Матричные вычисления

По лекциям Максима Рахубы

Содержание

1 О курсе

Большую часть сказанного можно найти в вики.
Правда, кроме указанных на вики источников, было упомянуто ещё два:

1. Gilbert (неразборчиво) — Matrix Methods in Data Science (скорее всего, Gilbert Strang — Matrix Methods in Data Analysis, Signal Processing, and Machine Learning)
2. Ivan Oseledets @ github. Скорее всего, имеются в виду репозитории с названиями `nla20XX`.

2 Основы матричного анализа

2.1 Векторные нормы

Определение 2.1. Векторная норма — функция $f : \mathbb{F}^n \to \mathbb{R}$ такая, что:

- $f(x) \geqslant 0; \quad f(x) = 0 \Leftrightarrow x = 0;$
- $f(\alpha x) = |\alpha|\, f(x);$
- $f(x + y) \leqslant f(x) + f(y).$

Обозначается $\|x\|$.
Примеры:

- L_1-норма (Единичная окружность — ромб, TODO: нарисовать):
$$\|x\|_1 = \sum_{i=1}^{n} |x_i|$$
- L_2-норма (Единичная окружность — окружность):
$$\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} = \sqrt{x^* x}$$
- A-норма:
$$\|x\|_A = \sqrt{x^* A x}, \quad A = A^*, \quad \forall x \neq 0 : x^* A x > 0$$
- L_∞-норма (Единичная окружность — квадрат):
$$\|x\|_\infty = \max_{1 \leqslant i \leqslant n} |x_i|$$
- L_p-норма:
$$\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}, \quad p \geqslant 1.$$

2.1.1 Разреженность в L1-норме

$$Ax = b, \quad A \in \mathbb{R}^{m \times n}, \quad m < n$$

Минимизируем x по L_2- и L_1-норме, в случае L_1 получим **разреженное** решение (с большим кол-вом нулей) (TODO: нарисовать).

2.1.2 Скалярное произведение

Определение 2.2. Скалярное произведение $(x, y) = x^* y$.

Теорема 2.1. *(Неравенство Коши-Буняковского-Шварца).*

$$|(x, y)| \leqslant \|x\| \cdot \|y\|\,.$$

Теорема 2.2. *(Неравенство Гёльдера).*

$$|(x, y)| \leqslant \|x\|_p \cdot \|y\|_q \quad \Leftarrow \quad \begin{cases} p, q \geqslant 1; \\ \frac{1}{p} + \frac{1}{q} = 1. \end{cases}$$

2.1.3 Унитарная инвариантность L2-нормы

Определение 2.3. Унитарная матрица U — $U \in \mathbb{C}^{n \times n}$:

$$U^{-1} = U^* \quad (\Leftrightarrow I = U^* U = U U^*)$$

Утверждение 2.1. Если U — унитарная матрица, то $\|Ux\|_2 = \|x\|_2$.

Доказательство.

$$\|Ux\|_2 = \sqrt{(Ux)^* Ux} = \sqrt{x^* U^* U x} = \sqrt{x^* x} = \|x\|_2\,.$$

□

2.2 Матричные нормы

Определение 2.4. Норма $\|\cdot\|$ называется матричной, если

1. $\|\cdot\|$ — векторная норма на пространстве матриц;
2. $\|AB\| \leqslant \|A\| \cdot \|B\|$ (субмультипликативность).

Примеры:

- Норма Фробениуса:

$$\|A\|_F = \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{trace(A^* A)}.$$

- Операторные нормы. Если $\|\cdot\|_*$, $\|\cdot\|_{**}$ — векторные нормы, то соответствующей им операторной нормой будет

$$\|A\|_{*,**} = \sup_{x \neq 0} \frac{\|Ax\|_*}{\|x\|_{**}} = \sup_{\|y\|_{**} = 1} \|Ay\|_*\,.$$

- Например, операторной нормой, соответствующей L_2-норме, является

$$\|A\|_2 = \sqrt{\lambda_{max}(A^* A)} = \sigma_1(A).$$

Утверждение 2.2. Для любой матрицы A и для любых унитарных матриц U, V верно

$$\|UAV\|_F = \|A\|_F$$
$$\|UAV\|_2 = \|A\|_2$$

Доказательство. Для $\|\cdot\|_2$:

$$\|UAV\|_2 = \sup_{x\neq 0} \frac{\|UAVx\|_2}{\|x\|_2} = \sup_{x\neq 0} \frac{\left\|U^*(UAVx)\right\|_2}{\|Vx\|_2}$$

Заменим Vx на y. В силу обратимости V это будет равно

$$\sup_{y\neq 0} \frac{\|Ay\|_2}{\|y\|_2} = \|A\|_2 .$$

□

2.3 Разложение Шура

Собственное разложение (существует не всегда):

$$A = S\Lambda S^{-1}, \quad \Lambda = diag(\lambda_1, \dots, \lambda_n)$$

Жорданова форма (всегда существует, но неустойчива при вычислениях):

$$A = PJP^{-1}$$

Для вычислений используют разложение Шура.

Теорема 2.3. *(О разложении Шура)*
Для всякой $A \in \mathbb{C}^{n\times n}$ существуют такие унитарная U и верхнетреугольная T, что $A = UTU^*$.

Доказательство. Индукцией по размерности A.
База. $n = 1$: $U = I$, $T = A$.
Переход. $(n-1) \to n$.

Т.к. $\mathbb{C}$ алгебраически замкнуто, у характеристического многочлена A есть хотя бы один корень, т.е. у A есть хотя бы одно собственное значение λ_1, т.е. всегда найдётся хотя бы один ненулевой собственный вектор v_1 единичной длины.

Дополним v_1 до ортонормированного базиса $v_1, \dots, v_n$ и положим $U_1 = (v_1|\dots|v_n)$.

v_1 — собственный вектор, так что $Av_1 = \lambda_1 v_1$. Тогда, в силу ортогональности v_i и v_j,

$$v_i^* A v_1 = \begin{cases} \lambda, & i = 1; \\ 0, & i \neq 1. \end{cases}$$

Поумножаем пару матриц:

$$U_1^* A U_1 = \begin{pmatrix} v_1^* \\ \vdots \\ v_n^* \end{pmatrix} \cdot \begin{pmatrix} Av_1 & \dots & Av_n \end{pmatrix} = \begin{pmatrix} \lambda & v_1^* A v_2 & \dots \\ 0 & & \\ \vdots & A_1 & \\ 0 & & \end{pmatrix}$$

По индукции разложим A_1 как $V_1 T_1 V_1^*$. Запишем $U_1^* A U_1$ с помощью блочного умножения:

$$\begin{pmatrix} 1 & 0 \\ 0 & V_1 \end{pmatrix} \cdot \begin{pmatrix} \lambda & \dots \\ 0 & T_1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & V_1^* \end{pmatrix}$$

В силу обратимости V_1^* мы можем так сделать (иначе вектора-строки для $\dots$ над T_1 могло бы и не существовать).

Получаем, что

$$T = \begin{pmatrix} \lambda & \dots \\ 0 & T_1 \end{pmatrix};$$

$$U = U_1 \cdot \begin{pmatrix} 1 & 0 \\ 0 & V_1 \end{pmatrix}.$$

Действительно, T верхнетреугольная по построению, а U унитарна как произведение двух унитарных матриц. □

2.4 Нормальные матрицы

Определение 2.5. Матрица A называется нормальной, если $A^*A = AA^*$.

Утверждение 2.3. Матрица диагонализуема в унитарном базисе тогда и только тогда, когда она является нормальной.

Доказательство. .

- $(\Rightarrow)$:
$$A^*A = U\Lambda^*U^*U\Lambda U^* = U\Lambda^*\Lambda U^* = U\Lambda\Lambda^*U^* = AA^*.$$

- $(\Leftarrow)$: Разложение Шура для A: UTU^*.
$$A^*A = AA^* \Rightarrow T^*T = TT^*.$$

  Оставшаяся часть доказательства $(\Leftarrow)$ будет в качестве упражнения в ДЗ.

□