Questions for the course

Numerical Linear Algebra

TMA265/MMA600

Date: October 30, 2018, Time: 14.00 - 18.00

Place: Samhällsbyggnaden (SB)

- Examiner: Larisa Beilina, tel. 070-4177036 or at work 031- 772 3567. Rond: Olof Zetterqvist, 031 - 772 5325.
- Results: results of examination can be received at the latest at 10 November at the student's office at the Department of Mathematics, daily 12.30-13.00;
- Grades: to pass (get G) requires 15 points together with points from homework assignments and computer exercises. Grades are set according to the table on the course homepage.
- Solutions will be announced at the end of exam (placed on the homepage of course).
- Aids: you can use written by hand notes on the one side of A4 sheet. Easy (not advanced) calculators are also allowed to use.

Instructions

- Answer to the question carefully and clearly.
- Write on the one side of the sheet. Do not use a red pen. Do not answer more than to the one question for one page.
- Sort your answers by the order of appearance of questions. Mark on the cover answered questions. Count the number of sheets you have and fill the number of every page on the cover.

Question 1

- 1. Using definition of singular values σ, find singular values of a matrix A which is defined as

$$\mathbf{A} = \begin{bmatrix} -i & 0 \\ 10 & i \end{bmatrix}$$

  Compute $\|A\|_2$ of a matrix A.

  (2p)
- 2. Prove that $\|A\|_2 = \|A^T\|_2$ for all matrices A of the size $n \times n$.

  (0.5p)
- 3.

  Suppose that you want to solve $(x-7)^7 = 0$ in Matlab. Explain why computation of roots of this polynomial is ill-posed problem.

  (2p)

Question 2

- 1. Prove that if the matrix B is nonsingular, then the matrix A is symmetric positive definite if and only if $B^T AB$ is symmetric positive definite.

(1.5p)

- 2. Prove that if there exists a unique unit lower triangular L and non-singular upper triangular U such that $A = LU$ then all leading principal submatrices of A are non-singular.
  (2p)

Question 3

- 1. Let us consider the problem of solution of linear system of equations $Ax = b$. Let $\tilde{x}$ be approximate solution of this equation such that $\delta x = \tilde{x} - x$ and δb be perturbation of the right hand side b. Assume that we don't have perturbations in elements of the matrix A. Derive the upper estimate for the relative change $\frac{||\delta x||}{||x||}$ in terms of the condition number $k(A)$ of the matrix A and relative change of the data $\frac{||\delta b||}{||b||}$.
  (1p)
- 2. Suppose that A is an invertible square matrix and u, v are vectors. Suppose furthermore that $1 + v^T A^{-1} u \neq 0$. Then the Sherman-Morrison formula states that

$$(A + uv^T)^{-1} = A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u}.$$

Here, uv^T is the outer product of two vectors u and v.

Verify that the matrix $Y = A^{-1} - \frac{A^{-1} uv^T A^{-1}}{1 + v^T A^{-1} u}$ (the right-hand side of the Sherman - Morrison formula) is the inverse of a matrix $X = A + uv^T$ (the inverse of the left hand-side of the Sherman - Morrison formula) if and only if $XY = YX = \mathrm{I}$.
(2p)

Question 4

- 1. Let A has full rank and $A = QR = U\Sigma V^T$ be the QR and SVD decompositions, respectively, of the m-by-n matrix A, where $m \geq n$. Prove that the Moore-Penrose pseudoinverse of A can be computed as

$$A^+ = R^{-1} Q^T = V \Sigma^{-1} U^T.$$

(2p)
- 2. Let $Q^T A Q = \Lambda$ be the eigendecomposition of a symmetric matrix A, Q is orthogonal matrix. Prove that the Rayleigh quotient $\rho(x, A)$ of a matrix A and nonzero vector x is equivalent to the Rayleigh quotient $\rho(\hat{x}, \Lambda)$ of a matrix Λ and nonzero vector $\hat{x} = Q^T x$.
  (1p)

Question 5

- Perform the first step of Hessenberg reduction for the matrix A

$$\mathbf{A} = \begin{bmatrix} 1 & 3 & 0 & 4 & 1 \\ 1 & 1 & 1 & -2 & 1 \\ -1 & 3 & 1 & 0 & 1 \\ 1 & -1 & 1 & -2 & 1 \\ -1 & 3 & 2 & 0 & -1 \end{bmatrix}$$

In other words, choose Q_1 such that

$$Q_1 A = \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix} \text{ and } A_1 \equiv Q_1 A Q_1^T = \begin{bmatrix} x & x & x & x & x \\ x & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \\ 0 & x & x & x & x \end{bmatrix}.$$

Confirm that Q_1 leaves the first row of $Q_1 A$ unchanged, and Q_1^T leaves the first column of $Q_1 A Q_1^T$ unchanged, including the zeros.
(3p)

- Transform the given matrix A to the lower Hessenberg matrix using Given's rotation

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

(1p)

Question 6

- 1. Using the Gerschgorin's theorem estimate eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & -4 & 3 & 5 \\ 3 & 4 & 0 & 1 \\ 4 & 0 & 2 & 2 \\ 3 & 2 & 5 & 7 \end{bmatrix}$$

(1p)

- 2.
Let $A - \lambda B$ be regular pencil. Prove that if B is nonsingular then all eigenvalues of $A - \lambda B$ are finite and the same as the eigenvalues of AB^{-1} or $B^{-1}A$. **(2p)**

Question 7

- 1. Prove that the Power method converges to the largest eigenvalue of the matrix A and to the corresponding eigenvector.
(2p)
- 2. Briefly describe the Jacobi method for the solution of the symmetric eigenproblem. Explain how the classical Jacobi algorithm differs from cyclic-by-row Jacobi algorithm.
(2p)

Numerical Linear Algebra

TMA265/MMA600
Solutions to the examination at 30 October 2018

Question 1

1. By definition of singular values $\sigma = \sqrt{\lambda(A^* A)}$ we have:

$$A^* = \overline{A^T} = \begin{bmatrix} i & 10 \\ 0 & -i \end{bmatrix}, \quad A^* A = \begin{bmatrix} 101 & 10i \\ -10i & 1 \end{bmatrix},$$

and characteristic equation to solve for λ is $\lambda^2 - 102\lambda + 1 = 0$.

Solving above equation we get eigenvalues $\lambda_{1,2} = \frac{102 \pm \sqrt{10400}}{2}$, or $\lambda_1 = 101.9902, \lambda_2 = 0.0098$.

Then singular values will be: $\sigma_1 = \sqrt{\lambda_1} = 10.0990, \sigma_2 = \sqrt{\lambda_2} = 0.0990$, and by properties of the norm we have $\|A\|_2 = \max(\sigma_1, \sigma_2) = 10.099$.

2. Using definition of SVD of A we have: $A = U\Sigma V^T$ and $A^T = V\Sigma^T U^T = V\Sigma U^T$. Then $\|A\|_2 = \|A^T\|_2 = \sigma_1$.

3. Let us analyze the problem

$$(\tilde{x} - 7)^{15} = \varepsilon,$$

for $\varepsilon = 10^{-15}$.

Then solving the above equation for $\tilde{x}$ we get:

$$\tilde{x} - 7 = \varepsilon^{1/15},$$
$$\tilde{x} = 7 + \varepsilon^{1/15}.$$

If $\varepsilon = 10^{-15}$ then $\varepsilon^{1/15} = 10^{-15/15} = 10^{-1}$ and thus computed $\tilde{x}$ will be $\tilde{x} = 7 + \varepsilon^{1/15} = 7.1$ with big changes in output data $\delta x = \tilde{x} - x = 0.1$. We observe that roots of polynomial $(x-7)^{15} = 0$ are very sensitive to the changes in the coefficients : small changes in the coefficients $|\delta c|/|c|$ will give big changes in the output data, or in $|\delta x|/|x| = 0.1/7 \approx 0.0143$, and thus the problem is ill-posed.

Question 2

1.

$\Rightarrow$ If B is nonsingular then we can define the vector $y = Bx \neq 0$ for all $x \neq 0$, such that $y^T A y = (x^T B^T) A (Bx) = x^T (B^T A B) x > 0$ for all $x \neq 0$. We know that $y^T A y > 0$ for all $y \neq 0$ since A is symmetric positive definite, and thus, using the equation above we conclude that $B^T A B$ is also symmetric positive definite.

$\Leftarrow$ If $B^T A B$ is symmetric positive definite then

$$x^T (B^T A B) x > 0 \tag{0.1}$$

for all $x \neq 0$. Let us define $y = Bx \neq 0$. Then we can rewrite (0.1) for all $y \neq 0$ as

$$y^T A y > 0, \tag{0.2}$$

and thus, A is also symmetric positive definite.

2. We can write $A = LU$ as

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} L_{11} & 0 \\ L_{21} & L_{22} \end{bmatrix} \times \begin{bmatrix} U_{11} & U_{12} \\ 0 & U_{22} \end{bmatrix} = \begin{bmatrix} L_{11}U_{11} & L_{11}U_{12} \\ L_{21}U_{11} & L_{21}U_{12} + L_{22}U_{22} \end{bmatrix}$$

where A_{11} is a j-by-j leading principal submatrix, as well as L_{11} and U_{11}. Therefore $det A_{11} = det(L_{11}U_{11}) = det L_{11} det U_{11} = 1 \cdot det U_{11} \neq 0$, as well as $det A_{22} = det(L_{21}U_{12} + L_{22}U_{22}) \neq 0$ since L is unit triangular and U is non-singular. Thus, all leading submatrices of A are non-singular.

Question 3

1. We have: $Ax = b$, $A(x + \delta x) = b + \delta b$, $\delta x = A^{-1}\delta b$. Take norms: $\|\delta x\| \leq \|A^{-1}\|\|\delta b\|$. Then $\frac{\|\delta x\|}{\|x\|} \leq \|A^{-1}\|\|A\|\frac{\|\delta b\|}{\|A\|\|x\|} = k(A)\frac{\|\delta b\|}{\|A\|\|x\|} \leq k(A)\frac{\|\delta b\|}{\|b\|}$

2. We first verify that the right hand side (Y) satisfies $XY =$ I.

$$\begin{aligned} XY &= (A + uv^T)\left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right) \\ &= AA^{-1} + uv^TA^{-1} - \frac{AA^{-1}uv^TA^{-1} + uv^TA^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u} \\ &= I + uv^TA^{-1} - \frac{uv^TA^{-1} + uv^TA^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u} \\ &= I + uv^TA^{-1} - \frac{u(1 + v^TA^{-1}u)v^TA^{-1}}{1 + v^TA^{-1}u}. \end{aligned}$$

Note that $v^TA^{-1}u$ is a scalar, so $(1 + v^TA^{-1}u)$ can be factored out, leading to:

$$XY = I + uv^TA^{-1} - uv^TA^{-1} = I.$$

In the same way, it is verified that

$$YX = \left(A^{-1} - \frac{A^{-1}uv^TA^{-1}}{1 + v^TA^{-1}u}\right)(A + uv^T) = I.$$

Question 4

1. We start from the normal equations:

$$\begin{aligned} A^+ &= (A^TA)^{-1}A^T \\ &= (R^TQ^TQR)^{-1}R^TQ^T = (R^TR)^{-1}R^TQ^T \\ &= R^{-1}R^{-T}R^TQ^T = R^{-1}Q^T. \end{aligned}$$

$$\begin{aligned} A^+ &= (A^TA)^{-1}A^T = ((U\Sigma V^T)^TU\Sigma V^T)^{-1}(U\Sigma V^T)^T = (V\Sigma U^TU\Sigma V^T)^{-1}(U\Sigma V^T)^T \\ &= (V\Sigma^2V^T)^{-1}V\Sigma U^T = V^{-1}\Sigma^{-2}V^{-T}V\Sigma U^T = V^{-1}\Sigma^{-2}V^2\Sigma U^T = V\Sigma^{-1}U^T. \end{aligned}$$

2. We change variables in Rayleigh quotient iteration to $\hat{x} = Q^Tx$. Then

$$\rho = \rho(x, A) = \frac{x^TAx}{x^Tx} = \frac{\hat{x}^TQ^TAQ\hat{x}}{\hat{x}^TQ^TQ\hat{x}} = \frac{\hat{x}^T\Lambda\hat{x}}{\hat{x}^T\hat{x}} = \rho(\hat{x}, \Lambda)$$

Question 5

- 1. To perform Hessenberg reduction we use Hauseholder transformation in following steps:
  - Choose $x = (1, -1, 1, -1)^T$ and compute
$$u = x + \alpha e_1,$$
where $\alpha = -sign(1) \cdot ||x||$, $\|x\| = \sqrt{4} = 2$, and thus $\alpha = -2$.
  - Construct $u = x + \alpha e_1 = (1, -1, 1, -1)^T - (2, 0, 0, 0)^T = (-1, -1, 1, -1)^T$.
  - Construct
$$v = \frac{u}{\|u\|}$$
with $\|u\| = \sqrt{4} = 2$.
Therefore $v = (-1/2, -1/2, 1/2, -1/2)^T$.
  - Compute
$$Q' = I - 2vv^T = \begin{pmatrix} 1/2 & -1/2 & 1/2 & -1/2 \\ -1/2 & 1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ -1/2 & -1/2 & 1/2 & 1/2 \end{pmatrix}.$$
  - Construct the matrix of the Householder transformation as:
$$Q_1 = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0.5 & -0.5 & 0.5 & -0.5 \\ 0 & -0.5 & 0.5 & 0.5 & -0.5 \\ 0 & 0.5 & 0.5 & 0.5 & 0.5 \\ 0 & -0.5 & -0.5 & 0.5 & 0.5 \end{pmatrix}$$
  - The first step in Hessenberg reduction will be:
$$Q_1 A = \begin{pmatrix} 1 & 3 & 0 & 4 & 1 \\ 2 & -3 & -0.5 & -2 & 1 \\ 0 & -1 & -0.5 & 0 & 1 \\ 0 & 3.0 & 2.5 & -2 & 1 \\ 0 & -1 & 0.5 & 0 & -1 \end{pmatrix}.$$
such that Q_1 leaves the first row of $Q_1 A$ unchanged, and then
$$Q_1 A Q_1^T = \begin{pmatrix} 1 & 3 & 0 & 4 & 1 \\ 2 & -2.75 & -0.25 & -2.25 & 1.25 \\ 0 & -0.75 & -0.25 & -0.25 & 1.25 \\ 0 & -1.25 & -1.75 & 2.25 & -3.25 \\ 0 & -0.25 & 1.25 & -0.75 & -0.25 \end{pmatrix}.$$
Q_1^T leaves the first column of $Q_1 A Q_1^T$ unchanged, including the zeros.
- 2. To obtain the lower Hessenberg matrix of the matrix
$$\mathbf{A} = \begin{bmatrix} 3 & 4 & 3 \\ 1 & 5 & 4 \\ 2 & 1 & 6 \end{bmatrix}$$
using Given's rotation we have to zero out $(1, 3)$ element of the matrix A.
To do that we compute elements of Givens matrix G and then compute $A * G^T$.

We compute G for A^T and thus, values of c, s are computed from the known $a = 4$ and $b = 3$ as

$$\begin{bmatrix} c & -s \\ s & c \end{bmatrix} \cdot \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} r \\ 0 \end{bmatrix}$$

to get formulas:

$$r = \sqrt{a^2 + b^2} = \sqrt{(4)^2 + 3^2} = 5,$$
$$c = \frac{a}{r} = 0.8,$$
$$s = \frac{-b}{r} = -0.6.$$

The Given's matrix will be

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & c & -s \\ 0 & s & c \end{bmatrix}$$

or

$$\mathbf{G} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0.8 & 0.6 \\ 0 & -0.6 & 0.8 \end{bmatrix}$$

Then the lower Hessenberg matrix will be computed as

$$\mathbf{AG^T} = \begin{bmatrix} 3 & 5 & 0 \\ 1 & 6.4 & 0.2 \\ 2 & 4.4 & 4.2 \end{bmatrix}$$

Question 6

1. By Gershgorin's Theorem, the eigenvalues λ of A are located in the union of the n disks

$$|\lambda - a_{kk}| \leq \sum_{j \neq k} |a_{kj}|.$$

Thus, elements on diagonal will be centers of these discs and for radius we use the formula

$$\sum_{j \neq i} |a_{ij}| = R_i.$$

Eigenvalues of A are located in union of the following four discs:
$D(0, 12), D(4, 4), D(2, 6),$ and $D(7, 10)$

2. If B is nonsingular and λ is an eigenvalue, then $0 = \det(A - \lambda B) = \det(B(AB^{-1} - \lambda I)) = \underbrace{\det B}_{\neq 0} \det(AB^{-1} - \lambda I)$. Thus, $0 = \det(AB^{-1} - \lambda I) = \det(B^{-1}A - \lambda I)$ and thus λ is also an eigenvalue of AB^{-1} and $B^{-1}A$.

Question 7

1. See Lecture 10:

ALGORITHM. Power method: Given x_0, we iterate

$$
\begin{array}{ll}
i = 0 & \\
\textit{repeat} & \\
\quad y_{i+1} = Ax_i & \\
\quad x_{i+1} = y_{i+1}/||y_{i+1}||_2 & (\textit{approximate eigenvector}) \\
\quad \tilde{\lambda}_{i+1} = x_{i+1}^T A x_{i+1} & (\textit{approximate eigenvalue}) \\
\quad i = i+1 & \\
\textit{until convergence} &
\end{array}
$$

Assume that $A = S\Lambda S^{-1}$ is diagonalizable, with $\Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n)$ and the eigenvalues sorted so that $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$. Write $S = [s_1, \ldots, s_n]$, where the columns s_i are the corresponding eigenvectors and also satisfy $||s_i||_2 = 1$. This lets us write $\mathbf{x_0} = \mathbf{S}(\mathbf{S^{-1}x_0}) \equiv \mathbf{S}([\xi_1, \ldots, \xi_n]^\mathbf{T})$. Also, since $A = S\Lambda S^{-1}$, we can write

$$A^i = \underbrace{(S\Lambda S^{-1}) \cdots (S\Lambda S^{-1})}_{i\,\text{times}} = S\Lambda^i S^{-1}$$

since all the $S^{-1} \cdot S$ pairs cancel.

This finally lets us write

$$A^i x_0 = \underbrace{S\Lambda^i S^{-1}}_{A^i} \cdot \underbrace{S \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}}_{x_0} = S \begin{bmatrix} \xi_1 \lambda_1^i \\ \xi_2 \lambda_2^i \\ \vdots \\ \xi_n \lambda_n^i \end{bmatrix} = \xi_1 \lambda_1^i S \begin{bmatrix} 1 \\ \frac{\xi_2}{\xi_1}(\frac{\lambda_2}{\lambda_1})^i \\ \vdots \\ \frac{\xi_n}{\xi_1}(\frac{\lambda_n}{\lambda_1})^i \end{bmatrix}.$$

The vector j in brackets converges to e_1, so $A^i x_0$ gets closer and closer to a multiple of $Se_1 = s_1$, the eigenvector corresponding to λ_1. Therefore, $\tilde{\lambda}_i = x_i^T A x_i$ converges to $s_1^T A s_1 = s_1^T \lambda_1 s_1 = \lambda_1$.

2. See Lecture 14, slides 8-19.

Given a symmetric matrix $A = A_0$, Jacobi's method produces a sequence A_1, A_2, ... of orthogonally similar matrices, which eventually converge to a diagonal matrix with the eigenvalues on the diagonal. A_{i+1} is obtained from A_i by the formula $A_{i+1} = J_i^T A_i J_i$, where J_i is an orthogonal matrix called a *Jacobi rotation*. Thus

$$
\begin{array}{rcl}
A_m & = & J_{m-1}^T A_{m-1} J_{m-1} \\
& = & J_{m-1}^T J_{m-2}^T A_{m-2} J_{m-2} J_{m-1} = \cdots \\
& = & J_{m-1}^T \cdots J_0^T A_0 J_0 \cdots J_{m-1} \\
& = & J^T A J.
\end{array}
$$

If we choose each J_i appropriately, A_m approaches a diagonal matrix Λ for large m. Thus we can write $\Lambda \approx J^T A J$ or $J \Lambda J^T \approx A$. Therefore, the columns of J are approximate eigenvectors.

ALGORITHM. *Classical Jacobi's algorithm:*

while off$(A) >$ tol *(where* tol *is the stopping criterion set by user)*
  choose j and k so a_{jk} is the largest off-diagonal entry in magnitude
  call Jacobi-Rotation(A, j, k)
end while

In practice, we do not use the classical Jacobi's algorithm because searching for the largest entry is too slow. Instead, we use the following simple method to choose j and k.

ALGORITHM. Cyclic-by-row-Jacobi: Sweep through the off diagonals of A rowwise.

```
repeat
  for j = 1 to n − 1
    for k = j + 1 to n
      call Jacobi-Rotation (A, j, k)
    end for
  end for
until A is sufficiently diagonal
```