

Midterm 2 Solutions

EECS 245, Spring 2026 at the University of Michigan

Name: _____

uniqname: _____

UMID: _____

Room: 1690 BBB (1-3PM) 1690 BBB (extended time)

Instructions

- This exam consists of 7 problems, worth 100 points, spread across 14 pages (7 sheets of paper).
- You have 120 minutes to complete this exam, unless you have extended-time accommodations through SSD.
- Write your uniqname in the top right corner of each page.
- For free response problems, **you must show all of your work**, and circle your final answer. We will not grade work that appears elsewhere, and you may lose points if your work is not shown.
- For multiple choice problems, completely fill in bubbles and square boxes; if we cannot tell which option(s) you selected, you may lose points.
 - A bubble means that you should only select one choice.
 - A square box means that you should select all that apply.
- You may refer to **two double-sided 8.5x11" handwritten notes sheets**. Other than that, you may not refer to any other resources or technology during the exam (no phones, watches, or calculators).

You are to abide by the University of Michigan/Engineering Honor Code. To receive a grade, please sign below to signify that you have kept the Honor Code pledge.

I have neither given nor received aid on this exam, nor have I concealed any violations of the Honor Code.

This page has been intentionally left blank.

Tip: Skim through the entire exam before starting to work on it.

Problem 1 (12 pts)

Suppose k is a real number. Let

$$A = \begin{bmatrix} 1 & k + 1 \\ 1 & 2k + 3 \end{bmatrix}$$

In each part, you are provided with information about A . **Your job is to find the value of k that satisfies the given condition.** Show your work in the space provided, and write your final answer in the bottom-right corner of the box. Your answers should be numbers with no variables.

a) (4 pts) $\det(A) = 14$.

$k =$ 12

Solution: Since A is a 2×2 matrix, its determinant is

$$\begin{aligned} \det(A) &= 1(2k + 3) - 1(k + 1) \\ &= k + 2 \end{aligned}$$

We're told that $\det(A) = 14$, so

$$\begin{aligned} k + 2 &= 14 \\ k &= 12 \end{aligned}$$

b) (4 pts) A is not invertible.

$k =$ -2

Solution: If A is not invertible, then $\det(A) = 0$. From part a),

$$\det(A) = k + 2$$

so

$$k + 2 = 0$$

$$k = -2$$

c) (4 pts) The bottom-right entry of A^{-1} is $1/4$.

$$k = \boxed{2}$$

Solution: The inverse of a 2×2 matrix is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad - bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

Here, $\det(A) = k + 2$, so

$$A^{-1} = \frac{1}{k + 2} \begin{bmatrix} 2k + 3 & -(k + 1) \\ -1 & 1 \end{bmatrix}$$

The bottom-right entry is $\frac{1}{k+2}$, and we're told that this equals $\frac{1}{4}$. So,

$$\frac{1}{k + 2} = \frac{1}{4}$$

$$k + 2 = 4$$

$$k = 2$$

Problem 2 (16 pts)

Suppose A is a 3×3 matrix whose null space is the plane

$$5x - y + 3z = 0$$

In other words, $\text{nullsp}(A) = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid 5x - y + 3z = 0 \right\}$.

a) (3 pts) Determine the following values. Give your answers as integers with no variables.

rank(A) = dim(nullsp(A)) =

Solution: The null space is a plane in \mathbb{R}^3 , which is a 2-dimensional subspace of \mathbb{R}^3 . So,

$$\dim(\text{nullsp}(A)) = 2$$

Since A has 3 columns, the rank-nullity theorem gives

$$\begin{aligned} \text{rank}(A) + \dim(\text{nullsp}(A)) &= 3 \\ \text{rank}(A) + 2 &= 3 \\ \text{rank}(A) &= 1 \end{aligned}$$

b) (4 pts) State one basis for nullsp(A). Your answer should be a list of vectors with no variables.

one basis for nullsp(A) =

Solution: The null space consists of all vectors satisfying

$$5x - y + 3z = 0$$

A basis for the null space, then, consists of any two linearly independent vectors that satisfy this equation. $\begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}$ satisfies it, since $5 \cdot 1 - (5) + 3 \cdot 0 = 0$, and similarly $\begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix}$ satisfies it. Therefore, one basis is

$$\left\{ \begin{bmatrix} 1 \\ 5 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 3 \\ 1 \end{bmatrix} \right\}$$

though there are infinitely many possible answers.

- c) (3 pts) State one basis for $\text{colsp}(A^T)$, the row space of A . Your answer should be a list of vectors with no variables.

one basis for $\text{colsp}(A^T) =$

$$\left\{ \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \right\}$$

Solution: A key fact to remember here is that the row space and null space of a matrix are orthogonal complements, as discussed in [Chapter 5.4](#). This means that every element in the row space must be orthogonal to every element in the null space.

We're given that the null space consists of all vectors $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ such that $5x - y + 3z = 0$.

Equivalently, this means that $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \cdot \vec{x} = 0$. So, this means that every element in the

null space is orthogonal to $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$, so $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$ must be in the row space. The row space is

1-dimensional, since $\text{rank}(A) = 1$. So, one basis is $\left\{ \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \right\}$.

If you didn't remember this key fact, no problem — you could have arrived at this conclusion from scratch on the exam. We know A is 3×3 , so suppose it looks like

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$$

If $\vec{x} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$ is in the null space, it must mean that $A\vec{x} = \vec{0}$:

$$A\vec{x} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} a_{11}x + a_{12}y + a_{13}z \\ a_{21}x + a_{22}y + a_{23}z \\ a_{31}x + a_{32}y + a_{33}z \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

From here, you see that the dot product of each row of A with \vec{x} must be 0, so each row of A must be orthogonal to \vec{x} . The plane form of the null space, $5x - y + 3z = 0$, tells you

that $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$ is orthogonal to every vector in the null space, so putting these facts together

gives us that $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$ is in the row space. Together with the fact that the row space is

1-dimensional, since $\text{rank}(A) = 1$, we have that a basis is $\left\{ \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \right\}$.

All other possible answers involve (non-zero) scalar multiples of $\begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix}$.

Recall, A is a 3×3 matrix whose null space is the plane $5x - y + 3z = 0$.

d) (6 pts) Suppose that

$$A \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 15 \\ 30 \\ 0 \end{bmatrix}$$

Find A . Show your work, and circle your final answer, which should be a matrix with no variables.

Solution: Since $\text{rank}(A) = 1$ and the row space is

$$\text{span} \left(\left\{ \begin{bmatrix} 5 \\ -1 \\ 3 \end{bmatrix} \right\} \right),$$

each row of A must be a scalar multiple of $[5 \ -1 \ 3]$. So, for some constants a , b , and c ,

$$A = \begin{bmatrix} 5a & -a & 3a \\ 5b & -b & 3b \\ 5c & -c & 3c \end{bmatrix}$$

We're told that

$$A \begin{bmatrix} 3 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 15 \\ 30 \\ 0 \end{bmatrix}$$

The left-hand side is 3 times the first column of A , so the first column of A is

$$\begin{bmatrix} 5 \\ 10 \\ 0 \end{bmatrix}$$

This gives

$$5a = 5, \quad 5b = 10, \quad 5c = 0$$

so $a = 1$, $b = 2$, and $c = 0$. Therefore,

$$A = \begin{bmatrix} 5 & -1 & 3 \\ 10 & -2 & 6 \\ 0 & 0 & 0 \end{bmatrix}$$

Problem 3 (12 pts)

Suppose A is an $n \times n$ matrix.

For each statement below, determine whether it is true or false. If true, prove that it is true. If false, give a counterexample or a short explanation.

- a) (4 pts) If A is symmetric, then A^2 must be symmetric.

True False

Solution: This is true. Since A is symmetric, $A^T = A$. So,

$$(A^2)^T = (AA)^T = A^T A^T = AA = A^2$$

Therefore, A^2 is symmetric.

- b) (4 pts) If A^2 is symmetric, then A must be symmetric.

True False

Solution: This is false. For example, let

$$A = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

This matrix is not symmetric, but

$$A^2 = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

which is symmetric.

- c) (4 pts) If $\vec{x} \in \text{nullsp}(A^T)$ and $\vec{y} \in \text{colsp}(A)$, then \vec{x} and \vec{y} are orthogonal.

True False

Solution: This is true. If $\vec{y} \in \text{colsp}(A)$, then $\vec{y} = A\vec{v}$ for some vector \vec{v} . Since $\vec{x} \in \text{nullsp}(A^T)$, we know $A^T\vec{x} = \vec{0}$. So,

$$\vec{x} \cdot \vec{y} = \vec{x}^T A\vec{v} = (A^T\vec{x})^T \vec{v} = \vec{0}^T \vec{v} = 0$$

Therefore, \vec{x} and \vec{y} are orthogonal.

Problem 4 (14 pts)

Suppose X is a matrix such that

$$X^T X = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix} \quad X X^T = \begin{bmatrix} 1 & \sqrt{3} & 0 & 0 \\ \sqrt{3} & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

a) (3 pts) Fill in each blank with an integer with no variables.

X has rows, columns, and $\text{rank}(X) =$.

Solution: Recall that if X is an $n \times d$ matrix, then $X^T X$ is an $d \times d$ matrix containing the dot products of all pairs of X 's columns, and $X X^T$ is an $n \times n$ matrix containing the dot products of all pairs of X 's rows.

Here, since $X^T X$ is 2×2 , X must have 2 columns and since $X X^T$ is 4×4 , X must have 4 rows. So X is 4×2 .

Also, recall that $\text{rank}(X) = \text{rank}(X^T X) = \text{rank}(X X^T)$, as proven [here](#). Since $\text{rank}(X^T X) = 2$ (as it is a diagonal matrix with 2 non-zero entries), we have that $\text{rank}(X) = 2$.

b) (4 pts) For each statement below, determine whether it is true or false.

(i) The columns of X are all orthogonal to each other.

True False

(ii) The columns of X are orthonormal.

True False

Solution:

(i) This is true. The entries of $X^T X$ are the dot products of the columns of X with each other. Since the off-diagonal entries are 0, the columns of X are orthogonal to each other.

(ii) This is false. The diagonal entries of $X^T X$ are the squared lengths of the columns of X . Since both diagonal entries are 4, both columns have length 2, not 1.

c) (7 pts) Suppose P is the matrix that projects onto the column space of X . In other words, for any \vec{y} of the appropriate shape, $P\vec{y}$ is the projection of \vec{y} onto $\text{colsp}(X)$. **Find** P . Show your work, and your final answer, which should be a matrix with no variables.

Solution: Since X has linearly independent columns, the projection matrix onto $\text{colsp}(X)$ is

$$P = X(X^T X)^{-1} X^T$$

Here,

$$X^T X = 4I \quad \implies \quad (X^T X)^{-1} = \frac{1}{4}I$$

So,

$$P = X \left(\frac{1}{4}I \right) X^T = \frac{1}{4} X X^T$$

Using the given value of $X X^T$,

$$P = \begin{bmatrix} 1/4 & \sqrt{3}/4 & 0 & 0 \\ \sqrt{3}/4 & 3/4 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Note that we're only able to answer this problem because $X^T X$ is a multiple of the identity matrix, so its inverse is just a multiple of the identity matrix. If $X^T X$ was not a multiple of the identity matrix, even if it was diagonal, we wouldn't be able to find P using just the information in this problem.

Problem 5 (19 pts)

Suppose we're given a dataset with $n = 5$ rows, and we use it to fit a multiple linear regression model with two features and an intercept term.

$$h(\vec{x}_i) = w_0 + w_1x_i^{(1)} + w_2x_i^{(2)}$$

Let X be the corresponding 5×3 design matrix and $\vec{y} \in \mathbb{R}^5$ be the corresponding observation vector. Suppose the matrix P that projects onto the column space of X is

$$P = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

a) (4 pts) **In parts a) and b) only**, suppose the projection of \vec{y} onto $\text{colsp}(X)$ is $\vec{p} = \begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$. There

are infinitely many such vectors \vec{y} . State one possible vector \vec{y} **whose five components are all different**. Give your answer as a vector with no variables.

one possible vector $\vec{y} =$

$$\begin{bmatrix} 0 \\ 1 \\ 5 \\ 6 \\ 3 \end{bmatrix}$$

Solution: For any vector \vec{y} , multiplying by P averages the first four components of \vec{y} and leaves the fifth component unchanged:

$$P\vec{y} = \begin{bmatrix} \frac{y_1 + y_2 + y_3 + y_4}{4} \\ \frac{y_1 + y_2 + y_3 + y_4}{4} \\ \frac{y_1 + y_2 + y_3 + y_4}{4} \\ \frac{y_1 + y_2 + y_3 + y_4}{4} \\ y_5 \end{bmatrix}$$

We want this to equal $\begin{bmatrix} 3 \\ 3 \\ 3 \\ 3 \\ 3 \end{bmatrix}$, so the first four components of \vec{y} need to have average 3, and the fifth component needs to be 3.

One possible choice is

$$\vec{y} = \begin{bmatrix} 0 \\ 1 \\ 5 \\ 6 \\ 3 \end{bmatrix}$$

The first four components have average 3, and all five components are different. There are infinitely many possible answers, though.

b) (3 pts) Let \vec{y} and \vec{p} be as defined in part (a). True or false: $X^T(\vec{p} - \vec{y}) = \vec{0}$.

True False

Solution: This is true. If \vec{p} is the projection of \vec{y} onto $\text{colsp}(X)$, then the error vector $\vec{y} - \vec{p}$ is orthogonal to every vector in $\text{colsp}(X)$. This is how we arrived at the normal equations, $X^T X \vec{w} = X^T \vec{y}$. Here, this means

$$X^T(\vec{y} - \vec{p}) = \vec{0}$$

Multiplying by -1 gives

$$X^T(\vec{p} - \vec{y}) = \vec{0}$$

For the rest of the problem, suppose that both $\vec{w}^* = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ and $\vec{w}' = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ are both optimal parameter vectors that minimize mean squared error.

c) (4 pts) Which of these vectors are in $\text{nullsp}(X)$? **Select all** that apply.

- $\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$
 $\begin{bmatrix} 5 \\ 4 \\ 1 \end{bmatrix}$
 $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$
 $\begin{bmatrix} 4 \\ 6 \\ 2 \end{bmatrix}$
 $\begin{bmatrix} -2 \\ 4 \\ 2 \end{bmatrix}$

Solution: If two parameter vectors are both solutions to the normal equation, their difference is in $\text{nullsp}(X)$. So,

$$\vec{w}' - \vec{w}^* = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \in \text{nullsp}(X)$$

Where did this come from? The fact that \vec{w}' and \vec{w}^* are both optimal parameter vectors means that they both result in the same projection of \vec{y} onto $\text{colsp}(X)$, so

$$X\vec{w}^* = X\vec{w}'$$

But, this means $X(\vec{w}' - \vec{w}^*) = \vec{0}$, which says that $\vec{w}' - \vec{w}^*$ is in $\text{nullsp}(X)$.

Also, P projects onto $\text{colsp}(X)$, and P has rank 2. Therefore $\text{rank}(X) = 2$ (the logic behind this is described [here](#)). Since X has 3 columns, the rank-nullity theorem gives

$$\dim(\text{nullsp}(X)) = 3 - 2 = 1$$

So $\text{nullsp}(X)$ is exactly

$$\text{nullsp}(X) = \text{span} \left(\left\{ \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \right\} \right)$$

Among the listed choices, the vectors in this span are $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$ and $\begin{bmatrix} -2 \\ 4 \\ 2 \end{bmatrix}$.

The rank-nullity logic wasn't strictly necessary to answer the question; I've included it here for completeness, as it fully justifies why none of the other listed vectors are in $\text{nullsp}(X)$.

The information stated below, above part d), is the same as the information stated on the previous page. It's provided for your convenience.

Recall, X is a 5×3 design matrix for the model

$$h(\vec{x}_i) = w_0 + w_1 x_i^{(1)} + w_2 x_i^{(2)}$$

Additionally, $\vec{y} \in \mathbb{R}^5$ is an observation vector, both $\vec{w}^* = \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$ and $\vec{w}' = \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix}$ are both optimal parameter vectors that minimize mean squared error, and the matrix P that projects onto the column space of X is

$$P = \begin{bmatrix} 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 1/4 & 1/4 & 1/4 & 1/4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

- d) (8 pts) Find one possible design matrix X , consistent with all of the information above. Show your work, and circle your final answer, which should be a matrix with no variables.

Solution: Since P projects onto $\text{colsp}(X)$, we need $\text{colsp}(X) = \text{colsp}(P)$. Notice that the result $P\vec{y}$ for any vector $\vec{y} \in \mathbb{R}^5$ will have equal first four components (resulting from averaging the original first four components of \vec{y}) and the fifth component will be unchanged. If we think of the space of possible values of $P\vec{y}$, we realize that any $P\vec{y}$ is of the form

$$\begin{bmatrix} a \\ a \\ a \\ a \\ b \end{bmatrix} = a \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} + b \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

This means

$$\text{colsp}(X) = \text{span} \left(\left(\begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix} \right) \right)$$

Now, the problem boils down to finding a design matrix X with the above column space, that also meets the other requirements. Here are the other relevant requirements:

1. Since the model has an intercept term, the first column of X should be $\vec{1} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$.

2. From part c), we need $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} \in \text{nullsp}(X)$.

If the columns of X are $\vec{x}^{(0)}$, $\vec{x}^{(1)}$, and $\vec{x}^{(2)}$ (we're told X has 3 columns), the first requirement states

$$\vec{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

The second requirement states

$$\underbrace{\begin{bmatrix} | & | & | \\ \vec{x}^{(0)} & \vec{x}^{(1)} & \vec{x}^{(2)} \\ | & | & | \end{bmatrix}}_X \begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix} = \vec{0}$$

or, in other words, $\vec{x}^{(0)} - 2\vec{x}^{(1)} - \vec{x}^{(2)} = \vec{0}$.

To guarantee $\text{colsp}(X)$ is the span we set out before,

$$\text{colsp}(X) = \text{span} \left(\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\} \right)$$

let's just pick $\vec{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix}$. Since $\vec{x}^{(0)} - \vec{x}^{(1)} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$, we have accomplished the goal of

finding a design matrix X with the desired column space. With our choices of $\vec{x}^{(0)}$ and $\vec{x}^{(1)}$ out of the way, $\vec{x}^{(2)}$ is fully determined for us:

$$\vec{x}^{(0)} - 2\vec{x}^{(1)} - \vec{x}^{(2)} = \vec{0} \implies \vec{x}^{(2)} = \vec{x}^{(0)} - 2\vec{x}^{(1)} = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - 2 \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

Therefore, one possible design matrix is

$$X = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 1 & -1 \\ 1 & 0 & 1 \end{bmatrix}$$

This design matrix has a column space of span $\left(\left\{ \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\} \right)$, which is the same as the

column space of P . It also has the required null space, which is why it would be wrong to just pick, say,

$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

— the above matrix has a null space spanned by $\begin{bmatrix} 1 \\ -1 \\ -1 \end{bmatrix}$, not $\begin{bmatrix} 1 \\ -2 \\ -1 \end{bmatrix}$.

Problem 6 (12 pts)

Suppose A is an $n \times d$ matrix and $\vec{x} \in \mathbb{R}^d$. Consider the function $f : \mathbb{R}^d \rightarrow \mathbb{R}$ given by

$$f(\vec{x}) = \|A\vec{x}\|$$

a) (2 pts) True or False: $f(\vec{x})$ is a linear transformation.

True False

Solution: This is false. Recall, a linear transformation must satisfy $f(c\vec{x}) = cf(\vec{x})$ for any scalar c . But, suppose we pick $n = d = 1$, and let $A = [1]$ (here we're thinking of a 1×1 matrix as a scalar). Then, $f(x)$ is just the absolute value of the scalar x .

$$f(x) = |x|$$

But, $f(-2) = 2$ is not the same as $-2f(1) = -2$. So, this $f(x)$ is not a linear transformation, and thus in general $f(\vec{x}) = \|A\vec{x}\|$ is not a linear transformation.

Another way to think about why $f(\vec{x})$ is not linear is to use the fact that $\|A\vec{x}\|^2 = \vec{x}^T A^T A \vec{x}$:

$$f(\vec{x}) = \sqrt{\vec{x}^T A^T A \vec{x}}$$

$f(\vec{x})$ is the square root of a quadratic form, which is not linear.

b) (10 pts) Find $\nabla f(\vec{x})$. Assume that $A\vec{x} \neq \vec{0}$. Show your work, and write your final answer in the bottom-right corner of the box. Your answer should be an expression in terms of A , \vec{x} , and/or constants. *Hint: Start by taking the gradient of $\|A\vec{x}\|^2$, then apply the chain rule.*

$$\nabla f(\vec{x}) = \frac{A^T A \vec{x}}{\|A \vec{x}\|}$$

Solution: As the hint suggests, let's start by writing

$$\|A\vec{x}\|^2 = (A\vec{x})^T (A\vec{x}) = \vec{x}^T A^T A \vec{x}$$

Using the quadratic-form gradient rule,

$$\nabla \|A\vec{x}\|^2 = 2A^T A \vec{x}$$

Now,

$$f(\vec{x}) = \|A\vec{x}\| = \sqrt{\|A\vec{x}\|^2}$$

The chain rule from [Chapter 8.2](#) states that if $f(\vec{x}) = h(g(\vec{x}))$, where $h : \mathbb{R} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^d \rightarrow \mathbb{R}$ are both differentiable, then $\nabla f(\vec{x}) = h'(g(\vec{x})) \nabla g(\vec{x})$.

Here, $h(x) = \sqrt{x}$ (so $h'(x) = \frac{1}{2\sqrt{x}}$) and $g(\vec{x}) = \|A\vec{x}\|^2$, so

$$\nabla f(\vec{x}) = \frac{1}{2\sqrt{\|A\vec{x}\|^2}} (2A^T A \vec{x}) = \frac{A^T A \vec{x}}{\|A\vec{x}\|}$$

Problem 7 (15 pts)

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ given by

$$f(\vec{x}) = cx_1^2 + dx_2^2$$

where c and d are constants. We'd like to use gradient descent to minimize $f(\vec{x})$. For some values of c and d , and some initial guess $\vec{x}^{(0)}$ and learning rate/step size α , we find that

$$\vec{x}^{(1)} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}, \quad \nabla f(\vec{x}^{(1)}) = \begin{bmatrix} 6 \\ -2 \end{bmatrix}, \quad \vec{x}^{(2)} = \begin{bmatrix} 2.8 \\ 1.4 \end{bmatrix}$$

- a) (5 pts) Find the value of α . Show your work, and write your final answer in the bottom-right corner of the box. Your answer should be a number with no variables.

$\alpha =$ 1/5

Solution: Gradient descent uses the update

$$\vec{x}^{(2)} = \vec{x}^{(1)} - \alpha \nabla f(\vec{x}^{(1)})$$

Substituting the given values,

$$\begin{bmatrix} 2.8 \\ 1.4 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} - \alpha \begin{bmatrix} 6 \\ -2 \end{bmatrix} = \begin{bmatrix} 4 - 6\alpha \\ 1 + 2\alpha \end{bmatrix}$$

Using either component,

$$4 - 6\alpha = 2.8$$

$$6\alpha = 1.2$$

$$\alpha = \frac{1}{5}$$

- b) (5 pts) Find the value of d (**not** c). Show your work, and write your final answer in the bottom-right corner of the boxes. Your answer should be a number with no variables.

$$d = \boxed{-1}$$

Solution: The gradient of

$$f(\vec{x}) = cx_1^2 + dx_2^2$$

is

$$\nabla f(\vec{x}) = \begin{bmatrix} 2cx_1 \\ 2dx_2 \end{bmatrix}$$

At $\vec{x}^{(1)} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}$, we're told that

$$\nabla f(\vec{x}^{(1)}) = \begin{bmatrix} 6 \\ -2 \end{bmatrix}$$

Using the second component (because we're only asked for d),

$$2d(1) = -2$$

$$d = -1$$

- c) (5 pts) Your friend claims that gradient descent always converges to a minimum because each iteration moves in the direction of steepest decrease. Based on the information in this problem, is your friend correct? State “yes” or “no”, and briefly explain your reasoning.

Solution: No. From part **b**), $d = -1$, so

$$f(\vec{x}) = cx_1^2 - x_2^2$$

This function does not have a minimum, because we can make $f(\vec{x})$ arbitrarily negative by making $|x_2|$ arbitrarily large. So, in this problem, gradient descent cannot converge to a minimum.

This page has been intentionally left blank. Feel free to use it for scratch work.

username: _____

Congrats on finishing Midterm 2!

Feel free to draw us a picture about EECS 245 in the box below.