MATH557 — Applied Linear Algebra
Final Exam

Dr. Slim Belhaiza
Department of Mathematics
KFUPM

Duration: 2h30
No Computer, Regular Calculators Allowed

Instructions

- Answer **all questions**. Total: **100 points**.
- Show all essential steps. Exact values are preferred; keep radicals as needed.
- You may state (without proof) standard theorems about eigen-decomposition, Cholesky, QR, and SVD.
- Unless specified, vectors are columns. Use clear notation for orthogonality and projections.

Question 1 — Matrix Calculus Warm-up (10 points)

Let

$$A(t) = \begin{pmatrix} t+2 & 1 \\ 1 & 2t+1 \end{pmatrix}, \qquad t > 0.$$

1. (4 pts) Compute $\dfrac{d}{dt} A(t)^{-1}$ using the identity $\dfrac{d}{dt}A^{-1} = -A^{-1}A'A^{-1}$.
2. (3 pts) Compute the eigenvalues $\lambda_1(t), \lambda_2(t)$ of $A(t)$.
3. (3 pts) Differentiate your eigenvalues to obtain $\lambda_1'(t), \lambda_2'(t)$.

Question 2 — Cholesky Factorization and SPD Check (12 points)

Consider

$$S = \begin{pmatrix} 4 & 2 & 0 \\ 2 & 3 & 1 \\ 0 & 1 & 2 \end{pmatrix}.$$

1. (4 pts) Determine whether S is symmetric positive definite (SPD).
2. (8 pts) If S is SPD, compute the Cholesky factorization $S = LL^T$ with L lower-triangular and diagonal entries > 0.

Question 3 — QR / Gram–Schmidt (12 points)

Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \\ 1 & 3 \end{pmatrix}.$$

1. (8 pts) Perform (classical) Gram–Schmidt to compute a reduced QR factorization $A = QR$, where $Q \in \mathbb{R}^{3\times 2}$ has orthonormal columns and $R \in \mathbb{R}^{2\times 2}$ is upper-triangular.
2. (4 pts) Explain briefly why solving $\min_x \|Ax - b\|_2$ is easy once $A = QR$ is known (write the normal equation in QR form).

Question 4 — Least Squares and Pseudoinverse (14 points)

Let

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

1. (9 pts) Compute the least-squares solution $x^* = \arg\min_x \|Ax - b\|_2$ using the pseudoinverse of A.
2. (5 pts) Compute the residual $r = b - Ax^*$ and verify $A^T r = 0$.

.

Question 5 — One Iteration: GD vs Newton (14 points)

Let

$$f(x_1, x_2) = (x_1 - 1)^2 + (x_2 + 1)^4 + x_1 x_2.$$

1. (6 pts) Compute ∇f and the Hessian H.
2. (4 pts) Perform one iteration of Gradient Descent from $(0, 0)$ with step size $\alpha = \frac{1}{4}$.
3. (4 pts) Perform one iteration of Newton-Raphson's method from $(0, 0)$ (assume $H(0, 0)$ is invertible). State whether your iterate is a stationary point (justify briefly).

.

Question 6 — Matrix Completion (12 points)

1. (6 pts) Find the matrix $S = \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}$, with the smallest l_2 norm $||S||_2$ and such that:

$$S \begin{bmatrix} 1 \\ 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

2. (6 pts) Find the best possible real-valued matrix A^* that completes $A = \begin{pmatrix} * & 2 \\ 1 & * \end{pmatrix}$ and minimizes:

$$\min_{C \text{ and } R} \frac{1}{2}||(A - CR^t)_{known}||_2^2 + \frac{1}{2}(||C||_F^2 + ||R||_F^2)$$

where $CR^t = (2 \times 1)(1 \times 2)$.

Question 7 — Small Neural Net: Gradient and Hessian Structure (12 points)

We observe data $\{(p_{1t}, p_{2t}, y_t)\}_{t=1}^{T}$ and fit the (non-linear) one-neuron model

$$\hat{y}_t(w_1, w_2, w_3, w_4, w_5, b) = w_1 p_{1t}^2 + w_2 p_{2t}^2 + w_3 p_{1t} p_{2t} + w_4 p_{1t} + w_5 p_{2t} + b.$$

We also define the squared error

$$E(w_1, w_2, w_3, , w_4, w_5, b) = \sum_{t=1}^{T} (y_t - \hat{y}_t)^2.$$

1. (6 pts) Derive the gradient vector ∇E in a compact matrix form.
2. (4 pts) Derive the Hessian matrix $H = \nabla^2 E$ and explain why it is symmetric positive semi-definite.
3. (2 pts) Give one numerical linear algebra reason why the structure of H is helpful when iterating using Newton-Raphson's optimizer.

.

Question 8 — SVD and Best Rank-One Approximation (14 points)

Let

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix},$$

with

$$A^T A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

whose eigenvalues and corresponding unit eigenvectors are given by

$$\lambda_1 = 3, \quad v_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \lambda_2 = 1, \quad v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

Hint: You may use these facts directly.

1. (8 pts) Compute an SVD $A = U\Sigma V^T$ (you may work through $A^T A$). Keep radicals and do not over-simplify.
2. (4 pts) Write A as a sum of rank-one matrices using your SVD.
3. (2 pts) Write the closest rank-one approximation to A in Frobenius norm.