Лекция 8. Вероятностный метод. Малые вариации. Непрерывное время. Метод второго момента

Лектор — Нагорный Александр Степанович
anagorny@list.ru

факультет ВМК МГУ имени М.В. Ломоносова

Лекции на сайте http://mk.cs.msu.ru

Непрерывное время. Идея (вариант метода малых вариаций)

Дискретные случайные процессы могут иногда изучаться путём помещения их в рамки **непрерывного времени**. Это позволяет применить мощные методы математического анализа (например, интегрирование). Такой подход оказывается наиболее эффективным, когда речь идёт о случайных упорядочениях.

В ходе лекции приводятся пример, где эта идея работает.

Непрерывное время. Свойство B

Свойство *B*. Модифицируем доказательство оценки $m(n) = \Omega\left(2^n (n/\ln n)^{1/2}\right),\ n \to \infty$ из предыдущей лекции. Припишем каждой вершине $v \in V$ «момент рождения» x_v. Тогда x_v – это независимые действительные переменные, каждая из которых является случайной величиной, равномерно распределённой на отрезке $[0,1]$. При этом порядок на множестве V задаётся упорядочением (относительно $\leq$) величин x_v.

Теорема 1.

$$\Pr[B_{ef}] \leq \sum_{l=0}^{n-1} \binom{n-1}{l} 2^{1-2n} \int_0^1 x^l p^{l+1} (1-xp)^{n-1} dx.$$

Непрерывное время. Свойство B

$$\Pr[B_{ef}] \leq \sum_{l=0}^{n-1} \binom{n-1}{l} 2^{1-2n} \int_0^1 x^l p^{l+1} (1-xp)^{n-1} dx.$$

Доказательство. Для $T \subseteq e \setminus \{v\}$ обозначим через B_{efT} следующее событие: «выполнено B_{ef}, и при первой раскраске ребро e содержит в точности $T \cup \{v\}$ синих вершин». Имеется $\binom{n-1}{l}$ способов выбора l – множества T при l, меняющемся от 0 до $n-1$. Тогда первая раскраска множества $e \cup f$ определена и имеет вероятность появления, равную 2^{1-2n}. Предположим, что v имеет момент рождения $x_v = e$.

Непрерывное время. Свойство B

$$\Pr[B_{ef}] \le \sum_{l=0}^{n-1} \binom{n-1}{l} 2^{1-2n} \int_0^1 x^l p^{l+1} (1-xp)^{n-1} dx.$$

Доказательство (окончание). Для всех $w \in T \cup \{v\}$ при втором подбрасывании монеты должен выпадать герб. Вероятность этого равна p^{l+1}. Все $w \in T$ должны родиться раньше v, т. е. $x_v < x$, что выполняется с вероятностью x^l. Ни для какой вершины $w \in f \setminus \{v\}$, рождающейся прежде v, при бросании монеты не может выпасть герб. Каждая такая w имеет вероятность рождения xp, а вероятность того, что ни одна из w не родится, равна $(1-xp)^{n-1}$. Величина $x_v < x$ равномерно распределена в $[0,1]$. Интегрируем по x. ■

Непрерывное время. Свойство B

Теорема 1.

$$\Pr[B_{ef}] \leq \sum_{l=0}^{n-1} \binom{n-1}{l} 2^{1-2n} \int_0^1 x^l p^{l+1} (1-xp)^{n-1} dx.$$

Следствие 1.

$$\Pr[B_{ef}] \leq 2^{1-2n} p.$$

Упражнение. Докажите следствие 1 самостоятельно.

Определения

Определение 1. Длину кратчайшего цикла в графе G назовём **обхватом графа** G и обозначим через $girth(G)$.

Определение 2. Максимальный размер независимого множества [вершин] в графе G обозначим через $\alpha(G)$.

Определение 3. [Вершинное] хроматическое число графа G обозначим через $\chi(G)$.

Большой обхват и большое хроматическое число

Докажем методом малых вариаций утверждение, которое часто рассматривается как одно из самых красивых применений вероятностного метода.

Теорема 2 ([2] Эрдёш, 1959).

Для любых натуральных k, l существует граф G с обхватом $girth(G) > l$ и хроматическим числом $\chi(G) > k$.

Большой обхват и большое хроматическое число

Доказательство. Зафиксируем $0 < \theta < 1/l$ и положим $G \sim G(n,p)$ с $p = n^{\theta-1}$ (то есть, G — случайный граф на n вершинах, в котором для каждой пары вершин ребро выбирается случайно и независимо с вероятностью p). Пусть X — количество циклов длины, не превосходящей l. Тогда

$$\mathbf{E}[X] = \sum_{i=3}^{l} \frac{(n)_i}{2i} p^i \le \sum_{i=3}^{l} \frac{n^{\theta i}}{2i} = o(n), \qquad n \to \infty,$$

так как $\theta l < 1$. В частности,

$$\Pr[X \ge n/2] = o(1), \qquad n \to \infty.$$

Большой обхват и большое хроматическое число

Доказательство (продолжение).

Положим $x = \left\lceil \frac{3}{p} \ln n \right\rceil$, тогда (**убедитесь в этом сами!**)

$$\Pr[\alpha(G) \geq x] \leq \binom{n}{x} (1-p)^{\binom{x}{2}} < \left[ne^{-p(x-1)/2}\right]^x = o(1), \qquad n \to \infty.$$

Пусть n настолько велико, что оба события имеют вероятность, меньшую, чем $1/2$. Тогда существует граф G с менее чем $n/2$ циклами длины, не большей l, и с $\alpha(G) < 3n^{1-\theta} \ln n$.

Большой обхват и большое хроматическое число

Доказательство (окончание). Удалим из G по вершине из каждого цикла длины, не превышающей l. Получим граф G^* по крайней мере с $n/2$ вершинами. При этом G^* имеет обхват больше чем l и $\alpha(G^*) \leq \alpha(G)$. Таким образом,

$$\chi(G^*) \geq \frac{|V(G^*)|}{\alpha(G^*)} \geq \frac{n/2}{3n^{1-\theta}\ln n} = \frac{n^\theta}{6\ln n} \to \infty, \qquad n \to \infty.$$

Чтобы завершить доказательство, возьмём n достаточно большим с тем, чтобы последнее было больше k. ■

Метод второго момента. Основы

После математического ожидания наиболее используемым статистическим понятием для случайной величины X является **дисперсия** $\mathrm{Var}[X]$. По определению,

$$\mathrm{Var}[X] = \mathbf{E}[(X - \mathbf{E}[x])^2].$$

Другими словами, дисперсия является мерой отклонения величины X от своего математического ожидания. Следуя обычной практике, будем обозначать через μ математическое ожидание, а через σ^2 — дисперсию. Арифметический квадратный корень σ из дисперсии называется **стандартным** (или **среднеквадратическим**) **отклонением**. При этих обозначениях нашим основным инструментом будет следующее утверждение.

Метод второго момента. Основы

Теорема 3 (неравенство Чебышёва). Пусть X — случайная величина с конечными математическим ожиданием μ и дисперсией σ^2. Для любого $\lambda > 0$ верно

$$\Pr[|X - \mu| \geq \lambda\sigma] \leq \frac{1}{\lambda^2}.$$

Доказательство. Имеем

$$\sigma^2 = \mathrm{Var}[X] = \mathbf{E}[(X-\mu)^2] \geq \lambda^2\sigma^2 \Pr[|X - \mu| \geq \lambda\sigma]. \quad \blacksquare$$

Использование неравенства Чебышёва принято называть **методом второго момента**.

Метод второго момента. Основы

Замечание 1. Неравенство Чебышёва неулучшаемо (если нет никаких дополнительных ограничений на X).

Например, если X принимает значения $\mu - \lambda\sigma$ и $\mu + \lambda\sigma$ с вероятностью $\lambda^2/2$, а значение μ с вероятностью $1 - \lambda^2$, то неравенство Чебышёва обращается в равенство (**проверьте это!**).

Метод второго момента. Основы

Замечание 2. Если X имеет нормальное распределение со средним μ и стандартным отклонением σ, то при $\lambda \to \infty$ имеем

$$\Pr[|X - \mu| \geq \lambda\sigma] \sim \sqrt{2/\pi}\, e^{-\lambda^2/2}/\lambda,$$

что существенно меньше $\frac{1}{\lambda^2}$.

Метод второго момента. Основы

Предположим, что $X = X_1 + \ \ldots \ + X_m$.

Тогда $\mathrm{Var}[X]$ может быть подсчитана по формуле

$$\mathrm{Var}[X] = \sum_{i=1}^{m} \mathrm{Var}[X_i] + \sum_{i \neq j} \mathrm{Cov}\left[X_i, X_j\right].$$

Вторая сумма берётся по упорядоченным парам $(i, j) \neq (j, i)$, а **ковариация** $\mathrm{Cov}\left[X_i, X_j\right]$ определяется следующим образом:

$$\mathrm{Cov}[Y, Z] = \mathbf{E}[YZ] - \mathbf{E}[Y]\mathbf{E}[Z].$$

Если случайные величины Y и Z независимы, то $\mathrm{Cov}[Y, Z] = 0$. Это часто существенно упрощает вычисление дисперсии.

Метод второго момента. Основы

В дальнейшем мы предполагаем (это всегда будет выполняться в последующем), что X_i являются индикаторами некоторых событий, т. е.

$$X_i = \begin{cases} 1, & \text{если событие } A_i \text{ происходит,} \\ 0, & \text{иначе.} \end{cases}$$

Если $\Pr[X_i = 1] = \Pr[A_i] = p_i$, то

$$\operatorname{Var}[X_i] = p_i(1 - p_i) \le p_i = \mathbf{E}[X_i],$$

и поэтому

$$\operatorname{Var}[X] \le \mathbf{E}[X] + \sum_{i \ne j} \operatorname{Cov}[X_i, X_j].$$

Метод второго момента в теории чисел

Метод второго момента оказывается эффективным в теории чисел. Пусть $\nu(n)$ обозначает количество простых чисел p, делящих n. (Мы не учитываем кратность, хотя это приводит к некоторым отличиям.) Следующий результат утверждает, грубо говоря, что «почти все» n имеют число простых делителей «весьма близкое» к $\ln \ln n$. Это впервые было обнаружено Харди и Рамануджаном в 1920 г. с помощью весьма сложного доказательства. Мы представим здесь удивительно простое доказательство, принадлежащее Турану **[3]** и сыгравшее ключевую роль в развитии вероятностных методов в теории чисел.

Метод второго момента в теории чисел

Теорема 4. Пусть $w(n) \to \infty$ произвольно медленно. Тогда число тех x из множества $\{1, \dots, n\}$, для которых

$$|v(x) - \ln\ln n| > w(n)\sqrt{\ln\ln n},$$

есть $o(n)$, $n \to \infty$.

Доказательство. Пусть x случайно выбирается из множества $\{1, \dots, n\}$. Для простого p положим

$$X_p = \begin{cases} 1, \text{если } p|x, \\ 0 \quad \text{иначе.} \end{cases}$$

Пусть $M = n^{1/10}$ и $X = \sum X_p$, где суммирование ведётся по всем простым $p \le M$. Поскольку никакое $x \le n$ не может иметь более десяти простых делителей, больших M, мы имеем

Метод второго момента в теории чисел

Доказательство (продолжение).

$$v(x) - 10 \le X(x) \le v(x),$$

и, тем самым, границы больших уклонений для X переходят в асимптотически равные им границы для v. (Здесь число 10 может быть заменено любой другой достаточно большой константой.) Имеем

$$\mathbf{E}[X_p] = \frac{\lfloor n/p \rfloor}{n}.$$

Поскольку $y - 1 < \lfloor y \rfloor \le y$, то $\mathbf{E}[X_p] = 1/p + O(1/n),\ n \to \infty$..

В силу линейности математического ожидания

$$\mathbf{E}[X] = \sum_{p \le M} \left(\frac{1}{p} + O\left(\frac{1}{n}\right) \right) = \ln\ln n + O(1).$$

Метод второго момента в теории чисел

Доказательство (продолжение). Здесь мы использовали хорошо известный факт, что $\sum_{p\le x}\frac{1}{p}=\ln\ln n+O(1)$, который может быть доказан комбинированием формулы Стирлинга с суммированием по Абелю.

Теперь мы найдём асимптотическое выражение для

$$\mathrm{Var}[X]=\sum_{p\le M}\mathrm{Var}[X_p]+\sum_{p\ne q}\mathrm{Cov}[X_p,X_q].$$

Поскольку $\mathrm{Var}[X_p]=\frac{1}{p}\left(1-\frac{1}{p}\right)+O\left(\frac{1}{n}\right)$, то

$$\sum_{p\le M}\mathrm{Var}[X_p]=\left(\sum_{p\le M}\frac{1}{p}\right)+O(1)=\ln\ln n+O(1).$$

Метод второго момента в теории чисел

Доказательство (продолжение). Для различных простых p и q равенство $X_p X_q = 1$ выполняется тогда и только тогда, когда $p|x$ и $q|x$, что эквивалентно $pq|x$. Следовательно,

$$\operatorname{Cov}[X_p, X_q] = \mathbf{E}[X_p X_q] - \mathbf{E}[X_p]\mathbf{E}[X_q] = \frac{\lfloor n/pq \rfloor}{n} - \frac{\lfloor n/p \rfloor}{n}\frac{\lfloor n/q \rfloor}{n} \le$$

$$\le \frac{1}{pq} - \left(\frac{1}{p} - \frac{1}{n}\right)\left(\frac{1}{q} - \frac{1}{n}\right) \le \frac{1}{n}\left(\frac{1}{p} + \frac{1}{q}\right).$$

Таким образом,

$$\sum_{p \ne q} \operatorname{Cov}[X_p, X_q] \le \frac{1}{n}\sum_{p \ne q}\left(\frac{1}{p} + \frac{1}{q}\right) \le \frac{2M}{n}\sum \frac{1}{p}.$$

Откуда

$$\sum_{p \ne q} \operatorname{Cov}[X_p, X_q] \le O\left(n^{-9/10} \ln\ln n\right) = o(1).$$

Метод второго момента в теории чисел

Доказательство (окончание). И, аналогично,

$$\sum_{p \neq q} \operatorname{Cov}[X_p, X_q] \geq -o(1).$$

То есть, ковариации не влияют на дисперсию.

$$\operatorname{Var}[X] = \ln\ln n + O(1),$$

и неравенство Чебышёва приводит к соотношению

$$\Pr\left[|X - \ln\ln n| \geq \lambda\sqrt{\ln\ln n}\right] < \frac{1}{\lambda^2} + o(1), \qquad n \to \infty.$$

для любой константы $\lambda > 0$. Так как $|X - v| \leq 10$, то же самое справедливо и для v. ■

Метод второго момента в теории чисел

В классической работе **[4]** показано, что распределение случайной величины v приблизительно нормально с математическим ожиданием и дисперсией, равными $\ln\ln n$. Сформулируем точный результат.

Теорема 5 ([4] Erdós and Kac, 1940, без доказательства).

Пусть λ — произвольное фиксированное число (положительное, отрицательное или ноль). Тогда

$$\lim_{n\to\infty}\frac{1}{n}\left|\{x : 1 \le x \le n, v(x) \ge \ln\ln n + \lambda\sqrt{\ln\ln n}\}\right| = \int_{\lambda}^{\infty}\frac{1}{\sqrt{2\pi}}e^{-t^2/2}dt.$$

И одна цитата напоследок…

Мы процитируем здесь известное изречение Харди:

«Число 317 является простым вовсе не потому, что мы так решили или наш мозг устроен так, а не иначе, но потому, что это **на самом деле так**, потому, что такова математическая реальность.»

Может показаться странным (хотя это ничему и не противоречит), что именно методы теории вероятностей позволяют прийти к лучшему пониманию свойств разложения чисел на простые множители.

Литература к лекции

1. Алон Н., Спенсер Дж. Вероятностный метод. М.: БИНОМ. Лаборатория знаний, 2007, С. 53-54, 59-64.
2. Erdós P. (1959) Graph theory and probability. // *Canad. J. Math.* **11:** P. 34-38.
3. Turan P. (1934) On a theorem of Hardy and Ramanujan. // *J. London Math Soc.* **9**: P. 274-276.
4. Erdós P. and Kac M. (1940) The Gaussian law of errors in the theory of additive number theoretic functions. // *Amer. J. Math.* **62**: P. 738-742.

СПАСИБО ЗА ВНИМАНИЕ!