WBMA063-05 – Numerical Linear Algebra
Mock Exam
2025-2026

Instructions:

1. **Write your name and student number of the top of each sheet of writing paper!**
2. Use the writing (lined) and scratch (blank) paper provided, raise your hand if you need more paper.
3. Start each question on a new page.

This exam consists of 4 questions for a total of 90 points. 10 points are free.

1. Let $A \in \mathbb{R}^{n\times n}$ be a non-singular matrix.
   (a) (4 points) Define the LU decomposition of A. Also define the pivoted LU decomposition of A.
   (b) (8 points) Assume the LU decomposition of $A = LU$ exists. Show that the dominant computational cost of computing L and U is $\mathcal{O}(\frac{2}{3}n^3)$ flops.
   (c) (5 points) Suppose you have computed a pivoted LU decomposition of A, $PA = LU$. Describe an efficient algorithm to explicitly compute the inverse A^{-1}. What is the computational cost of your algorithm?
   Hint: consider building A^{-1} one column at a time.
   (d) (10 points) Let $A = A^T \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix. Suppose we computed a Cholesky decomposition of A in floating point arithmetic with a backward stable algorithm. Then, still in floating poitn arithmetic, we use the computed Cholesky factors to solve the linear system $A\boldsymbol{x} = \boldsymbol{b}$. Show the computed solution $\hat{\boldsymbol{x}}$ is backward stable. You may use the fact that a triangular solve algorithm is backward stable.
   (e) (4 points) An alternative approach to solve symmetric positive definite linear systems is to use Conjugate Gradients (CG). Reflect on the different contexts in which you might prefer a Cholesky-based approach to CG, and on the different contexts in which you might prefer CG to Cholesky. Explain your answer.

2. Let $A \in \mathbb{R}^{m\times n}$ with $m > n$ be a matrix with singular value decomposition $A = U\Sigma V^T$, where Σ is square.

   (a) (3 points) For each of the factors U, Σ, and V in the singular value decomposition, provide their dimension and their special structure.

   (b) (3 points) Assume A has full column rank. Express the solution of the least squares problem

   $$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2$$

   in terms of the singular value decomposition factors. You need not prove the result.

   (c) (8 points) Again assume A has full rank. Show that

   $$\|A\boldsymbol{x}^* - \boldsymbol{b}\|_2 = \|U_\perp^T \boldsymbol{b}\|_2,$$

   where $U_\perp \in \mathbb{R}^{m\times(m-n)}$ is the orthogonal complement of the left singular vectors U. That is, $U_\perp \in \mathbb{R}^{m\times(m-n)}$ is such that

   $$U_{\mathrm{F}} = \begin{bmatrix} U & U_\perp \end{bmatrix}$$

   is a square orthogonal matrix.

   (d) (4 points) Computing the solution to a least squares problem

   $$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2$$

   via the singular value decomposition of A is backward stable. Suppose this algorithm is used in floating point arithmetic to compute the solution $\hat{\boldsymbol{x}}$ to the least squares problem. What does it mean for $\hat{\boldsymbol{x}}$ to be backward stable? Provide a mathematical expression.

   (e) (3 points) Suppose $\hat{\boldsymbol{x}}$ is a backward stable solution to

   $$\min_{\boldsymbol{x}\in\mathbb{R}^n} \|A\boldsymbol{x} - \boldsymbol{b}\|_2,$$

   which has exact solution $\boldsymbol{x}^*$. What can you say about the size of

   $$\frac{\|\hat{\boldsymbol{x}} - \boldsymbol{x}^*\|}{\|\boldsymbol{x}^*\|}?$$

   Is it always, sometimes, or never close to machine precision? Support your reasoning.

3. Let $A \in \mathbb{R}^{n \times n}$ be diagonalizable with eigenpairs $(\lambda_i, \boldsymbol{v}_i)$ where all $\boldsymbol{v}_i$ have unit norm. Assume the eigenvalues are all real and are ordered as

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > 0.$$

The power method starts by initializing a vector $\boldsymbol{x}_0 \in \mathbb{R}^n$ with unit norm, and then iterating as:

$$\begin{aligned} &\text{For } k = 0, 1, 2, \ldots \\ &\quad \boldsymbol{y}_k = A\boldsymbol{x}_k \\ &\quad \boldsymbol{x}_{k+1} = \boldsymbol{y}_k / \|\boldsymbol{y}_k\|_2. \end{aligned}$$

(a) (6 points) Expand $\boldsymbol{x}_0$ in the basis formed by the eigenvectors $\boldsymbol{v}_i$:

$$\boldsymbol{x}_0 = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i.$$

Show

$$\boldsymbol{x}_k = \text{sign}(\lambda_1)^k \frac{\alpha_1 \boldsymbol{v}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{v}_i}{\|\alpha_1 \boldsymbol{v}_1 + \sum\limits_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1}\right)^k \boldsymbol{v}_i\|_2}.$$

(b) (4 points) Explain what happens to $\boldsymbol{x}_k$ as k tends to infinity. Additionally, explain how the dominant eigenvalue λ_1 can be approximated from $\boldsymbol{x}_k$.

(c) (6 points) Assume A is symmetric. Show

$$|\boldsymbol{x}_k^* A \boldsymbol{x}_k - \lambda_1| \approx \left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right),$$

using the simplification that, as $k \to \infty$,

$$\boldsymbol{x}_k \approx \text{sign}(\lambda_1)^k \frac{\alpha_1 \boldsymbol{v}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{v}_2}{\|\alpha_1 \boldsymbol{v}_1 + \alpha_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k \boldsymbol{v}_2\|_2}.$$

(d) (3 points) How would you adapt the algorithm to instead find the eigenvector corresponding to the eigenvalue of A closest to -2? Explain your reasoning.

4. Consider a non-singular diagonalizable matrix $A \in \mathbb{R}^{n \times n}$.

   (a) (4 points) The k-dimensional Krylov subspace associated to the linear system $Ax = b$ is

   $$\mathcal{K}_k(A, \boldsymbol{b}) = \operatorname{span}\left\{\boldsymbol{b}, A\boldsymbol{b}, A^2\boldsymbol{b}, \dots, A^{k-1}\boldsymbol{b}\right\}.$$

   Define a matrix $K_k \in \mathbb{R}^{n \times k}$ whose columns are the natural basis vectors for the Krylov subspace normalized; that is,

   $$K_k = \begin{bmatrix} \frac{\boldsymbol{b}}{\|\boldsymbol{b}\|_2} & \frac{A\boldsymbol{b}}{\|A\boldsymbol{b}\|_2} & \frac{A^2\boldsymbol{b}}{\|A^2\boldsymbol{b}\|_2} & \cdots & \frac{A^{k-1}b}{\|A^{k-1}\boldsymbol{b}\|_2} \end{bmatrix}.$$

   Qualitatively, explain why the matrix K_k becomes increasingly ill-conditioned as k increases.

   (b) (8 points) It is instead advised to use the Arnoldi algorithm to build a basis for the Krylov subspace $\mathcal{K}_k(A, \boldsymbol{b})$. This algorithm computes an orthogonal basis $\{\boldsymbol{q}_j\}$ for the Krylov subspace:

   $$\mathcal{K}_k(A, \boldsymbol{b}) = \operatorname{span}\{\boldsymbol{q}_1, \boldsymbol{q}_2, \dots, \boldsymbol{q}_k\}, \quad \boldsymbol{q}_i^T \boldsymbol{q}_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

   The matrix

   $$Q_k = \begin{bmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_k \end{bmatrix}$$

   is then well-conditioned. The Arnoldi algorithm (in exact arithmetic) can be written as

   $$\begin{aligned} &\boldsymbol{q}_1 = \boldsymbol{b}/\|\boldsymbol{b}\|_2 \\ &\text{For } j = 1, 2, \dots, k \\ &\qquad \boldsymbol{v} = A\boldsymbol{q}_j \\ &\qquad \boldsymbol{w} = \boldsymbol{v} - \sum_{i=1}^{j} (\boldsymbol{q}_i^T \boldsymbol{v}) \boldsymbol{q}_i \\ &\qquad \boldsymbol{q}_{j+1} = \boldsymbol{w}/\|\boldsymbol{w}\|_2 \end{aligned}$$

   This is not how the algorithm is implemented in floating point arithmetic. Describe the algorithm as it is implemented. Then, show that after k steps, we have the identity

   $$AQ_k = Q_{k+1}\tilde{H}_k,$$

   for a $(k+1) \times k$ matrix $\tilde{H}_k$. Describe the structure of this matrix.

   (c) (4 points) Let the eigendecomposition of A be given by $A = X\Lambda X^{-1}$. The residual at the kth GMRES iterate $\boldsymbol{x}_k$ can be bounded by

   $$\frac{\|\boldsymbol{b} - A\boldsymbol{x}_k\|_2}{\|\boldsymbol{b}\|_2} \leq \kappa(X) \min_{p \in \mathcal{P}_k} \max_{\lambda \in \Lambda(A)} |p(\lambda)|,$$

where we assumed $\boldsymbol{x}_0 = \mathbf{0}$, $\mathcal{P}_k$ is a subset of all degree k polynomials p with $p(0) = 1$, and $\Lambda(A)$ is the set of all eigenvalues of A. Comment on the expected convergence behaviour based on the eigenvalues of A. Distinguish between the case when A is normal and when A is not normal.
Hint: A is normal when $A^T A = AA^T$. The eigenvectors of a normal matrix form an orthogonal basis.

(d) (3 points) Let A be symmetric matrix with $m < n$ distinct eigenvalues. Will GMRES converge to the exact solution in at most m iterations? Explain your reasoning.