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

$$2mn^2 - \frac{2}{3}n^3 + \mathcal{O}(mn), \qquad H_i \in \mathbb{R}^{m\times m}$$

арифметических операций.

$H(v) = I - 2vv^T$ $\quad R = H_n \cdot \ldots \cdot H_1 A$, H_i - матрица Хаусхолдера

$m \geq n \Rightarrow$ для каждого столбца

① $H = H_1 \in \mathbb{R}^{m\times m}$

v^TA - $(m+m-1)n = 2mn - n$ операций

умножить вектор-строку $1\times n$ на 2. n операций

$$HA = (I - 2vv^T)A = A - \underset{m\times 1\ 1\times m\ m\times n}{2vv^TA} = A - \underset{m\times 1\quad 1\times n}{2v(v^TA)} = A - v(2v^TA)$$

mn операций, разность – mn операций

$$= A - (v \cdot 2v^TA) = HA$$

Итого первой итерации: $4mn$

② $H_2 = \begin{pmatrix} 1 & 0 \\ 0 & H(v_2) \end{pmatrix} \in \mathbb{R}^{m\times m}$, $H(v_2) \in \mathbb{R}^{(m-1)\times(m-1)}$

$[0\ H(v_2)]\,(H_1A)$ - $4(m-1)(n-1) = 4mn - 4n - 4m + 4$

Далее на k-ой итерации $H_k = \begin{pmatrix} I_k & \\ & H(v_k) \end{pmatrix}$ и кол-во операций:

$$4(m-k)(n-k) = 4mn - 4kn - 4km + 4k^2$$

В итоге $\sum_{k=0}^{n-1}(4mn - 4kn - 4km + 4k^2) = 4mn^2 - 4n\sum_{k=0}^{n-1} k - 4m\sum_{k=0}^{n-1} k + 4\sum_{k=0}^{n-1} k^2 =$

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

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

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

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

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

$$\underset{m\times n}{Q} = \underset{m\times m}{H_1 \ldots H_n} \cdot \begin{bmatrix} I \\ 0 \end{bmatrix}_{\substack{n\times n \\ (m-n)\times n}}$$

$H\begin{bmatrix} I \\ 0 \end{bmatrix} = \check{H}$ - блок $m\times n$ из H

В нулевой итерации происходит обрезание H_n.

$(k = 1, \ldots, n)$ $\quad$ из п.а)

Далее на k-ой итерации $H_{n-k} = \begin{pmatrix} I_{n-k} & 0 \\ 0 & H(v_{n-k}) \end{pmatrix}$ и $4(m-n+k)\cdot k$ итераций, чтобы умножить $\underset{m\times m}{H_{n-k}} \cdot \underset{m\times n}{\check{H}_{n-k+1}}$. По сути то же самое, что в п. а), но в обратном порядке.

Всего итераций: $\sum_{k=1}^{n} 4mk - 4nk + 4k^2 = 4m\frac{n(n+1)}{2} - 4n\frac{n(n+1)}{2} + 4\frac{n(n+1)(2n+1)}{6} =$

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

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:

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

$$A^+ = V\Sigma^+ U^T \quad , \Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{m\times n}$$

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

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

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

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

$$= \|\underbrace{(\Sigma^2 + \mu I)^{-1}}\Sigma - \Sigma^+\|_2 \; = \text{ старшее сингулярное значение} \Rightarrow$$

$\det \Sigma_r \neq 0 \Rightarrow \det \Sigma_r^2 \neq 0 \Rightarrow \det(\Sigma^2 + \mu I) \neq 0 \quad , \mu > 0 \Rightarrow$ обратима

$$\Sigma^2 + \mu I = \begin{pmatrix} \sigma_1^2+\mu & & & & & 0 \\ & \ddots & & & & \\ & & \sigma_r^2+\mu & & & \\ & & & \mu & & \\ & & & & \ddots & \\ 0 & & & & & \mu \end{pmatrix} \quad (\Sigma^2 + \mu I)^{-1} = \begin{pmatrix} \frac{1}{\sigma_1^2+\mu} & & & & & 0 \\ & \ddots & & & & \\ & & \frac{1}{\sigma_r^2+\mu} & & & \\ & & & \frac{1}{\mu} & & \\ & & & & \ddots & \\ 0 & & & & & \frac{1}{\mu} \end{pmatrix}$$

$$(\Sigma^2 + \mu I)^{-1}\Sigma = \begin{pmatrix} \frac{\sigma_1}{\sigma_1^2+\mu} & & & & \\ & \ddots & & & \\ & & \frac{\sigma_r}{\sigma_r^2+\mu} & & \\ & & & 0 & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix} \quad (\Sigma^2 + \mu I)^{-1}\Sigma - \Sigma^+ = \begin{pmatrix} \frac{\sigma_1}{\sigma_1^2+\mu} - \frac{1}{\sigma_1} & & & \\ & \ddots & & \\ & & \frac{\sigma_r}{\sigma_r^2+\mu} - \frac{1}{\sigma_r} & \\ & & & 0 \\ & & & & \ddots \\ & & & & & 0 \end{pmatrix} = (*)$$

$$\Rightarrow \text{SVD}(*) = I\,(*)\,(-I) \Rightarrow \quad = \boxed{\frac{\mu}{(\sigma_r^2 + \mu)\sigma_r}}$$

$$\frac{\sigma_r}{\sigma_r^2+\mu} - \frac{1}{\sigma_r} = \frac{\sigma_r^2 - \sigma_r^2 - \mu}{(\sigma_r^2+\mu)\sigma_r} = \frac{-\mu}{(\sigma_r^2+\mu)\sigma_r}$$

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

1) $B(\mu) \to A^+, \mu \to +0 \iff \lim\limits_{\mu\to+0} \|B(\mu) - A^+\|_2 = 0$, а $\lim\limits_{\mu\to+0} \frac{\mu}{(\sigma_r^2+\mu)\sigma_r} = \lim\limits_{\mu\to+0} \frac{0}{\sigma_r^3} = 0$

2) $x_\mu = B(\mu)b \Rightarrow \underbrace{B(\mu)}b \to \underbrace{A^+}b, \mu \to +0$

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 $\|x\|_2$.

Знаем, что все решения имеют вид $x = A^+b + (I - A^+A)y \quad \forall y \in \mathbb{R}^n$

и $x_* = A^+b$ имеет минимальное значение $\|x\|_2$

Надо показать: $\|x\|_2^2 = \|A^+b\|_2^2 + \|(I-A^+A)y\|_2^2 \iff$

$\iff \langle x,x\rangle = \langle A^+b + (I-A^+A)y, A^+b + (I-A^+A)y\rangle = \langle A^+b, A^+b\rangle + \langle (I-A^+A)y, (I-A^+A)y\rangle$

$\Rightarrow A^+b \perp (I-A^+A)y$, т.к. $I-A^+A$ - ортопроектор на $\ker(A) \Rightarrow$

$\Rightarrow (I-A^+A)y \in \ker(A)$. $\ker(A) \perp Im(A^T) = Im(V) = Im(A^+)$

$(A^T = V\Sigma U^T,\ A^+ = V\Sigma^+U^T)$. $A^+b \in Im(A^+) \Rightarrow A^+b \perp (I-A^+A)y \Rightarrow$

$\Rightarrow \langle A^+b, (I-A^+A)y\rangle = 0.$

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.

(b) Запишите ядро $G \in \mathbb{R}^{2 \times 2 \times 1}$ и факторы U, V, W из разложения Таккера A.

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

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

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

$U = [a \;\; 3b] \quad V = [a + 2b \;\; b] \quad W = [a \;\; a]$

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

$G = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \qquad U = [a \; b] \quad V = [a \; b] \quad W = [a]$

(c) Выпишем развертки и посчитаем ранги:

$A_{(1)} = U G_{(1)} (W \otimes V)^T = \underbrace{[a \; b]}_{rk=2} \underbrace{\begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}}_{rk=2} [\underbrace{a \otimes [a \; b]}_{rk=2}]^T$ (✱) $\quad rk A_{(1)} = 2$

$A_{(2)} = V G_{(2)} (W \otimes U)^T = \underbrace{[a \; b]}_{rk=2} \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}}_{rk=2} [\underbrace{\overset{n\times1}{a} \otimes \overset{n\times2}{[a \; b]}}_{rk=2}]^T \quad rk A_{(2)} = 2$

$A_{(3)} = W G_{(3)} (V \otimes U)^T = \underbrace{[a]}_{n\times1} \underbrace{[1 \; 0 \; 2 \; 3] ([a \, b] \otimes [a \, b])^T}_{1\times n^2} \quad rk A_{(3)} = 1$

(✱) $a \otimes [a \; b] = \begin{pmatrix} a_1 a & a_1 b \\ a_2 a & a_2 b \\ \vdots & \vdots \\ a_n a & a_n b \end{pmatrix} \in n^2 \times 2$, столбцы линейно независимы

$\text{mlrank}\, A = (rk A_{(1)}, rk A_{(2)}, rk A_{(3)}) = (2, 2, 1)$

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

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

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

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

$$\parallel$$

$$\begin{pmatrix} a_{11}B & a_{12}B & \dots & a_{1n}B \\ \vdots & & & \\ a_{n1}B & \dots & & a_{nn}B \end{pmatrix}$$

Каждый элемент вычисляется за n^2 операций, всего элементов n^2

$$\underset{n^2\times n^2}{y=(A\otimes B)}\underset{n^2\times 1}{x} \; - \text{ вычисляется за } \; (\underbrace{n^2+n^2-1}_{(A\otimes B)_{(i)}x})\underbrace{n^2}_{\text{всего строк в } A\otimes B} = 2n^4-n^2$$

В итоге: $n^4 + 2n^4 - n^2 = 3n^4 - n^2 = O(n^4)$

(b) $(A\otimes B)x = vec(BXA^T)$, $X = x.reshape((n,n))$

Покажем равенство:

$$(A\otimes B)\,x = \underset{n^2\times n^2}{\begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & & \\ a_{n1}B & \dots & a_{nn}B \end{pmatrix}} \underset{n^2\times 1}{\begin{pmatrix} x_1 \\ \vdots \\ x_{n^2} \end{pmatrix}} = \left[\begin{pmatrix} x_1 \\ \vdots \\ x_{n^2} \end{pmatrix} = \begin{pmatrix} \boxed{\begin{matrix} x_{11} \\ \vdots \\ x_{n1} \end{matrix}} \\ \vdots \\ \boxed{\begin{matrix} x_{1n} \\ \vdots \\ x_{nn} \end{matrix}} \end{pmatrix} \right] =$$

$$= \begin{pmatrix} a_{11}BX^{(1)} + a_{12}BX^{(2)} + \dots + a_{1n}BX^{(n)} \\ \vdots \\ a_{n1}BX^{(1)} + a_{n2}BX^{(2)} + \dots + a_{nn}BX^{(n)} \end{pmatrix} \xrightarrow{reshape((n,n))} \left(\sum a_{1i}BX^{(i)} \;\dots\; \sum a_{ni}BX^{(i)} \right) =$$

$$= \begin{pmatrix} \sum a_{1i}B_{(1)}X^{(i)} & \dots & \sum a_{ni}B_{(1)}X^{(i)} \\ \vdots & & \vdots \\ \sum a_{1i}B_{(n)}X^{(i)} & \dots & \sum a_{ni}B_{(n)}X^{(i)} \end{pmatrix} = \underset{n\times n}{\begin{pmatrix} B_{(1)}X^{(1)} & \dots & B_{(1)}X^{(n)} \\ \vdots & & \\ B_{(n)}X^{(1)} & \dots & B_{(n)}X^{(n)} \end{pmatrix}} \underset{n\times n}{\begin{pmatrix} a_{11} & \dots & a_{n1} \\ a_{1n} & \dots & a_{nn} \end{pmatrix}} =$$

$$= \underset{n\times n}{\begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \\ b_{n1} & \dots & b_{nn} \end{pmatrix}} \underset{n\times n}{\begin{pmatrix} x_{11} & \dots & x_{n1} \\ \vdots & & \\ x_{1n} & \dots & x_{nn} \end{pmatrix}} \underset{n\times n}{\begin{pmatrix} a_{11} & \dots & a_{n1} \\ \vdots & & \\ a_{1n} & \dots & a_{nn} \end{pmatrix}} = BXA^T$$

Сколько операций требует подсчет BXA^T:

$B_{(i)}X^{(j)}$ занимает $2n-1$ операций $\Rightarrow \forall i,j: (2n-1)n^2 = 2n^3 - n^2$

$BX\cdot A^T$ аналогично $2n^3 - n^2$. В итоге $4n^3 - 2n^2 = O(n^3)$

(c) C - циркулянт $\Rightarrow$ $C = F_n^{-1}\,\mathrm{diag}(F_n c)\,F_n$, F_n - матрица Фурье

A циркулянт $\Rightarrow$ A^T - циркулянт.

Умножение BX: $BX^{(i)}$ - умножение вектора на цирк. - $O(n\log n)$

т.е. всего $n\,O(n\log n) = O(n^2\log n)$.

Пусть $BX = Y$ $YA^T = (AY^T)^T$: $A(Y^T)^{(i)}$ - так же за $O(n\log n)$ и всего $O(n^2\log n)$

Итоговая сложность получается $O(n^2\log n)$.