

Data Science Preliminary Exam (September 17, 2025)

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

Instructions:

1. All problems are worth 10 points.
2. Explain your answers clearly. Unclear answers will not receive credit.
State results and theorems you are using.
3. Use first the front and back of each page to write the solution of each problem.
4. If you need extra pages, please do not use the same sheet for different problems.
5. Write your name and problem number on any additional page you use.

PROBLEM 1. Recall two matrices $A, B \in \mathbb{R}^{n \times n}$ are orthogonally equivalent if there exists an orthogonal matrix Q with $A = QBQ^T$.

- a) Do orthogonally equivalent matrices have the same eigenvalues? Why?
- b) Is it true or false that A, B are orthogonally equivalent if and only if they have the same singular values?
- c) Suppose further that the nullspace of B is contained in the nullspace of A . Is it true that this implies that the range of B contains the range of A ?

PROBLEM 2. We consider now problems about norms.

a) Show that for any nonzero vector x ,

$$\text{card}(x) \geq \frac{\|x\|_1^2}{\|x\|_2^2},$$

where $\text{card}(x)$ is the cardinality of the vector x , i.e., the number of nonzero elements in x . Find vectors x for which the lower bound is attained.

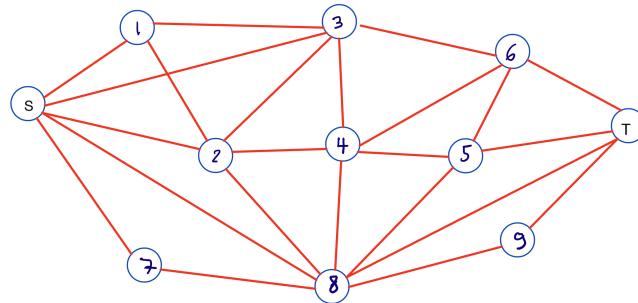
b) The spectral radius of a square real matrix A is the maximum absolute value of an eigenvalue of A . Prove that the operator 2-norm (also called the spectral norm) of A is bounded below by the spectral radius.

c) Show that the Frobenius norm of a matrix A depends only on its singular values. More precisely show that

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

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

PROBLEM 3. A company has to transport toxic waste from city s to city t . There are several roads connecting s and t passing through intermediate cities (an example of the situation is shown in the following graph where cities are nodes and edges represent roads).



Due to danger of driving the material it requires a crew of 2 staff and the same crew member cannot be the designated driver on consecutive stretches of travel (edges of the graph), thus the trips from s to t must pass an even number of roads. We call them *even trips*. The use of each road from city i to city j has a cost associated C_{ij} and the goal is to find the cheapest even trip possible.

a) Formulate the even trip optimization problem for a general st -graph as an integer linear program. (HINT: It may help you to think how to do it in the example graph or smaller situation).

b) Is there an efficient solution to your model? Discuss, but you do not have to compute anything explicitly.

PROBLEM 4. a) Let $a_1, a_2, \dots, a_n$ be n distinct real numbers and consider the linear program (P)

$$\max\{x \in \mathbb{R} : x \leq a_i \quad i = 1 \dots n\}.$$

What is the optimal value of (P)? Find the dual (D) of (P) and explain what (D) means.

b) Is it possible to give an example of a linear program that is infeasible and its dual is also infeasible? Justify your answer accordingly.

PROBLEM 5. In data classification, we are given a dataset $\mathcal{D} = \{(x_1, y_1), \dots, (x_n, y_n)\}$ with $x_i \in \mathbb{R}^d$ and $y_i \in \{-1, +1\}$. y_i is the *label* of data point i . A dataset is (strictly) *linearly separable* if and only if there exists $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$ such that:

$$y_i(w^\top x_i - b) > 0, \quad 1 \leq i \leq n.$$

- a. Show that the condition of being linearly separable is equivalent to the following condition: there exists $w \in \mathbb{R}^d$ and $b \in \mathbb{R}$ such that

$$y_i(w^\top x_i - b) \geq 1, \quad 1 \leq i \leq n.$$

Can you give an example of a data set that is not linear separable?

- b. Let us define $X^+ = \{x_i : y_i = +1, 1 \leq i \leq n\}$ and $X^- = \{x_i : y_i = -1, 1 \leq i \leq n\}$. Using Farkas' lemma, show that if $\mathcal{D}$ is not linearly separable then $C(X^+) \cap C(X^-) \neq \emptyset$. Here $C(X)$ denotes the convex hull of X . Can you explain the intuition of this result?

PROBLEM 6. We consider now problems on non-linear optimization.

a) Let B_i , for $i = 1, \dots, m$, 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 square that contains the balls? Are both formulations convex?

b) Next, consider the optimization problem

$$\min_{x \in \mathbb{R}^n} \sum_{i=1}^n \left(\frac{1}{2} d_i x_i^2 + r_i x_i \right)$$

such that $a^T x = 1$, $x_i \in [-1, 1]$, for $i = 1, \dots, n$, where $a \neq 0$ and $d_i > 0$.

Verify if strong duality holds for this optimization problem, and write down the KKT optimality conditions.