Numerical methods, midterm test I (2017/18, autumn semester) Solutions

1. (6p) Let $\|.\|$ denote an arbitrary vector norm or the matrix norm induced by this vector norm. Prove the following statement for a quadratic matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$: If there exists a number $\alpha > 0$ such that $\|\mathbf{A}\overline{\mathbf{x}}\| \geq \alpha\|\overline{\mathbf{x}}\|$ for all vectors $\overline{\mathbf{x}} \in \mathbb{R}^n$, then $\mathbf{A}$ is non-singular and the estimate $\|\mathbf{A}^{-1}\| \leq 1/\alpha$ is valid!

Solution: If $\overline{\mathbf{x}}$ is a non-zero vector, then $\mathbf{A}\overline{\mathbf{x}} \neq \mathbf{0}$. This can be seen as follows. If $\mathbf{A}\overline{\mathbf{x}}$ was zero then we would have $0 = \|\mathbf{A}\overline{\mathbf{x}}\| \geq \alpha\|\overline{\mathbf{x}}\| > 0$, which is a contradiction. Thus zero is not an eigenvalue of $\mathbf{A}$, that is the matrix is non-singular.

Applying the above estimation, we have

$$\|\mathbf{A}^{-1}\| = \sup_{\overline{\mathbf{x}}\neq 0} \frac{\|\mathbf{A}^{-1}\overline{\mathbf{x}}\|}{\|\overline{\mathbf{x}}\|} = \sup_{\overline{\mathbf{y}}\neq 0} \frac{\|\mathbf{A}^{-1}\mathbf{A}\overline{\mathbf{y}}\|}{\|\mathbf{A}\overline{\mathbf{y}}\|} = \sup_{\overline{\mathbf{y}}\neq 0} \frac{\|\overline{\mathbf{y}}\|}{\|\mathbf{A}\overline{\mathbf{y}}\|} \leq \sup_{\overline{\mathbf{y}}\neq 0} \frac{(1/\alpha)\|\mathbf{A}\overline{\mathbf{y}}\|}{\|\mathbf{A}\overline{\mathbf{y}}\|} = \frac{1}{\alpha}.$$

The last estimation can be obtained also as follows. Let us apply the inequality of the problem to the vector $\mathbf{A}^{-1}\overline{\mathbf{x}}$: $\|\overline{\mathbf{x}}\| = \|\mathbf{A}\mathbf{A}^{-1}\overline{\mathbf{x}}\| \geq \alpha\|\mathbf{A}^{-1}\overline{\mathbf{x}}\|$. Then $1/\alpha \geq \|\mathbf{A}^{-1}\overline{\mathbf{x}}\|/\|\overline{\mathbf{x}}\|$, which shows the statement, since the maximum of the expression on the right hand side is $\|\mathbf{A}^{-1}\|$.

2. (6p) We would like to compute the value $f(x) = e^x - x - 1$ for $x_0 = 0.0005$ as accurately as we can. We use a computer that uses decimal numbers with 8-digit-long mantissas. Let us apply first the approximation $e^{0.0005} \approx 1.0005001$! Give a better approximation of $f(x_0)$ on the same computer!

Solution: With the first approximation we have $f(x_0) \approx 1.0005001 - 1 - 0.0005 = 1 \times 10^{-7}$. We have to give better estimation to the value of the expression than this. We have to avoid the subtraction of close numbers. Using the Taylor's expansion of the function e^x, we get that

$$f(x_0) = \frac{x_0^2}{2!} + \frac{x_0^3}{3!} + \ldots.$$

$x_0^2/2$ can be computed more accurately on the given computer: $x_0^2/2 = 1.25\times 10^{-7}$. Because this value is larger than the value obtained above, and the exact value is larger than this (we have omitted only positive terms from the series), this approximation will be closer the to exact value than the previous one.

Naturally, if we took more terms from the series then we would get better approximation. The best approximation on the given computer would be 1.2502084×10^{-7}. This is the exact function value rounded to 8-digit-long mantissa.

3. Show that if M is an M-matrix, then $\mathbf{AD}$ is also an M-matrix for all diagonal matrices $\mathbf{D}$ with positive diagonal. Using this statement, give an upper estimation for the maximum norm of the inverse of the matrix

$$\mathbf{B} = \begin{bmatrix} 9 & -1 & 0 \\ -3 & 3 & -3 \\ 0 & -1 & 9 \end{bmatrix} = \begin{bmatrix} 3\cdot 3 & -1 & 3\cdot 0 \\ 3\cdot(-1) & 3 & 3\cdot(-1) \\ 3\cdot 0 & -1 & 3\cdot 3 \end{bmatrix}.$$

Solution: Trivially we have offdiag$(\mathbf{AD}) \leq 0$. Thus, it is enough to show that there exists a positive vector $\overline{\mathbf{p}}$ such that $\mathbf{AD}\overline{\mathbf{p}}$ is positive. Because $\mathbf{A}$ is an M-matrix, there exists a vector $\overline{\mathbf{g}} > 0$ such that $\mathbf{A}\overline{\mathbf{g}} > 0$, thus $\overline{\mathbf{p}} = \mathbf{D}^{-1}\overline{\mathbf{g}}$ is a good choice since this is a positive vector, moreover $\mathbf{AD}\overline{\mathbf{p}} = \mathbf{ADD}^{-1}\overline{\mathbf{g}} = \mathbf{A}\overline{\mathbf{g}} > 0$. (Other method: $(\mathbf{AD})^{-1} = \mathbf{D}^{-1}\mathbf{A}^{-1}$, $\mathbf{A}$ is invertible because it is an M-matrix, $\mathbf{D}$ is also invertible because it is a diagonal matrix with positive diagonal. Both inverses are nonnegative thus their product is also nonnegative.)

Matrix $\mathbf{B}$ can be written in the form :

$$\mathbf{B} = \begin{bmatrix} 9 & -1 & 0 \\ -3 & 3 & -3 \\ 0 & -1 & 9 \end{bmatrix} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 3 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3 \end{bmatrix} =: \mathbf{AD},$$

where $\mathbf{A}$ is an M-matrix, since its offdiagonal is non-positive, and the all-one vector is a majorizing vecto. According to the known estimation we have

$$\|\mathbf{B}^{-1}\|_\infty \le \frac{\|\overline{\mathbf{p}}\|_\infty}{\min(\mathbf{B}\overline{\mathbf{p}})_i} = \frac{\|\mathbf{D}^{-1}\overline{\mathbf{g}}\|_\infty}{\min(\mathbf{A}\overline{\mathbf{g}})_i} = 1.$$

4. We would like to solve the linear system $\mathbf{A}\overline{\mathbf{x}} = \overline{\mathbf{b}}$, where $\mathbf{A} \in \mathbb{R}^{5\times 5}$ and $\overline{\mathbf{b}} = [1,2,3,4,5]^T$. We know that $\|\mathbf{A}^{-1}\|_\infty = 4$ and that the LU decomposition of $\mathbf{A}$ has the concise form

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

Compute the solution $\overline{\mathbf{x}}$ of the system. How much would the solution change in maximum norm if we changed the vector $\overline{\mathbf{b}}$ to the vector $\overline{\mathbf{b}}' = [0.99, 1.99, 3.05, 4.02, 5.1]^T$? (Hint: Do not compute matrix $\mathbf{A}$ explicitly! Let us estimate the unknown quantities.)

Solution: The LU decomposition of the matrix is

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

We get with simple back substitution that $\mathbf{U}\overline{\mathbf{x}} = [1,1,1,1,1]^T$, and with a back substitution again we obtain $\overline{\mathbf{x}} = [1/2, 1, 3/2, 0, 1/2]^T$.

Now matrix $\mathbf{A}$ does not change mátrix nem változik. Thus using the known error estimation formula we have

$$\|\delta\overline{\mathbf{x}}\|_\infty \le \|\overline{\mathbf{x}}\|_\infty \kappa_\infty(\mathbf{A}) \frac{\|\delta\overline{\mathbf{b}}\|_\infty}{\|\overline{\mathbf{b}}\|_\infty} = \|\overline{\mathbf{x}}\|_\infty \|\mathbf{A}\|_\infty \|\mathbf{A}^{-1}\|_\infty \frac{\|\delta\overline{\mathbf{b}}\|_\infty}{\|\overline{\mathbf{b}}\|_\infty} \le$$

$$\le \|\overline{\mathbf{x}}\|_\infty \|\mathbf{L}\|_\infty \|\mathbf{U}\|_\infty \|\mathbf{A}^{-1}\|_\infty \frac{\|\delta\overline{\mathbf{b}}\|_\infty}{\|\overline{\mathbf{b}}\|_\infty} = 3/2 \cdot 5 \cdot 3 \cdot 4 \cdot \frac{0.1}{5} = 1.8.$$

5. We solve system
$$\begin{aligned} 5x_1 - x_2 &= 7 \\ -x_1 + 3x_2 - x_3 &= 4 \\ -x_2 + 2x_3 &= 5 \end{aligned}$$
with an iterative solver. Show that the Gauss–Seidel method is applicable to this system. We start the iteration from the vector $\overline{\mathbf{x}}^{(0)} = [0,0,0]^T$, and we get the vector $\overline{\mathbf{x}}^{(1)} = [1.4000, 1.8000, 3.4000]^T$ in the first step (rounded to 4 decimal places). How many iterations does the iteration need to approximate the solution with a maximum error of 5×10^{-6} in 1-norm (the inverse of a lower triangular matrix is lower triangular!).

Solution: The Gauss–Seidel method is applicable because the coefficient matrix is strictly diagonally dominant (or M-matrix, or SPD).

We need to compute only the iteration matrix $\mathbf{B}$ ($\overline{\mathbf{x}}^{(1)}$ is given).

$$\mathbf{B} = (\mathbf{D} - \mathbf{L})^{-1}\mathbf{R} = \begin{bmatrix} 0 & 1/5 & 0 \\ 0 & 1/15 & 1/3 \\ 0 & 1/30 & 1/6 \end{bmatrix}.$$

The 1-norm of this matrix is 1/2. According to the error estimation formula we have

$$\|\overline{\mathbf{x}}^{(k)} - \overline{\mathbf{x}}^{\star}\|_1 \le \frac{\|\mathbf{B}\|_1^k}{1 - \|\mathbf{B}\|_1} \|\overline{\mathbf{x}}^{(1)} - \overline{\mathbf{x}}^{(0)}\|_1 \le \frac{(1/2)^k}{1 - 1/2} 6.60015 \le 5 \times 10^{-6}.$$

We estimated the norm $\|\overline{\mathbf{x}}^{(1)} - \overline{\mathbf{x}}^{(0)}\|_1$ from above, since the elements of the vector $\overline{\mathbf{x}}^{(1)}$ are rounded values.

We get that the error is less than the require value from the iteration step $k = 22$.

6. (7p) Give the QR decomposition of the matrix $\mathbf{A} = \begin{bmatrix} 0 & -4 \\ 0 & 0 \\ -5 & -2 \end{bmatrix}$ (it is enough to use two Givens rotations)!

Solution: The first givens rotation is constructed with the 1. and 3. element of the first column, the second one with the 2. and 3. elements of the second column.

$$\mathbf{G}_2\mathbf{G}_1\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 0 & 0 \\ -5 & -2 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ 0 & 4 \\ 0 & 0 \end{bmatrix} = \mathbf{R}.$$

The matrix $\mathbf{Q}$

$$\mathbf{Q} = \mathbf{G}_1^T\mathbf{G}_2^T = \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \end{bmatrix}.$$

The QR decomposition is not unique, thus other decompositions are also possible.

(Megjegyzés: In this special case, we can reduce the matrix to an upper triangular form with one permutation matrix (this is also orthogonal).

$$\mathbf{PA} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -4 \\ 0 & 0 \\ -5 & -2 \end{bmatrix} = \begin{bmatrix} -5 & -2 \\ 0 & -4 \\ 0 & 0 \end{bmatrix} = \mathbf{R},$$

moreover $\mathbf{Q} = \mathbf{P}^T$.)