

Спецкурс 2020/2021: “Геометрические и комбинаторные свойства матриц и аппроксимация”
Блок лекций “Сложность матриц и аппроксимация”
Лекция 6: “Аппроксимативный ранг”

1 декабря 2020 г.

Определение

Пусть A — вещественная матрица, и $\varepsilon > 0$. Определим аппроксимативный ранг (ε -ранг) матрицы A следующим образом:

$$\text{rank}_\varepsilon(A) := \min\{\text{rank } B : \max_{i,j} |A_{i,j} - B_{i,j}| \leq \varepsilon\}.$$

Разминка:

- пусть $A \in \{0, 1\}^{M \times N}$ — булева матрица, оцените $\text{rank}_{1/2}(A)$;
- если $\varepsilon < \min |A_{i,j}|$, то $\text{rank}_\varepsilon(A) \geq \text{rank}_\pm(A)$.

Понятие ε -ранга возникло в теории коммуникационной сложности, см. работы Buhrman, Wolf (2000), Lee, Shraibman (2008) и другие.

Вспомним, что в модели детерминированной коммуникации есть нижняя оценка сложности $C(f) \geq \log_2 \text{rank } f$.

В модели вероятностной коммуникации с неограниченной ошибкой $U(f) \approx \log \text{rank}_\pm f$.

В модели квантовой коммуникации с ограниченной ошибкой сложность оценивается через аппроксимативный ранг:

$$Q(f) \gg \log \text{rank}_{1/3}(f).$$

Квантовая коммуникация

Очень кратко про квантовые вычисления.

Обычные m бит определяют *состояние* системы — вектор $x \in \{0, 1\}^m$. Состояние системы с квантовыми m битами (qubits, кубиты) задаётся *амплитудой* $\alpha: \{0, 1\}^m \rightarrow \mathbb{C}$, задающей распределение вероятностей:

$$\sum_x |\alpha_x|^2 = 1.$$

При измерении состояния мы получаем вектор $y \in \{0, 1\}^m$ с вероятностью $P(y) = |\alpha_y|^2$.

Над состояниями можно выполнять унитарные преобразования $U: \alpha \mapsto U\alpha$, где $UU^* = I$.

Предполагается, что есть квантовая система из m кубит, Анна и Борис могут выполнять преобразования над своими кубитами, а также обмениваться кубитами.

Поперечник

Поперечником по Колмогорову порядка n множества W в нормированном пространстве X называется величина

$$d_n(W, X) := \inf_{\substack{L_n \subset X \\ \dim L_n \leq n}} \rho(W, L_n)_X,$$

где $\rho(W, L)_X := \sup_{x \in W} \inf_{y \in L} \|x - y\|_X$, а L_n — линейные пространства размерности не выше n .

Понятие поперечника возникло в теории аппроксимации в работе А.Н. Колмогорова ещё в 1936 году.

Мотивация: обобщение результатов об аппроксимации алгебраическими полиномами \mathcal{P}_n , тригонометрическими \mathcal{T}_n на произвольные системы из n функций.

Теория поперечников активно развивалась в 70-е – 80-е годы (Тихомиров, Исмагилов Кашин, Глускин, Майоров, Пинкус, Лоренц и другие).

Эквивалентность

Пусть $A \in \mathbb{R}^{M \times N}$. Обозначим через $W_A \subset \mathbb{R}^N$ множество векторов A_i – строк матрицы A ($i = 1, \dots, M$).

Функции rank_ε и d_n обратны друг к другу:

$$\text{rank}_\varepsilon(A) \leq n \iff d_n(W_A, \ell_\infty^N) \leq \varepsilon.$$

Действительно, если матрица B ранга n приближает A поэлементно, то строки B приближают вектора из W_A и лежат в n -мерном пространстве.

Замечания.

- вместо строк можно рассматривать столбцы матрицы;
- множество W_A можно заменить на $\text{conv}(\pm W_A)$.

Оценки для rank_ε

Мы будем пользоваться соглашением: если $\chi(A)$ — некоторая характеристика матриц, то через $\chi_\varepsilon(A)$ обозначим минимум χ в ε -окрестности:

$$\chi_\varepsilon(A) := \inf\{\chi(B) : \|A - B\|_\infty \leq \varepsilon\}.$$

Например, rank_ε . Другими важными для нас характеристиками будут

$$\gamma_{2,\varepsilon}(A) := \inf\{\gamma_2(B) : \|A - B\|_\infty \leq \varepsilon\}$$

и

$$\|A\|_{S_{p,\varepsilon}} := \inf\{\|B\|_{S_p} : \|A - B\|_\infty \leq \varepsilon\},$$

где $\|A\|_{S_p} = \|\sigma(A)\|_{\ell_p}$ — норма Шаттена.

Спектральный метод

Вектор $\sigma(A)$ сингулярных чисел матрицы A имеет $\text{rank } A$ ненулевых координат. Следовательно,

$$\|\sigma(A)\|_p \leq (\text{rank } A)^{1/p-1/q} \|\sigma(A)\|_q, \quad p \leq q.$$

При $p = 1, q = 2$ получаем неравенство

$$\|A\|_{\Sigma} \leq \sqrt{\text{rank } A} \cdot \|A\|_F. \quad (*)$$

Нормы Шаттена при $p = 1, 2, \infty$ обладают особыми свойствами и названиями:

- $\|A\|_{S_1} = \|A\|_{\Sigma}$ — следовая норма;
- $\|A\|_{S_2} = \|A\|_F = (\sum A_{i,j}^2)^{1/2}$ — норма Фробениуса;
- $\|A\|_{S_{\infty}} = \|A\|_{2 \rightarrow 2}$ — спектральная норма.

Перейдём к ε -окрестности: подставим в неравенство (*) матрицу A_{ε} , на которой достигается ε -ранг:

$$\|A\|_{\Sigma, \varepsilon} \leq \sqrt{\text{rank}_{\varepsilon} A} \cdot (\|A\|_F + \varepsilon \sqrt{MN}).$$

Мы доказали оценку снизу для ε -ранга: $\text{rank}_{\varepsilon}(A) \geq \frac{\|A\|_{\Sigma, \varepsilon}^2}{(\|A\|_F + \varepsilon \sqrt{MN})^2}$.

Утверждение

$$\gamma_{2,\varepsilon}(A) \leq \sqrt{\text{rank}_\varepsilon A} \cdot (\|A\|_\infty + \varepsilon).$$

Доказательство. Ранее (см. лекцию №4) мы доказали неравенство

$$\gamma_2(A) \leq \sqrt{\text{rank } A} \cdot \|A\|_\infty.$$

Перейдём в этом неравенстве к ε -окрестностям:

$$\gamma_2(A_\varepsilon) \leq \sqrt{\text{rank}_\varepsilon A} \cdot \|A_\varepsilon\|_\infty.$$

Ч.т.д.

Мы получили нижнюю оценку для ε -ранга:

$$\text{rank}_\varepsilon(A) \geq \frac{\gamma_{2,\varepsilon}(A)^2}{(\|A\|_\infty + \varepsilon)^2}.$$

Оценка сверху

Утверждение

$$\text{rank}_{2\varepsilon}(A) \leq C\varepsilon^{-2}\gamma_{2,\varepsilon}(A)^2 \log(M + N).$$

Докажем неравенство

$$\text{rank}_\varepsilon(A) \leq C\varepsilon^{-2}\gamma_2(A)^2 \log(M + N), \quad (1)$$

после чего подставим в него \tilde{A}_ε , на которой достигается $\gamma_{2,\varepsilon}(A)$

$$\text{rank}_{2\varepsilon}(A) \leq \text{rank}_\varepsilon(\tilde{A}_\varepsilon) \leq C\varepsilon^{-2}\gamma_{2,\varepsilon}(A) \log(MN).$$

Левая и правая части неравенства (1) однородны по A , поэтому можно заменить A на $A' = A/\gamma_2(A)$ и считать, что $\gamma_2(A) = 1$. **На самом деле нет!** Но однородность есть по ε ! Положим $\delta = \varepsilon/\gamma_2(A)$, тогда

$$\text{rank}_\varepsilon(A) = \text{rank}_\delta(A'), \quad \varepsilon^{-2}\gamma_2(A)^2 = \delta^{-2}\gamma_2(A')^2.$$

Значит, мы можем считать, что $\gamma_2(A) = 1$ и доказывать неравенство $\text{rank}_\varepsilon(A) \leq C\varepsilon^{-2} \log(M + N)$.

Итак, пусть есть представление $A_{i,j} = \langle x_i, y_j \rangle$, где $|x_i|, |y_j| \leq 1$.
Воспользуемся леммой Johnson–Lindenstrauss

Утверждение

Пусть R — матрица $d \times N$ со стандартными гауссовыми элементами, т.е. $R_{i,j} \sim \mathcal{N}(0, 1)$. Тогда для любых векторов $x, y \in \mathbb{R}^N$, $|x|, |y| \leq 1$ и $\varepsilon \in (0, 1/2)$ имеем

$$P(|\langle \frac{1}{\sqrt{d}}Rx, \frac{1}{\sqrt{d}}Ry \rangle - \langle x, y \rangle| \geq \varepsilon) \leq 2 \exp(-d\varepsilon^2/8).$$

Применяем это утверждение для векторов x_i, y_j и находим, что если $2MN \exp(-d\varepsilon^2/8) < 1$, то найдётся матрица R , такая что вектора $x'_i = Rx_i/\sqrt{d}, y'_j = Ry_j/\sqrt{d} \in \mathbb{R}^d$ и

$$\max_{i,j} |\langle x'_i, y'_j \rangle - \langle x_i, y_j \rangle| \leq \varepsilon.$$

Мы построили требуемую аппроксимацию матрицы A матрицей ранга $d \asymp \varepsilon^{-2} \log(MN)$ с погрешностью ε .

Утверждение доказано.

Следствие. Пусть S — сигнум $N \times N$ матрица, $\varepsilon = 1/2$. Тогда

$$\frac{1}{3}\gamma_{2,\frac{1}{2}}^2(S) \leq \text{rank}_{1/2}(S) \leq C\gamma_{2,\frac{1}{4}}(S)^2 \log N.$$

Связь ε -ранга с margin.

Из определения максимального отступа $\min \frac{\langle x_i, y_j \rangle}{|x_i| \cdot |y_j|}$ следует неравенство

$$\text{margin}(\text{sign } B) \geq \frac{\min_{i,j} |B_{i,j}|}{\gamma_2(B)}.$$

Отсюда для произвольной сигнум-матрицы S и $\varepsilon \in [0, 1)$ получим

$$\text{margin}(S) \geq \frac{1 - \varepsilon}{\gamma_{2,\varepsilon}(S)}.$$

Комбинируя это с полученной ранее оценкой

$$\text{rank}_\varepsilon(A) \geq \frac{\gamma_{2,\varepsilon}(A)^2}{(\|A\|_\infty + \varepsilon)^2},$$

выводим неравенство

$$\text{rank}_\varepsilon(S) \geq \left(\frac{1 - \varepsilon}{1 + \varepsilon} \right)^2 \cdot \frac{1}{\text{margin}^2(S)}.$$

Пример 1

Пусть U — ортогональная матрица $N \times N$. В нулевой лекции мы упоминали теорему Эккарта–Юнга:

$$\min_{\text{rank } B \leq n} \|A - B\|_F^2 = \sum_{j>n} \sigma_j(A)^2. \quad (\star)$$

Все сингулярные числа U равны чему? единице. Применим (\star) : $\|U - B\|_F^2 \geq N - n$, если $\text{rank } B \leq n$. В теории поперечников — “теорема о поперечнике октаэдра”:

$$d_n(B_1^N, \ell_2^N) = \inf_{L_n} (N^{-1} \sum_{j=1}^N \rho(e_j, L_n)^2)^{1/2} = \sqrt{1 - n/N}.$$

(Нижняя оценка следует из матричной формы; интересно, что она точна для поперечника.) Пусть $n = N/2$, тогда $\|U - B\|_F^2 \geq N/2$. Очевидно, $\|U - B\|_F^2 \leq N^2 \|U - B\|_\infty^2$, откуда $\|U - B\|_\infty \geq (2N)^{-1/2}$,
 $\text{rank}_\varepsilon(U) \geq N/2$, если $\varepsilon < \frac{1}{\sqrt{2N}}$.

Следствие: для матрицы Адамара $\text{rank}_\varepsilon(H) \geq N/2$ для $\varepsilon = 1/\sqrt{2}$.

Пример 2

Пусть E — единичная $N \times N$ матрица. Оценим $\text{rank}_{1/3}(E)$.

Из оценки $\text{rank}_\varepsilon(A) \ll \varepsilon^{-2} \gamma_2(A) \log N$ и равенства $\gamma_2(E) = 1$ следует, что $\text{rank}_{1/3}(E) \ll \log n$.

Оценку снизу удобно доказывать в терминах поперечников.

$\text{conv } W_E = \text{conv}\{\pm e_i\} = B_1^N = \{x: \|x\|_{\ell_1^N} \leq 1\}$.

$$\text{rank}_\varepsilon(E_N) \leq n \iff d_n(B_1^N, \ell_\infty^N).$$

Мы докажем, что $\text{rank}_{1/3}(E) \gg \log N$. Эквивалентно, нужно показать, что если n -мерное пространство L_n приближает октаэдр в ℓ_∞^N с точностью $\leq 1/3$, то $n \gg \log N$.

Используем соображения *энтропии*. Сформулируем идею в общем виде, для оценки поперечника $d_n(W, X)$.

Пусть тело W лежит в единичном шаре X и в нём нашлось M точек x_1, \dots, x_M с попарными расстояниями $\|x_i - x_j\| \geq 3\varepsilon$.

Пусть также n -мерное пространство L_n приближает W в норме X с точностью ε .

Сопоставим каждой точке x_i элемент $y_i \in L_n$, $\|x_i - y_i\| \leq \varepsilon$. Тогда $\|y_i - y_j\| \geq \varepsilon$. При этом $\|y_i\| \leq 1 + \varepsilon$.

Но в маломерном шаре $B_X \cap L_n$ не может быть слишком много попарно удалённых точек: открытые шары $B(y_i, r = \varepsilon/2)$ не пересекаются лежат в шаре радиуса $(1 + 3\varepsilon/2)$.

$$M \cdot (\varepsilon/2)^n \leq (1 + 3\varepsilon/2)^n, \quad n \geq \frac{\log M}{(3 + 2/\varepsilon)}.$$

В случае B_1^N берём просто вектора e_1, \dots, e_N , т.е. $M = N$, и получаем оценку $n \gg \log N$. Однако метод работает и для других тел, например, шара B_2^N .

Основные результаты по поперечнику октаэдра в ℓ_∞ были получены в работах Кашина и Глускина.

Пример 3

Пусть Δ^N — верхнетреугольная 0/1 матрица, т.е. $\Delta_{i,j} = 1$ при $j \geq i$ и $\Delta_{i,j} = 0$ в остальных случаях.

Проблема. Оценить $\text{rank}_{1/3}(\Delta^N)$. **Порядок этой величины неизвестен!**

Мы докажем, что

$$c \log^2 N \leq \text{rank}_{1/3}(\Delta^N) \leq C \log^3 N.$$

Для оценки сверху воспользуемся неравенством $\text{rank}_{1/3} \ll \gamma_2^2 \log N$ и соотношением

$$\gamma_2(\Delta^N) \asymp \log N. \quad (**)$$

Докажем оценку сверху в (**).

Вспомним эквивалентное выражение для γ_2 -нормы:

$$\gamma_2(\Delta^N) = \max_{\|A\|_{2 \rightarrow 2} \leq 1} \|A \circ \Delta^N\|_{2 \rightarrow 2}.$$

Требуется доказать, что если $\|A\|_{2 \rightarrow 2} \leq 1$, то $\|A \circ \Delta^N\|_{2 \rightarrow 2} \ll \log N$.
 Без ограничения общности считаем, что $N = 2^k$. Треугольник $\{(i, j) : j \geq i\}$ мы разрежем на части R_i , $i = 1, \dots, k$, так что

$$\|A \circ R_i\|_{2 \rightarrow 2} \leq 1.$$

Тогда $\|A \circ \Delta^N\|_{2 \rightarrow 2} \leq k$.

Первый блок R_1 это верхняя правая четверть — квадрат $\{(i, j) : 1 \leq i \leq 2^{k-1}, 2^{k-1} < j \leq 2^k\}$.

Для множества $I \times J$ имеем

$$|A \circ (I \times J)x| = |(Ax_J)_I| \leq |Ax_J| \leq |x_J| \leq |x|,$$

так что $\|A \circ (I \times J)\| \leq 1$.

В качестве R_2 берём объединение двух квадратов в оставшихся двух прямоугольниках: $R_2 = (I' \times J') \sqcup (I'' \times J'')$, тогда

$$\begin{aligned} |A \circ (I' \times J' \sqcup I'' \times J'')x|^2 &= |(Ax_{J'})_{I'} + (Ax_{J''})_{I''}|^2 = \\ &= |(Ax_{J'})_{I'}|^2 + |(Ax_{J''})_{I''}|^2 \leq |x_{J'}|^2 + |x_{J''}|^2 \leq |x|^2. \end{aligned}$$

И т.д. (нужно нарисовать картинку).

Теперь нужно оценить $1/3$ -ранг снизу. Воспользуемся оценкой

$$\text{rank}_{1/3}(A) \gg \gamma_{2,1/3}(A)^2.$$

Рассмотрим матрицу Гильберта

$$G_{i,j} = \begin{cases} 0, & i = j, \\ \frac{1}{i-j}, & i \neq j. \end{cases}$$

Хорошо известно, что $\|G\| \leq C$. С другой стороны, при вырезании верхнего треугольника в $G \circ \Delta$ остаётся только отрицательная часть, получается матрица нормы $\asymp \log N$.






Действительно, при действии $G \circ \Delta$ на вектор из одних единиц половина координат будет по модулю не меньше $\sum_1^{N/2} 1/k \asymp \log N$.

Отсюда $\gamma_2(G) \gg \log N$.

Если $\|\Delta^N - \Delta^N\|_\infty \leq 1/3$, то работают аналогичные соображения, поэтому также $\gamma_{2,1/3}(\Delta^N) \gg \log N$.

Итак, мы доказали, что

$$c \log^2 N \leq \text{rank}_{1/3}(\Delta^N) \leq C \log^3 N.$$

-  H. Buhrman, R. Wolf, “Communication Complexity Lower Bounds by Polynomials” (2000), arXiv:cs/9910010.
-  T. Lee, A. Shraibman, “An approximation algorithm for approximation rank”, arXiv:0809.2093.
-  N. Alon, T. Lee, A. Shraibman, S. Vempala, “The approximate rank of a matrix and its algorithmic applications”, *Proc. of the 2013 ACM Symposium on Theory of Computing*, 675–684.
-  S.V. Lokam, “Complexity Lower Bounds using Linear Algebra” (2009).
-  Б.С. Кашин, Ю.В. Малыхин, К.С. Рютин, “Поперечник по Колмогорову и аппроксимативный ранг” (2018).