ТЕОРЕМА СПЕНСЕРА О РАССТАНОВКЕ ЗНАКОВ

Сообщение Ф.Л.Назарова

Нашей целью будет доказательство следующего утверждения:

Теорема Спенсера о расстановке знаков.

Для любых n линейных форм $L_j(x_1, ..., x_m) = \sum_{k=1}^{m} a_{j,k} x_k$ от m переменных с коэффициентами $a_{j,k}$, по модулю не превосходящими 1, найдется такой вектор $(\varepsilon_1, ..., \varepsilon_m)$ из $+1$ и -1 ("расстановка знаков"), что

$$|L_j(\varepsilon_1, ..., \varepsilon_m)| \leqslant A\sqrt{n}$$

для всех $j = 1, ..., n$ ($A > 0$ - некоторая абсолютная постоянная).

Упражнение 1.

Рассмотрите n форм $L_j(x_1, ..., x_n) \stackrel{\text{def}}{=} \sum_{k=1}^{n} e^{2\pi i j \frac{k}{n}} x_k$ от $m = n$ переменных. Проверьте, что для любой расстановки знаков $(\varepsilon_1, ..., \varepsilon_n)$ выполняется неравенство $\max\limits_{1 \leqslant j \leqslant n} |L_j(\varepsilon_1, ..., \varepsilon_n)| \geq \sqrt{n}$, то есть результат точен с точностью до значения постоянной A.

Подсказка: сосчитайте сумму $\sum_{j=1}^{n} |L_j(\varepsilon_1, ..., \varepsilon_n)|^2$.

Упражнение 2.

Выведите из теоремы Спенсера, что для любых n подмножеств $X_1, ... X_n$ конечного множества X существует раскраска множества X в два цвета (например, белый и синий), для которой разность числа синих и белых точек в X_j не превосходит $A\sqrt{n}$ при всех $j = 1, ..., n$ (в таком виде задача была поставлена Эрдешем).

Подсказка: пусть $X = \{1, ..., m\}$. Рассмотрите формы

$$L_j(x_1, ..., x_m) = \sum_{k \in X_j} x_k \quad (j = 1, ..., n)\,.$$

Упражнение 3.

Докажите (то есть выведите из теоремы Спенсера), что для любого набора $x_1, ..., x_m \in (-1,1)$ найдется такая расстановка знаков $(\varepsilon_1, ..., \varepsilon_m)$, что

$$|L_j(\varepsilon_1, ..., \varepsilon_m) - L_j(x_1, ..., x_m)| \leqslant 2A\sqrt{n}.$$

Подсказка: рассмотрите сначала случай двоично-рациональных x_k, то есть случай $x_k = 2^{-a}p_k \;:\; p_k \in \mathbb{Z},\ a \in \mathbb{Z}_+$. Индукцией по a проверьте, что существует расстановка знаков $(\varepsilon_1, ..., \varepsilon_m)$, для которой

$$|L_j(\varepsilon_1, ..., \varepsilon_m) - L_j(x_1, ..., x_m)| \leqslant (1 + \tfrac{1}{2} + ... + \tfrac{1}{2^a})\ A\sqrt{n}.$$

База (то есть $x_k \in \{-1,0,1\}$) по существу совпадает с теоремой Спенсера. Чтобы перейти от знаменателя 2^{-a-1} к знаменателю 2^{-a}, надо рассмотреть все нечетные p_k и заметить, что (опять же по теореме Спенсера) к каждому из них можно так добавить $+1$ или -1, что значение каждой из форм изменится не более чем на $2^{-a}\ A\ \sqrt{n}$. Но теперь все числители четные и все дроби можно сократить на 2.

Случай произвольных $x_k \in (-1,1)$ получается очевидным предельным переходом.

Упражнение 4. (полиномы Рудина-Шапиро).

Покажите, что для каждого $n \in \mathbb{N}$ найдется такой полином $p(z) = \sum\limits_{k=1}^{n} \varepsilon_k\ z^k$, что $\varepsilon_k \in \{-1,1\}$ для всех k и $|p(z)| \leqslant \widetilde{A}\sqrt{n}$ при всех $z \in \mathbb{T}$ ($\mathbb{T} = \{\zeta \in \mathbb{C} : |\zeta| = 1\}$ – единичная окружность с центром в 0).

Подсказка: попробуйте воспользоваться "хорошо известным" неравенством С.Н.Бернштейна $\max\limits_{z\in\mathbb{T}} |\frac{d}{dz}\ p(z)| \leqslant n\ \max\limits_{z\in\mathbb{T}} |p(z)|$, чтобы показать, что для любого полинома $p(z)$ степени n максимум модуля на всей единичной окружности не более чем вдвое превышает максимальное значение $p(z)$ в корнях степени $10n$ из 1, а этих корней ровно $10n$.

Замечание. Существует и явная конструкция полиномов Рудина-Шапиро. Положим

$$P_0(z) = Q_0(z) = z;\quad P_{s+1}(z) = z^{2^s}\ P_s(z) + Q_s(z),\quad Q_{s+1}(z) = z^{2^s}\ P_s(z) - Q_s(z).$$

Индукцией по s легко проверить, что P_s и Q_s – полиномы степени 2^s с коэффициентами ±1 и что $|P_s|^2 + |Q_s|^2 \equiv 2^{s+1}$ на $\mathbb{T}$. Так как всякое число n является суммой нескольких различных степеней двойки $n = 2^{m_1} + ... + 2^{m_r}$ $(m_1 > m_2 > ... > m_r \geq 0)$, то мы можем построить полином

$$p(z) = P_{m_1}(z) + z^{2^{m_1}}\ P_{m_2}(z) + z^{2^{m_1}+2^{m_2}}\ P_{m_3}(z) + ... + z^{2^{m_1}+2^{m_2}+...+2^{m_{r-1}}}\ P_{m_r}(z)$$

степени n с коэффициентами ± 1, для которого

$$|p(z)| \leqslant |P_{m_1}(z)| + ... + |P_{m_r}(z)| \leqslant \sqrt{2}\ (2^{\frac{m_1}{2}} + ... + 2^{\frac{m_r}{2}}) \leqslant \frac{2}{\sqrt{2}-1}\ 2^{\frac{m_1}{2}} \leqslant 5\sqrt{n}$$

при всех $z \in \mathbb{T}$.

Подход Спенсера имеет, однако, то преимущество, что позволяет строить полиномы, в которых участвуют не все степени z от z^1 до z^n, а, скажем, только $\frac{n}{2}$ наперед заданных. В этом случае никакая явная формула (или хотя бы полиномиально быстрый алгоритм) для нахождения таких полиномов неизвестен.

Доказательство теоремы Спенсера.

ШАГ 1: (сведение случая $m > n$ к случаю $m \leqslant n$).

Поскольку n форм с комплексными коэффициентами суть не что иное как $2n$ форм с вещественными коэффициентами, то мы в дальнейшем всюду будем предполагать, что все коэффициенты $a_{j,k}$ вещественны.

Лемма

Пусть $L_j(x_1, ..., x_m) = \sum\limits_{k=1}^{m} a_{j,k}x_k$ – какие угодно n форм с вещественными коэффициентами. Тогда существует такой набор чисел $y_1, ..., y_m \in [-1, 1]$, что не более чем n из чисел $y_1, ..., y_m$ отличны от ± 1 и $L_j(y_1, ..., y_m) = 0$ для всех $j = 1, ..., n$.

Доказательство леммы.

Пусть у нас есть какой-нибудь набор $y_1, ..., y_m \in [-1, 1]$, для которого все формы равны нулю – $L_j(y_1, ..., y_m) = 0$ и в котором по крайней мере $n+1$ переменных (скажем, $y_1, ..., y_{n+1}$) отличны от ± 1. Система n уравнений $\sum\limits_{k=1}^{n+1} a_{j,k}\ x_k\ = 0$ с $n+1$ неизвестными $x_1, ..., x_{n+1}$ имеет нетривиальное решение $z_1, ..., z_{n+1}$. Замена $y_1, ..., y_{n+1}$ на $y_1 + tz_1, ..., y_{n+1} + tz_{n+1}$ оставляет все формы L_j равными нулю, и нетрудно видеть, что параметр $t \in \mathbb{R}$ можно выбрать таким образом, чтобы по крайней мере одно из чисел $y_1 + tz_1, ..., y_{n+1} + tz_{n+1}$ стало равным ± 1, а все остальные по-прежнему лежали в промежутке $[-1, 1]$. Начав с $y_1 = ... = y_m = 0$ и повторив эту процедуру не более $m - n$ раз, мы придем к набору, существование которого требовалось доказать.

В силу результата упражнения 3 (вернее его модификации, в которой заключение требуется доказать лишь для $m \leqslant n$, но и теорема Спенсера предполагается верной лишь при $m \leqslant n$) это позволяет свести случай $m \geq n$ к случаю $m \leqslant n$ переменных с ухудшением константы не более чем в два раза.

ШАГ 2: (если не знаешь как делать, то делай как попало!).

Рассмотрим m независимых случайных величин $\varepsilon_1, ..., \varepsilon_m$, принимающих значения $+1$ и -1 с вероятностью $1/2$ каждое.

Читателю, незнакомому с терминологией теории вероятностей, пугаться в этом месте не следует: реально мы просто собираемся рассмотреть все возможные расстановки ± 1. Вероятность $\mathcal{P}\{A\}$ какого-либо "события" A (то есть утверждения о расстановке $(\varepsilon_1, ..., \varepsilon_m)$) это просто доля тех расстановок, для которых это утверждение верно. Например,

$$\mathcal{P}\{\text{ ровно } k \text{ из чисел } \varepsilon_1, ..., \varepsilon_m \text{ равны } +1\} = 2^{-m} C_m^k = 2^{-m} \tfrac{m!}{k!(m-k)!}.$$

Математическое ожидание $\mathcal{E} f(\varepsilon_1, ..., \varepsilon_m)$ функции f, зависящей от расстановки $\varepsilon_1, ..., \varepsilon_m$ – это среднее значение f по всем возможным расстановкам знаков, то есть $\mathcal{E} f = 2^{-m} \sum\limits_{\varepsilon_1, \ldots, \varepsilon_m} f(\varepsilon_1, ..., \varepsilon_m)$. "Независимость" же означает в точности то несложное наблюдение, что если есть два события A и B или две функции f и g, реально зависящие от непересекающихся наборов переменных $\varepsilon_1, ..., \varepsilon_m$ (например, A состоит в том, что произведение всех ε с нечетными номерами равно 1, а B – в том, что сумма всех ε с четными номерами равна 0; $f(\varepsilon_1, ..., \varepsilon_m) = {\varepsilon_1}^2 + \varepsilon_2, g(\varepsilon_1, ..., \varepsilon_m) = {\varepsilon_3}^3 + {\varepsilon_4}^4$), то $\mathcal{P}\{A \& B\} = \mathcal{P}\{A\}\mathcal{P}\{B\}$ и $\mathcal{E}(fg) = \mathcal{E}(f)\mathcal{E}(g)$ (как обычно, $A \& B$ означает, что верно и A, и B).

Отметим простое неравенство, используемое в дальнейшем: $\mathcal{P}\{f > y\} < e^{-y} \mathcal{E} e^f$.

Упражнение 5: Сосчитайте $\mathcal{P}\{A\}, \mathcal{P}\{B\}, \mathcal{E} f, \mathcal{E} g$ в этих примерах.

Пусть $L(x_1, ..., x_m) = \sum\limits_{k=1}^{m} a_k x_k$, причем $|a_k| \leqslant 1$ для всех k. Докажем, что

$$\mathcal{P}\{|L(\varepsilon_1, ..., \varepsilon_m)| > \lambda\sqrt{m}\} < 2e^{-\frac{\lambda^2}{2}} \text{ для всякого } \lambda > 0.$$

Упражнение 6: Докажите неравенство $\frac{e^y + e^{-y}}{2} \leqslant e^{\frac{y^2}{2}}$ для всех $y \in \mathbb{R}$.

Подсказка: разложите обе части неравенства в степенные ряды по y и сравните коэффициенты при равных степенях.

Заметим, что $\mathcal{P}\{|L| > \lambda\sqrt{m}\} = \mathcal{P}\{L > \lambda\sqrt{m}\} + \mathcal{P}\{-L > \lambda\sqrt{m}\}$. Для всякого $t > 0$ имеем

$$\mathcal{P}\{L > \lambda\sqrt{m}\} < e^{-t\lambda\sqrt{m}}\mathcal{E}e^{tL(\varepsilon_1,\dots,\varepsilon_m)} = e^{-t\lambda\sqrt{m}}\prod_{k=1}^{m}\frac{e^{ta_k}+e^{-ta_k}}{2} \leqslant$$

$$e^{-t\lambda\sqrt{m}}\prod_{k=1}^{m}e^{\frac{1}{2}t^2 {a_k}^2} \leqslant e^{-t\lambda\sqrt{m}}e^{m\frac{t^2}{2}}.$$

Подставляя $t = \frac{\lambda}{\sqrt{m}}$, находим, что $\mathcal{P}\{L > \lambda\sqrt{m}\} \leqslant e^{-\frac{\lambda^2}{2}}$. Аналогично, $\mathcal{P}\{-L > \lambda\sqrt{m}\} < e^{-\frac{\lambda^2}{2}}$.

Пусть теперь мы имеем не одну, а n линейных форм L_j. Полагая $\lambda = \sqrt{2\log(2n)}$, заключаем, что $\mathcal{P}\{|L_j(\varepsilon_1,\dots,\varepsilon_m)| > \sqrt{2m\log(2n)}\} < \frac{1}{n}$ для всякого $j = 1,\dots,n$ и, стало быть, $\mathcal{P}\{\max\limits_{1\leqslant j\leqslant n}|L_j| > \sqrt{2m\log(2n)}\} < n\frac{1}{n} = 1$, то есть найдется расстановка знаков $\varepsilon_1,\dots,\varepsilon_m$, для которой $|L_j(\varepsilon_1,\dots,\varepsilon_m)| \leqslant \sqrt{2m\log(2n)}$ при всех $j = 1,\dots,n$.

Упражнение 7: Покажите, что расстановки, для которых $\max|L_j| \leqslant 2\sqrt{m\log n}$ не просто существуют, но составляют подавляющее большинство, а именно:

$$\mathcal{P}\{\max_{1\leqslant j\leqslant n}|L_j| > 2\sqrt{m\log n}\} < \frac{2}{n} \underset{n\to\infty}{\longrightarrow} 0.$$

Подсказка: возьмите $\lambda = 2\sqrt{\log n}$.

Хотя "делание как попало" привело нас к весьма неплохому результату (при любом количестве переменных существует расстановка знаков, для которой все формы не превышают по модулю $2\sqrt{2n\log(2n)}$), мы все же имеем пока лишний множитель $\sqrt{\log n}$ по сравнению с тем, что нам нужно. Следующий (и последний) шаг позволяет нам от него избавиться.

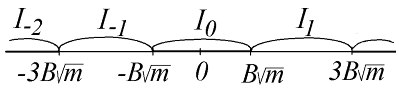
ШАГ 3: Принцип Дирихле.

Хотя было бы наивно надеяться, что при любом выборе n форм от $m \leqslant n$ переменных велика вероятность того, что все эти формы не превышают $A\sqrt{n}$ (это и на самом деле неверно: можно подобрать n форм от n переменных, для которых эта вероятность очень мала при больших n; кто хочет, может подумать как это сделать), мы все же можем надеяться, что с большой вероятностью количество "больших" форм мало.

В самом деле, для всякого $\lambda > 0$ и $j = 1, ..., n$ вероятность $\mathcal{P}\{|L_j| > \lambda\sqrt{m}\} \leqslant 2e^{-\frac{\lambda^2}{2}}$. Стало быть, математическое ожидание количества форм по модулю больших $\lambda\sqrt{m}$ не превышает $2e^{-\frac{\lambda^2}{2}}n$ и, тем самым,

$$\mathcal{P}\{\text{ по крайней мере } e^{-\frac{\lambda^2}{4}}n \text{ форм по модулю больше } \lambda\sqrt{m}\} \leqslant 2e^{-\frac{\lambda^2}{4}}.$$

Разобьем вещественную ось на интервалы $I_k, k \in \mathbb{Z}$, длины $2B\sqrt{m}$ как показано на рисунке (параметр $B > 0$ выберем позже)



Каждой расстановке знаков $(\varepsilon_1, ..., \varepsilon_m)$ поставим в соответствие последовательность из n номеров интервалов, в которые попали значения форм $L_1(\varepsilon_1, ..., \varepsilon_m), ..., L_n(\varepsilon_1, ..., \varepsilon_m)$ соответственно. В силу сделанного ранее замечания вероятность того, номер $k \neq 0$ встречается в этой последовательности больше $e^{-B^2\frac{k^2}{4}}n$ раз, не превышает $2e^{-B^2\frac{k^2}{4}}$. При $B^2 > 32$

$$\sum_{k\neq 0} 2e^{-B^2\frac{k^2}{4}} \leqslant 4\sum_{k=1}^{\infty} e^{-8k^2} \leqslant 4\sum_{k=1}^{\infty} e^{-8k} = \frac{4e^{-8}}{1-e^{-8}} < 8e^{-8} < 8\cdot 2^{-8} = \frac{1}{32}$$

и, стало быть, для по крайней мере $\frac{31}{32}2^m$ расстановок знаков последовательность соответствующих номеров такова, что всякое целое число k встречается в ней не более $ne^{-B^2\frac{k^2}{4}}$ раз.

Сколько всего последовательностей из n целых чисел обладает этим свойством?

Упражнение 8: Докажите неравенство $\log t \leqslant \sqrt{t-1}$ при $t \geq 1$.

Упражнение 9. Пусть $0 \leqslant \alpha \leqslant \frac{1}{2}$. Проверьте, что

$$\sum_{0\leqslant k\leqslant \alpha n} C_n^k \leqslant e^{2\sqrt{\alpha}n}.$$

Подсказка: Воспользуйтесь тождеством $\sum\limits_{1\leqslant k\leqslant n} C_n^k \alpha^k(1-\alpha)^{n-k} = 1$ и заметьте, что $\alpha^k(1-\alpha)^{n-k} \geq \alpha^{\alpha n}(1-\alpha)^{(1-\alpha)n}$ при всех $k \in [0, \alpha n]$. Наконец, используя результат упражнения 8, проверьте, что

$$\log \frac{1}{\alpha^\alpha(1-\alpha)^{1-\alpha}} = \alpha \log \frac{1}{\alpha} + (1-\alpha)\log\frac{1}{1-\alpha} \leqslant \alpha\sqrt{\frac{1-\alpha}{\alpha}} + (1-\alpha)\sqrt{\frac{\alpha}{1-\alpha}} \leqslant 2\sqrt{\alpha}.$$

Упражнение 10: Покажите, что количество интересующих нас последовательностей не превышает

$$\prod_{k\neq 0}\Bigl(\sum_{0\leqslant s\leqslant ne^{-\frac{B^2k^2}{4}}} C_n^s\Bigr) \leqslant \prod_{k\neq 0}\exp\left(2ne^{-\frac{B^2k^2}{8}}\right) = \exp\left(4n\sum_{k=1}^{\infty} e^{-\frac{B^2k^2}{8}}\right).$$

Подсказка: Последовательность полностью определяется тем, на каких местах в ней стоят числа $\pm 1, \pm 2, \ldots$.

Положим $B = \sqrt{8\log\frac{n}{m} + 32} \geq \sqrt{32}$ при $m \leqslant n$. Тогда

$$\sum_{k=1}^{\infty} e^{-\frac{B^2k^2}{8}} \leqslant \sum_{k=1}^{\infty} e^{-k^2(\log\frac{n}{m}+4)} \leqslant \frac{m}{n}\sum_{k=1}^{\infty} e^{-4k} = \frac{e^{-4}}{1-e^{-4}}\frac{m}{n} < \frac{m}{16n},$$

и, стало быть, число возможных последовательностей не превышает $e^{\frac{m}{4}} < \frac{31}{32}2^{\frac{m}{2}}$. Но это означает, что найдутся по крайней мере $2^{\frac{m}{2}}$ расстановок знаков, для которых соответствующие последовательности номеров совпадают.

Упражнение 11: Покажите, что среди любых $2^{\frac{m}{2}}$ (или больше) расстановок m знаков найдутся две, различающиеся по крайней мере в $\frac{m}{64}$ позициях.

Подсказка: количество расстановок, отличающихся от данной в не более чем $\frac{m}{64}$ местах (включая ее саму), равно $\sum\limits_{0\leqslant k\leqslant \frac{m}{64}} C_m^k \leqslant e^{\frac{m}{4}} < 2^{\frac{m}{2}}$.

Пусть теперь $(\varepsilon_1', ..., \varepsilon_m')$ и $(\varepsilon_1'', ..., \varepsilon_m'')$ – две расстановки знаков, различающиеся в $\frac{m}{64}$ или больше позициях, для которых последовательность номеров соответствующих интервалов одна и та же. Тогда

$$\left|L_j\left(\frac{\varepsilon_1'-\varepsilon_1''}{2}, ..., \frac{\varepsilon_m'-\varepsilon_m''}{2}\right)\right| \leqslant B\sqrt{m} = \sqrt{m}\ \sqrt{32+8\log\frac{n}{m}} \text{ при всех } j=1,...,n.$$

Заметим, что числа $\frac{\varepsilon_k'-\varepsilon_k''}{2}$ принимают только значения 0,-1 и $+1$, причем количество нулей m_1 не превышает $\frac{63}{64}m$. Стало быть, случай m переменных свелся к случаю $m_1 \leqslant \frac{63}{64}m$ переменных с потерей в оценке сверху $\sqrt{m}\sqrt{32+8\log\frac{n}{m}}$ или меньше. Продолжив этот процесс редукции, мы получим последовательность неотрицательных целых чисел $m = m_0 > m_1 > ... > m_q > m_{q+1} = 0$, в которой $m_{k+1} \leqslant \frac{63}{64}m_k$ при всех $k = 0, ..., q$, и расстановку знаков, для которой все формы не превосходят $\sum\limits_{k=0}^{q}\sqrt{m_k}\ \sqrt{32+8\log\frac{n}{m_k}}$. Поскольку функция $u(32+8\log\frac{n}{u})$ возрастает при $0 \leqslant u \leqslant n$, то эта сумма не превышает

$$\sum_{k=0}^{\infty}\sqrt{\left(\frac{63}{64}\right)^k m}\ \sqrt{32+8\log\frac{n}{m}+8k\log\frac{64}{63}} \leqslant \qquad (\text{так как } \sqrt{a+b} \leqslant \sqrt{a}+\sqrt{b})$$

$$\leqslant \sqrt{m}\ \sqrt{32+8\log\frac{n}{m}}\sum_{k=0}^{\infty}\left(\frac{63}{64}\right)^{k/2} + \sqrt{m}\ \sum_{k=0}^{\infty}\left(\frac{63}{64}\right)^{k/2}\sqrt{8k\log\frac{64}{63}} \leqslant$$

$$\leqslant A\ \sqrt{m}\ \sqrt{1+\log\frac{n}{m}}$$

при достаточно больших $A > 0$. При $m = n$ получаем $|L_j| \leqslant A\sqrt{n}$, а при $m \leqslant n$ и того меньше.

Замечание. Полученная оценка $A\ \sqrt{m}\ \sqrt{1+\log\frac{n}{m}}$ в действительности точна (с точностью до значения постоянной A) при всех $m \leqslant n \leqslant 2^m$, хотя показать это совсем непросто. При $n \geq 2^m$ тривиальное неравенство $|L_j| \leqslant m$ является одновременно и наилучшим возможным (почему?).

Литература. Joel Spencer. Six standard deviations suffice. Transactions of the Amer. Math. Soc., v.289, no 2, June 1985 (такое странное название возникло потому, что наилучшее известное значение A для случая $m = n$ равно $5,32 < 6$).