Лекция 1. Вероятностный метод в комбинаторике. Первое знакомство. Два примера применения метода

Лектор — Нагорный Александр Степанович
anagorny@list.ru

факультет ВМК МГУ имени М.В. Ломоносова

Лекции на сайте http://mk.cs.msu.ru

Развитие комбинаторики как науки

Область комбинаторной математики в последние десятилетия переживает период бурного развития. Уже не являясь более «деловым подвалом» топологии, комбинаторика сегодня составляет самостоятельную часть стандартной математической программы. Исследования по этой тематике ведутся во всех крупных математических центрах.

Связанная со многими областями, комбинаторика повсеместно используется математикой, естественными и гуманитарными науками. Уже построены прочные основы и развит разнообразный методологический аппарат комбинаторной математики. Наряду с математическим анализом возникла и быстро развивается новая дисциплина – комбинаторный анализ.

Разделы комбинаторной науки

- **Перечислительная комбинаторика** (подсчёт количества различных конфигураций, образуемых элементами конечных множеств, на которые могут накладываться определённые ограничения)
- **Структурная комбинаторика** (теория графов и матроидов)
- **Экстремальная комбинаторика,** включающая **теорию Рамсея** (насколько большим (маленьким) может быть множество конечных объектов (числа, графы, векторы, наборы, и т.д.), удовлетворяющее определенным ограничениям?)
- **Вероятностная комбинаторика** (какова вероятность присутствия определённого свойства у заданного множества?)
- **Топологическая комбинаторика** (идеи и методы комбинаторики в топологии, при изучении дерева принятия решений, частично упорядоченных множеств, раскрасок графа и др.)
- **Инфинитарная комбинаторика** (применение идей и методов комбинаторики к бесконечным (в том числе, несчётным) множествам)

Вероятностный (неконструктивный) метод

Мы будем рассматривать некий метод доказательства теорем, который мы будем называть **вероятностным** или **неконструктивным**. Этот метод используется для доказательства того, что некоторый экземпляр исследуемого класса объектов обладает определённым свойством, без непосредственного конструирования такого объекта. Мы показываем, что в некотором вероятностном пространстве вероятность того, что такой объект существует, положительна.

Для большей наглядности продемонстрируем этот метод на двух примерах.

Пример 1. Транзитивные турниры

Определение 1. Турниром $T = (V, E)$ на множестве V из n игроков назовём результат ориентации всех рёбер графа K_n (с множеством вершин V).

Определение 2. Турнир $T = (V, E)$ называется **транзитивным**, если для всех $i, j, k \in V$ из $(i, j), (j, k) \in E$ следует $(i, k) \in E$.

Упражнение 1. Докажите, что турнир $T = (V, E)$ является транзитивным тогда и только тогда, когда существует перестановка σ его игроков, такая, что $(i, j) \in E \Leftrightarrow \sigma(i) < \sigma(j)$.

Пример 1. Транзитивные турниры

Пусть $v(n)$ есть наибольшее целое, для которого всякий турнир из n игроков $\{1, \dots, n\}$ содержит [хотя бы один] транзитивный подтурнир с $v(n)$ игроками.

Теорема 1. ([2] Эрдёш, Мозер, 1964). $v(n) \leq 1 + [2 \log_2 n]$.

Другими словами, утверждается, что найдётся турнир T с n игроками, не обладающий транзитивным подтурниром с $v(n) = 2 + [2 \log_2 n]$ игроками.

Замечание. Р. Стирнз при помощи математической индукции показал, что $v(n) \geq 1 + [\log_2 n]$ (доказательство см. в книге Муна **[3]**).) Эрдёш и Мозер считали, что $v(n) = 1 + [\log_2 n]$ при всех n. Однако Рейд и Паркер эту гипотезу опровергли **[4]**.

Доказательство Теоремы 1

Мы продемонстрируем сейчас целых два доказательства теоремы 1 – мощностное (оценочное) и вероятностное – в целях для сравнения и обсуждения.

Оба они являются неконструктивными, поскольку ни для одного n не дают ответ на вопрос, как именно устроен турнир T с n игроками, не обладающий транзитивным подтурниром с $v(n) = 2 + [2\log_2 n]$ игроками. В обоих случаях доказывается лишь существование такого турнира.

Мощностное доказательство Теоремы 1

Пусть $\mathcal{T} = \mathcal{T}(n)$ обозначает класс всех турниров на $\{1, \dots, n\}$, а $\mathcal{T}'$ есть класс всех турниров на $\{1, \dots, n\}$, которые содержат транзитивный подтурнир с v игроками. Тогда

$$\mathcal{T}' = \bigcup_{A} \bigcup_{\sigma} \mathcal{T}_{A,\sigma},$$

где $A \subseteq \{1, \dots, n\}$, $|A| = v$, σ – перестановка элементов множества A и $\mathcal{T}_{A,\sigma}$ – множество турниров T таких, что индуцированный турнир $T|A$ генерирован σ. Таким образом,

$$|\mathcal{T}_{A,\sigma}| = 2^{\binom{n}{2} - \binom{v}{2}}$$

Мощностное доказательство Теоремы 1 (окончание)

и, согласно элементарным оценкам, при $v = 2 + [2\log_2 n]$ имеем

$$|\mathcal{T}'| \le \sum_{A,\sigma} 2^{\binom{n}{2}-\binom{v}{2}} = \binom{n}{v} v!\, 2^{\binom{n}{2}-\binom{v}{2}} < 2^{\binom{n}{2}} = |\mathcal{T}|.$$

Итак, $\mathcal{T} \setminus \mathcal{T}' \neq \emptyset$; это означает, что существует турнир $T \in \mathcal{T} \setminus \mathcal{T}'$, не содержащий транзитивного подтурнира с v игроками. ■

Концовка доказательства подчёркивает его неконструктивный характер. Мы доказали, что множество не пусто и, следовательно, содержит требуемый элемент T.

О борьбе с неконструктивностью

После того как теорема доказана неконструктивными методами, всё же представляет интерес получение доказательства при помощи конструкций. В данном примере было бы интересно «сконструировать» турнир T_n, не содержащий транзитивного подтурнира с $100\log_2 n$ игроками, хотя из приведённого выше доказательства и вытекает, что «почти все» турниры T_n обладают этим свойством.

Приложение вероятностного метода к «реальному миру» задач может вызвать дополнительные трудности. Пусть, например, какому-то вычислителю дано задание построить на вычислительной машине матрицу турнира со 127 игроками, не содержащего транзитивного подтурнира с 15 игроками.

О борьбе с неконструктивностью

Представим себе, что если матрица будет содержать такой подтурнир, то его фирма (или страна) понесёт большие финансовые потери. Из нашего курса он узнает, что такой турнир точно существует. Но этого недостаточно руководство требует от него конкретную матрицу. Наш вычислитель подсчитывает, что её нахождение методом полного перебора может обойтись в миллионы долларов, а он имеет ограниченный бюджет. Как ему поступить?

Самый лучший выход в данной ситуации – построить T_{127} наобум (случайным образом), уповая на судьбу, однако на случай спешного отъезда следует упаковать чемоданы.

О борьбе с неконструктивностью

Любопытно, что «реальная» ситуация кардинально меняется, если число 15 заменить на 16. То есть, в случае, если необходимо построить турнир на 127 вершинах, который не содержит транзитивного подтурнира с 16 игроками. Тогда

$$\frac{|\mathcal{T}'|}{|\mathcal{T}|} \leq \binom{127}{16} 2^{-\binom{16}{2}} < 0{,}0013.$$

Теперь «случайный» турнир будет обладать требуемым свойством с вероятностью $> 0{,}998$. Такая вероятность успеха вполне достаточна для многих практических ситуаций – наш вычислитель может хорошо рандомизировать и успокоиться.

О формулировках теорем

Обычно мы будем формулировать теоремы в следующей манере:

«Существует турнир (или граф, или матрица, и т. п.) с данным свойством».

Однако, во многих случаях они могут быть сформулированы и так:

««случайный» турнир (или граф, или матрица, и т. п.) обладает данным свойством с вероятностью p», где $p > 0$ (а ещё лучше, $p \to 1$ при $n \to \infty$).

Формулировки второго типа могут иметь бóльшую аппликативность.

Вероятностное доказательство Теоремы 1

Пусть $\mathbf{T} = \mathbf{T}_n$ есть случайная величина, чьи значения суть элементы множества $\mathcal{T}$, для каждого из которых выполнено $\Pr[\mathbf{T} = T] = 2^{-\binom{n}{2}}, \quad T \in \mathcal{T}$. То есть, все элементы множества $\mathcal{T}$ суть равновероятные значения случайной величины $\mathbf{T}$. Можно взглянуть на $\mathbf{T}$ и несколько по-другому. Для каждой пары (i, j) игроков справедливо равенство $\Pr[(i, j) \in \mathbf{T}] = 1/2,$ и эти вероятности, взятые по каждой игре, взаимно независимы. Перечисленные свойства определяют $\mathbf{T}$. Тогда

$$\Pr[\mathbf{T} \text{ содержит транзитивный подтурнир с } v \text{ игроками}] \leq$$

$$\leq \sum_{A} \sum_{\boldsymbol{\sigma}} \Pr[\mathbf{T} | A \text{ генерируется } \boldsymbol{\sigma}] = \binom{n}{v} v!\, 2^{-\binom{v}{2}} < 1.$$

Вероятностное доказательство Теоремы 1

Таким образом, некоторое значение T случайной величины **T** не содержит транзитивного замыкания подтурнира с v игроками. ■

По существу, мы просто переписали первоначальное доказательство, которое служит примером «оценочного рассуждения». Второе доказательство – это «вероятностное рассуждение», оно более ясное и интуитивное. В более сложных результатах вероятностный подход допускает непосредственное применение вероятностных теорем.

Отметим также, что далеко не все результаты, получаемые вероятностным методом, могут быть получены без использования «вероятностей».

Пример 2. Число гамильтоновых путей в турнире

Определение 3. Гамильтоновым путём в графе называется остовный путь, в котором все вершины попарно различны.

Замечание. Гамильтонов путь P в турнире $T = (V, E)$ на n вершинах обычно определяется по-другому – как такая перестановка $P(1), \dots, P(n)$ игроков турнира, для которой при всех $1 \leq i < n$ выполнено условие $\big(P(i), P(i+1)\big) \in E$.

Следующий результат, возможно, был первым применением неконструктивных методов.

Теорема 2. ([5] Селе, 1943). Существует турнир T на n вершинах, содержащий не менее $n!\,2^{1-n}$ гамильтоновых путей.

Пример 2. Число гамильтоновых путей в турнире

Доказательство. Пусть $\mathcal{T} = \mathcal{T}(n)$ определено, как и ранее. Положим $U = \{(T, P)\colon T \in \mathcal{T}(n), P$ – гамильтонов путь в $T\}$, тогда

$$|U| = \sum_T |\{P\colon (T, P) \in U\}| = \tag{1}$$

$$= \sum_P |\{T\colon (T, P) \in U\}|.$$

Зафиксируем произвольный путь P. Тогда $(T, P) \in U$ в том и только в том случае, когда $(n-1)$ определённых игр, а именно, $(n-1)$ игр, составляющих путь P, принадлежат T. Таким образом, $(T, P) \in U$ в точности для $2^{\binom{n}{2}-(n-1)}$ турниров T. А так как имеется $n!$ путей P, то

$$|U| = n!\, 2^{\binom{n}{2}-(n-1)}.$$

Пример 2. Число гамильтоновых путей в турнире

Доказательство (окончание). Правая часть **(1)** имеет в точности $2^{\binom{n}{2}}$ слагаемых. Следовательно, для некоторого T имеем

$$|\{P: (T, P) \in U\}| \geq n!\, 2^{1-n}. \qquad \blacksquare$$

Замечание. Селе показал конструктивными методами, что наибольшее число гамильтоновых путей в турнире на n вершинах ограничено сверху величиной $n!\, 2^{-3n/4}$.

Перепишем и это доказательство как вероятностное.

Вероятностное доказательство Теоремы 2

Доказательство. Пусть снова $\mathbf{T} = \mathbf{T}_n$ есть случайный элемент множества $\mathcal{T}$ с тем же распределением. Определим индикаторную функцию f_P на $\mathcal{T}$ следующим образом:

$$f_P(T) = \begin{cases} 1, \text{ если } P \text{ – гамильтонов путь в } T, \\ 0, \qquad\qquad \text{в противном случае.} \end{cases}$$

Положим $N(T) = \sum_P f_p(T) =$ {число гамильтоновых путей в T}.

Тогда

$$\mathbf{E}[N(T)] = \mathbf{E}\left[\sum_P f_p(T)\right] = \tag{2}$$

$$= \sum_P \mathbf{E}\left[f_p(T)\right] = \tag{3}$$

$$= \sum_P 2^{1-n} = n!\, 2^{1-n},$$

где $\mathbf{E}$ – это математическое ожидание случайной величины.

Вероятностное доказательство Теоремы 2

Доказательство (окончание). Таким образом, для некоторого T имеет место требуемое неравенство $N(T) \geq n!\,2^{1-n}$. ■

Замечание. Решающий переход от **(2)** к **(3)** обусловлен следующей теоремой.

Теорема 3. Если $X_1, \dots, X_n$ суть случайные величины с конечными математическим ожиданиями, то математическое ожидание их суммы существует и равно сумме их математических ожиданий.

Доказательство теоремы 3 можно найти практически в любом руководстве по теории вероятностей (см., например, **[6]**).

Обсуждение формулировки Теоремы 3

Заметим, что зависимость или независимость случайных величин X_i в теореме 3 не оговаривается.

Такое использование математического ожидания называется **методом первых моментов**.

Ещё четыре упражнения

Упражнение 2. Найти значения $v(n)$ при $2 \leq n \leq 7$ (можно использовать вычислительные возможности компьютера).

Упражнение 3*. Доказать, что следующие утверждения для турнира T с n вершинами эквивалентны:

1) турнир T транзитивен;
2) турнир T является бесконтурным;
3) турнир T не содержит контуров длины 3;
4) множество полустепеней исхода (выигрышей) турнира T есть $\{0, 1, \dots, n-1\}$;
5) турнир T содержит ровно один гамильтонов путь.

Ещё четыре упражнения

Определение 4. Гамильтоновым циклом в турнире T на n вершинах называется перестановка P на множестве $\{1, \dots, n\}$, такая, что $(P(i), P(i+1)) \in T$ $(1 \leq i < n)$ и $(P(n), P(1)) \in T$.

Упражнение 4. Доказать, что существует турнир на n вершинах, содержащий по крайней мере $n!\,2^{-n}$ гамильтоновых циклов.

Упражнение 5. Гардеробщице ночного клуба было сдано 30 шляп. Она поссорилась с управляющим и, досаждая ему, выдала все 30 шляп случайным образом. Каково среднее число людей, получивших обратно свои собственные шляпы?

Литература к лекции

1. Эрдёш П., Спенсер Дж. Вероятностные методы в комбинаторике. М.: Мир, 1976, с. 6-12.
2. Erdős P., Moser L. On the representation of directed graphs as unions of orderings // *Magyar Tud. Akad. Mat. Kutató Int. Közl.* — 1964. — T. 9. — p. 125-132.
3. Moon J.W. (1968) Topics on Tournaments // Holt, Reinhart and Winston, New York.
4. Reid K.B., Parker E.T. (1970) Disproof of a conjecture of Erdös and Moser // *Journal of Combinatorial Theory*. — Vol. 9, Issue 3. — p. 225-238.
5. Szele T. (1943) Kombinatorikai vizsgálatok az irányitott teljes gráffal kapcsolatban, *Mat. Fiz. Lapok*, **50**, p. 223-256.
6. Феллер В. Введение в теорию вероятностей и её приложения. В 2 томах. М.: Мир, 1984.

СПАСИБО ЗА ВНИМАНИЕ!