Конспект к лекции 5. (Казань, 6 апреля 2017 г.)

13 Случайное блуждание на экспандерах

Мы уже видели, что спектральные экспандеры обладают свойствами «хорошего перемешивания». Даже один шаг случайного блуждания на спектральном экспандере заметно приближает исходное распределение вероятностей на вершинах к равномерному (см. лемму о перемешивании). В этом параграфе мы изучим свойства многошагового случайного блуждания на спектральном экспандере.

Теорема 13.1 *Пусть граф G является спектральным (n, d, γ)-экспандером и A — некоторое множество вершин графа, состоящее из αn вершин (для некоторого $\alpha > 0$). Тогда все собственные числа индуцированного подграфа*[1] *на вершинах A не превосходят $(\gamma + \alpha(1-\gamma))d$.*

Замечание: Даже если граф G однороден (все вершины имеют степень d), его индуцированный подграф может быть неоднородным. Однако матрица индуцированного подграфа симметрична, а значит, имеет собственный ортонормированный базис. Собственные числа графа оцениваются сверху максимальным значением отношения Рэлея. При этом нет необходимости отдельно выписывать в явном виде матрицу подграфа — поскольку мы рассматриваем *индуцированный* подграф в G, достаточно ограничить множество ненулевых координат распределения $\mathbf{f}$ на вершинах графа. Таким образом, если нас интересует индуцированный подграф, заданный множеством вершин A, мы должны изучить отношение Рэлея

$$\frac{|\mathbf{f}M\mathbf{f}^{\perp}|}{||\mathbf{f}||^2}$$

для всех ненулевых векторов $\mathbf{f}$, сосредоточенных на вершинах выбранного подграфа (координаты f для вершин, не принадлежащих A, должны быть равны нулю).

Лемма 13.1 *Пусть граф G является спектральным (n, d, γ)-экспандером с матрицей M и $\mathbf{f} = (f_1, \ldots, f_n)$ некоторый вектор. Тогда*

$$\mathbf{f}M\mathbf{f}^{\perp} \leqslant \gamma d||f||^2 + \frac{d(1-\gamma)}{n}\left(\sum f_i\right)^2.$$

Доказательство: Обозначим $\mathbf{e}_1, \ldots, \mathbf{e}_n$ ортонормрованный собственный базис для M и $\lambda_1, \ldots, \lambda_n$ соответствующие собственные числа. При этом, как обычно, мы можем полагать первый собственный вектор $\mathbf{e}_1 = \frac{1}{\sqrt{n}}(1, \ldots, 1)$

[1] В индуцированном подграфе G_S множество вершин совпадает с A (некоторым подмножеством вершин исходного графа G), а в качестве рёбер берутся все рёбра графа G, оба конца которых принадлежат A.

и первое собственное число $\lambda_1 = d$). Далее, разложим вектор $\mathbf{f}$ в сумму $\mathbf{f} = \mathbf{f}_\| + \mathbf{f}_\perp$, где $\mathbf{f}_\|$ параллелен, $\mathbf{f}_\perp$ перпендикулярен базисному вектору $\mathbf{e}_1$.

Заметим, что

$$\mathbf{f}_\| = (\mathbf{f}, \mathbf{e}_1) \cdot \mathbf{e}_1 = \frac{\sum f_i}{\sqrt{n}}(1, \dots, 1).$$

Далее,

$$\mathbf{f}M\mathbf{f}^\perp = \mathbf{f}_\| M\mathbf{f}_\|^\perp + \mathbf{f}_\perp M\mathbf{f}_\perp^\perp.$$

Поскольку $\mathbf{f}_\perp$ лежит в подпространстве, натянутом на собственные векторы, собственные числа для которых не превосходят γd, мы получаем

$$|\mathbf{f}M\mathbf{f}^\perp| \leqslant d||\mathbf{f}_\||^2 + (\gamma d)||\mathbf{f}_\perp||^2.$$

По теореме Пифагора мы имеем $||\mathbf{f}||^2 = ||\mathbf{f}_\||^2 + ||\mathbf{f}_\perp||^2$. Следовательно,

$$|\mathbf{f}M\mathbf{f}^\perp| \leqslant d||\mathbf{f}_\||^2 + (\gamma d)(||\mathbf{f}||^2 - ||\mathbf{f}_\||^2).$$

Остаётся заметить, что $||\mathbf{f}_\||^2 = \frac{(\sum f_i)^2}{n}$, и лемма доказана.

Доказательство теоремы 13.1: Рассмотрим вектор $\mathbf{f}$, у которого все координаты кроме тех, что соответствуют вершинам множества A, равны нулю. Согласно лемме 13.1 имеем

$$\mathbf{f}M\mathbf{f}^\perp \leqslant \gamma d||f||^2 + \frac{d(1-\gamma)}{n}\left(\sum f_i\right)^2.$$

Далее, из неравенства между средним арифметическим и средним квадратичным выкает следующая оценка для суммы $\sum f_i$ (в которой содержится не более αn ненулевых членов):

$$\sum f_i \leqslant \sqrt{[\text{число ненулевых членов в сумме}]} \cdot \sqrt{\sum f_i^2} \leqslant \sqrt{\alpha n}||\mathbf{f}||.$$

Соединяя два последних неравенства, мы получаем

$$\mathbf{f}M\mathbf{f}^\perp \leqslant \gamma d||f||^2 + d\alpha(1-\gamma)||f||^2.$$

Это позволяет нам оценить отношение Рэлея: для всех f с нулевыми координатами за пределами A мы имеем

$$\frac{|\mathbf{f}M\mathbf{f}^\perp|}{||\mathbf{f}||^2} \leqslant (\gamma + \alpha(1-\gamma))\, d,$$

что и даёт требуемое неравенство для собственных чисел индуцированного подграфа.

Теперь мы готовы доказать несколько утверждений о блуждании на экспандере.

Утверждение 13.1 *Пусть граф G является спектральным (n, d, γ)-экспандером и A — некоторое множество вершин графа, состоящее из αn вершин. Рассмотрим случайное блуждание по графу*

$$x_0 - x_1 - \ldots - x_t,$$

где вершина x_0 выбирается случайно (по равномерному распределению), а затем на каждом шаге $i = 1, \ldots, t$ следующая вершина x_i выбирается случайно (также равномерно) среди всех соседей x_{i-1}. Тогда

$$\mathrm{Prob}[x_i \in A \text{ для всех } i] \leqslant (\alpha + \gamma(1-\alpha))^t.$$

Доказательство: Общее число путей $x_0 - x_1 - \ldots - x_t$, в графе G равно nd^t (имеется n вариантов для выбора первой вершины x_0 и по d вариантов для выбора каждого из t шагов). Нужно подсчитать, сколько из этих путей полностью лежат в A.

Обозначим $\mathbf{f}^{(i)} = (f_1^{(i)}, \ldots, f_n^{(i)})$ такой вектор, где $f_j^{(i)}$ есть число путей длины i, проходящих только по вершинам A и заканчивающимся в j-ой вершине графа G. В частности, в векторе $\mathbf{f}^{(0)}$ в позициях вершин A стоят единицы, а в позициях вершин вне A стоят нули.

Каждый следующий вектор $\mathbf{f}^{(i+1)}$ получается из $\mathbf{f}^{(i)}$ умножением на матрицу индуцированного подграфа. Поскольку собственные числа матрицы этого графа не превосходят $(\alpha + \gamma - \alpha\gamma)d$, мы получаем

$$||\mathbf{f}^{(t)}|| \leqslant ((\alpha + \gamma - \alpha\gamma)d)^t ||\mathbf{f}^{(0)}|| = ((\alpha + \gamma - \alpha\gamma)d)^t \cdot \sqrt{\alpha n}.$$

Применяем неравенство Коши и получаем, что сумма координат (L_1-норма) вектора $\mathbf{f}^t$ не превосходит

$$\sqrt{n} \cdot ||\mathbf{f}^{(t)}|| \leqslant \alpha(\alpha + \gamma - \alpha\gamma)^t \cdot (d^t n) \leqslant (\alpha + \gamma - \alpha\gamma)^t \cdot (d^t n).$$

Утверждение доказано.

Обобщение 1: Рассмотрим случайное блуждание по экспандеру, состоящее из $k = 2t$ шагов,

$$x_0 - x_1 - \ldots - x_{2t}.$$

Как и раньше, вершина x_0 выбирается случайно (по равномерному распределению), а затем на каждом шаге $i = 1, \ldots, 2t$ следующая вершина x_i выбирается случайно (также равномерно) среди всех соседей x_{i-1}. Будем интересоваться вероятностью того, что все вершины с чётными номерами, т.е., $x_0, x_2, x_4, \ldots, x_{2t}$, попали в множество A. Вероятность этого события оценивается следующим образом:

$$\mathrm{Prob}[x_i \in A \text{ для всех чётных } i] \leqslant (\alpha + \gamma^2(1-\alpha))^t \leqslant (\alpha + \gamma(1-\alpha))^t.$$

В самом деле, нужно применить утверждение 13.1 не к исходному графу G, а к его 2-ой степени (к графу на n вершинах, рёбрами которого являются пути длины 2 в G).

Обобщение 2: Снова рассмотрим случайное блуждание по экспандеру, состоящее из t шагов,

$$x_0 - x_1 - \ldots - x_t.$$

На этот раз оценим вероятностью того, что в множество A попали все вершины с «контролируемыми» номерами $x_{i_1}, x_{i_1+i_2}, x_{i_1+i_2+i_3}, \ldots, x_{i_1+i_2+\ldots+i_r}$. Вероятность этого события не превосходит

$$(\alpha + \gamma^{i_1}(1-\alpha)) \cdot (\alpha + \gamma^{i_1}(1-\alpha)) \cdot \ldots \cdot (\alpha + \gamma^{i_r}(1-\alpha)) \leqslant (\alpha + \gamma(1-\alpha))^r.$$

В самом деле, в рассуждение из замечания 1 легко переносится на случай, когда расстояние между «контролируемыми» номерами шагов x_j варьируются. Таким образом, мы доказали следующий результат:

Утверждение 13.2 *Пусть граф G является спектральным (n, d, γ)-экспандером и A — некоторое множество вершин графа, состоящее из αn вершин. Рассмотрим случайное блуждание по графу*

$$x_0 - x_1 - \ldots - x_t,$$

где вершина x_0 выбирается случайно и равномерно, а затем на каждом шаге $i = 1, \ldots, t$ следующая вершина x_i выбирается случайно равномерно среди всех соседей x_{i-1}.

Пусть $I \subset \{0, \ldots, t\}$ некоторое подмножество номеров шагов. Тогда

$$\mathrm{Prob}[x_i \in A \text{ для всех } i \in I] \leqslant (\alpha + \gamma - \alpha\gamma)^{|I|-1}.$$

Упражнение 13.1 *Для распределения вероятностей $\mathbf{p} = (p_1, \ldots, p_n)$ рассмотрим три варианта меры его «неопределённости»:*

(i) энтропия Шеннона $H(\mathbf{p}) = \sum\limits_{p_i \neq 0} p_i \log \frac{1}{p_i}$,

(ii) энтропия Реньи $H_2(\mathbf{p}) = -\log\left(\sum\limits_{i=1}^{n} p_i^2\right)$,

(iii) min-энтропия $H_{\min}(\mathbf{p}) = \log\left(\min\limits_{p_i>0} \frac{1}{p_i}\right)$.

Докажите, что для любого распределения вероятностей на спектральном (n, d, γ)-экспандере величины энтропий H, H_2 и $H_{\min}$ не убывают на каждом шаге случайного блуждания.

14 Блуждание на экспандере как генератор псевдослучайных битов.

В этом параграфе мы увидим, как превратить случайное блуждание по экспандеру в генератор «псевдослучайных». Мы покажем, как радикально уменьшить вероятности ошибки вероятностного алгоритма и при этом (а) не сильно ухудшить сложность вычислений, и (б) сравнительно «экономно» расходовать случайные биты.

Предположим, что для решения некоторой задачи имеется полиномиальный вероятностный алгоритм, который на любом входе x с вероятностью не менее $1-\delta$ возвращает правильный ответ. Чтобы уменьшить вероятность ошибки алгоритма, можно последовательно выполнить имеющийся алгоритм t раз на независимых значениях датчика случайных битов, а затем из полученных t результатов выбрать наиболее часто случающийся. У нового алгоритма вероятностью ошибки не будет превосходить c^t для некоторого $c<1$. Таким образом, сделав число итераций t достаточно большим, можно сделать вероятность ошибки меньше любого наперёд заданного числа. Можно даже сделать вероятность ошибки экспоненциально убывающей (с ростом длины входа), если взять чисо итераций $t=t(n)$ сравнимым с длиной входа. При этом время работы алгоритма будет оставаться полиномиальным. Очевидным недостатком этого подхода является рост числа используемых случайных битов — их число умножается на t.

Мы покажем, что существует альтернатива простому повторению исходного алгоритма на независимых наборах случайных битов. Данный подход позволит радикально уменьшить вероятность ошибки и при этом незначительно увеличить расход случайные биты. Для этого мы будем генерировать с помощью экспандеров «псевдослучайные» биты. Набор псевдослучайных битов можно будет вырастить из короткого «зерна» — небольшого набора настоящих случайных битов. При этом полученные псевдослучайные биты, как мы увидим, можно использовать для независимого запуска многих копий вероятностного алгоритма, (почти) как если бы они были по-настоящему случайными независимыми.

Далее в этом параграфе мы рассматриваем алгоритмы с двусторонней ошибкой (считаем, что вероятностный алгоритм может выдавать как положительные, так и отрицательные ложные ответы). При этом мы полагаем, что для любого входа вероятность ошибки ограничена некоторым $\delta<1/4$.

Обозначим через k количество случайных битов, оторое требовалось исходному вероятностному алгоритму (при работе со входами длины n). Рассмотрим спектральный $(2^k,d,\gamma)$-экспандер. Определим следующий случайный процесс: выберем случайно (по равномерному распределению) исходную вершину графа x_0, а затем сделаем t шагов случайного блуждания по графу,

$$x_0-x_1-\ldots-x_t,$$

на каждом шаге выбирая случайного соседа x_{i+1} предыдущей вершины x_i. Затем запустим $t+1$ копию старого алгоритма, используя индексы вершин $x_0,x_1,\ldots,x_t$ как наборы случайных битов. Среди полученных ответов выберем самый часто встречающийся и объявим его результатом работы нового алгоритма.

Утверждение 13.2 позволяет оценить вероятность ошибки нового алгоритма. Она не превосходит

$$\sum_{I\subset\{0,\ldots,t\},\ |I|>t/2}(\delta+\gamma-\delta\gamma)^{|I|-1}\leqslant 2^{t+1}(\delta+\gamma-\delta\gamma)^{(t-1)/2},$$

т.е., для $\delta < 1/4$ и достаточно малых γ вероятность ошибки будет экспоненциально убывать с ростом t. При этом случайное блуждание длины t в d-регулярном графе задается $k + t \log d$ случайными битами (случайная первая вершина и t переходов от текущей вершины к случайному соседу). Это значительно меньше, чем $t \cdot k$ битов, нужных для выбора t по-настоящему независимых наборов по k битов.

Упражнение 14.1 *Предположим, что исходный вероятностный алгоритм ошибался с вероятностью не более δ в ответах «да» и никогда не ошибался в ответах «нет»* (*вероятностный алгоритм с* односторонней *ошибкой*). *Докажите, что метод последовательного повторения исходного алгоритма на «псевдо-независимых» наборах случайных битов, соответствующих случайному блужданию на графе*

$$x_0 - x_1 - \ldots - x_t,$$

(*аналогичный описанной выше конструкции для алгоритмов с двусторонней ошибкой*) *уменьшает итоговую вероятность ошибки до $\delta \cdot (\delta + \gamma - \delta\gamma)^t$.*

Упражнение 14.2 *Рассмотрим следующий алгоритм, тестирующий число n на простоту:*

1. *если $n = 2$, то объявляем n простым*
2. *если n чётно и не равно 2, то объявляем n составным*
3. *для нечётных $n > 2$ повторяем следующую процедуру для $i = 1, \ldots, t$*
   - 3.1. *выбираем случайное целое y_i от 2 до $n - 1$;*
   - 3.2. *если НОД$(y_i, n) \neq 1$ хотя бы для одного i, то объявляем n составным*
   - 3.3. *вычисляем $z_i = y_i^{\frac{n-1}{2}} \mod n$*
4. *если все значения z_i принадлежат $\{1, n-1\}$, причём оба значения 1 и $n-1$ встречаются хотя бы по одному разу, то объявляем n простым; иначе объявляем n составным*

(а) Покажите, что данный алгоритм выполняется за время, полиномиально зависящее от величны параметра t и от длины двоичной записи входа n.

(б) Докажите, что вероятность ошибки данного алгоритма не превосходит $O(c^{-t})$ для некоторого $c < 1$.