Дискретная оптимизация

МФТИ, осень 2015

Александр Дайняк

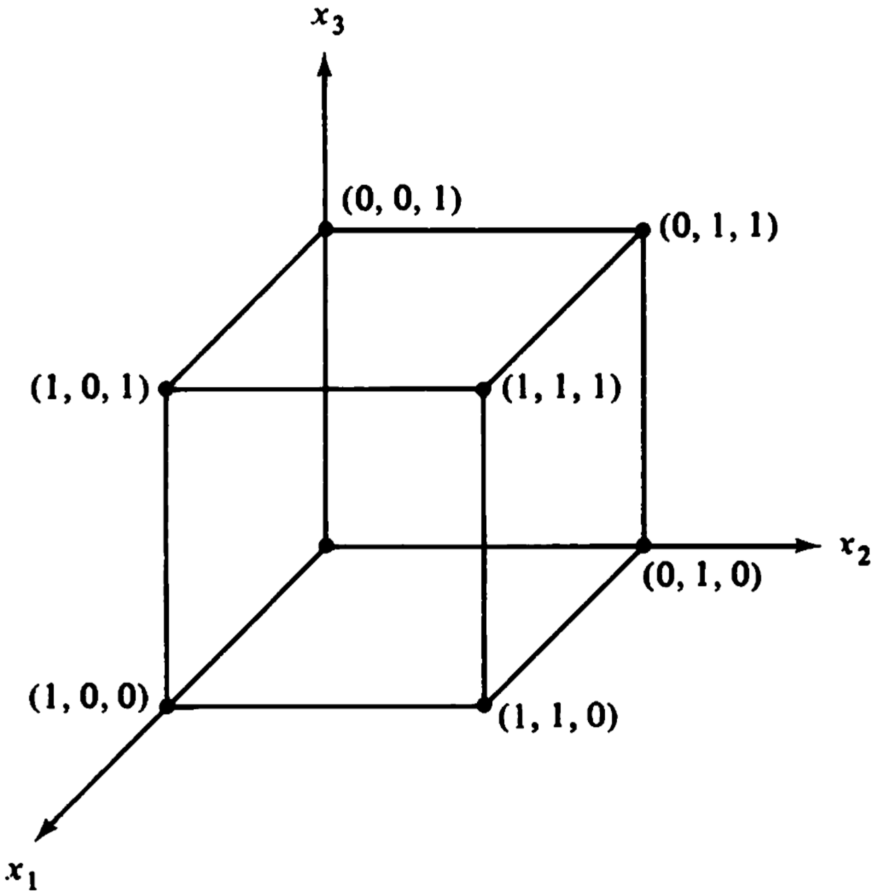
www.dainiak.com

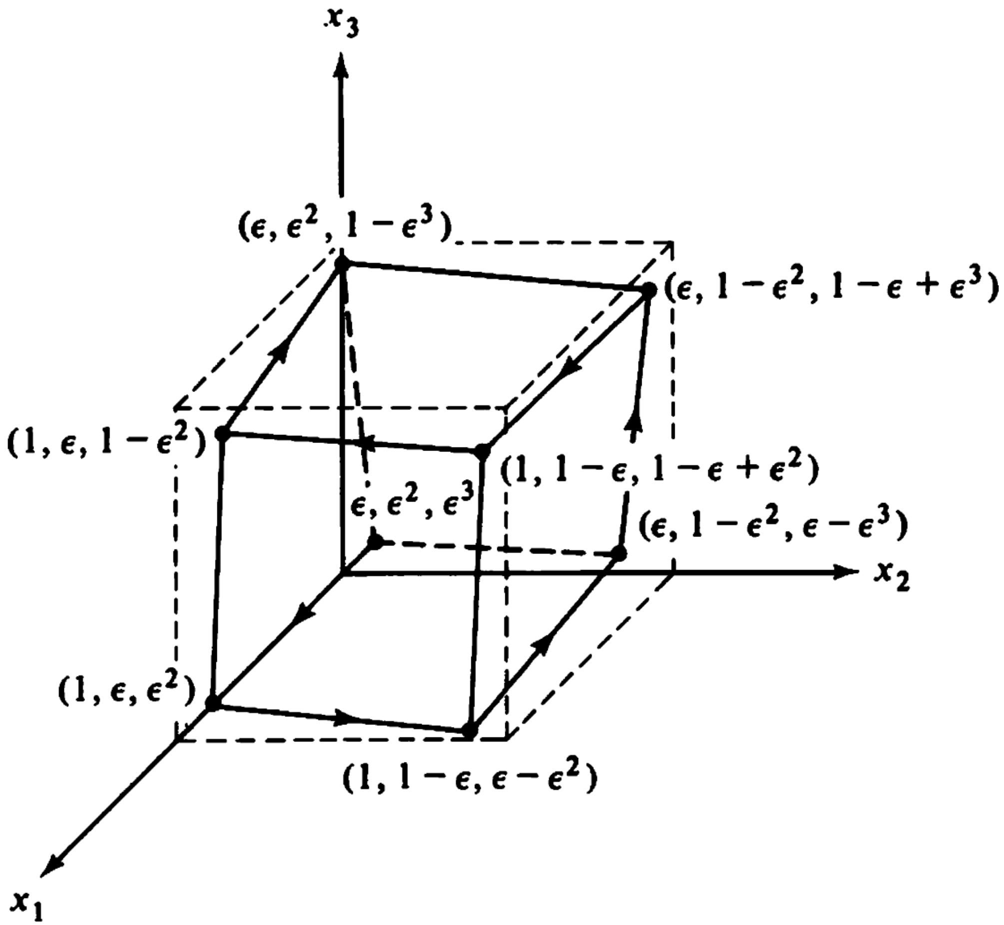
Неполиномиальность симплекс-метода

Теорема (V. Klee & G.J. Minty, 1972).

Для каждого $d > 1$ существует задача ЛП с $2d$ уравнениями, $3d$ переменными и целочисленными коэффициентами, такая, что симплекс-метод может проделать $2^d - 1$ шагов до получения ответа.

Возмущённый куб (perturbation of a cube)





Доказательство теоремы Кли—Минти

Рассмотрим для $\varepsilon < \frac{1}{2}$ задачу:

$$\begin{cases} x_d \to \min \\ x_1 - r_1 = \varepsilon \\ x_1 + s_1 = 1 \\ x_j - \varepsilon x_{j-1} - r_j = 0 \quad (j = 2, \dots, d) \\ x_j + \varepsilon x_{j-1} + s_j = 1 \quad (j = 2, \dots, d) \\ x_j, r_j, s_j \geq 0 \quad (j = 1, \dots, d) \end{cases}$$

Утверждение. Бдр этой задачи — всевозможные множества вида $\{x_1, \dots, x_d\} \cup R \cup S$, такие, что $(r_i \in R \Leftrightarrow s_i \notin S)$.

$$
\begin{cases}
x_d \to \min \\
x_1 - r_1 = \varepsilon \\
x_1 + s_1 = 1 \\
x_j - \varepsilon x_{j-1} - r_j = 0 \quad (j = 2, \dots, d) \\
x_j + \varepsilon x_{j-1} + s_j = 1 \quad (j = 2, \dots, d) \\
x_j, r_j, s_j \geq 0 \qquad (j = 1, \dots, d)
\end{cases}
$$

Пусть $\mathcal{B}$ — произвольное бдр.

Т.к. $x_1 \geq \varepsilon$ и $x_j \geq \varepsilon x_{j-1}$, то $x_j \geq \varepsilon^j > 0$ для каждого $j = 1, \dots, d$.
Поэтому все столбцы, соответствующие $x_1, \dots, x_d$, входят в $\mathcal{B}$.

Допустим, что $r_j = s_j = 0$ для некоторого j.
Не может быть $j = 1$, т.к. тогда было бы $x_1 = \varepsilon = 1$.
Если же $j > 1$, то $\varepsilon x_{j-1} = 1 - \varepsilon x_{j-1}$. Но тогда $\varepsilon x_{j-1} = \frac{1}{2}$, что противоречит неравенствам $\varepsilon < \frac{1}{2}$ и $x_{j-1} \leq 1$.

$$
\begin{cases}
\quad x_d \to \min \\
\quad x_1 - r_1 = \varepsilon \\
\quad x_1 + s_1 = 1 \\
x_j - \varepsilon x_{j-1} - r_j = 0 \quad (j = 2, \dots, d) \\
x_j + \varepsilon x_{j-1} + s_j = 1 \quad (j = 2, \dots, d) \\
\quad x_j, r_j, s_j \geq 0 \qquad (j = 1, \dots, d)
\end{cases}
$$

Бдр, в котором $r_{i_1}, \dots, r_{i_m} \neq 0$, будем обозначать $\boldsymbol{x}^{\{r_{i_1}, \dots, r_{i_m}\}}$.

Утверждение. Пусть все бдр занумерованы $\boldsymbol{x}^{R_1}, \dots, \boldsymbol{x}^{R_{2^d}}$ так, что

$$x_d^{R_1} \leq x_d^{R_2} \leq \dots \leq x_d^{R_{2^d}}.$$

Тогда все неравенства строгие и последовательные бдр в этой цепочке смежны.

Доказательство. Индукция по d.

База: $d = 1$.
Тогда есть два бдр: $(x_1, r_1, s_1) = (\varepsilon, 0, 1 - \varepsilon)$ и $(x_1, r_1, s_1) = (1, 1 - \varepsilon, 0)$.

$$\begin{cases} x_d \to \min \\ x_1 - r_1 = \varepsilon \\ x_1 + s_1 = 1 \\ x_j - \varepsilon x_{j-1} - r_j = 0 \quad (j = 2, \dots, d) \\ x_j + \varepsilon x_{j-1} + s_j = 1 \quad (j = 2, \dots, d) \\ x_j, r_j, s_j \geq 0 \qquad (j = 1, \dots, d) \end{cases}$$

Пусть теперь $d > 1$. Пусть

$$\boldsymbol{x}^{R'_1}, \dots, \boldsymbol{x}^{R'_{2^{d-1}}}$$

— нумерация бдр для $(d-1)$-мерного куба.

Т.к. $R'_i \subseteq \{1,2, \dots, d\}$, то можно рассмотреть соответствующие бдр в d-мерном кубе. Для них будет $x_d^{R'_i} = \varepsilon x_{d-1}^{R'_i}$ и $x_d^{R'_i \cup \{d\}} = 1 - x_d^{R'_i}$, отсюда

$$x_d^{R'_1} \leq x_d^{R'_2} \leq \dots \leq x_d^{R'_{2^{d-1}}} \leq x_d^{R'_{2^{d-1}} \cup \{d\}} \leq x_d^{R'_{2^{d-1}-1} \cup \{d\}} \leq \dots \leq x_d^{R'_1 \cup \{d\}}$$

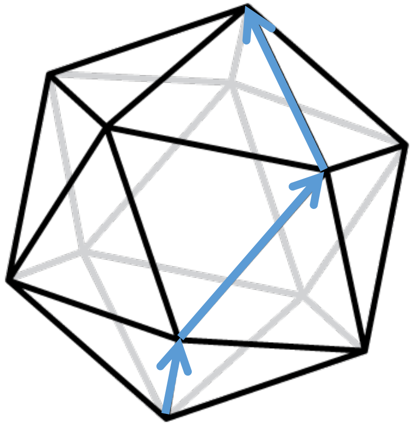
Теоремы a la Klee—Minty

Теорема (R.G. Jeroslow, 1973).
Существует конкретная последовательность задач ЛП (и начальные решения в этих задачах), на которой экспоненциальное число шагов будет делать симплекс-алгоритм, выбирающий из всех соседних бдр то, на котором значение функционала ближе всего к оптимальному.

Диаметр графа многогранника

$\mathrm{H}(n, d)$ — максимальный диаметр графа d-мерного n-гранника.

$\mathrm{H}(n, d)$ — нижняя оценка на число шагов симплекс-алгоритма в худшем случае при неудачном выборе стартовой вершины.



Гипотеза Хирша

$\mathrm{H}(n,d)$ — максимальный диаметр графа d-мерного n-гранника.

Гипотеза (Warren M. Hirsch, 1957).

$$\mathrm{H}(n,d) \leq n - d.$$

Гипотеза Хирша верна в следующих случаях:

- для $(0,1)$-многогранников,
- при $n - d \leq 6$,
- при $(n,d) \in \{(11,4),(12,4)\}$ (причём $\mathrm{H}(12,4) = 7$).

Теорема (Francisco Santos, 2010).
Гипотеза хирша неверна при $(n,d) = (86,43)$.

Известные оценки на $\mathrm{H}(n, d)$

Теорема (D. Barnette, 1974).

$$\mathrm{H}(n, d) \leq 2^d \cdot n.$$

Теорема (G. Kalai & D.J. Kleitman, 1992).

$$\mathrm{H}(n, d) \leq n^{\log_2 d+2}.$$

Теорема (M. J. Todd, 2014).

$$\mathrm{H}(n, d) \leq (n-d)^{\log_2 d+2}.$$

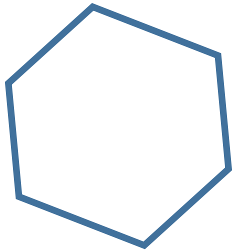
Доказательство теоремы Калаи—Клейтмана

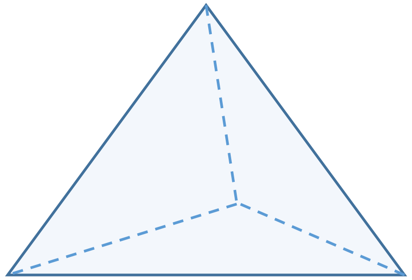
- $\mathrm{H}(n,d)$ — максимальный диаметр графа d-мерного n-гранника.
- $\mathrm{H}(n,d) \le n^{\log_2 d+2}$

Индукция по (n,d).

База:

- $d = 2$ — тогда $\mathrm{H}(n,d) \le n$
- $n = d + 1$ — тогда $\mathrm{H}(n,d) = 1$ (симплекс)





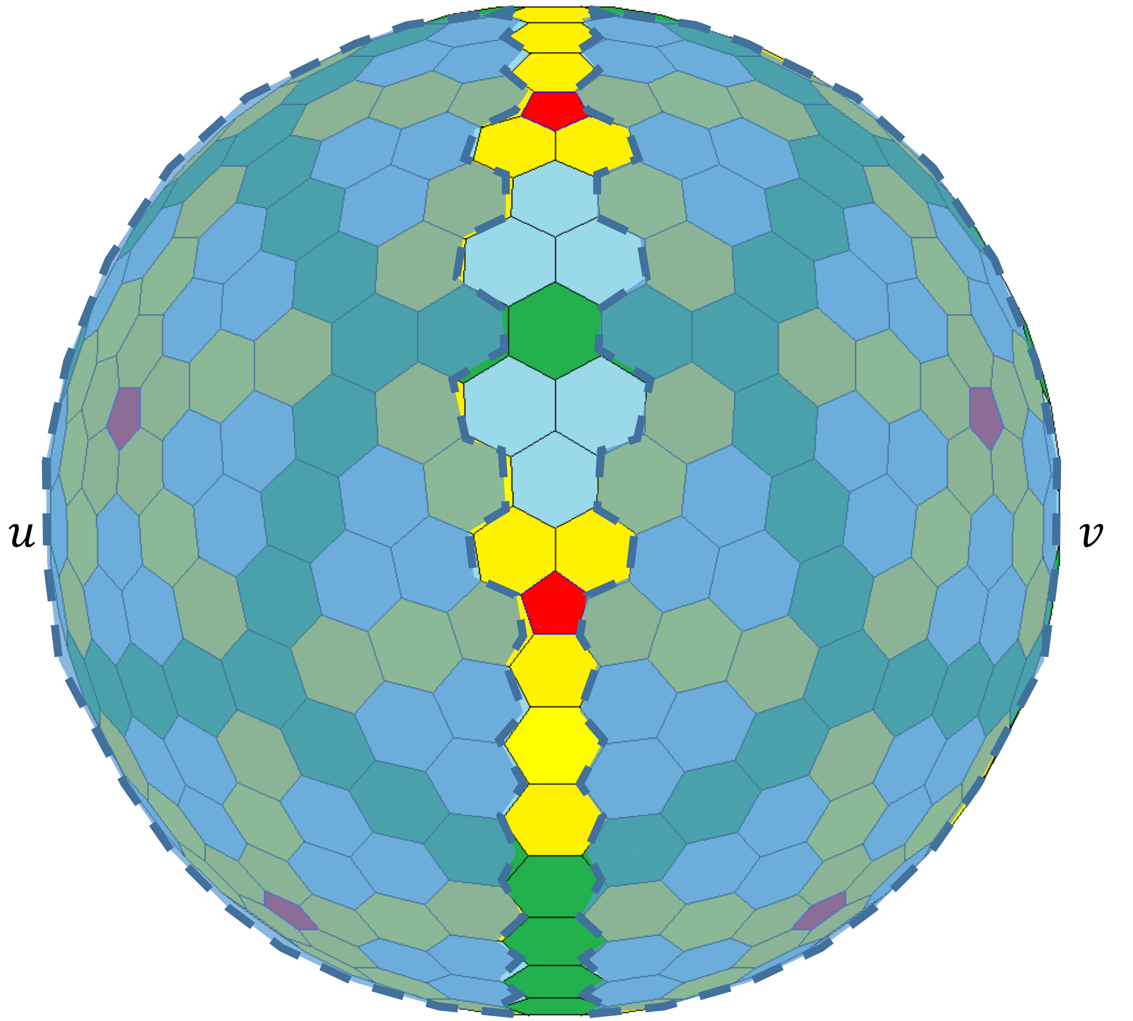
Доказательство теоремы Калаи—Клейтмана

Пусть u, v — пара вершин.

Введём:

k_u — такое максимальное число, что объединение всех цепей длины k_u с началом в u затрагивает не более $n/2$ граней.

k_v — аналогичное число для v.



Доказательство теоремы Калаи—Клейтмана

k_u — такое максимальное число, что объединение всех цепей длины k_u с началом в u затрагивает не более $\lfloor n/2 \rfloor$ граней.

Лемма. $k_u \leq \mathrm{H}(\lfloor n/2 \rfloor, d)$.

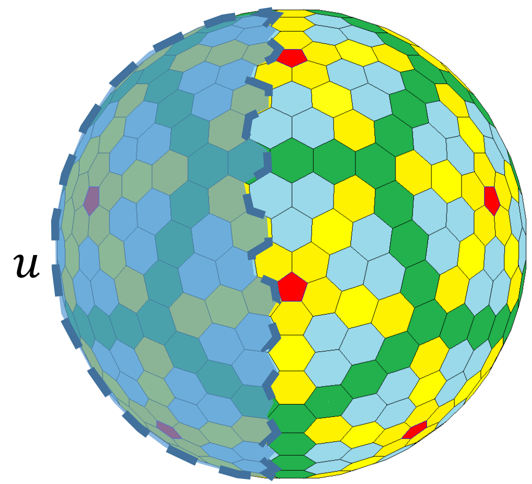
Обоснование.

Пусть P — исходный многогранник.

Пусть w — вершина в P на расстоянии k_u от u.

Пусть Q — многогранник, получаемый из P удалением всех граней, до которых нельзя добраться из u по цепям длины k_u.

Покажем, что расстояние в Q от u до w не меньше k_u.



Доказательство теоремы Калаи—Клейтмана

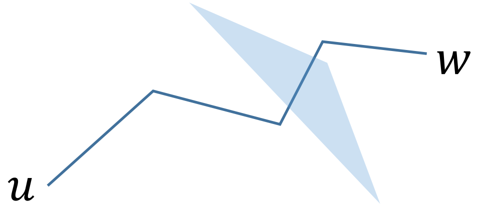
k_u — такое максимальное число, что объединение всех цепей длины k_u с началом в u затрагивает не более $\lfloor n/2 \rfloor$ граней.

Лемма. $k_u \leq \mathrm{H}(\lfloor n/2 \rfloor, d)$.

Обоснование.

Q получен из P удалением «далёких» граней.

w — вершина в P на расстоянии k_u от u.



Допустим, расстояние в Q от u до w меньше k_u.

Пусть e — первое ребро на пути в Q из u в w e, не входящее в P.

Должна найтись грань в P, которой нет в Q, и которая пересекает e.

Получается, что до этой грани можно было дойти в P за k_u шагов!

Доказательство теоремы Калаи—Клейтмана

- $\mathrm{H}(n, d)$ — максимальный диаметр графа d-мерного n-гранника.
- $\mathrm{H}(n, d) \leq n^{\log_2 d+2}$ — хотим доказать

k_u — такое максимальное число, что объединение всех цепей длины k_u с началом в u затрагивает не более $\lfloor n/2 \rfloor$ граней.

- $k_u \leq \mathrm{H}(\lfloor n/2 \rfloor, d)$.
- $k_v \leq \mathrm{H}(\lfloor n/2 \rfloor, d)$.

Найдётся грань, достижимая одновременно из u и v за $(k_u + 1)$ и $(k_v + 1)$ шагов соответственно. Отсюда

$$\mathrm{H}(n, d) \leq 2(\mathrm{H}(\lfloor n/2 \rfloor, d) + 1) + \mathrm{H}(n - 1, d - 1).$$

Доказательство теоремы Калаи—Клейтмана

- $\mathrm{H}(n,d)$ — максимальный диаметр графа d-мерного n-гранника.
- $\mathrm{H}(n,d) \le n^{\log_2 d+2}$ — хотим доказать

$$\mathrm{H}(n,d) \le 2(\mathrm{H}(\lfloor n/2 \rfloor, d) + 1) + \mathrm{H}(n-1, d-1)\,.$$

Используя предположение индукции, получаем:

$$\mathrm{H}(n,d) \le 2 \cdot \lfloor n/2 \rfloor^{\log_2 d+2} + (n-1)^{\log_2(d-1)+2} + 2.$$

Осталось доказать неравенство

$$2 \cdot (n/2)^{\log_2 d+2} + (n-1)^{\log_2(d-1)+2} + 2 \le n^{\log_2 d+2}.$$

Доказательство теоремы Калаи—Клейтмана

$$2 \cdot (n/2)^{\log_2 d+2} + (n-1)^{\log_2(d-1)+2} + 2 \overset{?}{\leq} n^{\log_2 d+2}$$

Эквивалентная форма:

$$\frac{1}{2d} + \frac{(n-1)^{\log_2(d-1)+2}}{n^{\log_2 d+2}} + \frac{2}{n^{\log_2 d+2}} \overset{?}{\leq} 1.$$

Имеем

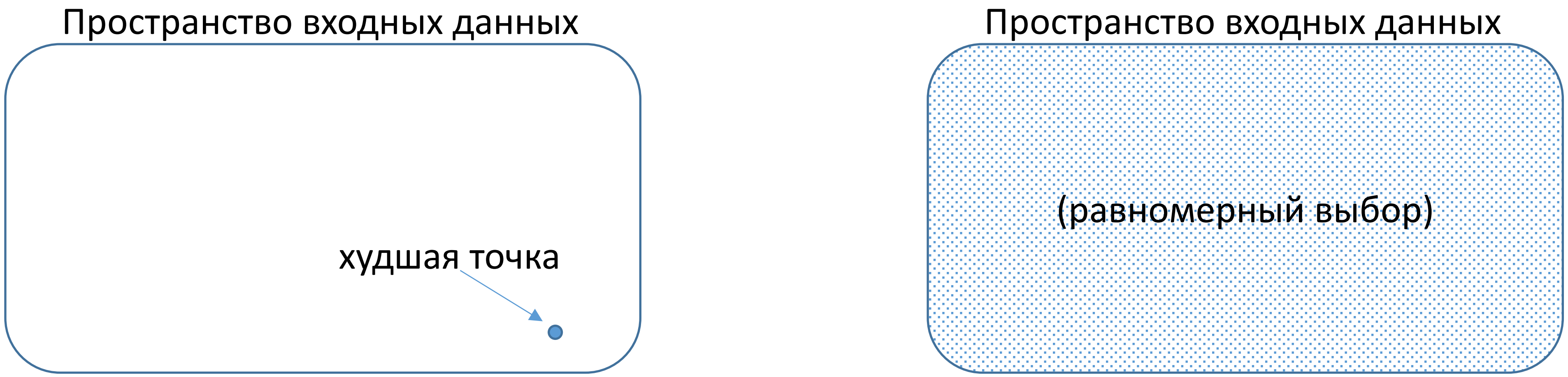
$$\frac{1}{2d} + \frac{(n-1)^{\log_2(d-1)+2}}{n^{\log_2 d+2}} + \frac{2}{n^{\log_2 d+2}} \leq \frac{1}{2d} + \frac{n^{\log_2(d-1)+2}}{n^{\log_2 d+2}} + \frac{2}{n^{\log_2 d+2}} =$$

$$= \frac{1}{2d} + (1-1/d)^{\log_2 n} + \frac{2}{d^{\log_2 n} \cdot n^2} \leq \frac{1}{2d} + 1 - \frac{1}{d} + \frac{1}{2d} \leq 1.$$

Понятие о сглаженном анализе алгоритмов (smoothed analysis of algorithms)

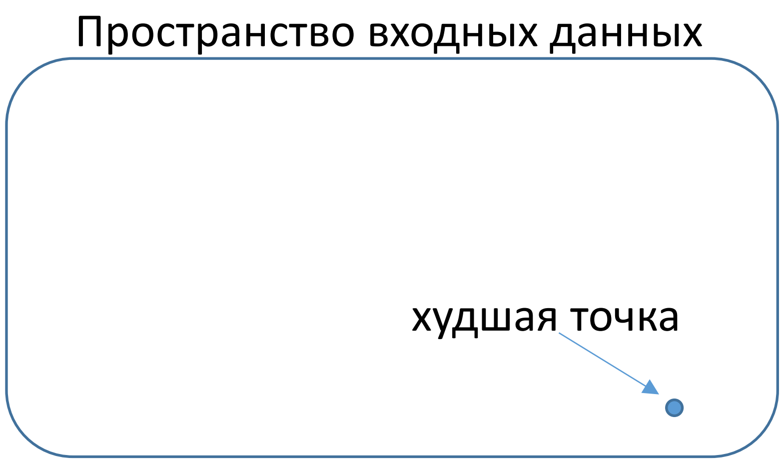
Две популярных методики анализа детерминированных алгоритмов:

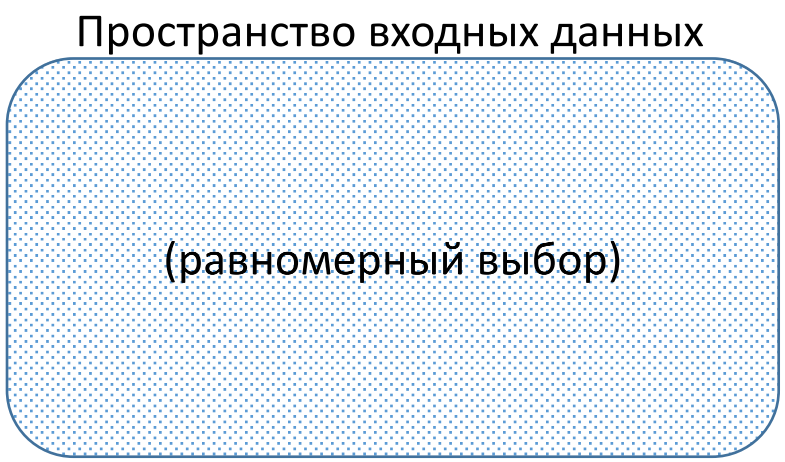
- Время работы алгоритма в худшем случае
- Время работы алгоритма в среднем на случайном входе



Сглаженный анализ алгоритмов

- Минус анализа худших случаев: чаще всего реальные входы алгоритма устроены вовсе не худшим возможным образом.
- Минус анализа на случайном (равномерно распределённом) входе: реальные данные обычно нельзя считать совершенно случайными.



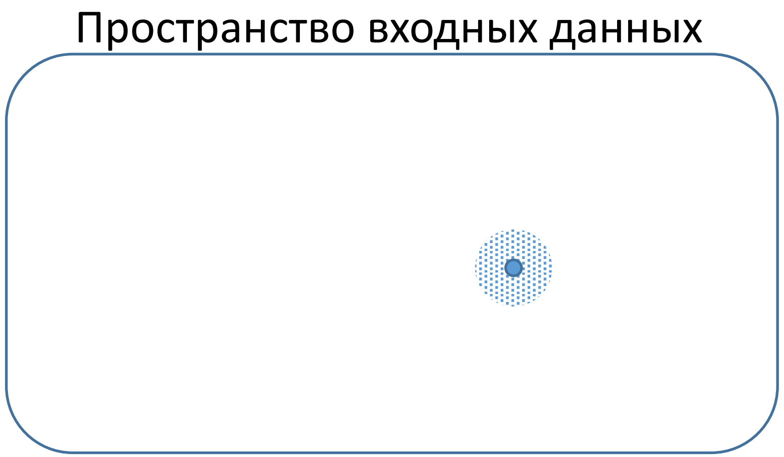


Сглаженный анализ алгоритмов

Smoothed analysis (Daniel A. Spielman & Shang-Hua Teng, 2001).

Предпосылки: реальные входные данные обычно имеют некоторую *неслучайную базовую структуру*, но зашумлены *небольшими случайными отклонениями*.

Идея: рассмотрим поведение алгоритма на точках, случайно выбираемых из малой окрестности фиксированной точки.



Сглаженный анализ алгоритмов

Рассмотрим задачу

$$\begin{cases} \boldsymbol{z}^{\mathrm{T}}\boldsymbol{x} \to \max \\ \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b} \end{cases}$$

Существует такой вариант симплекс-метода $\mathcal{A}$ и такой многочлен $\mathcal{P}$, что для любых $n > d \ge 3$ и любых $\boldsymbol{A}_0 = (\boldsymbol{a}_1, \dots, \boldsymbol{a}_n) \in \mathbb{R}^{n\times d}$, $\boldsymbol{b}_0 \in \mathbb{R}^n$, $\boldsymbol{z} \in \mathbb{R}^d$ и $\sigma > 0$ выполнено неравенство

$$\mathbb{E}_{\boldsymbol{A},\boldsymbol{y}}\big(\#\text{шагов } \mathcal{A} \text{ на входе } (\boldsymbol{A}, \boldsymbol{y}, \boldsymbol{z})\big) \le \mathcal{P}(d, n, \sigma^{-1}).$$

Здесь $\mathbb{E}_{\boldsymbol{A},\boldsymbol{y}}$ — это матожидание при случайном выборе $\boldsymbol{A}$ и $\boldsymbol{y}$ из гауссовских распределений с центрами $\boldsymbol{A}_0$ и $\boldsymbol{y}_0$ и стандартными отклонениями $\sigma \max\limits_i \|(\boldsymbol{a}_i, y_i)\|$.