

Optimization / Numerical Linear Algebra (ONLA)

INSTRUCTIONS FOR QUALIFYING EXAMS

Start each problem on a new sheet of paper. Write your university identification number at the top of each sheet of paper. **DO NOT WRITE YOUR NAME!**

Complete this sheet and staple to your answers. Read the directions of the exam carefully.

STUDENT ID NUMBER: _____

DATE: _____

EXAMINEES: DO NOT WRITE BELOW THIS LINE

1. _____

5. _____

2. _____

6. _____

3. _____

7. _____

4. _____

8. _____

Pass/fail recommend on this form.

Total score: _____

DO NOT FORGET TO WRITE YOUR SID NO. ON YOUR EXAM. PLEASE USE BLANK PAGES AT END FOR ADDITIONAL SPACE.

1. (10 points) Let $A \in \mathbb{R}^{n \times n}$ be symmetric positive definite, with extreme eigenvalues $0 < \lambda_{\min} \leq \lambda_{\max}$, and let $b \in \mathbb{R}^n$. Consider the stationary Richardson iteration

$$x^{(k+1)} = x^{(k)} + \omega(b - Ax^{(k)}), \quad \omega \in \mathbb{R}.$$

Determine all values of ω for which the iteration converges to the solution of $Ax = b$ for every initial guess. Among those values, determine the choice of ω that minimizes the asymptotic convergence factor, and give the resulting factor. Justify your answer.

OPTIMIZATION / NUMERICAL LINEAR ALGEBRA (ONLA)

2. (10 points) Let $A \in \mathbb{C}^{n \times n}$ and $v \in \mathbb{C}^n$ with $v \neq 0$. For $k \geq 1$ and any $z \in \mathbb{C}^n$, define

$$\mathcal{K}_k(A, z) = \text{span}\{z, Az, \dots, A^{k-1}z\}.$$

For each nonnegative integer d , let $\mathcal{P}_d$ denote the complex polynomials of degree at most d . Prove the following properties. Each part is worth 2 points. A^* denotes the conjugate transpose and orthogonality uses the standard Hermitian inner product.

(a) $\mathcal{K}_k(A, v) = \{p(A)v : p \in \mathcal{P}_{k-1}\}$.

(b) For every $u \in \mathbb{C}^n$,

$$\mathcal{K}_k(A, u + v) \subseteq \mathcal{K}_k(A, u) + \mathcal{K}_k(A, v).$$

(c) For every nonnegative integer d and every $p \in \mathcal{P}_d$, $p(A)\mathcal{K}_k(A, v) \subseteq \mathcal{K}_{k+d}(A, v)$.

(d) For every $w \in \mathbb{C}^n$, $w \perp \mathcal{K}_k(A, v)$ if and only if $v \perp \mathcal{K}_k(A^*, w)$.

(e) Let $n = n_1 + n_2$ and suppose

$$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}, \quad v = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix},$$

where $A_i \in \mathbb{C}^{n_i \times n_i}$ and $v_i \in \mathbb{C}^{n_i}$. If $\pi_i : \mathbb{C}^{n_1} \oplus \mathbb{C}^{n_2} \rightarrow \mathbb{C}^{n_i}$ is coordinate extraction from the i th block, prove that $\pi_i(\mathcal{K}_k(A, v)) = \mathcal{K}_k(A_i, v_i)$ for $i = 1, 2$.

OPTIMIZATION / NUMERICAL LINEAR ALGEBRA (ONLA)

3. (10 points) Let $A \in \mathbb{C}^{n \times n}$ be normal and suppose that its spectrum is contained in a disk that does not contain the origin:

$$\sigma(A) \subseteq \{z \in \mathbb{C} : |z - c| \leq R\}, \quad 0 < R < |c|.$$

In exact arithmetic, apply unrestarted GMRES to $Ax = b$ with $b \neq 0$ and $x_0 = 0$. Recall that at iteration k , GMRES chooses x_k from the affine space $x_0 + \text{span}\{r_0, Ar_0, \dots, A^{k-1}r_0\}$ to minimize $\|b - Ax\|_2$ over that space. Writing $r_k = b - Ax_k$, prove that

$$\frac{\|r_k\|_2}{\|r_0\|_2} \leq \left(\frac{R}{|c|}\right)^k.$$

Then give an explicit sufficient number of iterations for achieving $\|r_k\|_2/\|r_0\|_2 \leq \varepsilon$, where $0 < \varepsilon < 1$.

Hint. Use the affine map that sends 0 to 1 and the disk center c to 0 to construct an admissible degree- k residual polynomial. The residual estimate applies at each iteration before termination; if GMRES terminates earlier, the tolerance is already attained.

4. (10 points) Consider the function $f: \mathbb{R}^n \rightarrow \mathbb{R}$ defined by

$$f(x) = \max_{1 \leq i \leq m} \{a_i^\top x + b_i\}, \quad a_i \in \mathbb{R}^n, \quad b_i \in \mathbb{R}.$$

For $x \in \mathbb{R}^n$, define the active set

$$I(x) = \{i \in \{1, \dots, m\} : a_i^\top x + b_i = f(x)\}.$$

(a) [2 points] Prove that f is convex.

(b) [5 points] Show that

$$\partial f(x) = \text{conv}\{a_i : i \in I(x)\}.$$

(c) [3 points] Show that x^* minimizes f over $\mathbb{R}^n$ if and only if

$$0 \in \text{conv}\{a_i : i \in I(x^*)\}.$$

OPTIMIZATION / NUMERICAL LINEAR ALGEBRA (ONLA)

5. (10 points) Suppose $f: \mathbb{R}^n \rightarrow \mathbb{R}$ is differentiable, μ -strongly convex, and has L -Lipschitz continuous gradient, where $0 < \mu \leq L$. Let x^* denote its unique minimizer. Consider the gradient descent iteration

$$x_{k+1} = x_k - \alpha \nabla f(x_k),$$

where $\alpha > 0$.

- (a) [3 points] Show that for all $x, y \in \mathbb{R}^n$,

$$\langle \nabla f(x) - \nabla f(y), x - y \rangle \geq \mu \|x - y\|_2^2.$$

- (b) [4 points] Show that

$$\|x_{k+1} - x^*\|_2^2 \leq (1 - 2\alpha\mu + \alpha^2 L^2) \|x_k - x^*\|_2^2.$$

- (c) [3 points] Give the range of step sizes $\alpha > 0$ for which the bound in part (b) guarantees linear convergence of x_k to x^* . State the resulting contraction factor for $\|x_k - x^*\|_2$, and determine the step size that minimizes this factor.

6. (10 points) Consider the optimization problem

$$\min_{x, z \in \mathbb{R}^n} \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|z\|_1 \quad \text{subject to } x = z,$$

where $A \in \mathbb{R}^{m \times n}$, $b \in \mathbb{R}^m$, and $\lambda > 0$.

The scaled ADMM iterations with parameter $\rho > 0$ are

$$\begin{aligned} x^{k+1} &= \arg \min_x \left\{ \frac{1}{2} \|Ax - b\|_2^2 + \frac{\rho}{2} \|x - z^k + u^k\|_2^2 \right\}, \\ z^{k+1} &= \arg \min_z \left\{ \lambda \|z\|_1 + \frac{\rho}{2} \|x^{k+1} - z + u^k\|_2^2 \right\}, \\ u^{k+1} &= u^k + x^{k+1} - z^{k+1}. \end{aligned}$$

(a) [3 points] Derive an explicit formula for x^{k+1} .

(b) [4 points] Show that

$$z^{k+1} = S_{\lambda/\rho}(x^{k+1} + u^k),$$

where the soft-thresholding operator $S_\tau : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is defined componentwise by

$$[S_\tau(v)]_i = \text{sign}(v_i) \max\{|v_i| - \tau, 0\}.$$

(c) [3 points] Suppose that the iterates converge:

$$x^k \rightarrow x^*, \quad z^k \rightarrow z^*, \quad u^k \rightarrow u^*.$$

Show directly from the ADMM equations that $x^* = z^*$, and that x^* satisfies the first-order optimality condition for

$$\min_x \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1.$$

7. (10 points) Let $A, B \in \mathbb{R}^{n \times n}$, and let

$$\sigma_1(A) \geq \cdots \geq \sigma_n(A), \quad \sigma_1(B) \geq \cdots \geq \sigma_n(B)$$

denote their singular values.

(a) Prove that for each $i = 1, \dots, n$,

$$\sigma_i(AB) \leq \sigma_i(A)\sigma_1(B).$$

(b) Prove that for each $i = 1, \dots, n$,

$$\sigma_i(AB) \geq \sigma_i(A)\sigma_n(B).$$

OPTIMIZATION / NUMERICAL LINEAR ALGEBRA (ONLA)

8. (10 points) Let $A \in \mathbb{R}^{n \times n}$ and $b \in \mathbb{R}^n$ with $\|b\|_2 = 1$, and let $3 \leq k \leq n - 1$. Suppose that k steps of Arnoldi starting from b are performed without breakdown, giving $AQ_k = Q_{k+1}\tilde{H}$, where $Q_k = [q_1, \dots, q_k]$, $Q_{k+1} = [Q_k, q_{k+1}]$ has orthonormal columns, $q_1 = b$, and $\tilde{H} \in \mathbb{R}^{(k+1) \times k}$ is upper Hessenberg with entries $h_{i,j}$.

(a) Let $v = Q_k c$ for some $c \in \mathbb{R}^k$. Show that one can write

$$v = p_{k-1}(A)b$$

for some polynomial p_{k-1} of degree at most $k - 1$.

(b) Let $v = Ab$. Show that $v = Q_k c$ for some vector c (which you should identify using $h_{i,j}$).

(c) Let $v = A^2 b$. Find a vector c such that $v = Q_k c$.

OPTIMIZATION / NUMERICAL LINEAR ALGEBRA (ONLA)

9. (10 points) Let $A \in \mathbb{R}^{n \times n}$. Starting with $A_1 = A$, the unshifted QR algorithm computes a QR factorization $A_k = Q_k R_k$, where Q_k is orthogonal and R_k is upper triangular, and sets $A_{k+1} = R_k Q_k$. Define

$$Q^{(k)} = Q_1 Q_2 \cdots Q_k, \quad R^{(k)} = R_k R_{k-1} \cdots R_1.$$

- (a) Prove that

$$A^k = Q^{(k)} R^{(k)}$$

is the QR factorization of A^k , and that

$$A_{k+1} = (Q^{(k)})^T A Q^{(k)}.$$

- (b) Deduce that all A_k are similar to A , and hence have the same eigenvalues.
(c) Explain briefly why QR iteration can be viewed as a simultaneous power method.
(d) Suppose that $A \in \mathbb{R}^{n \times n}$ is symmetric and invertible. Show that

$$A^{-k} = (R^{(k)})^{-1} (Q^{(k)})^T.$$

By taking transposes, deduce that

$$A^{-k} = Q^{(k)} (R^{(k)})^{-T}.$$

What does this imply about the final column of $Q^{(k)}$? (The final column is its n -th column $q_n^{(k)} = Q^{(k)} e_n$, where $e_n = [0, \dots, 0, 1]^T$). Relate your answer to the power method applied to a suitable initial vector.