1. **(20 баллов)**. Пусть задана матрица $A \in \mathbb{R}^{m\times n}$, $m \geq n$.

(a) Покажите, что A можно привести к верхнетреугольной матрице R с помощью преобразований Хаусхолдера, используя

$$2mn^2 - \frac{2}{3}n^3 + \mathcal{O}(mn),$$

арифметических операций.

• $H(v) = I - 2vv^T$

• $R = H_n \cdot \ldots \cdot H_1 \cdot \underset{m\times n}{A}$, $H_i \in \mathbb{R}^{m\times m}$

$\exists H = H_1 \in \mathbb{R}^{m\times m}$. $HA = (I - 2vv^T)A = A - 2vv^TA$. Операции: вычитание - mn оп.

Умножение: умножение матриц ассоциативно $\Rightarrow 2vv^TA = 2v(\underbrace{v^TA}_{}) = v(2v^TA)$

На 2: n операций

Второе: mn штук, каждое считается за 1 операцию

i элем - $m-1$ сложений и m умножений. всего n штук $\Rightarrow 2mn - n$

Итого: $mn + mn + 2mn - n + n = 4mn$

$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & H(v_2) \end{pmatrix} \in \mathbb{R}^{m\times m}$, $H(v_2) \in \mathbb{R}^{m-1\times m-1}$.

$[0 \;\; H(v_2)](H_1A)$ - $4(m-1)(n-1)$ - операций.

$H_k = \begin{pmatrix} I & 0 \\ 0 & H(v_k) \end{pmatrix} \in \mathbb{R}^{m\times m}$, $H(v_2) \in \mathbb{R}^{m-k\times m-k}$ $\Rightarrow 4(m-k)(n-k)$ оп.

Overall число операций: $\sum_{k=0}^{n-1} 4(m-k)(n-k) = 4\sum_{k=0}^{n-1}(mn - k(m+n) + k^2) =$

$= 4mn^2 - 4(m+n)\cdot\frac{n(n-1)}{2} + 4\frac{(n-1)(n-2)(2n-3)}{6} =$

$= 4mn^2 - 2mn^2 - 2n^3 + 2mn + 2n^2 + \frac{2}{3}(2n^3 + O(n)) = 2mn^2 - 2n^3 + \frac{4}{3}n^3 +$

$+ O(mn) = 2mn^2 - \frac{2}{3}n^3 + O(mn)$

(b) Покажите, что количество арифметических операций для вычисления $Q \in \mathbb{R}^{m\times n}$ из тонкого QR будет:

$$2mn^2 - \frac{2}{3}n^3 + \mathcal{O}(mn).$$

$Q = H_1 \cdot \ldots \cdot H_n \begin{pmatrix} I \\ 0 \end{pmatrix}^{n\times n}_{(m-n)\times n}$ 0 операций - обрезали H_n.

На k шаге $H_{n-k} \cdot \begin{pmatrix} I_{n-k} & 0 \\ 0 & H(v_{n-k}) \end{pmatrix}$ и $4(m-n+k)\cdot k$ операций, чтобы умножить $H_{n-k} \cdot \hat{H}_{n-k+1}$

всего операций $\sum_{k=1}^{n} 4mk - 4nk + 4k^2 = 2mn^2 - 2n^3 + \frac{4}{3}n^3 + O(mn) = 2mn^2 - \frac{2}{3}n^2 + O(mn)$

2. **(20 баллов)**. Запишем решение x_μ задачи наименьших квадратов с ℓ_2-регуляризацией

$$\|Ax-b\|_2^2+\mu\|x\|_2^2 \to \min_x$$

для заданной матрицы $A\in\mathbb{R}^{m\times n}$ ранга r, вектора правой части $b \in \mathbb{R}^{m\times n}$ и константы $\mu\in\mathbb{R}$ в виде $x_\mu = B(\mu)b$ с матрицей $B(\mu)\in\mathbb{R}^{n\times m}$, которая выражается через A и μ (см. лекцию).

(a) Покажите, что для $\mu>0$ справедливо:

$$\|B(\mu)-A^+\|_2 = \frac{\mu}{(\mu+\sigma_r(A)^2)\,\sigma_r(A)}.$$

(b) Покажите, что $B(\mu)\to A^+$ и что $x_\mu \to A^+b$ при $\mu\to+0$.

а) $A^+ = V\Sigma^+ U^T$, $\Sigma^+ = \begin{pmatrix}\Sigma_r^{-1} & 0\\ 0 & 0\end{pmatrix}\in\mathbb{R}^{m\times n}$. $B(\mu) = (A^TA+\mu I)^{-1}A^T$.

$$\|B(\mu)-A^+\|_2 = \|(A^TA+\mu I)^{-1}A^T - A^+\|_2 =$$

$$= \|(V\Sigma^T\Sigma V^T + \mu V V^T)^{-1} V\Sigma U^T - V\Sigma^+ U^T\|_2 =$$

$$= \|(V(\Sigma^2 V^T + \mu V^T))^{-1} V\Sigma U^T - V\Sigma^+ U^T\|_2 =$$

$$= \|(V^T)^{-1}(\Sigma^2 - \mu I)^{-1} V^{-1} V\Sigma U^T - V\Sigma^+ U^T\|_2 =$$

$$= \|V((\Sigma^2+\mu I)^{-1}\Sigma U^T - \Sigma^+ U^T)\|_2 = \|((\Sigma^2+\mu I)^{-1}\Sigma - \Sigma^+)U^T\|_2 = \|\underbrace{(\Sigma^2+\mu I)}_{\text{обратима, т.к. } \det\Sigma_r\neq 0}^{-1}\Sigma - \Sigma^+\|_2.$$

Найдем $\sigma_1(\Sigma^2+\mu I)$: $\begin{pmatrix}\sigma_1^2+\mu & & & & \\ & \ddots & & & \\ & & \sigma_r^2+\mu & & \\ & & & \mu & \\ & & & & \ddots\ \mu\end{pmatrix}$ $(\Sigma^2+\mu I)^{-1} = \begin{pmatrix}1/\sigma_1^2+\mu & \dots & 0\\ \vdots & 1/\sigma_r^2+\mu & \\ 0 & & \ddots\ 1/\mu\end{pmatrix}$; $(\Sigma^2+\mu I)^{-1}\Sigma = \begin{pmatrix}\sigma_1/\sigma_1^2+\mu & \dots & 0\\ \vdots & \sigma_r/\sigma_r^2+\mu & \\ 0 & & \ddots\ 0\end{pmatrix}$ $(\Sigma^2+\mu I)^{-1}\Sigma-\Sigma^+=$

$$= \begin{pmatrix}\frac{\sigma_1}{\sigma_1^2+\mu}-\frac{1}{\sigma_1} & \dots & 0\\ \vdots & \frac{\sigma_r}{\sigma_r^2+\mu}-\frac{1}{\sigma_r} & \\ 0 & & \ddots\ 0\end{pmatrix} = (*) \Rightarrow SVD(*) = I(*)(-I) \Rightarrow \sigma_1 = \frac{\mu}{(\sigma_r^2+\mu)\sigma_r}$$ — ответ

б) $\lim\limits_{\mu\to+0}\|B(\mu)b - A^+b\|_2 = \lim\limits_{\mu\to+0}\|(B(\mu)-A^+)b\|_2 = 0$, т.к. $\lim\|B(\mu)\to A^+\|_2 = 0$ по усл.

$X(\mu) = B(\mu)B.$

3. **(15 баллов)**. Покажите, что для решений $x \in \mathbb{R}^n$ задачи $\|Ax - b\| \to \min_x$, где $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ заданы, справедливо:

$$\|x\|_2^2 = \|A^+b\|_2^2 + \|(I - A^+A)y\|_2^2,$$

где y – произвольный вектор (см. обозначения в лекции). Сделайте отсюда вывод, какое решение имеет наименьшую $\|x\|_2$.

Все решения мин. вида $x = A^+b + (I - A^+A)y$ $\forall\, y \in \mathbb{R}^n$ и $x_* = A^+b$ имеет мин. значение $\|x\|_2$.

$\|x\|_2^2 = \|A^+b\|_2^2 + \|(I - A^+A)y\|_2^2 \Rightarrow \exists\, c,d \in \mathbb{R}^n$: $x = c + d$ и $\langle c, d\rangle = 0$

Пусть $d = (I - A^+A)y \in \ker(A)$. $\ker(A) \perp Im(A^T)$, $Ad = (A_{(1)}d, \dots, A_{(m)}d)^T = 0$. $Im(A^T) = Im(V)$, т.к. $A^Tx = V(\Sigma U^Tx)$

$A^+b = V\Sigma^+U^Tb = V(\Sigma^+U^Tb) \in Im(V) = Im(A^T) \perp \ker(A) \Rightarrow \langle A^+b, (I - A^+A)y\rangle = 0$ и вып-ся равенство из усл.

4. **(25 баллов)**. Пусть ненулевые $a, b \in \mathbb{R}^n$, $n \geq 2$ ортогональны друг другу и

$$A = a \circ a \circ a + 2(a \circ b \circ a) + 3(b \circ b \circ a).$$

(a) Запишите матрицы $U, V, W \in \mathbb{R}^{n\times 2}$ из канонического разложения A.
Подсказка: используйте линейность тензорного произведения по каждому из аргументов.

$A = a\circ a\circ a + 2(a\circ b\circ a) + 3(b\circ b\circ a) = a\circ(a+2b)\circ a + 3b\circ b\circ a. \Rightarrow U = [a \;\; 3b]$ $V = [a+2b \;\; b]$ $W = [a \;\; a]$

(b) Запишите ядро $G \in \mathbb{R}^{2\times 2\times 1}$ и факторы U, V, W из разложения Таккера A. $G = \begin{bmatrix}1 & 2\\ 0 & 3\end{bmatrix}$. $U = [a\; b]$; $V = [a\; b]$, $W = [a]$.

(c) Докажите, что мультилинейный ранг тензора A равен $(2, 2, 1)$.

развертки: $A_{(1)} = UG_{(1)}(W \otimes V)^T = \overbrace{[a\; b]}^{rk2}\overbrace{\begin{bmatrix}1 & 2\\ 0 & 3\end{bmatrix}}^{rk2}\overbrace{[\underbrace{a \otimes [a\; b]}_{\begin{pmatrix} a_1a & a_1b\\ \vdots & \vdots\\ a_na & a_nb\end{pmatrix}}]^T}^{rk2}$.

rk $A_{(1)} = 2$

$A_{(2)} = VG_{(2)}(W \otimes U)^T = \underbrace{[a\; b]}_{rk2}\underbrace{\begin{bmatrix}1 & 0\\ 2 & 3\end{bmatrix}}_{rk2}\underbrace{[a \otimes [a\; b]]^T}_{rk2} \Rightarrow rk\, A_{(2)} = 2.$

mlrank$(A) = (2, 2, 1)$

$A_{(3)} = WG_{(3)}(V \otimes U)^T = \underbrace{[a]}_{rk1}\underbrace{[1\,0\,2\,3][[a\,b] \otimes [a\,b]]^T}_{rk1} \Rightarrow rk\, A_{(3)} = 1.$

5. **(20 баллов)**. Пусть $A, B \in \mathbb{R}^{n \times n}$ – некоторые заданные матрицы, и пусть стоит задача вычисления матрично-векторного произведения:

$$y = (A \otimes B)\, x, \quad x \in \mathbb{R}^{n^2}.$$

$A \otimes B \in \mathbb{R}^{n^2 \times n^2}$

$$A \otimes B = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \ddots & \\ a_{n1}B & & a_{nn}B \end{pmatrix}.$$

каждый эл-т – n^2 операций. всего – n^2 эл-тов.

(a) Каково асимптотическое число арифметических операций для вычисления y по x без учета дополнительной структуры матрицы $(A \otimes B)$?

$(A \otimes B)x$ – $(n^2 + n^2 - 1)n^2$ операций. $\Rightarrow O(n^4)$

(b) Предложите алгоритм вычисления y, имеющий асимптотическое число операций $\mathcal{O}(n^3)$. **Подсказка:** в этой задаче может помочь операция векторизации.

Покажем, что $(A \otimes B)x = vec(Bx'A^T)$.

Пусть $x = (x_{11} \dots x_{n1} \dots x_{1n} \dots x_{nn})^T$;

$x' = \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{n1} & \dots & x_{nn} \end{pmatrix}$ (просто делаем reshape.)

$$(A \otimes B)\,x = (A \otimes B) \cdot (x_{11} \dots x_{n1} \dots x_{1n} \dots x_{nn})^T = \begin{pmatrix} \sum a_{1i} B x^{(i)} \\ \vdots \\ \sum a_{ni} B x^{(i)} \end{pmatrix} \quad ①$$

$$BXA^T = \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \begin{pmatrix} x_{11} & \dots & x_{n1} \\ \vdots & & \\ x_{1n} & \dots & x_{nn} \end{pmatrix} \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \\ a_{1n} & \dots & a_{nn} \end{pmatrix} = \begin{pmatrix} B_{(1)}x^{(1)} & \dots & B_{(1)}x^{(n)} \\ \vdots & \ddots & \vdots \\ B_{(n)}x^{(1)} & \dots & B_{(n)}x^{(n)} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & \ddots & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix} =$$

$$= \begin{pmatrix} \sum a_{1i} B_{(1)} x^{(i)} & \dots & \sum a_{ni} B_{(1)} x^{(i)} \\ \vdots & \ddots & \vdots \\ \sum a_{1i} B_{(n)} x^{(i)} & \dots & \sum a_{ni} B_{(n)} x^{(i)} \end{pmatrix} = \left(\sum a_{1i} B x^{(i)} \dots \sum a_{ni} B x^{(i)} \right) \quad ②.$$

Заметим, что ① $\xrightarrow{\text{reshape}}$ ②.

#операций: $B_{(i)}x^{(j)}$ – $2n-1$ операц. всего – n^2 эл-тов $\Rightarrow 2n^3 - n^2$.

BXA^T – аналогично $2n^3 - n^2$. Итого: $O(4n^3 - 2n^2) = O(n^3)$

(c) Как получить число операций $\mathcal{O}(n^2 \log n)$, если A и B являются циркулянтами?

C – циркулянт $\Rightarrow C = F_n^{-1} \operatorname{diag}(F_n C) F_n$, где F_n – матр. Фурье

A, B – циркулянты $\Rightarrow A^T$ и B^T – циркулянты.

Bx: $Bx^{(i)}$ - умножение на циркулянт - $O(n\log n)$ => всего $O(n^2\log n)$. $BXA^T = (A(BX)^T)^T$. $(A(BX)^T)^{(i)}$ - за $O(n\log n)$

всего: $O(n^2\log n)$. Итого: $O(n^2\log n)$