SF2524 Matrix Computations for Large-scale Systems
Exam

Aids: None **Time: Four hours**

Grades: E: 16 points, D: 19 points, C: 22 points, B: 25 points, A: 28 points (out of the possible 35 points, including bonus points from homeworks).

Problem 1 (3p) Several iterative methods for linear systems of equations in this course generate iterates $x_1, x_2, \ldots \in \mathbb{R}^m$ which satisfy

$$\min_{x \in S_n} \|Ax - b\|_Z = \|Ax_n - b\|_Z$$

for some norm $\|\cdot\|_Z$ and space S_n. What is S_n and $\|\cdot\|_Z$ when the iterates $x_1, x_2, \ldots$ are generated by (a) GMRES, (b) CG, (c) CGN?

Problem 2 (5p)

(a) Prove that the result of one step of the shifted QR-method for a symmetric matrix is a symmetric matrix if the shift is real.

(b) What is the result of one step of the basic QR-method for the matrix $A = \begin{bmatrix} 4 & 0 \\ 3 & 0 \end{bmatrix}$?
Hint: You may want to show that the Q-matrix the QR-factorization of a two-by-two matrix is a Givens rotator $Q = \frac{1}{\sqrt{c^2+s^2}} \begin{bmatrix} c & -s \\ s & c \end{bmatrix}$

(c) Let $\tilde{A}$ be the result of one step of the QR-method applied to A. Derive a closed formula for the QR-method applied to the matrix $B = \begin{bmatrix} A & C \\ 0 & R \end{bmatrix}$ where the matrix R is upper triangular.

Problem 3 (4p) (a) Given an Arnoldi factorization $AQ_n = Q_{n+1}\underline{H}_n$, how are the Krylov approximations for matrix functions (of $f(A)b$) generated? (No derivation required.)

(b) Breakdown of the Arnoldi method corresponds to the case that $h_{n+1,n} = 0$. Prove that the Krylov method for matrix functions generates an exact result if this occurs (which means we have no approximation error). You may assume f is an entire function.

Problem 4 (4p)

(a) Suppose an Arnoldi factorization $AQ_n = Q_{n+1}\underline{H}_n$ is given, where $q_1 = b/\|b\|$. How is the GMRES approximation for $Ax = b$ computed from the Arnoldi relation?

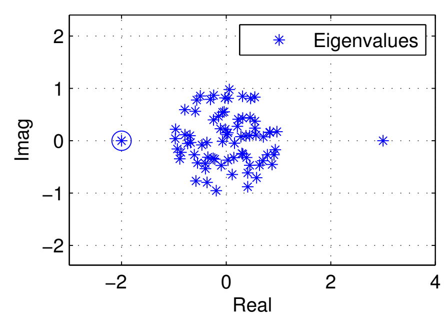
(b) Derive a closed formula for the approximation generated by one step ($n = 1$) of GMRES, only involving b and A.

Problem 5 (2p) Roxanne the rocket scientist needs to determine the trajectory of her space craft by solving a linear system of equations $Ax = b$, where A is a huge symmetric positive definite matrix. She knows that GMRES and CG have the same convergence factor $\rho = 0.1$ for her particular problem. She also knows that a matrix vector product takes 1 hour, a scalar product of two vectors takes 20 minutes, and the adding a linear combination of two vectors takes 10 minutes. What is the computation time for GMRES and CG to achieve full precision (10^{-16})? Which method should she select? Provide clear justifications of your reasoning and simplifications.

Problem 6 (5p) A theorem in this course states that the error indicator in Arnoldi's method for eigenvalue problems can (under appropriate conditions) be bounded as

$$\|(I - Q_n Q_n^T)x_j\| \leq \alpha \min_{\substack{p \in P_{n-1} \\ p(\lambda_j)=1}} \max_{i \neq j} |p(\lambda_i)|. \qquad (\star)$$

The eigenvalues of the matrix $A \in \mathbb{R}^{n \times n}$ are given in the figure to the right.



(a) The eigenvalue λ_1 is marked with a circle in the figure. Use ($\star$) to determine a convergence factor γ such that

$$\|(I - Q_n Q_n^T)x_1\| \leq \alpha \gamma^{n-1}$$

where $\gamma < 1$. Describe clearly what you identify in the figure.

(b) Consider the following generalization of the figure. Suppose eigenvalues $\lambda_3, \ldots, \lambda_n \in D(\rho, 0)$ and suppose $|\lambda_1| > |\lambda_2| > \rho$. Derive a formula for $\tilde{\gamma}$ such that

$$\|(I - QQ_n^T)x_1\| \leq \beta \tilde{\gamma}^n$$

such that we always have $\tilde{\gamma} < 1$.

Problem 7 (6p) (a) Compute $f(A)$ with the (simplified) Schur-Parlett method when $a < b < 10$ when

$$A = \begin{bmatrix} a & 1 & 0 \\ & b & 1 \\ & & 10 \end{bmatrix}$$

(b) What is $f(B)$ when $B = \begin{bmatrix} a & 1 \\ 0 & a \end{bmatrix}$ for some constant a?

(c) Suppose $B \in \mathbb{R}^{2\times 2}$, $c \in \mathbb{R}^2$ and $d \in \mathbb{R}$. Derive a formula for $f_c \in \mathbb{R}^2$ such

$$f(A) = \begin{bmatrix} f(B) & f_c \\ 0 & f(d) \end{bmatrix} \text{ when } A = \begin{bmatrix} B & c \\ 0 & d \end{bmatrix}.$$

The formula should be a linear system expressed in terms of $f(B)$, $f(d)$, c, d.

(d) The (simplified) Schur-Parlett method will fail for the matrix in (a) if $b = a$. Use (b)-(c) to derive formula for $f(A)$ when $b = a$. If you encounter a 2×2 linear system of equations, you do not need to explicitly solve it.