Department of Mathematical Sciences

Examination paper for **TMA4205 Numerical Linear Algebra**

Academic contact during examination: Markus Grasmair

Phone: 97580435

Examination date: December 16, 2015
Examination time (from–to): 09:00–13:00
Permitted examination support material: C: Specified, written and handwritten examination support materials are permitted. A specified, simple calculator is permitted. The permitted examination support materials are:

- Y. Saad: Iterative Methods for Sparse Linear Systems. 2nd ed. SIAM, 2003 (book or printout)
- L. N. Trefethen and D. Bau: Numerical Linear Algebra, SIAM, 1997 (book or photocopy)
- G. Golub and C. Van Loan: Matrix Computations. 3rd ed. The Johns Hopkins University Press, 1996 (book or photocopy)
- J. W. Demmel: Applied Numerical Linear Algebra, SIAM, 1997 (book or printout)
- E. Rønquist: Note on The Poisson problem in $\mathbb{R}^2$: diagonalization methods (printout)
- K. Rottmann: Matematisk formelsamling
- Your own lecture notes from the course (handwritten)

Language: English
Number of pages: 6
Number pages enclosed: 0

Checked by:

Date Signature

Problem 1

a) Let

$$A = \begin{pmatrix} 0 & 2 \\ 1 & 3 \end{pmatrix}.$$

Perform one iteration of QR-algorithm (for computing eigenvalues) with shift $\mu = 1$.

Solution: The shifted matrix

$$A - \mu I = \begin{pmatrix} -1 & 2 \\ 1 & 2 \end{pmatrix}$$

has orthogonal (but not orthonormal) columns. Thus Q in its QR-factorization is easily obtained by rescaling the columns, and R is diagonal with the length of the columns of A on the diagonal:

$$Q = \begin{pmatrix} -1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{pmatrix} \qquad R = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 2\sqrt{2} \end{pmatrix}$$

Furthermore $Q = Q^{\mathrm{T}}$, $R = R^{\mathrm{T}}$, therefore $RQ + \mu I = (QR + \mu I)^{\mathrm{T}} = A^{\mathrm{T}}$.

b) Let now $A \in \mathbb{R}^{n\times n}$ be an arbitrary square matrix. Assume that the shift μ in the QR-algorithm with shifts is equal to one of the eigenvalues of A. How can we easily detect this situation based on the QR factorization of the shifted matrix?

Solution: If $\mu \in \sigma(A)$ then $0 = \det(A - \mu I) = \det(QR) = \det(Q)\det(R)$. Since Q is unitary, $|\det(Q)| = 1$ and therefore $\det(R) = 0$. Now R is upper triangular and therefore must have a 0 element on the diagonal.

Problem 2 Let

$$A = \begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix}.$$

a) Find a singular value decomposition of A.

Solution: A is symmetric and therefore is unitarily diagonalizable. Its characteristic polynomial is

$$p_A(\lambda) = (1-\lambda)^2 - 9 = \lambda^2 - 2\lambda - 8$$

thus the eigenvalues are $\lambda_1 = -2$ and $\lambda_2 = 4$. The corresponding eigenvectors are for example

$$q_1 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix} \qquad q_2 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Therefore

$$A = [q_1, q_2]\mathrm{diag}(-2, 4)[q_1, q_2]^{\mathrm{T}} = [q_1, q_2]\mathrm{diag}(2, 4)[-q_1, q_2]^{\mathrm{T}},$$

the latter being an SVD of A.

b) Find the best, with respect to the $\|\cdot\|_2$-norm, rank 1 approximation of A^{-1}. That is, find some vectors $p, q \in \mathbb{R}^2$ minimizing the norm $\|pq^{\mathrm{T}} - A^{-1}\|_2$.

Solution: Let $U\Sigma V^{\mathrm{T}}$ be an SVD of A with non-singular Σ. Then $A^{-1} = V\Sigma^{-1}U^{\mathrm{T}}$ is an SVD of A^{-1}. The best rank-1 approximation of A^{-1} is $V_n\sigma_n^{-1}U_n^{\mathrm{T}}$, where σ_n is the smallest singular value of A (reciprocal of the largest singular value of A^{-1}) with the corresponding singular vectors U_n, V_n. Given the previously computed SVD we can take $p = -q_1$, $q = 1/2q_1$,

$$pq^{\mathrm{T}} = \begin{pmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix} 1/2 \begin{pmatrix} 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} = \begin{pmatrix} -1/4 & 1/4 \\ 1/4 & -1/4 \end{pmatrix}$$

c) Let $B \in \mathbb{R}^{n\times n}$ be an arbitrary non-singular matrix, and let $x_0, b \in \mathbb{R}^n$ be given vectors. Further, let $B_1 = pq^{\mathrm{T}}$ be a rank 1 approximation of B^{-1} for some $p, q \in \mathbb{R}^n$. Consider a general projection method with a search space $\mathcal{K}$ and a constraint space $\mathcal{L}$ for solving a left-preconditioned linear algebraic system $B_1Bx = B_1b$. Show that $\tilde{x} \in x_0 + \mathcal{K}$ satisfies Petrov–Galerkin conditions if and only if at least one of the following conditions hold:

$$(i) \quad b - B\tilde{x} \perp q \qquad \text{or} \qquad (ii) \quad p \perp \mathcal{L}.$$

Solution: Petrov–Galerkin conditions are:

$$\tilde{x} \in x_0 + \mathcal{K} \qquad \text{and} \qquad B_1b - B_1B\tilde{x} = pq^{\mathrm{T}}(b - B\tilde{x}) \perp \mathcal{L}.$$

The condition on the left is given; the condition on the right is a product of a vector p and a scalar $q^{\mathrm{T}}(b - B\tilde{x})$. Thus the resulting vector is orthogonal to $\mathcal{L}$ iff $p \perp \mathcal{L}$ or $q^{\mathrm{T}}(b - B\tilde{x}) = 0$.

Problem 3 Let $A \in \mathbb{C}^{n\times n}$ be an arbitrary square matrix. We put $B = (A + A^{\mathrm{H}})/2$, $C = (A - A^{\mathrm{H}})/2$; in particular $A = B + C$.

a) Show that $\sigma(B) \subset \mathbb{R}$ and $\sigma(C) \subset i\mathbb{R}$, where $\sigma(\cdot)$ is the spectrum of a matrix and $i^2 = -1$.

Solution: The short answer here is that B is Hermitian ($B = B^{\mathrm{H}}$ and C is skew-Hermitian ($C = -C^{\mathrm{H}}$). One can also argue directly like this: if $Bv = \lambda v$ for some $\lambda \in \mathbb{C}$ and $v \in \mathbb{C}^n \setminus \{0\}$ then

$$\lambda\|v\|^2 = (\lambda v, v) = (Bv, v) = (v, B^{\mathrm{H}}v) = (v, Bv) = (v, \lambda v) = \bar{\lambda}\|v\|^2.$$

Since $\|v\|^2 \neq 0$ we get $\lambda = \bar{\lambda}$, or $\lambda \in \mathbb{R}$.

b) Let $\alpha \in \mathbb{C}$ be an arbitrary scalar, and $I \in \mathbb{R}^{n\times n}$ be the identity matrix. Show that $B \pm \alpha I$ and $C \pm \alpha I$ are unitarily diagonalizable.

Solution: The matrix is unitarily diagonalizable iff it is normal. Normality follows from that of C, B, and the fact that they commute with $\pm\alpha I$. Alternatively, direct computation:

$$(B \pm \alpha I)^{\mathrm{H}}(B \pm \alpha I) = (B \pm \bar{\alpha} I)(B \pm \alpha I) = B^2 \pm (\alpha + \bar{\alpha})B + \alpha\bar{\alpha}I = (B \pm \alpha I)(B \pm \alpha I)^{\mathrm{H}},$$
$$(C \pm \alpha I)^{\mathrm{H}}(C \pm \alpha I) = (-C \pm \bar{\alpha} I)(C \pm \alpha I) = -C^2 \pm (\bar{\alpha} - \alpha)C + \alpha\bar{\alpha}I = (C \pm \alpha I)(C \pm \alpha I)^{\mathrm{H}}.$$

Assume now that that both $B + \alpha I$ and $C + \alpha I$ are non-singular. Consider the following iterative algorithm for solving the system $Ax = b$ starting with some initial approximaiton $x_0 \in \mathbb{C}^n$:

for $k = 0, 1, \ldots,$ **do**
  $x_{k+1/2} = (B + \alpha I)^{-1}[b - (C - \alpha I)x_k]$
  $x_{k+1} = (C + \alpha I)^{-1}[b - (B - \alpha I)x_{k+1/2}]$
end for

c) Show that if $\alpha = 0$ the algorithm converges, but the limits $\bar{x} = \lim_{k\to\infty} x_k$ and $\hat{x} = \lim_{k\to\infty} x_{k+1/2}$ do not necessarily coincide or solve the system $A\bar{x} = b$ (or $A\hat{x} = b$).

Solution: Direct computation: $x_{1/2} = B^{-1}[b - Cx_0]$, $x_1 = C^{-1}[b - Bx_{1/2}] = C^{-1}[b - BB^{-1}[b - Cx_0]] = x_0$. Thus algorithm converges, but not necessarily to the solution of the system $Ax = b$.

d) Show that the sequence of points x_k generated by the algorithm satisfies $\lim_{k\to\infty} \|x_k - A^{-1}b\|_2 = 0$ *for an arbitrary* $x_0 \in \mathbb{C}^n$ if and only if $\rho((C+\alpha I)^{-1}(B-\alpha I)(B+\alpha I)^{-1}(C-\alpha I)) < 1$, where $\rho(\cdot)$ is a spectral radius of a matrix.

Solution: Exactly as with matrix-splitting algorithms, one considers the error $e_k = x_k - A^{-1}b$. Then $e_{k+1/2} = (B+\alpha I)^{-1}(C-\alpha I)e_k$, and $e_{k+1} = (C+\alpha I)^{-1}(B-\alpha I)e_{k+1/2} = (C+\alpha I)^{-1}(B-\alpha I)(B+\alpha I)^{-1}(C-\alpha I)e_k$. Again, exactly as with matrix-splitting algorithms we obtain convergence from an arbitrary starting point iff the spectral radius of the iteration matrix is < 1.

e) Let A be Hermitian and positive definite. Show that the algorithm above converges for an arbitrary $\alpha > 0$ and $x_0 \in \mathbb{C}$.

Solution: If A is Hermitian then $C = 0$, $A = B$. If A is additionally positive definite then so is B. Thus the algorithm's iteration matrix is $(\alpha I)^{-1}(B-\alpha I)(B+\alpha I)^{-1}(-\alpha I) = -(B-\alpha I)(B+\alpha I)^{-1}$. If $\lambda_1, \dots, \lambda_n$ are eigenvalues of B, then the eigenvalues of the iteration matrix are $-(\lambda_i - \alpha)/(\lambda_i + \alpha)$. Since $\lambda_i > 0$, $\alpha > 0$, the spectral radius of the iteration matrix will be < 1.

Problem 4 Let $A \in \mathbb{R}^{n\times n}$ be a symmetric and positive definite matrix and $b \in \mathbb{R}^n$ be an arbitrary vector. Let $x_* = A^{-1}b$.

a) Show that the standard A-norm error estimate for the conjugate gradient algorithm:

$$\|x_m - x_*\|_A \le 2\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^m \|x_0 - x_*\|_A,$$

implies the 2-norm estimate

$$\|x_m - x_*\|_2 \le 2\sqrt{\kappa}\left[\frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}\right]^m \|x_0 - x_*\|_2,$$

where $\kappa = \lambda_{\max}(A)/\lambda_{\min}(A)$ is the spectral condition number of A.

Solution: Since A is normal (symmetric!) then it is unitarily diagonalizable. Let $q_1, \dots, q_n$ be an orthonormal basis composed of eigenvvectors of A and $\lambda_1, \dots, \lambda_n$ be the corresponding (positive) eigenvalues. Let $x = \sum_i \alpha_i q_i$. Then $\|x\|_2^2 = \sum_i |\alpha_i|^2$, $\|x\|_A^2 = \sum_i |\alpha_i|^2\lambda_i$. Thus $\min\{\lambda_i\}\|x\|_2^2 \le \|x\|_A^2 \le \max\{\lambda_i\}\|x\|_2^2$, from which the estimate follows.

Let V_{2m}, H_{2m} be the matrices produced after $2m$ iterations of Arnoldi process (without breakdown) applied to A, starting from some vector $r_0 = b - Ax_0$.

Let $\hat{V}_m := [V_{*,1}, V_{*,3}, \dots, V_{*,2m-1}]$ (i.e., the submatrix of V_{2m} corresponding to *odd* columns), and $\hat{H}_m = \hat{V}_m^{\mathrm{T}} A \hat{V}_m$.

b) Show that $\hat{H}_m$ is a non-singular diagonal matrix.

Solution: By construction $H_{2m} = V_{2m}^{\mathrm{T}} A V_{2m}$, thus $\hat{H}_m = \hat{V}_m^{\mathrm{T}} A \hat{V}_m$ is its submatrix corresponding to the odd rows/columns. As A is symmetric, then H_{2m} is tri-diagonal, and as a result $\hat{H}_m$ is diagonal. Furthermore, $(\hat{H}_m)_{ii} = (\hat{V}_m)_{*i}^{\mathrm{T}} A (\hat{V}_m)_{*i} > 0$ since A is positive definite and the columns of $\hat{V}_m$ (and V_{2m}) have length 1.

c) Consider now an orthogonal projection method for the system $Ax = b$ with $\mathcal{L} = \mathcal{K} = \mathrm{Ran}\hat{V}_m$, where we seek $\hat{x}_m \in x_0 + \mathcal{K}$ satisfying the Galerkin orthogonality condition. Show that $\hat{x}_m = x_0 + (r_0^{\mathrm{T}} r_0)(r_0^{\mathrm{T}} A r_0)^{-1} r_0$. (That is, the method takes one steepest descent step and then stops improving the solution.)

Solution: Since $\mathcal{K} = \mathrm{Ran}\hat{V}_m$ we can write $\hat{x}_m = x_0 + \hat{V}_m\hat{y}_m$ for some $\hat{y}_m$. Galerkin orthogonality condition is

$$0 = \hat{V}_m^{\mathrm{T}}(b - A\hat{x}_m) = \hat{V}_m^{\mathrm{T}}(r_0 - A\hat{V}_m\hat{y}_m) = \hat{V}_m^{\mathrm{T}}r_0 - \hat{H}_m\hat{y}_m = \|r_0\|_2 e_1 - \hat{H}_m\hat{y}_m,$$

since $(\hat{V}_m)_{*1} = r_0/\|r_0\|_2$. Therefore

$$\hat{y}_m = \begin{pmatrix} \|r_0\|_2/(\hat{H}_m)_{11} \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \|r_0\|_2/((\hat{V}_m)_{*1}^{\mathrm{T}}A(\hat{V}_m)_{*1}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} \|r_0\|_2^3/(r_0^{\mathrm{T}}Ar_0) \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

As a result $\hat{x}_m = x_0 + \|r_0\|_2^3(r_0^{\mathrm{T}}Ar_0)^{-1}(\hat{V}_m)_{*1} = x_0 + \|r_0\|_2^2(r_0^{\mathrm{T}}Ar_0)^{-1}r_0$.