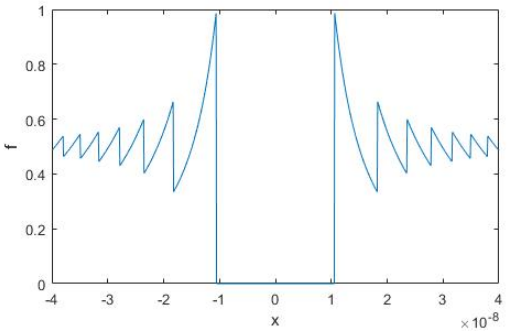
Numerical methods, midterm test I (2019/20 autumn, group A)

PROBLEM 1. (6p)

We have plotted the graph of the function $f(x) = (1 - \cos x)/x^2$ with Matlab on the interval $[-4 \times 10^{-8}, 4 \times 10^{-8}]$. The obtained graph can be seen in the figure. Let us explain the form of the graph and suggest a different form of the function that results in the correct graph. (Remember that $\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha$.)



Solution. The values of the cosine function are close to 1 for x values close to zero. The floating point representations of these values can be written in the form $fl(\cos(x)) = 1 - k\varepsilon_m/2$, $k = 0, 1, 2, \ldots$. This follows from the fact that the previous exactly representable number before 1 has the form $1 - \varepsilon_m/2$ ($1.111\ldots11 \times 2^{-1}$ in binary form), where $\varepsilon_m \approx 2.2 \times 10^{-16}$ is Matlab's machine epsilon. Thus Matlab will graph the values

$$\frac{k\varepsilon_m/2}{x^2}.$$

For x values for which $fl(\cos(x)) = 1$, the function value will be 0, for x values for which $fl(\cos(x)) = 1 - \varepsilon_m/2$, the function value will be $\varepsilon_m/2/x^2$, for x values for which $fl(\cos(x)) = 1 - 2\varepsilon_m/2$, the function value will be $2\varepsilon_m/2/x^2$, etc. This explains the zigzagged form of the graph.

We can avoid the subtraction in the numerator with the given trigonometric identity

$$\frac{1 - \cos x}{x^2} = \frac{\cos^2(x/2) + \sin^2(x/2) - (\cos^2(x/2) - \sin^2(x/2))}{x^2} = 2\frac{\sin^2(x/2)}{x^2}$$

(which gives function values 1/2 in the given interval).

PROBLEM 2. (6p) Let us decide whether a linear system with the coefficient matrix `A=-ones(6)+10*eye(6)` (9s in the diagonal, other elements are -1s) can be solved by the following methods: a) Cholesky, b) Gauss, c) relaxed Jacobi with $\omega = 1/2$, d) relaxed Gauss–Seidel with $\omega = 3$, e) conjugate gradient method. (We may use the known fact that symmetric M-matrices are positive definite matrices.)

Solution. The matrix is an M-matrix ($g = [1, 1, 1, 1, 1, 1]^T$ can be a majorizing vector). Moreover, since it is symmetric, it is also positive definite (SPD).

a) Cholesky: can be used because the matrix is SPD,

b) Gauss: can be used for M-matrices and SPD matrices,

c) relaxed Jacobi with $\omega = 1/2$: can be used for M-matrices with $\omega \in (0, 1]$,

d) relaxed Gauss–Seidel with $\omega = 3$: this method can be used only with $\omega \in (0, 2)$ for SPD matrices.

e) conjugate gradient method: can be used for SPD matrices.

PROBLEM 3. We solve the linear system $Ax = b$, where A is the matrix from the previous problem and b is a random vector with elements between 0.5 and 1. The solution of the system is denoted by x_1. Then we modify the system as follows. We add 0.001 to each element of the matrix A and subtract 0.01 from each element of b. The solution of the new system is denoted by x_2. Give an upper estimation for the relative error $\|x_1 - x_2\|_\infty / \|x_1\|_\infty$!

Solution. The maximum norm of the inverse matrix can be estimated using the learnt upper bound for M-matrices. We obtain that $\|A^{-1}\|_\infty \le 1/4$ (use the majorizing vector in the previous problem). Thus $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 14 \cdot 1/4$. Moreover we need the lower bound $\|b\|_\infty \ge 0.5$. Then (because $\|A^{-1}\|_\infty \cdot \|\delta A\|_\infty \le (1/4) \cdot 0.006 < 1$)

$$\frac{\|x_1 - x_2\|_\infty}{\|x_1\|_\infty} \le \frac{7/2}{1 - (7/2) \cdot (0.006/14)} \left(\frac{0.006}{14} + \frac{0.01}{0.5} \right) = 0.0716.$$

PROBLEM 4. (7p) Let us give the LU and Cholesky decompositions of the matrix (if they exist) and compute the value of $\det(A)$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 5 & 3 & 3 \\ 1 & 3 & 6 & 4 \\ 1 & 3 & 4 & 4 \end{bmatrix}!$$

Solution. LU decomposition can be obtained by the Gaussian elimination method.

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1/2 & 1 & 0 \\ 1 & 1/2 & 1/2 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 4 & 2 & 2 \\ 0 & 0 & 4 & 2 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

In order to get the G matrix of the Cholesky factorization we have to multiply the ith column of the matrix L by the value $\sqrt{u_{ii}}$, respectively. Thus the Cholesky factorization is $A = GG^T$, where

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

$\det(A) = \det(L)\det(U) = 1 \cdot 16 = 16$.

PROBLEM 5. (7p) Use the relaxed Jacobi method with relaxation parameter 1/2 to solve the linear system $\begin{aligned} 6x_1 - x_2 &= 1 \\ -x_1 + 3x_2 - x_3 &= 2 \\ -x_2 + 2x_3 &= 1. \end{aligned}$ Construct the iteration and estimate the number of iterations needed to approximate the exact solution of the system within the absolute error tolerance 10^{-6} in maximum norm. We start the iteration from the zero vector.

Solution. The form of the relaxed Jacobi method with $\omega = 1/2$ is

$$x_{k+1} = Bx_k + f = \begin{bmatrix} 1/2 & 1/12 & 0 \\ 1/6 & 1/2 & 1/6 \\ 0 & 1/4 & 1/2 \end{bmatrix} x_k + \begin{bmatrix} 1/12 \\ 1/3 \\ 1/4 \end{bmatrix}.$$

Because $\|B\|_\infty = 5/6$, $x_1 = f$ and $\|f\|_\infty = 1/3$, we have to solve the inequality

$$|x_k - x^\star| \le \frac{(5/6)^k}{1 - (5/6)} \cdot \frac{1}{3} \le 10^{-6}.$$

This shows that $k \ge 80$ steps are enough to obtain the solution within the given tolerance.

PROBLEM 6. (7p) Somebody has carried out three iteration steps for the system $Ax = b$, where A=[1,0,1;0,2,3;1,3,10] and b=[0;1;0], either with the gradient method or with the

conjugate gradient method. The obtained iteration vector is $x_3 = [3/38, 14/19, -3/20]^T$ and the residual is $r_3 = [27/380, -9/380, -300/380]^T$. Let us decide which method was used and calculate the next iteration step.

Solution. The method is the gradient method because the conjugate gradient method should terminate after the third step for a 3×3 system. Now this is not the case because r_3 is not zero.

$$\alpha_4 = \frac{781}{7748} = 0.1008, \quad x_4 = \begin{bmatrix} 442/5133 \\ 1441/1962 \\ -371/1616 \end{bmatrix} = \begin{bmatrix} 0.0861 \\ 0.7342 \\ -0.2296 \end{bmatrix}.$$

Numerical methods, midterm test I (2019/20 autumn, group B)

PROBLEM 1. (6p) Let us decide whether a linear system with the coefficient matrix `A=-ones(6)+11*eye(6)` (10s in the diagonal, other elements are -1s) can be solved by the following methods: a) Gauss, b) Cholesky, c) relaxed Gauss–Seidel with $\omega = 4$, d) relaxed Jacobi with $\omega = 1/3$, e) conjugate gradient method. (We may use the known fact that symmetric M-matrices are positive definite matrices.)

Solution. The matrix is an M-matrix ($g = [1, 1, 1, 1, 1, 1]^T$ can be a majorizing vector). Moreover, since it is symmetric, it is also positive definite (SPD).

a) Gauss: can be used for M-matrices and SPD matrices,

b) Cholesky: can be used because the matrix is SPD,

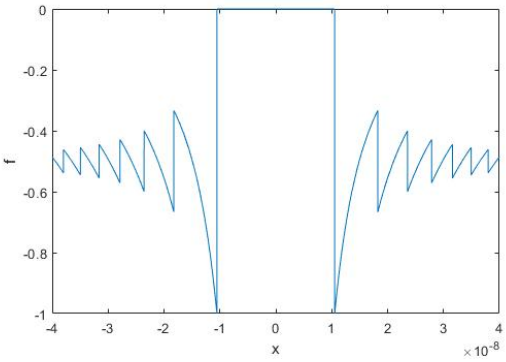
c) relaxed Gauss–Seidel with $\omega = 4$: this method can be used only with $\omega \in (0, 2)$ for SPD matrices,

d) relaxed Jacobi with $\omega = 1/3$: can be used for M-matrices with $\omega \in (0, 1]$,

e) conjugate gradient method: can be used for SPD matrices.

PROBLEM 2. (6p)

We have plotted the graph of the function $f(x) = (\cos x - 1)/x^2$ with Matlab on the interval $[-4 \times 10^{-8}, 4 \times 10^{-8}]$. The obtained graph can be seen in the figure. Let us explain the form of the graph and suggest a different form of the function that results in the correct graph. (Remember that $\cos 2\alpha = \cos^2 \alpha - \sin^2 \alpha$.)



Solution. The values of the cosine function are close to 1 for x values close to zero. The floating point representations of these values can be written in the form $fl(\cos(x)) = 1 - k\varepsilon_m/2$, $k = 0, 1, 2, \ldots$. This follows from the fact that the previous exactly representable number before 1 has the form $1 - \varepsilon_m/2$ ($1.111\ldots11 \times 2^{-1}$ in binary form), where $\varepsilon_m \approx 2.2 \times 10^{-16}$ is Matlab's machine epsilon. Thus Matlab will graph the values

$$-\frac{k\varepsilon_m/2}{x^2}.$$

For x values for which $fl(\cos(x)) = 1$, the function value will be 0, for x values for which $fl(\cos(x)) = 1 - \varepsilon_m/2$, the function value will be $-\varepsilon_m/2/x^2$, for x values for which $fl(\cos(x)) = 1 - 2\varepsilon_m/2$, the function value will be $-2\varepsilon_m/2/x^2$, etc. This explains the zigzagged form of the graph.

We can avoid the subtraction in the numerator with the given trigonometric identity

$$\frac{\cos x - 1}{x^2} = \frac{\cos^2(x/2) - \sin^2(x/2) - (\cos^2(x/2) + \sin^2(x/2))}{x^2} = -2\frac{\sin^2(x/2)}{x^2}$$

(which gives function values -1/2 in the given interval).

PROBLEM 3. (7p) We solve the linear system $Ax = b$, where A is the matrix from Problem 1 and b is a random vector with elements between 0.3 and 1. The solution of the system is denoted by x_1. Then we modify the system as follows. We subtract 0.01 from each element of the matrix A and add 0.001 to each element of b. The solution of the new system is denoted by x_2. Give an upper estimation for the relative error $\|x_1 - x_2\|_\infty / \|x_1\|_\infty$!

Solution. The maximum norm of the inverse matrix can be estimated using the learnt upper bound for M-matrices. We obtain that $\|A^{-1}\|_\infty \le 1/5$ (use the majorizing vector in Problem 1). Thus $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty = 15 \cdot 1/5 = 3$. Moreover we need the lower bound $\|b\|_\infty \ge 0.3$. Then (because $\|A^{-1}\|_\infty \cdot \|\delta A\|_\infty \le 1/5 \cdot 0.06 < 1$)

$$\frac{\|x_1 - x_2\|_\infty}{\|x_1\|_\infty} \le \frac{3}{1 - 3 \cdot (0.06/15)} \left(\frac{0.06}{15} + \frac{0.001}{0.3} \right) = 0.0223.$$

Problem 4. (7p) Let us give the LU and Cholesky decompositions of the matrix (if they exist) and compute the value of $\det(A)$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 5 & 3 & 3 \\ 1 & 3 & 3 & 3 \\ 1 & 3 & 3 & 7 \end{bmatrix}!$$

Solution. LU decomposition can be obtained by the Gaussian elimination method.

$$L = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1/2 & 1 & 0 \\ 1 & 1/2 & 1 & 1 \end{bmatrix}, \quad U = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 4 & 2 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

In order to get the G matrix of the Cholesky factorization we have to multiply the ith column of the matrix L by the value $\sqrt{u_{ii}}$, respectively. Thus the Cholesky factorization is $A = GG^T$, where

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

$\det(A) = \det(L)\det(U) = 1 \cdot 16 = 16$.

Problem 5. (7p) Use the relaxed Jacobi method with relaxation parameter 1/3 to solve the linear system $\begin{aligned} 5x_1 - x_2 &= 1 \\ -x_1 + 4x_2 - x_3 &= 3 \\ -x_2 + 2x_3 &= 1. \end{aligned}$ Construct the iteration and estimate the number of iterations needed to approximate the exact solution of the system within the absolute error tolerance 10^{-4} in maximum norm. We start the iteration from the zero vector.

Solution. The form of the relaxed Jacobi method with $\omega = 1/2$ is

$$x_{k+1} = Bx_k + f = \begin{bmatrix} 2/3 & 1/15 & 0 \\ 1/12 & 2/3 & 1/12 \\ 0 & 1/6 & 2/3 \end{bmatrix} x_k + \begin{bmatrix} 1/15 \\ 1/4 \\ 1/6 \end{bmatrix}.$$

Because $\|B\|_\infty = 5/6$, $x_1 = f$ and $\|f\|_\infty = 1/4$, we have to solve the inequality

$$|x_k - x^\star| \le \frac{(5/6)^k}{1 - (5/6)} \cdot \frac{1}{4} \le 10^{-4}.$$

This shows that $k \ge 53$ steps are enough to obtain the solution within the given tolerance.

Problem 6. (7p) Somebody has carried out three iteration steps for the system $Ax = b$, where A=[1,0,1;0,2,2;1,2,10] and b=[0;1;0], either with the gradient method or with the

conjugate gradient method. The obtained iteration vector is $x_3 = [1/18, 11/18, -1/10]^T$ and the residual is $r_3 = [2/45, -1/45, -5/18]^T$. Let us decide which method was used and calculate the next iteration step.

Solution. The method is the gradient method because the conjugate gradient method should terminate after the third step for a 3×3 system. Now this is not the case because r_3 is not zero.

$$\alpha_4 = \frac{645}{6274} = 0.1028, \quad x_4 = \begin{bmatrix} 299/4973 \\ 607/997 \\ -253/1968 \end{bmatrix} = \begin{bmatrix} 0.0601 \\ 0.6088 \\ -0.1286 \end{bmatrix}.$$