Математические основы Computer Science
Часть 2: Вероятностный метод. Лекция 7.

Дмитрий Ицыксон

ПОМИ РАН

15 ноября 2009

Содержание лекции

1. Условные вероятности. Независимость.
2. Локальная лемма Ловаса.
3. Раскраски.
4. k-SAT
5. Циклы в ориентированных графах.
6. Оценки Чернова.

Литература

1. Н. Алон, Дж. Спенсер. Вероятностный метод.
2. S. Jukhna. Extremal combinatorics.

Условные вероятности

- Ω, $B \subset \Omega$;
- $\Omega_B = \{A \cap B | A \subseteq \Omega\}$;
- $\Pr[A \mid B] = \frac{\Pr[AB]}{\Pr[B]}$;
- Пусть $A_1, A_2, \ldots, A_n$ — полная система несовместных событий ($A_i A_j = \emptyset$, $\bigcup A_i = \Omega$).
- $C = CA_1 \cup CA_2 \cup \cdots \cup CA_n$.
- (Формула полной вероятности)
  $\Pr[C] = \sum_i \Pr[CA_i] = \sum_i \Pr[C \mid A_i] \Pr[A_i]$

Независимость

- $A, B \subset \Omega$ называются независимыми, если $\Pr[AB] = \Pr[A]\Pr[B]$;
- $\Pr[A \mid B] = B$, $\Pr[B \mid A] = A$;
- Пример. $\Omega = \{00, 01, 10, 11\}$, все исходы равновероятны. A — первый бит равен 0, B — сумма битов четна. $\Pr\{A\} = \frac{1}{2}$, $\Pr\{B\} = \frac{1}{2}$, $\Pr\{AB\} = \frac{1}{4}$;
- События $\{A_i\}_{i \in I}$ называются взаимно независимыми, если для всех $T \subseteq I$ выполняется $\Pr[\bigcap_{i \in T} A_i] = \prod_{i \in T} \Pr[A_i]$.
- A взаимно независимо относительно $\{B_i\}_{i \in I}$, если для любого $T \subseteq I$ выполняется $\Pr[A] = \Pr[A \mid \bigcap_{i \in T} B_i]$
- (Дискретные) случайные величины ξ и η называются независимыми, если для всех $a, b \in \mathbb{R}$ выполняется $\Pr[\xi = a, \eta = b] = \Pr[\xi = a]\Pr[\eta = b]$.

Локальная лемма

- $A_1, A_2, \ldots, A_n$ — события.
- $G(V, E)$ — граф зависимостей:
  - $V = \{1, 2, \ldots, n\}$
  - A_i взаимно независимо со всеми A_j, что $(i, j) \notin E$.
- Лемма. [Эрдеш-Ловас]. G — граф зависимостей событий $A_1, A_2, \ldots, A_n$.
  - Степени всех вершин $\leq d$.
  - $\Pr[A_i] \leq p$
  - $4pd \leq 1$

  Тогда $\Pr[\overline{A_1} \cdot \overline{A_2} \cdots \overline{A_n}] > 0$
- Доказательство.
  - Индукцией по m докажем, что $\Pr[A_{i_1} \mid A_{i_2} \cdots A_{i_m}] \leq 2p$.
  - База $m = 1$ очевидна
  - Пусть 1 смежна с $2, 3, \ldots, k$ и не смежна с $k+1, \ldots, m$
  - $\Pr[A \mid BC] = \frac{\Pr[AB|C]}{\Pr[B|C]}$

Локальная лемма

Лемма. [Эрдеш-Ловас]. G — граф зависимостей событий $A_1, A_2, \ldots, A_n$. Степени всех вершин $\leq d$, $\Pr[A_i] \leq p$, $4pd \leq 1$. Тогда $\Pr[\overline{A_1} \cdot \overline{A_2} \cdots \overline{A_n}] > 0$

Доказательство.

- Индукцией по m докажем, что $\Pr[A_{i_1} \mid \overline{A}_{i_2} \cdots \overline{A}_{i_m}] \leq 2p$.
- База $m = 1$ очевидна
- Пусть 1 смежна с $2, 3, \ldots, k$ и не смежна с $k+1, \ldots, m$
- $\Pr[A \mid BC] = \frac{\Pr[AB|C]}{\Pr[B|C]}$
- $\Pr[A_1 \mid \overline{A}_2 \cdots \overline{A}_n] = \frac{\Pr[A_1 \overline{A}_2 \cdots \overline{A}_k | \overline{A}_{k+1} \cdots \overline{A}_m]}{\Pr[\overline{A}_2 \cdots \overline{A}_k | \overline{A}_{k+1} \cdots \overline{A}_m]}$
- $\Pr[A_1 \overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m] \leq \Pr[A_1 \mid \overline{A}_{k+1} \cdots \overline{A}_m] = \Pr[A_1] \leq p$
- $\Pr[\overline{A}_2 \cdots \overline{A}_k \mid \overline{A}_{k+1} \cdots \overline{A}_m] = 1 - \Pr[\bigcup_{i=2}^{k} A_i \mid \overline{A}_{k+1} \cdots \overline{A}_m] \geq 1 - \sum_{i=2}^{k} \Pr[A_i \mid \overline{A}_{k+1} \cdots \overline{A}_m] \geq 1 - 2p(k-1) \geq 1 - 2pd \geq \frac{1}{2}$
- $\Pr[A_1 \mid \overline{A}_2 \cdots \overline{A}_n] \leq \frac{p}{1/2} = 2p$.
- $\Pr[\overline{A}_1 \cdots \overline{A}_n] = \prod_{i=1}^{n} \Pr[\overline{A}_i \mid \overline{A}_1 \cdots \overline{A}_{i-1}] \geq (1-2p)^n > 0$.

О раскрасках

- $\mathcal{F}$ — семейство k-элементных множеств.
- $|\mathcal{F}| \leq 2^{k-1}$, тогда все элементы можно правильно раскрасить в два цвета ($\forall S \in \mathcal{F}$ в S должны быть элементы обоих цветов).
- Теорема. [Эрдеш-Ловас] Если каждый элемент $\mathcal{F}$ пересекает не больше 2^{k-3} других, то существует правильная раскраска $\mathcal{F}$ в два цвета.
  Доказательство.
  - $\mathcal{F} = \{S_1, \ldots, S_m\}$.
  - Раскрасим каждый элемент с вероятностью $\frac{1}{2}$ случайно.
  - A_i: все вершины A_i покрашены в один цвет
  - $\Pr[A_i] = 2^{1-k}$
  - Надо доказать $\Pr[\overline{A}_1 \cdots \overline{A}_m] > 0$
  - Граф зависимостей: соединяем A_i и A_j, если $S_i \cap S_j \neq \emptyset$.
  - $d = 2^{k-3}, 4pd = d2^{3-k} \leq 1$
  - Утверждение следует из локальной леммы.

k-SAT

Теорема. Формула в k-КНФ (в каждом дизъюнкте все переменные разные), в которой каждая переменная входит не более, чем в $\frac{2^{k-2}}{k}$ дизъюнктов, выполнима.

Доказательство.

- Рассмотрим случайный набор значений переменных.
- A_i — i-й дизъюнкт не выполнен.
- $\Pr[A_i] = 2^{-k}$.
- Граф зависимостей: соединяем A_i и A_j, если i-й дизъюнкт и j-й дизъюнкт имеют общие переменные.
- $d \leq 2^{k-2}, 4pd \leq 1$.
- Утверждение следует из локальной леммы.

Ориентированные циклы

- $G(V, E)$ — ориентированный граф, из каждой вершины исходит $\geq \delta$ ребер и входит $\leq \Delta$ ребер.
- Теорема. Если для некоторого k выполняется неравенство $4\Delta\delta(1 - \frac{1}{k})^{\delta} \leq 1$, то в G есть простой ориентированный цикл, длина которого кратна k.
  Доказательство.
  - Можно считать, что из каждой вершины исходит ровно δ ребер.
  - Пусть $f : V \to \{0, 1, \ldots, k-1\}$ — случайна раскраска. Цвет каждой вершины выбирается случайно и равновероятно.
  - $v \in V$, A_v — не существует вершины $u \in V$, что $(v, u) \in E$ и $f(u) = f(v) + 1 \bmod k$.
  - $\Pr[A_v] = \frac{(k-1)^{\delta}}{k^{\delta}} = (1 - \frac{1}{k})^{\delta}$.
  - Граф зависимостей: A_u и A_v соединяем ребром, если $(N^+(u) \cup \{u\}) \cap N^+(v) \neq \emptyset$
  - $d \leq \delta\Delta, 4pd \leq 1$
  - $\Pr[\bigwedge_{v \in V} \overline{A}_v] > 0$.

Произведение матожиданий

Теорема. $X_1, X_2, \ldots, X_n$ — взаимно независимы. Тогда $\mathrm{E}[X_1 X_2 \ldots X_n] = \mathrm{E}[X_1]\,\mathrm{E}[X_2] \ldots \mathrm{E}[X_n]$.

Доказательство.

$$\mathrm{E}[X_1 X_2 \ldots X_n] = \sum_x x \Pr\{X_1 X_2 \ldots X_n = x\} =$$

$$\sum_{x_1, x_2, \ldots, x_n} x_1 x_2 \cdots x_n \Pr\{X_1 = x_1, X_2 = x_2, \ldots, X_n = x_n\} \overset{\text{(независимость)}}{=}$$

$$\sum_{x_1, x_2, \ldots, x_n} x_1 \Pr\{X_1 = x_1\} x_2 \Pr\{X_2 = x_2\} \ldots x_n \Pr\{X_n = x_n\} =$$

$$\Big(\sum_{x_1} x_1 \Pr\{X_1 = x_1\}\Big)\Big(\sum_{x_2} x_2 \Pr\{X_2 = x_2\}\Big) \ldots \Big(\sum_{x_n} x_n \Pr\{X_n = x_n\}\Big) =$$

$$\prod_{i=1}^{n} \mathrm{E}[X_i]$$

Оценки Чернова-Хоефдинга

- $X_1, X_2, \ldots, X_n$ — взаимно независимые случайные величины, принимающие значения из $\{0, 1\}$;
- $m = \mathrm{E} \sum_{i=1}^{n} X_i$;
- Цель: получить
  $\Pr[|\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} m| \geq \delta] \leq$ что-то маленькое
- $\overline{X}_i = X_i - p_i$
- $\Pr[\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} m \geq \delta] = \Pr[\frac{1}{n} \sum_{i=1}^{n} \overline{X}_i \geq \delta] =$
  $\Pr[e^{t \sum_{i=1}^{n} \overline{X}_i} \geq e^{t\delta n}] \leq \frac{\mathrm{E}[e^{t \sum_{i=1}^{n} \overline{X}_i}]}{e^{t\delta n}}$
- $\mathrm{E}[e^{t \sum_{i=1}^{n} \overline{X}_i}] = \prod_{i=1}^{n} \mathrm{E}[e^{t\overline{X}_i}]$
- $\mathrm{E}[e^{t\overline{X}_i}] = p_i e^{t(1-p_i)} + (1 - p_i) e^{t(0-p_i)} \leq e^{t^2/8}$
- $\Pr[\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} m \geq \delta] \leq \frac{(e^{t^2/8})^n}{e^{t\delta n}} = e^{(t^2/8 - t\delta)n}$
- $t = 4\delta$
- $\Pr[|\frac{1}{n} \sum_{i=1}^{n} X_i - \frac{1}{n} m| \geq \delta] \leq 2e^{-2\delta^2 n}$

Оценки Чернова

- $X_1, X_2, \ldots, X_n$ — одинаково распределенные взаимно независимые случайные величины.
- $\Pr\{X_i = 1\} = p, \Pr\{X_i = 0\} = 1 - p$.
  $\mathrm{E}\, X_i = p, \mathrm{E}\, X = \mathrm{E} \sum_{i=1}^{n} X_i = np$;
- $\Pr\{|\frac{\sum X_i}{n} - p| \geq \varepsilon\} \leq 2e^{-2\varepsilon^2 n}$
- 1000 раз бросали монетку. Оценить вероятность того, что выпало больше 550 орлов?
- $e^{-2\frac{1000}{400}} = e^{-5}$.

Оценки Чернова: применение

- $f : \{0,1\}^n \to \{0,1\}$
- Как вычислить $\mu = \frac{1}{2^n} \sum_{x \in \{0,1\}^n} f(x)$?
- Сгенерировать m случайных строчек: $x_1, x_2, \ldots, x_m \in \{0,1\}^n$
- Посчитать $X = \frac{1}{m} \sum_{i=1}^{m} f(x_m)$

Оценки Чернова в мультипликативной форме

- $X_1, X_2, \ldots, X_n$ — взаимно независимые случайные величины, принимающие значения из $\{0, 1\}$;
- $X = \sum_{i=1}^{n} X_i$, $m = \mathrm{E}\, X$;
- Цель: получить $\Pr\{X \geq (1+\delta)m\} \leq$ что-то маленькое
- $\mathrm{E}\, e^{tX_i} = (1 - p_i) + e^t p_i$;
- $\mathrm{E}\, e^{tX} = \mathrm{E}\, e^{t\sum_i X_i} = \mathrm{E} \prod_i e^{tX_i} = \prod_i \mathrm{E}\, e^{tX_i} = \prod_i (1 + p_i(e^t - 1)) \leq \prod_i e^{p_i(e^t-1)} = e^{\sum_i p_i(e^t-1)} = e^{m(e^t-1)}$
- $\Pr\{X \geq (1+\delta)m\} = \Pr\{e^{tX} \geq e^{(1+\delta)mt}\} \leq \frac{\mathrm{E}\, e^{tX}}{e^{(1+\delta)mt}} \leq \frac{e^{m(e^t-1)}}{e^{(1+\delta)mt}}$
- $t = ln(1+\delta)$
- $\Pr\{X \geq (1+\delta)m\} \leq \left(\frac{e^{\delta}}{(1+\delta)^{(1+\delta)}}\right)^m$
- $\Pr\{X \leq (1-\delta)m\} \leq \left(\frac{e^{-\delta}}{(1-\delta)^{(1-\delta)}}\right)^m$
- Следствие. $\Pr[|X - m| \geq cm] \leq 2e^{-\min\{c^2/4; c/2\}m}$