Вестник Сыктывкарского университета.
Сер.1.Вып.3.1999

УДК 517.987

Замечания о неравенстве Громова–Мильмана [1]

С. Г. Бобков

Рассматриваются оптимальные оценки для функций концентрации и для уклонений липшицевых функций на метрических вероятностных пространствах, удовлетворяющих неравенствам типа Пуанкаре.

Пусть (M, ρ) – метрическое пространство, снабженное борелевской вероятностной мерой μ. Для функций g на M определим обобщенный модуль "градиента"

$$|\nabla g(x)| = \limsup_{\rho(x,y)\to 0} \frac{|g(x) - g(y)|}{\rho(x,y)}, \quad x \in M,$$

полагая $|\nabla g(x)| = 0$ для изолированных точек x в M. Легко видеть, что если ограничения g на шарах в M имеют конечную липшицеву константу (будем называть такие функции локально-липшицевыми), то функция $|\nabla g|$ конечна всюду и измерима по Борелю.

Говорят, что (M, ρ, μ) удовлетворяет неравенству типа Пуанкаре, если для всех локально-липшицевых функций g на M $\mathrm{E}|\nabla g|^2 < +\infty$ влечет $\mathrm{E}g^2 < +\infty$, и при этом

$$\lambda_1 \mathrm{Var}(g) \le \mathrm{E}|\nabla g|^2. \tag{1}$$

Здесь λ_1 – положительная постоянная, не зависящая от g, а математическое ожидание $\mathrm{E}g = \int g\, d\mu$ и дисперсия $\mathrm{Var}(g) = \mathrm{E}(g - \mathrm{E}g)^2$ понимаются в смысле меры μ.

Неравенства вида (1) широко распространены в римановой геометрии и теории пространств Соболева, главным образом в задачах, в

[1] При частичной поддержке Российского Фонда Фундаментальных Исследований, грант №96-01-00201

© Бобков С. Г., 1999.

которых наилучшее значение λ_1 интерпретируется как первое нетривиальное собственное число лапласиана. Для теории вероятностей и локальной теории банаховых пространств эти неравенства представляют интерес прежде всего ввиду их свойства аддитивности. Именно, без изменения в константе при некоторых условиях типа регулярности неравенство (1) распространяется на многомерные пространства M^n с продакт-мерой $\mu^n = \mu \times \ldots \times \mu$ и метрикой евклидового типа $\rho_n(x,y) = \sqrt{\sum_{i=1}^n \rho(x_i,y_i)^2}$, $x,y \in M^n$. Следовательно, любое соотношение между мерой μ и метрикой ρ, которое можно извлечь на основе (1) в терминах λ_1, будет справедливым для меры μ^n и метрики ρ_n. В частности, можно исследовать свойство концентрации продакт-мер μ^n, например, порядок убывания при $h \to \infty$ функции концентрации

$$\alpha(h) = \sup_n \sup_{\mu(A)\geq 1/2} 1 - \mu^n(A^h), \quad h > 0,$$

где (вторая) точная верхняя грань берется по всем борелевским множествам $A \subset M^n$ и

$$A^h = \{x \in M^n : \rho_n(x,a) < h \text{ для некоторого } a \in A\}$$

обозначает открытую h-окрестность A в (M^n, ρ_n). Впервые такой подход к исследованию концентрации был предложен Громовым и Мильманом [6], получившими неравенство

$$1 - \mu(A^h) \leq (1-p^2)e^{-\log(1+p)h\sqrt{\lambda_1}}, \quad p = \mu(A),\ A \subset M, \tag{2}$$

и вытекающую из нее оценку

$$\alpha(h) \leq Ke^{-ch\sqrt{\lambda_1}} \tag{3}$$

с постоянными $c = \log(3/2)$, $K = \frac{3}{4}$. Доказательство неравенства (2) (которое в силу упомянутого свойства аддитивности автоматически распространяется на многомерные пространства) было основано на применении (1) к функциям вида $g(x) = \min\{\rho(A^{k\delta}, x), \delta\}$ и последующей итерации по k с подходящим δ, зависящим от h и λ_1. Позднее с теми же рассуждениями Алон и Мильман [3] получили аналог неравенств (2)–(3) для графов.

Остался, однако, открытым вопрос о том, насколько оптимально неравенство (3). Тот факт, что экспоненциальное по h убывание в правой части (3) не может быть улучшено, легко обнаружить, рассматривая на прямой $M = \mathbf{R}$ экспоненциальные меры $\mu = \mu_{\lambda_1}$ с плотностью $\frac{d\mu_{\lambda_1}(x)}{dx} = \sqrt{\lambda_1}\exp\{-2\sqrt{\lambda_1}\,|x|\}$, $x \in \mathbf{R}$ (здесь λ_1 – положительный параметр). Известно, что мера μ_{λ_1} удовлетворяет неравенству типа Пуанкаре, причем с наилучшей постоянной λ_1 (впервые отмечено Боровковым, Утевым [5] и Клаассеном [7]). С другой стороны, так как множество $A = (-\infty, 0]$ имеет μ_{λ_1}-меру $1/2$, заключаем, что

$$\alpha(h) \geq 1 - \mu_{\lambda_1}(A^h) = \frac{1}{2}\, e^{-2h\sqrt{\lambda_1}}.$$

Следовательно, в неравенстве Громова–Мильмана с необходимостью $c \leq 2$. В данной заметке мы покажем, что при подходящем выборе постоянной K в (3) это неравенство остается справедливым для максимально возможного значения $c = 2$. Таким образом, в смысле концентрации экспоненциальные меры играют почти экстремальную роль в неравенствах типа Пуанкаре.

Теорема. *Предположим, что* (M, ρ, μ) *удовлетворяет неравенству типа Пуанкаре с постоянной* $\lambda_1 > 0$. *Тогда для всех борелевских множеств* $A \subset M$ *и всех* $h > 0$

$$1 - \mu(A^h) \leq \frac{9}{\mu(A)}\, e^{-2h\sqrt{\lambda_1}}. \tag{4}$$

Кроме того, для всех функций g *с липщицевой константой* $\|g\|_{\mathrm{Lip}} \leq 1$ *имеем* $\mathrm{E}|g| < +\infty$, *причем для всех* $h > 0$

$$\mu\{|g - \mathrm{E}g| \geq h\} \leq 6\, e^{-2h\sqrt{\lambda_1}}. \tag{5}$$

Очевидно, неравенства (4) и (5) эквивалентны друг другу с точностью до абсолютных множителей. В силу (5), для некоторых абсолютных постоянных $t_0 > 0$ и $C_0 > 0$

$$\mathrm{E}e^{2\sqrt{\lambda_1}t_0|g - \mathrm{E}g|} \leq C_0 \tag{6}$$

в классе всех функций с липшицевой константой $\|g\|_{\mathrm{Lip}} \leq 1$. Независимо от работы [6] такое неравенство с постоянными $t_0 = 1/24$, $C_0 = 2$ было доказано Боровковым и Утевым в [5], рассматривавшими, правда, неравенство Пуанкаре лишь для вероятностных мер μ на прямой $M = \mathbf{R}$. Используемый ими метод работает, однако, и в абстрактной ситуации и, в частности, дает (3) с худшей постоянной c в экспоненте: неравенство (1) применяется к функциям вида $|g - \mathrm{E}g|^p$ с условием $\|g\|_{\mathrm{Lip}} \leq 1$ и из получаемого рекуррентного соотношения выводятся оценки для моментов

$$(2\sqrt{\lambda_1}\,)^p\, \mathrm{E}|g - \mathrm{E}g|^p \leq (4p)^p, \quad p \geq 1,$$

откуда (6) сразу следует. Заметим, что в силу (5) имеет место правильная по порядку оценка $(2\sqrt{\lambda_1})^p \,\mathrm{E}|g - \mathrm{E}g|^p \le 6\,\Gamma(p+1)$ (где Γ – гамма-функция). Действительно:

$$\begin{aligned}(2\sqrt{\lambda_1})^p \,\mathrm{E}|g - \mathrm{E}g|^p &= (2\sqrt{\lambda_1})^p \int_0^\infty \mu\{|g - \mathrm{E}g| \ge h\}\, dh^p \\ &\le (2\sqrt{\lambda_1})^p \int_0^\infty 6\, e^{-2h\sqrt{\lambda_1}}\, dh^p = 6\,\Gamma(p+1).\end{aligned}$$

Вопрос об оптимальности неравенства Громова–Мильмана стал обсуждаться много позднее. Аида, Масуда и Шигекава [1], несколько модифицировав подход Громова–Мильмана, доказали конечность средних $\mathrm{E}e^{2t\sqrt{\lambda_1}|g-\mathrm{E}g|}$ при $t < t_0 = \frac{1}{2(e-1)}$. Дальнейшее уточнение было сделано в работе Аиды и Струка [2], ограничившимися, правда, рассмотрением только значения $t = 1/2$ при оценивании экспоненциальных моментов

$$u(t) = \mathrm{E}e^{2t\sqrt{\lambda_1}g}.$$

Предполагая, что $\|g\|_{\mathrm{Lip}} \le 1$ и $\mathrm{E}g = 0$, применение (1) к функциям вида $e^{t\sqrt{\lambda_1}g}$ приводит к функциональному неравенству

$$u(t) \le \frac{1}{1-t^2}\, u(t/2)^2,$$

и после n-кратного применения этого неравенства к значениям t, $\frac{t}{2}$, ..., $\frac{t}{2^{n-1}}$, получаем

$$u(t) \le \prod_{k=0}^{n-1} \frac{1}{(1-t^2/4^k)^{2^k}}\, u(t/2^n)^{2^n}. \tag{7}$$

Очевидно, произведение в (7) сходится при $|t| < 1$, и так как $u(\varepsilon) = 1 + o(\varepsilon)$ при $\varepsilon \to 0$ (в силу условия $\mathrm{E}g = 0$), приходим в пределе к оценке

$$u(t) \le U(t) = \prod_{n=0}^{\infty} \frac{1}{(1-t^2/4^n)^{2^n}}, \quad |t| < 1. \tag{8}$$

В частности, при $t = 1/2$ получаем неравенство

$$\mathrm{E}e^{\sqrt{\lambda_1}g} \le K_0 = \prod_{n=0}^{\infty} \frac{1}{(1-1/4^n)^{2^n}} = 1.720102\ldots. \tag{9}$$

Это неравенство, влекущее (6) с $t_0 = 1/2$ и (3) с $c = 1$, и есть результат Аиды и Струка (в общем же виде неравенство (8) появилось в

[4]). Чтобы рассуждения были совсем строгими, следует все выкладки проводить в предположении об ограниченности g (с тем, чтобы не возникало проблем с экспоненциальной интегрируемостью) и затем с помощью усечений легко получить (8)–(9) уже для всех липшицевых g на M с нулевым средним.

Таким образом, согласно (9), экстремальным значением в (6) является $t_0 = 1$. Впервые конечность экспоненциальных моментов $e^{2t\sqrt{\lambda_1}g}$ при $|t| < 1$ (как и выше – в предположении, что $\|g\|_{\mathrm{Lip}} \le 1$ и $\mathrm{E}g = 0$) была доказана Шмукеншлегером [8]. Применяя (1) к функциям вида $ge^{t\sqrt{\lambda_1}g}$ и анализируя получаемое дифференциальное неравенство относительно функции $u(t)$, он извлек оценку $u(t) \le (1-|t|)^{-4}$. Это, однако, несколько хуже, чем обобщенное неравенство Аиды и Струка (8): как показано в [4], $U(t) \le (1+|t|)/(1-|t|) \le (1-|t|)^{-4}$. В частности,

$$\mathrm{E}e^{2t\sqrt{\lambda_1}g} \le \frac{2}{1-|t|}, \quad |t| < 1. \tag{10}$$

Заметим, что правая часть в (10) с точностью до множителя эквивалентна функции U при $t \to 1$, и что с точностью до множителя это неравенство не может быть улучшено, как показывает все тот же пример экспоненциальных распределений μ_{λ_1} и функции $g(x) = x$, $x \in \mathbf{R}$. Однако неравенство (10) еще не влечет (5): если на вероятностном пространстве (M, μ) есть случайная величина ξ со свойствами $\mathrm{E}\xi = 0$, $\mathrm{E}e^{t\xi} \le \frac{C_1}{1-|t|}$ при $|t| < 1$, то по неравенству Чебышева можно получить оценки для уклонений вида $\mu\{|\xi| \ge h\} \le C_2(1+h)e^{-h}$, причем множитель $1+h$ не может быть опущен. Следовательно, нельзя надеяться получить оценки типа (5) и неравенство Громова–Мильмана (3) с наилучшей постоянной, основываясь только на оценках для экспоненциальных моментов, даже если они оптимальные.

Доказательство теоремы 1. При выводе (5) можно предполагать, что функция g ограничена. Применим неравенство типа Пуанкаре (1) к функции вида $g_h(x) = e^{\sqrt{\lambda_1}\min(g(x),h)}$, $x \in M$, $h \in \mathbf{R}$:

$$\lambda_1 \mathrm{E}g_h^2 \le \lambda_1 (\mathrm{E}g_h)^2 + \mathrm{E}|\nabla g_h|^2. \tag{11}$$

Чтобы оценить первый член в правой части (11), применим неравенство Аиды–Струка (9) к функции $\min(g,h) - \mathrm{E}\min(g,h)$:

$$\mathrm{E}g_h = \mathrm{E}e^{\sqrt{\lambda_1}\min(g,h)} \le K_0\, e^{\sqrt{\lambda_1}\mathrm{E}\min(g,h)} \le K_0\, e^{\sqrt{\lambda_1}\mathrm{E}g}.$$

Чтобы оценить второй член в правой части, заметим, что $|\nabla g_h| = 0$ на множестве $\{g > h\}$ (так как это множество открыто, а функция g_h на нем постоянна), и что на всем M

$$|\nabla g_h| \le \sqrt{\lambda_1} e^{\sqrt{\lambda_1} g_h} \le \sqrt{\lambda_1} e^{\sqrt{\lambda_1} g}$$

(так как $\|g_h\|_{\mathrm{Lip}} \le 1$). Следовательно, $\mathrm{E}|\nabla g_h|^2 \le \lambda_1 \mathrm{E} e^{2\sqrt{\lambda_1} g} 1_{\{g \le h\}}$. Используя эти оценки в (11) и записывая

$$\mathrm{E} g_h^2 = \mathrm{E} e^{2\sqrt{\lambda_1} g} 1_{\{g \le h\}} + e^{2\sqrt{\lambda_1} h} \mu\{g > h\},$$

приходим к неравенству $e^{2\sqrt{\lambda_1} h} \mu\{g > h\} \le K_0^2 e^{2\sqrt{\lambda_1} \mathrm{E} g}$. Так как его правая часть непрерывно зависит от h, строгое неравенство можно заменить на нестрогое:

$$\mu\{g \ge h\} \le K_0^2 e^{-2\sqrt{\lambda_1} h} e^{2\sqrt{\lambda_1} \mathrm{E} g}, \quad h \in \mathbf{R}. \tag{12}$$

Если записать это неравенство в виде $\mu\{g - \mathrm{E} g \ge h\} \le K_0^2 e^{-2\sqrt{\lambda_1} h}$ и сложить с аналогичным неравенством для функции $-g$, то при $h > 0$ мы придем к оценке

$$\mu\{|g - \mathrm{E} g| \ge h\} \le 2K_0^2 e^{-2\sqrt{\lambda_1} h}.$$

Это дает (5), так как $K_0^2 = 2.958750\ldots < 3$.

Чтобы вывести (4), предположим, что множество A непусто, и применим (12) при $h = 0$ к (липщицевой) функции $g(x) = -\rho(A, x)$, где, как обычно, $\rho(A, x) = \inf_{a \in A} \rho(a, x)$ обозначает кратчайшее расстояние от x до A. Так как $\mu\{g \ge 0\} = \mu(clos(A)) \ge \mu(A)$, получаем

$$e^{2\sqrt{\lambda_1} \mathrm{E} \rho(A,x)} \le \frac{K_0^2}{\mu(A)}.$$

Наконец, еще раз применим (12) с $h > 0$, но теперь к функции $g(x) = \rho(A, x)$. С учетом предыдущей оценки получаем:

$$\begin{aligned} 1 - \mu(A^h) &= \mu\{x \in M : \rho(A, x) \ge h\} \\ &\le K_0^2 e^{-2\sqrt{\lambda_1} h} e^{2\sqrt{\lambda_1} \mathrm{E} \rho(A,x)} \le \frac{K_0^4}{\mu(A)} e^{-2\sqrt{\lambda_1} h}. \end{aligned}$$

Остается заметить, что $K_0^4 < 9$. Теорема 1 доказана.

Замечание. При распространении неравенства (1) с пространства (M, ρ, μ) на многомерные пространства (M^n, ρ_n, μ^n) под свойствами типа "регулярности" мы понимаем следующие:

1) пространство M сепарабельно;

2) метрика ρ_n согласована с топологией произведения пространства M^n, причем для всех локально-липшицевых функций g на M^n в смысле ρ_n, для μ^n-почти всех $x = (x_1, \ldots, x_n) \in M^n$

$$|\nabla g(x)|^2 = \sum_{i=1}^{n} |\nabla_{x_i} g(x)|^2, \tag{13}$$

где $|\nabla g(x)|$ – модуль градиента в смысле ρ_n, а $|\nabla_{x_i} g|$ – модуль градиента по переменной $x_i \in M$ в смысле ρ для функции $x_i \to g(x_1, \ldots, x_n)$.

При этих условиях (M^n, ρ_n, μ^n) удовлетворяет неравенству типа Пуанкаре с той же постоянной λ_1. Доказательство можно провести по индукции, или просто можно воспользоваться одним общим свойством дисперсии: если (M, μ) – вероятностное пространство, и (M^n, μ^n) – его n-ая степень, то для всех μ^n-измеримых функций g на M^n

$$\mathrm{Var}_{\mu^n}(g) \leq \int_{M^n} \sum_{i=1}^{n} \mathrm{Var}_{x_i}(g)\, d\mu(x_1) \cdots d\mu(x_n). \tag{14}$$

Здесь $\mathrm{Var}_{x_i}(g)$ – μ-дисперсия функции $x_i \to g(x_1, \ldots, x_n)$ при фиксированных переменных x_j, $j \neq i$. Если теперь дана ρ_n–локально-липшицевая функция g на M^n, то дисперсию $\mathrm{Var}_{x_i}(g)$ можно оценить сверху, согласно "одномерному" неравенству (1), через $\frac{1}{\lambda_1}\mathrm{E}_{x_i}|\nabla_{x_i} g(x)|^2$, так что для всех $x = (x_1, \ldots, x_n) \in M^n$

$$\lambda_1 \sum_{i=1}^{n} \mathrm{Var}_{x_i}(g) \leq \sum_{i=1}^{n} |\nabla_{x_i} g(x)|^2.$$

Ввиду сепарбельности M и в силу локальной липшицевости g, правая часть этого неравенства будет представлять собой борелевскую функцию на M^n, и после интегрирования по мере μ^n, применяя (14), мы придем к неравенству

$$\lambda_1 \mathrm{Var}_{\mu^n}(g) \leq \int_{M^n} \sum_{i=1}^{n} |\nabla_{x_i} g(x)|^2\, d\mu(x_1) \cdots d\mu(x_n).$$

Согласно (13), мы получаем неравенство типа Пуанкаре для пространства (M^n, ρ_n, μ^n).

В представляющих интерес случаях свойство (13) выполняется для метрики $\rho_n(x,y) = \sqrt{\sum_{i=1}^{n} \rho(x_i, y_i)^2}$, $x, y \in M^n$. Например, это справедливо для $M = \mathbf{R}^d$ и абсолютно-непрерывных μ, так как по известной теореме Радемахера локально-липшицевые функции почти всюду дифференцируемы.

Литература

1. **Aida S., Masuda T., Shigekawa I.** Logarithmic Sobolev inequalities and exponential integrability //*J. Func. Anal. 1994. V.126. P. 83–101.*

2. **Aida S., Strook D.** Moment estimates derived from Poincaré and logarithmic Sobolev inequalities //*Math. Research Letters. 1994. V.1. P. 75–86.*

3. **Alon N., Milman V.D.** λ_1, isoperimetric inequalities for graphs, and superconcentrators // *J. Comb. Theory. Ser.B. 1985, V.38. P. 73–88.*

4. **Bobkov S.G., Ledoux M.** Poincare's inequalities and Talagrand's concentration phenomenon for the exponential distribution // *Probab. Theory Rel. Fields. 1997. V.107. P. 383–400.*

5. **Боровков А.А., Утев С.А.** Об одном неравенстве и связанном с ним характеризации нормального распределения // *Теория вероятн. и ее применен. 1983. Т.28. С. 209–218.*

6. **Gromov M., Milman V.** A topological application of the isoperimetric inequality //*Amer. J. Math. 1983. V.105. P. 843–854.*

7. **Klaassen C.A.J.** On an inequality of Chernoff //*Ann. Probab. 1985. V.13. P. 966–974.*

8. **Schmuckenschläger M.** Poincaré type inequalities and deviation inequalities and deviation inequalities. // *Preprint (1995), to appear in: J. Func. Anal.*

Summary

Bobkov S. G. Remarks on Gromov–Milman's inequality

We consider optimal estimates for the concentartion function and for deviations of Lipschitz functions on metric probability spaces satisfying Poincaré-type inequalities

Сыктывкарский государственный университет

Поступила 15.09.98