Лекция 10. Вероятностный метод. Метод второго момента. Локальная лемма Ловаса

Лектор — Нагорный Александр Степанович
anagorny@list.ru

факультет ВМК МГУ имени М.В. Ломоносова

Лекции на сайте http://mk.cs.msu.ru

Метод второго момента. Максимальный размер клики

Зафиксируем вероятность $p = \frac{1}{2}$ появления ребра, и пусть, как и раньше, $\omega(G)$ – максимальный размер клики. Обозначим через

$$f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$$

математическое ожидание числа k – клик. Функция $f(k)$ становится меньше единицы при $k \sim 2\log_2 n$ (Грубо говоря, $f(k)$ ведёт себя как $n^k 2^{-\frac{k^2}{2}}$.)

Теорема 1. Пусть $k = k(n)$ таково, что $k \sim 2\log_2 n$, и $f(k) \to \infty$ при $n \to \infty$. Тогда для почти всех графов $\omega(G) \geq k$.

Метод второго момента. Максимальный размер клики

Доказательство. Для каждого k – множества S пусть A_S – событие вида «S **является кликой**», а X_S – его индикатор. Положим

$$X = \sum_{|S|=k} X_S .$$

Заметим, что $\omega(G) \geq k$ тогда и только тогда, когда $X > 0$. Тогда $\mathbf{E}[X] = f(k) \to \infty$, и мы изучим Δ^* из следствия 4. Зафиксируем S, и пусть $T \sim S$ в том и только в том случае, когда $|T \cap S| = i$, где $2 \leq i \leq k-1$. Значит, для любого фиксированного S имеем

$$\Delta^* = \sum_{T \sim S} \Pr[A_T | A_S] = \sum_{i=2}^{k-1} \binom{k}{i} \binom{n-k}{k-i} 2^{\binom{i}{2} - \binom{k}{2}} .$$

Метод второго момента. Максимальный размер клики

Доказательство (продолжение). И поэтому

$$\frac{\Delta^*}{\mathbf{E}[X]} = \sum_{i=2}^{k-1} g(i), \text{ где } g(i) = \frac{\binom{k}{i}\binom{n-k}{k-i}}{\binom{n}{k}} \cdot 2^{\binom{i}{2}}.$$

Заметим, что $g(i)$ можно представлять себе как вероятность того, что случайно выбранное T пересекает фиксированное S в i точках, умноженную на возрастающую по $\Pr[A_T]$ функцию. Полагая $i = 2$, имеем

$$g(2) = \frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}} \cdot 2 \sim \frac{k^4}{n^2} \leq o(n^{-1}), \qquad n \to \infty.$$

Метод второго момента. Максимальный размер клики

Доказательство (окончание). На другом конце, при $i = k - 1$, имеем

$$g(k-1) = \frac{k(n-k)2^{-(k-1)}}{\binom{n}{k} 2^{-\binom{k}{2}}} \sim \frac{2kn2^{-k}}{\mathbf{E}[X]}, \qquad n \to \infty.$$

Поскольку $k \sim 2\log_2 n$, числитель равен $n^{-1+o(1)}$, $n \to \infty$. С другой стороны, знаменатель стремится к бесконечности «быстрее», чем $\log_2 n$, поэтому $g(k-1) \leq o(n^{-1})$, $n \to \infty$. Более детальные вычисления (которые мы опускаем) показывают, что остальными $g(i)$ и их суммой можно пренебречь, так что следствие 4 применимо. ■

Метод второго момента. Максимальный размер клики

Теорема 1 приводит к сильной концентрации для $\omega(G)$. При $k \sim 2\log_2 n$ имеем

$$\frac{f(k+1)}{f(k)} = \frac{n-k}{k+1} \cdot 2^{-k} = n^{-1+o(1)} = o(1).$$

Пусть $k_0 = k_0(n)$ – то значение, для которого выполняется $f(k_0) \geq 1 > f(k_0+1)$. Для «большинства» n функция $f(k)$ резко изменяет большое значение $f(k_0)$ на малое $f(k_0+1)$. Вероятность того, что G содержит клику размера k_0+1, не превосходит весьма малой величины $f(k_0+1)$. Когда $f(k_0)$ велико, из теоремы 1 следует, что G содержит клику размера k_0 с вероятностью, близкой к 1. Таким образом, с очень большой вероятностью верно $\omega(G) = k_0$.

Метод второго момента. Максимальный размер клики

Для некоторых n одна из величин $f(k_0)$, $f(k_0+1)$ может оказаться средних размеров. В этом случае эти соображения не применимы. Тем не менее можно доказать строгую концентрацию. Она получена независимо в работах **[2]** и **[3]**.

Приведём формулировку указанного результата без доказательства.

Следствие 1 ([2] Боллобаш, Эрдёш, 1976; [3] Матула, 1976). Существует $k = k(n)$, такое, что

$$\Pr[\omega(G) = k \text{ или } k+1] \to 1.$$

Метод второго момента. Различные суммы

Определение 1. Будем говорить, что **множество** $\{x_1, \dots, x_k\}$ натуральных чисел **имеет различные суммы**, если все суммы

$$\sum_{i \in S} x_i, \qquad \text{где } S \subseteq \{1, \dots, k\},$$

попарно различны. Через $f(n)$ обозначим максимальное k, для которого существует множество

$$\{x_1, \dots, x_k\} \subseteq \{1, \dots, n\}$$

с различными суммами.

Пример 1. Множество $\{2^i : 0 \le i \le \log_2 n\}$ является множеством с различными суммами. Этот пример показывает, что

$$f(n) \ge 1 + \lfloor \log_2 n \rfloor.$$

Метод второго момента. Различные суммы

Замечание 1. Пол Эрдёш обещал $300 за доказательство или опровержение того, что найдётся константа C такая, что

$$f(n) \leq \log_2 n + C.$$

Простую верхнюю оценку для $f(n)$ можно получить из мощностных соображений. Всего имеем $2^{f(n)}$ сумм, причём все они должны быть различны и меньше nk, а значит,

$$2^{f(n)} < nk = nf(n).$$

Отсюда следует (**докажите это!**), что

$$f(n) < \log_2 n + \log_2 \log_2 n + O(1), \qquad n \to \infty.$$

Метод второго момента даёт небольшое улучшение.

Метод второго момента. Различные суммы

Теорема 2. Справедливо соотношение

$$f(n) \le \log_2 n + \frac{1}{2}\log_2 \log_2 n + O(1), \qquad n \to \infty.$$

Доказательство. Зафиксируем множество $\{x_1, \dots, x_k\} \subseteq \{1, \dots, n\}$ с различными суммами. Пусть $\varepsilon_1, \dots, \varepsilon_k$ выбираются независимо и для каждого i

$$\Pr[\varepsilon_i = 1] = \Pr[\varepsilon_i = 0] = \frac{1}{2}.$$

Положим

$$X = \varepsilon_1 x_1 + \dots + \varepsilon_k x_k.$$

(Мы представляем X в виде случайной суммы.) Положим

$$\mu = \mathbf{E}[X] = \frac{x_1 + \dots + x_k}{2} \quad \text{и} \quad \sigma^2 = \mathrm{Var}[X].$$

Метод второго момента. Различные суммы

Доказательство (продолжение). Заметим, что

$$\sigma^2 = \frac{x_1^2 + \ \ldots + x_k^2}{4} \le \frac{n^2 k}{4},$$

откуда следует, что $\sigma \le n\sqrt{k}/2$. По неравенству Чебышёва для любого $\lambda > 1$ имеем

$$\Pr\left[|X - \mu| \ge \lambda\, n\sqrt{k}/2\right] \le \lambda^{-2}.$$

Тогда для вероятности противоположного события получаем неравенство

$$1 - \frac{1}{\lambda^2} \le \Pr\left[|X - \mu| < \lambda\, n\sqrt{k}/2\right].$$

Но X принимает каждое конкретное значение с вероятностью ноль либо 2^{-k}, поскольку каждое значение представляется в виде суммы не более чем одним способом (все суммы разные!).

Метод второго момента. Различные суммы

Доказательство (окончание). Отсюда

$$\Pr\left[|X-\mu| < \lambda\, n\sqrt{k}/2\right] \le 2^{-k}(\lambda n\sqrt{k}+1)$$

и, по транзитивности, получим

$$n \ge \frac{2^k(1-\lambda^{-2})-1}{\sqrt{k}\lambda}.$$

Хотя $\lambda = \sqrt{3}$ максимизирует правую часть последнего неравенства, легко видеть (**проверьте это!**), что при любом выборе $\lambda > 1$ из этого неравенства вытекает нужная верхняя оценка для k, то есть, для функции $f(n)$. ■

Локальная лемма. Введение

В типичном вероятностном доказательстве комбинаторного результата обычно требуется показать, что вероятность некоторого события положительна. Однако многие из таких доказательств в действительности дают больше и показывают, что вероятность рассматриваемого события не только положительна, но и достаточно большая.

Фактически, большинство вероятностных доказательств имеют дело с событиями, которые происходят с большой вероятностью (т. е. вероятность стремится к 1, когда размерность задачи возрастает).

Локальная лемма. Введение

Пример 2. Вспомним приведённое в лекции 2 доказательство того, что для каждого $k \geq 1$ существуют турниры, в которых для любого множества из k игроков существует игрок, побеждающий каждого из них. Доказательство, на самом деле, показывает, что для любого фиксированного k, при достаточно большом числе игроков n почти все турниры с n игроками обладают этим свойством, т. е. вероятность того, что случайный турнир с n игроками обладает требуемым свойством, стремится к 1 при n, стремящемся к бесконечности.

Локальная лемма. Введение

С другой стороны, существует тривиальный случай, в котором можно показать, что определённое событие происходит с положительной, хотя и очень малой, вероятностью.

Пример 3. Пусть мы имеем n взаимно независимых событий, и каждое из них происходит с вероятностью, не меньшей $p > 0$. Тогда вероятность того, что все события произойдут одновременно, равна по меньшей мере p^n, т. е. положительна, хотя может быть экспоненциально малой по n.

Локальная лемма. Введение

Естественно ожидать, что случай взаимной независимости может быть обобщён на слабые зависимости и дать более общий способ доказательства того, что определённые события происходят с положительной, хотя и малой вероятностью. Такое обобщение возможно, и это утверждается в следующей лемме, известной как локальная лемма Ловаса.

Эта простая лемма, впервые доказанная в 1975 г. в работе Эрдёша и Ловаса **[4]**, является чрезвычайно мощным инструментом, поскольку даёт возможность работать с редкими событиями.

Локальная лемма Ловаса

Определение 2. Пусть $A_1, A_2, \dots, A_n$ – события в произвольном вероятностном пространстве. Ориентированный граф $D = (V, E)$ на множестве вершин $V = \{1,2, \dots, n\}$ называется **орграфом зависимости** для событий $A_1, A_2, \dots, A_n$, если для всякого i, $1 \le i \le n$, каждое событие A_i взаимно независимо с событиями $\{A_j : (i, j) \notin E\}$.

Лемма 1 ([4] локальная лемма: общий случай) Предположим, что $D = (V, E)$ является орграфом зависимости для событий $A_1, A_2, \dots, A_n$, и пусть существуют действительные числа $x_1, \dots, x_n$, такие, что $0 \le x_i < 1$ и $\Pr[A_i] \le x_i \prod_{(i,j)\in E}(1 - x_j)$ для всех $1 \le i \le n$. Тогда $\Pr\left[\bigwedge_{i=1}^{n} \overline{A_i}\right] \ge \prod_{i=1}^{n}(1 - x_i)$. В частности, с положительной вероятностью ни одно из событий A_i не происходит.

Локальная лемма Ловаса

Доказательство. Сначала докажем индукцией по s что для любого набора чисел $S \subset \{1,2,\dots,n\}$, $|S| = s < n$ и для любого $i \notin S$ справедливо неравенство

$$\Pr\left[A_i \middle| \bigwedge_{j\in S} \overline{A_j}\right] \le x_i. \qquad (1)$$

Это, конечно, справедливо для $s = 0$.

Пусть неравенство (1) выполняется для всех чисел $s' < s$. Докажем его для s. Фиксируем $S \subset \{1,2,\dots,n\}$, $|S| = s$ и $i \notin S$. Положим $S_1 = \{j \in S: (i,j) \in E\}$, $S_2 = S \backslash S_1$. Тогда

$$\Pr\left[A_i \middle| \bigwedge_{j\in S} \overline{A_j}\right] = \frac{\Pr\left[A_i \wedge (\wedge_{j\in S_1} \overline{A_j}) \middle| \wedge_{l\in S_2} \overline{A_l}\right]}{\Pr\left[(\wedge_{j\in S_1} \overline{A_j}) \middle| \wedge_{l\in S_2} \overline{A_l}\right]}. \qquad (2)$$

Чтобы оценить числитель дроби, стоящей в правой части,

Локальная лемма Ловаса

Доказательство (продолжение). заметим, что так как каждое событие A_i взаимно независимо с событиями $\{A_l : l \in S_2\}$, то

$$\Pr\left[A_i \wedge \left(\bigwedge_{j \in S_1} \overline{A_j}\right) \middle| \bigwedge_{l \in S_2} \overline{A_l}\right] \leq \Pr\left[A_i \middle| \bigwedge_{l \in S_2} \overline{A_l}\right] =$$

$$= \Pr[A_i] \leq x_i \prod_{(i,j) \in E} (1 - x_j). \qquad (3)$$

С другой стороны, знаменатель дроби, стоящей в правой части (2), может быть оценен с использованием предположения индукции. В самом деле, предположим $S_1 = \{j_1, j_2, \dots, j_r\}$. Если $r = 0$, то знаменатель равен 1, и неравенство (1) выполняется. В противном случае (при $r \geq 1$) имеем

Локальная лемма Ловаса

Доказательство (продолжение).

$$\Pr\left[\overline{A}_{j_1} \wedge \overline{A}_{j_2} \wedge \cdots \wedge \overline{A}_{j_r} \mid \bigwedge_{l \in S_2} \overline{A_l}\right] =$$

$$= \left(1 - \Pr\left[A_{j_1} \mid \bigwedge_{l \in S_2} \overline{A_l}\right]\right) \cdot \left(1 - \Pr\left[A_{j_2} \mid \overline{A}_{j_1} \wedge \bigwedge_{l \in S_2} \overline{A_l}\right]\right) \times \cdots$$

$$\ldots \times \left(1 - \Pr\left[A_{j_r} \mid \overline{A}_{j_1} \wedge \cdots \wedge \overline{A}_{j_{r-1}} \wedge \bigwedge_{l \in S_2} \overline{A_l}\right]\right) \geq$$

$$\geq (1 - x_{j_1})(1 - x_{j_2}) \ldots (1 - x_{j_r}) \geq \prod_{(i,j) \in E} (1 - x_j). \qquad (4)$$

Локальная лемма Ловаса

Доказательство (окончание).
Подставляя соотношения (3) и (4) в равенство (2), заключаем, что $\Pr\left[A_i \middle| \wedge_{j\in S} \overline{A_j}\right] \le x_i$, что завершает индуктивное доказательство.

Теперь легко получить утверждение леммы 1, поскольку

$$\Pr\left[\bigwedge_{i=1}^{n} \overline{A}_i\right] = (1 - \Pr[A_1]) \cdot \left(1 - \Pr\left[A_2 \middle| \overline{A}_1\right]\right) \times \ \ldots$$

$$\ldots \times \left(1 - \Pr\left[A_n \middle| \wedge_{i=1}^{n} \overline{A}_i\right]\right) \ge \prod_{i=1}^{n} (1 - x_i)\,.$$

Лемма доказана. ■

Локальная лемма Ловаса

Следствие 2 (локальная лемма: симметричный случай). Пусть $A_1, A_2, \dots, A_n$ – события в произвольном вероятностном пространстве. Пусть каждое событие A_i взаимно независимо со всеми событиями, за исключением тех не более чем d событий A_j, для которых $\Pr[A_i|A_j] \le p$ при всех $1 \le i \le n$. Если

$$ep(d+1) \le 1, \tag{6}$$

то $\Pr\left[\bigwedge_{i=1}^{n} \overline{A}_i\right] > 0$.

Доказательство. Если $d = 0$, то утверждение тривиально. В противном случае рассмотрим орграф зависимости $D = (V, E)$ для событий $A_1, A_2, \dots, A_n$, в котором по условию для каждого j выполнено $|\{j: (i,j) \in E\}| \le d$. Результат теперь вытекает из леммы 1, если взять $x_i = 1/(d+1)$ (тогда $x_i < 1$ для всех i) и использовать тот факт, что $\left(1 - \frac{1}{d+1}\right)^d > 1/e$ при $d \ge 1$. ■

Локальная лемма Ловаса

Замечание 2. Как показал Ширер в 1985 г., константа «e» является наилучшей из возможных в неравенстве (6).

Замечание 3. Доказательство леммы 1 показывает, что утверждение остаётся справедливым, если заменить два предположения о том, что каждое событие A_i является взаимно независимым с событиями $\{A_j : (i,j) \notin E\}$ и что $\Pr[A_i] \leq x_i \prod_{(i,j)\in E}(1 - x_j)$ более слабым предположением о том, что для каждого i и любого $S_i \subset \{1, \dots, n\} \backslash \{j : (i,j) \in E\}$ выполнено $\Pr\left[A_i \mid \bigwedge_{j\in S} \overline{A_j}\right] \leq x_i \prod_{(i,j)\in E}(1 - x_j)$. Это оказывается полезным в некоторых приложениях.

Локальная лемма Ловаса

На следующей лекции мы рассмотрим различные применения локальной леммы Ловаса для получения комбинаторных результатов. Доказательства этих результатов, не использующие локальную лемму, не известны.

Литература к лекции

1. Алон Н., Спенсер Дж. Вероятностный метод. М.: БИНОМ. Лаборатория знаний, 2007, С. 70-72, 78-79, 83-86.
2. Bollobás B. and Erdós P. (1976) Cliques in random graphs. // *Math. Proc. Comb. Phil. Soc.* **80:** P. 419-427.
3. Matula D. W. (1976) The largest clique size in random graph. // *Technical Report. Southern Methodist University, Dallas.*
4. Erdós P. and Lovász L. (1975) Problems and results on 3-chromatic hypergraphs and some related questions, *in Infinite and Finite Sets* (A. Hajnal et al., eds.), Noth-Holland, Amsterdam, P. 609-628.

СПАСИБО ЗА ВНИМАНИЕ!