Qualifying Exam, Fall 2025

Optimization / Numerical Linear Algebra (ONLA)

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

---

1. (10 points) Let $A = [a_{ij}] \in \mathbb{R}^{n\times n}$ and use the standard splitting

$$A = D + L + U,$$

where $D = \operatorname{diag}(a_{11}, \ldots, a_{nn})$, L is strictly lower triangular with $L_{ij} = a_{ij}$ for $i > j$ (and 0 otherwise), and U is strictly upper triangular with $U_{ij} = a_{ij}$ for $i < j$ (and 0 otherwise). Assume A is *strictly diagonally dominant (SDD) by rows*:

$$|a_{ii}| \ > \sum_{j\neq i} |a_{ij}| \qquad (i = 1, \ldots, n).$$

Consider the Jacobi iteration for $Ax = b$,

$$x^{(k+1)} \ = \ -D^{-1}(L+U)\,x^{(k)} \ + \ D^{-1}b, \qquad G_J \ := \ I - D^{-1}A.$$

Prove that for every initial guess $x^{(0)} \in \mathbb{R}^n$ the iterates converge to the unique solution $x^\star$ of $Ax = b$.

2. (10 points) Let $A \in \mathbb{R}^{n\times n}$ and $v \in \mathbb{R}^n$, $v \neq 0$. For an integer $k \geq 1$, define

$$\mathcal{K}_k(A, v) \ := \ \operatorname{span}\{\, v,\ Av,\ A^2v,\ \ldots,\ A^{k-1}v\,\}.$$

Prove the following:

(3 points) (a) Scaling invariance. For any nonzero scalars $\alpha, \beta \in \mathbb{R}$,

$$\mathcal{K}_k(A, v) \ = \ \mathcal{K}_k(\alpha A,\ \beta v).$$

(3 points) (b) Nestedness. For every $k \geq 1$,

$$\mathcal{K}_k(A, v) \ \subseteq \ \mathcal{K}_{k+1}(A, v).$$

(4 points) (c) Stabilization. If there exists $m \geq 1$ with $\dim \mathcal{K}_{m+1}(A, v) = \dim \mathcal{K}_m(A, v)$, then

$$\mathcal{K}_k(A, v) = \mathcal{K}_m(A, v) \quad \text{for all } k \geq m.$$

3. (10 points) Let $A \in \mathbb{R}^{n\times n}$ and $b \in \mathbb{R}^n$, and define

$$\mathcal{K}_k(A,b) \ := \ \operatorname{span}\{\, b, Ab, \ldots, A^{k-1}b\,\} \qquad (k \ge 1).$$

Assume *exact arithmetic* and that the linear system $Ax = b$ is consistent (so $b \in \operatorname{range}(A)$). Suppose furthermore that $\dim \mathcal{K}_n(A,b) = n$ (i.e., $\mathcal{K}_n(A,b) = \mathbb{R}^n$). GMRES started from $x_0 = 0$ selects $x_k \in \mathcal{K}_k(A,b)$ that minimizes $\|r_k\|_2 = \|b - Ax_k\|_2$. Prove that GMRES produces an exact solution in at most n steps; in particular $r_n = 0$.

4. (10 points) Consider the constrained optimization problem

$$\text{minimize } f(x_1, x_2) = x_1^3 + 4x_2^2$$

subject to

$$g(x_1, x_2) = x_1^2 - x_2 - 1 \leq 0.$$

(a) First-order conditions (KKT):

- Write the Lagrangian for minimization and maximization.
- List all KKT candidates (for both minimization and maximization problems).

(b) Second-order conditions and classification:

- Check second-order conditions (necessary and sufficient) on the relevant tangent spaces.
- Identify which KKT points are strict local minima, strict local maxima, or saddle points.

(c) Global behavior:

- Prove or disprove that the problem attains the global minimum.
- Prove or disprove that the problem attains the global maximum.

5. (10 points) Consider the optimization problem

$$\begin{aligned} \text{minimize} \quad & f(x_1, x_2) = x_1^4 + 2x_2^2 - 3x_2 \\ \text{subject to} \quad & c(x_1, x_2) = x_1^2 + (x_2 + 1)^2 \ \leq \ 2. \end{aligned}$$

(a) Formulate the Lagrangian and derive the dual function. Include a description of the domain of the dual function.

(b) Solve the dual problem $\max_{\lambda \geq 0} g(\lambda)$.

(c) Use the dual solution to recover the primal minimizer $(x_1^\star, x_2^\star)$, and verify primal feasibility, complementary slackness, and strong duality.

Hint: Show first that the dual separates into a minimization in x_1 and in x_2.

6. (10 points) Consider the optimization problem

$$\min_{x\in\mathbb{R}^n} F(x) := f(x) + g(x),$$

where: $f : \mathbb{R}^n \to \mathbb{R}$ is smooth and convex with L-Lipschitz continuous gradient, $\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|$, and $g : \mathbb{R}^n \to \mathbb{R} \cup \{+\infty\}$ is proper, closed, convex, possibly nonsmooth.

The *proximal gradient method* with fixed step size $\alpha > 0$ is defined as

$$x^{k+1} = \operatorname{prox}_{\alpha g}\left(x^k - \alpha\nabla f(x^k)\right), \quad k = 0, 1, 2, \ldots,$$

where the proximal operator is

$$\operatorname{prox}_{\alpha g}(v) = \arg\min_{x\in\mathbb{R}^n}\left\{g(x) + \frac{1}{2\alpha}\|x - v\|^2\right\}.$$

(a) Firm nonexpansiveness of the proximal operator: Show that for any $u, v \in \mathbb{R}^n$ and convex g, the proximal operator satisfies the *firm nonexpansiveness property*:

$$\|\operatorname{prox}_{\alpha g}(u) - \operatorname{prox}_{\alpha g}(v)\|^2 \le \langle u - v, \operatorname{prox}_{\alpha g}(u) - \operatorname{prox}_{\alpha g}(v)\rangle.$$

(b) Descent lemma for proximal gradient: Show that if $0 < \alpha \le \frac{1}{L}$, the sequence $\{F(x^k)\}$ generated by the proximal gradient method monotonically decreases:

$$F(x^{k+1}) \le F(x^k) - \frac{1 - \alpha L}{2\alpha}\|x^{k+1} - x^k\|^2.$$

(c) Convergence to a minimizer: Suppose that F has at least one minimizer $x^\star$ and that the sub-level set $\{x : F(x) \le F(x^0)\}$ is bounded. Fix $\alpha \in (0, 1/L]$. Show that every *limit point* of the sequence $\{x^k\}$ is a minimizer of F.

Hint: Use the optimality condition of the proximal operator: $\frac{1}{\alpha}(v - \operatorname{prox}_{\alpha g}(v)) \in \partial g(\operatorname{prox}_{\alpha g}(v))$.

Apply the *descent lemma* for L-smooth functions: $f(y) \le f(x) + \langle\nabla f(x), y - x\rangle + \frac{L}{2}\|y - x\|^2$.

7. (10 points) Let $A \in \mathbb{R}^{n\times n}$, and denote by $A = U\Sigma V^T$ its SVD, where U, V are orthogonal and $\Sigma = \operatorname{diag}(\sigma_1, \ldots, \sigma_n)$, with $\sigma_1 \geqslant \sigma_2 \geqslant \cdots \geqslant \sigma_n \geqslant 0$. Let $\|A\|_2 = \sigma_1$ denote the spectral norm.

   (a) Let $A \in \mathbb{R}^{n\times n}$ be a rank-1 matrix $A = xy^T$ for $x, y \in \mathbb{R}^n$. Show that there exist Householder matrices $H_1, H_2 \in \mathbb{R}^{n\times n}$ (of the form $H = I - 2vv^T$, where $\|v\|_2 = 1$) such that $\tilde{A} = H_1 A H_2$ is nonzero only in the $(1,1)$ element $\tilde{A}_{1,1}$. What are the possible values of $\tilde{A}_{1,1}$ in terms of x, y ?

   (b) Alice wants to compute the SVD of the matrix $A = \begin{bmatrix} -1 & -3 \\ 3 & 1 \end{bmatrix} = U\Sigma V^T$. Bob claims that V should be equal to the matrix of eigenvectors of $A^T A$, which he computes correctly as $\tilde{V} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$; similarly, U is the matrix of eigenvectors of AA^T, which is $\tilde{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$. The singular values are the square roots of the eigenvalues of $A^T A$, so Σ should be $\tilde{\Sigma} = \begin{bmatrix} 4 & \\ & 2 \end{bmatrix}$. However, with these matrices $\tilde{U}\tilde{\Sigma}\tilde{V}^T = \begin{bmatrix} -3 & -1 \\ 1 & 3 \end{bmatrix} \neq A$.

   (i) Where is the error in the arguments above? Correct it and derive a valid SVD of A.

   (ii) Let $U_0, V_0 \in \mathbb{R}^{2025\times 2}$ be orthonormal, that is, $U_0^T U_0 = V_0^T V_0 = I_2$. What are the singular values of $U_0 A V_0^T$?

   (c) Suppose A has the form $A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}$ where $A_{11} \in \mathbb{R}^{r\times r}$ is nonsingular. Let $X = \begin{bmatrix} I_r \\ 0 \end{bmatrix}$ and consider the thin QR factorisation $AX = QR$.

   (i) Show that the matrix $QQ^T A$ has rank at most r.

   (ii) Give a formula for $QQ^T A - A$ in terms of A_{11}, A_{12} and A_{22}.

   (iii) Show that $\left\|QQ^T A - A\right\|_F \leqslant \|A\|_F$, where $\|A\|_F = \sqrt{\sum_{i,j} |A_{ij}|^2}$ denotes the Frobenius norm.

   (iv) Now suppose that

$$A_{11} = A_{12} = I_r, \quad A_{22} = \epsilon I_r, \quad \text{that is,} \quad A = \begin{bmatrix} I_r & I_r \\ 0 & \epsilon I_r \end{bmatrix} \tag{20}$$

   Denote by A_r the best rank- r approximation to A, and define $c = \frac{\left\|QQ^T A - A\right\|_2}{\|A - A_r\|_2}$. Find c, and show that $c \to \sqrt{2}$ as $\epsilon \to 0$.

8. (10 points) Let $A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \in \mathbb{R}^{n\times n}$ where $A_{11} \in \mathbb{R}^{k\times k}$ is a nonsingular matrix. Let $\sigma_1(A) \geqslant \sigma_2(A) \geqslant \cdots \geqslant \sigma_n(A) \geqslant 0$ denote the singular values of A, and let $\|\cdot\|_2$ denote the spectral norm, so $\|A\|_2 = \sigma_1(A)$.

   (a) Show that $\sigma_i(A) \geqslant \sigma_i(A_{11})$ for $i = 1, 2, \ldots, k$. [Here and below, you may use the Courant-Fischer theorem and Weyl's theorem without proof.]

   (b) Give a lower bound for $\operatorname{rank}(A)$ and an example where it is attained.

   (c) Give an upper bound for $\operatorname{rank}(A)$ and an example where it is attained.

   (d) By giving an example show that $\left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & 0 \end{bmatrix} \right\|_2 \leqslant \left\| \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \right\|_2$ does not always hold.

9. (10 points) (a) First, let $A, B \in \mathbb{R}^{n\times n}$ be square matrices with $\det(B) \neq 0$. Prove that the eigenvalues of AB and BA are the same.

   (b) Next let $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$ with $m \geq n$. Prove that

$$\begin{bmatrix} AB & A \\ 0 & 0_{n\times n} \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0_{m\times m} & A \\ 0 & BA \end{bmatrix} \tag{55}$$

   are similar. Hence show that the nonzero eigenvalues of AB and BA are again the same.

   (c) Do you see a connection between this problem and the QR algorithm?