Homework 5: Matrices

EECS 245, Fall 2025 at the University of Michigan

due Thursday, October 9th, 2025 at 11:59PM Ann Arbor Time

> Write your solutions to the following problems by either typing them up or handwriting them on another piece of paper. Homeworks are due to Gradescope by 11:59PM on the due date. See the **syllabus** for details on the slip day policy.
>
> Homework will be evaluated not only on the correctness of your answers, but on your ability to present your ideas clearly and logically. You should always explain and justify your conclusions, using sound reasoning. Your goal should be to convince the reader of your assertions. If a question does not require explanation, it will be explicitly stated.
>
> Before proceeding, make sure you're familiar with the **collaboration policy**.

Total Points: 15 + 12 + 11 + 14 + 15 = 67

Note: In some of the problems in this homework, we'll explicitly mention that you can use Python and `numpy` to perform some of the relevant calculations. For a reference on how to do so, consult **Chapter 2.7**. In other problems, we'll explicitly state that you must execute all calculations by hand.

Problem 1: Getting Started (15 pts)

Throughout this problem, let

$$A = \begin{bmatrix} 3 & 0 & 4 \\ 0 & 1 & 0 \\ 2 & -1 & -3 \\ 5 & 0 & -1 \\ 3 & 2 & 0 \end{bmatrix}$$

In each part, show all of your work. **You must execute all calculations by hand here; don't use Python or any other calculators.**

a) (4 pts) In each subpart, state whether the resulting object is a matrix, vector, or scalar. If the result is a matrix or vector, state its dimensions. If the result is not defined, state why. You don't need to actually compute the resulting objects.

(i) A^T

(ii) $A^T A$

(iii) AA^T

(iv) $A^T A + AA^T$

(v) $A^T \vec{x}$, where $\vec{x} \in \mathbb{R}^3$

(vi) $A^T \vec{x}$, where $\vec{x} \in \mathbb{R}^5$

(vii) $\vec{x}^T A^T A \vec{x}$, where $\vec{x} \in \mathbb{R}^3$

b) (3 pts) Evaluate $A \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix}$.

There are two interpretations of the resulting vector, based on what we've seen in Chapter 2.7 — what are they?

c) (5 pts) In both subparts, try and find a vector $\vec{x} \in \mathbb{R}^3$ such that $A\vec{x} = \vec{b}$. If it's not possible to do so, explain why.

(i) $\vec{b} = \begin{bmatrix} 0 \\ 5 \\ 3 \\ -1 \\ 4 \end{bmatrix}$

(ii) $\vec{b} = \begin{bmatrix} 10 \\ 1 \\ -17 \\ -14 \\ -4 \end{bmatrix}$

d) (3 pts) Explain why it's the case that — for this particular matrix A — if $A\vec{x}_1 = \vec{b}$ and $A\vec{x}_2 = \vec{b}$, then $\vec{x}_1 = \vec{x}_2$. *Hint: Think in terms of linear independence. If you want to claim that a collection of vectors is linearly independent, you'll need to use the formal definition.*

Problem 2: CR Decomposition (12 pts)

In Chapter 2.8, we'll introduce several key ideas involving matrices and the vector spaces that they define. This problem is meant to get you to start thinking about those ideas before Tuesday's lecture.

Consider the matrix A below.

$$A = \begin{bmatrix} 5 & 3 & 5 & 2 \\ 3 & 0 & -6 & 4 \\ -2 & 0 & 4 & 3 \\ 8 & 2 & -6 & -8 \\ 1 & 1 & 3 & 0 \end{bmatrix}$$

A might look familiar: it's the same matrix we used in Problem 5 of Homework 4. As we saw in Homework 4 (**solutions here, keep this open**), A has 3 linearly independent columns and 3 linearly independent rows.

In particular, columns 1, 2, and 4 of A are linearly independent; column 3 is a linear combination of columns 1 and 2.

Let's define the matrix C as containing the linearly independent columns of A, collected as we read from left to right.

$$C = \begin{bmatrix} 5 & 3 & 2 \\ 3 & 0 & 4 \\ -2 & 0 & 3 \\ 8 & 2 & -8 \\ 1 & 1 & 0 \end{bmatrix}$$

Note that C is a 5×3 matrix, where 5 is the number of rows of A and 3 is the number of linearly independent columns of A.

What we'd like to do is find a matrix R such that

$$A = CR$$

R would tell us how to "mix" the linearly independent columns of A (which are stored in C) to get back the original matrix A. In order for the dimensions of A and CR to match, R must be a 3×4 matrix — one row per linearly independent column of A, and one column per column of A.

This way of writing A is called the **CR decomposition** of A. Note that we specified that we're reading from left to right to construct C. If we instead read from right to left, we would have instead kept columns 4, 3, and 2, since then column 1 is a linear combination of columns 3 and 2. That would lead to a different —but still valid — C, and also a different R.

a) (3 pts) With some careful reading of the solutions of Homework 4, we find that

$$\underbrace{\begin{bmatrix} 5 & 3 & 2 \\ 3 & 0 & 4 \\ -2 & 0 & 3 \\ 8 & 2 & -8 \\ 1 & 1 & 0 \end{bmatrix}}_{C} \underbrace{\begin{bmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}}_{R} = A$$

Explain how we found R by reading the solutions to Problem 5 of Homework 4.

b) (9 pts) **By hand**, find a CR decomposition of the matrices below, by placing the linearly independent columns (reading from left to right) in C and the values needed to "mix" the linearly independent columns in C to get back the original matrix in R.

Hint: In at least one part, you'll find that there are multiple possible ways to write R.

(i) $A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$

(ii) $A = \begin{bmatrix} 3 & 5 \\ 1 & 1 \\ 2 & -4 \\ 30 & 0 \end{bmatrix}$

(iii) $A = \begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 4 & 1 & -5 \\ 3 & -6 & 4 & 2 \\ 0 & 0 & 5 & -5 \end{bmatrix}$

Problem 3: Correlation, Revisited (11 pts)

In this problem, we'll see how the correlation coefficient between two variables, r, can be expressed as a matrix multiplication.

Consider a dataset of n points, $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Define the data matrix D as follows:

$$D = \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix}$$

where $\bar{x}$ and $\bar{y}$ are the means of x and y, respectively. Note that D is an $n \times 2$ matrix, and it is mean-centered, meaning that the mean of each column is 0.

Define the matrix Σ as follows:

$$\Sigma = \frac{1}{n} D^T D$$

Σ is a 2×2 matrix. Its name is pronounced "sigma", just like in summation notation and standard deviation. Don't confuse it with summation notation; Σ is just a single matrix.

a) (4 pts) For this particular matrix D, find Σ. All four components of Σ should be expressions involving the points $x_1, x_2, \ldots, x_n$ and/or $y_1, y_2, \ldots, y_n$. Feel free to use summation notation in your answers.

b) (2 pts) In English, what do the two elements on the diagonal (top-left and bottom-right) of Σ represent?

c) (3 pts) You should notice that Σ is a **symmetric** matrix, meaning that $\Sigma^T = \Sigma$. The elements off the diagonal (top-right and bottom-left) are both equal, and are called the **covariance** of x and y. For that reason, Σ is often called the **covariance matrix**.

Find an expression for the off-diagonal elements of Σ in terms of the correlation coefficient, r, σ_x, and σ_y, but with no summation notation or other variables.

Hint: This only requires 1-2 lines of work. Remember the definition of r from Chapter 1.4.

d) (2 pts) In general, suppose $X \in \mathbb{R}^{n \times d}$ is a matrix containing n observations for each of d variables/features. The covariance matrix of X is defined similarly:

$$\Sigma = \frac{1}{n} X^T X$$

In English, explain what the element in row 3 and column 5 of this Σ represents.

Problem 4: Projections, Revisited (14 pts)

As we first saw in Chapter 2.3, the projection of $\vec{u}$ onto $\vec{v}$ is the vector

$$\vec{p} = \left(\frac{\vec{u} \cdot \vec{v}}{\vec{v} \cdot \vec{v}} \right) \vec{v}$$

If we assume that $\vec{v}$ is a unit vector, meaning $\|\vec{v}\| = 1$, then the projection of $\vec{u}$ onto $\vec{v}$ has a simpler form,

$$\vec{p} = (\vec{u} \cdot \vec{v})\vec{v}$$

For simplicity, assume that $\vec{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ is some arbitrary (not-necessarily unit) vector in $\mathbb{R}^2$, and $\vec{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$ is **a unit vector** in $\mathbb{R}^2$.

a) (6 pts) Find a 2×2 matrix P, called a **projection matrix**, such that

$$P\vec{u} = \vec{p} = (\vec{u} \cdot \vec{v})\vec{v}$$

Think of P as a matrix that transforms $\vec{u}$ into an approximation of it, in the direction of $\vec{v}$ (or "projects" $\vec{u}$ onto $\vec{v}$).

Hint: Start by writing $P = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ *and solve for* a, b, c, d *in terms of* v_1 *and* v_2*;* P *should not involve* u_1 *or* u_2*. Don't forget that* $\vec{v}$ *is a unit vector, and both* $\vec{u}, \vec{v} \in \mathbb{R}^2$*.*

b) (4 pts) Find the projection of $\vec{u} = \begin{bmatrix} 9 \\ -3 \end{bmatrix}$ onto the unit vector $\vec{v} = \begin{bmatrix} 3/5 \\ 4/5 \end{bmatrix}$ using:

(i) The formula for the projection of $\vec{u}$ onto $\vec{v}$

(ii) The projection matrix P you found in part **a)**

Feel free to use Python and `numpy` to compute the relevant products as we do in Chapter 2.7, but if you do so, include screenshots of your code and results, and also write out the final result by hand. If you just write the final result with no work shown, you will not receive any credit.

c) (4 pts) Show that P satisfies the following property:

$$P^2 = P$$

This means that P is an **idempotent** matrix, meaning that applying P twice (or three times, or four times, etc.) to a vector is the same as applying it once.

Hint: You'll likely end up with terms of the form v_1^4*. Remember that* $\vec{v}$ *is a unit vector; use this to help you simplify.*

Problem 5: Orthogonal Matrices (15 pts)

We say that vectors $\vec{q}_1, \vec{q}_2, \ldots, \vec{q}_d$ form an **orthonormal set** if:

- All vectors are unit vectors, i.e. $\|\vec{q}_i\| = 1$ for all i.
- All vectors are orthogonal to each other, i.e. $\vec{q}_i \cdot \vec{q}_j = 0$ for all $i \neq j$.

a) (3 pts) Let $A = \begin{bmatrix} 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \\ 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \end{bmatrix}$. Show that the columns of A form an orthonormal set, using just the two conditions above.

b) (6 pts) We say an $n \times n$ matrix Q is **orthogonal** if its columns form an orthonormal set. Note that for a matrix to be orthogonal, it's not enough that its columns are orthogonal to one another; they must also be unit vectors, and the matrix must also be square.

A above is an example of an orthogonal matrix.

The condition of being orthogonal is written as

$$Q^T Q = QQ^T = I$$

where I is the identity matrix, consisting of all 1s on the diagonal and 0s everywhere else.

For each of the following matrices, compute A^TA and AA^T; B^TB and BB^T; ... and use that to determine whether it is orthogonal. If a matrix is not orthogonal, explain which of the conditions for being orthogonal it does and does not satisfy.

$$A = \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 4 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 0 & 1 \\ 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \end{bmatrix}, \quad D = \begin{bmatrix} 3/5 & 0 & 1/\sqrt{2} \\ 4/5 & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix}$$

Note: Feel free to use Python and `numpy` to compute the relevant products as we do in Chapter 2.7, but if you do so, include screenshots of your code and results, and also write out the final result by hand. If you just write the final result with no work shown, you will not receive any credit.

c) (3 pts) Explain why the following statement is true: If Q is an orthogonal matrix, then the **rows** of Q form an orthonormal set, in addition to the columns.

Hint: Think about what Q^TQ and QQ^T each are.

d) (3 pts) Orthogonal matrices have many useful properties. One is that they **preserve the norm** of vectors. In other words, if $Q \in \mathbb{R}^{n \times n}$ is orthogonal and $\vec{x} \in \mathbb{R}^n$, then:

$$\|Q\vec{x}\| = \|\vec{x}\|$$

Prove the statement above. *Hint: Review this section of Chapter 2.7 if you're stuck on how to get started.*

e) (Optional, 0 pts) We'll comment more on how orthogonal matrices are useful in practice in Chapter 2.9. This optional part is meant to get you to start thinking about their uses.

At the end of **Chapter 2.7**, we presented the matrix

$$A = \begin{bmatrix} \frac{\sqrt{3}}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{\sqrt{3}}{2} \end{bmatrix}$$

and visualized three vectors, $\vec{u}$, $\vec{v}$, and $\vec{w}$ and the result of multiplying each one by A. We defined A as a rotation matrix; specifically, one that rotates vectors by $\theta = 30°$ counterclockwise. (Go and look at the picture there for context; we're intentionally not providing it here so that you have to look at the notes!)

In general, the 2×2 rotation matrix by an angle θ is given by

$$R = \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix}$$

Prove that the matrix R is orthogonal.

Hint: There's an identity on ***this page*** *that you'll need.*