Data Science Preliminary Exam

June 2026

First name : ______________	Student ID : ______________
Last name : ______________	Additional pages : ______________

Instructions:

1. All problems are worth 10 points.
2. Explain your answers clearly. Show your reasoning steps and state the results and theorems you used in your arguments.
3. Be organized and write clearly. Use separate sheets for the solution of each problem. Illegible work will not receive full credit.

PROBLEM 1. In what follows, suppose A is an $m \times n$ real matrix of rank r, and b is a real m-dimensional vector.

a) Suppose that $m = r < n$. Prove or disprove that $Ax = b$ has infinitely many solutions regardless of the choice of b.

b) Suppose that $m = r = n$. Prove or disprove that $Ax = b$ has a unique solution regardless of the choice of b.

c) Suppose that $r < m$ and $r < n$. Prove or disprove that $Ax = b$ has no solution or infinitely many solutions depending on the choice of b.

d) Let $\{q_1, q_2, \ldots, q_k\}$ form an orthonormal basis for $\mathbb{R}^k$, and let Q be a matrix with the q_i as columns. Let $b \in \mathbb{R}^k$. What is $QQ^T b$ equal to?

PROBLEM 2.

a) Construct a matrix that projects $b \in \mathbb{R}^3$ onto the line passing through the origin in the direction $(1, 1, 1)$.

b) Prove that for a real square $n \times n$ matrix A with eigenvalues $\lambda_1, \ldots, \lambda_n$,

$$\det(A) = \lambda_1 \lambda_2 \cdots \lambda_n \qquad \text{and} \qquad \operatorname{tr}(A) = \lambda_1 + \cdots + \lambda_n.$$

c) Suppose $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of A and $u_1, \ldots, u_k$ are associated eigenvectors; that is, u_i is a λ_i-eigenvector. Prove or disprove that the vectors $u_1, \ldots, u_k$ are linearly independent.

d) Show that for the Frobenius norm,

$$\|A\|_F = \|\sigma\|_2,$$

where σ is the r-dimensional vector formed with the singular values of A, and r is the rank of A.

e) Prove or disprove: the rank of A equals the number of singular values of A.

PROBLEM 3. Duality in linear programming.

1. Prove that if X is a set of $n+2$ points in $\mathbb{R}^n$, then X can be partitioned into two disjoint sets of points A and B such that their convex hulls intersect.

2. Prove the following:

   (a) Prove that if a variable x_j in the primal LP is unrestricted in sign, then the corresponding dual constraint is an equation.

   (b) Show that if the i-th constraint of the primal problem is an equation, then the corresponding dual variable is unrestricted in sign.

3. Consider the primal LP problem

$$P = \max\{c^T x : Ax \leq b,\ x \geq 0\}$$

   and denote by D its dual. Suppose that $\bar{x}$ is optimal for P and $\bar{y}$ is optimal for D. Now suppose you construct P' by replacing $Ax \leq b$ with $MAx \leq Mb$, where M is a square nonsingular matrix of suitable dimensions. Find conditions on M (besides non-singular) that allow you to find value of the optimal solutions of P' and its dual D' using the solutions of P, D. What type of M can make this possible?

PROBLEM 4.

1. A problem about chess rooks: Prove that in a matrix the maximum number of nonzero entries, no two in the same row or column, is equal to the minimum number of rows and columns that include all of the nonzero entries.

2. Formulate the problem of finding a minimum-cost perfect matching in a bipartite graph as a minimum-cost flow problem.

PROBLEM 5.

1. Consider the affine set
$$S = \{x \in \mathbb{R}^3 : x_1 + x_2 + x_3 = 1\}.$$
Find the point in S with minimum Euclidean norm.

2. Next, consider the optimization problem
$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^{n} \left(\frac{1}{2} d_i x_i^2 + r_i x_i \right)$$
subject to
$$a^T x = 1, \qquad x_i \in [-1, 1], \quad i = 1, \ldots, n,$$
where $a \neq 0$ and $d_i > 0$. Verify whether strong duality holds for this optimization problem, and write down the KKT optimality conditions.

PROBLEM 6. We now consider problems in nonlinear optimization.

a) Let B_i, for $i = 1, \ldots, m$, be m Euclidean balls in $\mathbb{R}^n$, with centers x_i and radii $r_i \geq 0$.

   First, we wish to find a ball B of minimum radius that contains all the m balls. Can you cast this problem as a convex optimization program? Second, how different is the formulation to find the smallest cubes with sides parallel to the axis that contains the balls? Are both formulations convex?

b) Suppose that $g_1, g_2, \ldots, g_m$ are convex functions from $\mathbb{R}^n$ into $\mathbb{R}$. Prove that

$$\bar{g}(x) = \max\{g_1(x), g_2(x), \ldots, g_m(x)\}$$

   is also a convex function.