Методы оптимизации. Семинар 1. Воспоминания из линейной алгебры.

Корнилов Никита Максимович

МФТИ ФИВТ

4 сентября 2025г

Матрицы и векторы

Мы будем работать с векторами и матрицами:

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n, \quad A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,m} \\ \vdots & \vdots & \vdots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,m} \end{pmatrix} \in \mathbb{R}^{n\times m}.$$

- Складывать можно только матрицы одинаковых размерностей $A, B \in \mathbb{R}^{n\times m} : A + B \in \mathbb{R}^{n\times m}$! Вектора тоже $x, y \in \mathbb{R}^n : x + y \in \mathbb{R}^n$!
- Перемножать матрицы A, B разных размерностей можно только если $A \in \mathbb{R}^{n\times m}$ и $B \in \mathbb{R}^{m\times k} : AB \in \mathbb{R}^{n\times k}$!!!
- В общем случае, переставлять квадратные матрицы $A, B \in \mathbb{R}^{n\times n}$ при умножении нельзя: $AB \neq BA$!!!
- Умножить матрицу $A \in \mathbb{R}^{n\times m}$ справа на вектор x можно только если $x \in \mathbb{R}^m : Ax \in \mathbb{R}^n$!!!
- Слева матрицу $A \in \mathbb{R}^{n\times m}$ можно умножить на строку $y^\top$, где $y \in \mathbb{R}^n$:$y^\top A \in \mathbb{R}^{1\times m}$.

След матрицы

След квадратной матрицы $A \in \mathbb{R}^{n \times n}$ обозначается как $Tr : \mathbb{R}^{n \times n} \to \mathbb{R}$ и считается по формуле:

$$Tr(A) = \sum_{i=1}^{n} a_{ii}.$$

Свойства следа:

1. $Tr(A^\top) = Tr(A), \quad A \in \mathbb{R}^{n \times n},$
2. $Tr(A + B) = Tr(A) + Tr(B), \quad A, B \in \mathbb{R}^{n \times n},$
3. $Tr(cA) = cTr(A), \quad A \in \mathbb{R}^{n \times n}, c \in \mathbb{R},$
4. Циклическое свойство:
   $Tr(A_1 A_2 \dots A_{k-1} A_k) = Tr(A_k A_1 A_2 \dots A_{k-1}), \quad A_1, \dots, A_k \in \mathbb{R}^{n \times n}.$

Скалярное произведение

- Стандартное скалярное произведение $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ считается по формуле
$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i = x^\top y, \quad x, y \in \mathbb{R}^n.$$
Для матриц скалярное произведение $\langle \cdot, \cdot \rangle : \mathbb{R}^{n \times m} \times \mathbb{R}^{n \times m} \to \mathbb{R}$ определено как
$$\langle X, Y \rangle = \sum_{i=1}^{n} \sum_{j=1}^{m} X_{ij} Y_{ij} = \langle Y, X \rangle = \mathrm{Tr}(X^\top Y), \quad X, Y \in \mathbb{R}^{n \times m}.$$
Перестановка матрицы A в скалярном произведении:
$$\langle Ax, y \rangle = \langle x, A^\top y \rangle \quad \langle AX, Y \rangle = \langle X, A^\top Y \rangle.$$
- Поэлементное умножение одинаковых по размерностям матриц обозначается как $\odot : \mathbb{R}^{n \times m} \times \mathbb{R}^{n \times m} \to \mathbb{R}^{n \times m}$:
$$(X \odot Y)_{ij} = X_{ij} * Y_{ij}, \quad X, Y \in \mathbb{R}^{n \times m}.$$

Ортогональные, симметричные и определенные матрицы

- Матрица $U \in \mathbb{R}^{n \times m}$ называется ортогональной или унитарной, если $U^\top U = I_m$. В случае квадратных матриц $U^{-1} = U^T$.
- Множество симметричных матриц $\mathbb{S}^n$:
$$A \in \mathbb{S}^n \quad \Longleftrightarrow \quad A = A^\top.$$
- Множество положительно определённых $\mathbb{S}^n_{++}$:
$$A \in \mathbb{S}^n_{++} \quad \Longleftrightarrow \quad A \in \mathbb{S}^n; \quad \forall x \neq 0 \in \mathbb{R}^n : \quad x^\top A x > 0.$$
Критерий Сильвестра: все угловые миноры имеют положительный определитель.
- Множество положительно полуопределённых $\mathbb{S}^n_{+}$:
$$A \in \mathbb{S}^n_{+} \quad \Longleftrightarrow \quad A \in \mathbb{S}^n; \quad \forall x \in \mathbb{R}^n : \quad x^\top A x \geq 0.$$
Критерий Шварценеггера: все *главные* миноры имеют неотрицательный определитель. Главным минором называется определитель подматрицы, симметричной относительно главной диагонали.

Собственные числа

Для квадратичной матрицы $A \in \mathbb{R}^{n \times n}$ существует n (возможно комплексных) собственных чисел $\{\lambda_i(A)\}_{i=1}^n$, где $\lambda_i(A)$ — i-ое по модулю собственное число.

- Собственное значение $\lambda \in \mathbb{C}$ и собственный вектор $x \neq 0 \in \mathbb{C}^d$:

$$Ax = \lambda x \Longleftrightarrow \det(A - \lambda I) = 0.$$

- Определитель и след матрицы A можно выразить через её собственные значения

$$\det(A) = \prod_{i=1}^n \lambda_i(A), \quad \mathrm{Tr}(A) = \sum_{i=1}^n \lambda_i(A).$$

- Для любой симметричной матрицы $A \in \mathbb{S}^n$ существует действительный ортонормированный базис из собственных векторов $S \in \mathbb{R}^{n \times n}$ и действительных собс. чисел:

$$A = S \Sigma S^\top, \quad S^\top S = I, \quad \Sigma - \text{диагональная с собс. числами.}$$

Векторные нормы

Норма $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}$ — функция, удовлетворяющая свойствам: Неотрицательность, Умножение на скаляр, Неравенство треугольника.

- p-норма $\|\cdot\|_p$ на $\mathbb{R}^n$ для $p \in [1, +\infty]$:

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p\right)^{\frac{1}{p}}, \quad \|x\|_\infty = \max_{i\in[1,n]} |x_i|.$$

  Частные случаи: $\|x\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} = \sqrt{\langle x, x\rangle}, \quad \|x\|_1 = \sum_{i=1}^n |x_i|.$

- Сопряженная норма $\|\cdot\|_*$ для любой нормы $\|\cdot\|$: $\|y\|_* := \sup_{\|x\|\le 1}\langle x, y\rangle$. Для p-нормы сопряженная это q-норма, $\frac{1}{p} + \frac{1}{q} = 1$.
- Скалярное произведение $\langle\cdot,\cdot\rangle_A$ и квадратичная норма $\|\cdot\|_A$, определенная пол. опр. матрицей $A \in \mathbb{S}^n_+$:

$$\langle x, y\rangle_A := x^\top Ay, \quad \|x\|_A = \sqrt{\langle x, x\rangle_A}.$$

- Неравенства Гёльдера и Коши-Буняковского:
  Для векторов $x, y \in \mathbb{R}^n$ и чисел $p \in [1, +\infty]$, $\frac{1}{q} + \frac{1}{p} = 1$ выполняется неравенство

$$|\langle x, y \rangle| \leq \sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \|y\|_q.$$

  В частности неравенство КБШ: $|\langle x, y \rangle| \leq \|x\|_2 \|y\|_2$.
- Все нормы в $\mathbb{R}^n$ эквивалентны, например, для $p < p' \in [1, +\infty]$:

$$\|x\|_{p'} \leq \|x\|_p \leq n^{\frac{1}{p} - \frac{1}{p'}} \|x\|_{p'}, \quad \forall x \in \mathbb{R}^n.$$

  В частности, имеем:

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2, \quad \forall x \in \mathbb{R}^n.$$

Матричные нормы

Норма $\|\cdot\|_p : \mathbb{R}^{n\times m} \to \mathbb{R}$ — функция, удовлетворяющая свойствам: Неотрицательность, Умножение на скаляр, Неравенство треугольника.

- Матричная норма $\|\cdot\|_p$ для $A \in \mathbb{R}^{n\times m}$, индуцированная векторной нормой $\|\cdot\|_p$, определятся как

$$\|A\|_p := \sup_{\|x\|_p=1} \|Ax\|_p.$$

Можно привести замкнутые формы для классических норм

  - $\|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^m |a_{ij}|$,
  - $\|A\|_1 = \max_{1\le j\le m} \sum_{i=1}^n |a_{ij}|$,
  - $\|A\|_2 = \sup\limits_{\langle x,x\rangle=1} \sqrt{\langle Ax, Ax\rangle} = \sqrt{\lambda_{\max}(A^\top A)}$.

- Норма Фробениуса для матрицы $A \in \mathbb{R}^{n\times m}$ определяется как

$$\|A\|_F^2 = \sum_{i=1}^n \sum_{j=1}^m A_{ij}^2 := \mathrm{Tr}(A^\top A) = \sum_{i=1}^m \lambda_i(A^\top A).$$

SVD разложение матрицы

Числа $\sigma_i(A) := \sqrt{\lambda_i(A^\top A)} \geq 0, i \in \overline{1, m}$ называются сингулярными числами $A \in \mathbb{R}^{n \times m}$. В случае симметричной матрицы $A \in \mathbb{S}^m$: $\sigma_i(A) = |\lambda_i(A)|, i \in \overline{1, m}$.

Lemma

Матрица $A \in \mathbb{R}^{n \times m}$ представима в виде SVD-разложения:

$$A = U \Sigma V^\top,$$

где $U \in \mathbb{R}^{n \times n}, V \in \mathbb{R}^{m \times m}$ — ортогональные, $\Sigma \in \mathbb{R}^{n \times m}$ — диагональная матрица, составленная из сингулярных чисел $\sigma_i(A) := \sqrt{\lambda_i(A^\top A)}, i \in \overline{1, m}$, расположенных в порядке убывания.

V - базис из собственных векторов $A^\top A$, U - базис из собственных векторов матрицы $AA^\top$.

Свойства матричных норм

- Индуцированные и Фробениусова нормы удовлетворяют свойству субмультипликативности (в общем случае - нет)

$$\|AB\| \leq \|A\|\|B\|, \quad A \in \mathbb{R}^{n\times m}, B \in \mathbb{R}^{m\times k}.$$

- Для любой $A \in \mathbb{R}^{n\times n}$ и индуцированной или Фробениусовой нормы $\|\cdot\|$ верно $|\lambda_{max}(A)| \leq \|A\|$.
- Для любой $A \in \mathbb{R}^{n\times m}$ верно ортогональной $U \in \mathbb{R}^{k\times n}$ и нормы Фробениуса верно

$$\|UA\|_F = \|A\|_F.$$

- Для любой $A \in \mathbb{R}^{n\times m}$ верно $\|A\|_2^2 \leq \|A\|_\infty \|A\|_1$.
- Для любой $A \in \mathbb{R}^{n\times m}$ верно $\|A\|_2 \leq \|A\|_F \leq \sqrt{m}\|A\|_2$.