности, для липшицевых f с $|Df| \leq 1$ имеем $\mathrm{E}_\mu f^2 - (\mathrm{E}_\mu f)^2 \leq \sigma^2$. Кроме того, если μ удовлетворяет (10) на классе всех выпуклых f, то и (17) выполняется на классе всех выпуклых f, и тогда для всех выпуклых липшицевых f, $||f||_{Lip} \leq 1$, получаем, что $\mathrm{E}_\mu f^2 - (\mathrm{E}_\mu f)^2 \leq \sigma^2$. Для $\mu = P_n$ $\sigma^2 = 2$, поэтому заключаем для $f = f_A$, что $\mathrm{E}\, f \leq \sqrt{\frac{1-p}{p}}\sqrt{2}$. Таким образом, согласно (16), получаем следующее утверждение.

Следствие 2. *Для любого выпуклого множества $A \subset \mathbb{R}^n$ меры $P_n(A) = p$, $0 < p < 1$ и любого $h \geq \sqrt{\frac{2p}{1-p}}$*

$$P_n(A^h) \geq 1 - \exp\left\{ -\frac{1}{4}\left(h - \sqrt{\frac{2p}{1-p}}\right)^2 \right\}. \tag{18}$$

При $p = \frac{1}{2}$ (18) лишь постоянной $\sqrt{2} > \sqrt{\log 2}$ хуже, чем (5). С другой стороны, вряд ли из (1), (2) или (5) можно вывести (14) или (15). Отметим еще без доказательства, что дискретным аналогом

(12) для квазивыпуклых f на $\overline{\Omega}_n$ с $|Df| \le 1$ на Ω_n является (с учетом того, что $\sigma^2 = 2$ для $\mu = P_n$) неравенство

$$\mathrm{E}\, e^{tf^2} \le \exp\left\{\frac{t}{1-4t}\mathrm{E}\, f^2\right\}, \qquad 0 \le t < \frac{1}{4}. \tag{19}$$

Если $f = f_A$, где A – выпукло и имеет P_n-меру p, то легко вывести из (17) оценку для $\mathrm{E}\, f^2$: $\mathrm{E}\, f^2 \le \frac{\sigma^2}{p} = \frac{2}{p}$. Тогда при $t = \frac{1}{8}$ (19) дает

$$\mathrm{E}\, e^{f_A^2/8} \le e^{1/(2p)},$$

что немногим хуже (1).

Литература

1. **Aida, S., Masuda, T., Shigekawa, I.** Logarithmic Sobolev inequalities and exponential integrability. Preprint (1993).

2. **Bonami, A.** Etude des coefficients de Fourier des fonctions de $L^p(G)$//*Ann. Inst. Fourier.* V.20,№2. 1970. P.335–402.

3. **Gross, L.** Logarithmic Sobolev inequalities// *Amer. J. Math.* V.97. 1975. P.1061–1083.

4. **Gross, L.** Logarithmic Sobolev inequalities and contractivity properties of semigroups. Varenna,1992// *Lect. Notes in Math.* V.1563. 1993. P.54–88.

5. **Davies J.-D., Simon B.** Ultracontractivity and the heat kernel for Schrödinger operators and Dirichlet Laplacians// *J. Funct. Anal.* 59. 1984. P.335–395.

6. **Harper, L.H.** Optimal numbering and isoperimetric problems on graphs//*J. Comb. Theor.* V.1. 1966. P.385–393.

7. **Ledoux, M.** Isoperimetry and Gaussian Analysis. *Ecole d'été de Probabilités de Saint-Flour.* (1994).

8. **Talagrand, M.** An isoperimetric theorem on the cube and the Khinchine-Kahane inequalities// *Proc. Amer. Math. Soc.* V.104. 1988. P.905–909.

9. **Talagrand, M.** Concentration of measure and isoperimetric inequalities in product spaces. Preprint (1994).

Summary

Bobkov S.G. On inequalities of Gross and Talagrand on the discrete cube

A Davies-Simons's approach is used to give estimates for large deviations of convex Lipschitz functionals of Bernoulli independent random variables. Connections between Gross' and Talagrand's inequalities on the discrete cube are investigated.

Сыктывкарский университет

Поступила 8.02.95