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: ______________________________

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 your favorite restaurant in Berkeley?**

   (b) (1 pts) **What's some music that makes you happy?**

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

4. Fun with Rank (12 pts)

Consider two matrices $A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times p}$. Let $\mathcal{R}(A)$ denote the range (i.e. column) space of the matrix A.

(a) (4 pts) Prove that $\mathcal{R}(AB) \subseteq \mathcal{R}(A)$.

HINT: $\mathcal{R}(A) = \{\vec{y} : \vec{y} = A\vec{x}$ *for some* $\vec{x} \in \mathbb{R}^n\}$.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(b) (4 pts) Prove that the following inequality holds:

$$0 \leq \text{rank}(AB) \leq \min\{\text{rank}(A), \text{rank}(B)\}.$$

HINT: Recall that the rank of a matrix is the dimension of its range space.

HINT: You may use the result of part (a), and the fact that the rank of any matrix is the same as the rank of its transpose.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(c) (4 pts) Give an example of matrices A, B such that $\operatorname{rank}(A) \neq 0, \operatorname{rank}(B) \neq 0$ but $\operatorname{rank}(AB) = 0$, by finding suitable values of $x, y, z \in \mathbb{R}$ in the following matrices:

$$A = \begin{bmatrix} x & y \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 1 & z \end{bmatrix}.$$

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

5. Compact SVD (4 pts)

Consider two sets of orthonormal vectors $\{\vec{p}_1, \vec{p}_2\} \subset \mathbb{R}^m$ with $\vec{p}_1 \perp \vec{p}_2$, and $\{\vec{q}_1, \vec{q}_2\} \subset \mathbb{R}^n$ with $\vec{q}_1 \perp \vec{q}_2$. Let $C \in \mathbb{R}^{m \times n}$ be defined as

$$C = \vec{p}_1 \vec{q}_1^\top + \frac{1}{2} \vec{p}_2 \vec{q}_2^\top .$$

Write the compact SVD representation of matrix C in terms of $\vec{p}_1, \vec{p}_2, \vec{q}_1, \vec{q}_2$. That is, compute the SVD matrices $U_r \in \mathbb{R}^{m \times r}$, $\Sigma_r \in \mathbb{R}^{r \times r}$, $V_r \in \mathbb{R}^{n \times r}$, such that $C = U_r \Sigma_r V_r^\top$.

6. Exploring SVD, Least Squares, and Min-Norm (13 pts)

(a) (4 pts) Let $A \in \mathbb{R}^{m \times n}$ be a matrix with rank $r > 0$. Consider the equation $A\vec{x} = \vec{b}$ for some $\vec{b} \in \mathcal{R}(A)$. Show that $\vec{x}_0 = V_r \Sigma_r^{-1} U_r^\top \vec{b}$ is a solution to $A\vec{x} = \vec{b}$, where the compact SVD of A is $A = U_r \Sigma_r V_r^\top$. *Show your work.*

HINT: Remember that $U_r U_r^\top$ isn't necessarily the identity, but $U_r U_r^\top \vec{d} = \vec{d}$ for any $\vec{d} \in \mathcal{R}(U_r)$.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(b) (4 pts) Let $A \in \mathbb{R}^{m \times n}$ be a matrix with $m < n$ and rank $r > 0$, and let $\vec{b} \in \mathcal{R}(A)$.

Also, let the compact and full SVD representations of A be, respectively,

$$\underbrace{A}_{m \times n} = \underbrace{U_r}_{m \times r} \underbrace{\Sigma_r}_{r \times r} \underbrace{V_r^\top}_{r \times n}, \qquad A = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \underbrace{\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}}_{m \times n} \begin{bmatrix} V_r^\top \\ V_{n-r}^\top \end{bmatrix},$$

where $U_{m-r} \in \mathbb{R}^{r \times (m-r)}$, $V_{n-r} \in \mathbb{R}^{r \times (n-r)}$.

Show that $V_r \Sigma_r^{-1} U_r^\top \vec{b} + V_{n-r} \vec{z}$ is a solution to $A\vec{x} = \vec{b}$, for any $\vec{z} \in \mathbb{R}^{n-r}$.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(c) (5 pts) Again, let $A \in \mathbb{R}^{m\times n}$ be a matrix with $m < n$ and rank r, and let $\vec{b} \in \mathcal{R}(A)$.

Also, let the compact and full SVD representations of A be, respectively,

$$\underbrace{A}_{m\times n} = \underbrace{U_r}_{m\times r}\underbrace{\Sigma_r}_{r\times r}\underbrace{V_r^\top}_{r\times n}, \qquad A = \begin{bmatrix} U_r & U_{m-r} \end{bmatrix} \underbrace{\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}}_{m\times n} \begin{bmatrix} V_r^\top \\ V_{n-r}^\top \end{bmatrix},$$

where $U_{m-r} \in \mathbb{R}^{r\times(m-r)}$, $V_{n-r} \in \mathbb{R}^{r\times(n-r)}$.

Let $\vec{x}^\star$ be the solution to the following problem:

$$\vec{x}^\star = \operatorname*{argmin}_{\vec{x}\in\mathbb{R}^n} \|\vec{x}\|_2^2 \quad \text{s.t.} \quad A\vec{x} = \vec{b}.$$

Find $\vec{x}^\star$ and justify your answer.

HINT: You may use the fact that $\{V_r\Sigma_r^{-1}U_r^\top\vec{b} + V_{n-r}\vec{z} : \vec{z} \in \mathbb{R}^{n-r}\}$ *is the set of all solutions to* $A\vec{x} = \vec{b}$*, without proof.*

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

7. Convexity (10 pts)

(a) (6 pts) Let $f : \mathbb{R}^n \to \mathbb{R}$ be defined as $f(\vec{x}) = \|A\vec{x}\|_2^2$, where $A \in \mathbb{R}^{m \times n}$ is a matrix. Is f a convex function? Prove or disprove.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(b) (4 pts) Let $g, h : \mathbb{R}^n \to \mathbb{R}$ be fixed twice-differentiable convex functions, and fix real numbers $a, b > 0$. Define $f : \mathbb{R}^n \to \mathbb{R}$ by $f(\vec{x}) = a \cdot g(\vec{x}) + b \cdot h(\vec{x})$ for each $\vec{x} \in \mathbb{R}^n$. Prove f is a convex function.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

8. Vector Calculus (12 pts)

(a) (6 pts) Compute the gradient and Hessian with respect to $\vec{x}$ of the function $f\colon \mathbb{R}^n \to \mathbb{R}$

$$f(\vec{x}) = 1 - (\vec{a}^\top \vec{x})^2, \tag{1}$$

where $\vec{a} \in \mathbb{R}^n$.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(b) (6 pts) Consider $f(\vec{x}) = \sum_{i=1}^{m} \log\left(b_i - \vec{a}_i^\top \vec{x}\right)$, where $\vec{a}_i \in \mathbb{R}^n$ for $i = 1, \ldots, m$, and $b_1, b_2, \ldots, b_m > 0$.

The domain of f is the set $\{\vec{x} \in \mathbb{R}^n \mid b_i - \vec{a}_i^\top \vec{x} > 0 \text{ for all } i = 1, \ldots, m\}$, which is assumed to be nonempty.

Compute the gradient of $f(\vec{x})$ with respect to $\vec{x}$.

HINT: Consider what happens in the special case of $f(x) = log(b - x)$ for a scalar variable $x \in \mathbb{R}$. Then, use the chain rule.

HINT: Recall that $\frac{d}{dx} \log(x) = \frac{1}{x}$.

9. Best Approximations (24 pts)

We start by recalling the Eckart-Young Theorem: Consider any square matrix $C \in \mathbb{R}^{n \times n}$ and assume that we can write its full singular value decomposition as $C = U\Sigma V^\top$ where Σ is the $n \times n$ diagonal matrix with distinct diagonal entries $\sigma_1 > \sigma_2 > \cdots > \sigma_n > 0$. Then, for $0 \leq k \leq n$, the Eckart-Young Theorem for Frobenius norm states that

$$C_k = U_k \Sigma_k V_k^\top = \underset{\substack{B \in \mathbb{R}^{n \times n} \\ \mathrm{rank}(B) \leq k}}{\mathrm{argmin}} \|C - B\|_F$$

where $U_k \in \mathbb{R}^{n \times k}$ is the matrix consisting of the first k columns of U, $V_k \in \mathbb{R}^{n \times k}$ is the matrix consisting of the first k columns of V, and Σ_k is the $k \times k$ diagonal matrix with the top-k singular values $\sigma_1 > \cdots > \sigma_k$ as its diagonal entries.

(a) (8 pts) Now, consider any square matrix $C \in \mathbb{R}^{n \times n}$ and let $C_k \in \mathbb{R}^{n \times n}$ denote its best rank-k approximation in the Frobenius norm. Then, for any orthonormal matrix $W \in \mathbb{R}^{n \times n}$, show that

$$WC_kW^\top = \underset{\substack{B \in \mathbb{R}^{n \times n} \\ \mathrm{rank}(B) \leq k}}{\mathrm{argmin}} \|WCW^\top - B\|_F.$$

 All Rights Reserved. This may not be publicly shared without explicit permission.

For the remainder of this problem, we consider a square matrix $A \in \mathbb{R}^{n \times n}$ and assume A has full rank. Using Gram-Schmidt Orthonormalization (GSO), we can write the matrix A as

$$A = QR \tag{2}$$

where $Q \in \mathbb{R}^{n \times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n \times n}$ is an upper triangular matrix.

(b) (4 pts) Find the best rank-n approximation to $AA^\top$ in the Frobenius norm. *Justify your answer.*

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(c) (6 pts) Recall that $A \in \mathbb{R}^{n\times n}$ is a square matrix with full rank, and we can write the matrix A as

$$A = QR \tag{3}$$

where $Q \in \mathbb{R}^{n\times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n\times n}$ is an upper triangular matrix.

Assume that $R = \text{diag}(r_1, r_2, \ldots, r_n) \in \mathbb{R}^{n\times n}$ is a diagonal matrix with

$$|r_1| > |r_2| > \cdots > |r_n|,$$

and all $r_1, r_2, \ldots, r_n$ are real numbers. Let $k < n$. Then, show that the best rank-k approximation to $AA^\top$ in the Frobenius norm is $QSQ^\top$, where S is a diagonal matrix defined as

$$S = \text{diag}(r_1^2, \ldots, r_k^2, 0, \ldots, 0) \in \mathbb{R}^{n\times n}.$$

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

(d) (6 pts) Recall that $A \in \mathbb{R}^{n\times n}$ is a square matrix with full rank, and we can write the matrix A as

$$A = QR \tag{4}$$

where $Q \in \mathbb{R}^{n\times n}$ is an orthonormal matrix and $R \in \mathbb{R}^{n\times n}$ is an upper triangular matrix.

Now, we no longer assume that R is diagonal. Let $k < n$ and assume that the best rank-k approximation to $RR^\top \in \mathbb{R}^{n\times n}$ in the Frobenius norm is given by $G \in \mathbb{R}^{n\times n}$. Then, using the result of part (a), show that the best rank-k approximation to $AA^\top$ in the Frobenius norm is given by $QGQ^\top$.

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

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

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

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

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

[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.]

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

[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.]

© UCB EECS 127/227AT, Fall 2023. All Rights Reserved. This may not be publicly shared without explicit permission.

Midterm Instructions

Read the following instructions before the exam.

There are 6 problems of varying numbers of points. You have 80 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 22 pages on the exam, so there should be 11 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” $\times$ 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.