EECS 127/227AT | Optimization Models in Engineering | UC Berkeley | Fall 2022

Midterm

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

   (a) (1 pts) What's your favorite building in Berkeley?

   (b) (1 pts) What is a hobby or activity that makes you happy?

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

3. Orthonormal Matrices (6 pts)

Prove the following identities.

(a) (3 pts) Let $\vec{x} \in \mathbb{R}^n$ be a vector, and let $U \in \mathbb{R}^{m \times n}$, where $m \geq n$, be an orthonormal matrix. **Prove the following equality:**

$$\|U\vec{x}\|_2 = \|\vec{x}\|_2 . \tag{1}$$

(b) (3 pts) Let $A \in \mathbb{R}^{n \times n}$ be a square matrix, and suppose $A = QR$ is a QR decomposition of A. **Compute $\|R\|_F$ in terms of $\|A\|_F$.**

[Extra page for scratch work that will not be graded.]

4. Vector Calculus (7 pts)

Let $\vec{x} \in \mathbb{R}^n$ and $A \in \mathbb{R}^{n \times n}$. Let $f(\vec{x}) \doteq \|A\vec{x}\|_2^2$.

(a) (4 pts) **Calculate the gradient of $f(\vec{x})$ with respect to $\vec{x}$.** *Show your work.*

(b) (3 pts) **Calculate the Hessian of $f(\vec{x}) \doteq \|A\vec{x}\|_2^2$ with respect to $\vec{x}$.** *You do not need to show your work.*

5. Convexity (9 pts)

(a) (4 pts) Let $f_1, \ldots, f_k \colon \mathbb{R}^n \to \mathbb{R}$ be convex functions. **Prove that**

$$\textbf{the set} \qquad S \doteq \{\vec{x} \in \mathbb{R}^n \mid f_i(\vec{x}) \leq 0 \quad \forall i = 1, \ldots, k\} \qquad \textbf{is convex.} \tag{2}$$

(b) (5 pts) Let $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$. **Prove that $f\colon \mathbb{R}^n \to \mathbb{R}$ given by $f(\vec{x}) \doteq \max_{1 \leq i \leq n} |x_i|$ is convex.**

6. Low-Rank Approximation (4 pts)

Let $A \in \mathbb{R}^{4\times 3}$ be a matrix whose full SVD is

$$A = \underbrace{\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ 0 & 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{bmatrix}}_{V^\top}. \tag{3}$$

Give the best rank-2 approximation to A, i.e., the solution to the problem

$$\operatorname*{argmin}_{\substack{B\in\mathbb{R}^{4\times 3} \\ \mathrm{rk}(B)\leq 2}} \|A - B\|_F^2 . \tag{4}$$

No justification is necessary.

NOTE: Please leave your answer in terms of a matrix product.

[Extra page for scratch work that will not be graded.]

7. Power Iteration and SVD (22 pts)

In this problem, we will discuss how to efficiently compute singular values and vectors using an algorithm called "power iteration" that provides eigenvalues and eigenvectors. You do not need any prior knowledge about the algorithm to complete this problem, other than the description below.

The "power iteration" algorithm, denoted by POWITER, operates as follows:

- For a symmetric positive semidefinite matrix $B \in \mathbb{S}_+^n$, $\text{PowIter}(B) = (\lambda, \vec{v})$, where λ is the largest eigenvalue of B and $\vec{v}$ is a corresponding unit eigenvector.
- For a non-square, non-symmetric, or non-positive semidefinite matrix C, $\text{PowIter}(C) = \text{Error}$.

(a) (4 pts) Let $A \in \mathbb{R}^{m \times n}$ be known to you. **Explain how to use POWITER to compute a top right singular vector, i.e., the first column $\vec{v}_1$ of V in an SVD of $A = U\Sigma V^\top$, as well as its corresponding singular value σ_1.** *A 1-2 sentence algorithm description or pseudocode will suffice.*

(b) (6 pts) Let $B \in \mathbb{S}^n_+$ be a symmetric positive semidefinite matrix with eigenpairs $(\lambda_1, \vec{w}_1), \ldots, (\lambda_n, \vec{w}_n)$ where $\lambda_1 \geq \cdots \geq \lambda_n \geq 0$. **Prove that the matrix $D \doteq B - \lambda_1 \vec{w}_1 \vec{w}_1^\top$ is a symmetric positive semidefinite matrix with eigenpairs $(0, \vec{w}_1), (\lambda_2, \vec{w}_2), \ldots, (\lambda_n, \vec{w}_n)$.**

(c) (5 pts) Let $A \in \mathbb{R}^{m \times n}$ and $k \leq \min\{m, n\}$ be known to you. **Explain how to use** POWITER **to compute the top k right singular vectors of A, i.e., the first k columns of V in an SVD $A = U\Sigma V^{\top}$, as well as their corresponding singular values.** *A 1-2 sentence algorithm description or pseudocode will suffice.*

NOTE: You may use the result from part (b), even if you haven't proved it.

(d) (4 pts) Suppose that you know how to compute any number of right singular vectors of any matrix using POWITER (regardless of whether or not you completed part (c)). Let $A \in \mathbb{R}^{m \times n}$ and $r \doteq \mathrm{rk}(A)$ be known to you. **Explain how to compute a basis for $\mathcal{R}(A^\top)$.** *A 1-2 sentence solution will suffice.*

(e) (3 pts) Let $A \in \mathbb{R}^{m \times n}$ be unknown to you (so you cannot compute its SVD or even use POWITER). Suppose that you are given a basis for $\mathcal{R}(A^\top)$. **Explain how to compute a basis for $\mathcal{N}(A)$.** *A one sentence solution will suffice.*

8. Matrix Square Root (9 pts)

Let $A, B \in \mathbb{S}^n_{++}$ be symmetric positive definite matrices.

As B is symmetric, it has an orthonormal eigendecomposition $B = V\Lambda V^\top$. Since B is positive definite, we can define its matrix square root as follows $B^{1/2} = V\Lambda^{1/2}V^\top$, where $\Lambda^{1/2}$ is a diagonal matrix whose entries are the square roots of the corresponding entries of Λ. We denote the inverse of $B^{1/2}$ as $B^{-1/2}$. Finally, define $C \doteq B^{-1/2}AB^{-1/2}$.

Prove that the maximum eigenvalue of C is $\lambda^\star$, where

$$\lambda^\star \doteq \max_{\vec{x} \neq \vec{0}} \frac{\vec{x}^\top A\vec{x}}{\vec{x}^\top B\vec{x}}. \tag{5}$$

[Extra page for scratch work that will not be graded.]

9. **Gradient Descent with A Wide Matrix (31 pts)**

Consider a matrix $X \in \mathbb{R}^{n \times d}$ with $n < d$ and a vector $\vec{y} \in \mathbb{R}^n$, both of which are known and given to you. Suppose X has full row rank.

(a) (3 pts) **Consider the following problem:**

$$X\vec{w} = \vec{y} \tag{6}$$

where $\vec{w} \in \mathbb{R}^d$ is unknown. How many solutions does Equation (6) have? *Justify your answer.*

(b) (5 pts) Consider the minimum-norm problem

$$\vec{w}_\star = \operatorname*{argmin}_{\substack{\vec{w} \in \mathbb{R}^d \\ X\vec{w} = \vec{y}}} \|\vec{w}\|_2^2 . \tag{7}$$

We know that the optimal solution to this problem is $\vec{w}_\star = X^\top (XX^\top)^{-1}\vec{y}$. Now let $X = U\Sigma V^\top = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} V^\top$ be the SVD of X, where $\Sigma_1 \in \mathbb{R}^{n\times n}$. Recall that this is possible because $n < d$ and X is full row rank. **Prove that $\vec{w}_\star$ is given by**

$$\vec{w}_\star = V \begin{bmatrix} \Sigma_1^{-1} \\ 0 \end{bmatrix} U^\top \vec{y}. \tag{8}$$

All steps must be shown and justified for full credit.

(c) (5 pts) Let $\eta > 0$, and I be the identity matrix of appropriate dimension. Using the SVD $X = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} V^\top$, **prove the following identity for all positive integers $i > 0$:**

$$(I - \eta X^\top X)^i = V \left(I - \eta \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} \right)^i V^\top. \tag{9}$$

All steps must be shown and justified for full credit.

(d) (9 pts) Recall that $X \in \mathbb{R}^{n\times d}$, and that we can write the SVD of X as $X = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} V^\top$. We will use gradient descent to solve the minimization problem

$$\min_{\vec{w}\in\mathbb{R}^d} \frac{1}{2} \|X\vec{w} - \vec{y}\|_2^2 \tag{10}$$

with step-size $\eta > 0$. Let $\vec{w}_0 = \vec{0}$ be the initial state, and $\vec{w}_k$ be the k^{th} iterate of gradient descent. **Use the identity:**

$$(I - \eta X^\top X)^i = V \left(I - \eta \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} \right)^i V^\top. \tag{11}$$

to prove that after k steps, we have

$$\vec{w}_k = \eta \sum_{i=0}^{k-1} V \left(I - \eta \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} \right)^i \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} U^\top \vec{y}. \tag{12}$$

HINT: Remember to set $\vec{w}_0 = \vec{0}$.

(e) (9 pts) Now let $0 < \eta < \frac{1}{\sigma_1^2}$, where σ_1 denotes the maximum singular value of $X = U \begin{bmatrix} \Sigma_1 & 0 \end{bmatrix} V^\top$. Let $\vec{w}_k$ be given as

$$\vec{w}_k = \eta \sum_{i=0}^{k-1} V \left(I - \eta \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix} \right)^i \begin{bmatrix} \Sigma_1 \\ 0 \end{bmatrix} U^\top \vec{y}. \tag{13}$$

and let $\vec{w}_\star$ be the minimum norm solution given as

$$\vec{w}_\star = V \begin{bmatrix} \Sigma_1^{-1} \\ 0 \end{bmatrix} U^\top \vec{y}. \tag{14}$$

Prove that $\lim_{k \to \infty} \vec{w}_k = \vec{w}_\star$**.**

HINT: You may use the following result without proof. When all eigenvalues of $A \in \mathbb{R}^{n \times n}$ have magnitude < 1, we have the identity $(I - A)^{-1} = I + A + A^2 + \dots$.

[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.]

[Extra page for scratch work that will not be graded.]

Midterm Instructions

Read the following instructions before the exam.

There are 9 problems of varying numbers of points. You have 120 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 24 pages on the exam, so there should be 12 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 ONE 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.

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.