OTUS: Basic ML
Модуль 4. Теоретический минимум для ML: линейная алгебра, начала мат.анализа и оптимизации, статистика.

1 Линейная алгебра

1. Даны матрицы:

$$A = \begin{pmatrix} 2 & -4 \\ 3 & 5 \\ -1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 2 & 7 \\ -3 & -4 & 0 \\ 5 & 2 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} 6 & -3 & 9 \\ 4 & -5 & 2 \\ 8 & 1 & 5 \end{pmatrix}.$$

   Посчитайте матрицу $D = A^\top C - 2A^\top B^\top$. Приведите полную последовательность вычислений.

2. Дано выражение:

$$3 \cdot \begin{pmatrix} x & 2 & 3 \\ -1 & y & 4 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 & 2 & -5 \\ 2 & -6 & z \end{pmatrix} = \begin{pmatrix} 8 & v & -1 \\ 1 & 6 & 4 \end{pmatrix}.$$

   Найдите значения x, y, z и v, при которых выражение верно.

3. Укажите те значения параметров p и q, при которых ранг матрицы $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & p \\ 5 & 10 & q \end{pmatrix}$ равен единице.

4. Относительно канонического (стандартного) базиса в $\mathbb{R}^2$ даны три вектора $a_1 = (2, -5)^\top$, $a_2 = (-1, 3)^\top$, и $x = (1, -4)^\top$. Примите векторы a_1, a_2 за новый базис B, предварительно проверив, что они линейно независимы.

   (a) Найдите координаты $[x]_B$ вектора x в новом базисе.

   (b) Предположим, что координаты вектора y в базисе B заданы $[y]_B = (1, 1)^T$. Найдите координаты вектора y в стандартном базисе.

5. Исследовательское задание: малоранговая аппроксимация матрицы. Сгенерируйте случайную квадратную матрицу $A \in \mathbb{R}^{n \times n}$, $n \geq 100$. Выполните сингулярное разложение этой матрицы, и получите три матрицы: U, S, $V^\top$. Выполняйте аппроксимацию матрицы A с рангом r, меняя его значение, например, от 2 до n:

$$\tilde{A} = U[:, \, : r]S[: r, \, : r]V^\top[: r, \, :],$$

   и каждый раз считайте ошибку апроксимации (как восстановленная матрица отличается от исходной):

$$E(r) = ||A - \tilde{A}||_F = \sqrt{\sum_{i=1}^{n}\sum_{j=1}^{n}(a_{ij} - \tilde{a}_{ij})^2}.$$

   Используя библиотеку *matplotlib*, постройте график зависимости ошибки аппроксимации матрицы от ранга r.

2 Начала мат.анализа и оптимизации

1. Посчитайте матрицу Гессе следующей функции:

$$f(x) = x_1^3 - 2x_1x_2 + x_2^2 - 3x_1 - 2x_2, \quad x \in \mathbb{R}^2.$$

   Найдите критические точки x_c, такие что $\nabla f(x_c) = 0$.

2. Проверьте, что функция $f = \ln(\sqrt{x_1} + \sqrt{x_2})$ удовлетворяет уравнению:

$$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} = \frac{1}{2}.$$

3. Предположим, задана функция $f : \mathbb{R}^3 \to \mathbb{R}^2$:

$$f(x, y, z) = (x + y + z,\ xyz)^\top.$$

   Найдите матрицу Якоби J_f функции f, и её численное значение в точке $v = (1, 2, 3)^\top$.

4. (Куб евклидовой нормы). Найти первый и второй дифференциалы $df(x)$ и $d^2f(x)$, а также градиент $\nabla f(x)$ и гессиан $\nabla^2 f(x)$ функции:

$$f(x) = \frac{1}{3}||x||_2^3, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

5. (Евклидова норма). Найдите первый дифференциал $df(x)$, и градиент $\nabla f(x)$ функции:

$$f(x) = ||x||_2, \quad x \in \mathbb{R}^n \setminus \{0\}.$$

6. Найдите первый дифференциал $df(x)$ и градиент $\nabla f(x)$ функции:

$$f(x) = ||Ax||_2, \quad x \in \mathbb{R}^n \setminus \{0\}, \quad A \in \mathbb{R}^{m \times n}.$$

7. Найдите первый дифференциал $df(x)$ и градиент $\nabla f(x)$ функции:

$$f(x) = -e^{-x^\top x}, \quad x \in \mathbb{R}^n.$$

1) $A=\begin{pmatrix}2 & -4\\ 3 & 5\\ -1 & 0\end{pmatrix}$, $B=\begin{pmatrix}1 & 2 & 7\\ -3 & -4 & 0\\ 5 & 2 & 1\end{pmatrix}$, $C=\begin{pmatrix}6 & -3 & 9\\ 4 & -5 & 2\\ 8 & 1 & 5\end{pmatrix}$

$$D=A^TC-2A^TB^T=\begin{pmatrix}2 & 3 & -1\\ -4 & 5 & 0\end{pmatrix}\begin{pmatrix}6 & -3 & 9\\ 4 & -5 & 2\\ 8 & 1 & 5\end{pmatrix}-2\begin{pmatrix}2 & 3 & -1\\ -4 & 5 & 0\end{pmatrix}\begin{pmatrix}1 & -3 & 5\\ 2 & -4 & 2\\ 7 & 0 & 1\end{pmatrix}=$$

$$=\begin{pmatrix}2\cdot6+3\cdot4+(-1)\cdot8 & 2\cdot(-3)+3\cdot(-5)+(-1)\cdot(1) & 2\cdot9+3\cdot2+(-1)\cdot5\\ -4\cdot6+5\cdot4+0\cdot8 & -4\cdot(-3)+5\cdot(-5)+0\cdot1 & -4\cdot9+5\cdot2+0\cdot5\end{pmatrix}-2\cdot$$

$$\cdot\begin{pmatrix}2\cdot1+3\cdot2+(-1)\cdot7 & 2\cdot(-3)+3\cdot(-4)+(-1)\cdot0 & 2\cdot5+3\cdot2+(-1)\cdot1\\ -4\cdot1+5\cdot2+0\cdot7 & -4\cdot(-3)+5\cdot(-4)+0\cdot0 & -4\cdot5+5\cdot2+0\cdot1\end{pmatrix}=$$

$$=\begin{pmatrix}\cancel{12+12-8} & \cancel{-6-12+0}\\ \cancel{-24+20+0} & \cancel{-12-2+0}\end{pmatrix}$$

$$=\begin{pmatrix}12+12-8 & -6-15-1 & 18+6-5\\ -24+20+0 & -12-25+0 & -36+10+0\end{pmatrix}-2\cdot\begin{pmatrix}2+6-7 & -6-12+0 & 10+6-1\\ -4+10+0 & 12-20+0 & -20+10+0\end{pmatrix}=$$

$$=\begin{pmatrix}16 & -22 & 19\\ -4 & -37 & -26\end{pmatrix}-2\cdot\begin{pmatrix}1 & -18 & 15\\ 6 & -8 & -10\end{pmatrix}=\begin{pmatrix}16-2 & -22+36 & 19-30\\ -4-12 & -37+16 & -26+20\end{pmatrix}=\begin{pmatrix}14 & 14 & -11\\ -16 & -21 & -6\end{pmatrix}$$

2) $3\cdot\begin{pmatrix}x & 2 & 3\\ -1 & y & 4\end{pmatrix}+2\begin{pmatrix}1 & 2 & -5\\ 2 & -6 & z\end{pmatrix}=\begin{pmatrix}8 & v & -1\\ 1 & 6 & 4\end{pmatrix}$

$\cancel{=(3x\ 6\ 6}$

$$\begin{pmatrix}3x+2 & 6+2 & 9-10\\ -3+4 & 3y-12 & 12+2z\end{pmatrix}=\begin{pmatrix}8 & v & -1\\ 1 & 6 & 4\end{pmatrix}\quad\left|\begin{array}{l}3x+2=8,\ x=2\\ 3y-12=6,\ y=6\\ 12+2\cdot z=4,\ z=-4\end{array}\right|\quad 6+2=v,\ v=8$$

Ответ: $x=2,\ y=6,\ z=-4,\ v=8$

3) $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 6 & p \\ 5 & 10 & q \end{pmatrix}$, $R(A) = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 0 & p-9 \\ 0 & 0 & q-15 \end{pmatrix} = 1 \Rightarrow$

$\begin{matrix} p-9=0 \\ q-15=0 \end{matrix} \Rightarrow p=9, q=15$

4) $a_1 = (2, -5)^T$, $a_1 = (-1, 3)^T$; $x = (1, -4)$

a) Найти ~~x~~ координаты x в базисе (a_1, a_2)

Решение: $x = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ -4 \end{pmatrix} \Rightarrow \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}^{-1} \cdot \begin{pmatrix} 1 \\ -4 \end{pmatrix}$;

$\begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix} = \dfrac{1}{\begin{vmatrix} 2 & -1 \\ -5 & 3 \end{vmatrix}} \cdot \begin{pmatrix} 2 & -1 \\ -5 & 3 \end{pmatrix}^T$, $\begin{vmatrix} 2 & -1 \\ -5 & 3 \end{vmatrix} = 2 \cdot 3 - (-5) \cdot (-1) = 0$

Определитель $\begin{vmatrix} 2 & -1 \\ -5 & 3 \end{vmatrix}$ равен нулю, определить координаты x по базису (a_1, a_2) не возможно! ~~так как они на~~

2.1) $f(x) = x_1^3 - 2x_1x_2 + x_2^2 - 3x_1 - 2x_2$

$\frac{\partial^2 f}{\partial x_1^2} = 3\cdot 2x_1$, $\frac{\partial^2 f}{\partial x_1 \partial x_2} = \frac{\partial}{\partial x_2}(3x_1^2 - 2x_2 - 3) = -2$

$\frac{\partial^2 f}{\partial x_2 \partial x_1} = \frac{\partial}{\partial x_1}(-2x_1 + 2x_2 - 2) = -2$, $\frac{\partial^2 f}{\partial x_2^2} = 2 \Rightarrow$

$\Rightarrow$ Матрица Гессе $\begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} \\ \frac{\partial f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} \end{pmatrix} = \begin{pmatrix} 6x_1 & -2 \\ -2 & 2 \end{pmatrix}$

Критические точки. Для поиск КТ, найдем $\frac{\partial f}{x_1} = 0$ и $\frac{\partial f}{\partial x_2} = 0$

$3x_1^2 - 2x_2 - 3 = 0$ | $x_2 - x_1 - 1 = 0$ | $3x_1^2 - 2(x_1+1) - 3 = 0$ | $3x_1^2 - 2x_1 - 5 = 0$

$-2x_1 + 2x_2 - 2 = 0$ | $x_2 = x_1 + 1$ | $3x_1^2 - 2x_1 - 2 - 3 = 0$ | $D = (-2)^2 - 4\cdot 3\cdot(-5) = 4 + 20\cdot 3 = 64$

$x_1^1 = \frac{-(-2) + \sqrt{64}}{2\cdot 3} = \frac{10}{6} = \frac{5}{3}$ | $x_2^1 = 1 + \frac{5}{3} = \frac{8}{3}$ | Ответ: критичес. точки равны:

$x_1^2 = \frac{-(-2) - \sqrt{64}}{2\cdot 3} = \frac{-6}{6} = -1$ | $x_2^2 = 1 + (-1) = 0$ | $(x_1^1, x_2^1) = \cancel{\left(\frac{5}{3}, \frac{8}{3}\right)} \left(\frac{5}{3}, \frac{8}{3}\right)$

$(x_1^2, x_2^2) = (-1, 0)$

2.2) $f = \ln(\sqrt{x_1} + \sqrt{x_2})$

$x_1 \frac{\partial f}{\partial x_1} + x_2 \frac{\partial f}{\partial x_2} = \frac{x_1 \cdot \frac{1}{2}\frac{1}{\sqrt{x_1}}}{(\sqrt{x_1} + \sqrt{x_2})} + \frac{x_2 \cdot \frac{1}{2}\frac{1}{\sqrt{x_2}}}{(\sqrt{x_1} + \sqrt{x_2})} =$

$= \frac{1}{2} \cdot \frac{1}{(\sqrt{x_1} + \sqrt{x_2})} \cdot \left(\frac{x_1}{\sqrt{x_1}} + \frac{x_2}{\sqrt{x_2}}\right) = \frac{1}{2} \frac{(\sqrt{x_1} + \sqrt{x_2})}{(\sqrt{x_1} + \sqrt{x_2})} =$

$= \frac{1}{2}$; Ответ: $x_1 \frac{\partial \ln(\sqrt{x_1} + \sqrt{x_2})}{\partial x_1} + \frac{\partial \ln(\sqrt{x_1} + \sqrt{x_2})}{\partial x_2} = \frac{1}{2}$

2.3) $f(x,y,z)=(x+y+z,\ xyz)^T$

$$J_f=\begin{pmatrix}\frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} & \frac{\partial f}{\partial z}\end{pmatrix}=\begin{pmatrix}1 & 1 & 1\\ yz & xz & xy\end{pmatrix}$$

$$J_f(1,2,3)=\begin{pmatrix}1 & 1 & 1\\ 2\cdot 3 & 1\cdot 3 & 1\cdot 2\end{pmatrix}=\begin{pmatrix}1 & 1 & 1\\ 6 & 3 & 2\end{pmatrix}$$

2.4) $f(x)=\frac{1}{3}\|x\|_2^3$; $df=\frac{1}{3}\cdot 3\|x\|_2^2\, d\|x\|$;

$d^2f=2\cdot\|x\|_2\cdot d^2\|x\|$;

$\nabla f(x)=\nabla\frac{1}{3}\|x\|_2^3=\frac{1}{3}\cdot 3\|x\|_2^2\,\nabla\|x\|_2=\|x\|_2^2\cdot \operatorname{sign}|x|$

$\nabla^2 f(x)=2\cdot\|x\|_2\cdot(\operatorname{sign}(x))^2=2\cdot\|x\|_2$

2.5) $f(x)=\|x\|_2$, $df(x)=d\|x\|_2=\operatorname{sing}(x)\cdot dx$

$\nabla f(x)=\operatorname{sing}(x)$

2.6) $f(x)=\|Ax\|_2$, $df=d(\|Ax\|_2)=\|A\|\cdot d\|x\|_2=\|A\|\cdot\operatorname{sing}(x)\cdot$

$\nabla f(x)=\|A\|\cdot\operatorname{sing}(x)$

2.7) $f(x)=-e^{-x^Tx}$, $df=-e^{-x^Tx}\cdot d(-x^T\cdot x)=e^{-x^Tx}\cdot(x^Tdx+dx^Tx)$

$\nabla f(x)=-e^{-x^Tx}\cdot\nabla(-x^T\cdot x)=e^{-x^Tx}\cdot(x^T\cdot\nabla x+(\nabla x)^Tx)$