EECS 127/227AT Optimization Models in Engineering UC Berkeley Spring 2023

# Final

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: ______________________________

Name and SID of the person in front of you: ______________________________

Name and SID of the person behind you: ______________________________

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 (3 pts)**

   **When the exam starts, write your SID at the top of every page.** *No extra time will be given for this task.*

3. **Favorites. Any answer, as long as you write it down, will be given full credit. (2 pts)**

   (a) (1 pts) **What's a movie you are looking forward to watching this summer?**

   (b) (1 pts) **If you could have any animal as a pet, what animal would you choose?**

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

**4. Convex Functions (7 pts)**

(a) (3 pts) **Prove that the function** $f\colon \mathbb{R}_{++} \to \mathbb{R}$ **given by** $f(x) \doteq \log(1/x)$ **is a strictly convex function**, where $\mathbb{R}_{++}$ is the set of strictly positive real numbers.

(b) (4 pts) **Is the function** $g\colon \mathbb{R}_{++} \to \mathbb{R}$ **given by** $g(x) \doteq \max\{(ax-b)^2, \log(1/x)\}$ **a convex function?** *Justify your answer.*

You may use the fact that the function $f(x) \doteq \log(1/x)$ is convex for $x \in \mathbb{R}_{++}$.

**5. Hyperplanes (7 pts)**

(a) (3 pts) **Give a hyperplane of the form** $\mathcal{H} \doteq \{\vec{x} \in \mathbb{R}^n \mid \vec{c}^\top(\vec{x} - \vec{x}_0) = 0\}$ **which goes through the point** $(2, 3)$ **and is orthogonal to the vector** $(1, 1)$. *No justification is necessary.*

(b) (4 pts) Let $\vec{c}_1 \neq \vec{0}$ and $\vec{c}_2 \neq \vec{0}$ be two vectors in $\mathbb{R}^n$. Define the two hyperplanes $\mathcal{H}_1 = \{\vec{x} \in \mathbb{R}^n \mid \vec{c}_1^\top \vec{x} = 0\}$ and $\mathcal{H}_2 = \{\vec{x} \in \mathbb{R}^n \mid \vec{c}_2^\top \vec{x} = 0\}$ where $\vec{c}_1^\top \vec{c}_2 = 0$. **Give any point** $\vec{x}^\star$ **in terms of** $\vec{c}_1$ **and** $\vec{c}_2$ **such that** $\vec{c}_1^\top \vec{x}^\star > 0$ **and** $\vec{c}_2^\top \vec{x}^\star > 0$.

*HINT: Draw a picture of the two hyperplanes in 2D.*

**6. Newton's Method (6 pts)**

(a) (3 pts) Let $f : \mathbb{R}^n \to \mathbb{R}$ be a twice-differentiable function that we are attempting to minimize using Newton's method. Suppose that at the $k^{\text{th}}$ iterate $\vec{x}_k \in \mathbb{R}^n$ we have $\nabla^2 f(\vec{x}_k) = \alpha_k I_n$, where $\alpha_k > 0$ is some positive constant and $I_n \in \mathbb{R}^{n\times n}$ is the identity matrix. **Write the Newton's method step for $\vec{x}_{k+1}$ in terms of $\vec{x}_k$, $\alpha_k$, and $\nabla f(\vec{x}_k)$.**

(b) (3 pts) Now suppose we are trying to minimize the same function $f$ via gradient descent. **Write the gradient descent step for $\vec{x}_{k+1}$ in terms of $\vec{x}_k$ and $\nabla f(\vec{x}_k)$, with some arbitrary step size $\eta_k > 0$ at time $k$. For what value of $\eta_k$ is the gradient descent update equation the same as the Newton's update equation from part (a)?**

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

**7. Solving a Quadratic Program (10 pts)**

Consider the quadratic program

$$p^\star = \min_{\vec{x} \in \mathbb{R}^3} \left( \vec{x}^\top M \vec{x} - 2\vec{b}^\top \vec{x} \right), \tag{1}$$

where the matrix $M \in \mathbb{R}^{3\times 3}$ is defined as follows

$$M = 4 \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}^\top + 3 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}^\top. \tag{2}$$

(a) (5 pts) **If $\vec{b} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$, then is $p^\star = -\infty$ or is it finite?** You do not need to calculate $p^\star$. *Justify your answer.*

(b) (5 pts) **If** $\vec{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$**, then is** $p^\star = -\infty$ **or is it finite?** You do not need to calculate $p^\star$. *Justify your answer.*

**8. Quadratically Constrained Linear Program (16 pts)**

Let $\vec{c}, \vec{x}_0 \in \mathbb{R}^n$ where $\vec{c} \neq \vec{0}$. Let $Q \in \mathbb{S}^n_{++}$ be a symmetric positive definite matrix. Let $\epsilon > 0$ be a positive scalar. Consider the following optimization problem

$$\begin{aligned} p^\star = \min_{\vec{x}} \quad & \vec{c}^\top \vec{x} \\ \text{s.t.} \quad & \frac{1}{2}(\vec{x} - \vec{x}_0)^\top Q(\vec{x} - \vec{x}_0) \leq \epsilon. \end{aligned} \tag{3}$$

(a) (4 pts) **Is this problem convex? Does strong duality hold here?** *Justify your answer.*

Recall the optimization problem

$$p^\star = \min_{\vec{x}} \quad \vec{c}^\top \vec{x} \tag{3}$$
$$\text{s.t.} \quad \frac{1}{2}(\vec{x}-\vec{x}_0)^\top Q(\vec{x}-\vec{x}_0) \leq \epsilon. \tag{4}$$

(b) (8 pts) **Show that the dual function associated with the primal problem in (3) is**

$$g(\lambda) = \begin{cases} -\infty & \text{if } \lambda = 0 \\ \vec{c}^\top \vec{x}_0 - \frac{1}{2\lambda}\vec{c}^\top Q^{-1}\vec{c} - \lambda\epsilon & \text{if } \lambda > 0, \end{cases} \tag{5}$$

for $\lambda \geq 0$, where $\lambda$ is the dual variable associated with the quadratic inequality constraint.

(c) (4 pts) Consider the dual problem of the primal problem in (3):

$$d^\star = \max_{\lambda \geq 0} g(\lambda), \tag{6}$$

where

$$g(\lambda) = \begin{cases} -\infty & \text{if } \lambda = 0 \\ \vec{c}^\top \vec{x}_0 - \frac{1}{2\lambda}\vec{c}^\top Q^{-1}\vec{c} - \lambda\epsilon & \text{if } \lambda > 0. \end{cases} \tag{5}$$

**Find the optimal dual variable $\lambda^\star$.**

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

**9. Support Vector Regression (11 pts)**

Let $\epsilon > 0$. Let $X \in \mathbb{R}^{n \times d}$ be a data matrix and $\vec{y} \in \mathbb{R}^n$ be a label vector, such that

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix}, \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}, \tag{7}$$

where $\vec{x}_i^\top$ are the rows of $X$ and $y_i$ are the entries of $\vec{y}$. Consider the problem

$$\begin{aligned} \min_{\vec{w} \in \mathbb{R}^d} \quad & \frac{1}{2} \left\| \vec{w} \right\|_2^2 \\ \text{s.t.} \quad & \left\| X\vec{w} - \vec{y} \right\|_\infty \leq \epsilon. \end{aligned} \tag{8}$$

(a) (5 pts) **Rewrite this problem as an equivalent quadratic program (i.e. with quadratic objective function and finitely many linear constraints).**

(b) (6 pts) Now consider the problem:

$$\min_{\vec{w}\in\mathbb{R}^d} \left\{ \frac{1}{2}\|\vec{w}\|_2^2 + \lambda \sum_{i=1}^{n} \max\{0, \left|\vec{x}_i^\top \vec{w} - y_i\right| - \epsilon\} \right\}. \tag{9}$$

**Rewrite this problem as an equivalent quadratic program (i.e. with quadratic objective function and finitely many linear constraints).**

*HINT: Introduce a new variable $\vec{z}$.*

**10. Candidate Solution of Linear Programs (12 pts)**

(a) (5 pts) Consider the linear program

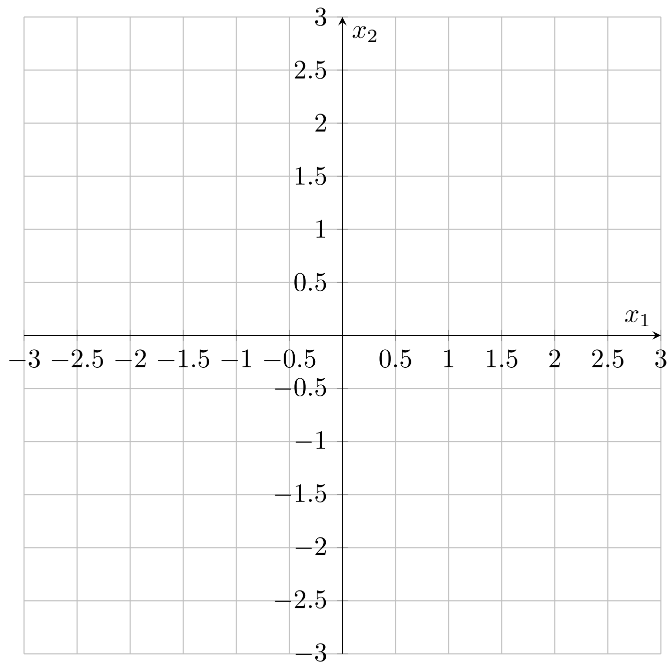
$$\min_{\vec{x}\in\mathbb{R}^2} \quad \begin{bmatrix} 2 \\ 1 \end{bmatrix}^\top \vec{x} \tag{10}$$

$$\text{s.t.} \quad \vec{x} \geq 0 \tag{11}$$

$$\begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \vec{x} \geq \begin{bmatrix} 2 \\ 1 \end{bmatrix}. \tag{12}$$

i. **Sketch and shade the feasible region of the above optimization problem in the graph provided below.**

ii. **Use your sketch to identify the optimal solution $\vec{x}^\star$ and write it in the box below.**



(b) (7 pts) Let $A \in \mathbb{S}^n$ be a *symmetric* matrix and $\vec{c} \neq \vec{0}$ be a vector in $\mathbb{R}^n$. Consider the linear program

$$\min_{\vec{x} \in \mathbb{R}^n} \quad \vec{c}^\top \vec{x} \tag{13}$$

$$\text{s.t.} \quad A\vec{x} \geq \vec{c} \tag{14}$$

$$\vec{x} \geq 0. \tag{15}$$

Consider a point $\vec{x}^\star > 0$ such that $A\vec{x}^\star = \vec{c}$. **Prove that $\vec{x}^\star$ is a minimizer of the above optimization problem.**

*HINT: Let $\vec{\lambda}$ be the dual variables associated with the constraints $A\vec{x} \geq \vec{c}$ and $\vec{\mu}$ be the dual variables associated with the constraints $\vec{x} \geq 0$. Use the KKT conditions.*

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

**11. Inscribed Box in a Polyhedron (8 pts)**

Let $\mathcal{P} := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} \leq \vec{b}\}$ be a bounded polyhedron with matrix $A \in \mathbb{R}^{m\times n}$ and vector and $\vec{b} \in \mathbb{R}^m$ such that

$$A = \begin{bmatrix} \vec{a}_1^\top \\ \vdots \\ \vec{a}_m^\top \end{bmatrix}, \qquad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}, \tag{16}$$

where $\vec{a}_i^\top$ are the rows of $A$ and $b_i$ are the entries of $\vec{b}$. In this problem, we wish to find the maximal radius box

$$\mathcal{B}(\vec{x}_0, r) = \{\vec{x}_0 + r\vec{u} \mid \|\vec{u}\|_\infty \leq 1\}, \tag{17}$$

such that $\mathcal{B}(\vec{x}_0, r) \subseteq \mathcal{P}$. In other words, we want to solve

$$\max_{\vec{x}_0, r} \quad r \tag{18}$$

$$\text{s.t.} \quad \mathcal{B}(\vec{x}_0, r) \subseteq \mathcal{P}. \tag{19}$$

**Express this problem as a linear program with at most $m$ constraints.** *Justify your answer.*

*HINT: Recall the $\ell^1$—$\ell^\infty$ duality:*

$$\|\vec{v}\|_1 = \max_{\vec{u}:\, \|\vec{u}\|_\infty \leq 1} \vec{u}^\top \vec{v}. \tag{20}$$

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

**12. Gambler's Destiny (16 pts)**

Aekus and Aditya notice that their local casino is offering great odds for betting on a marathon with $n$ athletes. They have $c > 0$ dollars with which they can place bets. For each athlete $i$, they can bet $b_i$ dollars. If athlete $i$ wins they receive $b_i r_i > 0$ dollars and receive nothing for all other athletes. Athlete $i$ wins with probability $p_i$. Exactly one athlete wins each race.

(a) (10 pts) Suppose Aekus and Aditya wish to *maximize* their expected profit after one bet in which they can bet up to $c$ dollars. Formally, they want to solve the optimization problem:

$$\begin{aligned} \min_{b_1,\ldots,b_n} \quad & -\left(\sum_{i=1}^{n} p_i b_i r_i + \left(c - \sum_{i=1}^{n} b_i\right)\right) \\ \text{s.t.} \quad & \sum_{i=1}^{n} b_i \leq c, \\ & b_i \geq 0 \; \forall i. \end{aligned} \tag{21}$$

**Write the Lagrangian and KKT conditions for (21) where $\lambda$ is the dual variable corresponding to the constraint $\sum_{i=1}^{n} b_i \leq c$ and $\mu_i$ is the dual variable corresponding to the constraint $b_i \geq 0$. Then, show that for the primal and dual optimizers $b_i^\star$, $\mu_i^\star$, and $\lambda^\star$, the following relations hold:**

i. if for any athlete $j$ we have that $p_j r_j > 1$, then $\sum_{i=1}^{n} b_i^\star = c$.

ii. if for any athlete $j$ we have that $p_j r_j < 1$, then $b_j^\star = 0$.

iii. if $p_i r_i > p_j r_j$, then $\mu_j^\star > \mu_i^\star$.

Recall the optimization problem

$$\begin{aligned} \min_{b_1,\ldots,b_n} \quad & -\left(\sum_{i=1}^{n} p_i b_i r_i + \left(c - \sum_{i=1}^{n} b_i\right)\right) \\ \text{s.t.} \quad & \sum_{i=1}^{n} b_i \leq c, \\ & b_i \geq 0 \; \forall i. \end{aligned} \tag{21}$$

Recall from part (a) that for the primal and dual optimizers $b_i^\star$, $\mu_i^\star$, and $\lambda^\star$, the following relations hold:

i. if for any athlete $j$ we have that $p_j r_j > 1$, then $\sum_{i=1}^{n} b_i^\star = c$.

ii. if for any athlete $j$ we have that $p_j r_j < 1$, then $b_j^\star = 0$.

iii. if $p_i r_i > p_j r_j$, then $\mu_j^\star > \mu_i^\star$.

(b) (6 pts) **Use these relations to argue that if there exists athlete $j$ such that $p_j r_j > 1$ and $p_j r_j > p_i r_i$ for all other athletes $i \neq j$, then $b_j^\star = c$.** In other words, Aekus and Aditya should allocate all their money to the most profitable bet in expectation, if one exists.

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

**13. Conjugate Gradient Method (39 pts)**

In this problem we explore a new descent method, called the conjugate gradient method[1], to solve the problem $A\vec{x} = \vec{b}$ where $A \in \mathbb{S}^n_{++}$ is a symmetric positive definite matrix, $\vec{x} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^n$.

(a) (5 pts) Consider a set of vectors $\{\vec{u}_1, \ldots, \vec{u}_k\}$, all of which are in $\mathbb{R}^n$. Suppose that for some $\vec{v} \in \mathbb{R}^n$ we have that $\vec{v}^\top A\vec{u}_i = 0$ for all $1 \leq i \leq k$. **Show that $\vec{v}^\top A\vec{w} = 0$ for any vector $\vec{w} \in \operatorname{span}(\vec{u}_1, \ldots, \vec{u}_k)$.**

(b) (5 pts) Let $\vec{v} \in \mathbb{R}^n$ and $\vec{w} \in \mathbb{R}^n$ with $\vec{v} \neq \vec{0}$ and $\vec{w} \neq \vec{0}$. **Show that if $\vec{v}$ and $\vec{w}$ are such that $\vec{v}^\top A\vec{w} = 0$, they must be linearly independent.**

[1]Two vectors $\vec{v}$ and $\vec{w}$ are "*conjugate in $A$*" if $\vec{v}^\top A\vec{w} = 0$, and such pairs of vectors are useful in the algorithm, giving it the name.

(c) (5 pts) Recall that $A \in \mathbb{S}^n_{++}$ is a symmetric positive definite matrix. Suppose $\{\vec{u}_1, \ldots, \vec{u}_n\}$ are a set of vectors in $\mathbb{R}^n$ such that $\vec{u}_i^\top A \vec{u}_j = 0$ for all $i, j$ where $i \neq j$. Consider the matrix,

$$U = \begin{bmatrix} \vec{u}_1 & \vec{u}_2 & \ldots & \vec{u}_n \end{bmatrix}. \tag{22}$$

**Show that** $\operatorname{rank}\left(U^\top A U\right) = n$**.**

This implies that $U$ has full rank (you do *not* need to show this), and hence $\{\vec{u}_1, \ldots, \vec{u}_n\}$ forms a basis for $\mathbb{R}^n$. We call this a *conjugate basis* in $A$ for $\mathbb{R}^n$.

(d) (7 pts) Recall that $A \in \mathbb{S}^n_{++}$ is a symmetric positive definite matrix, $\vec{x} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^n$. Consider

$$\vec{x}^\star = \sum_{i=1}^{n} \frac{\vec{u}_i^\top \vec{b}}{\vec{u}_i^\top A \vec{u}_i} \vec{u}_i, \tag{23}$$

where $\{\vec{u}_1, \ldots, \vec{u}_n\}$ form a conjugate basis in $A$ for $\mathbb{R}^n$ (i.e. $\vec{u}_i^\top A \vec{u}_j = 0$ for all $i, j$ where $i \neq j$) as defined in part (c). **Show that $\vec{x}^\star$ is a solution to $A\vec{x} = \vec{b}$.** *Show your work.*

*HINT: For any vectors $\vec{c} \in \mathbb{R}^n$ and $\vec{d} \in \mathbb{R}^n$ and any basis $\{\vec{z}_1, \ldots, \vec{z}_n\}$ of $\mathbb{R}^n$,*

$$\vec{c} = \vec{d} \iff \textit{for all } 1 \leq i \leq n, \textit{ we have } \vec{z}_i^\top \vec{c} = \vec{z}_i^\top \vec{d}. \tag{24}$$

(e) (9 pts) Recall that $A \in \mathbb{S}^n_{++}$ is a symmetric positive definite matrix, $\vec{x} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^n$. Consider a conjugate basis in $A$ given by $\{\vec{u}_1, \cdots, \vec{u}_n\}$. This implies that $\vec{u}_i^\top A\vec{u}_j = 0$ for all $i, j$ where $i \neq j$. Then the $(k+1)$-th iterate of conjugate gradient descent, $\vec{x}_{k+1}$, is calculated as:

$$\vec{x}_{k+1} = \sum_{i=1}^{k} \frac{\vec{u}_i^\top \vec{r}_i}{\vec{u}_i^\top A\vec{u}_i}\vec{u}_i, \tag{25}$$

where $\vec{r}_i = \vec{b} - A\vec{x}_i$. Recall from part (d) that

$$\vec{x}^\star = \sum_{i=1}^{n} \frac{\vec{u}_i^\top \vec{b}}{\vec{u}_i^\top A\vec{u}_i}\vec{u}_i. \tag{26}$$

**Show that $\vec{x}_{n+1} = \vec{x}^\star$.** This means that the conjugate gradient method converges to the solution of $A\vec{x} = \vec{b}$ in $n$ iterations.

*HINT: Prove and use the fact that $\vec{u}_k^\top \vec{r}_k = \vec{u}_k^\top \vec{b}$ for every $1 \leq k \leq n$.*

*HINT: Use part (a).*

(f) (8 pts) Recall that $A \in \mathbb{S}^n_{++}$ is a symmetric positive definite matrix, $\vec{x} \in \mathbb{R}^n$, and $\vec{b} \in \mathbb{R}^n$. Consider $f(\vec{x}) = \frac{1}{2}\vec{x}^\top A\vec{x} - \vec{b}^\top \vec{x}$. Note that the conjugate gradient iterates in (25) can be recursively expressed as

$$\vec{x}_{k+1} = \vec{x}_k + \frac{\vec{u}_k^\top \vec{r}_k}{\vec{u}_k^\top A\vec{u}_k}\vec{u}_k, \tag{27}$$

where $\vec{r}_k = \vec{b} - A\vec{x}_k$.

**Show that for all** $1 \leq k \leq n$ **we have that** $f(\vec{x}_{k+1}) \leq f(\vec{x}_k)$**.** Thus, if we use the conjugate gradient method to minimize the the function $f(\vec{x})$, the objective function will be non-increasing in every iteration.

*HINT: Use the first order condition of convexity and part (a).*

*HINT: Use the fact that $\vec{u}_k^\top \vec{r}_k = \vec{u}_k^\top \vec{b}$ for every $1 \leq k \leq n$.*

[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.]

# Final Instructions

## Read the following instructions before the exam.

**There are 13 problems of varying numbers of 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 each page. If a page is found without a student ID, and some pages from your exam go missing, we will have no way of giving you credit for those pages.** All exam pages will be separated during scanning.

You may consult TWO handwritten 8.5” $\times$ 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.