# 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 (2 pts)**

   (a) (1 pts) What is your favorite book or book series?

   (b) (1 pts) Who is the speaker or writer of your favorite inspirational quote?

3. **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.*

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

**4. Singular Values (10 pts)**

(a) (4 pts) Suppose $A \in \mathbb{R}^{3\times 2}$ is a matrix such that $A^\top A$ is given by

$$A^\top A = \begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}. \tag{1}$$

**What are the singular values of $A$?** *Justify your answer(s).*

(b) (6 pts) Suppose that $B \in \mathbb{R}^{3\times 2}$ has singular values $0$, $\sqrt{2}$, and $\sqrt{7}$. Let $C = \begin{bmatrix} B & -B & 3I_3 \end{bmatrix} \in \mathbb{R}^{3\times 7}$, where $I_3 \in \mathbb{R}^{3\times 3}$ is the $3 \times 3$ identity matrix. **What are the singular values of $C$?** *Show your work and justify your answer(s).*

*HINT: Consider the matrix $CC^\top \in \mathbb{R}^{3\times 3}$.*

**5. Convex Functions (10 pts)**

(a) (4 pts) **Show that the function $f: \mathbb{R}^n \to \mathbb{R}$ given by $f(\vec{x}) \doteq \|\vec{x}\|_2^2$ is convex.**

*NOTE*: You may use the gradient and Hessian of $f$, which were computed in lecture and homework, but the convexity of $f$ should be proved via "first principles" (zeroth/first/second order conditions, or other equivalent conditions for convexity).

(b) (6 pts) **Is the function $g: \mathbb{R}^n \to \mathbb{R}$ given by $g(\vec{x}) \doteq e^{\|\vec{x}\|_2^2}$ convex? If $g$ is convex, prove it; if $g$ is not convex, give an example $\vec{x}, \vec{y} \in \mathbb{R}^n$ and $\theta \in [0, 1]$ such that $g(\theta\vec{x} + (1-\theta)\vec{y}) > \theta g(\vec{x}) + (1-\theta)g(\vec{y})$.**

*NOTE*: One (short) solution to this problem does not use gradients or Hessians, but it is fine if yours does. In particular, the gradient and Hessian of $g$ were derived in homework; if you want to use these quantities, please derive them here. You may use without proof the gradient and Hessian of $f(\vec{x}) \doteq \|\vec{x}\|_2^2$.

*This is an extra page for scratch work that will not be graded unless you tell us in the original problem space.*
*The exam will continue on the following page.*

**6. Spectrahedron (7 pts)**

Let $F_1, \ldots, F_n \in \mathbb{R}^{m \times m}$ be symmetric matrices. Define the set $S \subseteq \mathbb{R}^n$, known as a *spectrahedron*, by

$$S \doteq \left\{ \vec{x} \in \mathbb{R}^n \;\middle|\; \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \quad \sum_{i=1}^{n} x_i F_i \succeq 0 \right\}. \tag{2}$$

Here $A \succeq 0$ means that $A$ is symmetric PSD. **Show that $S$ is a convex set.**

*HINT: You can use without proof that convex combinations of symmetric PSD matrices are symmetric PSD.*

**7. Gradient Descent on Quadratic (6 pts)**

Let $a, \eta \in \mathbb{R}$ be such that $\eta > 0$ and $0 < a < 1/\eta$. Define the function $f : \mathbb{R} \to \mathbb{R}$ by

$$f(x) \doteq \frac{1}{2}ax^2, \qquad \text{for all } x \in \mathbb{R}. \tag{3}$$

We run gradient descent on $f$ with constant step size $\eta$ and fixed initialization $x_0 = 1$ to get iterates $(x_t)_{t=0}^{\infty}$, i.e.,

$$x_{t+1} \doteq x_t - \eta\frac{\mathrm{d}f}{\mathrm{d}x}(x_t) \qquad \text{for all } t \geq 0, \qquad \text{and} \quad x_0 = 1. \tag{4}$$

**Complete the following tasks:**

- **compute the derivative of $f$ (denoted $\frac{\mathrm{d}f}{\mathrm{d}x}$ or $f'$);**
- **write the update rule for $x_{t+1}$ in terms of $x_t$, $a$, and $\eta$;**
- **write an expression for $x_t$ in terms of $x_0$, $a$, $\eta$, and $t$;**
- **and compute the limit $\lim_{t\to\infty} x_t$.**

*Show your work and justify your answer(s).*

**8. Vector Calculus (14 pts)**

(a) (8 pts) Let $\vec{a} \in \mathbb{R}^n$ be a fixed vector, and $b \in \mathbb{R}$ be a fixed scalar. **Compute the gradient and Hessian of the function** $f\colon \mathbb{R}^n \to \mathbb{R}$ **given by**

$$f(\vec{x}) \doteq \sin(\vec{a}^\top \vec{x} - b). \tag{5}$$

*Show your work and justify your answer(s).*

(b) (6 pts) Let $\vec{u} \in \mathbb{R}^n$ be a fixed vector. **Compute the Jacobian of the function $f\colon \mathbb{R}^n \to \mathbb{R}^n$ given by**

$$f(\vec{x}) \doteq (\vec{u}^\top \vec{x})\vec{u}. \tag{6}$$

*Show your work and justify your answer(s).*

*HINT: One (short) solution to this problem starts by rewriting $f(\vec{x})$ as a matrix-vector product, but you can do this problem any way you want.*

**9. Factorizations of PSD Matrices (16 pts)**

Let $k, n$ be positive integers, with $k \leq n$. In this problem, we prove that $A \in \mathbb{R}^{n \times n}$ is a symmetric PSD matrix of rank $k$ if and only if it can be written as $A = PP^\top$ for some matrix $P \in \mathbb{R}^{n \times k}$ which has full column rank.

(a) (8 pts) Let $A \in \mathbb{R}^{n \times n}$ be a symmetric PSD matrix with rank $k$. **Prove that there exists another matrix** $P \in \mathbb{R}^{n \times k}$ **with full column rank, i.e.,** $\operatorname{rank}(P) = k$**, such that** $A = PP^\top$**.**

*HINT: Recall that $A$ is a square and symmetric $n \times n$ matrix, while $P$ is a tall $n \times k$ matrix.*

Let $k, n$ be positive integers, with $k \leq n$.

(b) (8 pts) Let $P \in \mathbb{R}^{n \times k}$ be a matrix with full column rank, i.e., $\operatorname{rank}(P) = k$. **Prove that if we define $A \doteq PP^\top$, then $A \in \mathbb{R}^{n \times n}$ is a symmetric PSD matrix of rank $k$.**

*HINT: We know two ways to show that $\operatorname{rank}(A) = k$. One uses the rank-nullity theorem and that $\mathcal{N}\left(B^\top B\right) = \mathcal{N}(B)$ for any matrix $B$ in order to compute the rank of $A = PP^\top$. The other uses the SVD of $P$.*

**10.** $\ell^p$ **Norms (8 pts)**

Let $n$ be a positive integer. Recall that for $1 \leq p \leq \infty$ the $\ell^p$ norm on $\mathbb{R}^n$ is defined as

$$\|\vec{x}\|_p \doteq \begin{cases} \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} & \text{if } 1 \leq p < \infty \\ \max_{i \in \{1,\dots,n\}} |x_i| & \text{if } p = \infty, \end{cases} \qquad \text{for all } \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{R}^n. \tag{7}$$

Let $A \in \mathbb{R}^{m \times n}$ be a matrix. Let $\vec{r}_i \in \mathbb{R}^n$ be the $i^{\text{th}}$ row of $A$, i.e.,

$$A = \begin{bmatrix} \vec{r}_1^\top \\ \vdots \\ \vec{r}_m^\top \end{bmatrix}. \tag{8}$$

**Prove the identity**

$$\max_{\substack{\vec{v} \in \mathbb{R}^n \\ \|\vec{v}\|_2 = 1}} \|A\vec{v}\|_\infty = \max_{i \in \{1,\dots,m\}} \|\vec{r}_i\|_2 . \tag{9}$$

*HINT: The Cauchy-Schwarz inequality may be useful. Think about when equality holds.*

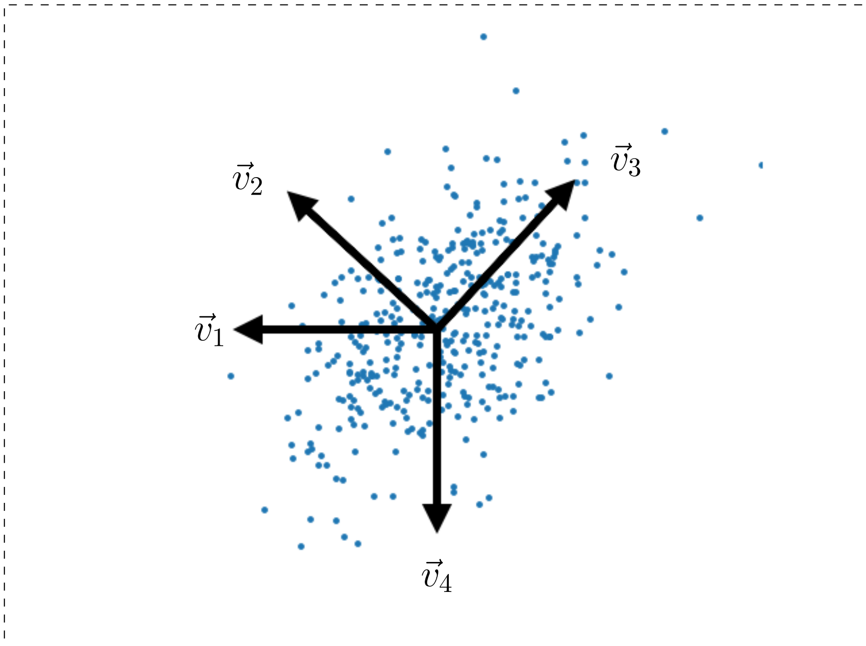
 

*This is an extra page for scratch work that will not be graded unless you tell us in the original problem space.*
*The exam will continue on the following page.*

**11. PCA and Regression (34 pts)**

(a) (4 pts) Given the following plot of data in $\mathbb{R}^2$ (i.e., each dot is a data point in $\mathbb{R}^2$) and candidate unit vectors $\vec{v}_1, \vec{v}_2, \vec{v}_3, \vec{v}_4 \in \mathbb{R}^2$, **identify the candidate vectors which could be the first principal component and second principal component of the data (and specify which is which).** *You do not need to show your work for this subpart.*



(b) (6 pts) Suppose we have pairs of data $(\vec{x}_1, y_1), \ldots, (\vec{x}_n, y_n) \in \mathbb{R}^d \times \mathbb{R}$, where $n > d$. As usual, we arrange these data points into a matrix and vector, i.e.,

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad \vec{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{R}^n. \tag{10}$$

*Assume that $X$ is centered, i.e., each column has mean zero: $(1/n)\sum_{i=1}^n \vec{x}_i = \vec{0}_d$*, where $\vec{0}_d$ is the zero vector in $\mathbb{R}^d$. Suppose that $X$ has compact SVD given by $X = U_d \Sigma_d V_d^\top$ where

$$U_d = \begin{bmatrix} \vec{u}_1, \ldots, \vec{u}_d \end{bmatrix} \in \mathbb{R}^{n \times d}, \qquad V_d = \begin{bmatrix} \vec{v}_1, \ldots, \vec{v}_d \end{bmatrix} \in \mathbb{R}^{d \times d}, \qquad \Sigma_d = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_d \end{bmatrix} \in \mathbb{R}^{d \times d} \tag{11}$$

where $\sigma_1 > \sigma_2 > \cdots > \sigma_d > 0$. **From this SVD, identify the top $k$ principal components of the data $\{\vec{x}_1, \ldots, \vec{x}_n\} \subseteq \mathbb{R}^d$, where $k \leq d$.** *You do not need to show your work for this subpart.*

*HINT: Recall that the first principal component solves the optimization problem* $\underset{\vec{w} \in \mathbb{R}^d \colon \|\vec{w}\|_2 = 1}{\operatorname{argmax}} \; \vec{w}^\top X^\top X \vec{w}$.

Let $k \leq d$ be a positive integer. Suppose the matrix $X \in \mathbb{R}^{n \times d}$ has rows $\vec{x}_i^\top$, i.e.,

$$X = \begin{bmatrix} \vec{x}_1^\top \\ \vdots \\ \vec{x}_n^\top \end{bmatrix} \in \mathbb{R}^{n \times d}. \tag{10}$$

(c) (4 pts) Suppose that $P = \begin{bmatrix} \vec{p}_1, \ldots, \vec{p}_k \end{bmatrix} \in \mathbb{R}^{d \times k}$ is a matrix with columns $\vec{p}_j$. Let $Z = XP$, and let the entries of $Z$ be $z_{ij}$, i.e.,

$$Z = \begin{bmatrix} z_{11} & \cdots & z_{1k} \\ \vdots & \ddots & \vdots \\ z_{n1} & \cdots & z_{nk} \end{bmatrix} \in \mathbb{R}^{n \times k}. \tag{12}$$

**Give an expression for $z_{ij}$ in terms of $\vec{x}_i$ and $\vec{p}_j$.** *You do not need to show your work for this subpart.*

Let $X \in \mathbb{R}^{n\times d}$ have compact SVD given by $X = U_d \Sigma_d V_d^\top$ where

$$U_d = \begin{bmatrix}\vec{u}_1, \dots, \vec{u}_d\end{bmatrix} \in \mathbb{R}^{n\times d}, \qquad V_d = \begin{bmatrix}\vec{v}_1, \dots, \vec{v}_d\end{bmatrix} \in \mathbb{R}^{d\times d}, \qquad \Sigma_d = \begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_d\end{bmatrix} \in \mathbb{R}^{d\times d} \tag{11}$$

where $\sigma_1 > \sigma_2 > \cdots > \sigma_d > 0$. Let $k \leq d$ be a positive integer, let $P \in \mathbb{R}^{d\times k}$ be a matrix, and define $Z = XP$.

(d) (10 pts) Define the matrices $U_k$, $V_k$, and $\Sigma_k$ as

$$U_k = \begin{bmatrix}\vec{u}_1, \dots, \vec{u}_k\end{bmatrix} \in \mathbb{R}^{n\times k}, \qquad V_k = \begin{bmatrix}\vec{v}_1, \dots, \vec{v}_k\end{bmatrix} \in \mathbb{R}^{d\times k}, \qquad \Sigma_k = \begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_k\end{bmatrix} \in \mathbb{R}^{k\times k}. \tag{13}$$

Suppose that $P = V_k$, so that $Z = XV_k$. Let $\lambda \geq 0$, and let $\vec{\beta}^\star \in \mathbb{R}^k$ solve the ridge regression problem

$$\vec{\beta}^\star \doteq \operatorname*{argmin}_{\vec{\beta}\in\mathbb{R}^k} \left[\|Z\vec{\beta} - \vec{y}\|_2^2 + \lambda\|\vec{\beta}\|_2^2\right]. \tag{14}$$

**Show that:**

$$\vec{\beta}^\star = (\Sigma_k^2 + \lambda I_k)^{-1}\Sigma_k U_k^\top \vec{y}, \tag{15}$$

where $I_k \in \mathbb{R}^{k\times k}$ is the $k \times k$ identity matrix.

Let $X \in \mathbb{R}^{n\times d}$ have compact SVD given by $X = U_d \Sigma_d V_d^\top$ where

$$U_d = \begin{bmatrix}\vec{u}_1, \ldots, \vec{u}_d\end{bmatrix} \in \mathbb{R}^{n\times d}, \qquad V_d = \begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_d\end{bmatrix} \in \mathbb{R}^{d\times d}, \qquad \Sigma_d = \begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_d\end{bmatrix} \in \mathbb{R}^{d\times d} \tag{11}$$

where $\sigma_1 > \sigma_2 > \cdots > \sigma_d > 0$. Let $k \le d$ be a positive integer, and define the matrices

$$U_k = \begin{bmatrix}\vec{u}_1, \ldots, \vec{u}_k\end{bmatrix} \in \mathbb{R}^{n\times k}, \qquad V_k = \begin{bmatrix}\vec{v}_1, \ldots, \vec{v}_k\end{bmatrix} \in \mathbb{R}^{d\times k}, \qquad \Sigma_k = \begin{bmatrix}\sigma_1 & & \\ & \ddots & \\ & & \sigma_k\end{bmatrix} \in \mathbb{R}^{k\times k}. \tag{13}$$

Let $Z = XV_k$, and let $\lambda \ge 0$. In part (d), we computed the ridge regression estimate $\vec{\beta}^\star \in \mathbb{R}^k$ as

$$\vec{\beta}^\star = (\Sigma_k^2 + \lambda I_k)^{-1}\Sigma_k U_k^\top \vec{y}, \tag{15}$$

where $I_k \in \mathbb{R}^{k\times k}$ is the $k \times k$ identity matrix.

(e) (10 pts) Let $\vec{\alpha}^\star \in \mathbb{R}^d$ solve the original ridge regression problem, i.e.,

$$\vec{\alpha}^\star \doteq \operatorname*{argmin}_{\vec{\alpha}\in\mathbb{R}^d} \left[\|X\vec{\alpha} - \vec{y}\|_2^2 + \lambda\|\vec{\alpha}\|_2^2\right] = V_d(\Sigma_d^2 + \lambda I_d)^{-1}\Sigma_d U_d^\top \vec{y}, \tag{16}$$

where $I_d \in \mathbb{R}^{d\times d}$ is the $d \times d$ identity matrix. (You can assume without proof that the above equality is true.)

**Compute**

$$\|X\vec{\alpha}^\star - Z\vec{\beta}^\star\|_2^2, \tag{17}$$

**in terms of the vectors $(\vec{u}_i)_{i=1}^d$ and $\vec{y}$, and the scalars $(\sigma_i)_{i=1}^d$ and $\lambda$.** *Show your work and justify your answer(s).*

*This is an extra page for scratch work that will not be graded unless you tell us in the original problem space.*

*This is an 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.*

# Midterm Instructions

Read the following instructions before the exam.

**There are 11 problems of varying numbers of points. There are 110 points on the exam, but you will be scored out of 100 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” $\times$ 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.