Midterm 2 Solutions

EECS 245, Winter 2026 at the University of Michigan

Name: _____

uniqname: _____

UMID: _____

Room: ◯ 1670 BBB (big room)   ◯ 1690 BBB   ◯ Other

Instructions

- This exam consists of 8 problems, worth a total of 100 points, spread across 12 pages (6 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 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

$$A = \begin{bmatrix} 3 & 2 \\ k & 4 \end{bmatrix}$$

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) $\mathrm{rank}(A) = 1$

$$k = \boxed{6}$$

Solution: If $\mathrm{rank}(A) = 1$, then A is not invertible, which means $\det(A) = 0$. So,

$$\det(A) = (3)(4) - (2)(k) = 12 - 2k = 0$$
$$2k = 12$$
$$k = 6$$

b) (4 pts) $\det(A) = 2$

$$k = \boxed{5}$$

c) (4 pts) $A^{-1} = \begin{bmatrix} 1 & -1/2 \\ -1 & 3/4 \end{bmatrix}$

$k = $ | 4

Solution: The inverse of a 2×2 matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$ is $\frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$. So,

$$A^{-1} = \frac{1}{12 - 2k} \begin{bmatrix} 4 & -2 \\ -k & 3 \end{bmatrix}$$

Since we're told that

$$A^{-1} = \begin{bmatrix} 1 & -1/2 \\ -1 & 3/4 \end{bmatrix},$$

we can match entries. For example, using the bottom-right entry,

$$\frac{3}{12 - 2k} = \frac{3}{4}$$
$$12 - 2k = 4$$
$$2k = 8$$
$$k = 4$$

Problem 2 (10 pts)

Suppose A is a matrix such that $A \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$ and $\left\{ \begin{bmatrix} 1 \\ 4 \\ -2 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \right\}$ is a basis for nullsp(A).

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

Solution: Let the columns of A be $\vec{c}_1, \vec{c}_2, \vec{c}_3, \vec{c}_4$. Since

$$A \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = 1\vec{c}_1 + 0\vec{c}_2 + 0\vec{c}_3 + 0\vec{c}_4 = \begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix},$$

we know that the first column of A must be $\begin{bmatrix} 6 \\ 0 \\ 1 \end{bmatrix}$.

Now, let's use the information given about the null space to find the other columns of A. Since

$$\begin{bmatrix} 0 \\ 1 \\ 1 \\ 0 \end{bmatrix} \in \text{nullsp}(A),$$

we have

$$\vec{c}_2 + \vec{c}_3 = \vec{0} \implies \vec{c}_3 = -\vec{c}_2$$

Also, since

$$\begin{bmatrix} 1 \\ 4 \\ -2 \\ 0 \end{bmatrix} \in \text{nullsp}(A),$$

we have

$$\vec{c}_1 + 4\vec{c}_2 - 2\vec{c}_3 = \vec{0}$$

Substituting $\vec{c}_3 = -\vec{c}_2$ gives

$$\vec{c}_1 + 6\vec{c}_2 = \vec{0}$$

$$\vec{c}_2 = -\frac{1}{6}\vec{c}_1 = \begin{bmatrix} -1 \\ 0 \\ -1/6 \end{bmatrix}.$$

So,

$$\vec{c}_3 = -\vec{c}_2 = \begin{bmatrix} 1 \\ 0 \\ 1/6 \end{bmatrix}$$

Finally, $\dim(\text{nullsp}(A)) = 2$ and A has 4 columns, so by rank-nullity,

$$\text{rank}(A) = 4 - 2 = 2$$

So we should choose $\vec{c}_4$ to be linearly independent from $\vec{c}_1$. One easy choice is

$$\vec{c}_4 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

This gives one possible matrix.

$$A = \begin{bmatrix} 6 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & -1/6 & 1/6 & 0 \end{bmatrix}$$

Problem 3 (11 pts)

Suppose A and B are both **non-zero** 6×6 matrices, such that $\text{rank}(A) = 4$ and that every column of B is in $\text{nullsp}(A)$.

a) (3 pts) Fill in the blanks: The third __(i)__ of A is __(ii)__ to the fourth __(iii)__ of B.

(i) ○ column ● row

(ii) ● orthogonal ○ parallel

(iii) ● column ○ row

> **Solution:** Every column of B is in $\text{nullsp}(A)$. From **Chapter 5.4 in the notes**, the row space of A and the null space of A are orthogonal complements. That means every row of A is orthogonal to every vector in $\text{nullsp}(A)$, and hence orthogonal to every column of B.
>
> So, the third **row** of A is **orthogonal** to the fourth **column** of B.

b) (4 pts) **Select all** possible values of $\text{rank}(AB)$.
☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6

> **Solution:** Let $\vec{b}_1, \vec{b}_2, \ldots, \vec{b}_6$ be the columns of B. Since every column of B is in $\text{nullsp}(A)$, we have
> $$A\vec{b}_j = \vec{0}$$
> for every j. But the jth column of AB is exactly $A\vec{b}_j$, so every column of AB is $\vec{0}$. Therefore,
> $$AB = 0_{6\times 6} \implies \text{rank}(AB) = 0_{6\times 6}$$

c) (4 pts) **Select all** possible values of $\text{rank}(B)$.
☐ 0 ☐ 1 ☐ 2 ☐ 3 ☐ 4 ☐ 5 ☐ 6

Solution: Since A is 6×6 with rank 4, the rank-nullity theorem gives

$$\dim(\text{nullsp}(A)) = 6 - 4 = 2$$

Every column of B lies in $\text{nullsp}(A)$, so

$$\text{colsp}(B) \subseteq \text{nullsp}(A)$$

Therefore,

$$\text{rank}(B) = \dim(\text{colsp}(B)) \leq 2$$

Also, B is non-zero, so $\text{rank}(B) \neq 0$.
So the only possible values are **1** and **2**.

Both are achievable: all columns of B could be multiples of one non-zero vector in $\text{nullsp}(A)$, or they could span all of $\text{nullsp}(A)$ (which is a 2-dimensional subspace of $\mathbb{R}^6$, since $\text{rank}(A) = 4$).

Problem 4 (13 pts)

Suppose X is some $3 \times d$ matrix, for some integer d. Let

$$\vec{y} = \begin{bmatrix} 9 \\ -5 \\ 3 \end{bmatrix}$$

a) (5 pts) Which of the following **could** be the projection of $\vec{y}$ onto colsp(X)?
Select an answer, then briefly justify your answer in the space provided using properties of projections. Correct answers without justification may not receive full credit.

○ $\begin{bmatrix} 5 \\ -7 \\ 4 \end{bmatrix}$　○ $\begin{bmatrix} 7 \\ -7 \\ 4 \end{bmatrix}$　● $\begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix}$　○ $\begin{bmatrix} 6 \\ -7 \\ 3 \end{bmatrix}$

> **Solution:** If $\vec{p}$ is the projection of $\vec{y}$ onto colsp(X), then the error
>
> $$\vec{y} - \vec{p}$$
>
> must be orthogonal to all vectors in colsp(X), and hence orthogonal to $\vec{p}$ itself.
> For the third option, $\vec{p} = \begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix}$, we have
>
> $$\vec{p} = \begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix} \implies \vec{y} - \vec{p} = \begin{bmatrix} 9 \\ -5 \\ 3 \end{bmatrix} - \begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix}$$
>
> The dot product of $\vec{p}$ and $\vec{y} - \vec{p}$ is
>
> $$\vec{p} \cdot (\vec{y} - \vec{p}) = \begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 2 \\ -1 \end{bmatrix} = 18 - 14 - 4 = 0$$
>
> So $\vec{p} = \begin{bmatrix} 6 \\ -7 \\ 4 \end{bmatrix}$ could be the projection of $\vec{y}$ onto colsp(X). If you repeat this calculation for the other three options, you'll find that $\vec{p}$ and $\vec{y} - \vec{p}$ are not orthogonal.

In each of the remaining parts, identify whether the statement is True or False and justify your answer in the space provided. Correct answers without justification may not receive full credit.

b) (4 pts) If the projection of $\vec{y}$ onto colsp(X) is $\vec{y}$ itself, then rank(X) must be 3.
○ True　● False

> **Solution:** This is false. If the projection of $\vec{y}$ onto colsp(X) is $\vec{y}$ itself, that only tells us that $\vec{y} \in$ colsp(X).
>
> But colsp(X) could still be a 1-dimensional or 2-dimensional subspace of $\mathbb{R}^3$ that happens to contain $\vec{y}$. For example, if colsp(X) = span ($\{\vec{y}\}$), then the projection of $\vec{y}$ is still $\vec{y}$, but rank(X) = 1, not 3.

c) (4 pts) If rank(X) = 3, then the projection of $\vec{y}$ onto colsp(X) must be $\vec{y}$ itself.

● True   ◯ False

> **Solution:** This is true. If rank(X) = 3 and X is a $3 \times d$ matrix, then colsp(X) is a 3-dimensional subspace of $\mathbb{R}^3$. The only 3-dimensional subspace of $\mathbb{R}^3$ is all of $\mathbb{R}^3$.
> But, this means every vector in $\mathbb{R}^3$, including $\vec{y}$, is in colsp(X). Therefore, the projection of $\vec{y}$ onto colsp(X) is just $\vec{y}$ itself.

Problem 5 (13 pts)

Let

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 2 & 1 & 0 & 0 & 4 \\ 3 & 1 & 0 & -7 & 4 \end{bmatrix}$$

a) (5 pts) Recall, a CR decomposition of an $n \times d$ matrix A is a product $A = CR$, where C is an $n \times r$ matrix with linearly independent columns and R is an $r \times d$ matrix with linearly independent rows, and $r = \text{rank}(A)$.

Provide a CR decomposition of A. Your answers should be matrices with no variables.

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & -7 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

Solution: The columns of A are

$$\vec{c}_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \quad \vec{c}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{c}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}, \quad \vec{c}_4 = \begin{bmatrix} 0 \\ 0 \\ -7 \end{bmatrix}, \quad \vec{c}_5 = \begin{bmatrix} 0 \\ 4 \\ 4 \end{bmatrix}$$

Reading left-to-right, columns 1, 2, and 4 are linearly independent, so we place them in C:

$$C = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & -7 \end{bmatrix}$$

Now we need to express each column of A as a linear combination of the columns of C:

$$\vec{c}_1 = 1\vec{c}_1 + 0\vec{c}_2 + 0\vec{c}_4, \quad \vec{c}_2 = 0\vec{c}_1 + 1\vec{c}_2 + 0\vec{c}_4, \quad \vec{c}_3 = \vec{0},$$

$$\vec{c}_4 = 0\vec{c}_1 + 0\vec{c}_2 + 1\vec{c}_4, \quad \vec{c}_5 = 0\vec{c}_1 + 4\vec{c}_2 + 0\vec{c}_4$$

The coefficients in each linear combination are the entries in the corresponding column of R. So,

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

Thus, one CR decomposition is

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 3 & 1 & -7 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix}$$

b) (3 pts) Find $\dim(\text{nullsp}(A^T))$. Give your answer as an integer with no variables.

$$\dim(\text{nullsp}(A^T)) = \boxed{0}$$

> **Solution:** The matrix A has 3 rows and rank 3. Applying rank-nullity to A^T, we get
>
> $$\text{rank}(A^T) + \dim(\text{nullsp}(A^T)) = \text{number of columns of } A^T = 3$$
>
> Since $\text{rank}(A^T) = \text{rank}(A) = 3$,
>
> $$3 + \dim(\text{nullsp}(A^T)) = 3$$
> $$\dim(\text{nullsp}(A^T)) = 0$$
>
> This means that A^T's null space is $\left\{\vec{0}\right\}$.

c) (5 pts) Suppose we apply the Gram-Schmidt process to the **rows** of A, and place the resulting orthonormal vectors into the **rows** of a new matrix, Q.

Let P be the matrix that projects vectors in $\mathbb{R}^5$ onto $\text{colsp}(Q^T)$ (the row space of Q). In other words, if $\vec{y} \in \mathbb{R}^5$, then $P\vec{y}$ is the projection of $\vec{y}$ onto $\text{colsp}(Q^T)$.

Find an expression for P in terms of Q and Q^T. Show your work, and ⬚ circle your final answer, which should be an expression in terms of Q and Q^T. Answers that aren't fully simplified will not be given credit.

$$P = \boxed{Q^TQ}$$

Solution: In general, the projection matrix onto colsp(X), where X is any matrix with linearly independent columns, is

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

Here, we want to project onto colsp(Q^T), so we should use $X = Q^T$:

$$P = Q^T((Q^T)^T Q^T)^{-1}(Q^T)^T = Q^T(QQ^T)^{-1}Q$$

But, since Q's rows are orthonormal, $QQ^T = I$. This is because QQ^T is a matrix containing the dot products of the rows of Q with each other (the same way $Q^T Q$ is a matrix containing the dot products of the columns of Q with each other). Since the rows of Q are orthonormal, the dot products are all 0 except for the diagonal, which is 1.
So,

$$P = Q^T I Q = Q^T Q$$

Problem 6 (20 pts)

Suppose we'd like to fit a multiple linear regression model **without** an intercept term to **predict the number of fans in attendance at a Michigan football home game** given various features.

For each row in the dataset, the corresponding feature vector is $\vec{x}_i = \begin{bmatrix} \text{tempF}_i \\ \text{tempC}_i \\ \text{night}_i \\ \text{day}_i \end{bmatrix}$, where:

- tempF_i is the temperature, in degrees **Fahrenheit**, at kickoff for game i
- tempC_i is the temperature, in degrees **Celsius**, at kickoff for game i
- night_i is 1 if game i is a night game and 0 otherwise
- day_i is 0 if game i is a night game and 1 otherwise

Important: Note that
$$\text{tempC}_i = \frac{5}{9}(\text{tempF}_i - 32)$$

So, our model is of the form
$$h(\vec{x}_i) = w_1 \cdot \text{tempF}_i + w_2 \cdot \text{tempC}_i + w_3 \cdot \text{night}_i + w_4 \cdot \text{day}_i$$

We find optimal model parameters, $\vec{w}^* = \begin{bmatrix} w_1^* \\ w_2^* \\ w_3^* \\ w_4^* \end{bmatrix}$, by solving the normal equation.

a) (4 pts) The first two rows of the dataset have the following information:

- Game 1: 77 degrees Fahrenheit, 25 degrees Celsius, not night game, 102,111 fans
- Game 2: 59 degrees Fahrenheit, 15 degrees Celsius, night game, 101,982 fans

Write the first two rows of the design matrix, X. Your answer should be a matrix with two rows and no variables.

$$X = \begin{bmatrix} 77 & 25 & 0 & 1 \\ 59 & 15 & 1 & 0 \end{bmatrix}$$

Solution: Each row of the design matrix is just the feature vector for that game:

$$\vec{x}_i = \begin{bmatrix} \text{tempF}_i \\ \text{tempC}_i \\ \text{night}_i \\ \text{day}_i \end{bmatrix}$$

So,

$$\text{Game 1}: \begin{bmatrix} 77 & 25 & 0 & 1 \end{bmatrix} \qquad \text{Game 2}: \begin{bmatrix} 59 & 15 & 1 & 0 \end{bmatrix}$$

Therefore, the first two rows of X are

$$X = \begin{bmatrix} 77 & 25 & 0 & 1 \\ 59 & 15 & 1 & 0 \end{bmatrix}$$

Recall, our model is of the form

$$h(\vec{x}_i) = w_1 \cdot \text{tempF}_i + w_2 \cdot \text{tempC}_i + w_3 \cdot \text{night}_i + w_4 \cdot \text{day}_i$$

where $\text{tempC}_i = \frac{5}{9}(\text{tempF}_i - 32)$.

b) (4 pts) Suppose $\vec{w}'$ is one solution to the normal equation for this model. Which option describes the **complete set** of solutions to the normal equation?

○ $\left\{ \vec{w}' + t \begin{bmatrix} 1 \\ -9/5 \\ 32 \\ -32 \end{bmatrix} , t \in \mathbb{R} \right\}$ ● $\left\{ \vec{w}' + t \begin{bmatrix} 1 \\ -9/5 \\ -32 \\ -32 \end{bmatrix} , t \in \mathbb{R} \right\}$

○ $\left\{ \vec{w}' + t \begin{bmatrix} 1 \\ 9/5 \\ 32 \\ -32 \end{bmatrix} , t \in \mathbb{R} \right\}$ ○ $\left\{ \vec{w}' + t \begin{bmatrix} 1 \\ 9/5 \\ 32 \\ -32 \end{bmatrix} , t \in \mathbb{R} \right\}$

Solution: If we add any vector in $\text{nullsp}(X)$ to one solution of the normal equation, we get another solution. So we just need to find a non-zero vector in $\text{nullsp}(X)$.

For any row of X, the following two statements must be true:

$$\text{tempC}_i = \frac{5}{9}(\text{tempF}_i - 32)$$
$$\text{night}_i + \text{day}_i = 1$$

The idea is to use this information to find a linear combination of X's columns that equals the zero vector.

Using the first equation, we have

$$\text{tempF}_i - \frac{9}{5}\text{tempC}_i - 32 = 0$$

In order to write this as a linear combination of X's columns, the 32 needs to come from a vector that is "constant" across all rows. Fortunately, that's true of the sum of the night and day columns, since $\text{night}_i + \text{day}_i = 1$ for all rows. So, this means

$$\text{tempF}_i - \frac{9}{5}\text{tempC}_i - 32(\text{night}_i + \text{day}_i) = 0$$
$$\text{tempF}_i - \frac{9}{5}\text{tempC}_i - 32\text{night}_i - 32\text{day}_i = 0$$

Meaning that

$$\begin{bmatrix} 1 \\ -9/5 \\ -32 \\ -32 \end{bmatrix} \in \text{nullsp}(X)$$

Therefore, if $\vec{w}'$ is one solution, the complete set of solutions is

$$\left\{ \vec{w}' + t \begin{bmatrix} 1 \\ -9/5 \\ -32 \\ -32 \end{bmatrix} , t \in \mathbb{R} \right\}$$

c) (6 pts) First, assume $h(\vec{x}_i)$ is the model at the top of the page.

(i) What is the **largest possible** rank of the design matrix, X? (Note that we're asking about the full design matrix, not just its first two rows.)

largest possible value of $\text{rank}(X) = \boxed{3}$

(ii) True or False: The sum of the errors of the model's predictions is 0.

⬤ True   ◯ False

(iii) True or False: The sum of the errors of the model's predictions **on just the rows of the dataset corresponding to night games** is 0.

● True   ○ False

Solution: Let $\vec{e} = \vec{y} - X\vec{w}^*$ be the error vector. Since $\vec{w}^*$ satisfies the normal equation, $\vec{e}$ is orthogonal to every column of X, and to every linear combination of those columns.

(i) The largest possible rank of X is 3, which happens when the tempF, tempC, and night columns are linearly independent. When the day column is added, the columns become linearly dependent. (Equivalently, the tempF, night, and day columns are linearly independent, but linearly dependent with the tempC column.)

(ii) This is true. Even though there is no explicit intercept term (and thus, no column of all ones), the all-ones vector is still in $\text{colsp}(X)$, because the night and day columns add up to 1 for each row.

$$\vec{1} = \text{night column} + \text{day column}$$

Since $\vec{e}$ is orthogonal to every vector in $\text{colsp}(X)$, it is orthogonal to $\vec{1}$, so

$$\vec{1}^T \vec{e} = \sum_{i=1}^{n} e_i = 0$$

(iii) This is also true. The night indicator is itself a column of X, so

$$\vec{e} \cdot (\text{night column}) = \vec{e} \cdot \begin{bmatrix} 0 \\ 1 \\ \vdots \end{bmatrix} = e_2 + \ldots = 0$$

But this dot product is exactly the sum of the errors for just the night games, because the night column has 1s on night rows and 0s elsewhere.

---

d) (6 pts) Now, suppose we remove the **day**$_i$ feature from our model, meaning our model is

$$h(\vec{x}_i) = w_1 \cdot \text{tempF}_i + w_2 \cdot \text{tempC}_i + w_3 \cdot \text{night}_i$$

(i) After removing the day column, what is the **largest possible** rank of the **new** design matrix?

largest possible value of rank(new design matrix) = $\boxed{3}$

(ii) True or False: The sum of the errors of the new model's predictions is 0.

○ True   ● False

(iii) True or False: The sum of the errors of the new model's predictions **on just the rows of the dataset corresponding to night games** is 0.

● True   ○ False

Solution: Let $\vec{e}_{\text{new}}$ be the error vector for the new model.

(i) After removing the day column, the new design matrix has 3 columns, and there is no longer a forced linear dependence among them. So the largest possible rank is still 3. What's new now is that X's columns are all linearly independent, meaning there is a unique solution to the normal equation.

(ii) This is false. The normal equations still tell us that $\vec{e}_{\text{new}}$ is orthogonal to each column of the new design matrix, but there is no guarantee that the all-ones vector is in the column space anymore. So the errors are not guaranteed to sum to 0.

(iii) This is true. The night indicator column is still present in the new design matrix, so $\vec{e}_{\text{new}}$ is orthogonal to that column. Therefore, the sum of the errors over the night-game rows is still 0.

Problem 7 (10 pts)

a) (6 pts) Suppose A is a 4×4 matrix and $\vec{x} \in \mathbb{R}^4$. Furthermore, suppose that the gradient of the function $f(\vec{x}) = \vec{x}^T A \vec{x}$ is given by

$$\nabla f(\vec{x}) = \begin{bmatrix} 2x_1 \\ -15x_2 \\ 10x_3 \\ x_4 \end{bmatrix}$$

Find one possible matrix A. Your answer should be a 4×4 matrix with no variables.

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -15/2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}$$

Solution: Recall that for
$$f(\vec{x}) = \vec{x}^T A \vec{x},$$
the gradient is
$$\nabla f(\vec{x}) = (A + A^T)\vec{x}$$

We want
$$(A + A^T)\vec{x} = \begin{bmatrix} 2x_1 \\ -15x_2 \\ 10x_3 \\ x_4 \end{bmatrix}$$

One easy way to make this happen is to choose A to be diagonal and symmetric. Then $A + A^T = 2A$, so we want

$$2A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & -15 & 0 & 0 \\ 0 & 0 & 10 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

Thus, one possible choice is

$$A = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -15/2 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1/2 \end{bmatrix}$$

b) (4 pts) Suppose A is an $n \times n$ matrix, $\vec{b} \in \mathbb{R}^n$, and that $g : \mathbb{R}^n \to \mathbb{R}$ is defined by

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

Which of the following is $\nabla g(\vec{x})$, the gradient of $g(\vec{x})$?

○ $2A^T\vec{b}$
○ $(\vec{b}^T A \vec{x}) A^T \vec{b}$
○ $2(\vec{b}^T A \vec{x})\vec{b}$
● $2(\vec{b}^T A \vec{x}) A^T \vec{b}$
○ $2(\vec{b}^T A \vec{x}) A^T \vec{x}$
○ None of the above

Solution: Let
$$f(\vec{x}) = \vec{b}^T A \vec{x} = (A^T \vec{b})^T \vec{x} = (A^T \vec{b}) \cdot \vec{x}$$

Then
$$g(\vec{x}) = (g(\vec{x}))^2$$

The gradient of $f(\vec{x})$ can be computed using the dot product "big three" rule, which tells us that
$$\nabla f(\vec{x}) = A^T \vec{b}$$

Applying the chain rule,
$$\nabla g(\vec{x}) = 2f(\vec{x})\nabla f(\vec{x})$$
$$= 2(\vec{b}^T A \vec{x})A^T \vec{b}$$

Problem 8 (11 pts)

Let $\vec{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$. Consider the function

$$f(\vec{x}) = (x_1 + x_2 - 4)^2$$

a) (3 pts) Fill in the blanks: The set of all vectors $\vec{x}^*$ that minimize $f(\vec{x})$ form a __(i)__ in $\mathbb{R}^2$. This set __(ii)__ a subspace of $\mathbb{R}^2$.

(i) ○ point ● line ○ plane

(ii) ● is not ○ is

> **Solution:** We have
> $$f(\vec{x}) = (x_1 + x_2 - 4)^2$$
> so the minimum value is 0, which happens exactly when
> $$x_1 + x_2 - 4 = 0 \iff x_1 + x_2 = 4$$
> The equation
> $$x_1 + x_2 = 4$$
> describes a **line** in $\mathbb{R}^2$.
> This line is **not** a subspace, because it does not pass through the origin. For example,
> $$\begin{bmatrix} 0 \\ 0 \end{bmatrix}$$
> is not a minimizer. Therefore, this set **is not** a subspace.

b) (8 pts) Suppose we use gradient descent to minimize $f(\vec{x})$ using an initial guess of $\vec{x}^{(0)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Find the learning rate/step size α that will cause gradient descent to converge to a global minimum of $f(\vec{x})$ **in one iteration**, i.e. such that $\vec{x}^{(1)}$ is a minimizer of $f(\vec{x})$.

Show your work in the space provided, and write your final answer in the bottom-right corner of the box. Your answer should be a number with no variables.

$$\alpha = \boxed{1/4}$$

Solution: First, we need to compute the gradient of $f(\vec{x})$:

$$f(\vec{x}) = (x_1 + x_2 - 4)^2 \implies \nabla f(\vec{x}) = \begin{bmatrix} 2(x_1 + x_2 - 4) \\ 2(x_1 + x_2 - 4) \end{bmatrix}$$

At

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

we have

$$x_1^{(0)} + x_2^{(0)} - 4 = 1 + 1 - 4 = -2$$

so

$$\nabla f(\vec{x}^{(0)}) = \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

One gradient descent step gives

$$\vec{x}^{(1)} = \vec{x}^{(0)} - \alpha \nabla f(\vec{x}^{(0)})$$

$$= \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \alpha \begin{bmatrix} -4 \\ -4 \end{bmatrix}$$

$$= \begin{bmatrix} 1 + 4\alpha \\ 1 + 4\alpha \end{bmatrix}$$

We want $\vec{x}^{(1)}$ to be a minimizer, so it must satisfy

$$x_1^{(1)} + x_2^{(1)} = 4$$

That gives

$$(1 + 4\alpha) + (1 + 4\alpha) = 4$$
$$2 + 8\alpha = 4$$
$$8\alpha = 2$$
$$\alpha = \frac{1}{4}$$

Congrats on finishing Midterm 2!

Feel free to draw us a picture about EECS 245 in the box below (or use it for scratch work).

Did you notice any violations of the Honor Code during the exam? If so, share details with us here. We will keep your identity anonymous when investigating any cases.