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.
   $H(V) = I - 2VV^T$
   суть преобразований состоит в последовательном умножении матрицы A на H_i - матрицы Хаусхолдера.
   Посмотрим на первое умножение:
   $HA = (I - 2VV^T)A = A - v(2v^T A)$
   посчитаем операции:
   $v^T A$ потребует $2mn - n$ операций
   умножение на 2 вектора потребует n операций, разность с матрицей A еще mn операций. Итого к этому моменту имеем $4mn$.
   На i-той итерации будет
   $H_i = \begin{pmatrix} I_k & 0 \\ 0 & H(V_k) \end{pmatrix}$
   потребуется $4(m-i)(n-i) = 4mn - 4ni - 4mi + 4i^2$
   Тогда итоговую сложность можно найти:
   $\sum_{i=0}^{n-1}(4mn - 4in - 4im + 4i^2) = 4mn^2 - 4n \cdot \binom{n}{2} - 4m \cdot \binom{n}{2} + 4 \cdot \frac{(n-1)\cdot(n-2)\cdot(2n-3)}{6} =$
   $2mn^2 - \frac{2n^3}{3} + O(mn)$

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

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

   $Q = H_1 \dots H_n \cdot \begin{pmatrix} I \\ 0 \end{pmatrix}$
   $\hat{H} = H \cdot \begin{pmatrix} I \\ 0 \end{pmatrix}$
   На i-той итерации $H_{n-i} = \begin{pmatrix} I_{n-i} & 0 \\ 0 & H(V_{n-i}) \end{pmatrix}$ уйдет $4(m-n+i)\cdot i$ из тех же соображений, что и в предыдущем пункте.
   Итого:
   $\sum_{i=1}^{n} 4mi - 4ni + 4i^2 = 2mn^2 - \frac{2n^3}{3} + O(mn)$

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:

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

$\|B(\mu)-A^+\|_2 = \|(A^TA+\mu I)^{-1}A^T - A^+\|_2 = \|(V\Sigma^T\Sigma V^T + \mu VV^T)^{-1}V\Sigma U^T - V\Sigma^+U^T\|_2 = \|V(\Sigma^2\mu I)^{-1}V^{-1}V\Sigma U^T - V\Sigma^+U^T\|_2 = \|V((\Sigma^2+\mu I)^{-1}\Sigma U^T - \Sigma^+U^T)\|_2 = \|(\Sigma^2+\mu I)^{-1}\Sigma - \Sigma^+\|_2$ - эта штука будет страшим сингулярным числом.

Обращать скобу можно, там определитель ненулевой.

$$Sigma^2 + \mu I = \begin{pmatrix} \sigma_1^2+\mu & \dots & 0 \\ \vdots & \dots & \vdots \\ 0 & \dots & \mu \end{pmatrix}$$

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

Тогда $\|(\Sigma^2+\mu I)^{-1}\Sigma - \Sigma^+\|_2 = \frac{\mu}{(\sigma_1^2+\mu)\sigma_r}$

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

$B(\mu) \to A^+$ as $\mu \to +0 \iff \lim\limits_{\mu \to x+0} \|B(\mu) - A^+\|_2 = 0$

$\lim\limits_{\mu \to x+0} \|B(\mu) - A^+\|_2 = \lim\limits_{\mu \to x+0} \frac{\mu}{(\sigma_r^2+\mu)\sigma_r} = 0$

$\lim\limits_{\mu \to x+0} \|B(\mu)b - A^+b\|_2 = \lim\limits_{\mu \to x+0} \|(B(\mu) - A^+)b\|_2 = \lim\limits_{\mu \to x+0} \|0b\|_2 = 0$

успех.

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 $\|x\|_2$.

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

тк из лекций знаем, что минимум второй нормы будет при $x = A^+b$

тогда $\exists t_1, t_2 \in \mathbb{R}^n : x = t_1 + t_2, \langle t_1, t_2 \rangle = 0$

$\sqsupset t_2 = (I - A^+A)y \in Ker(A)$

$A^+b = V\Sigma^+U^Tb = V(\Sigma^+U^Tb) \in Im(V) = Im(A^T)$

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

опять успех.

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.

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

$U = \begin{bmatrix} a & 3b \end{bmatrix}, V = \begin{bmatrix} a+2b & b \end{bmatrix}, W = \begin{bmatrix} a & a \end{bmatrix}$

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

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

$G = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$

$U = \begin{bmatrix} a & b \end{bmatrix}, V = \begin{bmatrix} a & b \end{bmatrix}, W = \begin{bmatrix} a \end{bmatrix}$

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

Можно узнать ранг из развертки;

$A_{(1)} = UG_{(1)}(W \otimes V)^T = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} a \otimes \begin{bmatrix} a & b \end{bmatrix} \end{bmatrix}^T$

$a \otimes \begin{bmatrix} a & b \end{bmatrix} = \begin{pmatrix} a_1 a & a_1 b \\ a_2 a & a_2 b \\ \vdots & \vdots \\ a_n a & a_n b \end{pmatrix}$

тогда

$rkA_{(1)} = 2$ тк ранги всех сомножителей равны двум

$A_{(2)} = VG_{(1)}(W \otimes U)^T = \begin{bmatrix} a & b \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \begin{bmatrix} a \otimes \begin{bmatrix} a & b \end{bmatrix} \end{bmatrix}^T$

$rkA_{(2)} = 2$ аналогично

$A_{(1)} = WG_{(1)}(V \otimes U)^T = \begin{bmatrix} a \end{bmatrix} \begin{bmatrix} 1023 \end{bmatrix} \begin{bmatrix} \begin{bmatrix} a & b \end{bmatrix} \otimes \begin{bmatrix} a & b \end{bmatrix} \end{bmatrix}^T$

$rkA_{(3)} = 1$ тк мы умножаем столбец на строку.

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 $(A \otimes B)$?

в $A \otimes B$ будет n^2 строк, на подсчет каждой в произведении с иксом тоже потребуется порядка n^2. Тк просят асимптотику, можно забить на константу.

Общая сложность будет $O(n^4)$

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

$(A \otimes B)x = \begin{pmatrix} a_{11}B & \dots & a_{1n}B \\ \vdots & \dots & \vdots \\ a_{n1}B & \dots & a_{nn}B \end{pmatrix} \begin{pmatrix} x_{11} \\ \vdots \\ x_n^2 \end{pmatrix} = \begin{pmatrix} B_{(1)}X^{(1)} & \dots & B_{(1)}X^{(n)} \\ \vdots & \dots & \vdots \\ B_{(n)}X^{(1)} & \dots & B_{(n)}X^{(n)} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \dots & \vdots \\ a_{1n} & \dots & a_{nn} \end{pmatrix} =$
$\begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & \dots & \vdots \\ b_{n1} & \dots & b_{nn} \end{pmatrix} \cdot \begin{pmatrix} x_{11} & \dots & x_{1n} \\ \vdots & \dots & \vdots \\ x_{n1} & \dots & x_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & x_{n1} \\ \vdots & \dots & \vdots \\ a_{1n} \dots & a_{nn} & \end{pmatrix} = BXA^T$

Считаем сложность. $\forall i, j B_{(i)}X^{(j)}$ требует $2n-1$ операций для каждого из n^2 умножений. BXA^T из аналогичных соображений требует столько же в асимптотике, а именно $O(n^3)$

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

$BXA^T = (A(BX)^T)^T$

Теперь заметим, что A^T будет циркулянтом, тк A такова.

$(A(BX)^T)^{(i)}$ можно посчитать за $O(n \log n)$ тк мы умножаем вектор на циркулянт.

Таких умножений требуется n штук, итого имеем сложность $O(n^2 \log n)$

Бонусные задачи

1. **(40 б. балла)**. Пусть $A \in \mathbb{C}^{m \times n}$, $m \geq n$ – матрица полного ранга. Верно ли, что в ее тонком QR-разложении все r_{kk}, $k = 1, \dots, n$ (диагональные элементы R) всегда можно выбрать вещественными положительными? Пусть $r_{kk} > 0$, $k = 1, \dots, n$. Покажите, что в таком случае разложение определяется единственным образом.

2. **(60 б. балла)**. Докажите или опровергните, что

$$\|X^+X - Y^+Y\|_F \le \left(\|X^+\|_2 + \|Y^+\|_2\right) \|X - Y\|_F,$$

для любых комплексных матриц X и Y одного размера.
Подсказка: подумайте над тем, что можно добавить и вычесть под знаком нормы.