Домашнее задание 1

Дедлайн 19 сентября 23:59

Матричные вычисления

В дальнейшем мы регулярно будем пользоваться различными матрично-векторными скалярными произведениями и нормами. В связи с этим, кратко напомним основные понятия и факты из этой области. Всюду в дальнейшем будут использоваться следующие стандартные обозначения:

- $\mathbb{R}$ обозначает множество вещественных чисел;
- $\mathbb{R}^n$ обозначает множество всех n-мерных вещественных вектор-столбцов;
- $\mathbb{R}^{m\times n}$ обозначает множество всех вещественных матриц с m строками и n столбцами;
- $\mathbb{S}^n$ обозначает множество всех $n\times n$ вещественных симметричных матриц;
- $\mathbb{S}^n_+$ и $\mathbb{S}^n_{++}$ обозначают множество всех $n\times n$ вещественных симметричных положительно полуопределенных и положительно определенных матриц соответственно;
- I_n обозначает единичную матрицу размера n.

Заметим, что под векторами из $\mathbb{R}^n$ всюду будут подразумеваться именно вектор-столбцы (а не, например, вектор-строки); таким образом, $\mathbb{R}^n = \mathbb{R}^{n\times 1}$, но $\mathbb{R}^n \neq \mathbb{R}^{1\times n}$. Напомним, что $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{R}^{m\times n}$ и $\mathbb{S}^n$ являются вещественными векторными пространствами (со стандартными операциями сложения и умножения на число).

Для матрицы $A \in \mathbb{R}^{n\times n}$ символ $Tr(A) := \sum_{i=1}^n a_{ii}$ обозначает ее след.

Задача 1. *(1 балл) Покажите, что для любых матриц* $A \in R^{m \times n}$, $B \in R^{n \times m}$ *выполнено*

$$Tr(AB) = Tr(BA)$$

.

Определение 0.1. *(Скалярное произведение)*
Пусть V *— вещественное векторное пространство. Функция* $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{R}$*, которая каждой паре* x*,* y *векторов в* V *ставит в соответствие вещественное число* $\langle x, y \rangle$*, называется вещественным скалярным произведением, если она удовлетворяет следующим аксиомам:*

- *(Положительность) Для любого* $x \in V$ *выполнено* $\langle x, x \rangle \geq 0$. *Более того,* $\langle x, x \rangle = 0$ *тогда и только тогда, когда* $x = 0$.
- *(Симметричность) Для любых* $x, y \in V$ *выполнено* $\langle x, y \rangle = \langle y, x \rangle$.
- *(Линейность) Для любых* $x, y, z \in V$ *выполнено* $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$. *Для любых* $x, y \in V$ *и любого* $\alpha \in \mathbb{R}$ *выполнено* $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$ *.*

Векторное пространство с заданным на нем скалярным произведением называется пространством со скалярным произведением или предгильбертовым пространством. Конечномерное вещественное пространство со скалярным произведением также называют евклидовым пространством.

Определение 0.2. *(Стандартное скалярное произведение в* $\mathbb{R}^n$*)*
В пространстве $\mathbb{R}^n$ *вещественных* n*-мерных вектор-столбцов стандартное скалярное произведение задается формулой*

$$\langle x, y \rangle := x^T y = \sum_{i=1}^{n} x_i y_i$$

.

Определение 0.3. *(Стандартное скалярное произведение в $\mathbb{R}^{m\times n}$). В пространстве $\mathbb{R}^{m\times n}$ матриц можно ввести фробениусово скалярное произведение*

$$\langle A, B\rangle := Tr(ATB) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}b_{ij}$$

. Это скалярное произведение называется стандартным скалярным произведением в $\mathbb{R}^{m\times n}$.

Определение 0.4. *(Стандартное скалярное произведение в Sn) Наследуя фробениусово скалярное произведение из пространства $\mathbb{R}^{n\times n}$ на подпространство симметричных матриц $\mathbb{S}^n$, получаем фробениусово скалярное произведение в $\mathbb{S}^n$:*

$$\langle A, B\rangle := Tr(AB)$$

. Это скалярное произведение называется стандартным скалярным произведением в $\mathbb{S}^n$.

Задача 2. *(1 балл) Докажите, что для любых матриц $A \in \mathbb{R}^{m\times k}$, $B \in \mathbb{R}^{k\times n}$, $C \in \mathbb{R}^{m\times n}$ имеет место*

$$\langle AB, C\rangle = \left\langle B, A^T C\right\rangle = \left\langle A, CB^T\right\rangle$$

.

Задача 3. *(1 балл) Пусть $x, y \in \mathbb{R}^n$. Покажите, что $\left\langle xx^T, yy^T\right\rangle = \langle x, y\rangle^2$.*

Задача 4. *(2 балла) (Неравенство Коши–Буняковского). Пусть V — вещественное векторное пространство со скалярным произведением $\langle\cdot,\cdot\rangle$. Тогда для любых $x, y \in V$ справедливо*

$$|\langle x, y\rangle| \leq \langle x, x\rangle^{1/2} \langle y, y\rangle^{1/2},$$

причем неравенство переходит в равенство тогда и только тогда, когда либо $y = 0$, либо $x = \alpha y$ для некоторого $\alpha \in \mathbb{R}$.

Докажите это утверждение.

Определение 0.5. *(Стандартная евклидова норма в $\mathbb{R}^n$).*

$$\|x\|_2 := \langle x, x\rangle^{1/2} = \left(\sum_{i=1}^{n} x_i^2\right)^{1/2}$$

Эта норма также известна как l_2-норма.

Определение 0.6. *(l_1-норма). Помимо евклидовой нормы в пространстве $\mathbb{R}^n$ также можно задать и неевклидову норму. Например, функция $\|\cdot\|_1 : \mathbb{R}^n \to R$, определенная по формуле*

$$\|\cdot\|_1 = \sum_{i=1}^{n} |x_i|,$$

задает норму, которая называется l_1-нормой.

Определение 0.7. *(l_∞-норма). Еще одной популярной нормой в пространстве Rn является l_∞-норма*

$$\|x\|_\infty := \max_{1\le i\le n} |x_i|$$

.

Эта норма также известна как равномерная норма или норма Чебышева.

Задача 5. *(2 балла) Пусть $x \in \mathbb{R}^n$. Докажите следующие неравенства:*

$$\|x\|_\infty \le \|x\|_2 \le \sqrt{n}\,\|x\|_\infty$$
$$\frac{1}{\sqrt{n}}\,\|x\|_1 \le \|x\|_2 \le \|x\|_1$$

. (Подсказка: для первой части второго неравенства используйте неравенство Коши–Буняковского.)

Определение 0.8. *(Фробениусова норма в $\mathbb{R}^{m\times n}$). Рассмотрим теперь в качестве V пространство матриц $\mathbb{R}^{m\times n}$ со стандартным скалярным произведением. Соответствующая евклидова норма в* *данном случае называется фробениусовой нормой и задается формулой*

$$\|A\|_F := \langle A, A\rangle^{1/2} = [Tr(A^T A)]^{1/2} = \left(\sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}^2\right)^{1/2}.$$

Эта норма также известна как норма Гильберта–Шмидта.

Определение 0.9. *(Операторная норма в $\mathbb{R}^{m\times n}$). Помимо фробениусовой нормы важным примером матричной нормы в пространстве $\mathbb{R}^{m\times n}$ является операторная норма:*

$$\|A\|_{op} := \max_{x\in\mathbb{R}^n : |x|=1} |Ax|$$

. Эта норма также известна как спектральная норма.

Задача 6. *(4 ,балла) Упростите каждое из следующих выражений:*

- $Det(AXB(C^{-T}X^TC)^{-T})$, *где* $A, B, C, X \in \mathbb{R}^{n\times n}$, $Det(C) \neq 0$, $Det(C^{-T}X^TC) \neq 0$.
- $\|uvT - A\|_F^2 - \|A\|_F^2$, *где* $u \in \mathbb{R}^m$, $v \in \mathbb{R}^n$, $A \in \mathbb{R}^{m\times n}$.
- $Tr((2I_n + aa^T)^{-1}(uv^T + vu^T))$, *где* $a, u, v \in \mathbb{R}^n$.
- $\sum_{i=1}^n \left\langle S^{-1}a_i, a_i\right\rangle$, *где* $a_1, \ldots, a_n \in \mathbb{R}^n$, $S := \sum_{i=1}^n a_i a_i^T$, $Det(S) \neq 0$.

Матрично-векторное дифференцирование

Для вычисления большинства производных, которые возникают на практике, достаточно лишь небольшой таблицы стандартных производных и правил преобразования. Удобнее всего оказывается работать в терминах «дифференциала» — с ним можно не задумываться о промежуточных размерностях, а просто применять стандартные правила. Здесь A, B — фиксированные матрицы; α — фиксированный скаляр; X, Y — произвольные дифферен- цируемые матричные функции (согласованные по размерностям, чтобы все операции имели смысл); ϕ — произвольная дифференцируемая скалярная функция.

Правила преобразования
$dA = 0$
$d(\alpha X) = \alpha(dX)$
$d(AXB) = A(dX)B$
$d(X+Y) = dX + dY$
$d(X^T) = (dX)^T$
$d(XY) = (dX)Y + X(dY)$
$d\langle X, Y\rangle = \langle dX, Y\rangle + \langle X, dY\rangle$
$d\left(\frac{X}{\phi}\right) = \frac{\phi dX - (d\phi)X}{\phi^2}$

Таблица стандартных производных
$d\langle A, X\rangle = \langle A, dX\rangle$
$d\langle Ax, x\rangle = \langle (A + A^T)x, dx\rangle$
$d\langle Ax, x\rangle = 2\langle Ax, dx\rangle$ **(если** $A = A^T$**)**
$d(\mathrm{Det}(X)) = \mathrm{Det}(X)\langle X^{-T}, dX\rangle$
$d(X^{-1}) = -X^{-1}(dX)X^{-1}$

Одним из самых важных является правило производной композиции. Пусть $g(Y)$ и $f(X)$ — две дифференцирумые функции, и мы знаем выражения для их дифференциалов: $dg(Y)$ и $df(X)$. Чтобы посчитать производную композиции $\phi(X) := g(f(X))$, как и в скалярном случае, нужно взять выражение посчитанного дифференциала $dg(Y)$, подставить в него вместо Y значение $f(X)$, а вместо dY значение $df(X)$.

Обычно, все возникающие на практике матрично-векторные функции составлены с помощью табличных функций и стандартных операций над ними. Благодаря универсальности приведенных правил, дифференцировать сколь угодно сложные функции такого типа становится настолько же просто, как и дифференцировать одномерные функции.

Объект $\nabla f(x)$ (вектор для функции векторного аргумента и матрица для функции матричного аргумента) называется градиентом. Матрица $J_f(x) = \left(\frac{\partial^2 f}{\partial x_i \partial x_j}\right)$ называется матрицей Якоби. Найти вторую производную функции f(X) можно по следующему «алгоритму»: посчитать первую производную функции; зафиксировать в выражениии для df(X) приращение dX — обозначить его как dX_1; посчитать производную для функции $g(X) = df(X)$, считая dX_1 фиксированным (константа). Новое приращение обозначать dX_2.

Для второй производной каноническая форма для скалярной функции векторного аргумента $d^2f(x) = \left\langle \nabla^2 f(x) dx_1, dx_2 \right\rangle$. Матрица $\nabla^2 f(x)$ называется гессианом. Для дважды непрерывно дифференцируемых функций гессиан является симметричной матрицей.

Задача 7. *(2 балла) Пусть f — одна из следующих функций:*

- $f : E \to R$ *— функция* $f(t) := Det(A - tI_n)$*, где* $A \in \mathbb{R}^{n \times n}$*,* $E := \{t \in \mathbb{R} : Det(A - tI_n) \neq 0\}$.
- $f : \mathbb{R}_{++} \to R$ *— функция* $f(t) := |(A + tI_n)^{-1} b|^2$*, где* $A \in \mathbb{S}^n_+$*,* $b \in R^n$.

Для каждого из указанных вариантов вычислите первую и вторую производные f' и f''.

Задача 8. *(3 балла) Пусть f — одна из следующих функций:*

- $f : \mathbb{R}^n \to \mathbb{R}$ *— функция* $f(x) := \frac{1}{2} \left\| xx^T - A \right\|_F^2$*, где* $A \in \mathbb{S}^n$.
- $f : R^n \setminus \{0\} \to \mathbb{R}$ *— функция* $f(x) := \frac{\langle Ax, x \rangle}{|x|^2}$*, где* $A \in \mathbb{S}^n$.
- $f : R^n \setminus \{0\} \to \mathbb{R}$ *— функция* $f(x) := \langle x, x \rangle^{\langle x, x \rangle}$

Для каждого из указанных вариантов вычислите вектор градиент ∇f и матрицу Гессе $\nabla^2 f$ (относительно стандартного скалярного произведения в пространстве $\mathbb{R}^n$).

Задача 9. *(3 балла) Пусть* $f : \mathbb{S}^n_{++} \to \mathbb{R}$ *— одна из следующих функций:*

- $f(X) := Tr(X^{-1})$.
- $f(X) := \left\langle X^{-1}v, v \right\rangle$, *где* $v \in \mathbb{R}^n$.
- $f(X) := (Det(X))^{1/n}$.

Для каждого из указанных вариантов покажите, что вторая производная $D^2 f(X)[H, H]$ *имеет постоянный знак для всех* $X \in \mathbb{S}^n_{++}$ *и всех* $H \in \mathbb{S}^n$.