NTNU
Norwegian University of
Science and Technology

Department of Mathematical Sciences

Examination paper for **TMA4205 Numerical Linear Algebra**

Academic contact during examination:

Phone:

Examination date: 29th November 2016

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).
- E. Rønquist: Note on The Poisson problem in $\mathbb{R}^2$: diagonalization methods (printout).
- M. Grasmair: The singular value decomposition (printout).
- Rottmann, Matematisk formelsamling.
- Your own lecture notes from the course (handwritten).

Language: English

Number of pages: 9

Number of pages enclosed: 0

Checked by:

Date Signature

Problem 1 Let

$$A = \begin{pmatrix} 6 & 3 \\ 3 & 4 \end{pmatrix}.$$

Perform one step of the QR-method (for computing eigenvalues) with the shift parameter $\mu = 2$.

We first shift the matrix by μ and obtain

$$A - \mu I = \begin{pmatrix} 4 & 3 \\ 3 & 2 \end{pmatrix}.$$

Next we compute a QR-factorisation of $A - \mu I$ (e.g. using the Gram–Schmidt process) and obtain (for instance — the QR decomposition is not unique)

$$Q = \frac{1}{5}\begin{pmatrix} 4 & -3 \\ 3 & 4 \end{pmatrix}$$

and

$$R = \frac{1}{5}\begin{pmatrix} 25 & 18 \\ 0 & -1 \end{pmatrix}.$$

Now we multiply the matrix R from the right by Q and obtain

$$RQ = \frac{1}{25}\begin{pmatrix} 154 & -3 \\ -3 & -4 \end{pmatrix}.$$

Finally we add back μI and obtain

$$A^{(1)} = \frac{1}{25}\begin{pmatrix} 204 & -3 \\ -3 & 46 \end{pmatrix}.$$

(Note that the diagonal entries of $A^{(1)}$ coincide with the actual eigenvalues of A (which are $5 \pm \sqrt{10}$) on the first three significant digits.)

Problem 2 Assume that $A \in \mathbb{R}^{m \times n}$ with $m < n$ has the non-zero singular values $\sigma_1, \dots, \sigma_m$.

a) Given $\lambda \geq 0$, we define the matrices

$$G_\lambda := \begin{pmatrix} \lambda I & A \\ A^T & -\lambda I \end{pmatrix} \qquad \text{and} \qquad H_\lambda := \begin{pmatrix} \lambda I & A \\ -A^T & \lambda I \end{pmatrix}.$$

Show that the singular values of the matrices G_λ and H_λ are precisely the values $\sqrt{\sigma_k^2 + \lambda^2}$, $k = 1, \dots, m$, and λ. Additionally, show that both matrices are invertible in case $\lambda > 0$, and determine whether the matrices G_λ or H_λ are positive (semi-)definite.

The singular values of G_λ are precisely the square roots of the eigenvalues of the matrix $G_\lambda G_\lambda^T$. Now

$$G_\lambda G_\lambda^T = \begin{pmatrix} \lambda^2 I + AA^T & 0 \\ 0 & \lambda^2 + A^T A \end{pmatrix},$$

and the eigenvalues of this matrix are the eigenvalues of the matrices $\lambda^2 I + AA^T$ and $\lambda^2 I + A^T A$. By assumption, the matrix A has the non-zero eigenvalues $\sigma_1,\ldots,\sigma_m$, which implies that the matrix AA^T has the eigenvalues $\sigma_1^2,\ldots,\sigma_m^2$, and therefore the matrix $\lambda^2 I + AA^T$ the eigenvalues $\lambda^2+\sigma_1^2,\ldots,\lambda^2+\sigma_m^2$. Similarly, the matrix $A^T A$ has the non-zero eigenvalues $\sigma_1^2,\ldots,\sigma_m^2$, and additionally the eigenvalue 0 with multiplicity $n-m$. Thus the eigenvalues of $\lambda^2 + A^T A$ are precisely the values $\lambda^2+\sigma_1^2,\ldots,\lambda^2+\sigma_m^2$ and λ^2. Taking the square roots of these terms, we obtain the claimed singular values of G_λ.

For the matrix H_λ we obtain

$$H_\lambda H_\lambda^T = \begin{pmatrix} \lambda^2 I + AA^T & 0 \\ 0 & \lambda^2 + A^T A \end{pmatrix},$$

which is precisely the same as for G_λ, showing that G_λ and H_λ have the same singular values.

Since the matrices G_λ and $H_\lambda \in \mathbb{R}^{(m+n)\times(m+n)}$ have $m+n$ non-zero singular values, they are invertible.

In case $\lambda > 0$, the (symmetric) matrix G_λ cannot be positive semi-definite because of the negative entries on the diagonal.

In the case of the non-symmetric matrix H_λ, we need to consider its symmetrisation $(H_\lambda + H_\lambda^T)/2$ for determining whether it is positive semi-definite. However, since this symmetrisation yields the matrix λI, it is obvious that H_λ is positive definite for $\lambda > 0$ and positive semi-definite for $\lambda = 0$.

Given $a \in \mathbb{R}^m$ and $b \in \mathbb{R}^n$, we want to simultaneously solve the linear systems

$$\begin{aligned} u - Av &= a, \\ A^T u + v &= b. \end{aligned} \tag{1}$$

b) Show that the Jacobi method for the solution of this system converges for all initial values $u^{(0)}$ and $v^{(0)}$ in case $\|A\|_2 < 1$.

The Jacobi method for this system reads

$$\begin{aligned} u^{(k+1)} &= a + Av^{(k)}, \\ v^{(k+1)} &= b - A^T u^{(k)}, \end{aligned},$$

which can also be written as

$$\begin{pmatrix} u^{(k+1)} \\ v^{(k+1)} \end{pmatrix} = \begin{pmatrix} 0 & A \\ -A^T & 0 \end{pmatrix} \begin{pmatrix} u^{(k)} \\ v^{(k)} \end{pmatrix} + \begin{pmatrix} a \\ b \end{pmatrix}.$$

Thus we see that the matrix defining this method is the matrix H_0 defined above, and the Jacobi method will converge provided that $\|H_0\|_2 < 1$. However, $\|H_0\|$ is equal to the largest singular value of H_0, which, as shown above, is σ_1, which in turn is equal to $\|A\|_2$. Thus the Jacobi method converges in case $\|A\|_2 < 1$.

c) Given $\mu > 0$, we now apply the iteration

$$\begin{aligned} u^{(k+1)} &= \frac{1}{\mu}\Big(a + (\mu - 1)u^{(k)} + Av^{(k)}\Big), \\ v^{(k+1)} &= \frac{1}{\mu}\Big(b + (\mu - 1)v^{(k)} - A^T u^{(k)}\Big). \end{aligned}$$

For which $\mu > 0$ does this iteration converge to a solution of (1) independent of the starting vectors $u^{(0)}$ and $v^{(0)}$?

How should one choose the parameter μ in order to obtain the fastest convergence?

The matrix defining this iteration is

$$J = \frac{1}{\mu}\begin{pmatrix} (\mu-1)I & A \\ -A^T & (\mu-1)I \end{pmatrix} = \frac{1}{\mu}H_{\mu-1},$$

which has the singular values

$$\frac{\sqrt{\sigma_1^2 + (\mu-1)^2}}{\mu}, \ldots, \frac{\sqrt{\sigma_m^2 + (\mu-1)^2}}{\mu} \text{ and } \frac{\mu - 1}{\mu}.$$

The largest of the singular values is the value $\sqrt{\sigma_1^2 + (\mu-1)^2}/\mu$, and the iteration will converge provided that this value is smaller than 1. This is the case, if

$$\sqrt{\sigma_1^2 + (\mu-1)^2} < \mu,$$

which (with $\mu > 0$) is equivalent to

$$\sigma_1^2 + (\mu-1)^2 < \mu^2,$$

which in turn is equivalent to

$$\sigma_1^2 - 2\mu + 1 < 0$$

or

$$\mu > \frac{\sigma_1^2 + 1}{2}.$$

In order to obtain the fastest convergence, we would need the value $\sqrt{\sigma_1^2 + (\mu - 1)^2}/\mu$ to be as small as possible. Defining

$$f(\mu) := \frac{\sqrt{\sigma_1^2 + (\mu - 1)^2}}{\mu},$$

this means that we should find the minimum of f on $\mathbb{R}_{>0}$. The derivative of f is

$$f'(\mu) = \frac{\mu - 1}{\mu\sqrt{\sigma_1^2 + (\mu - 1)^2}} - \frac{\sqrt{\sigma_1^2 + (\mu - 1)^2}}{\mu^2},$$

and we have $f'(\mu) = 0$ if

$$\mu^2 - \mu = \sigma_1^2 + (\mu - 1)^2$$

or

$$\mu = \sigma_1^2 + 1.$$

Since $f(\mu) \to \infty$ as $\mu \to 0$ and $f(\mu) \to 1$ as $\mu \to \infty$, and $f(\sigma_1^2 + 1) < 1$, this implies that the minimum is obtained at this point, and we should choose $\mu = \sigma_1^2 + 1$ in order to obtain the fastest convergence.

Problem 3

a) We are given a linear system of the form

$$(I + uu^T)x = b,$$

where $I \in \mathbb{R}^{n \times n}$ is the n-dimensional identity matrix, and $u \in \mathbb{R}^n \setminus \{0\}$ is some given non-zero vector. Assume we apply the CG-method for solving this system. How many iterations do you expect the method to take until convergence is reached? Justify your answer!

The matrix uu^T is of rank one and has the unique non-zero eigenvalue $\|u\|_2^2$ with eigenvector u (as $uu^T u = \|u\|_2^2 u$). Thus the matrix $I + uu^T$ has the eigenvalues 1 and $1 + \|u\|_2^2$. Since there are only two distinct eigenvalues, it follows that the CG-method will converge after at most two steps.

b) Consider now in particular the system

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

Use the CG method with initialisation $x^{(0)} = 0$ for solving this system. Iterate until you have reached convergence.

We start with $r_0 = p_0 = b = (2, 2, 1, 0)^T$ and compute $Ap_0 = (7, 7, 6, 5)^T$, thus

$$\alpha_0 = \|r_0\|^2/(Ap_0, p_0) = 9/34.$$

Thus

$$x_1 = \frac{9}{34} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix}$$

and

$$r_1 = \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix} - \frac{9}{34} \begin{pmatrix} 7 \\ 7 \\ 6 \\ 5 \end{pmatrix} = \frac{1}{34} \begin{pmatrix} 5 \\ 5 \\ -20 \\ -45 \end{pmatrix}$$

Now

$$\beta_1 = \frac{\|r_1\|^2}{\|r_0\|^2} = \frac{825}{3468}$$

Thus

$$p_1 = \frac{1}{34} \begin{pmatrix} 5 \\ 5 \\ -20 \\ -45 \end{pmatrix} + \frac{825}{3468} \begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix} = \frac{1}{1156} \begin{pmatrix} 720 \\ 720 \\ -405 \\ -1530 \end{pmatrix}$$

Now

$$Ap_1 = \frac{1}{1156} \begin{pmatrix} 225 \\ 225 \\ -900 \\ -2025 \end{pmatrix}$$

and

$$\alpha_1 = \frac{\|r_1\|^2}{(Ap_1, p_1)} = \frac{1156}{1530}.$$

Thus

$$x_2 = \frac{9}{34}\begin{pmatrix}2\\2\\1\\0\end{pmatrix} + \frac{1156}{1530}\frac{1}{1156}\begin{pmatrix}720\\720\\-405\\-1530\end{pmatrix} = \begin{pmatrix}1\\1\\0\\-1\end{pmatrix},$$

which is already the solution of the system.

Problem 4 We consider the two linear systems

$$Ax = b \qquad \text{and} \qquad \tilde{A}\tilde{x} = \tilde{b},$$

where

$$\tilde{A} = QAQ^T \qquad \text{and} \qquad \tilde{b} = Qb$$

for some orthogonal matrix Q. Denote by $x^{(k)}$ the k-th iterate for the GMRES method for solving the system $Ax = b$ with initialisation $x^{(0)} = 0$, and by $\tilde{x}^{(k)}$ the k-th iterate for the GMRES method for solving the system $\tilde{A}\tilde{x} = \tilde{b}$, again with initialisation $\tilde{x}^{(0)} = 0$. Show that $\tilde{x}^{(k)} = Qx^{(k)}$.

The k-th iterate for the GMRES method for $Ax = b$ with $x^{(0)} = 0$ is defined by the conditions

$$x^{(k)} \in \mathcal{K}_k(A, b) \qquad \text{and} \qquad b - Ax^{(k)} \perp A\mathcal{K}_k(A, b).$$

Similarly, the k-th iterate for GMRES in case of the equation $\tilde{A}x = \tilde{b}$ is defined be the conditions

$$\tilde{x}^{(k)} \in \mathcal{K}_k(\tilde{A}, \tilde{b}) \qquad \text{and} \qquad \tilde{b} - \tilde{A}\tilde{x}^{(k)} \perp \tilde{A}\mathcal{K}_k(\tilde{A}, \tilde{b}).$$

Now we note that

$$\mathcal{K}_k(\tilde{A}, \tilde{b}) = \operatorname{span}\{\tilde{b}, \tilde{A}\tilde{b}, \ldots, \tilde{A}^{k-1}\tilde{b}\} = \operatorname{span}\{Qb, QAb, \ldots, QA^{k-1}b\} = \\ Q\operatorname{span}\{b, Ab, \ldots, A^{k-1}b\} = Q\mathcal{K}_k(A, b).$$

Moreover

$$\tilde{A}\mathcal{K}_k(\tilde{A}, \tilde{b}) = QA\mathcal{K}_k(A, b).$$

Thus the conditions for the k-th iterate of the GMRES method for $\tilde{A}x = \tilde{b}$ can be rewritten as

$$\tilde{x}^{(k)} \in Q\mathcal{K}_k(A, b) \qquad \text{and} \qquad Qb - QAQ^T\tilde{x}^{(k)} \perp QA\mathcal{K}_k(A, b).$$

Moreover, as Q is orthogonal, the second condition is equivalent to

$$b - AQ^T\tilde{x}^{(k)} \perp A\mathcal{K}_k(A, b).$$

Now it immediately follows from the characterisation of $x^{(k)}$ that $Qx^{(k)}$ satisfies these properties, which shows that $Qx^{(k)} = \tilde{x}^{(k)}$.

Problem 5 We are given a linear system $Ax = b$ and some initial guess $x_0 \in \mathbb{R}^n$ of its solution. Denote by x^{SD} the result of one step of the steepest descent method starting with x_0, and by x^{MR} the result of one step of the MR iteration starting with x_0. Assume now that $x^{SD} = x^{MR}$.

Show that in this case the initial residual $b - Ax_0$ is an eigenvector of A. Conclude that $x^{SD} = x^{MR}$ solves the system $Ax = b$.

Denote by $r_0 = b - Ax_0$ the residual at the initial guess x_0 and assume without loss of generality that $r_0 \neq 0$ (else x_0 is a solution and $x_0 = x^{SD} = x^{MR}$). Then

$$x^{SD} = x_0 + \alpha^{SD} r_0 \qquad \text{with} \qquad \alpha^{SD} = \frac{\|r_0\|_2^2}{(r_0, Ar_0)},$$

and

$$x^{MR} = x_0 + \alpha^{MR} r_0 \qquad \text{with} \qquad \alpha^{MR} = \frac{(r_0, Ar_0)}{\|Ar_0\|_2^2}.$$

Thus $x^{SD} = x^{MR}$ if and only if $\alpha^{SD} = \alpha^{MR}$, which in turn is equivalent to

$$\frac{\|r_0\|_2^2}{(r_0, Ar_0)} = \frac{(r_0, Ar_0)}{\|Ar_0\|_2^2}$$

or

$$\|r_0\|_2^2 \|Ar_0\|_2^2 = (r_0, Ar_0)^2.$$

That is, the Cauchy–Schwarz inequality actually becomes an equality for the vectors r_0 and Ar_0, which is the case if and only if r_0 and Ar_0 are parallel, that is, $Ar_0 = \lambda r_0$ for some $\lambda \in \mathbb{R}$. In other words, r_0 is an eigenvector of A for the eigenvalue λ. This, however, implies that

$$Ax^{SD} = Ax_0 + \alpha^{SD} Ar_0 = b - r_0 + \frac{\|r_0\|_2^2}{(r_0, Ar_0)} \lambda r_0 = b - r_0 + \frac{\|r_0\|_2^2}{\lambda (r_0, r_0)} \lambda r_0 = b,$$

that is, x^{SD} (and therfore also x^{MR}) solves the equation $Ax = b$.

Problem 6 Consider the matrix

$$A = \begin{pmatrix} 1 & 0.2 & 1 & -0.2 \\ 0.5 & 1.1 & 0.5 & -1.1 \end{pmatrix}.$$

a) Compute the reduced singular value decomposition of the matrix A.

We start by computing the eigenvalues of the matrix

$$AA^T = \begin{pmatrix} 2.08 & 1.44 \\ 1.44 & 2.92 \end{pmatrix}.$$

Its characteristic polynomial is

$$p(\lambda) = (2.08 - \lambda)(2.92 - \lambda) - 1.44^2 = \lambda^2 - 5\lambda + 4,$$

which has the zeros 4 and 1. Thus the singular values of A are the values $\sigma_2 = 2$ and $\sigma_1 = 1$.

Next we note that, if $A = U\Sigma V^T$, then $AA^T = U\Sigma^2 U^T$. That is, the matrix U contains the normalised eigenvectors of AA^T. We note that the vector $(4, -3)^T$ is an eigenvector of AA^T for the eigenvalue 1. Thus we can set $u_2 = (4, -3)^T/5$. Since the eigenvectors of AA^T are orthogonal, we obtain as possible first eigenvector the vector $u_1 = (3, 4)^T/5$. Thus we may choose

$$U = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}.$$

Moreover, we have

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

Now the equation $A = U\Sigma V^T$ implies that $V^T = \Sigma^{-1}U^T A$, that is,

$$V^T = \frac{1}{10}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 & 0.2 & 1 & -0.2 \\ 0.5 & 1.1 & 0.5 & -1.1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{pmatrix}.$$

Thus we have the singular value decomposition

$$A = U\Sigma V^T = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} 1 & 1 & 1 & -1 \\ 1 & -1 & 1 & 1 \end{pmatrix}.$$

b) Use the singular value decomposition in order to compute the solution of the problem

$$\min_{x \in \mathbb{R}^4} \|x\|_2^2 \qquad \text{s.t. } Ax = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

The solution of this optimisation problem is given by $x = A^\dagger b$ with $b = (1, -2)^T$ and $A^\dagger = V\Sigma^{-1}U^T$ the pseudoinverse of A. We first compute

$$U^T b = \frac{1}{5}\begin{pmatrix} 3 & 4 \\ 4 & -3 \end{pmatrix}\begin{pmatrix} 1 \\ -2 \end{pmatrix} = \frac{1}{5}\begin{pmatrix} -5 \\ 10 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$$

Next,

$$\Sigma^{-1}U^T b = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} -1 \\ 2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} -1 \\ 4 \end{pmatrix}.$$

Finally,

$$x = V\Sigma^{-1}U^T b = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 1 & 1 \\ -1 & 1 \end{pmatrix}\frac{1}{2}\begin{pmatrix} -1 \\ 4 \end{pmatrix} = \frac{1}{4}\begin{pmatrix} 3 \\ -5 \\ 3 \\ 5 \end{pmatrix}.$$