

## Homework 5: Matrices Solutions

EECS 245, Spring 2026 at the University of Michigan

due Thursday, May 28th, 2026 at 11:59PM Ann Arbor Time

Write your solutions to the following problems either by writing them on a piece of paper or on a tablet and scanning your answers as a PDF. Note that you are not allowed to use LaTeX, Google Docs, or any other digital document creation software to type your answers. 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:  $10 + 15 + 11 + 14 + 15 + 9 = 74$

**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 5.1](#) or [this video](#). In other problems, we'll explicitly state that you must execute all calculations by hand.

### Problem 1: Midterm 1 Solutions Review (10 pts)

Review the solutions to Midterm 1. Pick **two problem parts** (for example, Problem 3b and Problem 7a) from Midterm 1 in which your solutions have the most room for improvement, i.e., where they have unsound reasoning, could be significantly more efficient or clearer, etc. **Include a screenshot of your solution to each problem part**, and in a few sentences, explain what was deficient and how it could be fixed.

Alternatively, if you think one of your solutions is significantly better than the posted one, copy it here and explain why you think it is better. If you didn't do Midterm 1, choose two problem parts from it that look challenging to you, and in a few sentences, explain the key ideas behind their solutions in your own words.

**Solution:**

## Problem 2: Getting Started (15 pts)

Let

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

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  $\vec{x} \in \mathbb{R}^3$

### Solution:

- (i)  $A^T$  is a matrix in  $\mathbb{R}^{3 \times 5}$ , since we replace the rows with the columns, and  $A$  is a  $\mathbb{R}^{5 \times 3}$  matrix.
- (ii)  $A^T A$  is a matrix in  $\mathbb{R}^{3 \times 3}$ . We can multiply  $A^T$  with  $A$  because their inner dimensions match ( $\mathbb{R}^{(3 \times 5)} \times \mathbb{R}^{(5 \times 3)}$ ).
- (iii)  $AA^T$  is a matrix in  $\mathbb{R}^{5 \times 5}$  because the inner dimensions match ( $\mathbb{R}^{(5 \times 3)} \times \mathbb{R}^{(3 \times 5)}$ ).
- (iv) Undefined.  $A^T A$  and  $AA^T$  have different shapes, so we can't add them together.
- (v) Undefined, because the inner dimensions of the product don't match ( $\mathbb{R}^{(3 \times 5)} \times \mathbb{R}^{(3 \times 1)}$ ).
- (vi)  $A^T \vec{x}$  is a vector in  $\mathbb{R}^3$ .
- (vii)  $\vec{x}^T A^T A \vec{x}$  is a scalar. The steps for this are below:

$$\begin{aligned} \vec{x}^T A^T A \vec{x} &= \vec{x}^T (A^T A) \vec{x} \quad \text{using associative property} \\ &= \vec{x}^T (\mathbb{R}^{3 \times 3}) \vec{x} \quad \text{substituting from part (ii)} \\ &= \mathbb{R}^{1 \times 3} (\mathbb{R}^{3 \times 3}) \vec{x} \quad \text{transpose } \vec{x} \\ &= \mathbb{R}^{1 \times 3} \vec{x} \quad \text{resolve the left product} \\ &= \mathbb{R}^{1 \times 1} \end{aligned}$$

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 5.1](#) — what are they?

**Solution:**

$$\begin{aligned} \begin{bmatrix} 3 & 0 & 4 \\ 0 & 1 & 0 \\ 2 & -1 & -3 \\ 5 & 0 & -1 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \\ -2 \end{bmatrix} &= \begin{bmatrix} 3 \cdot 3 + 0 \cdot 0 + 4 \cdot (-2) \\ 0 \cdot 3 + 1 \cdot 0 + 0 \cdot (-2) \\ 2 \cdot 3 + (-1) \cdot 0 + (-3) \cdot (-2) \\ 5 \cdot 3 + 0 \cdot 0 + (-1) \cdot (-2) \\ 3 \cdot 3 + 2 \cdot 0 + 0 \cdot (-2) \end{bmatrix} \\ &= \begin{bmatrix} 9 + 0 + (-8) \\ 0 + 0 + 0 \\ 6 + 0 + 6 \\ 15 + 0 + 2 \\ 9 + 0 + 0 \end{bmatrix} \\ &= \begin{bmatrix} 1 \\ 0 \\ 12 \\ 17 \\ 9 \end{bmatrix} \end{aligned}$$

The two key interpretations of matrix-vector multiplication from [Chapter 5.1](#) are:

- Dot product with the rows of  $A$ :**  $A\vec{x}$  represents the dot product of the rows of  $A$  with the vector  $\vec{x}$ .
- Linear combination of the columns of  $A$ :**  $A\vec{x}$  represents the linear combination of the columns of  $A$  with coefficients given by the components of  $\vec{x}$ . This is equivalent to thinking of  $A\vec{x}$  as a transformation of  $\vec{x}$  by the matrix  $A$ , sending it from  $\mathbb{R}^3$  to  $\mathbb{R}^5$ .

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

**Solution:**

(i)

$$A\vec{x} = \vec{b}$$
$$\begin{bmatrix} 3 & 0 & 4 \\ 0 & 1 & 0 \\ 2 & -1 & -3 \\ 5 & 0 & -1 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 3 \\ -1 \\ 4 \end{bmatrix}$$
$$x_1 \begin{bmatrix} 3 \\ 0 \\ 2 \\ 5 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 0 \\ -3 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \\ 3 \\ -1 \\ 4 \end{bmatrix}$$

Now, we try to see if we can make  $\vec{b}$  as a linear combination of the columns of  $A$ :

$$\begin{aligned} 3x_1 + 4x_3 &= 0 \\ x_2 &= 5 \\ 2x_1 - x_2 - 3x_3 &= 3 \\ 5x_1 - x_3 &= -1 \\ 3x_1 + 2x_2 &= 4 \end{aligned}$$

The second equation tells us that  $x_2 = 5$ , so we can plug that into the last one to solve for  $x_1$ :

$$\begin{aligned} 3x_1 + 2(5) &= 4 \\ 3x_1 &= -6 \\ x_1 &= -2 \end{aligned}$$

Using  $x_1$ , we can see if there's a contradiction in the first and fourth equations:

$$\begin{aligned} 3(-2) + 4x_3 &= 0 \\ 5(-2) - x_3 &= -1 \end{aligned}$$

In order for the equations to be true,  $x_3$  must both be negative and positive, so we don't have to continue. So, there is no possible  $\vec{x}$ , since  $\vec{b}$  is not in the column space of  $A$ .

(ii) We can do a similar process for this part, even down to the steps of solving the system of linear equations (since we'll always know  $x_2$ , plug into the last equation

to solve for  $x_1$ , and so on):

$$\begin{bmatrix} 3 & 0 & 4 \\ 0 & 1 & 0 \\ 2 & -1 & -3 \\ 5 & 0 & -1 \\ 3 & 2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ -17 \\ -14 \\ -4 \end{bmatrix}$$
$$x_1 \begin{bmatrix} 3 \\ 0 \\ 2 \\ 5 \\ 3 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \\ -1 \\ 0 \\ 2 \end{bmatrix} + x_3 \begin{bmatrix} 4 \\ 0 \\ -3 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 10 \\ 1 \\ -17 \\ -14 \\ -4 \end{bmatrix}$$

$$3x_1 + 4x_3 = 10$$

$$x_2 = 1$$

$$2x_1 - x_2 - 3x_3 = -17$$

$$5x_1 - x_3 = -14$$

$$3x_1 + 2x_2 = -4$$

$$3x_1 + 2(1) = -4$$

$$3x_1 = -6$$

$$x_1 = -2$$

Now we solve for  $x_3$ , starting with the first equation:

$$3(-2) + 4x_3 = 10$$

$$-6 + 4x_3 = 10$$

$$4x_3 = 16$$

$$x_3 = 4$$

Plugging that value of  $x_3$  into the fourth equation results in a valid equation, so all that's left is to check the third equation:

$$2(-2) - 1 - 3(4) = -17$$

$$-4 - 1 - 12 = -17$$

All the equations hold, so the vector  $\vec{x} = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}$  satisfies  $A\vec{x} = \vec{b}$

Since we were able to find an  $\vec{x}$  that satisfies  $A\vec{x} = \vec{b}$ ,  $\vec{b}$  is in the column space of  $A$ .

- 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  $\vec{x}_1 = \vec{x}_2$ .

**Solution:** The reason has to do with the fact that the column vectors of  $A$  are linearly independent. Recall,

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

We won't show here that the columns are linearly independent; this is something you should verify on your own.

Let the columns of  $A$  be  $\vec{v}^{(1)}$ ,  $\vec{v}^{(2)}$ , and  $\vec{v}^{(3)}$ . Since they are linearly independent, the only solution to

$$a_1\vec{v}^{(1)} + a_2\vec{v}^{(2)} + a_3\vec{v}^{(3)} = \vec{0}$$

is  $a_1 = a_2 = a_3 = 0$ .

We'll now show that if  $A\vec{x}_1 = \vec{b}$  and  $A\vec{x}_2 = \vec{b}$ , then  $\vec{x}_1 = \vec{x}_2$ . Let  $\vec{x}_1 = \begin{bmatrix} x_{11} \\ x_{12} \\ x_{13} \end{bmatrix}$  and  $\vec{x}_2 = \begin{bmatrix} x_{21} \\ x_{22} \\ x_{23} \end{bmatrix}$ .

Given that  $A\vec{x}_1 = \vec{b}$ , we have

$$A\vec{x}_1 = \vec{b} \implies x_{11}\vec{v}^{(1)} + x_{12}\vec{v}^{(2)} + x_{13}\vec{v}^{(3)} = \vec{b}$$

Similarly, given that  $A\vec{x}_2 = \vec{b}$ , we have

$$A\vec{x}_2 = \vec{b} \implies x_{21}\vec{v}^{(1)} + x_{22}\vec{v}^{(2)} + x_{23}\vec{v}^{(3)} = \vec{b}$$

Subtracting the two equations, we get

$$(x_{11} - x_{21})\vec{v}^{(1)} + (x_{12} - x_{22})\vec{v}^{(2)} + (x_{13} - x_{23})\vec{v}^{(3)} = \vec{0}$$

Since the columns are linearly independent, we know the three coefficients must all be zero. Therefore,

$$x_{11} - x_{21} = 0, x_{12} - x_{22} = 0, x_{13} - x_{23} = 0$$

Therefore,  $\vec{x}_1 = \vec{x}_2$ . The more intuitive phrasing of this result is that if a set of vectors are linearly independent, then any vector that can be written as a linear combination of them can only be written in one way, or "linear combinations of linearly independent vectors are unique."

### 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,  $(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ , and let

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

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  $\Sigma$  as follows.

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

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

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

**Solution:**

$$\begin{aligned} \Sigma &= \frac{1}{n} D^T D \\ &= \frac{1}{n} \underbrace{\begin{bmatrix} x_1 - \bar{x} & x_2 - \bar{x} & \cdots & x_n - \bar{x} \\ y_1 - \bar{y} & y_2 - \bar{y} & \cdots & y_n - \bar{y} \end{bmatrix}}_{D^T, \text{ shape } 2 \times n} \underbrace{\begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ x_2 - \bar{x} & y_2 - \bar{y} \\ \cdots & \cdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix}}_{D, \text{ shape } n \times 2} \\ &= \frac{1}{n} \begin{bmatrix} (x_1 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2 & (x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y}) \\ (x_1 - \bar{x})(y_1 - \bar{y}) + \cdots + (x_n - \bar{x})(y_n - \bar{y}) & (y_1 - \bar{y})^2 + \cdots + (y_n - \bar{y})^2 \end{bmatrix} \\ &= \frac{1}{n} \begin{bmatrix} \sum_{i=1}^n (x_i - \bar{x})^2 & \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^n (y_i - \bar{y})^2 \end{bmatrix} \end{aligned}$$

If we'd like, we can distribute the  $\frac{1}{n}$  to get

$$\Sigma = \begin{bmatrix} \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})^2 & \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \\ \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) & \frac{1}{n} \sum_{i=1}^n (y_i - \bar{y})^2 \end{bmatrix}$$

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

**Solution:** The top-left element represents the variance of  $x$ ,  $\sigma_x^2$ , and the bottom-right element represents the variance of  $y$ ,  $\sigma_y^2$ .

- c) (3 pts) You should notice that  $\Sigma$  is a *symmetric* matrix, meaning  $\Sigma^T = \Sigma$ . (See [Chapter 5.2](#) for more on symmetric matrices.) 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,  $\Sigma$  is often called the **covariance matrix**.

Find an expression for the off-diagonal elements of  $\Sigma$  in terms of the correlation coefficient,  $r$ ,  $\sigma_x$ , and  $\sigma_y$ , but with no summation notation or other variables.

*Hint: This only requires 1-2 lines of work. Remember the definition of  $r$  from [Chapter 2.4](#).*

**Solution:** The off-diagonal elements, both of which are equal to

$$\text{covariance} = \frac{1}{n} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})$$

are equal to

$$\text{covariance} = r\sigma_x\sigma_y$$

This comes from the definition of  $r$ .

$$r = \frac{1}{n} \sum_{i=1}^n \left( \frac{x_i - \bar{x}}{\sigma_x} \right) \left( \frac{y_i - \bar{y}}{\sigma_y} \right)$$

Multiplying both sides in the definition of  $r$  by  $\sigma_x\sigma_y$  gives us the desired result.

- 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  $\Sigma$  represents.

**Solution:** This is the covariance between feature 3 and feature 5, which is the same as their correlation, multiplied by the standard deviation of feature 3 and the standard deviation of feature 5.

#### Problem 4: Projections, Revisited (14 pts)

As we first saw in [Chapter 3.4](#), 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  $2 \times 2$  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  $\vec{u}, \vec{v} \in \mathbb{R}^2$ .*

#### Solution:

$$\begin{aligned} P\vec{u} &= (\vec{u} \cdot \vec{v})\vec{v} \\ \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} &= \begin{bmatrix} (u_1 v_1 + u_2 v_2) v_1 \\ (u_1 v_1 + u_2 v_2) v_2 \end{bmatrix} \\ &= \begin{bmatrix} u_1 v_1^2 + u_2 v_1 v_2 \\ u_1 v_1 v_2 + u_2 v_2^2 \end{bmatrix} \end{aligned}$$

$$u_1 a + u_2 b = u_1 v_1^2 + u_2 v_1 v_2$$

$$u_1 c + u_2 d = u_1 v_1 v_2 + u_2 v_2^2$$

$$a = v_1^2, \quad b = v_1 v_2, \quad c = v_1 v_2, \quad d = v_2^2$$

$$P = \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix}$$

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 5.2](#) and [this video](#), 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.

**Solution:**

$$\begin{aligned}\vec{p} &= (\vec{u} \cdot \vec{v})\vec{v} \\ &= \left(9 \cdot \frac{3}{5} + (-3) \cdot \frac{4}{5}\right)\vec{v} \\ &= \left(\frac{27}{5} - \frac{12}{5}\right)\vec{v} \\ &= 3\vec{v} \\ &= \begin{bmatrix} 9/5 \\ 12/5 \end{bmatrix}\end{aligned}$$

$$\begin{aligned}P &= \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix} \\ &= \begin{bmatrix} (9/25) & (12/25) \\ (12/25) & (16/25) \end{bmatrix}\end{aligned}$$

$$\begin{aligned}P\vec{u} &= \begin{bmatrix} (9/25) & (12/25) \\ (12/25) & (16/25) \end{bmatrix} \begin{bmatrix} 9 \\ -3 \end{bmatrix} \\ &= \begin{bmatrix} 9(9/25) + (-3)(12/25) \\ 9(12/25) + (-3)(16/25) \end{bmatrix} \\ &= \begin{bmatrix} 81/25 - 36/25 \\ 108/25 - 48/25 \end{bmatrix} \\ &= \begin{bmatrix} 45/25 \\ 60/25 \end{bmatrix} \\ &= \begin{bmatrix} 9/5 \\ 12/5 \end{bmatrix}\end{aligned}$$

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.*

**Solution:**

$$\begin{aligned} PP &= \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix} \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix} \\ &= \begin{bmatrix} v_1^4 + v_1^2 v_2^2 & v_1^3 v_2 + v_1 v_2^3 \\ v_1^3 v_2 + v_1 v_2^3 & v_1^2 v_2^2 + v_2^4 \end{bmatrix} \\ &= \begin{bmatrix} v_1^2(v_1^2 + v_2^2) & v_1 v_2(v_1^2 + v_2^2) \\ v_1 v_2(v_1^2 + v_2^2) & v_2^2(v_1^2 + v_2^2) \end{bmatrix} \end{aligned}$$

$$\begin{aligned} v_1^2 + v_2^2 &= \vec{v} \cdot \vec{v} \\ &= \|\vec{v}\|^2 = 1 \text{ because } \vec{v} \text{ is a unit vector} \end{aligned}$$

$$PP = \begin{bmatrix} v_1^2 & v_1 v_2 \\ v_1 v_2 & v_2^2 \end{bmatrix} = P$$

### Problem 5: Orthogonal Matrices (15 pts)

Read [Orthogonal Matrices section of Chapter 5.2](#) before starting this problem.

- a) (6 pts) For each of the following matrices, compute  $A^T A$  and  $AA^T$ ;  $B^T B$  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 5.2](#) and [this video](#), 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.*

**Solution:**

$$\begin{aligned} A^T A &= \begin{bmatrix} 3 & 0 & 4 \\ 0 & 2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 4 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 25 & 0 \\ 0 & 4 \end{bmatrix} \neq I \end{aligned}$$

$$\begin{aligned} A A^T &= \begin{bmatrix} 3 & 0 \\ 0 & 2 \\ 4 & 0 \end{bmatrix} \begin{bmatrix} 3 & 0 & 4 \\ 0 & 2 & 0 \end{bmatrix} \\ &= \begin{bmatrix} 9 & 0 & 12 \\ 0 & 4 & 0 \\ 12 & 0 & 16 \end{bmatrix} \neq I \end{aligned}$$

$A$ 's columns are orthogonal to one another, but they're not unit vectors. Also, since  $A^T A$  and  $A A^T$  have different shapes, they can't possibly be equal, which gives us another implicit condition: only square matrices can be orthogonal.

$$\begin{aligned} B^T B &= \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix} \\ &= \begin{bmatrix} 9/25 + 16/25 & -12/25 + 12/25 \\ -12/25 + 12/25 & 16/25 + 9/25 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

$$\begin{aligned} B B^T &= \begin{bmatrix} 3/5 & -4/5 \\ 4/5 & 3/5 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 \\ -4/5 & 3/5 \end{bmatrix} \\ &= \begin{bmatrix} 9/25 + 16/25 & 12/25 - 12/25 \\ 12/25 - 12/25 & 16/25 + 9/25 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

$B$  satisfies both conditions, so it is orthogonal!

$$\begin{aligned}
C^T C &= \begin{bmatrix} 0 & 3/5 & 4/5 \\ 0 & 4/5 & -3/5 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 9/25 + 16/25 & 12/25 - 12/25 & 0 \\ 12/25 - 12/25 & 16/25 + 9/25 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I
\end{aligned}$$

$$\begin{aligned}
C C^T &= \begin{bmatrix} 0 & 0 & 1 \\ 3/5 & 4/5 & 0 \\ 4/5 & -3/5 & 0 \end{bmatrix} \begin{bmatrix} 0 & 3/5 & 4/5 \\ 0 & 4/5 & -3/5 \\ 1 & 0 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 9/25 + 16/25 & 0 \\ 0 & 12/25 - 12/25 & 16/25 + 9/25 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I
\end{aligned}$$

$C$  satisfies both conditions, so it is orthogonal!

$$\begin{aligned}
D^T D &= \begin{bmatrix} 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \begin{bmatrix} 3/5 & 0 & 1/\sqrt{2} \\ 4/5 & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \\
&= \begin{bmatrix} 9/25 + 16/25 & 0 & 3/5\sqrt{2} + 4/5\sqrt{2} \\ 0 & 1 & 0 \\ 3/5\sqrt{2} + 4/5\sqrt{2} & 0 & 1/2 + 1/2 \end{bmatrix} \\
&= \begin{bmatrix} 1 & 0 & 7/5\sqrt{2} \\ 0 & 1 & 0 \\ 7/5\sqrt{2} & 0 & 1 \end{bmatrix} \neq I
\end{aligned}$$

$$\begin{aligned}
D D^T &= \begin{bmatrix} 3/5 & 0 & 1/\sqrt{2} \\ 4/5 & 0 & 1/\sqrt{2} \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 3/5 & 4/5 & 0 \\ 0 & 0 & 1 \\ 1/\sqrt{2} & 1/\sqrt{2} & 0 \end{bmatrix} \\
&= \begin{bmatrix} 9/25 + 1/2 & 12/25 + 1/2 & 0 \\ 12/25 + 1/2 & 16/25 + 1/2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} 43/50 & 49/50 & 0 \\ 49/50 & 57/50 & 0 \\ 0 & 0 & 1 \end{bmatrix} \neq I
\end{aligned}$$

$D$ 's columns are unit vectors, but they aren't orthogonal to one another, as evidenced by the non-diagonal elements being non-zero in  $D^T D$ . So,  $D$  is not orthogonal.

- b) (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  $Q^T Q$  and  $Q Q^T$  each are.*

**Solution:** For  $Q$  to be an orthogonal matrix,  $Q^T Q = Q Q^T = I$  must be true.  $Q^T Q$  is a matrix containing dot products of the columns, while  $Q Q^T$  has dot products of the rows. Recall that for any vector  $\vec{v}$ ,  $\vec{v} \cdot \vec{v} = \|\vec{v}\|^2$ . For a matrix  $Q^T Q$ , each diagonal value is the dot product of a column with itself, and for  $Q Q^T$  the diagonal has the dot products of rows with themselves. If the diagonal values are 1, then the length of the rows and columns must be 1.

Using similar logic, the off diagonal values of  $Q^T Q$  and  $Q Q^T$  are 0, meaning the rows and columns have to be orthogonal to each other.

- c) (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.

**Solution:**

$$\begin{aligned}\|Q\vec{x}\|^2 &= (Q\vec{x})^T (Q\vec{x}) \\ &= \vec{x}^T Q^T Q \vec{x} \\ &= \vec{x}^T I \vec{x} \\ &= \vec{x}^T \vec{x} \\ &= \|\vec{x}\|^2\end{aligned}$$

Since  $\|Q\vec{x}\|^2 = \|\vec{x}\|^2$ , we can take the square root of both sides to get  $\|Q\vec{x}\| = \|\vec{x}\|$ . (This is legal since  $\|Q\vec{x}\|$  and  $\|\vec{x}\|$  are both non-negative.)

- d) (3 pts) At the end of [Chapter 5.2](#), 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  $\theta = 30^\circ$  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  $2 \times 2$  rotation matrix by an angle  $\theta$  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.*

**Solution:** The key identity that helps us unlock the proof is

$$\cos^2(\theta) + \sin^2(\theta) = 1$$

So, given that, we have

$$\begin{aligned} R^T R &= \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & -\cos(\theta)\sin(\theta) + \sin(\theta)\cos(\theta) \\ -\sin(\theta)\cos(\theta) + \cos(\theta)\sin(\theta) & \sin^2(\theta) + \cos^2(\theta) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

$$\begin{aligned} R R^T &= \begin{bmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{bmatrix} \begin{bmatrix} \cos(\theta) & \sin(\theta) \\ -\sin(\theta) & \cos(\theta) \end{bmatrix} \\ &= \begin{bmatrix} \cos^2(\theta) + \sin^2(\theta) & \cos(\theta)\sin(\theta) - \sin(\theta)\cos(\theta) \\ \sin(\theta)\cos(\theta) - \cos(\theta)\sin(\theta) & \sin^2(\theta) + \cos^2(\theta) \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I \end{aligned}$$

$R$  satisfies both conditions, so it is an orthogonal matrix.

### Problem 6: CR Decomposition (9 pts)

Read the [CR Decomposition section of Chapter 5.4](#) before starting this problem. There, we introduce the concept of a CR decomposition. As another example, let  $A$  be the matrix from [Homework 4, Problem 5](#).

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

Its CR decomposition is:

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

To understand where the numbers in  $R$  came from, read the solutions to Homework 4, linked above.

**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: Most of these can be done quickly by eyeballing the relationships between columns.*

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

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

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

**Solution:** While we *could* use the full algorithm from Homework 4 or [Chapter 4.2](#), the three matrices provided are all small enough that we can eyeball their relationships.

(i)

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

The first two columns of  $A$  are linearly independent. You might notice that the middle column is the average of the first and third columns, i.e.

$$\frac{\text{column 1} + \text{column 3}}{2} = \text{column 2}$$

meaning that

$$\text{column 3} = -\text{column 1} + 2 \cdot \text{column 2}$$

So,  $C$  only needs to contain the first two columns, and  $R$  should start with the  $2 \times 2$  identity matrix and have a third column of  $-1$  and  $2$ .

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & 2 \\ 4 & 5 \\ 7 & 8 \end{bmatrix}}_C \underbrace{\begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \end{bmatrix}}_R$$

(ii)

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

The two columns of  $A$  are linearly independent, so all we have to do is make  $R$  the  $2 \times 2$  identity matrix.

$$A = \begin{bmatrix} 3 & 5 \\ 1 & 1 \\ 2 & -4 \\ 30 & 0 \end{bmatrix} = \underbrace{\begin{bmatrix} 3 & 5 \\ 1 & 1 \\ 2 & -4 \\ 30 & 0 \end{bmatrix}}_C \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_R$$

(iii)

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

$A$ 's second column is just  $-2$  times the first column, so we shouldn't include it in  $C$ .  $A$ 's third column is not a scalar multiple of  $A$ 's first column, so we should include it in  $C$ .

The question, then, is what to do with the fourth column, since it's not immediately obvious whether it's a linear combination of columns 1 and 3. If you notice that the last

component of column 4 is  $-5$  and the last component of column 3 is  $5$ , it exposes the fact that if column 4 were a linear combination of columns 1 and 3, the coefficient on column 3 would have to be  $-1$ , since column 1's last component is  $0$ . A quick guess and check confirms that

$$2 \cdot \text{column 1} - \text{column 3} = \text{column 4}$$

This gives us one CR decomposition:

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

Another possible CR decomposition is to use  $A$ 's 2nd and 3rd columns, rather than its 1st and 3rd columns. This gives us the following CR decomposition:

$$A = \begin{bmatrix} 1 & -2 & 3 & -1 \\ -2 & 4 & 1 & -5 \\ 3 & -6 & 4 & 2 \\ 0 & 0 & 5 & -5 \end{bmatrix} = \underbrace{\begin{bmatrix} -2 & 3 \\ 4 & 1 \\ -6 & 4 \\ 0 & 5 \end{bmatrix}}_C \underbrace{\begin{bmatrix} -\frac{1}{2} & 1 & 0 & -1 \\ 0 & 0 & 1 & -1 \end{bmatrix}}_R$$

These are not the only two possible CR decompositions of  $A$ ! We could also, for instance, use  $A$ 's 1st and 4th columns, or 2nd and 4th.