

Exam Location: _____

PRINT your student ID: _____

PRINT AND SIGN your name: _____,
(last) (first) (sign)

PRINT your discussion sections and (u)GSIs (the ones you attend): _____

Name and SID of the person to your left: _____

Name and SID of the person to your right: _____

1. Honor Code (0 pts)

Please copy the following statement in the space provided below and sign your name.

As a member of the UC Berkeley community, I act with honesty, integrity, and respect for others. I will follow the rules and do this exam on my own.

**IF YOU DO NOT COPY THE HONOR CODE AND SIGN YOUR NAME,
YOU WILL GET A 0 ON THE EXAM.**

2. SID and Exam Location (2 pts)

**BEFORE THE EXAM STARTS, WRITE YOUR SID AT THE TOP OF FIRST PAGE AND LAST PAGE.
ALSO MENTION YOUR EXAM LOCATION AT THE TOP OF FIRST PAGE.**

No extra time will be given for this task.

3. What is your favorite sport? Any answer, including no answer, will be given full credit. (2 pts)

Do not turn the page until your proctor tells you to do so.

4. Descent Methods (12 pts)

Consider the least squares problem

$$\min_{\vec{x} \in \mathbb{R}^n} \frac{1}{2} \|A\vec{x} - \vec{b}\|_2^2,$$

for some $\vec{b} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n \times n}$ such that $A \neq O_{n \times n}$.

- (a) (4 pts) What is the most restrictive class of optimization problem that this problem belongs to from the following options?

No justification is required.

NOTE: In this problem, choosing “I don’t know” in a multiple choice question merits **1 point** to discourage guessing. Picking more than one of the choices will automatically get 0 points.

HINT: Recall that $LP \subseteq QP \subseteq QCQP \subseteq SOCP \subseteq \text{Convex optimization problems}$.

- (1) LP
- (2) QP
- (3) QCQP
- (4) SOCP
- (5) Convex Optimization Problem
- (6) I don’t know

(b) (4 pts) Suppose

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \vec{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

Consider running gradient descent for the least squares problem

$$\min_{\vec{x} \in \mathbb{R}^2} \frac{1}{2} \|A\vec{x} - \vec{b}\|_2^2,$$

with step size $\eta = 1$ and initialization $\vec{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. What is the next iterate $\vec{x}^{(1)}$? Please provide a numerical value for each entry of $\vec{x}^{(1)}$.

HINT: Note that $A^\top \vec{b} = \begin{bmatrix} 22 \\ 32 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix}$, and $A^{-1} \vec{b} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

(c) (4 pts) Suppose

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}, \vec{b} = \begin{bmatrix} 4 \\ 6 \end{bmatrix}.$$

Consider running Newton's method for the least squares problem

$$\min_{\vec{x} \in \mathbb{R}^2} \frac{1}{2} \|A\vec{x} - \vec{b}\|_2^2,$$

with initialization $\vec{x}^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. What is the next iterate $\vec{x}^{(1)}$? Please provide a numerical value for each entry of $\vec{x}^{(1)}$.

HINT: Note that $A^\top \vec{b} = \begin{bmatrix} 22 \\ 32 \end{bmatrix}$, $A^\top A = \begin{bmatrix} 10 & 14 \\ 14 & 20 \end{bmatrix}$, and $A^{-1} \vec{b} = \begin{bmatrix} -2 \\ 3 \end{bmatrix}$.

5. Duality Trivia (20 pts)

Let $f_0, f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ be given. Suppose we have the primal problem

$$p^* = \min_{\vec{x} \in \mathbb{R}^n} f_0(\vec{x})$$

$$\text{s.t. } f_i(\vec{x}) \leq 0, \quad \text{for each } i \in \{1, \dots, m\},$$

where the dual function is $g(\vec{\lambda}) = \min_{\vec{x} \in \mathbb{R}^n} L(\vec{x}, \vec{\lambda})$ for the Lagrangian $L(\vec{x}, \vec{\lambda})$. For all subparts in this problem, no justification is required.

NOTE: In this problem, choosing “I don’t know” in a multiple choice question merits **1 point** to discourage guessing. On the other hand, “None of the above” means that we can’t say any of the above statements in generality (different scenarios could occur in different examples). In each subpart, picking more than one of the choices will automatically get 0 points.

(a) (4 pts) If $f_i(\vec{x})$ is a convex function for all $i = 0, \dots, m$, then what can we say about $g(\vec{\lambda})$?

NOTE: The constant functions $-\infty$ and $+\infty$ are considered to be both convex and concave.

- (1) $g(\vec{\lambda})$ is convex in $\vec{\lambda}$.
- (2) $g(\vec{\lambda})$ is concave in $\vec{\lambda}$.
- (3) $g(\vec{\lambda})$ is neither convex nor concave in $\vec{\lambda}$.
- (4) None of the above.
- (5) I don’t know.

(b) (4 pts) If $f_i(\vec{x})$ is neither convex nor concave, for all $i = 0, \dots, m$, then what can we say about $g(\vec{\lambda})$?

NOTE: The constant functions $-\infty$ and $+\infty$ are considered to be both convex and concave.

- (1) $g(\vec{\lambda})$ is convex in $\vec{\lambda}$.
- (2) $g(\vec{\lambda})$ is concave in $\vec{\lambda}$.
- (3) $g(\vec{\lambda})$ is neither convex nor concave in $\vec{\lambda}$.
- (4) None of the above.
- (5) I don’t know.

(c) (4 pts) Suppose $p^* = 0$, and $f_0(\vec{x})$ is linear, and $f_i(\vec{x})$ are affine for $i = 1, \dots, m$. What is the strongest statement we can make about d^* ?

- (1) $d^* = 0$.
- (2) $d^* \leq 0$.
- (3) $d^* \geq 0$.
- (4) None of the above.
- (5) I don’t know.

(d) (4 pts) Suppose $p^* = 0$, and $f_i(\vec{x})$ is a convex function for each $i \in \{0, 1, \dots, m\}$. Moreover, there exists a vector $\vec{c} \in \mathbb{R}^n$ that satisfies $\max_{i \in \{1, \dots, m\}} f_i(\vec{c}) = -3$. What is the strongest statement we can make about d^* ?

- (1) $d^* = 0$.
- (2) $d^* \leq 0$.
- (3) $d^* \geq 0$.
- (4) None of the above.
- (5) I don’t know.

- (e) (4 pts) Suppose $p^* = 0$, and $f_i(\vec{x})$ is a convex quadratic function of \vec{x} for each $i \in \{0, 1, \dots, m\}$. What is the strongest statement we can make about d^* ?
- (1) $d^* = 0$.
 - (2) $d^* \leq 0$.
 - (3) $d^* \geq 0$.
 - (4) None of the above.
 - (5) I don't know.

6. Norms and Linear Algebra (12 pts)

Let $A \in \mathbb{R}^{m \times n}$ with $m \leq n$, with columns \vec{a}_1 through \vec{a}_n , i.e., $A = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \dots & \vec{a}_n \end{bmatrix}$. Suppose $AA^\top = I_{m \times m}$, where $I_{m \times m}$ denotes the $m \times m$ identity matrix.

- (a) (4 pts) Find the average of the squared norms of the columns of A , i.e. $\frac{1}{n} \sum_{i=1}^n \|\vec{a}_i\|_2^2$.

(b) (8 pts) Show that the norm of every column of A is at most 1, i.e. $\|\vec{a}_i\|_2 \leq 1$ for all $i \in \{1, \dots, n\}$.

HINT: Recall that the spectral norm of a matrix equals its largest singular value.

7. Slater's Condition (12 pts)

Consider the following optimization problem

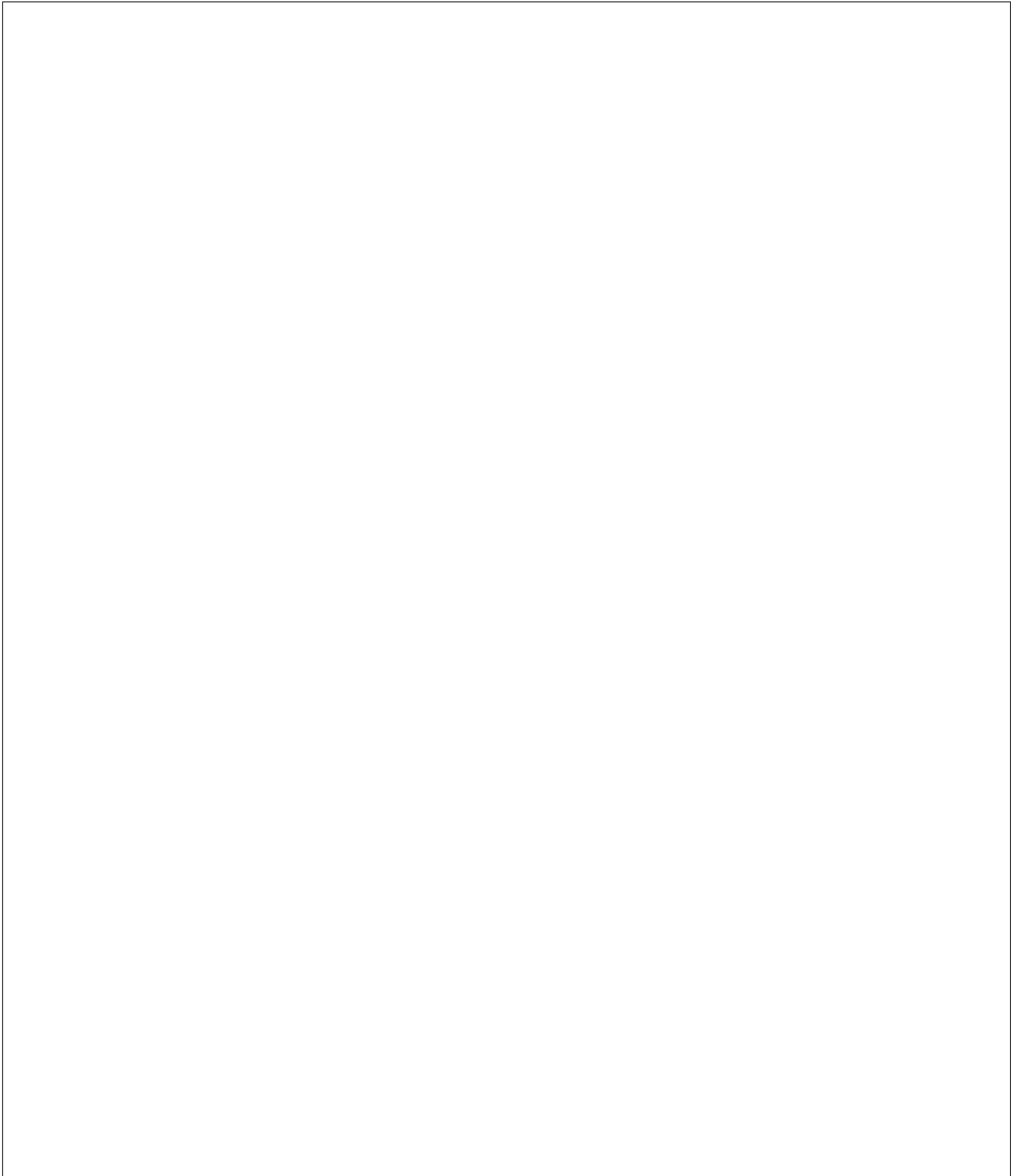
$$\begin{aligned} \min_{x,y \in \mathbb{R}} \quad & 2x + y \\ \text{s.t.} \quad & x^2 + y^2 \leq 1, \\ & (x - 2)^2 + y^2 \leq c, \end{aligned}$$

for some parameter $c > 0$.

- (a) (4 pts) If $c = 4$, does this problem satisfy Slater's condition? Justify your answer.

(b) (4 pts) If $c = 1$, what is the feasible set of the problem?

HINT: For each constraint, try drawing the set of all $(x, y) \in \mathbb{R}^2$ that satisfy the constraint.



(c) (4 pts) If $c = 1$, does the problem satisfy Slater's condition? Justify your answer.

8. An Optimization Problem over Matrices (8 pts)

Let $I_{n \times n}$ denote the $n \times n$ identity matrix. Is the following a convex optimization problem?

$$\begin{aligned} \min_{X \in \mathbb{R}^{n \times n}} \quad & \|X - 2I_{n \times n}\|_F^2 \\ \text{s.t.} \quad & \vec{e}_i^\top X \vec{e}_i \geq 0, \quad \text{for each } i \in \{1, \dots, n\}, \\ & \text{tr}(X) = n, \end{aligned}$$

where $\text{tr}(X)$ denotes the trace of the matrix X , $\|X - 2I_{n \times n}\|_F$ denotes the Frobenius norm of the matrix $X - 2I_{n \times n}$, and for each $i \in \{1, \dots, n\}$ we have used $\vec{e}_i \in \mathbb{R}^n$ to denote the vector with 1 in the i -th coordinate and 0 elsewhere. Justify your answer.

9. Cones (12 pts)

Let $K := \{\vec{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 2x_1 \geq 0\}$.

(a) (4 pts) Prove that K is a cone.

HINT: A set K is a cone if and only if for all $\vec{x} \in K$ and $\alpha \in \mathbb{R}, \alpha \geq 0$, we have $\alpha\vec{x} \in K$.

(b) (8 pts) Let K^* denote the dual cone of K which is defined as follows

$$K^* = \{\vec{y} = (y_1, y_2) \in \mathbb{R}^2 : \vec{y}^\top \vec{x} \geq 0 \text{ for all } \vec{x} \in K\}.$$

Compute an expression for K^* of the form

$$K^* = \{\vec{y} = (y_1, y_2) \in \mathbb{R}^2 : a_1 y_1 + a_2 y_2 \geq 0, b_1 y_1 + b_2 y_2 \geq 0\},$$

for some scalars $a_1, a_2, b_1, b_2 \in \mathbb{R}$.

HINT: One can equivalently write $K = \{\vec{x} = (x_1, x_2) \in \mathbb{R}^2 : x_2 \geq 2x_1 \geq 0\}$ as $K = \{(x_1, 2x_1 + u) \in \mathbb{R}^2 : x_1 \geq 0, u \geq 0\}$.

10. Constrained Optimization (8 pts)

Consider the following optimization problem for fixed scalars $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} \min_{x \in \mathbb{R}} \quad & \frac{1}{2}(x - \alpha)^2 + \beta x \\ \text{s.t.} \quad & x \geq 0. \end{aligned}$$

Show that the optimal solution of this optimization problem is given by $x^* = \max\{0, \alpha - \beta\}$.

HINT: Try first to solve the unconstrained problem $\min_{x \in \mathbb{R}} \frac{1}{2}(x - \alpha)^2 + \beta x$.

11. Sum of Even Polynomials (12 pts)

Let $a_1, a_2 \in \mathbb{R}$. Consider the optimization problem:

$$\begin{aligned} \min_{x_1, x_2 \in \mathbb{R}} \quad & (x_1 + a_1)^4 + (x_2 + a_2)^6 \\ \text{s.t.} \quad & x_1 \geq 0, \\ & x_2 \geq 0, \\ & x_1 + x_2 = 1. \end{aligned}$$

The objective function and all the constraint functions are assumed to have domain \mathbb{R}^2 . Also note that the constraints are defined by affine functions.

(a) (4 pts) Show that the objective function is convex.

(b) (8 pts) We can write a Lagrangian for the problem in the form

$$\mathcal{L}(x_1, x_2, \lambda_1, \lambda_2, \nu) = (x_1 + a_1)^4 + (x_2 + a_2)^6 - \lambda_1 x_1 - \lambda_2 x_2 + \nu(x_1 + x_2 - 1).$$

Here λ_1 and λ_2 are the dual variables corresponding to the two inequality constraints respectively, and ν is the dual variable corresponding to the equality constraint. With this Lagrangian in mind, write down the KKT equations for the problem.

NOTE: It is not necessary to attempt to compute the dual objective function. Also, you are not asked to solve the KKT equations, just to write them down.

12. Zero-Sum Games (22 pts)

Consider a matrix $A \in \mathbb{R}^{m \times n}$. Let

$$p^* = \min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y},$$

where \mathcal{P}_m and \mathcal{P}_n are unit simplices in \mathbb{R}^m and \mathbb{R}^n , respectively, given by

$$\mathcal{P}_m = \left\{ \vec{x} \in \mathbb{R}^m : \sum_{i=1}^m x_i = 1, x_i \geq 0 \text{ for each } i \in \{1, 2, \dots, m\} \right\},$$

$$\mathcal{P}_n = \left\{ \vec{y} \in \mathbb{R}^n : \sum_{j=1}^n y_j = 1, y_j \geq 0 \text{ for each } j \in \{1, 2, \dots, n\} \right\}.$$

Furthermore, for each $j \in \{1, 2, \dots, n\}$, let \vec{e}_j denote the vector with 1 in its j -th coordinate and 0, otherwise. Define

$$\mathcal{E}_n = \{\vec{e}_j : j \in \{1, 2, \dots, n\}\}.$$

- (a) (6 pts) Show that for any $\vec{x} \in \mathcal{P}_m$, there exists $\vec{e}_{j^*} \in \mathcal{E}_n$ such that

$$\vec{x}^\top A \vec{e}_{j^*} = \max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y}.$$

HINT: What kind of optimization problem is $\max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y}$?

(b) (8 pts) Show that

$$\min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y} = \min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{e}_j \in \mathcal{E}_n} \vec{x}^\top A \vec{e}_j.$$

HINT: Try to show that $\min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y} \leq \min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{e}_j \in \mathcal{E}_n} \vec{x}^\top A \vec{e}_j$, and $\min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{y} \in \mathcal{P}_n} \vec{x}^\top A \vec{y} \geq \min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{e}_j \in \mathcal{E}_n} \vec{x}^\top A \vec{e}_j$.

(c) (8 pts) Formulate the optimization problem $\min_{\vec{x} \in \mathcal{P}_m} \max_{\vec{e}_j \in \mathcal{E}_n} \vec{x}^\top A \vec{e}_j$ as a linear program.

HINT: The set \mathcal{E}_n contains only finitely many vectors.

13. Gradient Descent for Convex Functions (12 pts)

Let $f : \mathbb{R} \rightarrow \mathbb{R}$ be a differentiable function, with $\text{dom}(f) = \mathbb{R}$. We carry out one step of gradient descent on f , with step size $\frac{1}{4}$, starting at $x_0 \in \mathbb{R}$. This leads us to

$$x_1 = x_0 - \frac{1}{4} \frac{df}{dx}(x_0).$$

Suppose the function $g(x) := 2x^2 - f(x)$ is convex. Show that

$$f(x_1) \leq f(x_0) - \frac{1}{8} \left(\frac{df}{dx}(x_0) \right)^2.$$

NOTE: You may not assume that either f or g satisfies any property other than the ones mentioned above.

14. Second-Order Cone Programming (14 pts)

Fix $A \in \mathbb{R}^{m \times n}$, $\vec{y} \in \mathbb{R}^m$, and $\lambda, \mu > 0$. Consider the following modified version of the LASSO problem with both an ℓ^2 regularization term and an ℓ^1 regularization term, i.e.,

$$\min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{y}\|_2^2 + \mu \|\vec{x}\|_2 + \lambda \|\vec{x}\|_1.$$

We will show that this problem can be equivalently written as an SOCP.

NOTE: The ℓ^2 regularization term is given as $\mu \|\vec{x}\|_2$, rather than $\mu \|\vec{x}\|_2^2$.

- (a) (4 pts) Find vectors $\vec{a}, \vec{b} \in \mathbb{R}^n$ and scalars $u, v \in \mathbb{R}$ (each of them in terms of μ and λ) such that the optimization problem is equivalent to solving

$$\begin{aligned} \min_{\substack{\vec{x}, \vec{r} \in \mathbb{R}^n \\ t, p \in \mathbb{R}}} & \vec{a}^\top \vec{x} + ut + vp + \vec{b}^\top \vec{r} \\ \text{s.t.} & \|A\vec{x} - \vec{y}\|_2^2 \leq t, \\ & \|\vec{x}\|_2 \leq p, \\ & |x_i| \leq r_i, \quad \forall i \in \{1, \dots, n\}. \end{aligned}$$

(b) (6 pts) Find a second-order cone constraint over the variables (\vec{x}, t) that is equivalent to the constraint $\|A\vec{x} - \vec{y}\|_2^2 \leq t$.

HINT: You may use without proof the fact that for any $\vec{z} \in \mathbb{R}^n$ and $t \in \mathbb{R}$

$$\|\vec{z}\|_2^2 \leq t \text{ if and only if } \left\| \begin{bmatrix} 2\sqrt{2}\vec{z} \\ 1 - 2t \end{bmatrix} \right\|_2 \leq 1 + 2t.$$

- (c) (4 pts) For each $i \in \{1, \dots, n\}$, find a second-order cone constraint over the variables (x_i, r_i) that is equivalent to the constraint $|x_i| \leq r_i$.

NOTE. We know that $\|x\|_2 \leq t$ is a second-order cone constraint. Using this and parts (a)-(c), we can write the optimization problem

$$\min_{\vec{x} \in \mathbb{R}^n} \quad \|A\vec{x} - \vec{y}\|_2^2 + \mu\|\vec{x}\|_2 + \lambda\|\vec{x}\|_1,$$

as an SOCP. You are not expected to do this.

15. Support Vector Machines (12 pts)

Recall the maximum margin support vector machine problem:

$$\begin{aligned} \min_{\vec{w} \in \mathbb{R}^k, b \in \mathbb{R}} \quad & \frac{1}{2} \|\vec{w}\|_2^2 \\ \text{s.t.} \quad & y_i(\vec{w}^\top \vec{x}_i + b) \geq 1 \quad \forall i \in \{1, \dots, n\}, \end{aligned}$$

where the data points (\vec{x}_i, y_i) , with features $\vec{x}_i \in \mathbb{R}^k$ and labels $y_i \in \{+1, -1\}$ for $i \in \{1, \dots, n\}$, are given.

- (a) (8 pts) Consider the pairs of features $\vec{x}_i \in \mathbb{R}^2$ and labels $y_i \in \{+1, -1\}$ given in Figure 1. The maximum margin hyperplane for this data along with the support vectors are depicted in Figure 2. Find the vector \vec{w} and scalar b that solve this problem.

Index i	Features $(x_{i1}, x_{i2}) \in \mathbb{R}^2$	Label $y_i \in \{+1, -1\}$
1	(1, 1)	+1
2	(3, 4)	+1
3	(3, 5)	+1
4	(4, 0)	-1
5	(5, 1)	-1
6	(6, 6)	-1

Figure 1: Data points and their labels

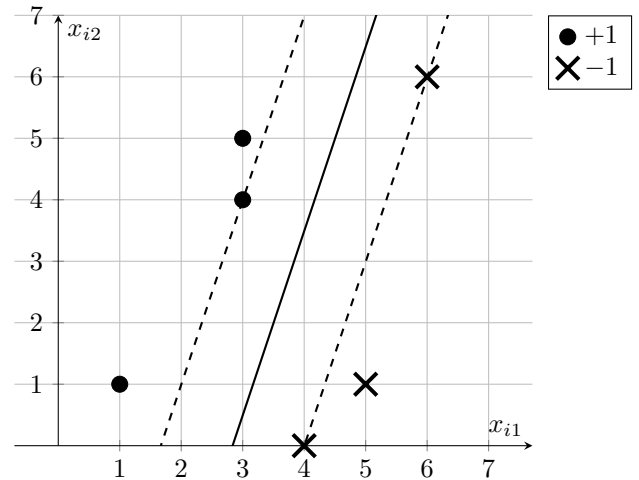


Figure 2: Maximum margin hyperplane and support vectors

HINT: Note that the constraints in the maximum margin support vector machine problem must be satisfied with equality at the support vectors.

HINT: You are likely to find at least one of these two calculations to be useful:

$$\begin{bmatrix} 3 & 4 & 1 \\ 4 & 0 & 1 \\ 6 & 6 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -3/7 & 1/7 & 2/7 \\ 1/7 & -3/14 & 1/14 \\ 12/7 & 3/7 & -8/7 \end{bmatrix}, \quad \begin{bmatrix} 1 & 1 & 1 \\ 3 & 5 & 1 \\ 5 & 1 & 1 \end{bmatrix}^{-1} = \begin{bmatrix} -1/4 & 0 & 1/4 \\ -1/8 & 1/4 & -1/8 \\ 11/8 & -1/4 & -1/8 \end{bmatrix}.$$



(b) (4 pts) Now, consider the pairs of features $\vec{x}_i \in \mathbb{R}^2$ and labels $y_i \in \{+1, -1\}$ given in Figure 3, and depicted visually in Figure 4:

Index i	Features $(x_{i1}, x_{i2}) \in \mathbb{R}^2$	Label $y_i \in \{+1, -1\}$
1	(1, 1)	+1
2	(4.5, 1)	+1
3	(4, 6)	+1
4	(4, 0)	-1
5	(4, 2)	-1
6	(5, 1)	-1

Figure 3: Data points and their labels

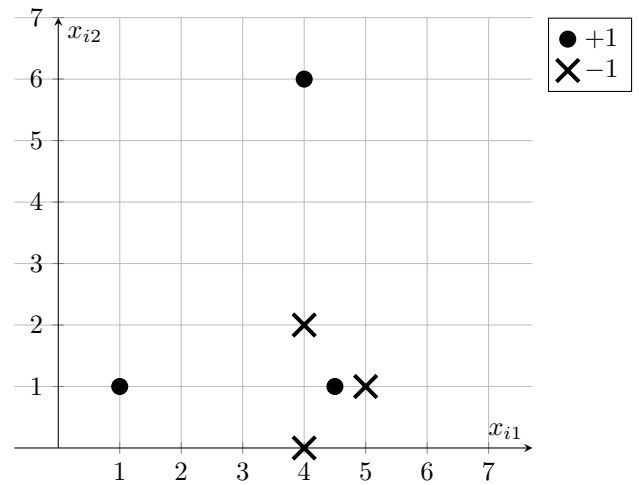


Figure 4: Visual depiction of data points and labels

If possible, find a separating hyperplane that solves the maximum margin support vector machine problem with this data, or provide a justification why such a hyperplane cannot be found.

[Extra page for scratch work that will not be graded unless you tell us in the original problem space.]

[Extra page for scratch work that will not be graded unless you tell us in the original problem space.]

[Extra page for scratch work that will not be graded unless you tell us in the original problem space.]

[Doodle page! Draw us something if you want or give us suggestions or complaints. You can also use this page to report anything suspicious that you might have noticed. You can also use this page to write solutions if you need the space, but please tell us in the original problem space.]

Read the following instructions before the exam.

PRINT your student ID: _____

There are 15 problems of varying numbers of points, with a total of 160 points. You have 180 minutes for the exam. The problems are of varying difficulty, so pace yourself accordingly, do easier problems first, and avoid spending too much time on any one question until you have gotten all of the other points you can. Problems are not necessarily ordered in terms of difficulty, so be sure to read all the problems.

There are 32 pages on the exam, so there should be 16 sheets of paper in the exam. The exam is printed double-sided. Do not forget the problems on the back sides of the pages! Notify a proctor immediately if a page is missing. **Do not tear out or remove any of the pages. Do not remove the exam from the exam room.**

No collaboration is allowed, and do not attempt to cheat in any way. Cheating will not be tolerated.

Write your student ID on pgs. 1 and 32. Each exam has a unique identifying number, which will be tied to your SID to enable us to determine the SID corresponding to each page. Additionally, also write the exam location on page 1.

You may consult TWO handwritten 8.5" × 11" note sheet(s) (front and back). No phones, calculators, tablets, computers, other electronic devices, or scratch paper are allowed.

Please write your answers legibly in the boxed spaces provided on the exam. The space provided should be adequate. **If you still run out of space, please use a blank page and clearly tell us in the original problem space where to look for your solution.**

Unless otherwise specified, show all of your work in order to receive full credit. Partial credit will be given for substantial progress on each problem.

We will not be able to answer most questions or offer clarifications during the exam.

If you need to use the restrooms during the exam, bring your student ID card, your phone, and your exam to a proctor. You can collect them once you return from the restrooms.

Our advice to you: if you can't solve the problem, state and solve a simpler one that captures at least some of its essence. You might get some partial credit, and more importantly, you will perhaps find yourself on a path to the solution.

Good luck!

Do not turn the page until your proctor tells you to do so.