HW5

1. **(15 баллов)**. Посчитайте (аналитически) LU-разложение матрицы

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

с помощью элементарных преобразований Z_k. Объясните, почему в данном случае существование LU-разложения не противоречит теореме о существовании LU-разложения из лекций.

Знаем, что шаг метода Гаусса выглядит как

$$\underbrace{\begin{pmatrix} 1 & 0 \\ -\frac{1}{a}b & I \end{pmatrix}}_{= Z_1} \underbrace{\begin{pmatrix} a & c^T \\ b & D \end{pmatrix}}_{= A} = \begin{pmatrix} a & c^T \\ 0 & D - \frac{1}{a} b c^T \end{pmatrix} = Z_1 A$$

В общем случае Z_i выглядит как
$$\begin{array}{c} \quad\quad i \\ \begin{pmatrix} 1 & & & & \\ & 1 & & & 0 \\ & & 1 & & \\ 0 & & \left[-\frac{1}{a}b\right] & 1 & \\ & & 0 & & 1 \end{pmatrix} \end{array}$$

Начнем:

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix}}_{Z_1} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}}_{A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix} \; Z_1 A$$

$$\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}}_{Z_2} \underbrace{\begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & -6 & -12 \end{pmatrix}}_{Z_1 A} = \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix} \; Z_2 Z_1 A = U$$

$Z_2 Z_1 A = U$

$A = \underbrace{Z_1^{-1} Z_2^{-1}}_{L} U$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ -4 & 1 & 0 \\ -7 & 0 & 1 \end{pmatrix}^{-1} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} =$$

$$= \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix}$$

$$A = LU = \begin{pmatrix} 1 & 0 & 0 \\ 4 & 1 & 0 \\ 7 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 0 & -3 & -6 \\ 0 & 0 & 0 \end{pmatrix}$$

Теорема в лекции выглядела как

Th Невырожд. A имеет LU разлож. <=> A строго регулярна

В нашем случае матрица вырождена, можем это заметить здесь.
Получается, что наша матрица не подпадает под условия теоремы из лекции.

2. **(35 баллов)**. Пусть симметричная положительно определенная матрица $A \in \mathbb{R}^{n \times n}$ задана в следующем виде:

$$A = \begin{bmatrix} a & c^\top \\ c & D \end{bmatrix},$$

где $a \in \mathbb{R}$, $c \in \mathbb{R}^{n-1}$, $D \in \mathbb{R}^{(n-1)\times(n-1)}$. Исключением Холецкого первой строки назовем следующую операцию:

$$A = \begin{bmatrix} a & c^\top \\ c & D \end{bmatrix} = \begin{bmatrix} l & 0 \\ \frac{c}{l} & I \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & D - \frac{cc^\top}{a} \end{bmatrix} \begin{bmatrix} l & \frac{c^\top}{l} \\ 0 & I \end{bmatrix} = L_0 A_1 L_0^\top, \quad l = \sqrt{a}. \tag{1}$$

Применив разложение аналогичное (1) к дополнению по Шуру $D - \frac{cc^\top}{a}$ и т.д., получим разложение Холецкого.

(a) **(9 баллов)**. Докажите, что $D - \frac{cc^\top}{a}$ будет симметричной положительно определенной.

(b) (4 баллов). Для матрицы

Симметричность

Мы знаем, что матрица D симметрична, т.к. это угловой блок в изначальной симметричной матрице. Матрица $\frac{cc^T}{a}$ тоже симметрична, т.к. $c_i \cdot c_j = c_j \cdot c_i$, что в матрице будет $(cc^T)_{ij} = (cc^T)_{ji}$. Разность симметричных матриц - тоже симметричная матрица.

Положительная определенность

Пусть $D - \frac{cc^T}{a} = B$. Хотим доказать, что $v^T B v > 0 \ \forall v$

Знаем, что $v^T A v > 0 \ \forall v$

$$0 < v^T A v = \underbrace{v^T L_0}_{v'} A_1 \underbrace{L_0^T v}_{v'} = [v'_1 \dots v'_n] \begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} v'_1 \\ \vdots \\ v'_n \end{bmatrix}$$

L_0 - невырожд. матрица, на ее строки натянуто пространство. Значит с помощью $v^T L_0$ мы переберем все векторы из этого пространства.

По критерию Сильвестра все миноры у $\begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$ невырождены. Это значит, что все миноры у B тоже больше 0: $0 < \delta_i = 1 \cdot \det B_i$

$$\begin{bmatrix} 1 & 0 \\ 0 & B \end{bmatrix}$$

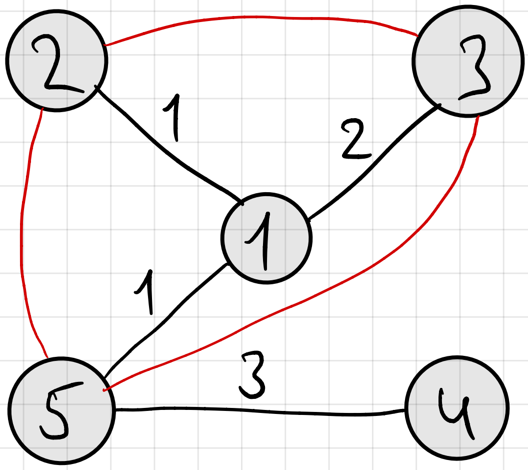
По критерию Сильвестра $D - \frac{cc^T}{a}$ положительно определена.

(b) **(4 балла)**. Для матрицы

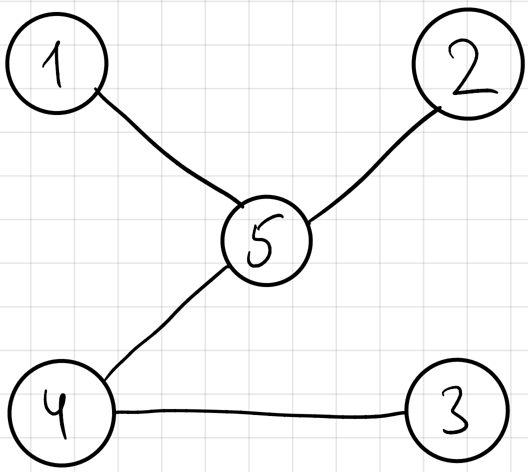
$$A = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}\begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} 4 & 1 & 2 & 0 & 1 \\ 1 & 4 & 0 & 0 & 0 \\ 2 & 0 & 5 & 0 & 0 \\ 0 & 0 & 0 & 1 & 3 \\ 1 & 0 & 0 & 3 & 4 \end{bmatrix} \end{array} \tag{2}$$

постройте ее граф G(A).

(c) **(8 баллов)**. Нарисуйте на графе $G(A)$ (выделите цветом) возникающие заполнения у дополнения по Шуру $D - \frac{cc^\top}{a}$ и продолжите эту процедуру рекурсивно. Находить конкретные числа не обязательно.



(d) **(8 баллов)**. Примените к графу G(A) алгоритм minimal degree ordering (вычислять разложение для матрицы не нужно, этот пункт подразумевает работу только с графом) и соответствующую матрицу $PAP^\top$. На сколько сократилось число ребер заполнения при исключении вершин в новом порядке?



Количество вершин сократилось на 3

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$A' = PAP^T = \begin{pmatrix} 4 & 0 & 0 & 0 & 1 \\ 0 & 5 & 0 & 0 & 2 \\ 0 & 0 & 1 & 3 & 0 \\ 0 & 0 & 3 & 4 & 1 \\ 1 & 2 & 0 & 1 & 4 \end{pmatrix}$$

вершин в новом порядке?

(e) **(6 баллов)**. Запишите матрицу A из (2) в CSR формате.

val = [4, 1, 2, 1, 1, 4, 2, 5, 1, 3, 1, 3, 4]

row = [0, 4, 6, 8, 10, 12]

column = [0, 1, 2, 4, 0, 1, 0, 2, 3, 4, 0, 3, 4]

3. (25 баллов).

(a) Пусть $A \in \mathbb{R}^{n\times n}$ – произвольная невырожденная матрица. Пусть рассматривается итерационный процесс вида $x_{k+1} = x_k + \tau_k r_k$, $r_k = b - Ax_k$. Получите оптимальные параметры $\tau_k \in \mathbb{R}$, минимизирующие функционал невязки $J(x) = \|b - Ax\|_2^2$ на каждом шаге итерационного процесса. Убедитесь, что полученное выражение зависит только от A и r_k.

(b) В случае $A = A^\top > 0$ покажите, что для полученного процесса выполняется:

$$\|r_k\|_2 \le \left(\frac{\mathrm{cond}_2(A) - 1}{\mathrm{cond}_2(A) + 1}\right)^k \|r_0\|_2.$$

a) $J(x_{k+1}) = \|b - Ax_{k+1}\|_2^2 = \|b - A(x_k + \tau_k r_k)\|^2 = \|\underbrace{b - Ax_k}_{r_k} - A\tau_k r_k\|^2 =$

$= \|r_k - A\tau_k r_k\|^2 = (r_k - A\tau_k r_k)^T(r_k - A\tau_k r_k) = r_k^T r_k - r_k^T A\tau_k r_k - (A\tau_k r_k)^T r_k +$

$+ (A\tau_k r_k)^T A\tau_k r_k = \|r_k\|^2 - r_k^T A\tau_k r_k - r_k^T\tau_k A^T r_k + \|A\tau_k r_k\|^2 =$

$= \|r_k\|^2 - r_k^T(A\tau_k + \tau_k A^T)r_k + \|A\tau_k r_k\|^2 = \|r_k\|^2 - \tau_k r_k^T(A + A^T)r_k + \|A\tau_k r_k\|^2 \to \min\limits_{\tau_k}$

$-\tau_k r_k^T(A + A^T)r_k + \tau_k^2\|Ar_k\|^2 \to \min\limits_{\tau_k}$

$\frac{\partial}{\partial \tau_k}(\ldots) = -r_k^T(A + A^T)r_k + 2\tau_k\|Ar_k\|^2 = 0$

$$\tau_k = \frac{r_k^T(A + A^T)r_k}{2\|Ar_k\|^2}$$

b) $\|r_{k+1}\|_2 \overset{\text{из пункта а)}}{=} \|r_k - A\tau_k r_k\|_2 = \|(I - A\tau_k)r_k\|_2 \le \|I - A\tau_k\|_2\|r_k\|_2 = \left\{\tau_k = \frac{2}{\lambda_1 + \lambda_n}\right\} =$

$= \max_i \frac{|\lambda_1 + \lambda_n - 2\lambda_i|}{\lambda_1 + \lambda_n}\|r_k\|_2 = \frac{\lambda_1 - \lambda_n}{\lambda_1 + \lambda_n}\|r_k\|_2 = \frac{\frac{\lambda_1}{\lambda_n} - 1}{\frac{\lambda_1}{\lambda_n} + 1}\|r_k\|_2 = \frac{\mathrm{cond}(A) - 1}{\mathrm{cond}(A) + 1}\|r_k\|_2 =$

$= \{\text{рекурсивно}\} = \left(\frac{\mathrm{cond}(A) - 1}{\mathrm{cond}(A) + 1}\right)^k \|r_0\|_2$

4. **(25 баллов)**. Покажите, что для достижения точности $\frac{\|e_k\|_2}{\|e_0\|_2} \le \varepsilon$ для заданного $0 < \varepsilon < 1$ в методе Чебышева необходимо сделать не больше

$$N(\varepsilon) = 1 + \frac{\sqrt{\mathrm{cond}_2(A)}}{2}\ln(2\varepsilon^{-1})$$

итераций (при достаточно большом $\mathrm{cond}_2(A)$).

$$\|e_k\|_2 \le 2\left(\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1}\right)^k \|e_0\|_2$$

$$\frac{\|e_k\|_2}{\|e_0\|_2} = 2\left(\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1}\right)^k \le \varepsilon$$

$$\ln 2 + k\ln\left(\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1}\right) \le \ln\varepsilon$$

$$k\ln\left(\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1}\right) \le \ln\left(\frac{\varepsilon}{2}\right)$$

$$k \ge \frac{\ln\left(\frac{\varepsilon}{2}\right)}{\ln\left(\frac{\sqrt{\mathrm{cond}_2(A)}-1}{\sqrt{\mathrm{cond}_2(A)}+1}\right)} = \frac{\ln\left(\frac{\varepsilon}{2}\right)}{\ln\left(1-\frac{2}{\sqrt{\mathrm{cond}(A)}+1}\right)} =$$

$$= \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln\left(\frac{\sqrt{\mathrm{cond}_2(A)}+1}{\sqrt{\mathrm{cond}_2(A)}-1}\right)} = \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\ln\left(1+\frac{2}{\sqrt{\mathrm{cond}_2(A)}-1}\right)} \approx \frac{\ln\left(\frac{2}{\varepsilon}\right)}{\frac{2}{\sqrt{\mathrm{cond}_2(A)}-1}} = \frac{\sqrt{\mathrm{cond}_2(A)}-1}{2}\ln\left(\frac{2}{\varepsilon}\right) \quad ♡$$

$\ln(1+x) \sim x$
при $x \to 0$

Кол-во итераций должно быть больше или равно ♡, $N(\varepsilon) = 1 + \frac{\sqrt{\mathrm{cond}_2(A)}}{2}\ln(2\varepsilon^{-1})$ подходит под это условие.