Lecture Notes

for

ICMA 223: Linear Algebra A

Pakawut Jiradilok

Mahidol University International College

Salaya, Phutthamonthon, Nakhon Pathom

THAILAND

Lecture Notes Version: 0.1.2025.05.09 (May 9, 2025)

Contents

1. The First Week

1.1. **Preliminaries.** Let us recall some standard notations. Throughout this course,

- $\mathbb{Z}$ denotes the set of all integers,
- $\mathbb{Q}$ denotes the set of all rational numbers,
- $\mathbb{R}$ denotes the set of all real numbers,
- $\mathbb{C}$ denotes the set of all complex numbers, and
- $\varnothing$ and $\{\ \}$ denote the empty set.

We might also consider certain subsets of sets above such as the following.

- $\mathbb{Z}_{\geq 0}$ is the set of all *nonnegative* integers.
- $\mathbb{Z}_{>0}$ (which is the same as $\mathbb{Z}_{\geq 1}$) is the set of all *positive* integers.

The notation $\in$ means "to be an element of". If S is a set, then the notation

$$x \in S$$

means "x is an element of S". Similarly, $\notin$ means "to be not an element of". We might also write $S \ni x$ and $S \not\ni x$ (which mean the same as $x \in S$ and $x \notin S$, respectively.)

The notation $\forall$ reads "for all" (or "for every"). For example, the sentence

$$\forall x \in \mathbb{Z}_{\geq 0},\ x + 5 \in \mathbb{Z}$$

means "for all elements x of $\mathbb{Z}_{\geq 0}$, the number $x + 5$ is an element of $\mathbb{Z}$", or equivalently, "for every nonnegative integer x, the number $x + 5$ is an integer". This is a true statement.

The notation $\exists$ reads "there exists". When we write something like

$$\exists x \in S,\ P(x),$$

the comma immediately after $\exists x \in S$ reads "such that". For example, the sentence

$$\exists x \in \mathbb{Q},\ 2x \in \mathbb{Z}_{\geq 3}$$

means "there exists a rational number x such that $2x$ is an integer at least 3". This is a true statement.

Throughout this course, the notation $:=$ means "to be defined as", or "to be equal by definition". For example, the equation

$$x := 1,$$

say that x equals 1 by definition, or we define x to be 1. Thus, something like

$$x := y,$$

might indicate that y is a variable that was defined earlier, and we are currently defining a new variable x to be equal to y.

1.2. **One linear equation in one variable.** We consider the following equation

$$a \cdot x = b, \tag{1}$$

where x is a variable, and a and b are fixed constants. The main question in this subsection is "what is the set A of all real numbers $x \in \mathbb{R}$ which satisfy Equation (1)?"

Below we list some examples when a and b are real numbers ($a, b \in \mathbb{R}$), and x is a real variable.

(i) When $a = 0$ and $b = 0$, Equation (1) becomes $0 \cdot x = 0$.

(ii) When $a = 2$ and $b = 6$, Equation (1) becomes $2 \cdot x = 6$.

(iii) When $a = 0$ and $b = 1$, Equation (1) becomes $0 \cdot x = 1$.

Let us quickly analyze the above three equations. The first equation $0 \cdot x = 0$ is a true statement, for every real number x. This equation is true regardless of the value of $x \in \mathbb{R}$. In other words, we can write

$$\forall x \in \mathbb{R},\ 0 \cdot x = 0.$$

This statement is true. (Note that the statement $\exists x \in \mathbb{R}, 0 \cdot x = 0$ is also true.) For the first equation, the desired set A is $A = \mathbb{R}$, since every real number $x \in \mathbb{R}$ satisfies Equation (1).

For the second equation $2 \cdot x = 6$, as we plug in different values of $x \in \mathbb{R}$, we might obtain equations which could be true or false. For instance, if we substitute x with 1, we obtain $2 \cdot 1 = 6$, which is false. If we substitute x with 3, we obtain $2 \cdot 3 = 6$, which is true. Experience from elementary arithmetic tells us that the value 3 is the only value of $x \in \mathbb{R}$ for which the equation becomes a true statement. So the statement

$$\forall x \in \mathbb{R},\ 2x = 6$$

is false. On the other hand, the statement

$$\exists x \in \mathbb{R},\ 2x = 6$$

is true. Indeed, there exists a real number $x \in \mathbb{R}$ (namely, $x = 3 \in \mathbb{R}$) such that $2x = 6$. For the second equation, the desired set A is $A = \{3\}$, the set containing exactly one real number 3 as its element.

Finally, the third equation is $0 \cdot x = 1$. Observe that this equation is false, for every real number $x \in \mathbb{R}$. There are no real numbers $x \in \mathbb{R}$ which satisfy $0 \cdot x = 1$. For the third equation, the desired set A is $A = \varnothing$.

We have seen three different behaviors of the equation $ax = b$:

(i) it can be true for infinitely many values of x,

(ii) it can be true for exactly one value of x, or

(iii) it cannot be true for any values of x.

It turns out, as we will see in this class, that this ∞-or-1-or-0 phenomenon holds not just for one equation in one variable, but for any system of many equations in many variables.

1.3. **Linear equations in many variables.** A **linear equation** in the *variables* $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{2}$$

where $a_1, a_2, \ldots, a_n, b$ are constants. Some authors (e.g. [AR14]) might require that $a_1, a_2, \ldots, a_n$ are not simultaneously zero (but it's okay if some of them are zero as long as at least one is not zero). We are not going to worry too much about this point.

One way we can think about Equation (2) is as follows. We think of $a_1, a_2, \ldots, a_n, b$ as "known", or "fixed", or "already given to you", and think of the variables $x_1, x_2, \ldots, x_n$ as "quantities we would like to solve for". The numbers $a_1, a_2, \ldots, a_n$ are called **coefficients**. We might want to be more precisely and say that a_1 is the coefficient of x_1, a_2 is the coefficient of x_2, and so on.

Example 1.1. Consider the equation

$$3x + 4y = 5.$$

This is a linear equation in the variables x, y with coefficients $3, 4$, respectively.

One point to note is that the equation above is just the same as the equation

$$3 \cdot x + 4 \cdot y + 0 \cdot z + 0 \cdot d = 5,$$

so we could also say that it is a linear equation in the variables x, y, z, d with coefficients $3, 4, 0, 0$ as well. (We could even *permute* the variables and say that it is a linear equation in the variables y, d, z, x with coefficients $4, 0, 0, 3$.)

Let us consider Equation (2) again. We say that the equation is a **homogeneous**[1] linear equation if $b = 0$; otherwise, we say that it is **inhomogeneous**.

[1]The word "homogeneous" is pronounced "homoGEneous" with the stress on the third syllable. Its synonym "homogenous", on the other hand, is pronounced "hoMOgenous" with the stress on the second syllable.

Example 1.2. The following are some examples of homogeneous linear equations:

$$x + y = 0, \quad -2x = 0, \quad z = 0, \quad -x + 5z = 0, \quad x = 0, \quad \text{and} \quad x + y + z = 0.$$

The following are some examples of inhomogeneous linear equations:

$$x + y = 2, \quad -2x = 1, \quad z = 7, \quad -x + 5z = -1, \quad x = 3, \quad \text{and} \quad x + y + z = -5.$$

Recall that $\mathbb{R}$ denotes the set of real numbers. For each positive integer n, let us denote by $\mathbb{R}^n$ the set of all finite lists of n (not necessarily distinct) real numbers. For instance,

$$(-2, 0, 3, 3, 14, 0) \in \mathbb{R}^6,$$

because $(-2, 0, 3, 3, 14, 0)$ is a list of 6 real numbers.

Two elements

$$(x_1, x_2, \ldots, x_n) \text{ and } (y_1, y_2, \ldots, y_n)$$

of $\mathbb{R}^n$ are considered *the same* if and only if

$$x_1 = y_1, x_2 = y_2, \ldots, x_n = y_n.$$

(For each $i \in [n]$, the i^{th} entries from the left must match.)

Example 1.3. The elements

$$(0, 0) \text{ and } (2, 3)$$

of $\mathbb{R}^2$ are different. The elements

$$(0, 1, 2) \text{ and } (2, 1, 0)$$

of $\mathbb{R}^3$ are also different.

The elements

$$(\sqrt{4}, 6/2, 2) \text{ and } (2, \sqrt{9}, 1 + 1)$$

in $\mathbb{R}^3$ are the same.

In general, the elements of $\mathbb{R}^n$ are called **n-tuples of real numbers** or **real n-tuples**. When $n = 1$, we simply think of (x) as x, so a 1-tuple is just a real number. When $n = 2$, a 2-tuple is also called an **ordered pair of real numbers**. When $n = 3$, a 3-tuple is also called a **triple of real numbers**.

A **system of linear equations** or a **linear system** in the variables $x_1, x_2, \ldots, x_n$ is a collection of linear equations in the variables $x_1, x_2, \ldots, x_n$.

A **solution** to a linear system in the variables $x_1, x_2, \ldots, x_n$ is an n-tuple $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ such that for every equation in the linear system, when we plug

$$x_1 = a_1, x_2 = a_2, \ldots, x_n = a_n$$

into the equation, we obtain a true statement.

The **solution set** of a linear system in the variables $x_1, \ldots, x_n$ is the set of all solutions $(a_1, a_2, \ldots, a_n) \in \mathbb{R}^n$ to the linear system. We say that two linear systems are **equivalent** if they have the same solution set.

Example 1.4. Consider the following system of linear equations

$$\begin{cases} -y + z = 2, \\ x + 2z = 5, \end{cases}$$

with 2 linear equations in the 3 variables x, y, z.

(a) Is $(0, 0, 0)$ a solution to the linear system?

(b) Is $(5, -2, 0)$ a solution to the linear system?

(c) Is $(-1, 1, 3)$ a solution to the linear system?

(d) Is $(1, 3, 5)$ a solution to the linear system?

The answers to these questions are given in the footnote.[2]

Example 1.5. Let us consider the linear system from Example 1.4. Now let $t \in \mathbb{R}$ be an arbitrary real number. Note that the triple

$$(-2t + 5, t - 2, t)$$

is a solution to the linear system, *no matter what* the value of t is.

Hence, the linear system has infinitely many solutions.

Recall that we have seen the "∞ or 1 or 0" phenomenon from Subsection 1.2. This is in fact a general phenomenon. Let us state it as a theorem.

Theorem 1.1. *Every linear system has either*

(i) *infinitely many solutions, or*

(ii) *exactly one solution, or*

(iii) *no solutions.*

A linear system is said to be **consistent** if it has at least one solution; otherwise, it is said to be **inconsistent**.

[2](a) No. (b) Yes. (c) Yes. (d) No.

Example 1.6. Consider the following system

$$\begin{cases} 2x + y = 5, \\ x - 2y = -2, \end{cases}$$

of two linear equations in the variables x and y.

(a) Determine whether $(2, 1)$ is a solution to the linear system.

(b) Determine the set of all real numbers t for which $(t - 2, t - 3)$ is a solution to the linear system.

(c) Determine the set of all real numbers u for which $(2u, u + 1)$ is a solution to the linear system.

(a). To determine whether $(2, 1)$ is a solution to the linear system, we plug in $x = 2$ and $y = 1$ into each equation. The first equation becomes

$$2(2) + 1 = 5,$$

which is true. The second equation becomes

$$2 - 2(1) = -2,$$

which is false. A solution to the linear system has to make *both* equations hold true. In this case, $(2, 1)$ makes *only one* equation true. Therefore, $(2, 1)$ is *not* a solution to the linear system. The answer is $\boxed{\text{No}}$.

(b). Let us plug in $x = t - 2$ and $y = t - 3$ and see which value of t would make $(t - 2, t - 3)$ a solution. After the substitution, the first equation becomes

$$2(t - 2) + (t - 3) = 5,$$

which is equivalent to

$$3t = 12,$$

which implies $t = 4$. On the other hand, the second equation becomes

$$(t - 2) - 2(t - 3) = -2,$$

which is equivalent to

$$-t = -6,$$

which implies $t = 6$. We observe that for a value of t to make both equations hold true, we must have *both* $t = 4$ and $t = 6$. Since any value t cannot simultaneously be both 4 and 6, we conclude that there is no real number t for which $(t-2, t-3)$ is a solution to the linear system. Thus, the set we would like to determine is $\boxed{\varnothing}$, the empty set.

(c). In a manner similar to part (b) above, we plug in $x = 2u$ and $y = u + 1$. The first equation becomes

$$2(2u) + (u + 1) = 5,$$

which is equivalent to $u = 4/5$. The second equation becomes

$$(2u) - 2(u + 1) = -2,$$

which is equivalent to $0 = 0$. The first equation is true if and only if $u = 4/5$. The second is true for any value of u. Thus, the set we would like to determine is $\boxed{\{4/5\}}$, the set containing exactly one element $4/5$.

If we'd like, we can check if $u = 4/5$ does make $(2u, u + 1)$ a solution. When $u = 4/5$, the pair $(2u, u + 1)$ becomes $(8/5, 9/5)$. Plugging in $x = 8/5$ and $y = 9/5$ to the first equation, we find

$$2(8/5) + (9/5) = 5,$$

which is true. Plugging in $x = 8/5$ and $y = 9/5$ to the second equation, we find

$$(8/5) - 2(9/5) = -2,$$

which is true. Hence, $(8/5, 9/5)$ is indeed a solution to the linear system.

1.4. **Solving a system of linear equations by eliminating variables.** Let us consider the following linear system.

$$\begin{cases} x + y - z = 0, \\ x + 2z = 7, \\ 2x - y + 3z = 9. \end{cases}$$

This linear system can be solved by eliminating variables. From the first equation, we find

$$x = -y + z.$$

Now we can substitute x with $-y + z$ in the remaining equations. This results in

$$-y + 3z = 7,$$
$$-3y + 5z = 9.$$

We have eliminated the variable x. We can continue. From the top equation, we have

$$y = 3z - 7.$$

Substitute y with $3z-7$ in the final equation to obtain

$$-4z+12=0.$$

This gives $z=3$. Substitute back to find $y=2$ and $x=1$. We have obtained

$$(x,y,z)=(1,2,3).$$

We can verify that $(1,2,3)$ is a solution to the original system by plugging the numbers back in.

Since we start from the three equations as our assumptions on the variables x,y,z and we obtain the conclusion that (x,y,z) *must be* $(1,2,3)$, this shows that *any* solution, if it exists, must be $(1,2,3)$. On the other hand, we can verify easily that $(1,2,3)$ is indeed a solution. A logical conclusion is that

$$\text{the solution set } = \{(1,2,3)\}$$

is the set with exactly one element that is $(1,2,3)$.

In this situation, we have solved the linear system explicitly.

1.5. **Problems and Solutions.**

Problem 1. For each of the following formal mathematical sentences, do two things: (i) translate it into a "simpler" (or "less formal") version in English language, and (ii) decide whether it is true statement or a false statement, with some explanations of how you made the decision.

(1) $\forall x \in \mathbb{R},\ x \in \mathbb{R}$.

(2) $\exists y \in \mathbb{R},\ y \notin \mathbb{R}$.

(3) $\forall z \in \mathbb{R},\ z \in \mathbb{Q}$.

(4) $\exists w \in \mathbb{Q},\ 2w + 1 \in \mathbb{Z}$.

(5) $\forall (x, y) \in \mathbb{R}^2,\ 0 \cdot x = 0 \cdot y$.

(6) $\exists (x, y) \in \mathbb{R}^2,\ x + y = 0$.

Problem 2. Let $a_1, a_2, \ldots, a_6$ be real numbers. Suppose that

$$(a_1, a_2, 3, a_3 + 1) \text{ and } (2, 7, a_4 - 1, a_1 + a_2)$$

are the same 4-tuple in $\mathbb{R}^4$. Suppose also that

$$(a_1 + a_2, a_3 + a_4) \text{ and } (a_3 + a_4 + a_5, a_4 + a_5 + a_6 + 6)$$

are the same ordered pair in $\mathbb{R}^2$. How many equations in the following system

$$\begin{cases} a_1 x + y - z &= a_1 + a_2, \\ -x + a_2 y &= a_2 - a_4 + a_5, \\ a_3 y - 2z &= -a_1 + a_2 - a_6, \\ x + a_4 z &= a_2 - a_3 + a_4 + a_5 \end{cases}$$

of 4 linear equations in the 3 variables x, y, z are homogeneous? Which ones?

Problem 3. Consider the following system of linear equations

$$\begin{cases} 2x - 3y &= -1, \\ 6y + 5z &= 11, \\ 4x + 5z &= 9, \end{cases}$$

in the variables x, y, z.

Answer the following questions.

(a) Is $(1, 1, 1)$ a solution to the linear system?

(b) Is $(4, 3, 2)$ a solution to the linear system?

(c) Is $(10, 7, -31)$ a solution to the linear system?

(d) Suppose that t is a real number such that the triple $(-4, t, 5)$ is a solution to the linear system. What is the value of t?

Problem 4. Consider the following system of linear equations

$$\begin{cases} \frac{1}{2}x + \frac{1}{3}y &= -1, \\ -6x - 4y &= 12, \end{cases}$$

in the variables x, y.

Answer the following questions.

(a) Is $(0, -3)$ a solution to the linear system?

(b) Is $(-2, 0)$ a solution to the linear system?

(c) What is the size (number of elements) of the solution set of the linear system?

Problem 5. Consider the following system of linear equations

$$\begin{cases} -x + y + z = 6, \\ x - y + z = 4, \end{cases}$$

in the variables x, y, z.

(a) Determine all real numbers t such that the triple (t, t, t) is a solution to the linear system.

(b) Determine all real numbers t such that the triple $(t - 1, t, t + 1)$ is a solution to the linear system.

(c) Determine all real numbers t such that the triple $(t, t+1, 5)$ is a solution to the linear system.

Solution to Problem 1. (1)(i) "For every real number x, the number x is an element of $\mathbb{R}$." (ii) This is $\boxed{\text{true}}$, because for every real number x, the real number x is an element of the set $\mathbb{R}$ of all real numbers.

(2)(i) "There exists a real number y such that y is not an element of $\mathbb{R}$." (ii) This is $\boxed{\text{false}}$. The set $\mathbb{R}$ contains all real numbers. There cannot be a real number y for which y is not in $\mathbb{R}$.

(3)(i) "For every real number z, the number z is a rational number." Or, simply, "every real number is rational." (ii) This is $\boxed{\text{false}}$. While every rational number is a real number, it's not true that every real number is a rational number. For example, $\pi, \sqrt{2}, \sqrt{3}, \ldots$ are real numbers which are not rational.

(4)(i) "There exists a rational number w such that $2w+1$ is an integer." (ii) This is $\boxed{\text{true}}$. One such rational number is $w = 1$. Note that $w = 1$ is a rational number and $2w + 1 = 3$ is an integer.

(5)(i) "For every pair (x, y) of real numbers, the product $0 \cdot x$ equals the product $0 \cdot y$." (ii) This is $\boxed{\text{true}}$. Indeed, for every pair $(x, y) \in \mathbb{R}^2$, the product $0 \cdot x$ is equal to 0 and the product $0 \cdot y$ is also equal to 0. Thus, $0 \cdot x = 0 = 0 \cdot y$.

(6)(i) "There exists a pair (x, y) of real numbers such that $x + y = 0$." (ii) This is $\boxed{\text{true}}$. One such pair is $(x, y) = (1, -1)$. This is a pair of real numbers. Note that $x + y = 1 + (-1) = 0$. □

Solution to Problem 2. From $(a_1, a_2, 3, a_3 + 1) = (2, 7, a_4 - 1, a_1 + a_2)$, we find

$$a_1 = 2, \quad a_2 = 7, \quad a_3 = a_1 + a_2 - 1 = 8, \quad \text{and } a_4 = 3 + 1 = 4.$$

From $(a_1 + a_2, a_3 + a_4) = (a_3 + a_4 + a_5, a_4 + a_5 + a_6 + 6)$, we find

$$a_5 = a_1 + a_2 - a_3 - a_4 = 2 + 7 - 8 - 4 = -3 \quad \text{and} \quad a_6 = (a_3 + a_4) - (a_4 + a_5 + 6) = a_3 - a_5 - 6 = 5.$$

Therefore, the system of linear equations is

$$\begin{cases} 2x + y - z & = 2 + 7, \\ -x + 7y & = 7 - 4 + (-3), \\ 8y - 2z & = -2 + 7 - 5, \\ x + 4z & = 7 - 8 + 4 + (-3), \end{cases}$$

which is

$$\begin{cases} 2x + y - z & = 9, \\ -x + 7y & = 0, \\ 8y - 2z & = 0, \\ x + 4z & = 0. \end{cases}$$

We find that $\boxed{3}$ of these equations are homogeneous: $\boxed{\text{the second, the third, and the fourth}}$. □

Solution to Problem 3. For a triple to be a solution to a linear system, when we substitute (x, y, z) with the triple in the linear system, all equations have to be true simultaneously. This means if at least one equation is not true after the substitution, the triple is immediately not a solution.

(a) We plug in $(x, y, z) = (1, 1, 1)$ in the system and find that all three equations

$$\begin{cases} 2(1) - 3(1) & = -1, \\ 6(1) + 5(1) & = 11, \\ 4(1) + 5(1) & = 9, \end{cases}$$

are true. Therefore, $\boxed{\text{yes}}$, $(1, 1, 1)$ is a solution to the linear system.

(b) Let us plug in $(x, y, z) = (4, 3, 2)$ in the second equation:

$$6(3) + 5(2) = 11,$$

which says $28 = 11$. This is a false statement. Thus, $\boxed{\text{no}}$, $(4, 3, 2)$ is *not* a solution to the linear system.

(c) Let us plug in $(x, y, z) = (10, 7, -31)$ in the second equation:

$$6(7) + 5(-31) = 11,$$

which says $-113 = 11$. This is a false statement. Thus, $\boxed{\text{no}}$, $(10, 7, -31)$ is *not* a solution to the linear system.

(d) If $(-4, t, 5)$, then when we plug in $(x, y, z) = (-4, t, 5)$ to the second equation, the resulting equation has to be true. Let's plug in:

$$6 \cdot t + 5 \cdot 5 = 11,$$

which implies $t = -7/3$.

Now let's check whether $(-4, -7/3, 5)$ is a solution or not by plugging this triple into the linear system:

$$\begin{cases} 2(-4) - 3(-7/3) & = -1, \\ 6(-7/3) + 5(5) & = 11, \\ 4(-4) + 5(5) & = 9, \end{cases}$$

We find that all three equations are true. Therefore, $(-4, -7/3, 5)$ is a solution to the linear system. Hence, $t = \boxed{-\frac{7}{3}}$ is the answer. □

Solution to Problem 4.

(a) We plug $(0,-3)$ into the linear system:

$$\begin{cases} \frac{1}{2}(0)+\frac{1}{3}(-3) & =-1, \\ -6(0)-4(-3) & =12. \end{cases}$$

Both equations are true. Therefore, $\boxed{\text{yes}}$, $(0,-3)$ is a solution to the linear system.

(b) We plug $(-2,0)$ into the linear system:

$$\begin{cases} \frac{1}{2}(-2)+\frac{1}{3}(0) & =-1, \\ -6(-2)-4(0) & =12. \end{cases}$$

Both equations are true. Therefore, $\boxed{\text{yes}}$, $(-2,0)$ is a solution to the linear system.

(c) Since we have found two different solutions to the linear system, we can conclude, by the "0-or-1-or-∞ theorem", that the size of the solution set has to be $\boxed{\text{infinity}}$. $\square$

Solution to Problem 5.

(a) Plug (t,t,t) into the linear system. We obtain

$$\begin{cases} t=6, \\ t=4. \end{cases}$$

Recall that for (t,t,t) to be a solution, both equations have to hold true simultaneously. This means we need

$$t=6 \quad \text{and} \quad t=4$$

simultaneously. But this cannot happen, since $6\neq 4$. Therefore, there are $\boxed{\text{no such real numbers } t}$.

(b) Plug $(t-1,t,t+1)$ into the linear system. We obtain

$$\begin{cases} t=4, \\ t=4. \end{cases}$$

Both equations are true simultaneously when $t=4$. Let's check whether this works. When $t=4$, the triple is $(3,4,5)$. We plug this triple back into the linear system:

$$\begin{cases} -3+4+5=6, \\ 3-4+5=4. \end{cases}$$

We see that both equations are true. Hence, $t=\boxed{4}$ is the answer.

(c) Plug $(t,t+1,5)$ into the linear system. We obtain

$$\begin{cases} -t+(t+1)+5=6, \\ t-(t+1)+5=4. \end{cases}$$

Note that both equations in the linear system are true for every real number t. Hence, the set of all real numbers $t \in \mathbb{R}$ for which $(t, t+1, 5)$ is a solution to the linear system is $\mathbb{R}$, the set of $\boxed{\text{all real numbers}}$ itself. $\square$

2. The Second Week

2.1. **A quick refresher.** We recall from the last week that a *linear system* or a *system of linear equations* in the variables $x_1, x_2, \ldots, x_n$ is simply a collection of linear equations in the variables $x_1, x_2, \ldots, x_n$. We say that a linear system is *consistent* if it has either one or infinitely many solutions. On the other hand, if a linear system has zero solutions, we say that it is *inconsistent*.

2.2. **What does it mean to "solve" a linear system?** What does the verb "solve" mean when we say we would like to *solve* a system of equations? For instance, suppose we have a number of equations with variables x and y. Perhaps one interpretation of *solving* is "to find what x is and what y is so that all the equations become true simultaneously". Let's consider the following example.

Example 2.1. Let us consider the following system

$$\begin{cases} 2x + 3y = 1, & \text{(3a)} \\ 6x + 11y = 1 & \text{(3b)} \end{cases}$$

with two linear equations in the variables x and y. Suppose we would like to solve for $(x, y) \in \mathbb{R}^2$ that satisfy (3).

When we multiply both sides of (3a) with 3, we obtain

$$6x + 9y = 3. \tag{4}$$

Now we subtract (4) from (3b) to obtain

$$2y = -2,$$

which reduces to

$$y = -1.$$

Now we substitute $y = -1$ in either of the two equations in (3). This gives us

$$x = 2.$$

Therefore, $(x, y) = (2, -1)$. On the other hand, we can verify that $(2, -1)$ is indeed a solution to the linear system by substituting the values of x and y in (3a) and (3b).

In this case, the result "$(x, y) = (2, -1)$" seems satisfactory: we have obtained the values of x and y such that both equations hold simultaneously, and moreover we know from our reasoning (mathematical deduction) that there are no other solutions.

In other words, we can say the following.

$$\text{Both (3a) and (3b) hold true}$$
$$\text{if and only if}$$
$$(x, y) = (2, -1).$$

The solution set to linear system (3) is

$$\{(2, -1)\}.$$

Now let's consider another example.

Example 2.2. Consider the following system

$$\begin{cases} x - y + z = 4, & (5a) \\ 2x - 2y - z = -1, & (5b) \end{cases}$$

of two linear equations in the variables x, y, z.

Multiply both sides of (5a) with 2 to obtain

$$2x - 2y + 2z = 8. \tag{6}$$

Subtract (6) from (5b) to obtain

$$-3z = -9,$$

which reduces to

$$z = 3.$$

Now we substitute $z = 3$ back in either (5a) or (5b). We find that

$$x - y = 1.$$

Let's consider the equations we have obtained so far in solving linear system (5). It seems like the two variables x and y are always "together". We end up with the following system:

$$\begin{cases} z = 3, & (7a) \\ x - y = 1. & (7b) \end{cases}$$

But what should we do if we would like to answer the question "what is y?" Can we pinpoint one value of y like in the previous example?

It turns out we cannot. In fact, y can be *anything*.

Suppose that

$$y = \texttt{any},$$

short for `anything`. Then (7b) says that

$$x = \texttt{any} + 1.$$

From (7a) we know $z = 3$, and so we have

$$(x, y, z) = (\texttt{any} + 1, \texttt{any}, 3).$$

Indeed, we can verify easily that once we replace `any` by *any* real number in the formula above, the output becomes a triple that is a solution to the linear system.

In other words,

$$\forall t \in \mathbb{R},\ \big((t+1, t, 3) \text{ is a solution to linear system (5)}\big).$$

That means, for instance, all of

$$(1, 0, 3), (2, 1, 3), (235, 234, 3), \text{ and } (-5, -6, 3)$$

are solutions! On the other hand, from our mathematical reasoning, we know that *any* solution to linear system (5) needs to be in the form

$$(t+1, t, 3)$$

for a certain real number t.

In some sense, this is the *best* we could do: we cannot pinpoint exactly one triple that is *the* solution, but we give a formula (which depends on a *parameter* t) such that

- when we substitute the parameter t with any real number, we obtain a solution, and
- any solution must be realizable by substituting the parameter t with some number. (In other words, we obtained all the solutions.)

The solution set to linear system (5) is

$$\{(t+1, t, 3) \mid t \in \mathbb{R}\}.$$

Let's consider another example.

Example 2.3. Consider the following system

$$\begin{cases} z - x = 1, & \text{(8a)} \\ x - y = 2, & \text{(8b)} \\ y - z = 3. & \text{(8c)} \end{cases}$$

of three linear equations in the variables x, y, z.

Suppose we would like to solve for real numbers x, y, z which satisfy (8a), (8b), and (8c) simultaneously.

Let's add the three equations together. We obtain

$$0 = 6.$$

This means that if such x, y, z exist, then 0 would be equal to 6. But 0 is not equal to 6, so we deduce that no solutions exist for linear system (8). The solution set is

$$\varnothing,$$

the empty set.

Indeed, in the three examples above, we just observed a reemergence of what we called the "∞ or 1 or 0" phenomenon. Example 2.1 exhibits a linear system with exactly one solution (i.e. the solution set has size 1). Example 2.2 exhibits a linear system with infinitely many solutions (i.e. the solution set has size infinity). Example 2.3 exhibits a linear system without solutions (i.e. the solution set is the empty set).

When the solution set has size 1 or 0, it is perhaps convenient to describe the set. If the size is 0, the solution set is empty. Nothing is an element of the set.[3] If the size is 1, the solution set contains exactly one item. In our context it contains some n-tuple of real numbers, but just exactly one tuple, no more, no less. When the size is ∞, it could be a little tricky to describe the set.

Consider Example 2.2 for instance. The solution set is an infinite set. Since the linear system was considered to have three variables x, y, z, the solution set is an infinite subset of $\mathbb{R}^3$. Even though the solution set is infinite, it does not mean that it contains everything in $\mathbb{R}^3$. For example, if we haphazardly plug in

$$x = 0, y = 0, z = 0$$

in linear system (5), we find that the triple

$$(0, 0, 0) \in \mathbb{R}^3$$

is *not* a solution to the linear system. Hence, the solution set is infinite, but somehow strictly smaller than the infinite set of all triples of real numbers. So how do we describe it? How do we describe the solution set explicitly?

A convenient way to describe the infinite set is what we already did in Example 2.2: we gave a formula with a parameter that describes *all the solutions*.

A **parametric solution** to a linear system in the variables $x_1, x_2, \ldots, x_n$ is an n-tuple in which each entry is a function of variables $t_1, t_2, \ldots, t_r$ (where r is the number of *parameters*) such that

[3] This sentence means the set is empty: it is $\{\}$. It does not mean that the set is $\{\texttt{nothing}\}$ which contains something called "nothing".

- when we substitute $t_1, t_2, \ldots, t_r$ with any r real number simultaneously, then the n-tuple becomes a solution to the linear system, and
- every solution to the linear system can be obtained by substituting $t_1, t_2, \ldots, t_r$ by some r real numbers.

The variables $t_1, t_2, \ldots, t_r$ are called **parameters**.

For instance, in Example 2.2, the triple

$$(t + 1, t, 3)$$

is a parametric solution to linear system (5).

From these discussions above, we can now describe what we mean by "solving" a linear system (at least in the context of this lecture). **Solving a linear system** entails:

(i) Decide whether the solution set of the linear system has size 0, 1, or ∞, and then

(ii) proceed to do

  (a) if the size is 0, conclude that there are no solutions.

  (b) if the size is 1, determine the n-tuple that is the unique[4] solution.

  (c) if the size is ∞, give a parametric solution to the linear system. (Or, equivalently, describe the solution set parametrically.)

Example 2.4. Let us solve the following system

$$\begin{cases} x - y = 1, \\ y - 2z = 3, \end{cases}$$

of two linear equations in the variables x, y, z. Note that if x, y, z satisfy the two equations in the system, we must have that

$$y = 2z + 3 \quad \text{and } x = y + 1.$$

Suppose for a moment that we let z be equal to some *parameter* $t \in \mathbb{R}$. Then from the above, we must have

$$y = 2t + 3,$$

and also

$$x = 2t + 4.$$

This shows that (x, y, z) is of the form $(2t + 4, 2t + 3, t)$, for a real number $t \in \mathbb{R}$.

[4]In mathematics, "unique" means *only one*.

Conversely, for every choice of the real number $t \in \mathbb{R}$, if we plug in $x = 2t + 4$, $y = 2t + 3$, and $z = t$, we find that the two equations in the system become

$$\begin{cases} (2t+4) - (2t+3) = 1, \\ (2t+3) - 2t = 3, \end{cases}$$

which are both true, regardless of the value of $t \in \mathbb{R}$. This shows that the solution set of the linear system is exactly

$$\{(2t+4, 2t+3, t) \mid t \in \mathbb{R}\},$$

which is the set of all triples $(2t+4, 2t+3, t)$ where t is a real number. The parameter $t \in \mathbb{R}$ is *free* in the sense that for *any* choice of t, the triple $(2t+4, 2t+3, t)$ is a solution to the linear system.

We have described the solution set parametrically (using the *parameter* t).

Exercise 2.5. Show that the solution set to the system

$$\begin{cases} x + y + z = 5, \\ y + z = 1, \end{cases}$$

of two linear equations in the variables x, y, z is the set

$$A := \{(4, 1-t, t) \mid t \in \mathbb{R}\}.$$

A solution to this exercise is given in the footnote.[5]

2.3. **Gaussian elimination.** Gaussian elimination provides a method for solving any given linear system. Suppose we have a system

$$\begin{aligned}
&a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\
&a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\
&\cdots \\
&a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m,
\end{aligned}$$

[5] *Solution.* First, we show that any solution must be in the form $(4, 1-t, t)$ for a certain $t \in \mathbb{R}$. Indeed, if we think of z as a parameter $t \in \mathbb{R}$, the second equation gives $y = 1 - t$. Using $y = 1 - t$ and $z = t$ in the first equation, we find that $x = 4$. Hence any solution is in the form $(4, 1-t, t)$. Second, we show conversely that any triple $(4, 1-t, t)$, where t is an arbitrary real number, is a solution to the system. This can be done by plugging in directly. The first equation becomes $x4 + (1-t) + t = 5$, which is true for any value of t. The second equation becomes $(1-t) + t = 1$, which is also true for any value of t. Hence, the triple $(4, 1-t, t)$ is a solution to the system for any value of t.

Since we have established both inclusions (one that A is a subset of the claimed set, and the other that the claimed set is a subset of A), we conclude that A is equal to the claimed solution set as desired. ∎

of m linear equations in the variables $x_1, x_2, \ldots, x_n$, where $a_{11}, a_{12}, \ldots, a_{mn}, b_1, \ldots, b_m$ are constants. From the system, we construct the following array

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

of $m \times (n+1)$ numbers, with m rows and $n+1$ columns. Such an object is called a *matrix*. In particular, it is an $m \times (n+1)$-matrix, indicating that there are m rows and $n+1$ columns.

The matrix that is constructed from a system of linear equations like above is called the **augmented matrix** of the system.

Example 2.6. Consider the following system

$$\begin{cases} 2x + 5y + 3z = 7, & \text{(9a)} \\ -y - 2z = 0, & \text{(9b)} \\ x + 4z = 1. & \text{(9c)} \end{cases}$$

of 3 linear equations in the variables x, y, z. In this case, the *augmented matrix* is

$$\begin{bmatrix} 2 & 5 & 3 & 7 \\ 0 & -1 & -2 & 0 \\ 1 & 0 & 4 & 1 \end{bmatrix}.$$

Example 2.7. Let us revisit the linear system from Example 2.3:

$$\begin{cases} z - x = 1, & \text{(10a)} \\ x - y = 2, & \text{(10b)} \\ y - z = 3. & \text{(10c)} \end{cases}$$

In this case, the augmented matrix (when we consider the variables to be x, y, z in this order) is

$$\begin{bmatrix} -1 & 0 & 1 & 1 \\ 1 & -1 & 0 & 2 \\ 0 & 1 & -1 & 3 \end{bmatrix}.$$

Suppose we have an $m \times n$ matrix

$$A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

Note that we have m rows, each of which is a list of n numbers. Let $R_1, R_2, \ldots, R_m$ denote the m rows, from the top to the bottom. Each R_i (for $i = 1, 2, \ldots, m$) can be thought of as an n-tuple. That is, we write

$$R_1, R_2, \ldots, R_m \in \mathbb{R}^n.$$

Suppose R and R' are two rows. We can *add* them by adding the corresponding entries together. For example, in the matrix

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

we can write the first row as

$$R_1 = (1, 2, 0, -1) \in \mathbb{R}^4,$$

the second row as

$$R_2 = (0, -1, 3, -2) \in \mathbb{R}^4,$$

and the third row as

$$R_3 = (2, 5, 0, -2) \in \mathbb{R}^4.$$

We can add R_1 and R_2 to obtain

$$R_1 + R_2 = (1, 1, 3, -3) \in \mathbb{R}^4.$$

We can also multiply each row with a real number by multiplying that real number to each entry. For example, if we take $c = 2$, then

$$c \cdot R_1 = (2, 4, 0, -2) \in \mathbb{R}^4.$$

Gaussian elimination is a process which transforms an augmented matrix of a linear system into a certain "simplified" matrix. In Gaussian elimination, three operations are allowed:

- Pick any row R, pick any nonzero real number c, and multiply c to R. (In other words, replace R with $c \cdot R$.)

- Swap any two rows.

- Pick any two rows (say R_1 and R_2), pick any real number $c \in \mathbb{R}$, and add $c \cdot R_2$ to R_1. (In other words, replace R_1 with $R_1 + c \cdot R_2$.

These operations are called **elementary row operations** on a matrix. The goal of Gaussian elimination is to transform a given matrix into a matrix in *reduced row echelon form*.

Remark 2.1. The transformation of replacing R_2 with $R_1 - R_2$ is *not* an elementary row operation. For example, the transformation

$$\begin{bmatrix} 8 & 9 \\ 1 & 2 \end{bmatrix} \xmapsto{R_2 \mapsto R_1 - R_2} \begin{bmatrix} 8 & 9 \\ 7 & 7 \end{bmatrix}$$

is not an elementary row operation. While $R_1 \mapsto R_1 - R_2$ and $R_2 \mapsto R_2 - R_1$ are elementary row operations, $R_2 \mapsto R_1 - R_2$ is not. It is possible, however, to transform

$$\begin{bmatrix} 8 & 9 \\ 1 & 2 \end{bmatrix}$$

into

$$\begin{bmatrix} 8 & 9 \\ 7 & 7 \end{bmatrix}$$

using *two* steps of elementary row operations:

$$\begin{bmatrix} 8 & 9 \\ 1 & 2 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - R_1} \begin{bmatrix} 8 & 9 \\ -7 & -7 \end{bmatrix} \xmapsto{R_2 \mapsto (-1) \cdot R_2} \begin{bmatrix} 8 & 9 \\ 7 & 7 \end{bmatrix}.$$

Before giving the definition of matrices in reduced row echelon form, let us work with an example. Consider the matrix

$$A_1 := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 2 & 4 & 6 & 8 \\ 1 & 4 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

Let us introduce *Property P1*. We say that a matrix satisfies **Property P1** if

(P1) For any row with at least one nonzero entry, the first nonzero entry from the left in the row is 1. This 1 in the row is called the **leading** 1 of the row.

Note that the matrix A_1 above does not satisfy Property P1. However, we can perform a sequence of elementary row operations to transform A_1 into a matrix satisfying Property P1. By multiplying the second row of A_1 by $1/2$, we obtain

$$A_2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{bmatrix}.$$

The new matrix A_2 now satisfies (P1).

Next, we introduce *Property P2*. We say that a matrix satisfies **Property P2** if

(P2) All the rows that contain only zeroes are grouped together at the bottom of the matrix. (In other words, if a certain row has only zeroes in it, then all rows beneath it also have only zeroes in them.)

The matrix A_2 does not satisfy Property P2. Let us swap R_1 and R_4 in the matrix A_2. We obtain

$$A_3 = \begin{bmatrix} 0 & 0 & 1 & 1 \\ 1 & 2 & 3 & 4 \\ 1 & 4 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

The matrix A_3 satisfies both (P1) and (P2).

Next, we introduce *Property P3*. We say that a matrix satisfies **Property P3** if

(P3) For any two different rows which have leading 1's, the leading 1 of the higher row is *strictly to the left* of the leading 1 of the lower row. (This means the two leading 1's cannot be in the same column. The top leading 1 is in a column to the left of the column the bottom leading 1 is in.)

The matrix A_3 does not satisfy Property P3. Let us swap R_1 and R_3 in A_3. We obtain

$$A_4 = \begin{bmatrix} 1 & 4 & 0 & 1 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Still, A_4 does not satisfy (P3). Let us subtract R_1 from R_2. We obtain

$$A_5 = \begin{bmatrix} 1 & 4 & 0 & 1 \\ 0 & -2 & 3 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now there is another issue with A_5: it does not satisfy (P1)! Let us multiply R_2 of A_5 by $-1/2$. We obtain

$$A_5 = \begin{bmatrix} 1 & 4 & 0 & 1 \\ 0 & 1 & -3/2 & -3/2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Now, A_5 satisfies (P1), (P2), and (P3).

Next, we introduce *Property P4*. We say that a matrix satisfies **Property P4** if

(P4) For any column that contains a leading 1, the leading 1 is the only nonzero entry in the column. (That is, all other entries are zero.)

We add $(3/2) \cdot R_3$ to R_2 in A_5 to obtain

$$A_5 = \begin{bmatrix} 1 & 4 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Then we add $(-4) \cdot R_2$ to R_1 to obtain

$$A_6 = \begin{bmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that A_6 satisfies (P1), (P2), (P3), and (P4).

Definition 2.8 (see e.g. [AR14, § 1.2]). Let m and n be positive integers. Let A be an $m \times n$-matrix. We say that A is in **row echelon form** if it satisfies the following three properties:

(P1) For any row with at least one nonzero entry, the first nonzero entry from the left in the row is 1. This 1 in the row is called the **leading** 1 of the row.

(P2) All the rows that contain only zeroes are grouped together at the bottom of the matrix. (In other words, if a certain row has only zeroes in it, then all rows beneath it also have only zeroes in them.)

(P3) For any two different rows which have leading 1's, the leading 1 of the higher row is *strictly to the left* of the leading 1 of the lower row. (This means the two leading 1's cannot be in the same column. The top leading 1 is in a column to the left of the column the bottom leading 1 is in.)

Furthermore, if A is already in row echelon form, we say that A is in **reduced row echelon form** if it satisfies the following fourth property:

(P4) For any column that contains a leading 1, the leading 1 is the only nonzero entry in the column. (That is, all other entries are zero.)

Example 2.9. The following matrix

$$\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$$

is *not* in row echelon form. The first row violates Property (P1). In particular, the matrix is not in reduced row echelon form.

Example 2.10. The following matrix

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

is *not* in row echelon form. While Property (P1) is satisfied, Property (P2) is not. The second row is an all-zero row, but the third row has a nonzero entry. In particular, the matrix is not in reduced row echelon form.

Example 2.11. The following matrix

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

is *not* in row echelon form. While Properties (P1) and (P2) are satisfied, Property (P3) is not. The second and the third rows both have their leading 1's in the second column. In particular, the matrix is not in reduced row echelon form.

Example 2.12. The following matrix

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

is in row echelon form, but it is *not* in reduced row echelon form. The third column contains a leading 1 in the second row, but it also contains another nonzero entry in the first row.

Example 2.13. The following matrix

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

is in row echelon form, but it is also *not* in reduced row echelon form. The third column contains a leading 1 in the second row, but it also contains another nonzero entry in the first row. Even though the other nonzero entry (the third entry on the first row) is a 1, Property (P4) is still violated, since (P4) would require that there is only one nonzero entry in the column.

Example 2.14. The following matrix

$$\begin{bmatrix} 1 & 2 & 0 & 3 & 0 \\ 0 & 0 & 1 & 4 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

is in reduced row echelon form. All the four conditions (P1)–(P4) are satisfied.

Example 2.15. The following matrix

$$\begin{bmatrix} 1 & 0 & 1 & 5 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 5 \end{bmatrix}$$

is in reduced row echelon form. All the four conditions (P1)–(P4) are satisfied. Note that the 1 in the first row and the third column is *not* a leading 1, since a leading 1 has to be the leftmost nonzero entry in a row. Similarly, the 1 in the first row and the sixth column is not a leading 1 either.

Example 2.16. The following matrix

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

is also a matrix in reduced row echelon form. All the four conditions (P1)–(P4) are satisfied. Note that the two 1's in the rightmost column are not leading 1's.

What's the point of matrices being in reduced row echelon form? We have the following theorem.

Theorem 2.1. *Suppose that we start with an $m \times n$ matrix A. Suppose that, after performing a sequence of elementary row operations with A, we obtain a matrix B. Then the linear systems associated with A and B have the same solution set.*

Therefore, suppose we would like solve a linear system. We can begin with writing the augmented matrix for the system. Call it A. Transform A using a sequence of elementary row operations to obtain a matrix B in reduced row echelon form. Transform B back to a system of linear equations. Then the original system and the new system have the same solution set. Note that linear systems associated with matrices in reduced row echelon form are easy to solve parametrically (or find the unique solution in the case the solution set has size 1, or declare the solutions are nonexistent in the case the solution set is empty). By solving the easy system, we solve the original system as well.

Example 2.17. Suppose we would like to solve the following system

$$\begin{cases} y - z = -1, & \text{(11a)} \\ x + 2y = 6, & \text{(11b)} \\ x - y + z = 3, & \text{(11c)} \end{cases}$$

of 3 linear equations in the variables x, y, z. We write the augmented matrix:

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

After a sequence of elementary row operations, we obtain

$$B := \begin{bmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{bmatrix}.$$

The linear system associated to B is

$$\begin{cases} x = 2, & \text{(12a)} \\ y = 2, & \text{(12b)} \\ z = 3. & \text{(12c)} \end{cases}$$

But the solution set is evident! The solution set of the system associated to B is

$$\{(2, 2, 3)\}.$$

By Theorem 2.1, the solution set of the original system is also

$$\{(2, 2, 3)\}.$$

Example 2.18. Suppose we would like to solve the following system

$$\begin{cases} x + y = 5, & \text{(13a)} \\ y + z = 6, & \text{(13b)} \\ x - z = 7. & \text{(13c)} \end{cases}$$

The augmented matrix is

$$A := \begin{bmatrix} 1 & 1 & 0 & 5 \\ 0 & 1 & 1 & 6 \\ 1 & 0 & -1 & 7 \end{bmatrix}.$$

After a sequence of elementary row operations, we obtain

$$B := \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

which is the augmented matrix for the system

$$\begin{cases} x - z = 0, & \text{(14a)} \\ y + z = 0, & \text{(14b)} \\ 0 = 1. & \text{(14c)} \end{cases}$$

The solution set is empty. Therefore, the original linear system (13) has no solutions.

Example 2.19. Let us solve the following system

$$\begin{cases} x_1 - x_2 + x_4 + 2x_5 = 1, & \text{(15a)} \\ x_1 - x_2 + x_3 + x_4 + 5x_5 = 0, & \text{(15b)} \\ 2x_3 + 6x_5 = -2, & \text{(15c)} \end{cases}$$

of 3 linear equations in the variables x_1, x_2, x_3, x_4, x_5. The augmented matrix is

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

After a sequence of elementary row operations, we obtain

$$B := \begin{bmatrix} 1 & -1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 3 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

which is in reduced row echelon form. The system associated to B is

$$\begin{cases} x_1 - x_2 + x_4 + 2x_5 = 1, & \text{(16a)} \\ x_3 + 3x_5 = -1. & \text{(16b)} \\ 0 = 0. & \text{(16c)} \end{cases}$$

Note that for any choice of x_2, x_4, x_5, we can find x_1 and x_3 (as functions of x_2, x_4, x_5) so that the three equations in the system hold. Thus, we obtain a parametric solution

$$(t - u - 2v + 1, t, -3v - 1, u, v).$$

It can be easily checked that for any choice of $t, u, v \in \mathbb{R}$, the 5-tuple above is indeed a solution to the system. Therefore, the solution set is

$$\{(t - u - 2v + 1, t, -3v - 1, u, v) \mid t, u, v \in \mathbb{R}\}.$$

Let us conclude with a general formula. In Gaussian elimination, we do the following steps.

(i) Start with a linear system. Write down the augmented matrix for the system.

(ii) Use elementary row operations to successively transform the matrix into a matrix in reduced row echelon form.

(iii) Now there are three cases.

(0) If there is a row of the form

$$0 \quad 0 \quad \cdots \quad 0 \quad 0 \quad 1$$

(or equivalently, if the rightmost column has a leading 1), then the linear system is inconsistent: there are no solutions.

(1) If the matrix in reduced row echelon form becomes a matrix of the form

$$\begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & a_1 \\ 0 & 1 & 0 & \cdots & 0 & a_2 \\ 0 & 0 & 1 & \cdots & 0 & a_3 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_n \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

(where there can be any number, including 0, of all-zero rows), then there is only one solution:

$$(a_1, a_2, \ldots, a_n)$$

is the unique solution to the linear system.

(∞) If the rightmost column does not have a leading 1 and there is another column (that is not the rightmost column) without a leading 1, then for the variables corresponding to columns without leading 1's, we make them parameters, and write the variables corresponding to columns with leading 1's as functions of the parameters. The result is a parametric solution to the linear system.

This completes the algorithm to solve any linear system.

Let us provide more examples.

Example 2.20. Suppose that after a sequence of elementary row operations, we obtain the following matrix

$$\begin{bmatrix} 1 & 0 & -2 & 5 \\ 0 & 1 & 1 & 7 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

for a linear system in the variables x, y, z. Then because the columns corresponding to x and y have leading 1's in them, and the column corresponding to z does not, we make $z = t$ a parameter. The solution set of the linear system is

$$\{(2t+5, -t+7, t) \mid t \in \mathbb{R}\}.$$

Example 2.21. Suppose that after a sequence of elementary row operations, we obtain the following matrix

$$\begin{bmatrix} 1 & 0 & -4 & 0 & 0 & 5 & 0 \\ 0 & 1 & 4 & 0 & 0 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & -5 & 3 \\ 0 & 0 & 0 & 0 & 1 & 0 & 7 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

for a linear system in the variables $x_1, x_2, x_3, x_4, x_5, x_6$. The columns corresponding to x_3 and x_6 do not have leading 1's in them, so we make x_3 and x_6 *free* parameters t_3 and t_6, respectively. Now we find

$$\begin{aligned} x_1 &= 4t_3 - 5t_6, & x_2 &= -4t_3 + 2, & x_3 &= t_3, \\ x_4 &= 5t_6 + 3, & x_5 &= 7, & x_6 &= t_6. \end{aligned}$$

Therefore, the solution set of the corresponding linear system can be described parametrically as

$$\{(4t_3 - 5t_6, -4t_3 + 2, t_3, 5t_6 + 3, 7, t_6) \mid t_3, t_6 \in \mathbb{R}\}.$$

2.4. **Problems and Solutions.**

Problem 1. For each of the following matrices, determine (i) whether it is in row echelon form, and (ii) whether it is in reduced row echelon form. Explain your reasoning.

(a)

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

(b)

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

(c)

$$\begin{bmatrix} 1 & 0 & 2 & 0 & 4 & 0 \\ 0 & 1 & 0 & 3 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 2. For each of the following matrices, determine (i) whether it is in row echelon form, and (ii) whether it is in reduced row echelon form. Explain your reasoning.

(a)

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

(b)

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

(c)

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

Problem 3. For each of the following matrices, determine (i) whether it is in row echelon form, and (ii) whether it is in reduced row echelon form. Explain your reasoning.

(a)
$$\begin{bmatrix} 1 & 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

(b)
$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

(c)
$$\begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

Problem 4. For each of the following matrices, determine (i) whether it is in row echelon form, and (ii) whether it is in reduced row echelon form. Explain your reasoning.

(a)
$$\begin{bmatrix} 1 & 2 & 3 \\ 0 & 1 & 4 \\ 0 & 1 & 5 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(b)
$$\begin{bmatrix} 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix}$$

(c)
$$\begin{bmatrix} 1 & -1 & 0 & 0 & 2 \\ 0 & 0 & 1 & -1 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 5. For each of the following matrices, determine (i) whether it is in row echelon form, and (ii) whether it is in reduced row echelon form. Explain your reasoning.

(a)
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

(b)
$$\begin{bmatrix} 1 & 2 & 3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 & 3 \end{bmatrix}$$

(c)
$$\begin{bmatrix} 1 & 2 & 1 & 2 & 1 \\ 0 & 0 & 1 & 2 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Problem 6. Consider the following system

$$\begin{cases} x + 3y + 4z = 5, & \text{(17a)} \\ 2x - y = 1, & \text{(17b)} \\ 3x + y + 2z = 3, & \text{(17c)} \end{cases}$$

of 3 linear equations in the variables x, y, z.

(a) Write the augmented matrix for the linear system.

(b) Perform a sequence of elementary row operations to transform the augmented matrix into a matrix in reduced row echelon form.

(c) Solve the linear system.

Problem 7. Consider the following system

$$\begin{cases} 2x - y - z = 1, & \text{(18a)} \\ x + 2y - 3z = 1, & \text{(18b)} \\ -x + z = 1, & \text{(18c)} \end{cases}$$

of 3 linear equations in the variables x, y, z.

(a) Write the augmented matrix for the linear system.

(b) Perform a sequence of elementary row operations to transform the augmented matrix into a matrix in reduced row echelon form.

(c) Solve the linear system.

Problem 8. Consider the following system

$$\begin{cases} w + 2x + 3y - z = 7, & \text{(19a)} \\ 2w - 3x - y - 2z = 0, & \text{(19b)} \\ w + y - z = 3, & \text{(19c)} \\ -w + 3x + 2y + z = 3, & \text{(19d)} \end{cases}$$

of 4 linear equations in the variables w, x, y, z.

(a) Write the augmented matrix for the linear system.

(b) Perform a sequence of elementary row operations to transform the augmented matrix into a matrix in reduced row echelon form.

(c) Solve the linear system.

Problem 9. Consider the following system

$$\begin{cases} x_1 + x_2 + x_3 + x_4 + x_5 + x_6 = 1, & \text{(20a)} \\ x_3 + x_5 + x_6 = 3, & \text{(20b)} \\ x_4 + x_6 = 5, & \text{(20c)} \end{cases}$$

of 3 linear equations in the variables $x_1, x_2, x_3, x_4, x_5, x_6$.

(a) Write the augmented matrix for the linear system.

(b) Perform a sequence of elementary row operations to transform the augmented matrix into a matrix in reduced row echelon form.

(c) Solve the linear system.

Solution to Problem 1. *Answers.* (a) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

(b) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

(c) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

Explanation. (a) The second row has a nonzero entry, and the first nonzero entry is not 1. So it's not in row echelon form. In particular, it's not in reduced row echelon form.

(b) All four conditions are satisfied.

(c) All four conditions are satisfied. ☐

Solution to Problem 2. *Answers.* (a) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

(b) (i) Yes, it's in row echelon form. (ii) No, it's not in reduced row echelon form.

(c) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

Solution. (a) The second row is an all-zero row, but it is above the third row, which is not an all-zero row. So it's not in row echelon form. In particular, it's not in reduced row echelon form.

(b) All the first three conditions are satisfied, so it's in row echelon form. However, the leading 1 in the fifth column is not the only nonzero entry in the column, so it's not in reduced row echelon form.

(c) All four conditions are satisfied. ☐

Solution to Problem 3. *Answers.* (a) (i) Yes, it's in row echelon form. (ii) No, it's not in reduced row echelon form.

(b) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

(c) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

Solution. (a) All the first three conditions are satisfied, so it's in row echelon form. However, the leading 1 in the second column is not the only nonzero entry in the column, so it's not in reduced row echelon form.

(b) The third row has a nonzero entry, and the first nonzero entry is 2, not 1. So it's not in row echelon form. In particular, it's not in reduced row echelon form.

(c) The first row is an all-zero row, but it is above the second row, which is not an all-zero row. So it's not in row echelon form. In particular, it's not in reduced row echelon form. ☐

Solution to Problem 4. *Answers.* (a) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

(b) (i) No, it's not in row echelon form. (ii) No, it's not in reduced row echelon form.

(c) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

Solution. (a) The leading 1's in the second and the third rows are in the same column. So it's not in row echelon form. In particular, it's not in reduced row echelon form.

(b) The leading 1 in the first row is to the right of the leading 1 in the second row. So it's not in row echelon form. In particular, it's not in reduced row echelon form.

(c) All four conditions are satisfied. □

Solution to Problem 5. *Answers.* (a) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

(b) (i) Yes, it's in row echelon form. (ii) Yes, it's in reduced row echelon form.

(c) (i) Yes, it's in row echelon form. (ii) No, it's not in reduced row echelon form.

Solution. (a) All four conditions are satisfied.

(b) All four conditions are satisfied.

(c) All the first three conditions are satisfied, so it's in row echelon form. However, the leading 1 in the third column is not the only nonzero entry in the column, so it's not in reduced row echelon form. □

Solution to Problem 6. (a)

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

(b)

$$\begin{bmatrix} 1 & 3 & 4 & 5 \\ 2 & -1 & 0 & 1 \\ 3 & 1 & 2 & 3 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - 2 \cdot R_1} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & -7 & -8 & -9 \\ 3 & 1 & 2 & 3 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 - 3 \cdot R_1} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & -7 & -8 & -9 \\ 0 & -8 & -10 & -12 \end{bmatrix}$$

$$\xmapsto{R_2 \mapsto R_2 - R_3} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & -8 & -10 & -12 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 + 8 \cdot R_2} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 6 & 12 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto (1/6)\cdot R_3} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\xmapsto{R_2 \mapsto R_2 - 2\cdot R_3} \begin{bmatrix} 1 & 3 & 4 & 5 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\xmapsto{R_1 \mapsto R_1 - 4\cdot R_3} \begin{bmatrix} 1 & 3 & 0 & -3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$$

$$\xmapsto{R_1 \mapsto R_1 - 3\cdot R_2} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 2 \end{bmatrix}.$$

(c) The solution set is $\{(0, -1, 2)\}$. □

Solution to Problem 7. (a)

$$\begin{bmatrix} 2 & -1 & -1 & 1 \\ 1 & 2 & -3 & 1 \\ -1 & 0 & 1 & 1 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 2 & -1 & -1 & 1 \\ 1 & 2 & -3 & 1 \\ -1 & 0 & 1 & 1 \end{bmatrix} \xmapsto{R_1 \rightleftharpoons R_2} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 2 & -1 & -1 & 1 \\ -1 & 0 & 1 & 1 \end{bmatrix}$$

$$\xmapsto{R_2 \mapsto R_2 - 2\cdot R_1} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & -5 & 5 & -1 \\ -1 & 0 & 1 & 1 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 + R_1} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & -5 & 5 & -1 \\ 0 & 2 & -2 & 2 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto (1/2)\cdot R_3} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & -5 & 5 & -1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

$$\xmapsto{R_2 \mapsto R_2 + 5\cdot R_3} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

$$\xmapsto{R_2 \rightleftharpoons R_3} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 4 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto (1/4) \cdot R_3} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xmapsto{R_2 \mapsto R_2 - R_3} \begin{bmatrix} 1 & 2 & -3 & 1 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xmapsto{R_1 \mapsto R_1 - R_3} \begin{bmatrix} 1 & 2 & -3 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xmapsto{R_1 \mapsto R_1 - 2 \cdot R_2} \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(c) The solution set is $\varnothing$. □

Solution to Problem 8. (a)

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

(b)

$$\begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 2 & -3 & -1 & -2 & 0 \\ 1 & 0 & 1 & -1 & 3 \\ -1 & 3 & 2 & 1 & 3 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - 2 \cdot R_1} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & -7 & -7 & 0 & -14 \\ 1 & 0 & 1 & -1 & 3 \\ -1 & 3 & 2 & 1 & 3 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 - R_1} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & -7 & -7 & 0 & -14 \\ 0 & -2 & -2 & 0 & -4 \\ -1 & 3 & 2 & 1 & 3 \end{bmatrix}$$

$$\xrightarrow{R_4 \mapsto R_4 + R_1} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & -7 & -7 & 0 & -14 \\ 0 & -2 & -2 & 0 & -4 \\ 0 & 5 & 5 & 0 & 10 \end{bmatrix}$$

$$\xrightarrow{R_4 \mapsto R_4 + (5/7) \cdot R_2} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & -7 & -7 & 0 & -14 \\ 0 & -2 & -2 & 0 & -4 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_3 \mapsto R_3 - (2/7) \cdot R_2} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & -7 & -7 & 0 & -14 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_2 \mapsto (-1/7) \cdot R_2} \begin{bmatrix} 1 & 2 & 3 & -1 & 7 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$\xrightarrow{R_1 \mapsto R_1 - 2 \cdot R_2} \begin{bmatrix} 1 & 0 & 1 & -1 & 3 \\ 0 & 1 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

(c) The solution set is

$$\{(-t+u+3, -t+2, t, u) \,|\, t, u \in \mathbb{R}\}.$$

□

Solution to Problem 9. (a)

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 & 5 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 & 5 \end{bmatrix} \xrightarrow{R_1 \mapsto R_1 - R_3} \begin{bmatrix} 1 & 1 & 1 & 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 & 5 \end{bmatrix}$$

$$\xrightarrow{R_1 \mapsto R_1 - R_2} \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & -1 & -7 \\ 0 & 0 & 1 & 0 & 1 & 1 & 3 \\ 0 & 0 & 0 & 1 & 0 & 1 & 5 \end{bmatrix}.$$

(c) The solution set is

$$\{(-t+v-7, t, -u-v+3, -v+5, u, v) \mid t, u, v \in \mathbb{R}\}.$$

$\square$

3. The Third Week

3.1. **Quick review of solving linear systems.** Let us recall that in the previous lecture we learned how to solve linear systems. The general outline can be described as follows.

(i) Identify the variables (say $x_1, x_2, \ldots, x_n$) in the linear system.

(ii) Write the augmented matrix A. (If there are n variables, then the number of columns of A is $n+1$. The number of rows of A is equal to the number of equations in the system.

(iii) Perform Gaussian elimination: use a sequence of elementary row operations to transform A into a matrix, say M, in reduced row echelon form.

(iv) Check whether M has a leading 1 in the rightmost column. If so, the solution set is $\varnothing$ and we finish solving the system. If not, proceed to the next step.

(v) For each column that is (i) not the rightmost column and (ii) does not have a leading 1, we introduce a "free" variable.

(vi) Write the solution set. Finish solving the system.

3.2. **Rank and nullity.** Suppose we start with a matrix A with m rows and n columns. We can perform Gaussian elimination on A to transform A into a matrix, say M, in reduced row echelon form. This can be done even without the discussion of the underlying linear system which corresponds to A.

Now in this context, the number n, which is the number columns of A, can be called the **dimension of the domain** of A. The number of leading 1's in M is called the **rank** of A, which is also called the **dimension of the image** of A, and is denoted $\operatorname{rank}(A)$. The number of columns without leading 1's in M is called the **nullity** of A, which is also called the **dimension of the kernel** of A, and is denoted $\operatorname{nullity}(A)$.

From this discussion, we obtain a version of the "rank–nullity theorem":

$$\operatorname{rank}(A) + \operatorname{nullity}(A) = n, \tag{$*$}$$

where n is the number of columns of A.

Let us remark that the way we introduce this version of the rank–nullity theorem makes it seem *trivial*: indeed Equation $(*)$ seems to follow immediately from our definitions! Our remark is that the actual rank–nullity theorem is not as trivial as how we describe it here. Our presentation hides many nontrivial steps along the way.

3.3. **Matrices.** We have already seen one way in which matrices occur. When a system of linear equations (together with an ordered list of variables, such as "x, y, z") is given, we can construct the *augmented matrix* corresponding to the linear system. If the system has m linear equations in n variables, then the resulting augmented matrix has m rows and $n+1$ columns. The n columns from the left of the augmented matrix correspond to the n variables (in the same order: if the ordered list of variables is x, y, z, for instance, then the leftmost column corresponds to x, the second-leftmost, to y, and the third-leftmost, to z). The last column—the rightmost one—corresponds to the "right-hand sides" of the linear equations.

Now let us study matrices as algebraic objects in their own right. (That is, we can work with matrices, even when we do not have to consider the linear systems corresponding to them.)

Definition 3.1. Let m and n be positive integers. An **$m \times n$-matrix** A is an array of $m \times n$ numbers with m rows and n columns, written with an open bracket to the left of the first column and a closing bracket to the right of the last column. In this case, we say that the **size** (or the **dimension**) of A is $m \times n$.

The set of all $m \times n$-matrices with real number entries is denoted $\mathbb{R}^{m\times n}$.

Example 3.2. Here is a 2×3 matrix:

$$\begin{bmatrix} 1 & 2 & 0 \\ -1 & 0 & 4 \end{bmatrix} \in \mathbb{R}^{2\times 3}.$$

Here is a 3×2 matrix:

$$\begin{bmatrix} 3 & -1 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{R}^{3\times 2}.$$

3.3.1. *Extracting rows and columns from a matrix.* There is a convenient way to extract rows and columns from a matrix. Suppose we have the following matrix.

$$A := \begin{bmatrix} 5 & 2 & 0 & -1 \\ -2 & 7 & 0 & 0 \\ 0 & 1 & 0 & 8 \end{bmatrix}$$

The matrix A has dimension 3×4: it has 3 rows and 4 columns. We can write the following equation:

$$A = \begin{bmatrix} \text{——} & R_1 & \text{——} \\ \text{——} & R_2 & \text{——} \\ \text{——} & R_3 & \text{——} \end{bmatrix}. \tag{21}$$

The above equation says that A has 3 rows, named R_1, R_2, R_3, from the top to the bottom. We don't actually need the long dashes to the left and to the right of each R_i when we write the matrix on the right-hand side: we can simply write

$$A = \begin{bmatrix} R_1 \\ R_2 \\ R_3 \end{bmatrix}. \tag{22}$$

However, when we include the dashes, they help us remember that R_1, R_2, R_3 are *rows*, not just *numbers*. The matrix A is still a 3×4-matrix, not a 3×1-matrix as Equation (22) might mislead us.

Back to our example, if we write Equation (21) or (22), then we have *extracted* the rows from the matrix. The implication is that

$$R_1 = \begin{bmatrix} 5 & 2 & 0 & -1 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} -2 & 7 & 0 & 0 \end{bmatrix},$$

and

$$R_3 = \begin{bmatrix} 0 & 1 & 0 & 8 \end{bmatrix}.$$

Each of R_1, R_2, and R_3 is a 1×4 matrix.

Similarly, we can extract the columns. If we write

$$A = \begin{bmatrix} | & | & | & | \\ C_1 & C_2 & C_3 & C_4 \\ | & | & | & | \end{bmatrix},$$

then we have

$$C_1 = \begin{bmatrix} 5 \\ -2 \\ 0 \end{bmatrix},$$

$$C_2 = \begin{bmatrix} 2 \\ 7 \\ 1 \end{bmatrix},$$

$$C_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix},$$

and

$$C_4 = \begin{bmatrix} -1 \\ 0 \\ 8 \end{bmatrix}.$$

Each of C_1, C_2, C_3, and C_4 is a matrix of dimension 3×1.

3.3.2. *Some definitions.* Now let us give some definitions. Suppose that A is an $m \times n$-matrix. For each $i \in \{1, 2, \ldots, m\}$, the i^{th} **row** of A is the $1 \times n$-matrix that is the i^{th} row from the top of A. Similarly, for each $j \in \{1, 2, \ldots, n\}$, the j^{th} **column** of A is the $m \times 1$-matrix that is the j^{th} column from the left of A.

Definition 3.3. Let A be an $m \times n$ matrix.

(a) We say that A is a **row vector** if $m = 1$.

(b) We say that A is a **column vector** if $n = 1$.

(c) We say that A is a **square matrix** if $m = n$.

The **(i, j)-entry** of the matrix A is the number that is in the i^{th} row and in the j^{th} column of A.

Example 3.4. Consider

$$A = \begin{bmatrix} 2 & 3 \\ 5 & 7 \end{bmatrix}.$$

Then A is a square matrix of dimension 2×2. The $(1, 1)$-entry of A is 2. The $(1, 2)$-entry of A is 3. The $(2, 1)$-entry of A is 5. The $(2, 2)$-entry of A is 7.

A matrix is said to be a **zero matrix** if all its entries are zero.

Definition 3.5. Let A be a square $n \times n$ matrix. Then the **diagonal entries** of A are the $(1, 1)$, $(2, 2)$, $\ldots$, (n, n)-entries.

Example 3.6. The diagonal entries of the matrix

$$\begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & -1 & 0 \\ 0 & 4 & -2 & 1 \\ -4 & 6 & 0 & 0 \end{bmatrix}$$

are $1, 1, -2, 0$.

Definition 3.7. Let A and B be two matrices. We say that A and B are **equal** (or that they are the **same** matrix) if they have the same number of rows (say m), they have the same number of columns (say n), and for every $i \in \{1, 2, \ldots, m\}$, and for every $j \in \{1, 2, \ldots, n\}$, they have the same (i, j)-entry.

In other words, if A is an $m \times n$-matrix and B is an $m' \times n'$-matrix, then $A = B$ if and only if the following three conditions hold:

(i) $m = m'$,

(ii) $n = n'$, and

(iii) $\forall i \in \{1, 2, \ldots, m\}, \forall j \in \{1, 2, \ldots, n\}$, $a_{i,j} = b_{i,j}$,

where $a_{i,j}$ and $b_{i,j}$ denote the (i, j)-entries of A and B, respectively.

3.4. **Matrix operations.**

3.4.1. *Addition, subtraction, and scalar multiplication.* We can add and subtract matrices when their dimensions match. We do so by adding and subtracting their corresponding entries.

Example 3.8. Let

$$A := \begin{bmatrix} -1 & 0 & 0 \\ 1 & 2 & 0 \end{bmatrix} \qquad \text{and} \qquad B := \begin{bmatrix} 0 & -4 & 5 \\ 12 & 8 & 4 \end{bmatrix}.$$

Then

$$A + B = \begin{bmatrix} -1 & -4 & 5 \\ 13 & 10 & 4 \end{bmatrix} \qquad \text{and} \qquad A - B = \begin{bmatrix} -1 & 4 & -5 \\ -11 & -6 & -4 \end{bmatrix}.$$

We can also multiply a matrix by a scalar.[6] This is done by multiplying the scalar to every entry of the matrix.

Example 3.9. Let

$$A := \begin{bmatrix} -4 & 0 \\ 2 & 1 \end{bmatrix}.$$

Then

$$-2 \cdot A = \begin{bmatrix} 8 & 0 \\ -4 & -2 \end{bmatrix} \qquad \text{and} \qquad \frac{1}{2} \cdot A = \begin{bmatrix} -2 & 0 \\ 1 & 0.5 \end{bmatrix}.$$

3.4.2. *Matrix multiplication.* Here comes an interesting operation. Let us start with a simpler case. Suppose we have a $1 \times m$-matrix

$$R = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix},$$

and an $m \times 1$-matrix

$$C = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

Note that R and C have the same number of entries. Then we define the result of the multiplication

$$R \cdot C$$

[6] Here, a *scalar* just means a number, not a vector and not a matrix.

to be the 1×1-matrix

$$R \cdot C := \begin{bmatrix} a_1 \cdot b_1 + a_2 \cdot b_2 + \cdots + a_m \cdot b_m \end{bmatrix}.$$

Example 3.10. Consider

$$A := \begin{bmatrix} 4 & -4 & 2 \end{bmatrix} \qquad \text{and} \qquad B := \begin{bmatrix} 0 \\ 1 \\ -4 \end{bmatrix}.$$

Then since

$$4 \cdot 0 + (-4) \cdot 1 + 2 \cdot (-4) = -12,$$

we have

$$A \cdot B = \begin{bmatrix} -12 \end{bmatrix}.$$

More generally, suppose that $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. When the number of *columns* of A matches the number of *rows* of B, we can define

$$A \cdot B.$$

Let us write

$$A = \begin{bmatrix} \text{——} & R_1 & \text{——} \\ \text{——} & R_2 & \text{——} \\ & \vdots & \\ \text{——} & R_m & \text{——} \end{bmatrix},$$

and

$$B = \begin{bmatrix} | & | & & | \\ C_1 & C_2 & \cdots & C_p \\ | & | & & | \end{bmatrix}.$$

Note that each R_i is a $1 \times n$-matrix and each C_j is an $n \times 1$-matrix, so we can consider $R_i \cdot C_j$ as we did above. We define the result of the **matrix multiplication** (or the **product**) $A \cdot B$ to be the $m \times p$-matrix in which the (i, j)-entry is the unique entry in the 1×1-matrix $R_i \cdot C_j$.

Example 3.11. Let

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

and

$$B := \begin{bmatrix} 1 & 5 \\ 0 & -1 \end{bmatrix}.$$

Then if we write

$$A = \begin{bmatrix} \text{——} & R_1 & \text{——} \\ \text{——} & R_2 & \text{——} \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} | & | \\ C_1 & C_2 \\ | & | \end{bmatrix},$$

we find that

$$R_1 = \begin{bmatrix} -1 & 0 \end{bmatrix},$$

$$R_2 = \begin{bmatrix} 1 & 1 \end{bmatrix},$$

$$C_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix},$$

and

$$C_2 = \begin{bmatrix} 5 \\ -1 \end{bmatrix}.$$

Now from the rule above, we compute

$$R_1 \cdot C_1 = \begin{bmatrix} -1 \end{bmatrix},$$

$$R_1 \cdot C_2 = \begin{bmatrix} -5 \end{bmatrix},$$

$$R_2 \cdot C_1 = \begin{bmatrix} 1 \end{bmatrix},$$

and

$$R_1 \cdot C_1 = \begin{bmatrix} 4 \end{bmatrix}.$$

Therefore,

$$A \cdot B = \begin{bmatrix} -1 & -5 \\ 1 & 4 \end{bmatrix}.$$

Exercise 3.12. Let

$$A := \begin{bmatrix} 0 & -1 & -1 \\ 1 & 2 & 0 \end{bmatrix} \qquad \text{and} \qquad B := \begin{bmatrix} 1 & 0 \\ 1 & 2 \\ -2 & 0 \end{bmatrix}.$$

(a) Compute $A \cdot B$.

(b) Compute $B \cdot A$.

(c) Are $A \cdot B$ and $B \cdot A$ equal?

A solution to this exercise is given in the footnote.[7]

---

[7]**(a)**

$$A \cdot B = \begin{bmatrix} 1 & -2 \\ 3 & 4 \end{bmatrix}.$$

3.5. **Linear systems.** Let us revisit linear systems. Suppose we have the following system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2, \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m, \end{cases} \tag{23}$$

of m linear equations[8] in the n variables $x_1, x_2, \dots, x_n$. Previously, we have considered the *augmented matrix* of this system, which is the $m \times (n+1)$-matrix given by

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}.$$

Now for the same system and the same ordered list of variables $(x_1, x_2, \dots, x_n)$, we define the **coefficient matrix** (or the **matrix of coefficients**) of the system as the $m \times n$-matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Observation 3.1. The coefficient matrix is simply the augmented matrix with the rightmost column removed.

Now consider the linear system (23) again. Let A denote the coefficient matrix of the system, let us denote by

$$\mathbf{x} := \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

(b)

$$B \cdot A = \begin{bmatrix} 0 & -1 & -1 \\ 2 & 3 & -1 \\ 0 & 2 & 2 \end{bmatrix}.$$

(c) No, they are not equal. Their dimensions do not match.

[8]The subscript 11 in a_{11}, for example, is a shorthand for $(1,1)$. It is not the number eleven (11). So a_{11} should actually be $a_{1,1}$ or $a_{(1,1)}$, but writing a_{ij} instead of $a_{i,j}$ (or $a_{(i,j)}$) is a common practice. Similarly, the $2n$ and the mn in the subscripts refer to $(2,n)$ and (m,n), not the products $2 \cdot n$ and $m \cdot n$.

the column vector of the variables, and let us denote by

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix}$$

the column vector of the right-hand sides of (23).

Note that the linear system (23) can now be written succinctly as

$$A \cdot \mathbf{x} = \mathbf{b}.$$

This one simple-looking equation contains the whole linear system.

For example, we can also discuss *consistency* of a linear system in this new language. For the above linear system with A, $\mathbf{x}$, and $\mathbf{b}$ as defined above, the linear system is *consistent* if and only if

$$\exists \mathbf{x} \in \mathbb{R}^{n\times 1},\ A \cdot \mathbf{x} = \mathbf{b}.$$

3.6. **Transpose and trace.** Let us give a little more definitions.

Definition 3.13. Let $A \in \mathbb{R}^{m\times n}$. The **transpose** of A, denoted A^{T}, is the $n \times m$-matrix whose (i, j)-entry is the (j, i)-entry of A.

Example 3.14. If

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

then the transpose of A is

$$A^{\mathsf{T}} = \begin{bmatrix} 0 & 3 \\ 1 & 4 \\ 2 & 5 \end{bmatrix}.$$

Observation 3.2. For any matrix A, we have $(A^{\mathsf{T}})^{\mathsf{T}} = A$. In other words, if we transpose a matrix twice, we obtain the original matrix back.

Definition 3.15. Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. Let a_{ij} denote the (i, j)-entry of A. Then the **trace** of A is defined as

$$\operatorname{tr}(A) := a_{11} + a_{22} + \cdots + a_{nn}.$$

In other words, the trace is the sum of diagonal entries.

Example 3.16. Let

$$A := \begin{bmatrix} 0 & 1 & 0 \\ -1 & -1 & -1 \\ -2 & -3 & 3 \end{bmatrix}.$$

Then the trace of A is

$$\operatorname{tr}(A) = 0 + (-1) + 3 = 2.$$

Example 3.17. Note that the *trace* and the *transpose* are different concepts! For a square matrix $A \in \mathbb{R}^{n\times n}$, the trace $\operatorname{tr}(A)$ is a real number, while the transpose A^{T} is a matrix!

For instance, if

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

then the *trace* of M is the real number

$$\operatorname{tr}(M) = 1 + 5 + 9 = 15 \in \mathbb{R},$$

while the *transpose* of M is the matrix

$$M^{\mathsf{T}} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix} \in \mathbb{R}^{3\times 3}.$$

Exercise 3.18. Let

$$N := \begin{bmatrix} -4 & -3 \\ 0 & 2 \end{bmatrix}.$$

What is $\operatorname{tr}(N) \cdot N^{\mathsf{T}}$?

A solution to this exercise is given in the footnote.[9]

---

[9]**Solution.** Note that the trace of N is $\operatorname{tr}(N) = (-4) + 2 = -2$. The transpose of N is

$$N^{\mathsf{T}} = \begin{bmatrix} -4 & 0 \\ -3 & 2 \end{bmatrix}.$$

Therefore,

$$\operatorname{tr}(N) \cdot N^{\mathsf{T}} = (-2) \cdot \begin{bmatrix} -4 & 0 \\ -3 & 2 \end{bmatrix} = \begin{bmatrix} 8 & 0 \\ 6 & -4 \end{bmatrix}$$

is the desired matrix. ■

3.7. **Problems and Solutions.**

Problem 1. Let $A \in \mathbb{R}^{3\times 3}$ be a 3×3-matrix with the following property:

for every $i \in \{1,2,3\}$ and for every $j \in \{1,2,3\}$, the (i,j)-entry of A is equal to $2 \cdot i + j$.

(a) Write A explicitly (showing all the entries of the matrix).

(b) What is A^{T}? (Recall that A^{T} denotes the *transpose* of A.)

(c) Compute $\operatorname{tr}(A)$ and $\operatorname{tr}(A^{\mathsf{T}})$. Are the two traces equal?

(d) Perform a sequence of elementary row operations on A to transform A into a matrix in reduced row echelon form.

Problem 2. Let $B \in \mathbb{R}^{3\times 4}$ be a matrix with the following property:

for every $i \in \{1,2,3\}$ and for every $j \in \{1,2,3,4\}$, the (i,j)-entry of B is equal to $i^2 \cdot j$.

(a) Write B explicitly (showing all the entries of the matrix).

(b) What is B^{T}?

(c) Perform a sequence of elementary row operations on B to transform B into a matrix in reduced row echelon form.

(d) Perform a sequence of elementary row operations on B^{T} to transform B^{T} into a matrix in reduced row echelon form.

Problem 3. Let us consider the following 1×1-matrices:

$$A := \begin{bmatrix} 3 \end{bmatrix}, \qquad B := \begin{bmatrix} -2 \end{bmatrix}, \qquad \text{and} \qquad C := \begin{bmatrix} 9 \end{bmatrix}.$$

(a) Compute $A + B$.

(b) Compute $B - 2 \cdot C$.

(c) Compute $-3 \cdot A + B + 4 \cdot C$.

(d) Compute $A \cdot B$.

(e) Compute $B \cdot C$.

(f) Compute $\operatorname{tr}(A + B) + \operatorname{tr}(C^{\mathsf{T}})$.

(g) Compute $(A + B) \cdot (B + C)$.

(h) Compute $2 \cdot A \cdot B - 3 \cdot A \cdot C + 4 \cdot B \cdot C$.

Problem 4. Let us consider the following 1×4-matrices:

$$A := \begin{bmatrix} 0 & 1 & -1 & -2 \end{bmatrix}, \qquad B := \begin{bmatrix} 2 & -1 & 3 & -5 \end{bmatrix}, \qquad \text{and} \qquad C := \begin{bmatrix} 3 & 0 & -2 & 3 \end{bmatrix}.$$

(a) Compute $-A + B - 2 \cdot C$.

(b) Compute $-3 \cdot A^{\mathsf{T}} + 4 \cdot C^{\mathsf{T}}$.

(c) Consider A, B, C as matrices. What are their dimensions? (How many rows and how many columns do they each have?)

(d) Consider $A^{\mathsf{T}}, B^{\mathsf{T}}, C^{\mathsf{T}}$ as matrices. What are their dimensions? (How many rows and how many columns do they each have?)

(e) Compute the 1×1-matrix $A \cdot B^{\mathsf{T}}$.

(f) Compute the 1×1-matrix $B \cdot A^{\mathsf{T}}$.

(g) Compute the 4×4-matrix $B^{\mathsf{T}} \cdot C$.

(h) Compute the 4×4-matrix $C^{\mathsf{T}} \cdot B$.

(i) Compute $(A + B) \cdot (B + C)^{\mathsf{T}}$.

(j) Compute $(A + B)^{\mathsf{T}} \cdot (B + C)$.

Problem 5. Define

$$X := \begin{bmatrix} 2 & -5 & 0 & 1 \\ 0 & 7 & -3 & -2 \end{bmatrix} \qquad \text{and} \qquad Y := \begin{bmatrix} -1 & 8 & 0 & 4 \\ -2 & -6 & 3 & 2 \end{bmatrix}.$$

(a) Compute $X + Y$.

(b) Compute $-X^{\mathsf{T}} + 2 \cdot Y^{\mathsf{T}}$.

Problem 6. Consider

$$M := \begin{bmatrix} 0 & 2 & -3 \\ -1 & 0 & 2 \\ 2 & 0 & 1 \end{bmatrix} \qquad \text{and} \qquad N := \begin{bmatrix} -3 & 0 & 1 \\ 0 & 2 & 2 \\ -1 & 0 & 0 \end{bmatrix}.$$

(a) Compute $M \cdot N$.

(b) Compute $N \cdot M$.

(c) Compute $\mathrm{tr}((M + N^{\mathtt{T}}) \cdot (M^{\mathtt{T}} + N))$.

Problem 7. Let

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

(a) Compute $A \cdot A^{\mathtt{T}}$.

(b) Compute $A^{\mathtt{T}} \cdot A$.

(c) Compute $\mathrm{tr}(A \cdot A^{\mathtt{T}})$ and $\mathrm{tr}(A^{\mathtt{T}} \cdot A)$. Are the two traces equal?

Solution to Problem 1. (a)

$$\begin{bmatrix} 3 & 4 & 5 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix}.$$

(b)

$$\begin{bmatrix} 3 & 5 & 7 \\ 4 & 6 & 8 \\ 5 & 7 & 9 \end{bmatrix}.$$

(c) $\mathrm{tr}(A) = \mathrm{tr}(A^{\mathsf{T}}) = 3 + 6 + 9 = 18$. The two traces are equal.

(d)

$$\begin{aligned}
\begin{bmatrix} 3 & 4 & 5 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix} &\xmapsto{R_1 \mapsto R_1 - R_2} \begin{bmatrix} -2 & -2 & -2 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix} \\
&\xmapsto{R_1 \mapsto (-1/2) \cdot R_1} \begin{bmatrix} 1 & 1 & 1 \\ 5 & 6 & 7 \\ 7 & 8 & 9 \end{bmatrix} \\
&\xmapsto{R_2 \mapsto R_2 - 5 \cdot R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 7 & 8 & 9 \end{bmatrix} \\
&\xmapsto{R_3 \mapsto R_3 - 7 \cdot R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 1 & 2 \end{bmatrix} \\
&\xmapsto{R_3 \mapsto R_3 - R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \\
&\xmapsto{R_1 \mapsto R_1 - R_2} \begin{bmatrix} 1 & 0 & -1 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix}.
\end{aligned}$$

□

Solution to Problem 2. (a)

$$B = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 8 & 12 & 16 \\ 9 & 18 & 27 & 36 \end{bmatrix}.$$

(b)

$$B^{\mathsf{T}} = \begin{bmatrix} 1 & 4 & 9 \\ 2 & 8 & 18 \\ 3 & 12 & 27 \\ 4 & 16 & 36 \end{bmatrix}.$$

(c)

$$\begin{bmatrix} 1 & 2 & 3 & 4 \\ 4 & 8 & 12 & 16 \\ 9 & 18 & 27 & 36 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - 4 \cdot R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 9 & 18 & 27 & 36 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 - 9 \cdot R_1} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

(d)

$$\begin{bmatrix} 1 & 4 & 9 \\ 2 & 8 & 18 \\ 3 & 12 & 27 \\ 4 & 16 & 36 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - 2 \cdot R_1} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ 3 & 12 & 27 \\ 4 & 16 & 36 \end{bmatrix}$$

$$\xmapsto{R_3 \mapsto R_3 - 3 \cdot R_1} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 4 & 16 & 36 \end{bmatrix}$$

$$\xmapsto{R_4 \mapsto R_4 - 4 \cdot R_1} \begin{bmatrix} 1 & 4 & 9 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

$\square$

Solution to Problem 3. The point of this problem is that 1×1-matrices behave just like real numbers. Moreover, if A is a 1×1-matrix, then A^{T} is just A itself. The trace of a 1×1-matrix is just the number inside the matrix.

(a) $A + B = [3 + (-2)] = [1]$.

(b) $B - 2 \cdot C = [(-2) - 2 \cdot 9] = [-20]$.

(c) $-3 \cdot A + B + 4 \cdot C = [(-3) \cdot 3 + (-2) + 4 \cdot 9] = [-9 - 2 + 36] = [25]$.

(d) $A \cdot B = [3 \cdot (-2)] = [-6]$.

(e) $B \cdot C = [(-2) \cdot 9] = [-18]$.

(f) $\operatorname{tr}(A+B)+\operatorname{tr}(C^{\mathsf{T}}) = 3+(-2)+9 = 10$. Note that the answer to this part is the number 10, not a matrix.
(g) $(A+B)\cdot(B+C) = [1]\cdot[7] = [7]$.
(h) $2\cdot A\cdot B - 3\cdot A\cdot C + 4\cdot B\cdot C = [-12-81-72] = [-165]$. $\square$

Solution to Problem 4. **(a)** Addition and subtraction are done entry-wise. Thus,

$$-A+B-2\cdot C = \begin{bmatrix} -4 & -2 & 8 & -9 \end{bmatrix}.$$

(b)

$$-3\cdot A^{\mathsf{T}} + 4\cdot C^{\mathsf{T}} = \begin{bmatrix} 0 \\ -3 \\ 3 \\ 6 \end{bmatrix} + \begin{bmatrix} 12 \\ 0 \\ -8 \\ 12 \end{bmatrix} = \begin{bmatrix} 12 \\ -3 \\ -5 \\ 18 \end{bmatrix}.$$

(c) A, B, C have the same dimension, which is 1×4: each has 1 row and 4 columns.
(d) $A^{\mathsf{T}}, B^{\mathsf{T}}, C^{\mathsf{T}}$ have the same dimension, which is 4×1: each has 4 rows and 1 column.
(e)

$$A\cdot B^{\mathsf{T}} = \begin{bmatrix} 0 & 1 & -1 & -2 \end{bmatrix}\cdot \begin{bmatrix} 2 \\ -1 \\ 3 \\ -5 \end{bmatrix} = \begin{bmatrix} 6 \end{bmatrix}.$$

(f)

$$B\cdot A^{\mathsf{T}} = \begin{bmatrix} 2 & -1 & 3 & -5 \end{bmatrix}\cdot \begin{bmatrix} 0 \\ 1 \\ -1 \\ -2 \end{bmatrix} = \begin{bmatrix} 6 \end{bmatrix}.$$

(g)

$$B^{\mathsf{T}}\cdot C = \begin{bmatrix} 6 & 0 & -4 & 6 \\ -3 & 0 & 2 & -3 \\ 9 & 0 & -6 & 9 \\ -15 & 0 & 10 & -15 \end{bmatrix}.$$

(h)

$$C^{\mathsf{T}}\cdot B = \begin{bmatrix} 6 & -3 & 9 & -15 \\ 0 & 0 & 0 & 0 \\ -4 & 2 & -6 & 10 \\ 6 & -3 & 9 & -15 \end{bmatrix}.$$

Note that $(B^{\mathsf{T}}\cdot C)^{\mathsf{T}} = C^{\mathsf{T}}\cdot B$.

(i)

$$(A+B)\cdot(B+C)^{\mathtt{T}} = \begin{bmatrix} 2 & 0 & 2 & -7 \end{bmatrix} \cdot \begin{bmatrix} 5 \\ -1 \\ 1 \\ -2 \end{bmatrix} = \begin{bmatrix} 26 \end{bmatrix}.$$

(j)

$$(A+B)^{\mathtt{T}}\cdot(B+C) = \begin{bmatrix} 2 \\ 0 \\ 2 \\ -7 \end{bmatrix} \cdot \begin{bmatrix} 5 & -1 & 1 & -2 \end{bmatrix} = \begin{bmatrix} 10 & -2 & 2 & -4 \\ 0 & 0 & 0 & 0 \\ 10 & -2 & 2 & -4 \\ -35 & 7 & -7 & 14 \end{bmatrix}.$$

Note that $(A+B)\cdot(B+C)^{\mathtt{T}}$ and $(A+B)^{\mathtt{T}}\cdot(B+C)$ are different. They do not even have the same dimensions! $\square$

Solution to Problem 5. **(a)** Addition happens entry-wise, so $X+Y = \begin{bmatrix} 1 & 3 & 0 & 5 \\ -2 & 1 & 0 & 0 \end{bmatrix}$.

(b) We have

$$-X^{\mathtt{T}} + 2\cdot Y^{\mathtt{T}} = -\begin{bmatrix} 2 & 0 \\ -5 & 7 \\ 0 & -3 \\ 1 & -2 \end{bmatrix} + 2\cdot \begin{bmatrix} -1 & -2 \\ 8 & -6 \\ 0 & 3 \\ 4 & 2 \end{bmatrix} = \begin{bmatrix} -4 & -4 \\ 21 & -19 \\ 0 & 9 \\ 7 & 6 \end{bmatrix}.$$

$\square$

Solution to Problem 6. **(a)**

$$M\cdot N = \begin{bmatrix} 3 & 4 & 4 \\ 1 & 0 & -1 \\ -7 & 0 & 2 \end{bmatrix}.$$

(b)

$$N\cdot M = \begin{bmatrix} 2 & -6 & 10 \\ 2 & 0 & 6 \\ 0 & -2 & 3 \end{bmatrix}.$$

Note that $M\cdot N \neq N\cdot M$.

(c)

$$(M+N^{\mathtt{T}})\cdot(M^{\mathtt{T}}+N) = \begin{bmatrix} 29 & -1 & -9 \\ -1 & 9 & 3 \\ -9 & 3 & 14 \end{bmatrix}.$$

Therefore, $\operatorname{tr}((M+N^{\mathtt{T}})\cdot(M^{\mathtt{T}}+N)) = 29+9+14 = 52$. $\square$

Solution to Problem 7. (a)

$$A \cdot A^{\mathsf{T}} = \begin{bmatrix} 5 & 2 & 1 \\ 0 & 4 & 3 \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 2 & 4 \\ 1 & 3 \end{bmatrix} = \begin{bmatrix} 30 & 11 \\ 11 & 25 \end{bmatrix}.$$

(b)

$$A^{\mathsf{T}} \cdot A = \begin{bmatrix} 5 & 0 \\ 2 & 4 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 5 & 2 & 1 \\ 0 & 4 & 3 \end{bmatrix} = \begin{bmatrix} 25 & 10 & 5 \\ 10 & 20 & 14 \\ 5 & 14 & 10 \end{bmatrix}.$$

(c) We have

$$\operatorname{tr}(A \cdot A^{\mathsf{T}}) = 30 + 25 = 55,$$

and

$$\operatorname{tr}(A^{\mathsf{T}} \cdot A) = 25 + 20 + 10 = 55.$$

The two traces are equal. $\square$

4. The Fourth Week

4.1. **Algebraic Properties of Matrices.** The main theme of this week's lecture is that we consider what we know to do with the arithmetic of real numbers and see if it can be generalized to matrices.

Here is an example.

Proposition 4.1. *For every real number $a \in \mathbb{R}$, we have $a + 0 = 0 + a = a$. In other words,*

$$\forall a \in \mathbb{R},\ a + 0 = 0 + a = a.$$

If we consider the equation

$$a + 0 = 0 + a = a,$$

we notice that there is the very special real number "zero" (denoted 0) with the property that when we add this zero either to the right or to the left of a real number a, the result remains the same real number a. For real numbers, we say that zero is the **additive identity**.

Now we consider whether there is something similar for matrices. For example, we might consider

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

We note that

$$\begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 0 \\ 0 & 3 \end{bmatrix}.$$

This is something that holds in general.

Proposition 4.2. *Let m and n be positive integers. Let $\mathbf{0}_{m\times n}$ denote the $m \times n$-zero matrix. For every $m \times n$-matrix $A \in \mathbb{R}^{m\times n}$, we have*

$$A + \mathbf{0}_{m\times n} = \mathbf{0}_{m\times n} + A = A.$$

In other words, $\mathbf{0}_{m\times n} \in \mathbb{R}^{m\times n}$ is the additive identity for $m \times n$-matrices.

Let's continue with this theme. Recall that for every real number $a \in \mathbb{R}$, we have

$$a + (-a) = (-a) + a = 0.$$

For matrices, we also have the analogous result.

Proposition 4.3. *Let m and n be positive integers. Let $\mathbf{0}_{m\times n}$ denote the $m \times n$-zero matrix. For every $m \times n$-matrix $A \in \mathbb{R}^{m\times n}$, we have*

$$A + (-A) = (-A) + A = \mathbf{0}_{m\times n}.$$

For any two real numbers a and b, we have

$$a + b = b + a.$$

This property is called **commutativity for addition** of real numbers. We say that the addition is **commutative** for real numbers. This also holds for matrices.

Proposition 4.4. *For any two matrices $A, B \in \mathbb{R}^{m\times n}$,*

$$A + B = B + A.$$

For any three real numbers $a, b, c \in \mathbb{R}$, we have

$$a + (b + c) = (a + b) + c.$$

This is called **associativity for addition** of real numbers. We say that the addition is **associative** for real numbers. This also holds for matrices.

Proposition 4.5. *For any three matrices $A, B, C \in \mathbb{R}^{m\times n}$,*

$$A + (B + C) = (A + B) + C.$$

Next, let us move on to multiplication. Recall that for any three real numbers $a, b, c \in \mathbb{R}$, we have

$$a \cdot (b \cdot c) = (a \cdot b) \cdot c.$$

Now, let us consider the following example.

Example 4.1. Let

$$A = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 & 1 \\ 1 & 4 \end{bmatrix}, \qquad \text{and} \qquad C = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}.$$

Observe that

$$B \cdot C = \begin{bmatrix} 1 & 1 \\ 5 & 4 \end{bmatrix},$$

and so

$$A \cdot (B \cdot C) = \begin{bmatrix} 1 & 2 \\ 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 5 & 4 \end{bmatrix} = \begin{bmatrix} 11 & 9 \\ 10 & 8 \end{bmatrix}.$$

Observe that

$$A \cdot B = \begin{bmatrix} 2 & 9 \\ 2 & 8 \end{bmatrix},$$

and so

$$(A \cdot B) \cdot C = \begin{bmatrix} 2 & 9 \\ 2 & 8 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 11 & 9 \\ 10 & 8 \end{bmatrix}.$$

Note that

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C.$$

Example 4.1 is a special case of the more general property that is **associativity for multiplication** of matrices.

Proposition 4.6. *Let m, n, p, q be positive integers. For any $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$, and $C \in \mathbb{R}^{p\times q}$, we have*

$$A \cdot (B \cdot C) = (A \cdot B) \cdot C \in \mathbb{R}^{m\times q}.$$

Because of the associativity, it now makes sense to write a product of more than two terms such as

$$A \cdot B \cdot C,$$

just like how it makes sense to write $2 \cdot 3 \cdot 5 = 30$. It does not matter whether we do $A \cdot B$ first, or do $B \cdot C$ first. The final result of the product is the same by the proposition above.

So far matrices seem to behave quite similarly to how the real numbers behave. Here's something different.

Recall that for any two real numbers $a, b \in \mathbb{R}$, we have

$$a \cdot b = b \cdot a.$$

In other words, we have **commutativity for multiplication** of real numbers.

It turns out, however, that there is no commutative for multiplications of matrices in general:

$$A \cdot B \neq B \cdot A \qquad \text{(in general)}.$$

Example 4.2. Consider

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix}.$$

Note that

$$A \cdot B = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix},$$

but

$$B \cdot A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 1 \end{bmatrix}.$$

Note that $A \cdot B \neq B \cdot A$.

Definition 4.3. We say that two matrices A and B **commute** (or, synonymously, that A **commutes with** B, or that B **commutes with** A) if $A \cdot B = B \cdot A$.

Example 4.4. Let

$$M = \begin{bmatrix} 1 & 3 \\ 0 & 1 \end{bmatrix} \qquad \text{and} \qquad N = \begin{bmatrix} 1 & 4 \\ 0 & 1 \end{bmatrix}.$$

Note that both $M \cdot N$ and $N \cdot M$ equal

$$\begin{bmatrix} 1 & 7 \\ 0 & 1 \end{bmatrix}.$$

Therefore, M and N commute.

Recall that for real numbers, we have the **distributive property**, which says that for $a, b, c \in \mathbb{R}$,

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad (a + b) \cdot c = a \cdot c + b \cdot c.$$

We have analogous results for matrices.

Proposition 4.7. *Let m, n, p be positive integers. Let $A \in \mathbb{R}^{m \times n}$ and $B, C \in \mathbb{R}^{n \times p}$. Then*

$$A \cdot (B + C) = A \cdot B + A \cdot C.$$

Proposition 4.8. *Let m, n, p be positive integers. Let $A, B \in \mathbb{R}^{m \times n}$ and $C \in \mathbb{R}^{n \times p}$. Then*

$$(A + B) \cdot C = A \cdot C + B \cdot C.$$

Definition 4.5. A square matrix is said to be an **identity matrix** if all its diagonal entries are 1, and all its off-diagonal entries are 0.

For example, the following are identity matrices:

$$[1] \in \mathbb{R}^{1\times 1}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \in \mathbb{R}^{2\times 2}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{3\times 3}. \quad \text{and} \quad \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \in \mathbb{R}^{4\times 4}.$$

The following are *not* identity matrices:

$$[0], \quad [-1], \quad \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

For each positive integer n, there is a unique[10] identity matrix of dimension $n \times n$. We denote it by

$$I_n \in \mathbb{R}^{n \times n}.$$

When the dimension $(n \times n)$ can be understood from the context, we say that I_n is *the* identity matrix. Let us also make a remark that we do not define an identity matrix when the dimension is not a square: all identity matrices are square matrices.

[10] In mathematics, the adjective "unique" means "only one".

Exercise 4.6. What is the trace of I_n?

A solution to this exercise is given in the footnote.[11]

Recall that for real numbers, the very special number 1 has the property that for every real number a,

$$1 \cdot a = a \cdot 1 = a \in \mathbb{R}.$$

We have something similar for matrices.

Proposition 4.9. *Let m and n be positive integers. For any $A \in \mathbb{R}^{m\times n}$, we have*

$$I_m \cdot A = A \cdot I_n = A \in \mathbb{R}^{m\times n}.$$

Example 4.7. Observe that

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ 2 & 0 & 3 \end{bmatrix}.$$

Recall that for real numbers $a, b, c \in \mathbb{R}$ such that $a \neq 0$, if

$$a \cdot b = a \cdot c,$$

then we can "cancel" a from both sides, and conclude that

$$b = c.$$

What about matrices? If A, B, C are matrices (where $A \cdot B$ and $A \cdot C$ are well-defined), do we always have that $A \cdot B = A \cdot C$ implies $B = C$?

Example 4.8. Let

$$A := \begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix}, \quad B := \begin{bmatrix} 2 & 0 \\ 1 & 4 \end{bmatrix}, \quad \text{and} \quad C := \begin{bmatrix} 0 & 2 \\ 2 & 3 \end{bmatrix}.$$

Observe that

$$A \cdot B = \begin{bmatrix} 4 & 8 \\ 12 & 24 \end{bmatrix},$$

and

$$A \cdot C = \begin{bmatrix} 4 & 8 \\ 12 & 24 \end{bmatrix}.$$

Thus, $A \neq \mathbf{0}_{2\times 2}$ and $A \cdot B = A \cdot C$, but $B \neq C$.

The previous example shows that even when A is not the zero matrix, the equation $A \cdot B = A \cdot C$ in general does not imply $B = C$. In general, you cannot simply "cancel" A from both sides!

---

[11] *Solution.* There are n entries on the diagonal of I_n, each of which is a number 1. Therefore, $\mathrm{tr}(I_n)$ is $1+1+\cdots+1$, with n copies of 1, and so $\mathrm{tr}(I_n) = n$. ■

4.2. **Matrix Inverses.** Throughout this subsection, let m and n denote positive integers.

We give the following definition.

Definition 4.9. Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{n\times m}$. We say that A and B are **inverses** (or, synonymously, A is an **inverse** of B, or, synonymously, B is an **inverse** of A) if

$$A \cdot B = I_m \qquad \text{and} \qquad B \cdot A = I_n.$$

In the definition above, we say A is *an* inverse of B, where we have used the article "an". This is because a priori we do not claim that an inverse is unique. A matrix inverse, however, if it does exist, is unique. This is the following proposition.

Proposition 4.10. *Suppose that $A \in \mathbb{R}^{m\times n}$ is a matrix with an inverse. Then the inverse of A is unique.*

In other words, suppose B is a matrix such that

$$A \cdot B = I_m \qquad \textit{and} \qquad B \cdot A = I_n,$$

and simultaneously suppose B' is a matrix such that

$$A \cdot B' = I_m \qquad \textit{and} \qquad B' \cdot A = I_n.$$

(This is saying that both B and B' are inverses of A.) Then $B = B'$.

Proof. The main idea of this proof is to consider the product

$$M := B \cdot A \cdot B'$$

of three matrices. Recall that such a product is well-defined by Proposition 4.6 above. There are now two ways to compute the product. First, we can group the first two factors:

$$M = (B \cdot A) \cdot B'.$$

Since $B \cdot A = I_n$, using Proposition 4.9, we find that

$$M = I_n \cdot B' = B'.$$

Second, we can group the last two factors:

$$M = B \cdot (A \cdot B').$$

Since $A \cdot B' = I_m$, using Proposition 4.9, we find that

$$M = B \cdot I_m = B.$$

Because both computations are for the same matrix M, we deduce that

$$B = M = B',$$

showing that B and B' must be the same matrix. $\square$

In light of Proposition 4.10, if A has an inverse, then we can now talk about *the* inverse of A. In such a case, we denote the inverse of A by A^{-1}.

Definition 4.10. Let $A \in \mathbb{R}^{m\times n}$. We say that A is **invertible** if A has an inverse. Otherwise, if A does not have an inverse, we say that A is **non-invertible**.

Remark 4.1. If A is a square matrix, we say that A is **non-singular** if and only if A is *invertible*. If A is a square matrix, we say that A is **singular** if and only if A is *non-invertible*. So, for square matrices,

$$\text{invertible} \;\Leftrightarrow\; \text{non-singular},$$

and

$$\text{non-invertible} \;\Leftrightarrow\; \text{singular}.$$

Proposition 4.11. *Let $A \in \mathbb{R}^{m\times n}$. If $m \neq n$, then A is non-invertible. In other words,* non-square *matrices are always not invertible. Every invertible matrix is a square matrix. (But not every square matrix is an invertible matrix!)*

We can also define **powers** of a matrix as follows. Suppose $A \in \mathbb{R}^{n\times n}$ is a square $n \times n$-matrix. Then for each positive integer k, we let

$$A^k := \underbrace{A \cdot A \cdots A}_{k \text{ copies of } A}.$$

Moreover, if A is invertible, then for each positive integer k, we let

$$A^{-k} := \underbrace{A^{-1} \cdot A^{-1} \cdots A^{-1}}_{k \text{ copies of } A^{-1}},$$

and let

$$A^0 := I_n \in \mathbb{R}^{n\times n}.$$

Exercise 4.11. Consider the matrix

$$B := \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}.$$

Verify that

$$B^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix},$$

and compute $B^{-2} + B^{-1} + B^0 + B^1 + B^2$.

A solution to this exercise is given in the footnote.[12]

[12] *Solution.* To verify that $\begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ is the inverse of B, we directly compute:

$$\begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \cdot \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2 \qquad \text{and} \qquad \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Just like how for every real number $a \neq 0$ and for every pair of integers k and ℓ, we have

$$a^{k+\ell} = a^k \cdot a^\ell,$$

we have the analog of this for invertible matrices.

Proposition 4.12. *Let A be an invertible matrix. For every pair of integers k and ℓ, we have*

$$A^{k+\ell} = A^k \cdot A^\ell.$$

Recall that for every real number $a \neq 0$, we have $(a^{-1})^{-1} = a$. We also have the same thing for invertible matrices.

Proposition 4.13. *Let A be an invertible matrix. Then the inverse of the inverse of A is A itself. That is,*

$$(A^{-1})^{-1} = A.$$

Exercise 4.12. Is the following statement true or false: "if A is an invertible matrix, then for every integer k, the matrix A^k is also invertible"?

A solution to this exercise is given in the footnote.[13]

4.3. **Invertibility of 2×2-matrices.** In this subsection, we focus on 2×2-matrices.

Proposition 4.14. *Let*

$$A := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

be a 2×2-matrix. Then A is invertible if and only if

$$ad - bc \neq 0.$$

This shows that $\begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$ is indeed the inverse B^{-1} of B.

Now we have

$$B^2 = B \cdot B = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 7 & 16 \\ 24 & 55 \end{bmatrix},$$

and

$$B^{-2} = B^{-1} \cdot B^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} 55 & -16 \\ -24 & 7 \end{bmatrix}.$$

Therefore,

$$B^{-2} + B^{-1} + B^0 + B^1 + B^2 = \begin{bmatrix} 55 & -16 \\ -24 & 7 \end{bmatrix} + \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} + \begin{bmatrix} 7 & 16 \\ 24 & 55 \end{bmatrix} = \begin{bmatrix} 71 & 0 \\ 0 & 71 \end{bmatrix}.$$

■

[13]*Solution.* The statement is true. Say $A \in \mathbb{R}^{n\times n}$. Note that $(A^k) \cdot (A^{-k}) = (A^{-k}) \cdot A^k = A^0 = I_n$. This shows that A^{-k} is the inverse of A^k, and hence A^k is invertible. ■

Moreover, if A is invertible, then its inverse is

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

Example 4.13. The matrix

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

is invertible. Its inverse is

$$A^{-1} = \frac{1}{2}\begin{bmatrix} 7 & -3 \\ -4 & 2 \end{bmatrix} = \begin{bmatrix} 7/2 & -3/2 \\ -2 & 1 \end{bmatrix}.$$

We can verify directly that

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

Example 4.14. Let

$$B := \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}.$$

What are $B^{-3}, B^{-2}, B^{-1}, B^0, B^1, B^2, B^3$?

First note that $2 \cdot 4 - 1 \cdot 7 = 1 \neq 0$. Thus, B is invertible, and its inverse is

$$B^{-1} = \frac{1}{1} \cdot \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix} = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix}.$$

We have

$$B^{-2} = B^{-1} \cdot B^{-1} = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix} \cdot \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix} = \begin{bmatrix} 23 & -6 \\ -42 & 11 \end{bmatrix}.$$

Next, we have

$$B^{-3} = B^{-1} \cdot B^{-2} = \begin{bmatrix} 4 & -1 \\ -7 & 2 \end{bmatrix} \cdot \begin{bmatrix} 23 & -6 \\ -42 & 11 \end{bmatrix} = \begin{bmatrix} 134 & -35 \\ -245 & 64 \end{bmatrix}.$$

Now,

$$B^0 = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$B^1 = B = \begin{bmatrix} 2 & 1 \\ 7 & 4 \end{bmatrix}.$$

Hence,

$$B^2 = B \cdot B = \begin{bmatrix} 11 & 6 \\ 42 & 23 \end{bmatrix}.$$

Finally,

$$B^3 = B \cdot B^2 = \begin{bmatrix} 64 & 35 \\ 245 & 134 \end{bmatrix}.$$

We have computed the powers B^k, for $k = -3, -2, \ldots, 1, 2, 3$.

4.4. **Properties of Transpose.** In this subsection, we list some properties of transpose.

Proposition 4.15. *For any matrix A, we have $(A^{\mathtt{T}})^{\mathtt{T}} = A$.*

Proposition 4.16. *For any matrix A and for any real number $t \in \mathbb{R}$, we have $(t \cdot A)^{\mathtt{T}} = t \cdot A^{\mathtt{T}}$. In particular, this implies that $(-A)^{\mathtt{T}} = -(A^{\mathtt{T}})$.*

Proposition 4.17. *For any two matrices A and B of the same dimension, $(A + B)^{\mathtt{T}} = A^{\mathtt{T}} + B^{\mathtt{T}}$.*

Proposition 4.18. *Let m, n, p be positive integers. For any $A \in \mathbb{R}^{m \times n}$ and for any $B \in \mathbb{R}^{n \times p}$, we have*

$$(A \cdot B)^{\mathtt{T}} = B^{\mathtt{T}} \cdot A^{\mathtt{T}}.$$

Proposition 4.19. *If A is an invertible matrix, then $A^{\mathtt{T}}$ is also an invertible matrix and*

$$(A^{\mathtt{T}})^{-1} = (A^{-1})^{\mathtt{T}}.$$

That is, the inverse of the transpose of A is the transpose of the inverse of A.

4.5. **Problems and Solutions.**

Problem 1. Let

$$A = \begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}, \qquad \text{and} \qquad C = \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix}.$$

(a) Compute $A \cdot B$.

(b) Compute $B \cdot C$.

(c) Using the result from part **(a)**, compute $(A \cdot B) \cdot C$.

(d) Using the result from part **(b)**, compute $A \cdot (B \cdot C)$. Verify that

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Problem 2. Let

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix}, \qquad \text{and} \qquad C = \begin{bmatrix} 1 & 5 & 3 & 1 \end{bmatrix}.$$

(a) Compute $A \cdot B$.

(b) Compute $B \cdot C$.

(c) Using the result from part **(a)**, compute $(A \cdot B) \cdot C$.

(d) Using the result from part **(b)**, compute $A \cdot (B \cdot C)$. Verify that

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

Problem 3. Let

$$A = \begin{bmatrix} 8 & 18 \\ -3 & -7 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} -3 & -12 \\ 2 & 7 \end{bmatrix}.$$

(a) Compute $A \cdot B$.

(b) Compute $B \cdot A$.

(c) Do A and B commute?

(d) Do $A^{\mathtt{T}}$ and $B^{\mathtt{T}}$ commute?

(e) Do A and B^{T} commute?

Problem 4. Let

$$A = \begin{bmatrix} 4 & 9 \\ 3 & 7 \end{bmatrix} \qquad \text{and} \qquad B = \begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix}.$$

(a) Compute $A \cdot B$.

(b) Compute $B \cdot A$.

(c) Do A and B commute?

Problem 5. Determine all ordered pairs (x, y) of real numbers for which the matrices

$$\begin{bmatrix} 1 & 0 \\ 1 & x \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 2 \\ y & 4 \end{bmatrix}$$

commute.

Problem 6. Determine all ordered pairs (x, y) of real numbers for which the matrices

$$\begin{bmatrix} 1 & -1 & 1 \\ -2 & -3 & 2 \\ -2 & x & 4 \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 1 & 0 & y \\ -1 & -1 & 1 \\ -1 & -2 & 2 \end{bmatrix}$$

commute.

Problem 7. Define

$$M := \begin{bmatrix} -1 & 2 & -3 \\ 0 & 1 & -1 \\ 2 & 3 & -2 \end{bmatrix} \qquad \text{and} \qquad N := \begin{bmatrix} 1 & -5 & 1 \\ -2 & 8 & -1 \\ -2 & 7 & -1 \end{bmatrix}.$$

(a) Compute $M \cdot N$.

(b) Compute $N \cdot M$.

(c) Are M and N inverses? Why? Explain.

Problem 8. Define

$$U := \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 0 \end{bmatrix} \qquad \text{and} \qquad V := \begin{bmatrix} 3 & -1 \\ -5 & 2 \\ 0 & 0 \end{bmatrix}.$$

Are U and V inverses? Why? Explain.

Problem 9. Let

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

(a) Compute A^2.

(b) Compute A^3.

(c) Compute A^6.

(d) Compute A^{12}.

(e) Compute $(I_2 + 4 \cdot (A^{-3}))^{-1}$, where $I_2 \in \mathbb{R}^{2\times 2}$ denotes the 2×2-identity matrix.

Problem 10. For each of the following matrices, do: (i) decide whether it is invertible, and explain how you arrive at the decision, and (ii) if it is invertible, find its inverse.

(a) `(3 points)`

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

(b) `(3 points)`

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

(c) `(3 points)`

$$\begin{bmatrix} -2 & -4 \\ 1 & 2 \end{bmatrix}$$

(d) `(3 points)`

$$\begin{bmatrix} 2 & 3 \\ 5 & 6 \end{bmatrix}$$

(e) `(3 points)`

$$\begin{bmatrix} 1/3 & 1/4 \\ 1/2 & -1/8 \end{bmatrix}$$

(f) `(3 points)`

$$\begin{bmatrix} 6/7 & -2 \\ 5/28 & -5/12 \end{bmatrix}$$

(g) `(3 points)`

$$\begin{bmatrix} 0 & 0 \\ 0 & 25 \end{bmatrix}$$

Problem 11. Let

$$C := \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \quad \text{and} \quad D := \begin{bmatrix} 0 & -1 \\ 2 & 2 \end{bmatrix}.$$

(a) Compute $(C \cdot D)^{\mathtt{T}}$.

(b) Compute $C^{\mathtt{T}} \cdot D^{\mathtt{T}}$. Do we have that $C^{\mathtt{T}} \cdot D^{\mathtt{T}} = (C \cdot D)^{\mathtt{T}}$?

(c) Compute $D^{\mathtt{T}} \cdot C^{\mathtt{T}}$. Do we have that $D^{\mathtt{T}} \cdot C^{\mathtt{T}} = (C \cdot D)^{\mathtt{T}}$?

Solution to Problem 1. **(a)** We have

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

(b) We have

$$B \cdot C = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1 & 1 \\ -5 & 2 \end{bmatrix}.$$

(c) We have

$$(A \cdot B) \cdot C = \begin{bmatrix} 0 & 3 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ -4 & 4 \end{bmatrix}.$$

(d) We have

$$A \cdot (B \cdot C) = \begin{bmatrix} -2 & 1 \\ 4 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1 & 1 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ -4 & 4 \end{bmatrix},$$

which shows that indeed

$$(A \cdot B) \cdot C = A \cdot (B \cdot C).$$

This is *associativity* of multiplication of matrices. $\square$

Solution to Problem 2. **(a)**

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

(b)

$$B \cdot C = \begin{bmatrix} 2 \\ 4 \\ 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 2 & 10 & 6 & 2 \\ 4 & 20 & 12 & 4 \\ 3 & 15 & 9 & 3 \end{bmatrix}.$$

(c)

$$(A \cdot B) \cdot C = \begin{bmatrix} 5 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 25 & 15 & 5 \\ 1 & 5 & 3 & 1 \end{bmatrix}.$$

(d)

$$A \cdot (B \cdot C) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 10 & 6 & 2 \\ 4 & 20 & 12 & 4 \\ 3 & 15 & 9 & 3 \end{bmatrix} = \begin{bmatrix} 5 & 25 & 15 & 5 \\ 1 & 5 & 3 & 1 \end{bmatrix},$$

which shows that indeed $(A \cdot B) \cdot C = A \cdot (B \cdot C)$. $\square$

Solution to Problem 3. **(a)**

$$A \cdot B = \begin{bmatrix} 8 & 18 \\ -3 & -7 \end{bmatrix} \cdot \begin{bmatrix} -3 & -12 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 12 & 30 \\ -5 & -13 \end{bmatrix}.$$

(b)

$$B \cdot A = \begin{bmatrix} -3 & -12 \\ 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 8 & 18 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} 12 & 30 \\ -5 & -13 \end{bmatrix}.$$

(c) Since $A \cdot B = B \cdot A$, we find that A and B do commute.
(d) Note that

$$A^{\mathsf{T}} \cdot B^{\mathsf{T}} = \begin{bmatrix} 8 & -3 \\ 18 & -7 \end{bmatrix} \cdot \begin{bmatrix} -3 & 2 \\ -12 & 7 \end{bmatrix} = \begin{bmatrix} 12 & -5 \\ 30 & -13 \end{bmatrix},$$

and

$$B^{\mathsf{T}} \cdot A^{\mathsf{T}} = \begin{bmatrix} -3 & 2 \\ -12 & 7 \end{bmatrix} \cdot \begin{bmatrix} 8 & -3 \\ 18 & -7 \end{bmatrix} = \begin{bmatrix} 12 & -5 \\ 30 & -13 \end{bmatrix}.$$

Therefore, $A^{\mathsf{T}} \cdot B^{\mathsf{T}} = B^{\mathsf{T}} \cdot A^{\mathsf{T}}$, and thus A^{T} and B^{T} commute.

Note that actually we did not have to multiply these matrices again to obtain the conclusion above. If A and B commute, then A^{T} and B^{T} also commute. This is because

$$A^{\mathsf{T}} \cdot B^{\mathsf{T}} = (B \cdot A)^{\mathsf{T}} = (A \cdot B)^{\mathsf{T}} = B^{\mathsf{T}} \cdot A^{\mathsf{T}}.$$

(e)

$$A \cdot B^{\mathsf{T}} = \begin{bmatrix} 8 & 18 \\ -3 & -7 \end{bmatrix} \cdot \begin{bmatrix} -3 & 2 \\ -12 & 7 \end{bmatrix} = \begin{bmatrix} -240 & 142 \\ 93 & -55 \end{bmatrix},$$

while

$$B^{\mathsf{T}} \cdot A = \begin{bmatrix} -3 & 2 \\ -12 & 7 \end{bmatrix} \cdot \begin{bmatrix} 8 & 18 \\ -3 & -7 \end{bmatrix} = \begin{bmatrix} -30 & -68 \\ -117 & -265 \end{bmatrix}.$$

We conclude that A and B^{T} do not commute.

Observe that even though A commutes with B, and A^{T} commutes with B^{T}, the matrices A and B^{T} do not commute. $\square$

Solution to Problem 4. **(a)**

$$A \cdot B = \begin{bmatrix} 4 & 9 \\ 3 & 7 \end{bmatrix} \cdot \begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix} = \begin{bmatrix} 22 & 79 \\ 17 & 61 \end{bmatrix}.$$

(b)

$$B \cdot A = \begin{bmatrix} 1 & 4 \\ 2 & 7 \end{bmatrix} \cdot \begin{bmatrix} 4 & 9 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 16 & 37 \\ 29 & 67 \end{bmatrix}.$$

(c) Since $A \cdot B \neq B \cdot A$, the two matrices do not commute. $\square$

Solution to Problem 5. The set of all such ordered pairs (x, y) is $\varnothing$, the empty set. In other words, no $(x, y) \in \mathbb{R}^2$ has the property in the problem.

Note that

$$\begin{bmatrix} 1 & 0 \\ 1 & x \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ y & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 + xy & 2 + 4x \end{bmatrix},$$

and

$$\begin{bmatrix} 1 & 2 \\ y & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 1 & x \end{bmatrix} = \begin{bmatrix} 3 & 2x \\ 4 + y & 4x \end{bmatrix}.$$

For any choice of $(x, y) \in \mathbb{R}^2$, the two products cannot be equal, since their $(1, 1)$-entries are 1 and 3, which are different numbers. $\square$

Solution to Problem 6. The set of all such ordered pairs (x, y) is $\{(-5, 0)\}$.

Suppose that $(x, y) \in \mathbb{R}^2$ is such that the two matrices commute. We have that the two products

$$P_1 := \begin{bmatrix} 1 & -1 & 1 \\ -2 & -3 & 2 \\ -2 & x & 4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & y \\ -1 & -1 & 1 \\ -1 & -2 & 2 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 + y \\ -1 & -1 & 1 - 2y \\ -6 - x & -8 - x & 8 + x - 2y \end{bmatrix}$$

and

$$P_2 := \begin{bmatrix} 1 & 0 & y \\ -1 & -1 & 1 \\ -1 & -2 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & 1 \\ -2 & -3 & 2 \\ -2 & x & 4 \end{bmatrix} = \begin{bmatrix} 1 - 2y & -1 + xy & 1 + 4y \\ -1 & 4 + x & 1 \\ -1 & 7 + 2x & 3 \end{bmatrix}$$

have to be equal. Consider the $(1, 1)$-entries. We find

$$1 = 1 - 2y,$$

which implies $y = 0$. Consider the $(3, 1)$-entries. We find

$$-6 - x = -1,$$

which implies $x = -5$. We conclude that $(x, y) = (-5, 0)$.

Conversely, let us check if $(x, y) = (-5, 0)$ does make the two matrices commute. When $(x, y) = (-5, 0)$, we have

$$P_1 = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & -3 & 3 \end{bmatrix}$$

and

$$P_2 = \begin{bmatrix} 1 & -1 & 1 \\ -1 & -1 & 1 \\ -1 & -3 & 3 \end{bmatrix},$$

whence $P_1 = P_2$. Therefore, the two matrices commute. $\square$

Solution to Problem 7. (a)

$$M \cdot N = \begin{bmatrix} -1 & 2 & -3 \\ 0 & 1 & -1 \\ 2 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -5 & 1 \\ -2 & 8 & -1 \\ -2 & 7 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

(b)

$$N \cdot M = \begin{bmatrix} 1 & -5 & 1 \\ -2 & 8 & -1 \\ -2 & 7 & -1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 2 & -3 \\ 0 & 1 & -1 \\ 2 & 3 & -2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

(c) Since $M \cdot N = N \cdot M = I_3$, we conclude that M and N are inverses. $\square$

Solution to Problem 8. No, they are not inverses. Note that

$$U \cdot V = \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 0 \end{bmatrix} \cdot \begin{bmatrix} 3 & -1 \\ -5 & 2 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2,$$

and

$$V \cdot U = \begin{bmatrix} 3 & -1 \\ -5 & 2 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \neq I_3.$$

For U and V to be inverses, *both* products $U \cdot V$ and $V \cdot U$ have to be identity matrices. In this case, $U \cdot V$ is an identity matrix, but $V \cdot U$ is not an identity matrix. $\square$

Remark. In fact, every non-square matrix does not have an inverse in this sense. For instance, if $W \in \mathbb{R}^{3\times 4}$ is a 3×4-matrix, then there does not exist a matrix $X \in \mathbb{R}^{4\times 3}$ such that

$$X \cdot W = I_4.$$

Or, as another example, consider the matrix

$$U = \begin{bmatrix} 2 & 1 & 0 \\ 5 & 3 & 0 \end{bmatrix}$$

from above. There does not exist[14] a matrix $A \in \mathbb{R}^{3\times 2}$ such that

$$A \cdot U = I_3.$$

[14]Why? At least for this matrix U, we can prove this claim here. Suppose, for the sake of contradiction, that such a matrix A for which $A \cdot U = I_3$ exists. Let us multiply both sides of the equation by the matrix V from the problem. We find

$$A \cdot U \cdot V = I_3 \cdot V.$$

Note that since $U \cdot V = I_2$, the left-hand side is A, and note that the right-hand side is V. We obtain the conclusion that $A = V$. So such a matrix has to be the matrix V. On the other hand, we just saw in the solution that $V \cdot U \neq I_3$, a contradiction. $\blacksquare$

The matrix U does not have an inverse (in the sense defined in this class).

Solution to Problem 9. (a)

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

(b)

$$A^3 = A \cdot A^2 = \begin{bmatrix} 0 & -1 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} -2 & -1 \\ 2 & -1 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ -2 & -3 \end{bmatrix}.$$

(c)

$$A^6 = A^3 \cdot A^3 = \begin{bmatrix} -2 & 1 \\ -2 & -3 \end{bmatrix} \cdot \begin{bmatrix} -2 & 1 \\ -2 & -3 \end{bmatrix} = \begin{bmatrix} 2 & -5 \\ 10 & 7 \end{bmatrix}.$$

(d)

$$A^{12} = A^6 \cdot A^6 = \begin{bmatrix} 2 & -5 \\ 10 & 7 \end{bmatrix} \cdot \begin{bmatrix} 2 & -5 \\ 10 & 7 \end{bmatrix} = \begin{bmatrix} -46 & -45 \\ 90 & -1 \end{bmatrix}.$$

(e) Using the formula given in class, the inverse of A is

$$A^{-1} = \frac{1}{2} \cdot \begin{bmatrix} 1 & 1 \\ -2 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 1/2 \\ -1 & 0 \end{bmatrix}.$$

Thus,

$$A^{-2} = A^{-1} \cdot A^{-1} = \begin{bmatrix} 1/2 & 1/2 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1/2 & 1/2 \\ -1 & 0 \end{bmatrix} = \begin{bmatrix} -1/4 & 1/4 \\ -1/2 & -1/2 \end{bmatrix},$$

and

$$A^{-3} = A^{-1} \cdot A^{-2} = \begin{bmatrix} 1/2 & 1/2 \\ -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} -1/4 & 1/4 \\ -1/2 & -1/2 \end{bmatrix} = \begin{bmatrix} -3/8 & -1/8 \\ 1/4 & -1/4 \end{bmatrix}.$$

This shows that

$$I_2 + 4 \cdot A^{-3} = \begin{bmatrix} -1/2 & -1/2 \\ 1 & 0 \end{bmatrix}.$$

Using the formula for inverses again, we find

$$(I_2 + 4 \cdot A^{-3})^{-1} = \frac{1}{1/2} \cdot \begin{bmatrix} 0 & 1/2 \\ -1 & -1/2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -1 \end{bmatrix}.$$

$\square$

Solution to Problem 10. (a)

$$1 \cdot 1 - 0 \cdot 0 = 1 \neq 0,$$

so the matrix is invertible. Its inverse is

$$\frac{1}{1} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

(b)

$$0 \cdot 0 - 1 \cdot (-1) = 1 \neq 0,$$

so the matrix is invertible. Its inverse is

$$\frac{1}{1} \cdot \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

(c)

$$(-2) \cdot 2 - (-4) \cdot 1 = 0,$$

so the matrix is not invertible. There is no inverse.

(d)

$$2 \cdot 6 - 3 \cdot 5 = -3 \neq 0,$$

so the matrix is invertible. Its inverse is

$$\frac{1}{-3} \cdot \begin{bmatrix} 6 & -3 \\ -5 & 2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 5/3 & -2/3 \end{bmatrix}.$$

(e)

$$(1/3)(-1/8) - (1/4)(1/2) = (-1/24) - (1/8) = -1/6 \neq 0,$$

so the matrix is invertible. Its inverse is

$$\frac{1}{-1/6} \cdot \begin{bmatrix} -1/8 & -1/4 \\ -1/2 & 1/3 \end{bmatrix} = \begin{bmatrix} 3/4 & 3/2 \\ 3 & -2 \end{bmatrix}.$$

(f)

$$(6/7)(-5/12) - (-2)(5/28) = 0,$$

so the matrix is not invertible. There is no inverse.

(g)

$$0 \cdot 25 - 0 \cdot 0 = 0,$$

so the matrix is not invertible. There is no inverse. $\square$

Solution to Problem 11. (a)

$$C \cdot D = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 \\ 2 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 3 \\ 2 & 2 \end{bmatrix},$$

and therefore,

$$(C \cdot D)^{\mathtt{T}} = \begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix}.$$

(b)

$$C^{\mathtt{T}} \cdot D^{\mathtt{T}} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 2 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -1 & 6 \end{bmatrix}.$$

No, we do not have that $C^{\mathtt{T}} \cdot D^{\mathtt{T}} = (C \cdot D)^{\mathtt{T}}$.

(c)

$$D^{\mathtt{T}} \cdot C^{\mathtt{T}} = \begin{bmatrix} 0 & 2 \\ -1 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 3 & 2 \end{bmatrix}.$$

Yes, we do have that $D^{\mathtt{T}} \cdot C^{\mathtt{T}} = (C \cdot D)^{\mathtt{T}}$. □

Remark. The point of this is that if we have square matrices X and Y of the same dimension, then

$$(X \cdot Y)^{\mathtt{T}} = Y^{\mathtt{T}} \cdot X^{\mathtt{T}}.$$

(Note the switched order.) On the other hand, $X^{\mathtt{T}} \cdot Y^{\mathtt{T}}$ and $(X \cdot Y)^{\mathtt{T}}$ are different in general.

5. The Fifth Week

5.1. **Elementary matrices.** At this point of the class, we should be familiar with Gaussian elimination, where we perform a sequence of elementary row operations to a matrix to transform the matrix into a matrix in reduced row echelon form.

Let us consider the following example. Suppose we start with the matrix

$$A_1 := \begin{bmatrix} 0 & 1 \\ 3 & 15 \end{bmatrix}.$$

First, we swap R_1 and R_2 to obtain

$$A_1 = \begin{bmatrix} 0 & 1 \\ 3 & 15 \end{bmatrix} \xmapsto{R_1 \rightleftharpoons R_2} \begin{bmatrix} 3 & 15 \\ 0 & 1 \end{bmatrix} =: A_2.$$

Second, we multiply R_1 by $1/3$ to obtain

$$A_2 = \begin{bmatrix} 3 & 15 \\ 0 & 1 \end{bmatrix} \xmapsto{R_1 \mapsto (1/3) \cdot R_1} \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} =: A_3.$$

Finally, we subtract $5 \cdot R_2$ from R_1 to obtain

$$A_3 = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \xmapsto{R_1 \mapsto R_1 - 5 \cdot R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} =: A_4.$$

The matrix A_4, which the identity matrix $I_2 \in \mathbb{R}^{2\times 2}$, is in reduced row echelon form, and so we have finished.

Now let us consider the four matrices that occur in the process above:

$$A_1 = \begin{bmatrix} 0 & 1 \\ 3 & 15 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 3 & 15 \\ 0 & 1 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad A_4 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Recall that we went from A_1, to A_2, to A_3, and finally to A_4. Note that we can obtain A_2 from A_1 in the following way. Let us multiply the matrix

$$E_1 := \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

to the *left* of A_1:

$$E_1 \cdot A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 3 & 15 \end{bmatrix} = \begin{bmatrix} 3 & 15 \\ 0 & 1 \end{bmatrix} = A_2.$$

Similarly, we can go from A_2 to A_3: if we define

$$E_2 := \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix},$$

then observe that
$$E_2 \cdot A_2 = \begin{bmatrix} 1/3 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 15 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} = A_3.$$
To go from A_3 to A_4, define
$$E_3 := \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix},$$
and note that
$$E_3 \cdot A_3 = \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = A_4.$$
So we have found that
$$A_2 = E_1 \cdot A_1,$$
$$A_3 = E_2 \cdot A_2 = E_2 \cdot E_1 \cdot A_1,$$
and
$$A_4 = E_3 \cdot A_3 = E_3 \cdot E_2 \cdot A_2 = E_3 \cdot E_2 \cdot E_1 \cdot A_1.$$
Thus, the effect of the sequence of elementary row operations on A_1 above is the same as the effect of multiplying matrices to the left of A_1. This is a general phenomenon: each elementary row operation on a matrix $A \in \mathbb{R}^{m\times n}$ is equivalent to a multiplication to the left of the matrix $A \in \mathbb{R}^{m\times n}$ by a certain square matrix $E \in \mathbb{R}^{m\times m}$.

But how do we find the square matrix E, for each given elementary row operation? It turns out that the matrix E is the result of doing the *same* elementary row operation to the identity matrix $I_m \in \mathbb{R}^{m\times m}$.

Here are some examples.

Example 5.1. Consider
$$A := \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix}.$$
We can perform the following elementary row operation:
$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix} \xmapsto{R_2 \rightleftharpoons R_3} \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 0 \\ 4 & 5 & 6 \end{bmatrix} =: B.$$
Now let us find a matrix E such that
$$E \cdot A = B,$$
following the description above. What should be the dimension of E? Note that we want E to be a square matrix which can be multiplied to the *left* of A. Since A has dimension 3×3, the dimension

of E should be 3×3. So let us perform the same elementary row operation to I_3:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xmapsto{R_2 \rightleftharpoons R_3} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} =: E.$$

Note that

$$E \cdot A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 8 & 0 \\ 4 & 5 & 6 \end{bmatrix} = B.$$

Example 5.2. Let us consider the following matrix

$$C := \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 16 \end{bmatrix}.$$

Perform the following elementary row operation:

$$C = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 16 \end{bmatrix} \xmapsto{R_2 \mapsto (1/4) \cdot R_2} \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix} = D.$$

We want to compute the square matrix E according to the description above. What should be the dimension of E? Note that we want E to be a square matrix which can be multiplied to the *left* of C. Since C has dimension 2×3, the dimension of E should be 2×2. Now perform the same elementary row operation to I_2:

$$I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \xmapsto{R_2 \mapsto (1/4) \cdot R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1/4 \end{bmatrix} =: E.$$

Observe that

$$E \cdot C = \begin{bmatrix} 1 & 0 \\ 0 & 1/4 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 4 & 8 & 16 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \end{bmatrix} = D.$$

Example 5.3. Let

$$X := \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 4 & 8 \end{bmatrix}.$$

Perform the following elementary row operation:

$$X = \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 4 & 6 \end{bmatrix} \xmapsto{R_3 \mapsto R_3 - 4 \cdot R_1} \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 0 & 2 \end{bmatrix} =: Y.$$

Let us compute the square matrix E. What should be the dimension of E? We want to multiply E to the *left* of X. Since X has dimension 3×2, the dimension of E should be 3×3. Now perform the same elementary row operation on I_3:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xmapsto{R_3 \mapsto R_3 - 4 \cdot R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} =: E.$$

Observe that

$$E \cdot X = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 4 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 3 & 5 \\ 0 & 2 \end{bmatrix} = Y.$$

Now we give a definition.

Definition 5.4. Let m be a positive integer. A square matrix $E \in \mathbb{R}^{m \times m}$ is said to be an **elementary matrix** if it is the result of performing one elementary row operation to the $m \times m$-identity matrix I_m.

For example,

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

is an elementary matrix: it is obtained by swapping R_1 and R_2 in I_3.

The matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & -5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is also an elementary matrix: it is obtained by multiplying (-5) to the second row of I_3.

The matrix

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 7 & 1 \end{bmatrix}$$

is also an elementary matrix: it is obtained by adding $7 \cdot R_2$ to R_3 in I_3.

Exercise 5.5. Determine all real numbers $a \in \mathbb{R}$ for which the 1×1-matrix

$$[a]$$

is an elementary matrix.

A solution to this exercise is given in the footnote.[15]

[15] *Solution.* For any $a \neq 0$, we can obtain the matrix $[a]$ by multiplying a to the only row of I_1. On the other hand, there is no way to transform $I_1 = [1]$ into $[0]$ using elementary row operations. Therefore, the set of answers is $\mathbb{R} - \{0\}$, the set of all nonzero real numbers. ∎

Exercise 5.6. Is the identity matrix

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

an elementary matrix?

A solution to this exercise is given in the footnote.[16]

Exercise 5.7. Determine all real numbers t for which the matrix

$$\begin{bmatrix} 1 & 2-t & 0 \\ 0 & 1 & 3t+3 \\ 0 & 0 & 1 \end{bmatrix}$$

is an elementary matrix.

A solution to this exercise is given in the footnote.[17]

We have the following proposition.

Proposition 5.1. *Suppose that $A \in \mathbb{R}^{m \times n}$. Suppose that after a sequence of elementary row operations on A, we obtain the matrix B in reduced row echelon form. Then there exists a sequence of elementary matrices*

$$E_1, E_2, \ldots, E_k$$

such that

$$E_k \cdot E_{k-1} \cdots E_2 \cdot E_1 \cdot A = B.$$

---

[16] *Solution.* Yes, it is. We can obtain the identity matrix from the identity matrix itself by multiplying 1 to the first row, for example. This operation is allowed. ■

[17] *Solution.* The answer set is $\{-1, 2\}$. Suppose that the matrix in the exercise is an elementary matrix. Since all the diagonal entries are 1 (and there are no 0 on the diagonal), we deduce that the elementary is not obtained from swapping two rows of I_3. It has to be either from (i) multiplying a row of I_3 by a nonzero constant, or from (ii) adding some $c \cdot R_j$ to R_i. In either case, there can be *at most one* nonzero off-diagonal entry: either $2-t$ or $3t+3$ has to be zero. Therefore, either $t = 2$ or $t = -1$.

Conversely, if $t = -1$, the matrix is

$$\begin{bmatrix} 1 & 3 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

which can be obtained from I_3 by adding $3 \cdot R_2$ to R_1. And, if $t = 2$, the matrix is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 9 \\ 0 & 0 & 1 \end{bmatrix},$$

which can be obtained from I_3 by adding $9 \cdot R_3$ to R_2. In either case, the resulting matrix is an elementary matrix. Thus, the answer set is $\{-1, 2\}$, as claimed. ■

Remark 5.1. We have seen that multiplying an elementary matrix E to the *left* of a matrix A is equivalent to performing one step of an elementary row operation on A. One can ask what happens if we instead multiply E to the *right* of A? The answer is that it would be equivalent to performing one step of an elementary *column* operation on A. But for our purposes at the moment, let us only focus on *row* operations.

5.2. **Inverses of elementary matrices.** Elementary matrices are invertible. It is not difficult to compute their inverses. Suppose that $E \in \mathbb{R}^{m\times m}$ is an elementary matrix. The following describes how to find the inverse of E. The idea is to apply the "reverse operation".

Proposition 5.2. *The inverse of an elementary matrix $E \in \mathbb{R}^{m\times m}$ can be obtained by the following rules.*

(i) *First, if E comes from swapping two rows of I_m, then E is the inverse of itself:*

$$E \cdot E = I_m, \qquad \text{and therefore} \qquad E^{-1} = E.$$

(ii) *Second, if E comes from multiplying the i^{th} row of I_m with a constant $c \neq 0$, then let E' be the elementary matrix coming from multiplying the i^{th} row of I_m by $1/c$. We have*

$$E \cdot E' = E' \cdot E = I_m, \qquad \text{and therefore} \qquad E^{-1} = E'.$$

(iii) *Third, if E comes from adding $c \cdot R_j$ to R_i in I_m, then let E' be the elementary matrix coming from adding $(-c) \cdot R_j$ to R_i in I_m. We have*

$$E \cdot E' = E' \cdot E = I_m, \qquad \text{and therefore} \qquad E^{-1} = E'.$$

Proof. (i) In the first case, note that $E \cdot E$ is simply E multiplied to the left of E. That means $E \cdot E$ can be obtained from swapping two rows, say R_i and R_j, in E. However, E itself comes from swapping R_i and R_j in I_m. Thus, $E \cdot E$ is the result of swapping R_i and R_j in I_m and then swapping it back. Therefore, $E \cdot E = I_m$.

(ii) In the second case, note that the entries of E agree with those in the identity matrix almost in all entries, except that its (i, i)-entry is c. The matrix $E' \cdot E$ is the result of multiplying the i^{th} row of E by $1/c$, and thus the (i, i)-entry in $E' \cdot E$ becomes $(1/c) \cdot c = 1$. Hence, $E' \cdot E = I_m$. Similarly, we have that $E \cdot E' = I_m$.

(iii) In the third case, note that the entries of E agree with those in the identity matrix almost in all entries, except that its (i, j)-entry is c. The matrix $E' \cdot E$ is the result of adding $-c$ times the j^{th} row of E to the i^{th} row of E. So the (i, j)-entry of $E' \cdot E$ becomes $c + (-c) \cdot 1 = 0$. Hence, $E' \cdot E = I_m$. Similarly, we have that $E \cdot E' = I_m$. □

The following corollary is an immediate consequence of the above proposition.

Corollary 5.3. *Every elementary matrix is invertible. Moreover, its inverse is an elementary matrix.*

Example 5.8. Here are some examples.

The inverse of

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

is

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

itself.

The inverse of

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 5 \end{bmatrix}$$

is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1/5 \end{bmatrix}.$$

The inverse of

$$\begin{bmatrix} 1 & 0 & 0 \\ 8 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is

$$\begin{bmatrix} 1 & 0 & 0 \\ -8 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

We have the following proposition.

Proposition 5.4. *Every matrix $A \in \mathbb{R}^{m\times n}$ can be written in the form*

$$A = X_1 \cdot X_2 \cdots X_k \cdot B,$$

where $X_1, X_2, \ldots, X_k$ are elementary matrices and $B \in \mathbb{R}^{m\times n}$ is a matrix in reduced row echelon form.

Proof. From Proposition 5.1, we can write

$$E_k \cdot E_{k-1} \cdots E_2 \cdot E_1 \cdot A = B.$$

Now multiply both sides of the equation by

$$E_1^{-1} \cdot E_2^{-1} \cdots E_k^{-1}$$

to obtain

$$A = E_1^{-1} \cdot E_2^{-1} \cdots E_k^{-1} \cdot B.$$

Since each E_i^{-1} is an elementary matrix, we have finished the proof. $\square$

Example 5.9. By performing elementary row operations on

$$A := \begin{bmatrix} 0 & 1 & 3 \\ 1 & 3 & 5 \end{bmatrix},$$

we find

$$\begin{bmatrix} 1 & -3 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 3 \\ 1 & 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 3 \end{bmatrix}.$$

Therefore,

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

The last equation expresses A as a product of elementary matrices and a matrix in reduced row echelon form.

5.3. **Computing the inverse of an invertible matrix.** Suppose we start with any square matrix $A \in \mathbb{R}^{m \times m}$. By a sequence of elementary row operations, we can reduce A to a matrix in reduced row echelon form. Call this matrix B. From the above discussion, we know that there are elementary matrices $X_1, X_2, \ldots, X_k$ such that

$$A = X_1 \cdot X_2 \cdots X_k \cdot B.$$

Note that B is a matrix in reduced row echelon form of the same dimension as A: it is an $m \times m$-matrix. It is not hard to see that every square matrix in reduced row echelon form (such as B) either

- is the identity matrix I_m, or
- has an all-zero row.

Proposition 5.5. *Let A and B be as above. Namely, B is a matrix in reduced row echelon form coming from A. Then we have the following.*

(i) *If $B = I_m$ is the identity matrix, then A is invertible.*

(ii) *If B has an all-zero row, then A is non-invertible.*

Now for the rest of this subsection, let us suppose that $A \in \mathbb{R}^{m\times m}$ is an invertible matrix. We can write

$$A = X_1 \cdot X_2 \cdots X_k \cdot B,$$

in the sense of Proposition 5.4. Now from Proposition 5.5, we have that B must be the identity I_m. This means

$$A = X_1 \cdot X_2 \cdots X_k.$$

If we let E_i denote the inverse of X_i, then we have

$$(E_k \cdot E_{k-1} \cdots E_2 \cdot E_1) \cdot A = I_m.$$

Thus,

$$E_k \cdot E_{k-1} \cdots E_2 \cdot E_1$$

is the inverse of A. That is,

$$A^{-1} = E_k \cdot E_{k-1} \cdots E_2 \cdot E_1.$$

Example 5.10. Recall the matrix

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

from Example 5.13.

We have that

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}.$$

Therefore,

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

5.4. **Determinants.** Let us fix a positive integer m throughout this subsection. The *determinant*, which we are defining below, is a function that takes an $m \times m$-matrix as an input, and returns a number as an output. The determinant is an important quantity associated to each square matrix. It has applications in many areas of mathematics and science.

Definition 5.11. Let m be a positive integer. The **determinant** is a function

$$\det : \mathbb{R}^{m\times m} \to \mathbb{R}$$

which satisfies the following properties:

(i) for every pair $A, B \in \mathbb{R}^{m\times m}$, we have

$$\det(A \cdot B) = \det(A) \cdot \det(B),$$

(ii) if E is an elementary matrix coming from swapping two rows of I_m, then

$$\det(E) = -1,$$

(iii) if E is an elementary matrix coming from multiplying a row of I_m by a nonzero constant $c \neq 0$, then

$$\det(E) = c,$$

(iv) if E is an elementary matrix coming from adding $c \cdot R_j$ to R_i in I_m, then

$$\det(E) = 1,$$

and

(v) for every matrix $A \in \mathbb{R}^{m\times m}$, if A has an all-zero row, then

$$\det(A) = 0.$$

The value $\det(A)$ is called the **determinant** of A.

Example 5.12. What is $\det(I_m)$?

Since I_m is an elementary matrix obtained by multiplying the first row of I_m with $c = 1$, the item (iii) in Definition 5.11 implies that

$$\det(I_m) = c = 1.$$

To compute the determinant of a matrix $A \in \mathbb{R}^{m\times m}$ in general, we can first perform elementary row operations to write

$$A = X_1 \cdot X_2 \cdots X_k \cdot B,$$

in the sense of Proposition 5.4.

Then by (i) in Definition 5.11, we obtain

$$\det(A) = \det(X_1) \cdot \det(X_2) \cdots \det(X_k) \cdot \det(B).$$

It remains to compute $\det(X_1), \ldots, \det(X_k)$, and $\det(B)$. Since each X_i is an elementary matrix, we can compute $\det(X_i)$ using (ii), (iii), or (iv) in Definition 5.11. Finally, the matrix B is either the identity I_m, which has determinant $\det(B) = 1$, or a matrix with an all-zero row, which has determinant $\det(B) = 0$.

Example 5.13. To compute the determinant of

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix},$$

we write

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Let us write

$$A = X_1 \cdot X_2 \cdot B,$$

where

$$X_1 = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}, \quad X_2 = \begin{bmatrix} 1 & 2 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

We have

$$\det(A) = \det(X_1) \cdot \det(X_2) \cdot \det(B).$$

From Definition 5.11, we find that

$$\det(X_1) = \det(X_2) = \det(B) = 1,$$

and thus

$$\det(A) = 1 \cdot 1 \cdot 1 = 1.$$

Example 5.14. Let us compute the determinant of

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

We can write

$$A = \begin{bmatrix} 1 & 0 & 5 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Therefore,

$$\det(A) = 1 \cdot (-2) \cdot (-1) \cdot 1 = 2.$$

Now suppose that A is an invertible matrix. We have

$$A \cdot A^{-1} = I_m.$$

Therefore,

$$\det(A) \cdot \det(A^{-1}) = \det(I_m) = 1.$$

We obtain two things: first,

$$\det(A) \neq 0,$$

and second,

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

We have the following proposition.

Proposition 5.6. *Let* $A \in \mathbb{R}^{m\times m}$ *be any square matrix. There are two cases.*

(i) *If* A *is non-invertible, then* $\det(A) = 0$.

(ii) *If* A *is invertible, then* $\det(A) \neq 0$. *Moreover, in this case,*

$$\det(A^{-1}) = \frac{1}{\det(A)}.$$

So from the above proposition, we have that a square matrix $A \in \mathbb{R}^{m\times m}$ is invertible if and only if $\det(A) \neq 0$.

5.5. **Problems and Solutions.**

Problem 1. Compute the following products.

(a)
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}.$$

(b)
$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

(c)
$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 & 1 \\ 0 & 1 & 2 \\ 1 & 2 & 3 \end{bmatrix}.$$

Problem 2. Consider the following matrices.

$$A_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 4 \\ 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad A_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 \\ 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \qquad A_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$
$$A_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \qquad A_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(a) For each $i = 1, 2, 3, 4$, there is an elementary row operation which transforms A_i to A_{i+1}. Let E_i denote the corresponding elementary matrix so that $E_i \cdot A_i = A_{i+1}$. Find E_1, E_2, E_3, E_4.

(b) Compute E_1^{-1}, E_2^{-1}, E_3^{-1}, E_4^{-1}.

(c) Compute $E_1^{-1} \cdot E_2^{-1}$.

(d) Compute $(E_1^{-1} \cdot E_2^{-1}) \cdot E_3^{-1}$.

(e) Compute $(E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1}) \cdot E_4^{-1}$. Verify that

$$E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1} = A_1.$$

(f) Compute $E_4 \cdot E_3$.

(g) Compute $(E_4 \cdot E_3) \cdot E_2$.

(h) Determine A_1^{-1}.

(i) Compute $\det(E_1)$ and $\det(E_1^{-1})$.

(j) Compute $\det(E_2)$ and $\det(E_2^{-1})$.

(k) Compute $\det(E_3)$ and $\det(E_3^{-1})$.

(l) Compute $\det(E_4)$ and $\det(E_4^{-1})$.

(m) Compute $\det(A_1)$ and $\det(A_1^{-1})$.

(n) Compute $\det(A_2)$, $\det(A_3)$, $\det(A_4)$, and $\det(A_5)$.

Problem 3. Let

$$A := \begin{bmatrix} 0 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 \\ 5 & -5 & 0 & 0 \\ -10 & 10 & 0 & -7 \end{bmatrix}.$$

(a) Find elementary matrices X_1, X_2, X_3, X_4, X_5 and a matrix B in reduced row echelon form such that

$$A = X_1 \cdot X_2 \cdot X_3 \cdot X_4 \cdot X_5 \cdot B.$$

(b) Evaluate $\det(A)$. Determine whether A is singular.

Problem 4. Let

$$C := \begin{bmatrix} 2 & 0 & 6 & 1 & 0 \\ 3 & 0 & 9 & 1 & 1 \end{bmatrix}.$$

Find elementary matrices Y_1, Y_2, Y_3, Y_4, Y_5 and a matrix D in reduced row echelon form such that

$$C = Y_1 \cdot Y_2 \cdot Y_3 \cdot Y_4 \cdot Y_5 \cdot D.$$

Problem 5. Let

$$M := \begin{bmatrix} -1 & 5 \\ -7 & -11 \end{bmatrix}.$$

(a) Find elementary matrices Z_1, Z_2, Z_3, Z_4 such that

$$M = Z_1 \cdot Z_2 \cdot Z_3 \cdot Z_4.$$

(b) For the elementary matrices Z_1, Z_2, Z_3, Z_4 from part (a), compute $\det(Z_1)$, $\det(Z_2)$, $\det(Z_3)$, and $\det(Z_4)$.

(c) For the elementary matrices Z_1, Z_2, Z_3, Z_4 from part (a), compute

$$Z_4^{-1} \cdot Z_3^{-1} \cdot Z_2^{-1} \cdot Z_1^{-1}.$$

(d) Recall that there is the following formula for finding the inverse of an invertible 2×2-matrix: as long as $ad - bc \neq 0$, we have

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

Use this formula to find the inverse of M. Is the result the same as your answer from part (c)?

(e) Using your answers from part (b), evaluate $\det(M)$.

Solution to Problem 1. (a) The left matrix is the elementary matrix corresponding to swapping R_2 and R_3:

$$R_2 \rightleftharpoons R_3.$$

Thus, the product is

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 \\ 5 & 6 & 7 & 8 \end{bmatrix}.$$

(b) The left matrix is the elementary matrix corresponding to multiplying R_3 by -4:

$$R_3 \mapsto (-4) \cdot R_3.$$

Thus, the product is

$$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \\ 7 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -20 & -24 \\ 7 & 8 \end{bmatrix}.$$

(c) The left matrix is the elementary matrix corresponding to adding $2 \cdot R_2$ to R_3:

$$R_3 \mapsto R_3 + 2 \cdot R_2.$$

Therefore, the product is

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

□

Solution to Problem 2. (a) The four steps of elementary row operations are

1) Multiplying R_2 by $1/4$.

2) Adding R_4 to R_2: $R_2 \mapsto R_2 + R_4$.

3) Swapping R_2 and R_4.

4) Multiplying R_3 by $1/5$.

Therefore,

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and

$$E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(b) We do the reverse operation in each case. We have

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

and

$$E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(c)

$$E_1^{-1} \cdot E_2^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & -4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

(d)

$$(E_1^{-1} \cdot E_2^{-1}) \cdot E_3^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 4 & 0 & -4 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

(e)

$$(E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1}) \cdot E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 4 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & -4 & 0 & 4 \\ 0 & 0 & 5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix},$$

which is equal to A_1.

(f)

$$E_4 \cdot E_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/5 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

(g)

$$(E_4 \cdot E_3) \cdot E_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}.$$

(h) Since $A_1^{-1} = E_4 \cdot E_3 \cdot E_2 \cdot E_1$, we have

$$A_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1/4 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1/5 & 0 \\ 0 & 1/4 & 0 & 1 \end{bmatrix}.$$

(i)

$$\det(E_1) = 1/4 \qquad \text{and} \qquad \det(E_1^{-1}) = 4.$$

(j)

$$\det(E_2) = 1 \qquad \text{and} \qquad \det(E_2^{-1}) = 1.$$

(k)

$$\det(E_3) = -1 \qquad \text{and} \qquad \det(E_3^{-1}) = -1.$$

(l)

$$\det(E_4) = 1/5 \qquad \text{and} \qquad \det(E_4^{-1}) = 5.$$

(m) From $A_1 = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1}$, we find

$$\det(A_1) = \det(E_1^{-1}) \cdot \det(E_2^{-1}) \cdot \det(E_3^{-1}) \cdot \det(E_4^{-1}) = 4 \cdot 1 \cdot (-1) \cdot 5 = -20.$$

Then

$$\det(A_1^{-1}) = \frac{1}{\det(A_1)} = -\frac{1}{20}.$$

(n) We have

$$\det(A_2) = \det(E_1) \cdot \det(A_1) = (1/4) \cdot (-20) = -5,$$
$$\det(A_3) = \det(E_2) \cdot \det(A_2) = 1 \cdot (-5) = -5,$$
$$\det(A_4) = \det(E_3) \cdot \det(A_4) = (-1) \cdot (-5) = 5,$$

and

$$\det(A_5) = \det(E_4) \cdot \det(A_5) = (1/5) \cdot 5 = 1.$$

$\square$

Solution to Problem 3. Note! Since there are many ways to do Gaussian elimination, the matrices $Y_1, Y_2, \dots, Y_5$ that you find might be different (and it's okay as long as you perform Gaussian elimination correctly)!

(a) We do Gaussian elimination:

$$\begin{bmatrix} 0 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 \\ 5 & -5 & 0 & 0 \\ -10 & 10 & 0 & -7 \end{bmatrix} \xmapsto{R_4 \mapsto R_4 + 2R_3} \begin{bmatrix} 0 & 0 & 0 & 8 \\ 0 & 0 & 1 & 0 \\ 5 & -5 & 0 & 0 \\ 0 & 0 & 0 & -7 \end{bmatrix} \xmapsto{R_1 \rightleftharpoons R_3} \begin{bmatrix} 5 & -5 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & -7 \end{bmatrix}$$
$$\xmapsto{R_1 \mapsto (1/5) \cdot R_1} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 8 \\ 0 & 0 & 0 & -7 \end{bmatrix} \xmapsto{R_3 \mapsto (1/8) \cdot R_3} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -7 \end{bmatrix}$$
$$\xmapsto{R_4 \mapsto R_4 + 7R_3} \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Note that the last matrix is in reduced row echelon form. Let us call it B.

Therefore, we have the following elementary matrices

$$E_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 2 & 1 \end{bmatrix}, \qquad E_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_3 = \begin{bmatrix} 1/5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad E_4 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1/8 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

and

$$E_5 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 7 & 1 \end{bmatrix}$$

with the property that

$$E_4 \cdot E_3 \cdot E_3 \cdot E_1 \cdot A = B$$

Hence,

$$A = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1} \cdot B,$$

where

$$E_1^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -2 & 1 \end{bmatrix}, \qquad E_2^{-1} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_3^{-1} = \begin{bmatrix} 5 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \qquad E_4^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 8 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix},$$

$$E_5^{-1} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -7 & 1 \end{bmatrix}, \qquad \text{and} \qquad B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

So we can let $X_1 = E_1^{-1}$, $X_2 = E_2^{-1}$, $X_3 = E_3^{-1}$, and $X_4 = E_4^{-1}$ to complete the task.

(b) Note that the matrix B has an all-zero row. Therefore,

$$\det(A) = \det(B) = 0.$$

The matrix A is singular. □

Solution to Problem 4. Note! Since there are many ways to do Gaussian elimination, the matrices $Y_1, Y_2, \ldots, Y_5$ that you find might be different (and it's okay as long as you perform Gaussian elimination correctly)!

We proceed similarly to the previous problem.

$$\begin{bmatrix} 2 & 0 & 6 & 1 & 0 \\ 3 & 0 & 9 & 1 & 1 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - (3/2)\cdot R_1} \begin{bmatrix} 2 & 0 & 6 & 1 & 0 \\ 0 & 0 & 0 & -1/2 & 1 \end{bmatrix} \xmapsto{R_2 \mapsto (-2)\cdot R_2} \begin{bmatrix} 2 & 0 & 6 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$
$$\xmapsto{R_1 \mapsto R_1 - R_2} \begin{bmatrix} 2 & 0 & 6 & 0 & 2 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix} \xmapsto{R_1 \mapsto (1/2)\cdot R_1} \begin{bmatrix} 1 & 0 & 3 & 0 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix} =: D.$$

We have

$$E_1 = \begin{bmatrix} 1 & 0 \\ -3/2 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}, \quad E_3 = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad E_4 = \begin{bmatrix} 1/2 & 0 \\ 0 & 1 \end{bmatrix}.$$

So we let

$$Y_1 = E_1^{-1} = \begin{bmatrix} 1 & 0 \\ 3/2 & 1 \end{bmatrix}, \quad Y_2 = E_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -1/2 \end{bmatrix},$$
$$Y_3 = E_3^{-1} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad Y_4 = E_4^{-1} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix},$$

and obtain that

$$C = Y_1 \cdot Y_2 \cdot Y_3 \cdot Y_4 \cdot D.$$

□

Solution to Problem 5. Note! Since there are many ways to do Gaussian elimination, the matrices Z_1, Z_2, Z_3, Z_4 that you find might be different (and it's okay as long as you perform Gaussian elimination correctly)!

(a) We proceed similarly to the previous problems.

$$\begin{bmatrix} -1 & 5 \\ -7 & -11 \end{bmatrix} \xmapsto{R_2 \mapsto R_2 - 7\cdot R_1} \begin{bmatrix} -1 & 5 \\ 0 & -46 \end{bmatrix} \xmapsto{R_2 \mapsto (-1/46)\cdot R_2} \begin{bmatrix} -1 & 5 \\ 0 & 1 \end{bmatrix}$$
$$\xmapsto{R_1 \mapsto (-1)\cdot R_1} \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix} \xmapsto{R_1 \mapsto R_1 + 5\cdot R_2} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2.$$

Write

$$E_1 = \begin{bmatrix} 1 & 0 \\ -7 & 1 \end{bmatrix}, \quad E_2 = \begin{bmatrix} 1 & 0 \\ 0 & -1/46 \end{bmatrix}, \quad E_3 = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \text{and} \quad E_4 = \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix}.$$

We let

$$Z_1 = E_1^{-1} = \begin{bmatrix} 1 & 0 \\ 7 & 1 \end{bmatrix},$$

$$Z_2 = E_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & -46 \end{bmatrix},$$

$$Z_3 = E_3^{-1} = \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$Z_4 = E_4^{-1} = \begin{bmatrix} 1 & -5 \\ 0 & 1 \end{bmatrix}.$$

Since

$$E_4 \cdot E_3 \cdot E_2 \cdot E_1 \cdot M = I_2,$$

we have

$$M = Z_1 \cdot Z_2 \cdot Z_3 \cdot Z_4.$$

(b) We have

$$\det(Z_1) = \frac{1}{\det(E_1)} = \frac{1}{1} = 1,$$

$$\det(Z_2) = \frac{1}{\det(E_2)} = \frac{1}{-1/46} = -46,$$

$$\det(Z_3) = \frac{1}{\det(E_3)} = \frac{1}{-1} = -1,$$

and

$$\det(Z_4) = \frac{1}{\det(E_4)} = \frac{1}{1} = 1.$$

(c) We have

$$\begin{aligned} Z_4^{-1} \cdot Z_3^{-1} \cdot Z_2^{-1} \cdot Z_1^{-1} = E_4 \cdot E_3 \cdot E_2 \cdot E_1 &= \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 0 & -1/46 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ -7 & 1 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 \\ 7/46 & -1/46 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 5 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -1 & 0 \\ 7/46 & -1/46 \end{bmatrix} = \begin{bmatrix} -11/46 & -5/46 \\ 7/46 & -1/46 \end{bmatrix} \end{aligned}$$

(d) Using the formula, we find

$$M^{-1} = \frac{1}{46} \cdot \begin{bmatrix} -11 & -5 \\ 7 & -1 \end{bmatrix} = \begin{bmatrix} -11/46 & -5/46 \\ 7/46 & -1/46 \end{bmatrix},$$

which agrees with what we found in part (c).

(e) Using the determinants we computed in part (b), we find

$$\det(M) = \det(Z_1) \cdot \det(Z_2) \cdot \det(Z_3) \cdot \det(Z_4) = 1 \cdot (-46) \cdot (-1) \cdot 1 = 46.$$

$\square$

6. The Sixth Week

6.1. **Block Matrices.** Suppose we have two matrices $A \in \mathbb{R}^{m\times n_1}$ and $B \in \mathbb{R}^{m\times n_2}$ with the same number of rows, but possibly with different numbers of columns. We can put the two matrices together with A on the left of B to create an $m \times (n_1 + n_2)$-matrix. For example, suppose

$$A := \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B := \begin{bmatrix} 5 & 6 & 7 \\ 8 & 9 & 10 \end{bmatrix}.$$

Then we can put A to the left of B and create the following 2×5-matrix

$$\begin{bmatrix} 1 & 2 & 5 & 6 & 7 \\ 3 & 4 & 8 & 9 & 10 \end{bmatrix}.$$

When a matrix is constructed in this way from A and B, the result can be denoted by

$$[A|B].$$

The smaller matrices A and B are referred to as **blocks** of the larger matrix $[A|B]$.

Proposition 6.1. *Let m, n, p, q be positive integers. Suppose that $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times p}$ and $C \in \mathbb{R}^{n\times q}$. Then*

$$A \cdot [B|C] = [A \cdot B|A \cdot C].$$

Proof. This follows directly from the definition of matrix multiplication. □

Example 6.1. Let

$$A := \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad B := \begin{bmatrix} 3 & 5 \\ 2 & 3 \end{bmatrix}, \quad \text{and} \quad C := \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Then

$$A \cdot B = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix} \quad \text{and} \quad A \cdot C = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

Observe that

$$A \cdot [B|C] = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 & 5 & 2 \\ 2 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 3 & 0 \end{bmatrix} = [A \cdot B|A \cdot C].$$

6.2. **A method for computing matrix inverses and determinants.** Recall from the last time that if we start with a square matrix $A \in \mathbb{R}^{n\times n}$ and perform Gaussian elimination to transform A into a matrix $B \in \mathbb{R}^{n\times n}$ in reduced row echelon form, then there are two cases:

(i) If B has an all-zero row, then A is non-invertible.

(ii) If $B = I_n$ is the $n \times n$-identity matrix, then A is invertible.

In the first case where B has an all-zero row, then we have $\det(A) = 0$.

In the second case where $B = I_n$ and A is invertible, we saw that there are elementary matrices $E_1, E_2, \ldots, E_k$ such that

$$E_k E_{k-1} \cdots E_2 E_1 \cdot A = B = I_n,$$

and hence

$$A^{-1} = E_k E_{k-1} \cdots E_2 E_1.$$

Now let us build upon this same idea and consider what happens when we perform Gaussian elimination on the matrix

$$[A \,|\, I_n] \in \mathbb{R}^{n \times 2n}.$$

Suppose that the result is the matrix

$$[B \,|\, C] \in \mathbb{R}^{n \times 2n}$$

in reduced row echelon form. This means that there exist elementary matrices $E_1, E_2, \ldots, E_k$ for which

$$E_k E_{k-1} \cdots E_2 E_1 \cdot [A \,|\, I_n] = [B \,|\, C] \,.$$

Now there are two cases, again.

Case 1. If B has an all-zero row, then from Proposition 6.1 above, we learn that

$$E_k E_{k-1} \cdots E_2 E_1 \cdot A = B,$$

and since B has an all-zero row, we conclude that in this case A is non-invertible and $\det(A) = 0$.

Case 2. If $B = I_n$ is the identity matrix, then again from Proposition 6.1 above, we have

$$E_k E_{k-1} \cdots E_2 E_1 \cdot A = I_n,$$

and thus

$$E_k E_{k-1} \cdots E_2 E_1$$

is the inverse of A. Now notice that Proposition 6.1 also implies that

$$E_k E_{k-1} \cdots E_2 E_1 \cdot I_n = C,$$

and therefore

$$C = A^{-1}.$$

Let us conclude what just happened. We learn that we can start with any square matrix $A \in \mathbb{R}^{n \times n}$. Then we create an $n \times 2n$-matrix $[A|I_n]$. Next we perform Gaussian elimination to transform $[A|I_n]$ into $[B|C]$ in reduced row echelon form, and there are two cases. If B has an all-zero row, then A is non-invertible. On the other hand, if $B = I_n$, then the right $n \times n$ block C of $[B|C]$ is the inverse of A. Furthermore, if we keep track of the value of $\det(E_i)$ in every step of Gaussian elimination, then in the case $B = I_n$, the product of $\det(E_i)$ is the determinant of C, which is $\det(C) = 1/\det(A)$.

Example 6.2. Suppose we start with

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

Is A invertible? What is $\det(A)$?

We consider

$$[A|I_2] = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix}.$$

Perform the following Gaussian elimination.

$$[A|I_2] = \begin{bmatrix} 1 & -1 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_1)=1]{R_2 \mapsto R_2 + R_1} \begin{bmatrix} 1 & -1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \xmapsto[\det(E_2)=1]{R_1 \mapsto R_1 - R_2} \begin{bmatrix} 1 & -1 & 0 & -1 \\ 0 & 0 & 1 & 1 \end{bmatrix}$$

So using the notations from the above discussion, we have

$$B = \begin{bmatrix} 1 & -1 \\ 0 & 0 \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix}.$$

Since B has an all-zero row, the matrix A is non-invertible, and thus $\det(A) = 0$.

Example 6.3. Consider

$$A := \begin{bmatrix} 0 & 1 & 0 \\ 1 & 4 & 1 \\ 0 & 5 & 1 \end{bmatrix}.$$

Is A invertible? What is $\det(A)$?

We have

$$[A|I_3] = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 5 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Swapping R_1 and R_2 yields

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 0 & 0 \\ 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 5 & 1 & 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_1)=-1]{R_1 \rightleftharpoons R_2} \begin{bmatrix} 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 1 \end{bmatrix}.$$

Subtracting $5R_2$ from R_3 yields

$$\begin{bmatrix} 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 5 & 1 & 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_2)=1]{R_3 \mapsto R_3 - 5R_2} \begin{bmatrix} 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -5 & 0 & 1 \end{bmatrix}.$$

Subtracting $4R_2$ from R_1 yields

$$\begin{bmatrix} 1 & 4 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -5 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_3)=1]{R_1 \mapsto R_1 - 4R_2} \begin{bmatrix} 1 & 0 & 1 & -4 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -5 & 0 & 1 \end{bmatrix}.$$

Subtracting R_3 from R_1 yields

$$\begin{bmatrix} 1 & 0 & 1 & -4 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -5 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_4)=1]{R_1 \mapsto R_1 - R_3} \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -5 & 0 & 1 \end{bmatrix}.$$

In the above notations, we have $B = I_3$ and

$$C = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ -5 & 0 & 1 \end{bmatrix}.$$

Since $B = I_3$, we find that A is invertible, and

$$A^{-1} = C = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 0 & 0 \\ -5 & 0 & 1 \end{bmatrix}.$$

Moreover,

$$\det(C) = (-1) \cdot 1 \cdot 1 \cdot 1 = -1,$$

and hence

$$\det(A) = \frac{1}{\det(C)} = \frac{1}{-1} = -1.$$

Example 6.4. Consider

$$M := \begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}.$$

Is M invertible? What is $\det(M)$?

We have

$$[M|I_2] = \begin{bmatrix} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{bmatrix}.$$

Subtracting $2R_1$ from R_2 yields

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 4 & 5 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_1)=1]{R_2 \mapsto R_2 - 2R_1} \begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{bmatrix}.$$

Multiplying R_2 with -1 yields

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & -1 & -2 & 1 \end{bmatrix} \xmapsto[\det(E_2)=-1]{R_2 \mapsto (-1)\cdot R_2} \begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & -1 \end{bmatrix}.$$

Subtracting $3R_2$ from R_1 yields

$$\begin{bmatrix} 2 & 3 & 1 & 0 \\ 0 & 1 & 2 & -1 \end{bmatrix} \xmapsto[\det(E_3)=1]{R_1 \mapsto R_1 - 3R_2} \begin{bmatrix} 2 & 0 & -5 & 3 \\ 0 & 1 & 2 & -1 \end{bmatrix}.$$

Multiplying R_1 with $1/2$ yields

$$\begin{bmatrix} 2 & 0 & -5 & 3 \\ 0 & 1 & 2 & -1 \end{bmatrix} \xmapsto[\det(E_4)=1/2]{R_1 \mapsto (1/2)\cdot R_1} \begin{bmatrix} 1 & 0 & -5/2 & 3/2 \\ 0 & 1 & 2 & -1 \end{bmatrix}.$$

Therefore, M is invertible, and

$$M^{-1} = \begin{bmatrix} -5/2 & 3/2 \\ 2 & -1 \end{bmatrix}.$$

Moreover,

$$\det(M^{-1}) = 1 \cdot (-1) \cdot 1 \cdot (1/2) = -\frac{1}{2},$$

and hence

$$\det(M) = -2.$$

6.3. **Rule of Sarrus.** In this subsection, we discuss explicit formulas to compute the determinant of a *small* square matrix.

First, if a matrix A is 1×1, say

$$A = [a] \in \mathbb{R}^{1\times 1},$$

then it is not hard[18] to see that

$$\det(A) = a.$$

Second, if a matrix A is 2×2, say

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

[18]Why does it hold that $\det([a]) = a$? If this is not clear to the reader, perhaps it might be worth some time thinking about it.

then we have the explicit formula

$$\det(A) = ad - bc.$$

Third, if a matrix A is 3×3, say

$$A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix},$$

then the **rule of Sarrus** gives a mnemonic to compute the determinant as follows. Copy the two leftmost columns of A and paste them to the right of A so that we have a 3×5 array:

$$\begin{array}{ccc|cc} a & b & c & a & b \\ d & e & f & d & e \\ g & h & i & g & h \end{array}.$$

Now if we start at the top-leftmost entry (a) and go in the $45°$ southeast direction, we meet e and i. Similarly, if we start at the second entry of the first row (b), we meet f and g. If we start at c, we meet d and h. The rule says that first we have aei, bfg, and cdh.

Next, we start from the bottom row and go in the $45°$ northeast direction instead. We then have gec, hfa, and idb.

The rule of Sarrus says

$$\det(A) = aei + bfg + cdh - gec - hfa - idb.$$

Example 6.5. Let us compute

$$\det \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 0 \\ 0 & 5 & 6 \end{bmatrix}.$$

Write the array

$$\begin{array}{ccc|cc} 0 & 1 & 2 & 0 & 1 \\ 3 & 4 & 0 & 3 & 4 \\ 0 & 5 & 6 & 0 & 5 \end{array}.$$

The rule of Sarrus says that the determinant is

$$0 \cdot 4 \cdot 6 + 1 \cdot 0 \cdot 0 + 2 \cdot 3 \cdot 5 - 0 \cdot 4 \cdot 2 - 5 \cdot 0 \cdot 0 - 6 \cdot 3 \cdot 1 = 30 - 18 = 12.$$

6.4. **Some special types of matrices.** In this subsection, we discuss some special types of matrices.

Definition 6.6. A square matrix is said to be **diagonal** if all off-diagonal entries are zero.

Example 6.7. The following matrices are diagonal:

$$[0], \quad [1], \quad \begin{bmatrix} -2 & 0 \\ 0 & 7 \end{bmatrix}, \quad \begin{bmatrix} 5 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 3 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 7 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The following matrices are *not* diagonal:

$$\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 9 \\ 0 & 8 & 0 \\ 7 & 0 & 0 \end{bmatrix}.$$

If $a_1, a_2, \ldots, a_n$ are n (not necessarily distinct) real numbers, then we use the notation

$$\operatorname{diag}(a_1, a_2, \ldots, a_n)$$

to denote the $n \times n$ diagonal matrix whose diagonal entries from the top left to the bottom right are

$$a_1, a_2, \ldots, a_n$$

in this order. For example,

$$\operatorname{diag}(0, 3, 0, -4) = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -4 \end{bmatrix}.$$

Proposition 6.2. *If*

$$A = \operatorname{diag}(a_1, a_2, \ldots, a_n),$$

then

$$\det(A) = a_1 \cdot a_2 \cdots a_n.$$

In particular, we have the following.

- *If at least one number in the list $a_1, a_2, \ldots, a_n$ is zero, then A is non-invertible.*
- *If all numbers in the list $a_1, a_2, \ldots, a_n$ are nonzero, then A is invertible.*

Proposition 6.3. *If $A = \operatorname{diag}(a_1, a_2, \ldots, a_n)$ is invertible, then*

$$A^{-1} = \operatorname{diag}\left(a_1^{-1}, a_2^{-1}, \ldots, a_n^{-1}\right).$$

Proof of Propositions 6.2 and 6.3. Consider the diagonal matrix

$$A := \operatorname{diag}(a_1, a_2, \ldots, a_n) \in \mathbb{R}^{n \times n}.$$

If there is some index $i \in \{1, 2, \ldots, n\}$ such that $a_i = 0$, then A has an all-zero row, and therefore, $\det(A) = 0$ and A is non-invertible. In this case, we finish.

For the rest of this proof, suppose that for every $i \in \{1, 2, \ldots, n\}$, the entry a_i is not zero. We consider the matrix

$$X := [A \mid I_n] \in \mathbb{R}^{n \times 2n}.$$

Now let us perform Gaussian elimination on X. The idea is to multiply the i^{th} row by the constant $1/a_i \neq 0$, for $i = 1, 2, \ldots, n$. We let

$$E_i := \operatorname{diag}(1, 1, \ldots, 1, 1/a_i, 1, \ldots, 1) \in \mathbb{R}^{n \times n}$$

be the elementary matrix corresponding to multiplying the i^{th} row by $1/a_i$. Note that

$$E_n E_{n-1} \cdots E_2 E_1 \cdot X = E_n E_{n-1} \cdots E_2 E_1 \cdot [A \mid I_n] = [I_n \mid \operatorname{diag}(1/a_1, 1/a_2, \ldots, 1/a_n)],$$

which is in reduced row echelon form. This shows that

$$A^{-1} = \operatorname{diag}(1/a_1, 1/a_2, \ldots, 1/a_n).$$

Now, to find the determinant of A, note that

$$\det(E_n) \cdot \det(E_{n-1}) \cdots \det(E_2) \cdot \det(E_1) \cdot \det(A) = \det(I_n) = 1.$$

Since for each $i \in \{1, 2, \ldots, n\}$, we have $\det(E_i) = 1/a_i$, we conclude that

$$\det(A) = a_1 \cdot a_2 \cdots a_n,$$

as desired. $\square$

Example 6.8. Note that

$$\begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix} \cdot \begin{bmatrix} 1/3 & 0 \\ 0 & -1/4 \end{bmatrix} = \begin{bmatrix} 1/3 & 0 \\ 0 & -1/4 \end{bmatrix} \cdot \begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and therefore

$$\begin{bmatrix} 1/3 & 0 \\ 0 & -1/4 \end{bmatrix}$$

is the inverse of

$$\begin{bmatrix} 3 & 0 \\ 0 & -4 \end{bmatrix}.$$

Definition 6.9. Let $A \in \mathbb{R}^{n \times n}$ be a square matrix. Let us denote the (i, j)-entry of A by a_{ij}.

(i) A is said to be **upper-triangular** if for every $i, j \in \{1, 2, \ldots, n\}$ with $i > j$, we have $a_{ij} = 0$.

(ii) A is said to be **lower-triangular** if for every $i, j \in \{1, 2, \dots, n\}$ with $i < j$, we have $a_{ij} = 0$.

In other words, a square matrix is *upper-triangular* when all entries *below* the diagonal entries are zero, and a square matrix is *lower-triangular* when all entries *above* the diagonal entries are zero.

We also say that a matrix $A \in \mathbb{R}^{n \times n}$ is **triangular** if it is either lower-triangular or upper-triangular.

Example 6.10. The matrix

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

is upper-triangular, and

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

is lower-triangular.

Here are some observations.

Observation 6.1. We have the following.

- The transpose of an upper-triangular matrix is lower-triangular.
- The transpose of a lower-triangular matrix is upper-triangular.
- A square matrix is diagonal if and only if it is both upper-triangular and lower-triangular.

Proposition 6.4. *We have the following.*

(a) *The product of upper-triangular matrices is upper-triangular.*

(b) *The product of lower-triangular matrices is lower-triangular.*

(c) *An upper-triangular matrix is invertible if and only if all its diagonal entries are nonzero.*

(d) *A lower-triangular matrix is invertible if and only if all its diagonal entries are nonzero.*

(e) *If an upper-triangular matrix is invertible, then its inverse is upper-triangular.*

(f) *If a lower-triangular matrix is invertible, then its inverse is lower-triangular.*

Proposition 6.5. *Let $A \in \mathbb{R}^{n \times n}$ be a triangular matrix. Then* $\det(A)$ *equals the product of the diagonal entries of A:*

$$\det(A) = a_{11} \cdot a_{22} \cdots a_{nn},$$

where a_{ij} denotes the (i, j)-entry of A.

Proof. Let us consider the case where A is upper-triangular. The case when A is lower-triangular is similar.

We distinguish two cases.

Case 1. Suppose that there is some index i such that $a_{ii} = 0$. Let M denote the *largest* index for which $a_{MM} = 0$. This means that for every index j such that $M < j \leq n$, we have $a_{jj} \neq 0$. It is not hard to see that by a sequence of elementary row operations of the type $R_i \mapsto R_i + c \cdot R_j$, we can transform A into a matrix B whose M^{th} row is an all-zero row. Note that these elementary matrices have determinants 1. This shows that

$$\det(A) = \det(B).$$

Since B has an all-zero row, we have that $\det(B) = 0$, and hence $\det(A) = 0$.

Case 2. Suppose that for every index i, we have $a_{ii} \neq 0$. By a sequence of elementary row operations of the type $R_i \mapsto R_i + c \cdot R_j$, we can transform A into

$$B := \operatorname{diag}(a_{11}, a_{22}, \ldots, a_{nn}).$$

In a similar manner to the previous case, we note that

$$\det(A) = \det(B).$$

By Proposition 6.2, we know

$$\det(B) = a_{11} \cdot a_{22} \cdots a_{nn},$$

from which the desired result follows. □

Example 6.11. Let

$$A := \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 0 & 1 & 2 & 3 \\ 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 7 \end{bmatrix} \in \mathbb{R}^{5\times 5}.$$

Let us compute the determinant of A.

Note that A is *almost* upper-triangular, but it is not. The idea is to transform it into an upper-triangular matrix. Let

$$E := \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -4 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

be the elementary matrix corresponding to the elementary row operation $R_4 \mapsto R_4 - 4R_3$. Observe that

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

is now an upper-triangular matrix. Hence, by Proposition 6.5, we find

$$\det(E \cdot A) = 1 \cdot 6 \cdot 1 \cdot (-3) \cdot 7 = -126.$$

Since $\det(E) = 1$, we have $\det(A) = \det(E) \cdot \det(A) = \det(E \cdot A) = -126$.

Exercise 6.12. Let

$$M := \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 7 & 8 & 9 & 10 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 & 11 \end{bmatrix} \in \mathbb{R}^{6\times 6}.$$

Evaluate $\det(M)$.

A solution to this exercise is given in the footnote.[19]

Definition 6.13. Let $A \in \mathbb{R}^{n\times n}$ be a square matrix. Let us denote the (i, j)-entry of A by a_{ij}.

(i) A is said to be **symmetric** if for every $i, j \in \{1, 2, \dots, n\}$, we have $a_{ij} = a_{ji}$.

---

[19]**Solution.** Note that M is *not* a triangular matrix, according to the definition given above. We can, however, transform M into an upper-triangular matrix by swapping rows. Let E_1, E_2, E_3 denote the elementary matrices of swapping $R_1 \rightleftharpoons R_6$, $R_2 \rightleftharpoons R_5$, $R_3 \rightleftharpoons R_4$, respectively. We find that

$$E_3 \cdot E_2 \cdot E_1 \cdot M = \begin{bmatrix} 6 & 7 & 8 & 9 & 10 & 11 \\ 0 & 1 & 2 & 3 & 4 & 5 \\ 0 & 0 & 7 & 8 & 9 & 10 \\ 0 & 0 & 0 & 4 & 5 & 6 \\ 0 & 0 & 0 & 0 & 2 & 3 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

which is now an upper-triangular matrix. We have

$$\det(E_3 \cdot E_2 \cdot E_1 \cdot M) = 6 \cdot 1 \cdot 7 \cdot 4 \cdot 2 \cdot 1 = 336.$$

Note that $\det(E_1) = \det(E_2) = \det(E_3) = -1$. Thus, $\det(M) = -336$. ■

(i) A is said to be **antisymmetric** (or, synonymously, **skew-symmetric**) if for every $i, j \in \{1, 2, \ldots, n\}$, we have $a_{ij} = -a_{ji}$.

Example 6.14. The matrix

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

is symmetric, while the matrix

$$\begin{bmatrix} 0 & 2 & -3 \\ -2 & 0 & 5 \\ 3 & -5 & 0 \end{bmatrix}$$

is antisymmetric.

Exercise 6.15. Is the matrix

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

symmetric? Is it antisymmetric?

A solution to this exercise is given in the footnote.[20]

Observation 6.2. We have the following.

- If a square matrix is antisymmetric, then all its diagonal entries are zero.
- A square matrix A is symmetric if and only if $A = A^\intercal$.
- A square matrix A is antisymmetric if and only if $A = -A^\intercal$.
- A square matrix is the zero matrix if and only if it is both symmetric and antisymmetric.

Proposition 6.6. *Let $A \in \mathbb{R}^{n \times n}$ be any square matrix. Then A can be* uniquely *written as a sum of a symmetric matrix and an antisymmetric matrix.*

Proof. Existence. Note that

$$A = \left(\frac{A + A^\intercal}{2}\right) + \left(\frac{A - A^\intercal}{2}\right).$$

If we write

$$B := \frac{A + A^\intercal}{2},$$

and

$$C := \frac{A - A^\intercal}{2},$$

[20] *Solution.* For each $i, j \in \{1, 2\}$, let a_{ij} denote the (i, j)-entry of A. The matrix A is *not* symmetric, because $a_{12} = 2$ is not equal to $a_{21} = -2$. The matrix A is *not* antisymmetric, because $a_{11} = 1$ is not equal to $-a_{11} = -1$. ■

then we have $A = B + C$. Now observe that

$$B^{\mathrm{T}} = \frac{A^{\mathrm{T}} + (A^{\mathrm{T}})^{\mathrm{T}}}{2} = \frac{A + A^{\mathrm{T}}}{2} = B,$$

and hence B is symmetric. Also,

$$C^{\mathrm{T}} = \frac{A^{\mathrm{T}} - (A^{\mathrm{T}})^{\mathrm{T}}}{2} = -\frac{A - A^{\mathrm{T}}}{2} = -C,$$

and hence C is antisymmetric.

Therefore, $A = B + C$ is the desired decomposition of A.

Uniqueness. Now we show that such a decomposition is unique. Suppose that $A = B + C$ and $A = B' + C'$, where B and B' are symmetric and C and C' are antisymmetric. We then have that $(B - B') + (C - C') = A - A = \mathbf{0}$ is the zero matrix. From this we can write $B - B' = C' - C$. Note that $B - B'$ is symmetric, while $C' - C$ is antisymmetric, and therefore, the matrix $B - B' = C' - C$ is simultaneously symmetric and antisymmetric. From the above observation, this implies that $B - B' = C' - C = \mathbf{0}$, whence $B = B'$ and $C = C'$. This shows that the decomposition $A = B' + C'$ is actually the same as the decomposition $A = B + C$: the decomposition is unique. □

When a matrix A is decomposed as $A = B + C$, where B is symmetric and C is antisymmetric, we say that B is the **symmetric part** of A, and that C is the **antisymmetric part** of A.

Example 6.16. Consider

$$A := \begin{bmatrix} 0 & 5 \\ -3 & -4 \end{bmatrix}.$$

Let us consider

$$B := \frac{A + A^{\mathrm{T}}}{2} = \begin{bmatrix} 0 & 1 \\ 1 & -4 \end{bmatrix},$$

and

$$C := \frac{A - A^{\mathrm{T}}}{2} = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}.$$

Note that B is symmetric, C is antisymmetric, and $A = B + C$. Hence,

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

is the symmetric part of A, and

$$C = \begin{bmatrix} 0 & 4 \\ -4 & 0 \end{bmatrix}$$

is the antisymmetric part of A.

6.5. **Problems and Solutions.**

Problem 1. Let

$$A = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix}, \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 0 \end{bmatrix}, \quad \text{and} \quad C = \begin{bmatrix} 2 & 0 \\ 3 & -1 \end{bmatrix}.$$

(a) Write out the matrix $[B|C] \in \mathbb{R}^{2\times 4}$ explicitly.

(b) Compute $A \cdot B$.

(c) Compute $A \cdot C$.

(d) Compute $A \cdot [B|C]$. Verify that

$$A \cdot [B|C] = [A \cdot B|A \cdot C].$$

Problem 2. Consider the matrix

$$D = \begin{bmatrix} 0 & 1 & -1 \\ 3 & 2 & 1 \\ -1 & 0 & -1 \end{bmatrix}.$$

(a) Write out the matrix $[D|I_3]$ explicitly. Perform a sequence of elementary row operations on $[D|I_3]$ to transform it into a matrix $[G|H]$ in reduced row echelon form, where $G, H \in \mathbb{R}^{3\times 3}$.

(b) Is D invertible? What is $\det(D)$? If D is invertible, what is D^{-1}?

(c) Evaluate $\operatorname{tr}(H)$ and $\det(H)$. (*Hint.* How is $\det(H)$ related to the product of determinants of the elementary matrices obtained from the Gaussian elimination?)

Problem 3. Consider the matrix

$$K = \begin{bmatrix} 0 & 1 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 1 & 0 & 3 & -3 \\ 0 & 0 & 6 & 1 \end{bmatrix}.$$

(a) Write out the matrix $[K|I_4]$ explicitly. Perform a sequence of elementary row operations on $[K|I_4]$ to transform it into a matrix $[L|M]$ in reduced row echelon form, where $L, M \in \mathbb{R}^{4\times 4}$.

Hint. If you do row operations correctly, the matrix M should have the following features:

- there should be exactly 3 zero entries in M,
- there should be exactly 6 negative entries in M, and

– the sum of the $(1,2)$-entry and the $(3,1)$-entry of M should be 4.

(b) Is K invertible? What is $\det(K)$? If K is invertible, what is K^{-1}?

(c) Evaluate $\operatorname{tr}(M)$ and $\det(M)$.

Problem 4.

(a) Use the rule of Sarrus to compute

$$\det \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 1 & 1 & 4 \end{bmatrix}.$$

(b) Determine the set of all real numbers x for which the matrix

$$\begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 1 & 1 & x \end{bmatrix}$$

is singular.

Problem 5. For each of the following matrices, determine (i) its trace, (ii) its determinant, (iii) whether it is invertible, and if it is invertible, also write out its inverse.

(a)

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & 2/3 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

(b)

$$\begin{bmatrix} 4/7 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3/7 \end{bmatrix}$$

(c) $\operatorname{diag}(-1/3, 1/2, -1/6)$

(d) $\operatorname{diag}(-2, -1, 0, 1, 2)$

(e) the identity matrix I_2.

Problem 6. It was discussed in class that the product of two upper-triangular matrices is an upper-triangular matrix. Let us try an example.

Let

$$N = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} \quad \text{and} \quad P = \begin{bmatrix} 1 & -1 & -3 \\ 0 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

(a) Compute $N \cdot P$. Is the product upper-triangular?

(b) Compute $P \cdot N$. Is the product upper-triangular?

(c) Let

$$X := N \cdot P - P \cdot N.$$

Compute X. Is X upper-triangular? Is X lower-triangular? Is X singular? Explain how you obtain the answers.

Problem 7.

(a) Evaluate

$$\det\left(\begin{bmatrix} 2 & 0 & 3 & -1 & 9 \\ 0 & 1 & 4 & 2 & 5 \\ 0 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} -5 & 1 & -1 & -1 & 4 \\ 0 & 3 & 2 & 1 & -4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}\right).$$

Hint. You should be able to compute the determinant *without* having to multiply the two matrices explicitly!

(b) Evaluate

$$\operatorname{tr}\left(\begin{bmatrix} 2 & 0 & 3 & -1 & 9 \\ 0 & 1 & 4 & 2 & 5 \\ 0 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} -5 & 1 & -1 & -1 & 4 \\ 0 & 3 & 2 & 1 & -4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}\right).$$

Problem 8.

(a) Write the matrix

$$Q := \begin{bmatrix} 1 & -2 & 0 \\ -4 & 2 & 3 \\ 8 & -7 & 3 \end{bmatrix}$$

explicitly as the sum of its symmetric part and its antisymmetric part.

(b) Consider the matrix

$$R := \begin{bmatrix} -1 & x \\ 5 & 6 \end{bmatrix}.$$

Suppose that

$$R = S + T,$$

where S is a symmetric matrix, and T is an antisymmetric matrix. Determine the set of all real numbers x for which the $(1,1)$-entry of $S \cdot T$ is 4.

Solution to Problem 1. (a)

$$[B|C] = \begin{bmatrix} 2 & 3 & 2 & 0 \\ 1 & 0 & 3 & -1 \end{bmatrix}.$$

(b) The product is

$$A \cdot B = \begin{bmatrix} -1 & 3 \\ 2 & -3 \end{bmatrix}.$$

(c) The product is

$$A \cdot C = \begin{bmatrix} -7 & 3 \\ 10 & -4 \end{bmatrix}.$$

(d) We have

$$A \cdot [B|C] = \begin{bmatrix} 1 & -3 \\ -1 & 4 \end{bmatrix} \cdot \begin{bmatrix} 2 & 3 & 2 & 0 \\ 1 & 0 & 3 & -1 \end{bmatrix} = \begin{bmatrix} -1 & 3 & -7 & 3 \\ 2 & -3 & 10 & -4 \end{bmatrix},$$

which is the same as $[A \cdot B | A \cdot C]$. □

Solution to Problem 2. (a) The matrix $[D|I_3]$ is

$$[D|I_3] = \begin{bmatrix} 0 & 1 & -1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{bmatrix}.$$

Now we perform Gaussian elimination as follows.

$$\begin{bmatrix} 0 & 1 & -1 & 1 & 0 & 0 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_1)=-1]{R_1 \rightleftharpoons R_3} \begin{bmatrix} -1 & 0 & -1 & 0 & 0 & 1 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 \end{bmatrix}$$

$$\xmapsto[\det(E_2)=-1]{R_1 \mapsto (-1)\cdot R_1} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 3 & 2 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & 1 & 0 & 0 \end{bmatrix}$$

$$\xmapsto[\det(E_3)=1]{R_2 \mapsto R_2 - 3\cdot R_1} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 2 & -2 & 0 & 1 & 3 \\ 0 & 1 & -1 & 1 & 0 & 0 \end{bmatrix}$$

$$\xmapsto[\det(E_4)=1]{R_3 \mapsto R_3 - (1/2)\cdot R_2} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 2 & -2 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 & -1/2 & -3/2 \end{bmatrix}$$

$$\xmapsto[\det(E_5)=1/2]{R_2 \mapsto (1/2)\cdot R_2} \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 & 1/2 & 3/2 \\ 0 & 0 & 0 & 1 & -1/2 & -3/2 \end{bmatrix}.$$

Note that the final matrix is in reduced row echelon form, so we have finished the Gaussian elimination.

In particular, we find

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

and

$$H = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1/2 & 3/2 \\ 1 & -1/2 & -3/2 \end{bmatrix}.$$

(b) Since G has an all-zero row, we conclude that D is non-invertible, and

$$\det(D) = \det(G) = 0.$$

The matrix D has no inverse, because it is non-invertible.

(c) We have

$$\operatorname{tr}(H) = 0 + \frac{1}{2} + \left(-\frac{3}{2}\right) = -1.$$

To find $\det(H)$, note that from the elementary row operations

$$E_5 \cdot E_4 \cdot E_3 \cdot E_2 \cdot E_1 \cdot I_3 = H.$$

This implies that

$$\begin{aligned} \det(H) &= \det(E_5) \cdot \det(E_4) \cdot \det(E_3) \cdot \det(E_2) \cdot \det(E_1) \cdot \det(I_3) \\ &= \frac{1}{2} \cdot 1 \cdot 1 \cdot (-1) \cdot (-1) \cdot 1 = \frac{1}{2}. \end{aligned}$$

□

Solution to Problem 3. (a) The matrix $[K|I_4]$ is

$$[K|I_4] = \begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 3 & -3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 6 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now we do Gaussian elimination as follows.

$$\begin{bmatrix} 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 3 & -3 & 0 & 0 & 1 & 0 \\ 0 & 0 & 6 & 1 & 0 & 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_1)=-1]{R_1 \rightleftharpoons R_3} \begin{bmatrix} 1 & 0 & 3 & -3 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 6 & 1 & 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\xrightarrow[\det(E_2)=-1]{R_2\mapsto(-1)\cdot R_2}\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&1&0&1&1&0&0&0\\0&0&6&1&0&0&0&1\end{bmatrix}$$

$$\xrightarrow[\det(E_3)=1]{R_3\mapsto R_3-R_2}\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&0&1&1&1&1&0&0\\0&0&6&1&0&0&0&1\end{bmatrix}$$

$$\xrightarrow[\det(E_4)=1]{R_4\mapsto R_4-6\cdot R_3}\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&0&1&1&1&1&0&0\\0&0&0&-5&-6&-6&0&1\end{bmatrix}$$

$$\xrightarrow[\det(E_5)=-1/5]{R_4\mapsto(-1/5)\cdot R_4}\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&0&1&1&1&1&0&0\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}.$$

This final matrix is now in row echelon form, but not yet in *reduced* row echelon form. We continue.

$$\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&0&1&1&1&1&0&0\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}\xrightarrow[\det(E_6)=1]{R_3\mapsto R_3-R_4}\begin{bmatrix}1&0&3&-3&0&0&1&0\\0&1&-1&0&0&-1&0&0\\0&0&1&0&-1/5&-1/5&0&1/5\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}$$

$$\xrightarrow[\det(E_7)=1]{R_1\mapsto R_1+3\cdot R_4}\begin{bmatrix}1&0&3&0&18/5&18/5&1&-3/5\\0&1&-1&0&0&-1&0&0\\0&0&1&0&-1/5&-1/5&0&1/5\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}$$

$$\xrightarrow[\det(E_8)=1]{R_2\mapsto R_2+R_3}\begin{bmatrix}1&0&3&0&18/5&18/5&1&-3/5\\0&1&0&0&-1/5&-6/5&0&1/5\\0&0&1&0&-1/5&-1/5&0&1/5\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}$$

$$\xrightarrow[\det(E_9)=1]{R_1\mapsto R_1-3\cdot R_3}\begin{bmatrix}1&0&0&0&21/5&21/5&1&-6/5\\0&1&0&0&-1/5&-6/5&0&1/5\\0&0&1&0&-1/5&-1/5&0&1/5\\0&0&0&1&6/5&6/5&0&-1/5\end{bmatrix}.$$

We have finished Gaussian elimination. In particular, we find that $L = I_4$, and

$$M = \begin{bmatrix} 21/5 & 21/5 & 1 & -6/5 \\ -1/5 & -6/5 & 0 & 1/5 \\ -1/5 & -1/5 & 0 & 1/5 \\ 6/5 & 6/5 & 0 & -1/5 \end{bmatrix}.$$

(b) From the Gaussian elimination, we find that

$$E_9 \cdot E_8 \cdots E_2 \cdot E_1 \cdot K = L = I_4.$$

Since L is the identity matrix, the matrix K is invertible. Taking the determinant on both sides of the above equation, we find

$$\det(E_9) \cdot \det(E_8) \cdot \cdots \cdot \det(E_1) \cdot \det(K) = \det(I_4) = 1.$$

Therefore,

$$1 \cdot 1 \cdot 1 \cdot 1 \cdot \left(-\frac{1}{5}\right) \cdot 1 \cdot 1 \cdot (-1) \cdot (-1) \cdot \det(K) = 1,$$

which implies that $\det(K) = -5$.

The inverse of K is

$$K^{-1} = M = \begin{bmatrix} 21/5 & 21/5 & 1 & -6/5 \\ -1/5 & -6/5 & 0 & 1/5 \\ -1/5 & -1/5 & 0 & 1/5 \\ 6/5 & 6/5 & 0 & -1/5 \end{bmatrix}.$$

(c) The trace of M is

$$\operatorname{tr}(M) = \frac{21}{5} + \left(-\frac{6}{5}\right) + 0 + \left(-\frac{1}{5}\right) = \frac{14}{5}.$$

Since M is the inverse of K, the determinant of M is

$$\det(M) = \frac{1}{\det(K)} = \frac{1}{-5} = -\frac{1}{5}.$$

$\square$

Solution to Problem 4. (a) The rule of Sarrus gives

$$\det \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 1 & 1 & 4 \end{bmatrix} = 0 + 6 + (-1) - 0 - 3 - (-8) = 10.$$

(b) Note that the determinant of the matrix is

$$\det \begin{bmatrix} 1 & 2 & 1 \\ -1 & 0 & 3 \\ 1 & 1 & x \end{bmatrix} = 0 + 6 + (-1) - 0 - 3 - (-2x) = 2x + 2,$$

by the rule of Sarrus. Therefore, the determinant is zero if and only if $x = -1$. The set of all such real numbers x is therefore $\{-1\}$. □

Solution to Problem 5. (a) The trace is

$$(-2) + \frac{2}{3} + (-1) = -\frac{7}{3}.$$

The determinant is

$$(-2) \cdot \frac{2}{3} \cdot (-1) = \frac{4}{3}.$$

The matrix is invertible, and its inverse is

$$\begin{bmatrix} -1/2 & 0 & 0 \\ 0 & 3/2 & 0 \\ 0 & 0 & -1 \end{bmatrix}.$$

(b) The trace is

$$\frac{4}{7} + 0 + \frac{3}{7} = 1.$$

The determinant is

$$\frac{4}{7} \cdot 0 \cdot \frac{3}{7} = 0.$$

The matrix is non-invertible.

(c) The trace is

$$\left(-\frac{1}{3}\right) + \frac{1}{2} + \left(-\frac{1}{6}\right) = 0.$$

The determinant is

$$\left(-\frac{1}{3}\right) \cdot \frac{1}{2} \cdot \left(-\frac{1}{6}\right) = \frac{1}{36}.$$

The matrix is invertible, and its inverse is

$$\operatorname{diag}(-3, 2, -6) = \begin{bmatrix} -3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -6 \end{bmatrix}.$$

(d) The trace is

$$(-2) + (-1) + 0 + 1 + 2 = 0.$$

The determinant is
$$(-2) \cdot (-1) \cdot 0 \cdot 1 \cdot 2 = 0.$$
The matrix is non-invertible.

(e) The trace is
$$1 + 1 = 2.$$
The determinant is
$$1 \cdot 1 = 1.$$
The matrix is invertible, and its inverse is itself
$$I_2^{-1} = I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$
$\square$

Solution to Problem 6. (a) The product is
$$N \cdot P = \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 1 & -1 & -3 \\ 0 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -4 & -7 \\ 0 & -6 & 5 \\ 0 & 0 & 4 \end{bmatrix},$$
which is indeed upper-triangular.

(b) The product is
$$P \cdot N = \begin{bmatrix} 1 & -1 & -3 \\ 0 & -2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 & -1 \\ 0 & 3 & 1 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & -2 & -8 \\ 0 & -6 & 0 \\ 0 & 0 & 4 \end{bmatrix},$$
which is indeed upper-triangular.

(c) We have
$$X = N \cdot P - P \cdot N = \begin{bmatrix} 0 & -2 & 1 \\ 0 & 0 & 5 \\ 0 & 0 & 0 \end{bmatrix}.$$
The matrix X is upper-triangular, but not lower-triangular.

Note that since X has an all-zero row, if we perform Gaussian elimination on X and obtain a matrix in reduced row echelon form, the resulting matrix also has an all-zero row. In fact, it is not hard to see that the resulting matrix is
$$\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

This implies that the matrix X is non-invertible: X is singular. $\square$

Solution to Problem 7. Let

$$U := \begin{bmatrix} 2 & 0 & 3 & -1 & 9 \\ 0 & 1 & 4 & 2 & 5 \\ 0 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix},$$

and

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

Consider the upper-triangular matrix U. By using a sequence of row operations of the type

$$R_i \mapsto R_i + c \cdot R_j$$

for certain indices i, j (with $i \neq j$) and a certain real number $c \in \mathbb{R}$, we can transform the matrix U into the diagonal matrix

$$U' := \begin{bmatrix} 2 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix}.$$

Since all of the elementary matrices involved have determinant 1, we find that

$$\det(U) = \det(U').$$

The matrix U' is diagonal, so its determinant is

$$\det(U') = 2 \cdot 1 \cdot (-3) \cdot 1 \cdot (-2) = 12,$$

which gives $\det(U) = 12$.

By a similar argument, we find that

$$\det(V) = (-5) \cdot 3 \cdot 1 \cdot (-1) \cdot 5 = 75.$$

Therefore,

$$\det(U \cdot V) = \det(U) \cdot \det(V) = 12 \cdot 75 = 900.$$

(b) To compute the trace $\mathrm{tr}(U \cdot V)$, we note that it suffices to evaluate only the diagonal entries of $U \cdot V$. We have

$$
\begin{aligned}
U \cdot V &= \begin{bmatrix} 2 & 0 & 3 & -1 & 9 \\ 0 & 1 & 4 & 2 & 5 \\ 0 & 0 & -3 & 0 & 4 \\ 0 & 0 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & -2 \end{bmatrix} \cdot \begin{bmatrix} -5 & 1 & -1 & -1 & 4 \\ 0 & 3 & 2 & 1 & -4 \\ 0 & 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix} \\
&= \begin{bmatrix} -10 & * & * & * & * \\ 0 & 3 & * & * & * \\ 0 & 0 & -3 & * & * \\ 0 & 0 & 0 & -1 & * \\ 0 & 0 & 0 & 0 & -10 \end{bmatrix}
\end{aligned}
$$

where each $*$ denotes some real number. Hence,

$$\mathrm{tr}(U \cdot V) = (-10) + 3 + (-3) + (-1) + (-10) = -21.$$

$\square$

Solution to Problem 8. (a) The symmetric part of Q is

$$\frac{Q + Q^{\mathrm{T}}}{2} = \frac{1}{2} \cdot \begin{bmatrix} 1 & -2 & 0 \\ -4 & 2 & 3 \\ 8 & -7 & 3 \end{bmatrix} + \frac{1}{2} \cdot \begin{bmatrix} 1 & -4 & 8 \\ -2 & 2 & -7 \\ 0 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 4 \\ -3 & 2 & -2 \\ 4 & -2 & 3 \end{bmatrix}.$$

The antisymmetric part of Q is

$$\frac{Q - Q^{\mathrm{T}}}{2} = \frac{1}{2} \cdot \begin{bmatrix} 1 & -2 & 0 \\ -4 & 2 & 3 \\ 8 & -7 & 3 \end{bmatrix} - \frac{1}{2} \cdot \begin{bmatrix} 1 & -4 & 8 \\ -2 & 2 & -7 \\ 0 & 3 & 3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & -4 \\ -1 & 0 & 5 \\ 4 & -5 & 0 \end{bmatrix}.$$

The desired decomposition of Q is

$$\begin{bmatrix} 1 & -2 & 0 \\ -4 & 2 & 3 \\ 8 & -7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & -3 & 4 \\ -3 & 2 & -2 \\ 4 & -2 & 3 \end{bmatrix} + \begin{bmatrix} 0 & 1 & -4 \\ -1 & 0 & 5 \\ 4 & -5 & 0 \end{bmatrix}.$$

(b) We have

$$S = \frac{R + R^{\mathrm{T}}}{2} = \begin{bmatrix} -1 & \frac{x+5}{2} \\ \frac{x+5}{2} & 6 \end{bmatrix} \qquad \text{and} \qquad T = \frac{R - R^{\mathrm{T}}}{2} = \begin{bmatrix} 0 & \frac{x-5}{2} \\ \frac{5-x}{2} & 0 \end{bmatrix}.$$

Therefore,

$$S \cdot T = \begin{bmatrix} \frac{25-x^2}{4} & \frac{5-x}{2} \\ 15 - 3x & \frac{x^2-25}{4} \end{bmatrix}.$$

Hence, the $(1,1)$-entry of $S \cdot T$ is $\frac{25-x^2}{4}$, which is 4 if and only if $x^2 = 9$, which is if and only if $x \in \{-3, 3\}$.

The set of all such real numbers x is $\{-3, 3\}$. $\square$

7. The Seventh Week

7.1. **Elements of $\mathbb{R}^n$ as column vectors.** In this subsection, we discuss different ways to think of elements of $\mathbb{R}^n$. Previously, we used to think of elements in $\mathbb{R}^n$ as *n-tuples* of real numbers. For instance, we can write

$$(1, 2, 3, 3, 3) \in \mathbb{R}^5.$$

The object $(1, 2, 3, 3, 3)$ in $\mathbb{R}^5$ is a finite list of 5 (not necessarily distinct) real numbers.

There is another common way to think about elements of $\mathbb{R}^n$. Sometimes it is convenient to think of $\mathbb{R}^n$ as $\mathbb{R}^{n\times 1}$, the set of $n \times 1$-matrices. Recall that we call these matrices "column vectors". For example, the elements

$$(1, 0, 1), (1, 2, 3), (0, 0, 0) \in \mathbb{R}^3$$

can be written as the *vectors*

$$\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \mathbb{R}^3.$$

There are special column vectors in $\mathbb{R}^n$ called the *standard basis vectors.* In $\mathbb{R}^n$, the **standard basis vectors** refer to the n vectors

$$e_1, e_2, \ldots, e_n \in \mathbb{R}^n,$$

where e_i denotes the column vector with 1 in the i^{th}-entry and 0 everywhere else. For example, in $\mathbb{R}^2$, the standard basis vectors are

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{and} \quad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

In $\mathbb{R}^3$, the standard basis vectors are

$$e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad e_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad \text{and} \quad e_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Sometimes, to save vertical space, people might write something like

$$e_1 = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^{\mathtt{T}}, \quad e_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\mathtt{T}}, \quad \text{and} \quad e_3 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathtt{T}}$$

instead, where $^{\mathtt{T}}$ is simply the transpose operation, which in this case transforms each 1×3-row vector into a 3×1-column vector.

Remark 7.1. Whether

$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \text{or} \quad e_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

or some other vector depends on the context. For instance, if e_1 is understood to be an element of $\mathbb{R}^2$, then e_1 is the column vector $\begin{bmatrix} 1 & 0 \end{bmatrix}^{\mathrm{T}}$. On the other hand, if e_1 is an element of $\mathbb{R}^7$, then e_1 would be a 7×1-column vector.

A crucial property of the standard basis vectors in $\mathbb{R}^n$ is the following.

Proposition 7.1. *For every vector $v \in \mathbb{R}^n$, there exist real numbers $a_1, a_2, \ldots, a_n \in \mathbb{R}$ such that*

$$v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n.$$

Furthermore, such an expression is unique: if there are real numbers $a_1, a_2, \ldots, a_n, b_1, b_2, \ldots, b_n$ such that

$$v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n$$

and

$$v = b_1 \cdot e_1 + b_2 \cdot e_2 + \cdots + b_n \cdot e_n,$$

then

$$a_1 = b_1, \quad a_2 = b_2, \quad \ldots, \quad a_n = b_n.$$

Proof. Existence. Since v is an element of $\mathbb{R}^n$, we can write

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix},$$

where $v_i \in \mathbb{R}$ is the real number in the i^{th} entry of v. Now the existence of $a_1, a_2, \ldots, a_n$ is clear: we simply take $a_1 = v_1$, $a_2 = v_2$, …, $a_n = v_n$. We have

$$v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot v_n,$$

as desired.

Uniqueness. If

$$v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n$$

and

$$v = b_1 \cdot e_1 + b_2 \cdot e_2 + \cdots + b_n \cdot e_n,$$

happen simultaneously, then

$$(a_1 - b_1) \cdot e_1 + (a_2 - b_2) \cdot e_2 + \cdots + (a_n - b_n) \cdot e_n = \mathbf{0}.$$

This means

$$\begin{bmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_n - b_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

whence $a_1 = b_1$, $a_2 = b_2$, …, $a_n = b_n$. □

When we express v as $v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n$ as in the above proposition, we say that we write v as a *linear combination* of the standard basis vectors $e_1, e_2, \ldots, e_n$.

Example 7.1. Let us write the vector

$$v := \begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix} \in \mathbb{R}^3$$

as a linear combination of the standard basis vectors.

Since

$$\begin{bmatrix} -3 \\ 0 \\ 4 \end{bmatrix} = (-3) \cdot \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} + 0 \cdot \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 4 \cdot \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix},$$

we find that

$$v = (-3) \cdot e_1 + 0 \cdot e_2 + 4 \cdot e_3.$$

7.2. **Linear transformations.** In this subsection, we discuss linear transformations. Throughout this subsection, we let m and n be positive integers.

Definition 7.2. A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is said to be a **linear transformation** if it satisfies the following properties:

(i) for every pair of vectors $u, v \in \mathbb{R}^n$, we have

$$T(u + v) = T(u) + T(v),$$

and

(ii) for every vector $w \in \mathbb{R}^n$ and for every real number $a \in \mathbb{R}$, we have

$$T(a \cdot w) = a \cdot T(w).$$

The following proposition gives an equivalent characterization of linear transformations. It can be used as an alternative definition for linear transformations.

Proposition 7.2. *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then T is a linear transformation if and only if for every pair of real numbers $a, b \in \mathbb{R}$ and for every pair of vectors $u, v \in \mathbb{R}^n$, we have*

$$T(a \cdot u + b \cdot v) = a \cdot T(u) + b \cdot T(v). \tag{24}$$

Proof. ($\Rightarrow$) Suppose that T is a linear transformation in the sense of Definition 7.2. Let us show that (24) is satisfied. Take any $a, b \in \mathbb{R}$ and any $u, v \in \mathbb{R}^n$. By (i) from Definition 7.2, we have

$$T(a \cdot u + b \cdot v) = T(a \cdot u) + T(b \cdot v).$$

Now by (ii), we have

$$T(a \cdot u) = a \cdot T(u) \quad \text{and} \quad T(b \cdot v) = b \cdot T(v).$$

Combining the three equations above, we have

$$T(a \cdot u + b \cdot v) = a \cdot T(u) + b \cdot T(v),$$

as desired.

($\Leftarrow$) On the other hand, suppose that T is a function which satisfies (24). Let us show that both (i) and (ii) from Definition 7.2 are satisfied. First, we show (i). Take any $u, v \in \mathbb{R}^n$. Using (24) with $a = b = 1$, we have

$$T(u + v) = T(1 \cdot u + 1 \cdot v) = 1 \cdot T(u) + 1 \cdot T(v) = T(u) + T(v),$$

as desired.

Second, we show (ii). Take any $a \in \mathbb{R}$ and any $w \in \mathbb{R}^n$. Substitute a, b in (24) with $a/2$ and $a/2$, and u, v with w. We find

$$T(a \cdot w) = T\Big(\frac{a}{2} \cdot w + \frac{a}{2} \cdot w\Big) = \frac{a}{2} \cdot T(w) + \frac{a}{2} \cdot T(w) = a \cdot T(w),$$

as desired. $\square$

Example 7.3. Suppose that $T : \mathbb{R}^2 \to \mathbb{R}^3$ is a linear transformation such that

$$T\left(\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad \text{and} \quad T\left(\begin{bmatrix} 3 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \\ 4 \end{bmatrix}.$$

What is

$$T\left(\begin{bmatrix} 1 \\ 4 \end{bmatrix}\right)?$$

Note that

$$\begin{bmatrix} 1 \\ 4 \end{bmatrix} = (-2) \cdot \begin{bmatrix} 1 \\ -1 \end{bmatrix} + 1 \cdot \begin{bmatrix} 3 \\ 2 \end{bmatrix}.$$

Therefore,

$$T\left(\begin{bmatrix}1\\4\end{bmatrix}\right) = (-2)\cdot T\left(\begin{bmatrix}1\\-1\end{bmatrix}\right) + 1\cdot T\left(\begin{bmatrix}3\\2\end{bmatrix}\right) = (-2)\cdot\begin{bmatrix}1\\2\\3\end{bmatrix} + 1\cdot\begin{bmatrix}0\\-1\\4\end{bmatrix} = \begin{bmatrix}-2\\-5\\-2\end{bmatrix}.$$

Now observe that if we know the values of $T(e_1), T(e_2), \ldots, T(e_n)$, where $e_1, e_2, \ldots, e_n$ is the standard basis vectors in $\mathbb{R}^n$, then we know the value of $T(v)$ for any $v \in \mathbb{R}^n$: we write v is a linear combination

$$v = a_1 \cdot e_1 + a_2 \cdot e_2 + \cdots + a_n \cdot e_n,$$

and then we find

$$T(v) = a_1 \cdot T(e_1) + a_2 \cdot T(e_2) + \cdots + a_n \cdot T(e_n).$$

Now let us consider the matrix $M \in \mathbb{R}^{n\times m}$ whose i^{th} column from the left is the column vector $T(e_i)$. We find that as

$$v = \begin{bmatrix}a_1\\a_2\\\vdots\\a_n\end{bmatrix},$$

we have

$$M \cdot v = a_1 \cdot T(e_1) + a_2 \cdot T(e_2) + \cdots + a_n \cdot T(e_n) = T(v).$$

Thus, a linear transformation $\mathbb{R}^n \to \mathbb{R}^m$ is exactly the same thing as multiplication on the left by an $m \times n$-matrix!

Example 7.4. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}-1\\0\\1\end{bmatrix} \quad \text{and} \quad T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}0\\2\\3\end{bmatrix}.$$

Let

$$v = \begin{bmatrix}4\\5\end{bmatrix}.$$

What is the value of $T(v)$?

We write

$$v = 4 \cdot e_1 + 5 \cdot e_2.$$

We then have

$$T(v) = 4 \cdot T(e_1) + 5 \cdot T(e_2) = 4 \cdot \begin{bmatrix}-1\\0\\1\end{bmatrix} + 5 \cdot \begin{bmatrix}0\\2\\3\end{bmatrix} = \begin{bmatrix}-4\\10\\19\end{bmatrix}.$$

On the other hand, let us construct the matrix $M \in \mathbb{R}^{3\times 2}$ whose i^{th} column is $T(e_i)$:

$$M = \begin{bmatrix} -1 & 0 \\ 0 & 2 \\ 1 & 3 \end{bmatrix}.$$

Note that

$$M \cdot v = \begin{bmatrix} -1 & 0 \\ 0 & 2 \\ 1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \begin{bmatrix} -4 \\ 10 \\ 19 \end{bmatrix},$$

which is the same as $T(v)$.

Under the above correspondence, every linear transformation corresponds to a matrix, and every matrix corresponds to a linear transformation!

7.3. **A bit of midterm review.** Note! Remember to show your work on the midterm! Write how you obtain your answers, not just answers!

Problem 7.5. Suppose

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

Then (a) what is $3 \cdot A - 2 \cdot A^{\mathsf{T}}$? (b) what is $\operatorname{tr}(233 \cdot A + 35 \cdot A^{\mathsf{T}})$? [21]

Problem 7.6. Consider the following system

$$\begin{cases} x + y + z = 5, & \text{(25a)} \\ x - y = 2, & \text{(25b)} \end{cases}$$

of two linear equations in the three variables x, y, z.

(a) Determine the set of all real numbers $t \in \mathbb{R}$ for which the triple $(2t, 4-t, 1-t)$ is a solution to the linear system.

(b) Determine the set of all real numbers $t \in \mathbb{R}$ for which the triple $(3t, t, -3)$ is a solution to the linear system.[22]

---

[21] *Answers.* (a) $\begin{bmatrix} 0 & 0 & 3 & -3 \\ 0 & -1 & 9 & 15 \\ -2 & -11 & 0 & 0 \\ 2 & -10 & 0 & 1 \end{bmatrix}$. (b) 0.

[22] *Answers.* (a) $\{2\}$. (b) $\varnothing$, the empty set.

Problem 7.7. Suppose we have a square matrix $X \in \mathbb{R}^{n \times n}$. Suppose that after a sequence of elementary row operations, we transform X into a square matrix $Y \in \mathbb{R}^{n \times n}$ in reduced row echelon form. How does the bottommost row of Y indicate whether X is invertible or not?[23]

Problem 7.8. Let

$$Z := \begin{bmatrix} 0 & 1 \\ 2 & 3 \end{bmatrix}.$$

Compute (a) Z^{-1}. (b) $\det(Z)$. (c) $\det(Z^{-3})$.[24]

[23] *Answer.* If the bottommost row of Y is an all-zero row, then X is singular (non-invertible). If the bottommost row of Y is $[0 \ 0 \ \cdots \ 0 \ 1]$, then X is invertible.

[24] *Answers.* (a) $Z^{-1} = \begin{bmatrix} -3/2 & 1/2 \\ 1 & 0 \end{bmatrix}$. (b) $\det(Z) = -2$. (c) $\det(Z^{-3}) = -1/8$.

7.4. **Problems and Solutions.**

Problem 1. Suppose that $T : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation such that

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}2\\1\\3\end{bmatrix}, \quad T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = \begin{bmatrix}-1\\0\\1\end{bmatrix}, \quad \text{and} \quad T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}-1\\-1\\-4\end{bmatrix}.$$

(a) Evaluate $T\left(\begin{bmatrix}4\\0\\0\end{bmatrix}\right)$.

(b) Evaluate $T\left(\begin{bmatrix}0\\2\\0\end{bmatrix}\right)$.

(c) Evaluate $T\left(\begin{bmatrix}4\\2\\1\end{bmatrix}\right)$.

(d) Evaluate $T\left(\begin{bmatrix}5\\5\\5\end{bmatrix}\right)$.

Now consider the matrix M corresponding to the linear transformation T. Recall from our discussion in class that to construct M, we simply write $T(e_1)$, $T(e_2)$, $T(e_3)$ from left to right:

$$M = \big[T(e_1)\,|\,T(e_2)\,|\,T(e_3)\big] = \begin{bmatrix}2 & -1 & -1\\1 & 0 & -1\\3 & 1 & -4\end{bmatrix}.$$

(e) Compute the following products directly:

$$\begin{bmatrix}2 & -1 & -1\\1 & 0 & -1\\3 & 1 & -4\end{bmatrix} \cdot \begin{bmatrix}4\\2\\1\end{bmatrix}$$

and

$$\begin{bmatrix}2 & -1 & -1\\1 & 0 & -1\\3 & 1 & -4\end{bmatrix} \cdot \begin{bmatrix}5\\5\\5\end{bmatrix}.$$

Verify that the results agree with parts (c) and (d).

Problem 2. Let $U : \mathbb{R}^2 \to \mathbb{R}^2$ be a linear transformation such that

$$U\left(\begin{bmatrix} -1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 3 \end{bmatrix} \quad \text{and} \quad U\left(\begin{bmatrix} 1 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 0 \end{bmatrix}.$$

(a) Evaluate

$$U\left(\begin{bmatrix} 11 \\ -22 \end{bmatrix}\right).$$

(b) Evaluate

$$U\left(\begin{bmatrix} 3 \\ 15 \end{bmatrix}\right).$$

(c) Evaluate

$$U\left(\begin{bmatrix} 14 \\ -7 \end{bmatrix}\right).$$

(d) Evaluate

$$U\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right).$$

(e) Evaluate

$$U\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right).$$

Problem 3. Suppose that $V : \mathbb{R}^3 \to \mathbb{R}^3$ is a linear transformation, and suppose that we have

$$V\left(\begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}\right) = \begin{bmatrix} 1/2 \\ 1/3 \\ 1/4 \end{bmatrix} \quad \text{and} \quad V\left(\begin{bmatrix} 3 \\ 4 \\ 5 \end{bmatrix}\right) = \begin{bmatrix} 1/3 \\ 1/4 \\ 1/5 \end{bmatrix}.$$

(a) Evaluate

$$V\left(\begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}\right).$$

(b) Evaluate $V\big([5\ \ 6\ \ 7]^{\mathtt{T}}\big)$.

Solution to Problem 1. (a)

$$T\left(\begin{bmatrix}4\\0\\0\end{bmatrix}\right) = 4\cdot T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = 4\cdot\begin{bmatrix}2\\1\\3\end{bmatrix} = \begin{bmatrix}8\\4\\12\end{bmatrix}.$$

(b)

$$T\left(\begin{bmatrix}0\\2\\0\end{bmatrix}\right) = 2\cdot T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) = 2\cdot\begin{bmatrix}-1\\0\\1\end{bmatrix} = \begin{bmatrix}-2\\0\\2\end{bmatrix}.$$

(c) Using the results from parts (a) and (b), we find

$$T\left(\begin{bmatrix}4\\2\\1\end{bmatrix}\right) = T\left(\begin{bmatrix}4\\0\\0\end{bmatrix}\right) + T\left(\begin{bmatrix}0\\2\\0\end{bmatrix}\right) + T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}8\\4\\12\end{bmatrix} + \begin{bmatrix}-2\\0\\2\end{bmatrix} + \begin{bmatrix}-1\\-1\\-4\end{bmatrix} = \begin{bmatrix}5\\3\\10\end{bmatrix}.$$

(d) Note that

$$T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) + T\left(\begin{bmatrix}0\\1\\0\end{bmatrix}\right) + T\left(\begin{bmatrix}0\\0\\1\end{bmatrix}\right) = \begin{bmatrix}2\\1\\3\end{bmatrix} + \begin{bmatrix}-1\\0\\1\end{bmatrix} + \begin{bmatrix}-1\\-1\\-4\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

Therefore,

$$T\left(\begin{bmatrix}5\\5\\5\end{bmatrix}\right) = 5\cdot T\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = 5\cdot\begin{bmatrix}0\\0\\0\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

(e)

$$\begin{bmatrix}2 & -1 & -1\\1 & 0 & -1\\3 & 1 & -4\end{bmatrix}\cdot\begin{bmatrix}4\\2\\1\end{bmatrix} = \begin{bmatrix}2\cdot 4 + (-1)\cdot 2 + (-1)\cdot 1\\ 1\cdot 4 + 0\cdot 2 + (-1)\cdot 1\\ 3\cdot 4 + 1\cdot 2 + (-4)\cdot 1\end{bmatrix} = \begin{bmatrix}5\\3\\10\end{bmatrix}$$

and

$$\begin{bmatrix}2 & -1 & -1\\1 & 0 & -1\\3 & 1 & -4\end{bmatrix}\cdot\begin{bmatrix}5\\5\\5\end{bmatrix} = \begin{bmatrix}2\cdot 5 + (-1)\cdot 5 + (-1)\cdot 5\\ 1\cdot 5 + 0\cdot 5 + (-1)\cdot 5\\ 3\cdot 5 + 1\cdot 5 + (-4)\cdot 5\end{bmatrix} = \begin{bmatrix}0\\0\\0\end{bmatrix}.$$

These results agree with parts (c) and (d). □

Solution to Problem 2. (a)

$$U\left(\begin{bmatrix}11\\-22\end{bmatrix}\right) = -11\cdot U\left(\begin{bmatrix}-1\\2\end{bmatrix}\right) = -11\cdot\begin{bmatrix}2\\3\end{bmatrix} = \begin{bmatrix}-22\\-33\end{bmatrix}.$$

(b)

$$U\left(\begin{bmatrix}3\\15\end{bmatrix}\right) = 3 \cdot U\left(\begin{bmatrix}1\\5\end{bmatrix}\right) = 3 \cdot \begin{bmatrix}-1\\0\end{bmatrix} = \begin{bmatrix}-3\\0\end{bmatrix}.$$

(c) Using the results from parts (a) and (b), we find

$$U\left(\begin{bmatrix}14\\-7\end{bmatrix}\right) = U\left(\begin{bmatrix}11\\-22\end{bmatrix}\right) + U\left(\begin{bmatrix}3\\15\end{bmatrix}\right) = \begin{bmatrix}-22\\-33\end{bmatrix} + \begin{bmatrix}-3\\0\end{bmatrix} = \begin{bmatrix}-25\\-33\end{bmatrix}.$$

(d) We have

$$U\left(\begin{bmatrix}2\\-1\end{bmatrix}\right) = \frac{1}{7} \cdot U\left(\begin{bmatrix}14\\-7\end{bmatrix}\right) = \frac{1}{7} \cdot \begin{bmatrix}-25\\-33\end{bmatrix} = \begin{bmatrix}-25/7\\-33/7\end{bmatrix}.$$

(e) We have

$$U\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = U\left(\begin{bmatrix}-1\\2\end{bmatrix}\right) + U\left(\begin{bmatrix}2\\-1\end{bmatrix}\right) = \begin{bmatrix}2\\3\end{bmatrix} + \begin{bmatrix}-25/7\\-33/7\end{bmatrix} = \begin{bmatrix}-11/7\\-12/7\end{bmatrix}.$$

$\square$

Solution to Problem 3. (a) Note that

$$V\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = V\left(\begin{bmatrix}3\\4\\5\end{bmatrix}\right) - V\left(\begin{bmatrix}2\\3\\4\end{bmatrix}\right) = \begin{bmatrix}1/3\\1/4\\1/5\end{bmatrix} - \begin{bmatrix}1/2\\1/3\\1/4\end{bmatrix} = \begin{bmatrix}-1/6\\-1/12\\-1/20\end{bmatrix}.$$

Therefore,

$$V\left(\begin{bmatrix}4\\5\\6\end{bmatrix}\right) = V\left(\begin{bmatrix}3\\4\\5\end{bmatrix}\right) + V\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = \begin{bmatrix}1/3\\1/4\\1/5\end{bmatrix} + \begin{bmatrix}-1/6\\-1/12\\-1/20\end{bmatrix} = \begin{bmatrix}1/6\\1/6\\3/20\end{bmatrix}.$$

(b) We have

$$V\left(\begin{bmatrix}5\\6\\7\end{bmatrix}\right) = V\left(\begin{bmatrix}4\\5\\6\end{bmatrix}\right) + V\left(\begin{bmatrix}1\\1\\1\end{bmatrix}\right) = \begin{bmatrix}1/6\\1/6\\3/20\end{bmatrix} + \begin{bmatrix}-1/6\\-1/12\\-1/20\end{bmatrix} = \begin{bmatrix}0\\1/12\\1/10\end{bmatrix}.$$

$\square$

8. The Eighth Week

8.1. Determinant by Laplace expansion. We have previously seen a method of computing the determinant of an $n \times n$ matrix by performing Gaussian elimination (row reduction) and keeping track of the elementary matrices involved. For smaller matrices, like 3×3 ones, we also have explicit formulas to compute determinants. Now let us discuss a different way to compute the determinant of an $n \times n$ matrix via a method called "Laplace expansion" (also called "cofactor expansion").

Suppose that A is an $n \times n$-matrix, where $n \geq 2$ is a positive integer. Let a_{ij} denote the (i, j)-entry of A. For each $i, j \in \{1, 2, \dots, n\}$, we let M_{ij} denote the $(n-1) \times (n-1)$-matrix obtained by removing the i^{th} row and the j^{th} column from A. Then the **(i, j)-minor** of A is defined to be the number

$$m_{ij} := \det(M_{ij}) \in \mathbb{R},$$

and the **(i, j)-cofactor** of A is defined to be the number

$$c_{ij} := (-1)^{i+j} \cdot m_{ij} \in \mathbb{R}.$$

Example 8.1. Consider the following matrix

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

The $(1, 1)$-minor of A is

$$m_{11} = \det \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = +1,$$

and so the $(1, 1)$-cofactor of A is

$$c_{11} = (-1)^{1+1} \cdot m_{11} = (+1) \cdot (+1) = 1.$$

The $(1, 2)$-minor of A is

$$m_{12} = \det \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} = +1,$$

and so the $(1, 2)$-cofactor of A is

$$c_{12} = (-1)^{1+2} \cdot m_{12} = (-1) \cdot (+1) = -1.$$

The $(1, 3)$-minor of A is

$$m_{13} = \det \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix} = +2,$$

and so the $(1, 3)$-cofactor of A is

$$c_{13} = (-1)^{1+3} \cdot m_{13} = (+1) \cdot (+2) = 2.$$

The $(2,1)$-minor of A is

$$m_{21} = \det \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = -2,$$

and so the $(2,1)$-cofactor of A is

$$c_{21} = (-1)^{2+1} \cdot m_{21} = (-1) \cdot (-2) = 2.$$

The $(2,2)$-minor of A is

$$m_{22} = \det \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} = +1,$$

and so the $(2,2)$-cofactor of A is

$$c_{22} = (-1)^{2+2} \cdot m_{22} = (+1) \cdot (+1) = 1.$$

The $(2,3)$-minor of A is

$$m_{23} = \det \begin{bmatrix} 0 & -1 \\ -1 & 1 \end{bmatrix} = -1,$$

and so the $(2,3)$-cofactor of A is

$$c_{23} = (-1)^{2+3} \cdot m_{23} = (-1) \cdot (-1) = 1.$$

The $(3,1)$-minor of A is

$$m_{31} = \det \begin{bmatrix} -1 & 1 \\ 1 & 0 \end{bmatrix} = -1,$$

and so the $(3,1)$-cofactor of A is

$$c_{31} = (-1)^{3+1} \cdot m_{31} = (+1) \cdot (-1) = -1.$$

The $(3,2)$-minor of A is

$$m_{32} = \det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -1,$$

and so the $(3,2)$-cofactor of A is

$$c_{32} = (-1)^{3+2} \cdot m_{32} = (-1) \cdot (-1) = 1.$$

The $(3,3)$-minor of A is

$$m_{33} = \det \begin{bmatrix} 0 & -1 \\ 1 & 1 \end{bmatrix} = +1,$$

and so the $(3,3)$-cofactor of A is

$$c_{33} = (-1)^{3+3} \cdot m_{33} = (+1) \cdot (+1) = 1.$$

To compute the determinant, we have the following theorem.

Theorem 8.1 (Determinant by Laplace expansion)**.** *Let $n \geq 2$ be a positive integer. Let $A \in \mathbb{R}^{n\times n}$. For every pair $i, j \in \{1, 2, \ldots, n\}$, let a_{ij} and c_{ij} denote the (i, j)-entry and the (i, j)-cofactor of A, respectively.*

- (expansion along the i^{th} row) *For every $i \in \{1, 2, \ldots, n\}$, we have*

$$\det(A) = a_{i1}c_{i1} + a_{i2}c_{i2} + \cdots + a_{in}c_{in}.$$

- (expansion along the j^{th} column) *For every $j \in \{1, 2, \ldots, n\}$, we have*

$$\det(A) = a_{1j}c_{1j} + a_{2j}c_{2j} + \cdots + a_{nj}c_{nj}.$$

Example 8.2. Let A be the 3×3-matrix from Example 8.1:

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

First, recall that from the rule of Sarrus, we can compute

$$\det(A) = 0 + 0 + 1 - (-1) - 0 - (-1) = 3.$$

Now let us compute $\det(A)$ by row and column expansions, in *six* different ways!

Using the first row, we find

$$\det(A) = 0 \cdot 1 + (-1) \cdot (-1) + 1 \cdot 2 = 0 + 1 + 2 = 3.$$

Using the second row, we find

$$\det(A) = 1 \cdot 2 + 1 \cdot 1 + 0 \cdot 1 = 2 + 1 + 0 = 3.$$

Using the third row, we find

$$\det(A) = (-1) \cdot (-1) + 1 \cdot 1 + 1 \cdot 1 = 1 + 1 + 1 = 3.$$

Using the first column, we find

$$\det(A) = 0 \cdot 1 + 1 \cdot 2 + (-1) \cdot (-1) = 0 + 2 + 1 = 3.$$

Using the second column, we find

$$\det(A) = (-1) \cdot (-1) + 1 \cdot 1 + 1 \cdot 1 = 1 + 1 + 1 = 3.$$

Using the third column, we find

$$\det(A) = 1 \cdot 2 + 0 \cdot 1 + 1 \cdot 1 = 2 + 0 + 1 = 3.$$

Example 8.3. Let

$$B := \begin{bmatrix} 1 & 2 & 0 & 0 \\ 0 & 3 & 4 & 0 \\ 0 & 0 & 5 & 6 \\ 0 & 0 & 7 & 8 \end{bmatrix}.$$

Let us compute $\det(B)$.

First, by expanding along the first column, we find

$$\det(B) = 1 \cdot (-1)^{1+1} \cdot \det(C),$$

where

$$C := \begin{bmatrix} 3 & 4 & 0 \\ 0 & 5 & 6 \\ 0 & 7 & 8 \end{bmatrix}.$$

Now, by expanding along the first column of C, we find

$$\det(C) = 3 \cdot (-1)^{1+1} \cdot \det(D),$$

where

$$D := \begin{bmatrix} 5 & 6 \\ 7 & 8 \end{bmatrix}.$$

The determinant of D is

$$\det(D) = 5 \cdot 8 - 6 \cdot 7 = -2.$$

Therefore,

$$\det(B) = 1 \cdot 1 \cdot \det(C) = 1 \cdot 1 \cdot 3 \cdot 1 \cdot \det(D) = 1 \cdot 1 \cdot 3 \cdot 1 \cdot (-2) = -6.$$

Example 8.4. Consider the matrix

$$K := \begin{bmatrix} 0 & 0 & 0 & 5 \\ 2 & 0 & 0 & 8 \\ 15 & 2 & 3 & 9 \\ 16 & 4 & 11 & 10 \end{bmatrix}.$$

Let us compute $\det(K)$.

By expanding along the first row, we find

$$\det(K) = 5 \cdot (-1)^{1+4} \cdot \det(L),$$

where

$$L := \begin{bmatrix} 2 & 0 & 0 \\ 15 & 2 & 3 \\ 16 & 4 & 11 \end{bmatrix}.$$

Next, we expand along the first row of L to find

$$\det(L) = 2 \cdot (-1)^{1+1} \cdot \det(M),$$

where

$$M := \begin{bmatrix} 2 & 3 \\ 4 & 11 \end{bmatrix}.$$

Note that

$$\det(M) = 2 \cdot 11 - 3 \cdot 4 = 10.$$

Therefore,

$$\det(K) = 5 \cdot (-1) \cdot 2 \cdot 1 \cdot 10 = -100.$$

Now consider an arbitrary $n \times n$-matrix A. The **cofactor matrix** of A is defined as the matrix

$$\mathrm{Cof}(A) \in \mathbb{R}^{n \times n}$$

whose (i, j)-entry is c_{ij}, the (i, j)-cofactor of A. The **adjugate** of A is defined as the matrix

$$\mathrm{adj}(A) := (\mathrm{Cof}(A))^{\mathsf{T}} \in \mathbb{R}^{n \times n}.$$

Namely, the *adjugate* of A is the transpose of the cofactor matrix of A.

Proposition 8.2. *For every square matrix $A \in \mathbb{R}^{n \times n}$, we have*

$$A \cdot \mathrm{adj}(A) = \mathrm{adj}(A) \cdot A = \det(A) \cdot I_n.$$

Example 8.5. Let A be the 3×3-matrix from Example 8.1:

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

The cofactor matrix of A is

$$\mathrm{Cof}(A) = \begin{bmatrix} 1 & -1 & 2 \\ 2 & 1 & 1 \\ -1 & 1 & 1 \end{bmatrix},$$

and the adjugate of A is

$$\mathrm{adj}(A) = \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix}.$$

Note that

$$A \cdot \mathrm{adj}(A) = \begin{bmatrix} 0 & -1 & 1 \\ 1 & 1 & 0 \\ -1 & 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & -1 \\ -1 & 1 & 1 \\ 2 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix} = \det(A) \cdot I_3.$$

In particular, this means that, for a general square matrix $X \in \mathbb{R}^{n\times n}$ with $n \geq 2$, if X is invertible, then

$$X^{-1} = \frac{1}{\det(X)} \cdot \operatorname{adj}(X).$$

On the other hand, if X is non-invertible, then

$$X \cdot \operatorname{adj}(X) = \mathbf{0}_n.$$

Remark 8.1. Recall that if the matrix

$$N := \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

is invertible, then we have the formula

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

Note that the denominator $ad - bc$ is the *determinant* of N, and the matrix

$$\begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

is the *adjugate* of N.[25]

Remark 8.2. There are other beautiful methods of computing determinants, such as via *Dodgson condensation* and via *Desnanot–Jacobi identity*, but we are not going to those at the moment.

8.2. **Some properties of determinants.** This subsection discusses some properties of determinants. Throughout this subsection, let $n \geq 2$ be a positive integer, let $i, j \in \{1, 2, \ldots, n\}$ be indices with $i \neq j$, and let $A, B \in \mathbb{R}^{n\times n}$.

We have seen some of these properties earlier already.

Proposition 8.3. *We have* $\det(A \cdot B) = \det(A) \cdot \det(B)$.

[25]Indeed, this is not hard to see. The minors of N are $m_{11} = d$, $m_{12} = c$, $m_{21} = b$, and $m_{22} = a$. Thus, the cofactors of N are $c_{11} = d$, $c_{12} = -c$, $c_{21} = -b$, and $c_{22} = a$. The cofactor matrix of N is then

$$\operatorname{Cof}(N) = \begin{bmatrix} d & -c \\ -b & a \end{bmatrix},$$

and therefore the adjugate of N is

$$\operatorname{adj}(N) = (\operatorname{Cof}(N))^{\mathsf{T}} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

Proposition 8.4. *If the matrix B is obtained from A by swapping the i^{th} row and the j^{th} row of A,*[26] *then*

$$\det(B) = -\det(A).$$

Proposition 8.5. *Suppose c is a nonzero real number. If the matrix B is obtained from A by multiplying the i^{th} row of A by c,*[27] *then*

$$\det(B) = c \cdot \det(A).$$

Proposition 8.6. *Suppose c is a real number. If the matrix B is obtained from A by adding c times the j^{th} row of A to the i^{th} row of A,*[28] *then*

$$\det(B) = \det(A).$$

Proposition 8.7. *We have* $\det(A^{\intercal}) = \det(A)$.

Proposition 8.8. *If A has an all-zero row, then* $\det(A) = 0$.

Proposition 8.9. *If A has an all-zero column, then* $\det(A) = 0$.

Proposition 8.10. *Let $A \in \mathbb{R}^{n\times n}$ with $n \geq 2$. If there are $i, j \in \{1, 2, \ldots, n\}$ with $i \neq j$ such that the i^{th} row of A and the j^{th} row of A are identical, then* $\det(A) = 0$.

Proposition 8.11. *Let $A \in \mathbb{R}^{n\times n}$ with $n \geq 2$. If there are $i, j \in \{1, 2, \ldots, n\}$ with $i \neq j$ such that the i^{th} column of A and the j^{th} column of A are identical, then* $\det(A) = 0$.

Proposition 8.12. *If A is upper-triangular or lower-triangular, then $\det(A)$ is the product of the diagonal entries of A. In other words, for an upper-triangular or a lower-triangular matrix A, if a_{ij} denote the (i, j)-entry of A, then*

$$\det(A) = a_{11} \cdot a_{22} \cdots a_{nn}.$$

Recall that if $\lambda \in \mathbb{R}$ is a real number, then $\lambda \cdot A$ denotes the matrix whose (i, j)-entry is λ times the (i, j)-entry of A. For instance, if

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

then

$$5 \cdot X = \begin{bmatrix} 5 & 10 & 15 \\ 20 & 25 & 30 \\ 35 & 40 & 45 \end{bmatrix}.$$

---

[26] $R_i \rightleftharpoons R_j$
[27] $R_i \mapsto c \cdot R_i$
[28] $R_i \mapsto R_i + c \cdot R_j$

Proposition 8.13. *Let $\lambda \in \mathbb{R}$. Then*

$$\det(\lambda \cdot A) = \lambda^n \cdot \det(A).$$

Proof. We obtain $\lambda \cdot A$ from A by n steps of elementary row operations where the i^{th} step is multiplying the i^{th} row by λ. As a result, $\det(\lambda \cdot A)$ equals λ^n times $\det(A)$. $\square$

Example 8.6. Let

$$Y := \begin{bmatrix} 1 & 2 & 0 \\ 2 & 0 & 3 \\ 0 & 3 & 4 \end{bmatrix}.$$

Note that by the rule of Sarrus, we have $\det(Y) = -25$. Consider

$$Z := \begin{bmatrix} 4 & 8 & 0 \\ 8 & 0 & 12 \\ 0 & 12 & 16 \end{bmatrix}.$$

Note that $Z = 4 \cdot Y$, and thus $\det(Z) = \det(4 \cdot Y) = 4^3 \cdot \det(Y) = -1{,}600$.

To introduce the next proposition, we consider the following example.

Example 8.7. Consider the following three matrices

$$M_1 := \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}, \qquad M_2 := \begin{bmatrix} 2 & 1 \\ 3 & 5 \end{bmatrix}, \qquad \text{and} \qquad M_3 := \begin{bmatrix} 2 & 1 \\ 4 & 6 \end{bmatrix}.$$

Note that

$$\det(M_1) = 1, \qquad \det(M_2) = 7, \qquad \text{and} \qquad \det(M_2) = 8.$$

The three matrices M_1, M_2, and M_3 have the same first rows, while their second rows are related by the following additive relation:

$$\begin{bmatrix} 1 & 1 \end{bmatrix} + \begin{bmatrix} 3 & 5 \end{bmatrix} = \begin{bmatrix} 4 & 6 \end{bmatrix}.$$

On the other hand, their *determinants* are also related by the analogous additive relations:

$$\det(M_1) + \det(M_2) = \det(M_3).$$

The example above is a special case of the following general phenomenon.

Proposition 8.14. *Let $i \in \{1, 2, \ldots, n\}$ be an index. Suppose that $A, B, C \in \mathbb{R}^{n \times n}$ are matrices with the following properties:*

- *for every index $k \neq i$, the k^{th} rows of A, B, and C are identical, and*

- *there are real numbers* λ_1 *and* λ_2 *such that*

$$R_i(C) = \lambda_1 \cdot R_i(A) + \lambda_2 \cdot R_i(B),$$

where $R_i(A), R_i(B), R_i(C)$ *denote the* i^{th} *rows of* A, B, C, *respectively.*

Then we have

$$\det(C) = \lambda_1 \cdot \det(A) + \lambda_2 \cdot \det(B).$$

Example 8.8. Consider

$$A := \begin{bmatrix} 1 & 0 & 1 \\ 3 & 4 & 5 \\ 0 & 1 & 2 \end{bmatrix}, \quad B := \begin{bmatrix} 1 & 0 & 1 \\ -3 & 2 & 31 \\ 0 & 1 & 2 \end{bmatrix}, \quad \text{and} \quad C := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 6 \\ 0 & 1 & 2 \end{bmatrix}.$$

Observe that for every index $k \neq 2$ (in this case, this means $k \in \{1,3\}$), the k^{th} rows of A, B, and C are identical. Also note that

$$\frac{1}{6} \cdot R_2(A) + \frac{1}{6} \cdot R_2(B) = R_2(C).$$

The proposition above implies that

$$\frac{1}{6}\det(A) + \frac{1}{6}\det(B) = \det(C).$$

Indeed, this is the case, since

$$\det(A) = 6, \quad \det(B) = -30, \quad \text{and} \quad \det(C) = -4.$$

We also have an analogous result for columns.

Proposition 8.15. *Let* $i \in \{1, 2, \dots, n\}$ *be an index. Suppose that* $X, Y, Z \in \mathbb{R}^{n\times n}$ *are matrices with the following properties:*

- *for every index* $k \neq i$, *the* k^{th} *columns of* X, Y, *and* Z *are identical, and*

- *there are real numbers* λ_1 *and* λ_2 *such that*

$$C_i(Z) = \lambda_1 \cdot C_i(X) + \lambda_2 \cdot C_i(Y),$$

where $C_i(X), C_i(Y), C_i(Z)$ *denote the* i^{th} *columns of* X, Y, Z, *respectively.*

Then we have

$$\det(Z) = \lambda_1 \cdot \det(X) + \lambda_2 \cdot \det(Y).$$

Example 8.9. Suppose that w, x, y, z are real numbers such that

$$\det \begin{bmatrix} 1 & 1 & x \\ -1 & w & y \\ 0 & 1 & z \end{bmatrix} = 4 \qquad \text{and} \qquad \det \begin{bmatrix} 7 & 1 & x \\ 9 & w & y \\ 2 & 1 & z \end{bmatrix} = 20.$$

What is the value of

$$\det \begin{bmatrix} 1 & 1 & x \\ 3 & w & y \\ 1/2 & 1 & z \end{bmatrix}?$$

Let

$$X := \begin{bmatrix} 1 & 1 & x \\ -1 & w & y \\ 0 & 1 & z \end{bmatrix}, \qquad Y := \begin{bmatrix} 7 & 1 & x \\ 9 & w & y \\ 2 & 1 & z \end{bmatrix}, \qquad \text{and} \qquad Z := \begin{bmatrix} 1 & 1 & x \\ 3 & w & y \\ 1/2 & 1 & z \end{bmatrix}.$$

Note that

$$C_1(Z) = -\frac{3}{4} \cdot C_1(X) + \frac{1}{4} \cdot C_1(Y).$$

Therefore,

$$\det(Z) = -\frac{3}{4} \det(X) + \frac{1}{4} \det(Y) = -\frac{3}{4} \cdot 4 + \frac{1}{4} \cdot 20 = 2.$$

Another way to phrase the proposition above is as follows. Suppose that A is an $n \times n$-matrix. Let us consider the function $T : \mathbb{R}^n \to \mathbb{R}^1$ where for every column vector $u \in \mathbb{R}^n$, the 1×1-matrix $T(u)$ (which can be thought of as just a real number) is the determinant of the matrix obtained by replacing the i^{th} column of A by u. Then T is a *linear transformation.*

For example, the function

$$T : \mathbb{R}^3 \to \mathbb{R}$$

given by

$$T\left(\begin{bmatrix} a \\ b \\ c \end{bmatrix} \right) = \det \begin{bmatrix} -1 & a & 5 \\ 0 & b & 12 \\ 1 & c & -7 \end{bmatrix}$$

is a linear transformation.

8.3. **Problems and Solutions.**

Problem 1. Consider the matrix

$$A := \begin{bmatrix} 7 & -2 & 0 & 5 \\ 8 & 4 & -1 & 0 \\ 0 & 7 & 7 & 2 \\ 2 & 0 & 3 & 0 \end{bmatrix}.$$

(a) Compute the $(1,1)$-minor of A.

(b) Compute the $(2,3)$-minor of A.

(c) Compute the $(3,4)$-minor of A^{T}.

Problem 2. Consider the matrix

$$B := \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}.$$

(a) First, use the rule of Sarrus to compute the determinant of B. Verify that B is invertible.

(b) Compute the cofactor matrix $\mathrm{Cof}(B)$ of B.

(c) Compute the adjugate $\mathrm{adj}(B)$ of B.

(d) Compute $B \cdot \mathrm{adj}(B)$ by direct multiplication. Verify that the product equals $\det(B) \cdot I_3$.

(e) Compute B^{-1}.

(f) Compute $\det(\mathrm{Cof}(B))$.

(g) Compute $\mathrm{adj}(\mathrm{adj}(B))$.

Problem 3. Consider the matrix

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

(a) Compute $\det(C)$ by Laplace expansion along the first row.

(b) Compute $\det(C)$ by Laplace expansion along the third column.

Problem 4. Compute the following determinant

$$\det \begin{bmatrix} 0 & 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 0 & 2 & 3 \\ 0 & 0 & 4 & 0 & 1 & 0 \\ 2 & 0 & 0 & 3 & 4 & 0 \\ 1 & 2 & 0 & 0 & 0 & 3 \end{bmatrix}.$$

Problem 5. Suppose that the matrix $D \in \mathbb{R}^{3\times 3}$ satisfies

$$\det(D) = \frac{8}{5}.$$

Evaluate $\det(\operatorname{adj}(5 \cdot D))$.

Problem 6. Suppose that a, b, c, d are real numbers for which

$$\det \begin{bmatrix} 0 & a & 2 \\ b & c & 1 \\ d & 1 & 3 \end{bmatrix} = -87 \qquad \text{and} \qquad \det \begin{bmatrix} 0 & a & 1 \\ b & c & 0 \\ d & 1 & 1 \end{bmatrix} = -33.$$

Evaluate the determinant

$$\det \begin{bmatrix} 0 & a & -4 \\ b & c & 5 \\ d & 1 & 1 \end{bmatrix}.$$

Problem 7. Consider the matrix

$$G := \begin{bmatrix} 2 & 3 & 2 & 3 & 2 \\ 3 & 2 & 3 & 2 & 3 \\ 2 & 3 & 2 & 3 & 2 \\ 3 & 2 & 3 & 2 & 3 \\ 2 & 3 & 2 & 3 & 2 \end{bmatrix}.$$

Compute the adjugate $\operatorname{adj}(G)$ of G.

Solution to Problem 1. (a) The $(1,1)$-minor of A is

$$\det\begin{bmatrix}4 & -1 & 0\\ 7 & 7 & 2\\ 0 & 3 & 0\end{bmatrix} = 0+0+0-0-24-0 = -24.$$

(b) The $(2,3)$-minor of A is

$$\det\begin{bmatrix}7 & -2 & 5\\ 0 & 7 & 2\\ 2 & 0 & 0\end{bmatrix} = 0+(-8)+0-70-0-0 = -78.$$

(c) The transpose of A is

$$A^{\mathtt{T}} = \begin{bmatrix}7 & 8 & 0 & 2\\ -2 & 4 & 7 & 0\\ 0 & -1 & 7 & 3\\ 5 & 0 & 2 & 0\end{bmatrix}.$$

Thus, the $(3,4)$-minor of $A^{\mathtt{T}}$ is

$$\det\begin{bmatrix}7 & 8 & 0\\ -2 & 4 & 7\\ 5 & 0 & 2\end{bmatrix} = 56+280+0-0-0-(-32) = 368.$$

$\square$

Solution to Problem 2. (a) By the rule of Sarrus,

$$\det(B) = 1+0+0-1-1-0 = -1.$$

(b) The minors of B are $m_{11}=0$, $m_{12}=1$, $m_{13}=-1$, $m_{21}=1$, $m_{22}=0$, $m_{23}=-1$, $m_{31}=-1$, $m_{32}=-1$, and $m_{33}=1$. Therefore, the cofactor matrix of B is

$$\operatorname{Cof}(B) = \begin{bmatrix}0 & -1 & -1\\ -1 & 0 & 1\\ -1 & 1 & 1\end{bmatrix}.$$

(c) The adjugate is the transpose of the cofactor matrix. In this case, note that the cofactor matrix is symmetric. Therefore,

$$\operatorname{adj}(B) = \operatorname{Cof}(B)^{\mathtt{T}} = \operatorname{Cof}(B) = \begin{bmatrix}0 & -1 & -1\\ -1 & 0 & 1\\ -1 & 1 & 1\end{bmatrix}.$$

(d) We have

$$B \cdot \operatorname{adj}(B) = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 0 & -1 & -1 \\ -1 & 0 & 1 \\ -1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} = -I_3,$$

which is indeed $\det(B) \cdot I_3$.

(e) We have

$$B^{-1} = \frac{\operatorname{adj}(B)}{\det(B)} = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & -1 \end{bmatrix}.$$

(f) Recall that $B \cdot \operatorname{adj}(B) = \det(B) \cdot I_3$. By taking the determinant on both sides, we have

$$\det(B) \cdot \det(\operatorname{adj}(B)) = \det(B)^3 \cdot 1,$$

and so $\det(\operatorname{adj}(B)) = \det(B)^2 = (-1)^2 = 1$. Since $\operatorname{Cof}(B) = \operatorname{adj}(B)^{\mathsf{T}}$, we have

$$\det(\operatorname{Cof}(B)) = \det(\operatorname{adj}(B)) = 1.$$

(g) Note that

$$\operatorname{adj}(B) \cdot \operatorname{adj}(\operatorname{adj}(B)) = \det(\operatorname{adj}(B)) \cdot I_3.$$

From part (f), we have that $\det(\operatorname{adj}(B)) = 1$. Therefore,

$$\operatorname{adj}(B) \cdot \operatorname{adj}(\operatorname{adj}(B)) = I_3.$$

Now recall that $\operatorname{adj}(B) = \det(B) \cdot B^{-1} = -B^{-1}$. This shows that

$$-B^{-1} \cdot \operatorname{adj}(\operatorname{adj}(B)) = I_3,$$

which means that $-\operatorname{adj}(\operatorname{adj}(B))$ must be the inverse of B^{-1}. Hence,

$$\operatorname{adj}(\operatorname{adj}(B)) = -B = \begin{bmatrix} -1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & -1 \end{bmatrix}.$$

$\square$

Solution to Problem 3. (a) We have

$$\det(C) = 2 \cdot \det \begin{bmatrix} -1 & 0 & 4 \\ 3 & 4 & 0 \\ 2 & 2 & 6 \end{bmatrix} - 3 \cdot \det \begin{bmatrix} -1 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 2 & 5 \end{bmatrix}.$$

By the rule of Sarrus, we have

$$\det \begin{bmatrix} -1 & 0 & 4 \\ 3 & 4 & 0 \\ 2 & 2 & 6 \end{bmatrix} = -24 + 0 + 24 - 32 - 0 - 0 = -32.$$

Since the matrix

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

is lower-triangular, we have

$$\det \begin{bmatrix} -1 & 0 & 0 \\ 3 & 4 & 0 \\ 2 & 2 & 5 \end{bmatrix} = (-1) \cdot 4 \cdot 5 = -20.$$

Therefore,

$$\det(C) = 2(-32) - 3(-20) = -4.$$

(b) We have

$$\det(C) = 2 \cdot \det \begin{bmatrix} -1 & 0 & 4 \\ 3 & 4 & 0 \\ 2 & 2 & 6 \end{bmatrix} - 5 \cdot \det \begin{bmatrix} 0 & 0 & 3 \\ -1 & 0 & 4 \\ 3 & 4 & 0 \end{bmatrix}.$$

We have

$$\det \begin{bmatrix} 0 & 0 & 3 \\ -1 & 0 & 4 \\ 3 & 4 & 0 \end{bmatrix} = (-4) \cdot \det \begin{bmatrix} 0 & 3 \\ -1 & 4 \end{bmatrix} = (-4)(3) = -12.$$

Therefore,

$$\det(C) = 2(-32) - 5(-12) = -4.$$

□

Solution to Problem 4. Let A be the 6×6-matrix in the problem statement. By Laplace expansion along the third column of A, we have

$$\det(A) = (-4) \cdot \det(B),$$

where

$$B = \begin{bmatrix} 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 3 & 0 & 4 \\ 0 & 1 & 0 & 2 & 3 \\ 2 & 0 & 3 & 4 & 0 \\ 1 & 2 & 0 & 0 & 3 \end{bmatrix}.$$

By Laplace expansion along the first row of B, we have

$$\det(B) = \det(C) - 2 \cdot \det(D),$$

where

$$C = \begin{bmatrix} 0 & 0 & 0 & 4 \\ 0 & 1 & 2 & 3 \\ 2 & 0 & 4 & 0 \\ 1 & 2 & 0 & 3 \end{bmatrix} \qquad \text{and} \qquad D = \begin{bmatrix} 0 & 0 & 3 & 4 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 3 & 0 \\ 1 & 2 & 0 & 3 \end{bmatrix}.$$

By Laplace expansion along the first row of C, we have

$$\det(C) = (-4) \cdot \det \begin{bmatrix} 0 & 1 & 2 \\ 2 & 0 & 4 \\ 1 & 2 & 0 \end{bmatrix} = (-4) \cdot (0 + 4 + 8 - 0 - 0 - 0) = -48.$$

By Laplace expansion along the first row of D, we have

$$\begin{aligned} \det(D) &= 3 \cdot \det \begin{bmatrix} 0 & 1 & 3 \\ 2 & 0 & 0 \\ 1 & 2 & 3 \end{bmatrix} - 4 \cdot \det \begin{bmatrix} 0 & 1 & 0 \\ 2 & 0 & 3 \\ 1 & 2 & 0 \end{bmatrix} \\ &= 3 \cdot (0 + 0 + 12 - 0 - 0 - 6) - 4 \cdot (0 + 3 + 0 - 0 - 0 - 0) \\ &= 18 - 12 = 6. \end{aligned}$$

Therefore,

$$\det(B) = -48 - 2 \cdot 6 = -60,$$

and hence

$$\det(A) = (-4) \cdot (-60) = 240.$$

□

Solution to Problem 5. From $\det(D) = \frac{8}{5}$, we have

$$\det(5 \cdot D) = 5^3 \cdot \det(D) = 5^3 \cdot \frac{8}{5} = 200.$$

Recall that

$$(5 \cdot D) \cdot \operatorname{adj}(5 \cdot D) = \det(5 \cdot D) \cdot I_3.$$

Hence,

$$\det(5 \cdot D) \cdot \det(\operatorname{adj}(5 \cdot D)) = \det(5 \cdot D)^3 \cdot \det(I_3),$$

and so

$$\det(\operatorname{adj}(5 \cdot D)) = \det(5 \cdot D)^2 = 200^2 = 40000.$$

$\square$

Solution to Problem 6. Let us define the function $T : \mathbb{R}^3 \to \mathbb{R}$ by

$$T\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \det \begin{bmatrix} 0 & a & x \\ b & c & y \\ d & 1 & z \end{bmatrix},$$

for every $x, y, z \in \mathbb{R}$. Note that T is a linear transformation. From the problem statement, we find that

$$T\left(\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}\right) = -87 \qquad \text{and} \qquad T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = -33.$$

We have

$$T\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = T\left(\begin{bmatrix} 2 \\ 1 \\ 3 \end{bmatrix}\right) - 2 \cdot T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) = (-87) - 2(-33) = -21.$$

Therefore,

$$T\left(\begin{bmatrix} -4 \\ 5 \\ 1 \end{bmatrix}\right) = -4 \cdot T\left(\begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}\right) + 5 \cdot T\left(\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right) = -4(-33) + 5(-21) = 27.$$

Remark. The assumption in the problem is not vacuous. An example of (a, b, c, d) which satisfies the assumption is $(a, b, c, d) = (3, 8, 17, 1)$. $\square$

Solution to Problem 7. We argue that for any $i, j \in \{1, 2, 3, 4, 5\}$, the (i, j)-minor of G is 0. To see this, we take arbitrary indices i and j. Note that G has 3 copies of the row

$$\begin{bmatrix} 2 & 3 & 2 & 3 & 2 \end{bmatrix},$$

and has 2 copies of the row

$$\begin{bmatrix} 3 & 2 & 3 & 2 & 3 \end{bmatrix}.$$

If i is 1, 3, or 5, then after removing the i^{th} row from G, the remaining 4×5-matrix has two rows of each type. If i is 2 or 4, then after removing the i^{th} row from G, the remaining 4×5-matrix has 3 copies of the row of the first type, and 1 copy of the row of the second type.

Regardless of which value i is, the remaining 4×5-matrix has two identical rows. Next, we remove the j^{th} column from this matrix. Note that as we remove a column, if the 4×5-matrix has two identical rows, the resulting 4×4-matrix still has to have two identical rows. This means that no matter what values i and j might take, the resulting 4×4-matrix obtained from removing the i^{th} row and the j^{th} column from G has two identical rows, which implies that it has determinant 0.

Since i and j were arbitrary, all 25 minors of G are zero, which implies that all 25 cofactors of G are zero. Hence,

$$\operatorname{adj}(G) = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

$\square$

9. The Ninth Week

9.1. **Column replacement.** Let m and n be positive integers. Let us consider an $m \times n$ matrix $A \in \mathbb{R}^{m\times n}$. Let $i \in \{1, 2, \dots, n\}$ be an index. Let $b \in \mathbb{R}^m$ be an $m \times 1$-column vector.

Let us denote by $A \xleftarrow{\mathrm{col}_i} b$ the $m \times n$-matrix obtained by replacing the i^{th} column of A by b.

Example 9.1. Let

$$A := \begin{bmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{bmatrix}, \quad b := \begin{bmatrix} 0 \\ 3 \\ 5 \end{bmatrix}, \quad \text{and} \quad v := \begin{bmatrix} -1 \\ 0 \\ -7 \end{bmatrix}.$$

Then

$$A \xleftarrow{\mathrm{col}_2} b = \begin{bmatrix} 1 & 0 & 3 & 4 \\ 5 & 3 & 7 & 8 \\ 9 & 5 & 11 & 12 \end{bmatrix},$$

$$A \xleftarrow{\mathrm{col}_4} v = \begin{bmatrix} 1 & 2 & 3 & -1 \\ 5 & 6 & 7 & 0 \\ 9 & 10 & 11 & -7 \end{bmatrix},$$

$$\left(A \xleftarrow{\mathrm{col}_1} (b+v)\right) \xleftarrow{\mathrm{col}_3} (-3v) = \begin{bmatrix} -1 & 2 & 3 & 4 \\ 3 & 6 & 0 & 8 \\ -2 & 10 & 21 & 12 \end{bmatrix},$$

and

$$\left(A \xleftarrow{\mathrm{col}_4} v\right) \xleftarrow{\mathrm{col}_4} b = \begin{bmatrix} 1 & 2 & 3 & 0 \\ 5 & 6 & 7 & 3 \\ 9 & 10 & 11 & 5 \end{bmatrix}.$$

Previously we saw that if A is an $n \times n$-square matrix, $i \in \{1, 2, \dots, n\}$ is an index, then the function $T : \mathbb{R}^n \to \mathbb{R}$ given by

$$T(b) := \det(A \xleftarrow{\mathrm{col}_i} b)$$

is a linear transformation.

Example 9.2. Suppose that $B \in \mathbb{R}^{3\times 3}$ and $u, v \in \mathbb{R}^3$ are such that

$$\det(B \xleftarrow{\mathrm{col}_3} u) = 1 \quad \text{and} \quad \det(B \xleftarrow{\mathrm{col}_3} v) = 5.$$

Evaluate

$$\det\Big((B \xleftarrow{\mathrm{col}_3} u) + (B \xleftarrow{\mathrm{col}_3} v)\Big).$$

Note that

$$(B \xleftarrow{\mathrm{col}_3} u) + (B \xleftarrow{\mathrm{col}_3} v) = (2 \cdot B) \xleftarrow{\mathrm{col}_3} (u+v) = 2 \cdot \left(B \xleftarrow{\mathrm{col}_3} \frac{u+v}{2}\right).$$

Therefore,

$$\det\Big((B \xleftarrow{\text{col}_3} u) + (B \xleftarrow{\text{col}_3} v)\Big) = \det\left(2 \cdot \left(B \xleftarrow{\text{col}_3} \frac{u+v}{2}\right)\right).$$

Since the matrix $B \xleftarrow{\text{col}_3} \frac{u+v}{2}$ has dimension 3×3, we find that

$$\det\left(2 \cdot \left(B \xleftarrow{\text{col}_3} \frac{u+v}{2}\right)\right) = 2^3 \cdot \det\left(B \xleftarrow{\text{col}_3} \frac{u+v}{2}\right).$$

Now, by linearity,

$$\det\left(B \xleftarrow{\text{col}_3} \frac{u+v}{2}\right) = \frac{1}{2}\det\Big(B \xleftarrow{\text{col}_3} u\Big) + \frac{1}{2}\det\Big(B \xleftarrow{\text{col}_3} v\Big) = 3.$$

Combining the above equations, we find

$$\det\Big((B \xleftarrow{\text{col}_3} u) + (B \xleftarrow{\text{col}_3} v)\Big) = 2^3 \cdot 3 = 24.$$

9.2. **Cramer's rule.** Recall that for the following system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1, \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2, \\ \qquad\qquad \cdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n = b_m, \end{cases}$$

of m linear equations in the variables $x_1, x_2, \ldots, x_n$, where $a_{11}, a_{12}, \ldots, a_{mn}, b_1, \ldots, b_m$ are constants, the *augmented matrix* of the linear system is the $m \times (n+1)$-matrix

$$A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix},$$

and the *coefficient matrix* of the linear system is the $m \times n$-matrix

$$C := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}.$$

When $m = n$ (that is, when there are as many linear equations as there are variables), the coefficient matrix C is a square matrix.

Theorem 9.1 (Cramer's Rule)**.** *Suppose that we have a system of n linear equations in the n variables, $x_1, x_2, \ldots, x_n$ (listed in this order), with the augmented matrix*

$$A := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} & b_n \end{bmatrix} \in \mathbb{R}^{n\times(n+1)}$$

and the coefficient matrix

$$C := \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \in \mathbb{R}^{n\times n}.$$

Let

$$\mathbf{b} := \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \in \mathbb{R}^{n\times 1}$$

denote the rightmost column of the augmented matrix A.

If the square matrix C is invertible, then the linear system has a unique solution[29] *given by*

$$x_1 = \frac{\det(C \xleftarrow{\mathrm{col}_1} \mathbf{b})}{\det(C)}, \quad x_2 = \frac{\det(C \xleftarrow{\mathrm{col}_2} \mathbf{b})}{\det(C)}, \quad \ldots \quad , \quad x_n = \frac{\det(C \xleftarrow{\mathrm{col}_n} \mathbf{b})}{\det(C)}.$$

Example 9.3. Consider the system

$$\begin{cases} 2x - y = 1, \\ x + y = 5, \end{cases}$$

of two linear equations in the two variables x and y.

Using the notations from the theorem above, we have

$$C = \begin{bmatrix} 2 & -1 \\ 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 5 \end{bmatrix}.$$

Note that

$$\det(C) = 3 \neq 0,$$

and thus C is invertible. Cramer's rule implies that the system has a unique solution with

$$x = \frac{\det(C \xleftarrow{\mathrm{col}_1} \mathbf{b})}{\det(C)} = \frac{1}{3}\det\begin{bmatrix} 1 & -1 \\ 5 & 1 \end{bmatrix} = \frac{1}{3}\cdot 6 = 2,$$

[29]This means that the size of the solution set is one.

and

$$y = \frac{\det(C \xleftarrow{\mathrm{col}_2} \mathbf{b})}{\det(C)} = \frac{1}{3}\det\begin{bmatrix} 2 & 1 \\ 1 & 5 \end{bmatrix} = \frac{1}{3} \cdot 9 = 3.$$

Hence, the solution set of the linear system is $\{(2,3)\}$.

Let us check that $(2,3)$ is indeed a solution. Plugging in $x = 2$ and $y = 3$, we find

$$2x - y = 2 \cdot 2 - 3 = 4 - 3 = 1,$$

and

$$x + y = 2 + 3 = 5.$$

Thus, $(x, y) = (2, 3)$ is indeed a solution.

Example 9.4. Consider the system

$$\begin{cases} 2x + z = 1, \\ y + 3z = 5, \\ x + y + z = 3 \end{cases}$$

of three linear equations in the variables x, y, z.

With the same notations, we have

$$C = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ 3 \end{bmatrix}.$$

Note that $\det(C) = -5 \neq 0$, and thus C is invertible.

By Cramer's rule, we find that $(x, y, z) \in \mathbb{R}^3$ where

$$x = \frac{\det(C \xleftarrow{\mathrm{col}_1} \mathbf{b})}{\det(C)} = \frac{1}{-5} \cdot \det\begin{bmatrix} 1 & 0 & 1 \\ 5 & 1 & 3 \\ 3 & 1 & 1 \end{bmatrix} = 0,$$

$$y = \frac{\det(C \xleftarrow{\mathrm{col}_2} \mathbf{b})}{\det(C)} = \frac{1}{-5} \cdot \det\begin{bmatrix} 2 & 1 & 1 \\ 0 & 5 & 3 \\ 1 & 3 & 1 \end{bmatrix} = 2,$$

and

$$z = \frac{\det(C \xleftarrow{\mathrm{col}_3} \mathbf{b})}{\det(C)} = \frac{1}{-5} \cdot \det\begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 5 \\ 1 & 1 & 3 \end{bmatrix} = 1,$$

is the unique solution. The solution set is $\{(0, 2, 1)\}$.

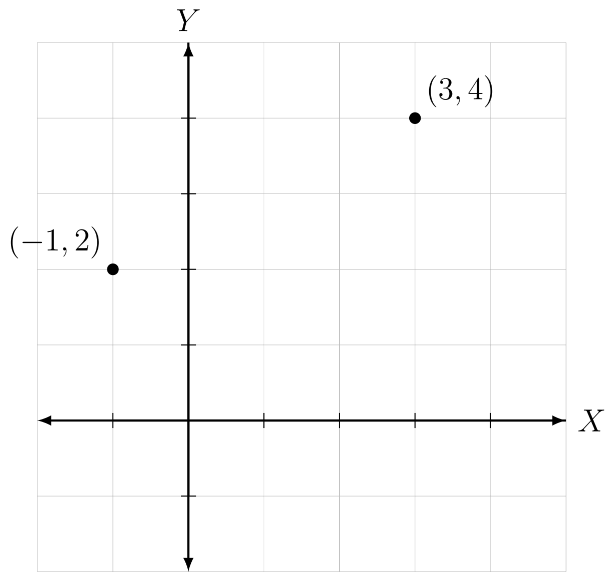

FIGURE 1. Points $(-1, 2)$ and $(3, 4)$ on $\mathbb{R}^2$

9.3. **Euclidean geometry.** In this class, we have seen that elements in the set $\mathbb{R}^n$ can be represented as either an n-tuple or an $n \times 1$-column vector. For instance, in $\mathbb{R}^2$, an element in it is the 2-tuple $(3, 4)$, which we can also view it as the column vector

$$\begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

Note that when we write this element as a 2-tuple, we use the *parentheses* instead of the *square brackets*. This helps us avoid thinking of it as a 1×2-row vector.

Using the Cartesian coordinates, we find that the set $\mathbb{R}^n$ can be used as a model for the n-dimensional space. For example, $\mathbb{R}^3$ is a model for the 3-dimensional physical space, and $\mathbb{R}^2$ is a model for the 2-dimensional flat space.

For instance, the two elements

$$(-1, 2), (3, 4) \in \mathbb{R}^2$$

can be thought of as two *points* on the plane. See Figure 1 for an illustration.

In general, when an element $P \in \mathbb{R}^n$—a point in the n-dimensional space—is represented by an n-tuple

$$P = (p_1, p_2, \ldots, p_n) \in \mathbb{R}^n,$$

we say that $(p_1, p_2, \ldots, p_n)$ is the **coordinates** of P. If $n = 2$, and $P = (p_1, p_2)$, we say that the **X-coordinate** of P is p_1 and the **Y-coordinate** of P is p_2. Similarly, if $n = 3$, and $P = (p_1, p_2, p_3)$, we say that the **X-coordinate** of P is p_1, the **Y-coordinate** of P is p_2, and the **Z-coordinate** of P is p_3.

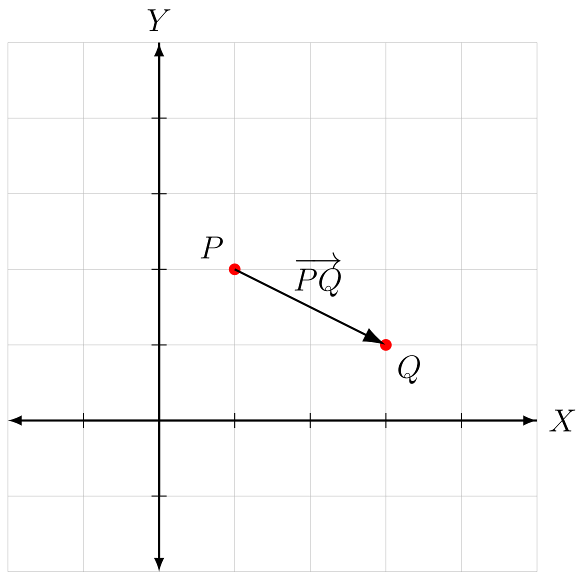

FIGURE 2. the vector $\overrightarrow{PQ}$

Now suppose that

$$P = (p_1, p_2, \ldots, p_n) \quad \text{and} \quad Q = (q_1, q_2, \ldots, q_n)$$

are two points in $\mathbb{R}^n$. We can define the column vector

$$\overrightarrow{PQ} := \begin{bmatrix} q_1 - p_1 \\ q_2 - p_2 \\ \vdots \\ q_n - p_n \end{bmatrix} \in \mathbb{R}^n.$$

Example 9.5. If $P = (1, 2) \in \mathbb{R}^2$ and $Q = (3, 1) \in \mathbb{R}^2$, then

$$\overrightarrow{PQ} = \begin{bmatrix} 2 \\ -1 \end{bmatrix} \in \mathbb{R}^2.$$

See Figure 2 for an illustration.

Proposition 9.2. *For every pair of points P and Q in $\mathbb{R}^n$, we have*

$$\overrightarrow{PQ} = -\overrightarrow{QP}.$$

Proof. Write $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$. We have

$$\overrightarrow{PQ} = \begin{bmatrix} q_1 - p_1 \\ q_2 - p_2 \\ \vdots \\ q_n - p_n \end{bmatrix},$$

and

$$\overrightarrow{QP} = \begin{bmatrix} p_1 - q_1 \\ p_2 - q_2 \\ \vdots \\ p_n - q_n \end{bmatrix}.$$

Thus, $\overrightarrow{PQ} = -\overrightarrow{QP}$. ☐

Proposition 9.3. *For points $P, Q, R \in \mathbb{R}^n$, we have*

$$\overrightarrow{PQ} + \overrightarrow{QR} = \overrightarrow{PR}.$$

Proof. Let us write $P = (p_1, p_2, \ldots, p_n)$, $Q = (q_1, q_2, \ldots, q_n)$, and $R = (r_1, r_2, \ldots, r_n)$. We have

$$\overrightarrow{PQ} + \overrightarrow{QR} = \begin{bmatrix} q_1 - p_1 \\ q_2 - p_2 \\ \vdots \\ q_n - p_n \end{bmatrix} + \begin{bmatrix} r_1 - q_1 \\ r_2 - q_2 \\ \vdots \\ r_n - q_n \end{bmatrix} = \begin{bmatrix} r_1 - p_1 \\ r_2 - p_2 \\ \vdots \\ r_n - p_n \end{bmatrix} = \overrightarrow{PR},$$

as desired. ☐

A vector $v \in \mathbb{R}^n$ is said to be the **zero vector** if all its entries are zero; in other words, it is the zero $n \times 1$-matrix. If v is not a zero vector, then we say that v is a **nonzero** vector.

Two nonzero vectors $u, v \in \mathbb{R}^n$ are said to be **parallel** if there is a nonzero real number $\lambda \in \mathbb{R}$ for which

$$u = \lambda \cdot v.$$

For example, the vectors

$$\begin{bmatrix} 1 \\ 0 \\ -2 \end{bmatrix} \text{ and } \begin{bmatrix} -3 \\ 0 \\ 6 \end{bmatrix}$$

are parallel.

Suppose that $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$ are two points in $\mathbb{R}^n$. The **midpoint** between P and Q (or the midpoint of the line segment $\overline{PQ}$) is the point $M \in \mathbb{R}^n$ with coordinates

$$M = \left(\frac{p_1 + q_1}{2}, \frac{p_2 + q_2}{2}, \ldots, \frac{p_n + q_n}{2} \right) \in \mathbb{R}^n.$$

Example 9.6. In $\mathbb{R}^3$, the midpoint between $(-3, 0, 2)$ and $(7, -2, 2)$ is the point $(2, -1, 2)$.

Proposition 9.4. *Let P and Q be two points in $\mathbb{R}^n$, and let M be the midpoint between P and Q. Then*

$$\overrightarrow{PM} = \overrightarrow{MQ}.$$

Proof. Write $P = (p_1, p_2, \ldots, p_n)$ and $Q = (q_1, q_2, \ldots, q_n)$. Note that both $\overrightarrow{PM}$ and $\overrightarrow{MQ}$ are the vector

$$\begin{bmatrix} (q_1 - p_1)/2 \\ (q_2 - p_2)/2 \\ \cdots \\ (q_n - p_n)/2 \end{bmatrix}.$$

Hence, $\overrightarrow{PM} = \overrightarrow{MQ}$. $\square$

Example 9.7. Let P and Q be two different points in $\mathbb{R}^n$. Let M be the midpoint of P and Q, and let N be the midpoint of P and M. Determine the real number λ for which

$$\overrightarrow{NQ} = \lambda \cdot \overrightarrow{NP}.$$

Note that

$$\overrightarrow{NM} = \overrightarrow{PN} = -\overrightarrow{NP}$$

and

$$\overrightarrow{MQ} = \overrightarrow{PM} = \overrightarrow{PN} + \overrightarrow{NM} = 2 \cdot \overrightarrow{PN} = -2 \cdot \overrightarrow{NP}.$$

Therefore,

$$\overrightarrow{NQ} = \overrightarrow{NM} + \overrightarrow{MQ} = -3 \cdot \overrightarrow{NP}.$$

It is not hard to see that $\overrightarrow{NP}$ is not the zero vector, so we can conclude that $\lambda = -3$.

9.4. **A bit of Quiz II review.**

Problem 9.8. Let

$$A := \begin{bmatrix} -1 & -2 \\ 4 & 7 \end{bmatrix}.$$

What is A^{T}? What is $\mathrm{tr}(A)$? What is $\det(A)$? What is $\mathrm{adj}(A)$? What is A^{-1}?

What are the symmetric and the antisymmetric parts of A?

Solution. Note that A^{T} denotes the *transpose* of A, while $\mathrm{tr}(A)$ denotes the trace of A. The transpose and the trace are different! The transpose is a matrix, while the trace is a number!

The *transpose* A^{T} of A is

$$A^{\mathsf{T}} = \begin{bmatrix} -1 & 4 \\ -2 & 7 \end{bmatrix}.$$

The *trace* $\mathrm{tr}(A)$ of A is

$$\mathrm{tr}(A) = (-1) + 7 = 6.$$

The *determinant* $\det(A)$ of A is

$$\det(A) = (-1)\cdot 7 - (-2)\cdot 4 = (-7) - (-8) = 1.$$

The *adjugate* $\operatorname{adj}(A)$ of A is

$$\operatorname{adj}(A) = \begin{bmatrix} 7 & 2 \\ -4 & -1 \end{bmatrix}.$$

The *inverse* A^{-1} of A is

$$A^{-1} = \frac{1}{\det(A)} \cdot \operatorname{adj}(A) = \frac{1}{1} \cdot \begin{bmatrix} 7 & 2 \\ -4 & -1 \end{bmatrix} = \begin{bmatrix} 7 & 2 \\ -4 & -1 \end{bmatrix}.$$

Recall that the *symmetric part* X and the *antisymmetric part* Y of A are the matrices such that

(i) X is symmetric,

(ii) Y is antisymmetric, and

(iii) $A = X + Y$.

The symmetric part of A is

$$X = \frac{A + A^{\intercal}}{2} = \begin{bmatrix} -1 & 1 \\ 1 & 7 \end{bmatrix},$$

and the antisymmetric part of A is

$$Y = \frac{A - A^{\intercal}}{2} = \begin{bmatrix} 0 & -3 \\ 3 & 0 \end{bmatrix}.$$

Note that all the diagonal entries of the antisymmetric part are zero. For every antisymmetric matrix, all diagonal entries are zero. □

Problem 9.9. Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be a linear transformation such that

$$T\left(\begin{bmatrix} 2 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 2 \\ 10 \end{bmatrix} \quad \text{and} \quad T\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 1 \\ 11 \end{bmatrix}.$$

Find the matrix M such that

$$\forall v \in \mathbb{R}^2,\ M \cdot v = T(v).$$

What is the dimension of M?

Solution. Note that since $M \cdot v = T(v)$ holds for every $v \in \mathbb{R}^2$, this means that when we multiply M to the left of a 2×1-column vector, we obtain a 3×1-column vector. Hence, the dimension of M has to be 3×2.

Now recall that we can find M by

$$M = \big[T(e_1) \,|\, T(e_2)\big], \tag{$*$}$$

where e_1 and e_2 are the standard basis vectors of $\mathbb{R}^2$.

Since T is a linear transformation, we have

$$T\left(\begin{bmatrix}3\\3\end{bmatrix}\right) = T\left(\begin{bmatrix}2\\1\end{bmatrix}\right) + T\left(\begin{bmatrix}1\\2\end{bmatrix}\right) = \begin{bmatrix}-1\\2\\10\end{bmatrix} + \begin{bmatrix}1\\1\\11\end{bmatrix} = \begin{bmatrix}0\\3\\21\end{bmatrix}.$$

Using that T is a linear transformation again, we find

$$T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \frac{1}{3} \cdot T\left(\begin{bmatrix}3\\3\end{bmatrix}\right) = \frac{1}{3} \cdot \begin{bmatrix}0\\3\\21\end{bmatrix} = \begin{bmatrix}0\\1\\7\end{bmatrix}.$$

Continuing, we have

$$T(e_1) = T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = T\left(\begin{bmatrix}2\\1\end{bmatrix}\right) - T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \begin{bmatrix}-1\\2\\10\end{bmatrix} - \begin{bmatrix}0\\1\\7\end{bmatrix} = \begin{bmatrix}-1\\1\\3\end{bmatrix}.$$

Similarly, we have

$$T(e_2) = T\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = T\left(\begin{bmatrix}1\\2\end{bmatrix}\right) - T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = \begin{bmatrix}1\\1\\11\end{bmatrix} - \begin{bmatrix}0\\1\\7\end{bmatrix} = \begin{bmatrix}1\\0\\4\end{bmatrix}.$$

Using $(*)$, we conclude

$$M = \begin{bmatrix}-1 & 1\\1 & 0\\3 & 4\end{bmatrix}.$$

This is our desired matrix. $\square$

9.5. **Problems and Solutions.**

Problem 1. Suppose that $A \in \mathbb{R}^{4\times 4}$ and $u, v \in \mathbb{R}^4$ satisfy

$$\det\Big(A \xleftarrow{\mathrm{col}_2} (u+v)\Big) = 5 \quad \text{and} \quad \det\Big(A \xleftarrow{\mathrm{col}_2} (u-v)\Big) = 1.$$

Evaluate

$$\det\Big(4 \cdot (A \xleftarrow{\mathrm{col}_2} u) - (A \xleftarrow{\mathrm{col}_2} v)\Big).$$

Problem 2. Use Cramer's rule to solve the following linear system.

$$\begin{cases} x + 2y = 1, \\ y + 2z = 10, \\ z + 2x = 16. \end{cases}$$

Problem 3. Consider the following three points

$$A = (0, 2, -1), \quad B = (-3, 1, 4), \quad \text{and} \quad C = (5, -1, 0)$$

in the three-dimensional space $\mathbb{R}^3$.

Let $X, Y, Z \in \mathbb{R}^3$ be such that

- X is the midpoint between A and Z,
- Y is the midpoint between B and X,
- Z is the midpoint between C and Y.

(a) Find the coordinates of the point X.

(b) Evaluate

$$\overrightarrow{XA} + \overrightarrow{XB} + \overrightarrow{XC} + \overrightarrow{YA} + \overrightarrow{YB} + \overrightarrow{YC} + \overrightarrow{ZA} + \overrightarrow{ZB} + \overrightarrow{ZC}.$$

Hint for part **(a)**. Here is an approach you might take. Let us use vectors to help us. Define

$$u := \overrightarrow{XY} \qquad \text{and} \qquad v := \overrightarrow{YZ}.$$

Note that since $\overrightarrow{XZ} = \overrightarrow{XY} + \overrightarrow{YZ}$, we can write $\overrightarrow{XZ}$ in terms of u and v. Next, because X is the midpoint between A and Z, we can now also write $\overrightarrow{AZ}$ in terms of u and v. Next, write $\overrightarrow{ZC}$ in terms of u and v.

Now note that $\overrightarrow{AC} = \overrightarrow{AZ} + \overrightarrow{ZC}$, so we can write $\overrightarrow{AC}$ in terms of u and v.

On the other hand, $\overrightarrow{AC}$ can be computed directly from the problem. So we learn something about u and v.

Now you can do similar things with $\overrightarrow{AB}$ and $\overrightarrow{BC}$ to learn more about u and v.

Then you should be able to find u and v. Use it to compute the coordinates of X.

Solution to Problem 1. Let us write

$$A = [\, C_1 \,|\, C_2 \,|\, C_3 \,|\, C_4 \,],$$

where C_i denotes the i^{th} column of A. Also write

$$X := 4 \cdot (A \xleftarrow{\text{col}_2} u) - (A \xleftarrow{\text{col}_2} v).$$

We would like to compute $\det(X)$.

Note that

$$\begin{aligned} X &= [\, 4 \cdot C_1 \,|\, 4 \cdot u \,|\, 4 \cdot C_3 \,|\, 4 \cdot C_4 \,] - [\, C_1 \,|\, v \,|\, C_3 \,|\, C_4 \,] \\ &= [\, 3 \cdot C_1 \,|\, 4 \cdot u - v \,|\, 3 \cdot C_3 \,|\, 3 \cdot C_4 \,] \\ &= 3 \cdot [\, C_1 \,|\, (4/3) \cdot u - (1/3)v \,|\, C_3 \,|\, C_4 \,]. \end{aligned}$$

Define the function $T : \mathbb{R}^4 \to \mathbb{R}$ by

$$T(w) := \det\left(A \xleftarrow{\text{col}_2} w\right),$$

for every $w \in \mathbb{R}^4$. Note that T is a linear transformation. The formula for X above implies that

$$\det(X) = \det\left(3 \cdot [\, C_1 \,|\, (4/3) \cdot u - (1/3)v \,|\, C_3 \,|\, C_4 \,]\right) = 3^4 \cdot T\left(\frac{4}{3} \cdot u - \frac{1}{3} \cdot v\right).$$

Because T is a linear transformation, we deduce

$$\det(X) = 3^4 \cdot \left(\frac{4}{3}T(u) - \frac{1}{3}T(v)\right).$$

From the assumption in the problem, we have

$$T(u+v) = 5 \qquad \text{and} \qquad T(u-v) = 1.$$

Therefore,

$$T(u) = \frac{1}{2}T(u+v) + \frac{1}{2}T(u-v) = \frac{5+1}{2} = 3,$$

and

$$T(v) = \frac{1}{2}T(u+v) - \frac{1}{2}T(u-v) = \frac{5-1}{2} = 2.$$

Using these values in the formula for $\det(X)$ above, we find

$$\det(X) = 3^4 \cdot \left(\frac{4}{3} \cdot 3 - \frac{1}{3} \cdot 2\right) = 270.$$

Remark. An example of A, u, v which satisfies the assumption in the problem is $A = I_4$, $u = \begin{bmatrix} 0 & 3 & 0 & 0 \end{bmatrix}^{\mathtt{T}}$, and $v = \begin{bmatrix} 0 & 2 & 0 & 0 \end{bmatrix}^{\mathtt{T}}$. $\square$

Solution to Problem 2. The coefficient matrix for the linear system is

$$C = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 1 & 2 \\ 2 & 0 & 1 \end{bmatrix}.$$

Note that by the rule of Sarrus, we have

$$\det(C) = 1 + 8 + 0 - 0 - 0 - 0 = 9.$$

Cramer's rule applies. We find that the linear system has a unique solution with

$$x = \frac{1}{\det(C)} \cdot \det \begin{bmatrix} 1 & 2 & 0 \\ 10 & 1 & 2 \\ 16 & 0 & 1 \end{bmatrix} = \frac{1}{9}(1 + 64 + 0 - 0 - 0 - 20) = 5,$$

$$y = \frac{1}{\det(C)} \cdot \det \begin{bmatrix} 1 & 1 & 0 \\ 0 & 10 & 2 \\ 2 & 16 & 1 \end{bmatrix} = \frac{1}{9}(10 + 4 + 0 - 0 - 32 - 0) = -2,$$

and

$$z = \frac{1}{\det(C)} \cdot \det \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 10 \\ 2 & 0 & 16 \end{bmatrix} = \frac{1}{9}(16 + 40 + 0 - 2 - 0 - 0) = 6.$$

Thus, we have found that $(x, y, z) = (5, -2, 6)$.

Let us check that $(x, y, z) = (5, -2, 6)$ is indeed a solution to the linear system. Note that

$$\begin{aligned} x + 2y &= 5 + 2(-2) = 1, \\ y + 2z &= (-2) + 2(6) = 10, \\ z + 2x &= 6 + 2(5) = 16. \end{aligned}$$

Hence, $(5, -2, 6)$ is indeed a solution.

The solution set to the linear system is $\{(5, -2, 6)\}$. $\square$

Solution to Problem 3. (a) Let $u := \overrightarrow{XY}$ and $v := \overrightarrow{YZ}$. Note that $\overrightarrow{XZ} = u + v$. We have $\overrightarrow{AX} = u + v$ and $\overrightarrow{ZC} = v$. Thus,

$$\overrightarrow{AC} = \overrightarrow{AX} + \overrightarrow{XZ} + \overrightarrow{ZC} = (u + v) + (u + v) + v = 2u + 3v.$$

On the other hand, from the assumption in the problem, we have

$$\overrightarrow{AC} = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}.$$

Therefore,

$$2u + 3v = \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix}. \tag{26}$$

Similarly, we have $\overrightarrow{AB} = 3u + v$. This gives

$$3u + v = \begin{bmatrix} -3 \\ -1 \\ 5 \end{bmatrix}. \tag{27}$$

Combining (26) and (27), we find

$$7u = 3 \cdot (3u + v) - (2u + 3v) = \begin{bmatrix} -14 \\ 0 \\ 14 \end{bmatrix},$$

whence

$$u = \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix}.$$

Plugging this value of u back in (27) above to obtain

$$v = \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix}.$$

Now let $O = (0, 0, 0)$ denote the origin in $\mathbb{R}^3$. We find that

$$\overrightarrow{OX} = \overrightarrow{OA} + \overrightarrow{AX} = \begin{bmatrix} 0 \\ 2 \\ -1 \end{bmatrix} + \begin{bmatrix} -2 \\ 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 3 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Therefore, $X = (1, 1, 0) \in \mathbb{R}^3$.

(b) Note that

$$\overrightarrow{XA} + \overrightarrow{YB} + \overrightarrow{ZC} = \overrightarrow{ZX} + \overrightarrow{XY} + \overrightarrow{YZ} = \overrightarrow{ZZ} = \mathbf{0},$$
$$\overrightarrow{XB} + \overrightarrow{YC} + \overrightarrow{ZA} = 2 \cdot \overrightarrow{XY} + 2 \cdot \overrightarrow{YZ} + 2 \cdot \overrightarrow{ZX} = 2 \cdot \overrightarrow{XX} = \mathbf{0},$$

and

$$\begin{aligned} \overrightarrow{XC} + \overrightarrow{YA} + \overrightarrow{ZB} &= \overrightarrow{XA} + \overrightarrow{AC} + \overrightarrow{YB} + \overrightarrow{BA} + \overrightarrow{ZC} + \overrightarrow{CB} \\ &= (\overrightarrow{XA} + \overrightarrow{YB} + \overrightarrow{ZC}) + (\overrightarrow{AC} + \overrightarrow{CB} + \overrightarrow{BA}) \\ &= \mathbf{0} + \mathbf{0} = \mathbf{0}. \end{aligned}$$

Therefore, the sum of the nine vectors in the problem is $\mathbf{0}$. □

10. The Tenth Week

10.1. **Vector geometry.** We continue our discussion of studying the geometry of Euclidean spaces using vectors. Throughout this section, let n be a positive integer. Suppose that $v \in \mathbb{R}^n$ is a column vector in the n-dimensional space $\mathbb{R}^n$ with

$$v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

We define the **norm** (or synonymously the **length**, or synonymously the **magnitude**) of v to be

$$\|v\| := \sqrt{\sum_{i=1}^{n} x_i^2}.$$

If P and Q are two points in $\mathbb{R}^n$, then the **distance** between P and Q is defined to be the norm of the vector $\overrightarrow{PQ}$.

Example 10.1. Consider Figure 3. Let $P = (0,-1) \in \mathbb{R}^2$ and $Q = (3,3) \in \mathbb{R}^2$. What is the distance between P and Q?

We note that

$$\overrightarrow{PQ} = \begin{bmatrix} 3-0 \\ 3-(-1) \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \end{bmatrix}.$$

The norm of the vector $\overrightarrow{PQ}$ is

$$\|\overrightarrow{PQ}\| = \sqrt{3^2+4^2} = 5.$$

Thus, the distance between P and Q is 5.

From the definition of norms, we have the following immediate consequences for every vector $v \in \mathbb{R}^n$ and every real number $\lambda \in \mathbb{R}$,

- the norm $\|v\|$ is a nonnegative real number.
- we have $\|\lambda \cdot v\| = |\lambda| \cdot \|v\|$.
- we have $\|v\| = 0$ *if and only if* v is the zero vector.

A vector $v \in \mathbb{R}^n$ is said to be a **unit vector** if $\|v\| = 1$.

Proposition 10.1. *Let $v \in \mathbb{R}^n$ be any nonzero vector. Then $\frac{v}{\|v\|}$ is a unit vector.*

Proof. Let us denote $\lambda := 1/\|v\|$ and $u := \frac{v}{\|v\|}$. Note that $\|u\| = \|\lambda \cdot v\| = |\lambda| \cdot \|v\| = \frac{\|v\|}{\|v\|} = 1$. $\square$

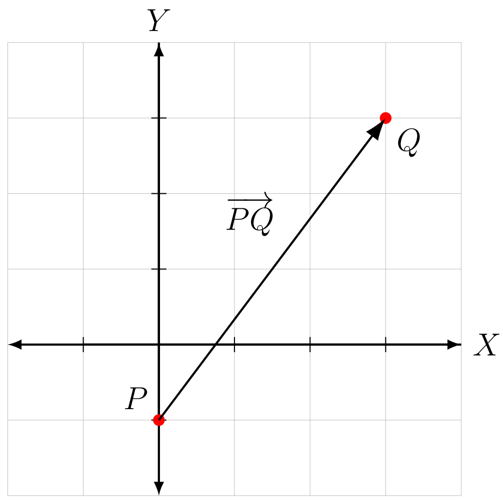

FIGURE 3. the vector $\overrightarrow{PQ}$

When $v \neq \mathbf{0}$, the vector $\frac{v}{\|v\|}$ is called the **unit vector in the same direction as** v.

Example 10.2. Recall that in $\mathbb{R}^n$, there are *standard basis vectors*

$$e_1, e_2, \ldots, e_n \in \mathbb{R}^n.$$

Each of the standard basis vectors is a unit vector.

Proposition 10.2. *Let $a_1, a_2, \ldots, a_n$ be real numbers. Let*

$$v := a_1e_1 + a_2e_2 + \cdots + a_ne_n,$$

where $e_1, e_2, \ldots, e_n$ are the standard basis vectors in $\mathbb{R}^n$. Then

$$\|v\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}.$$

Proof. Note that v can be written as

$$v = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}.$$

Therefore, $\|v\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2}$. $\square$

Definition 10.3. Let

$$u := \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{bmatrix} \quad \text{and} \quad v := \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

be vectors in $\mathbb{R}^n$. Then the **Euclidean inner product** (or just the **inner product**, or synonymously the **dot product**) between u and v is defined as

$$\langle u, v\rangle := u_1 \cdot v_1 + u_2 \cdot v_2 + \cdots + u_n \cdot v_n.$$

Example 10.4. The inner product between

$$u := \begin{bmatrix} 0 \\ 1 \\ 2 \\ 3 \end{bmatrix} \qquad \text{and} \qquad v := \begin{bmatrix} -1 \\ 3 \\ -1 \\ 0 \end{bmatrix}$$

is

$$\langle u, v\rangle = 0 \cdot (-1) + 1 \cdot 3 + 2 \cdot (-1) + 3 \cdot 0 = 1.$$

Note that for every vector $u \in \mathbb{R}^n$, we have

$$\|u\|^2 = \langle u, u\rangle.$$

Proposition 10.3. *Let $u, v, w \in \mathbb{R}^n$, and let $a, b \in \mathbb{R}$. We have the following.*

(i) $\langle u, v\rangle = \langle v, u\rangle$.

(ii) $\langle u, a \cdot v + b \cdot w\rangle = a \cdot \langle u, v\rangle + b \cdot \langle u, w\rangle$.

(iii) $\langle a \cdot u + b \cdot v, w\rangle = a \cdot \langle u, w\rangle + b \cdot \langle v, w\rangle$.

(iv) $a \cdot \langle u, v\rangle = \langle a \cdot u, v\rangle = \langle u, a \cdot v\rangle$.

The proposition above says that the inner product indeed behaves like a product. For instance, if $u, v \in \mathbb{R}^n$, then

$$\langle u + v, u + v\rangle = \langle u, u\rangle + 2 \cdot \langle u, v\rangle + \langle v, v\rangle,$$

and

$$\langle u + v, u - v\rangle = \langle u, u\rangle - \langle v, v\rangle.$$

Let us recall the definition of the *arccosine* function. The *arccosine* function,

$$\arccos : [-1, 1] \to [0, \pi],$$

is given by:

$$\arccos(x) = \theta \in [0, \pi] \qquad \text{if and only if} \qquad \cos(\theta) = x \in [-1, 1].$$

For example,

$$\arccos\left(-\frac{1}{2}\right) = \frac{2\pi}{3} \qquad \text{and} \qquad \arccos(-1) = \pi.$$

Remark 10.1. In other words, "$\arccos(x) = \theta$" means that $\cos\theta = x$ and $0° \leq \theta \leq 180°$.

Exercise 10.5. Simplify the following expressions. Provide your answers in both degrees and radians.

(a) $\arccos(\cos(-160°))$.

(b) $\arccos(\cos(-10°))$.

(c) $\arccos(\cos(100°))$.

(d) $\arccos(\cos(280°))$.

(e) $\arccos(\cos(720°))$.

(f) $\arccos(\cos(770°))$.

(g) $\arccos(\cos(900°))$.

(h) $\arccos(\sin(100°)) + \arccos(\sin(500°))$.

The answers are given in the footnote.[30]

Definition 10.6. Let $u, v \in \mathbb{R}^n$. The **angle** between u and v is given by

$$\arccos\left(\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}\right).$$

In particular, if the angle between u and v is θ, then we have

$$\langle u, v\rangle = \|u\| \cdot \|v\| \cdot \cos(\theta).$$

Example 10.7. Let

$$u := \begin{bmatrix} -2 \\ 2 \end{bmatrix} \qquad \text{and} \qquad v := \begin{bmatrix} 0 \\ -3 \end{bmatrix}.$$

What is the angle between u and v?

We have

$$\langle u, v\rangle = (-2) \cdot 0 + 2 \cdot (-3) = -6,$$
$$\|u\| = \sqrt{(-2)^2 + 2^2} = \sqrt{8} = 2 \cdot \sqrt{2},$$

and

$$\|v\| = \sqrt{0^2 + (-3)^2} = \sqrt{9} = 3.$$

Therefore, the angle between u and v is

$$\arccos\left(\frac{-6}{2\sqrt{2} \cdot 3}\right) = \arccos\left(-\frac{1}{\sqrt{2}}\right) = \frac{3\pi}{4} = 135°.$$

[30] *Answers.* (a) $160° = \frac{8\pi}{9}$. (b) $10° = \frac{\pi}{18}$. (c) $100° = \frac{5\pi}{9}$. (d) $80° = \frac{4\pi}{9}$. (e) $0° = 0$. (f) $50° = \frac{5\pi}{18}$. (g) $180° = \pi$. (h) $60° = \frac{\pi}{3}$. ■

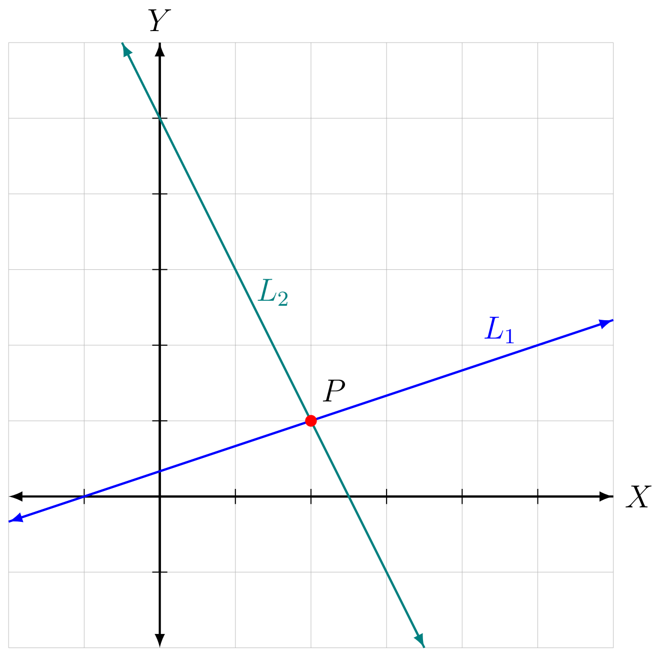

FIGURE 4. An intersection of two lines

Two vectors u and v are said to be **perpendicular** if $\langle u, v\rangle = 0$.

10.2. **Hyperplanes.** Let us define *hyperplanes* in $\mathbb{R}^n$.

In the following definition, let $O \in \mathbb{R}^n$ denote the *origin*:

$$O := (0, 0, \ldots, 0) \in \mathbb{R}^n.$$

Definition 10.8. A **hyperplane** H is a subset $H \subseteq \mathbb{R}^n$ of the form

$$H := \left\{X \in \mathbb{R}^n \,:\, \left\langle u, \overrightarrow{OX}\right\rangle = t\right\},$$

for a nonzero vector $u \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ and a real number $t \in \mathbb{R}$.

A hyperplane in $\mathbb{R}^n$ is a $(n-1)$-dimensional object: it has one dimension lower than the *ambient* space $\mathbb{R}^n$. For example, a hyperplane in $\mathbb{R}^1$ is a point. A hyperplane in $\mathbb{R}^2$ is a line. A hyperplane in $\mathbb{R}^3$ is a two-dimensional plane.

Example 10.9. In $\mathbb{R}^2$, hyperplanes are *lines*. Let

$$u := \begin{bmatrix} 1 \\ -3 \end{bmatrix} \qquad \text{and} \qquad v := \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Consider the following two lines:

$$L_1 := \left\{X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX}\right\rangle = -1\right\},$$

and

$$L_2 := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle v, \overrightarrow{OX} \right\rangle = 5 \right\}.$$

(See Figure 4 for an illustration.) Determine the intersection P between L_1 and L_2.

Let us write $P = (x_1, x_2) \in \mathbb{R}^2$. We have

$$\overrightarrow{OP} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

Since $P \in L_1$, we have

$$x_1 - 3 \cdot x_2 = -1.$$

Similarly, since $P \in L_2$, we have

$$2 \cdot x_1 + x_2 = 5.$$

Thus, we have the system

$$\begin{cases} x_1 - 3 \cdot x_2 = -1, \\ 2 \cdot x_1 + x_2 = 5, \end{cases}$$

of two linear equations in the variables x_1, x_2.

This system has a unique solution $(x_1, x_2) = (2, 1)$. Therefore, $P = (2, 1)$.

Suppose that H is an $(n-1)$-dimensional hyperplane in $\mathbb{R}^n$. Let $P \in \mathbb{R}^n$ be a point. The **image of the orthogonal projection** of P onto H is the unique point $Q \in H$ such that for every point $R \in H$, we have

$$\left\langle \overrightarrow{PQ}, \overrightarrow{QR} \right\rangle = 0.$$

The image of the orthogonal projection of P is the unique point on the hyperplane H that is closest to P. If Q is the image of the orthogonal projection of P onto H, then we define the **distance from the point P to the hyperplane H** to be the distance between P and Q.

We have the following formula.

Proposition 10.4. *Let $P \in \mathbb{R}^n$ be a point. Let $O \in \mathbb{R}^n$ denote the origin in $\mathbb{R}^n$. Suppose that $u \in \mathbb{R}^n$ is a vector, and $t \in \mathbb{R}$ is a real number. Let*

$$H := \left\{ x \in \mathbb{R}^n \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = t \right\}.$$

Then the distance from P to H is given by

$$\frac{\left| \left\langle u, \overrightarrow{OP} \right\rangle - t \right|}{\|u\|}.$$

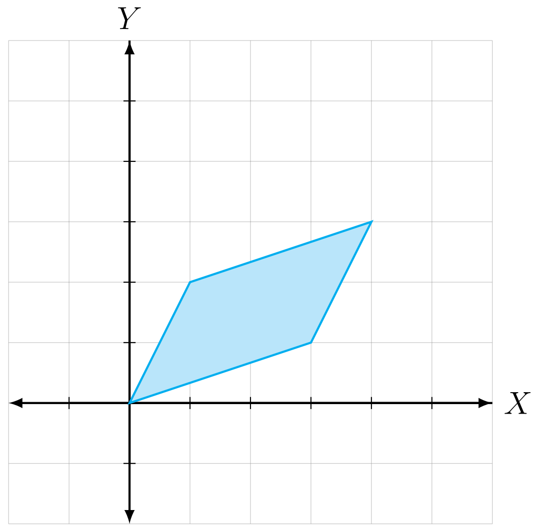

FIGURE 5. A parallelogram

Example 10.10. Let

$$u := \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

Consider the line

$$L := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 1 \right\}.$$

Then the distance from the points $(0, -1)$ and $(2, 2)$ to L are

$$\frac{|(-1) - 1|}{\sqrt{2}} = \sqrt{2} \qquad \text{and} \qquad \frac{|4 - 1|}{\sqrt{2}} = \frac{3}{2}\sqrt{2},$$

respectively.

10.3. **Parallelotopes.** Suppose that $v_1, v_2, \ldots, v_n \in \mathbb{R}^n$ are vectors. The **parallelotope generated by** $v_1, v_2, \ldots, v_n$ is the set

$$\mathcal{P} := \left\{ X \in \mathbb{R}^n \,:\, \overrightarrow{OX} = a_1 \cdot v_1 + a_2 \cdot v_2 + \cdots + a_n \cdot v_n, \ \text{ where } \forall i \in \{1, 2, \ldots, n\},\, a_i \in [0, 1] \right\}.$$

In $\mathbb{R}^2$, if u and v are nonzero nonparallel vectors, then the parallelotope generated by u and v is simply the *parallelogram* with vertices corresponding to the vectors $\mathbf{0}, u, v, u + v$.

The n-dimensional volume of a parallelotope generated by vectors $v_1, v_2, \ldots, v_n$ is given by

$$\text{vol}(\mathcal{P}) = |\det[v_1 \,|\, v_2 \,|\, \cdots \,|\, v_n]|\,.$$

Example 10.11. Let $\mathcal{P} \subseteq \mathbb{R}^2$ be a parallelogram whose vertices are $(0,0)$, $(3,1)$, $(4,3)$, and $(1,2)$. What is the two-dimensional volume of $\mathcal{P}$? (See Figure 5 for an illustration.)

Note that $\mathcal{P}$ is a parallelogram generated by

$$u := \begin{bmatrix} 3 \\ 1 \end{bmatrix} \qquad \text{and} \qquad v := \begin{bmatrix} 1 \\ 2 \end{bmatrix}.$$

Therefore, the volume of $\mathcal{P}$ is

$$\mathrm{vol}(\mathcal{P}) = \left| \det \begin{bmatrix} 3 & 1 \\ 1 & 2 \end{bmatrix} \right| = 5.$$

Note the absolute value function. The convention here is that the *volume* is always a nonnegative quantity.

10.4. **Problems and Solutions.**

Problem 1. Consider the two vectors

$$u := \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix} \qquad \text{and} \qquad v := \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}$$

in the three-dimensional space $\mathbb{R}^3$.

Determine the angle between u and v.

Note. Your answer should be in the form x° ("x degree"), where x is an integer.

Problem 2. Let

$$w := \begin{bmatrix} 8 \\ -4 \\ -1 \end{bmatrix} \in \mathbb{R}^3.$$

Consider the plane

$$H := \left\{ X \in \mathbb{R}^3 \,:\, \left\langle w, \overrightarrow{OX} \right\rangle = 20 \right\},$$

where $O = (0,0,0)$ is the origin in $\mathbb{R}^3$.

(a) Evaluate the distance from the point $(0,2,-1)$ to H.

(b) Evaluate the distance from the point $(4,1,8)$ to H.

Problem 3. Let $\mathcal{P}$ be the parallelepiped generated by

$$u_1 := \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad u_2 := \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}, \qquad \text{and} \qquad u_3 := \begin{bmatrix} 3 \\ 1 \\ 2 \end{bmatrix}.$$

Compute the three-dimensional volume $\operatorname{vol}(\mathcal{P})$ of $\mathcal{P}$.

Problem 4. Let $O = (0,0,0,0,0) \in \mathbb{R}^5$ be the origin. Determine the set of all real numbers t with the property that the two vectors

$$\overrightarrow{OX} \quad \text{and} \quad \overrightarrow{OY}$$

are perpendicular, where X and Y are two points in $\mathbb{R}^5$ given by

$$X := (t,t,t,4,-1) \quad \text{and} \quad Y := (t,-5,t,t,1).$$

Problem 5. In this problem, I would like you to find a physical ruler and make some physical measurements. You might want to read until the end of the problem statement first, so you know what you have to do.

Consider the vector

$$u := \begin{bmatrix} 5 \\ 7 \end{bmatrix} \in \mathbb{R}^2,$$

the line

$$L := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 24 \right\},$$

and the following points

$$A = (-1, 1), \quad B = (6, 1), \quad \text{and} \quad P = (5, 4).$$

A sketch of these objects L, A, B, P is given in Figure 6.

(a) Using the formula from class, compute the distance from P to L. Your answer should be of the form $\frac{a}{\sqrt{b}}$, where a and b are integers.

(b) Give an approximation of your answer in part (a) to 3 significant digits.

(c) Now look at Figure 6. You might want to zoom in or out so that the whole figure fits your screen (or you can print the figure out on a piece of paper). Find a physical ruler. Measure the distance from A to B. Report your measurement, rounded to a nearby integer, in millimeters. Let this integer be x.

(d) Using your ruler, measure the distance from P to L. (Try to find a point on the line L that is approximately closest to P, and measure the distance from that point to P.) Report your measurement, rounded to a nearby integer, in millimeters. Let this integer be y.

(e) Compute

$$d := \frac{y}{x} \cdot 7.$$

Give an approximation of d to 3 significant digits. Is the answer here close to your answer in part (b)? Discuss.

Problem 6. Consider the points $A = (-3, 0)$, $B = (4, 0)$, $C = (4, 3)$, $D = (-3, 3)$, $O = (0, 0)$, $P = (-3, 2)$, $Q = (4, 1)$, and $R = (1, 3)$ in $\mathbb{R}^2$. Let $u := \overrightarrow{OP}$ and $v := \overrightarrow{OQ}$. The parallelogram $\mathcal{P}$ generated by u and v is shown in Figure 7.

(a) Using the formula from class, compute the area of the parallelogram $\mathcal{P}$.

(b) Compute the areas of the following shapes:

(i) $a_1 := \text{Area}(\square ABCD)$,

(ii) $a_2 := \text{Area}(\triangle AOP)$,

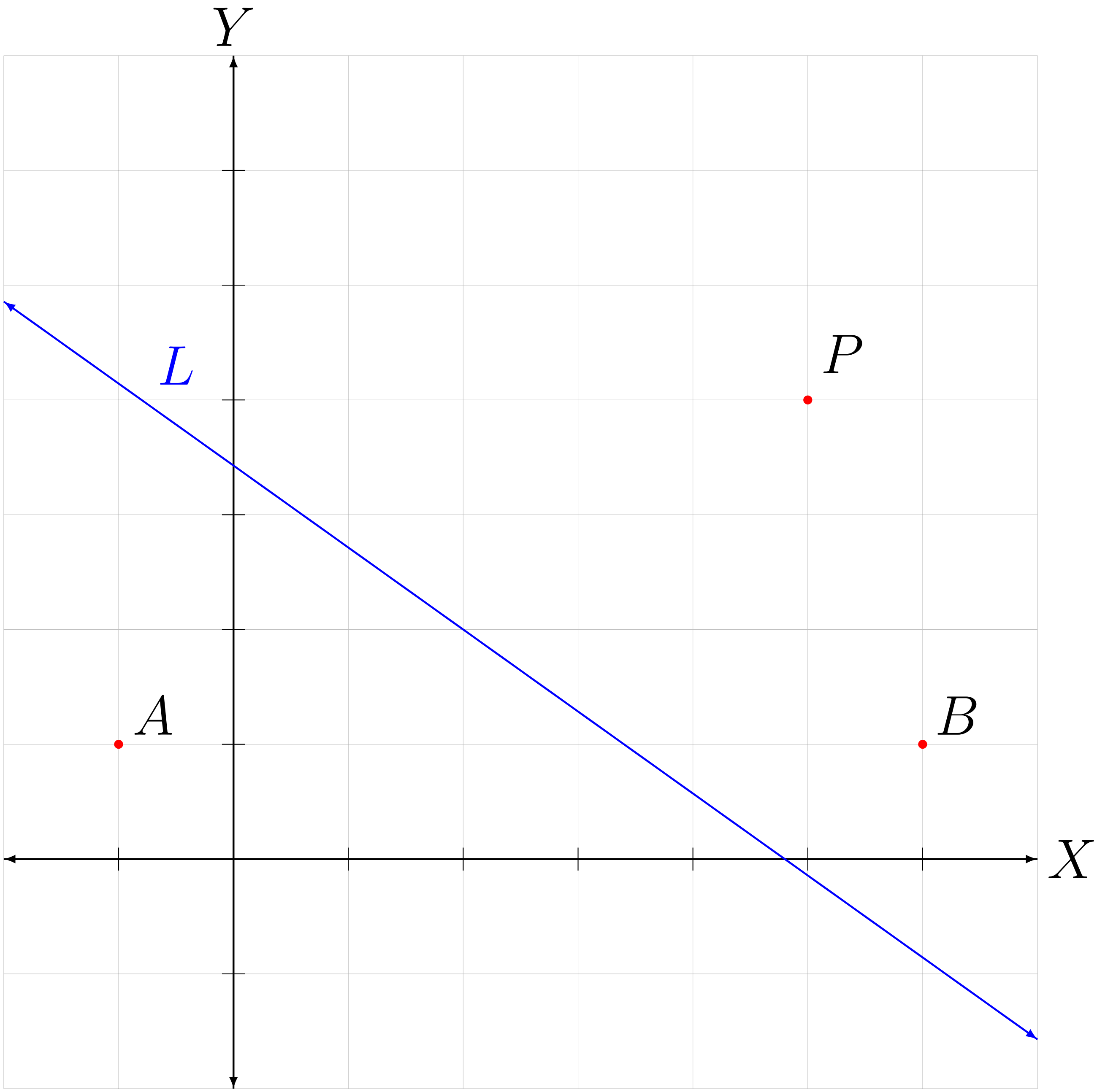

FIGURE 6. Diagram for Problem 5.

(iii) $a_3 := \text{Area}(\triangle OBQ)$,

(iv) $a_4 := \text{Area}(\triangle CQR)$, and

(v) $a_5 := \text{Area}(\triangle DRP)$.

(c) Evaluate $a_1 - a_2 - a_3 - a_4 - a_5$. Is your answer here the same as your answer in part (a)?

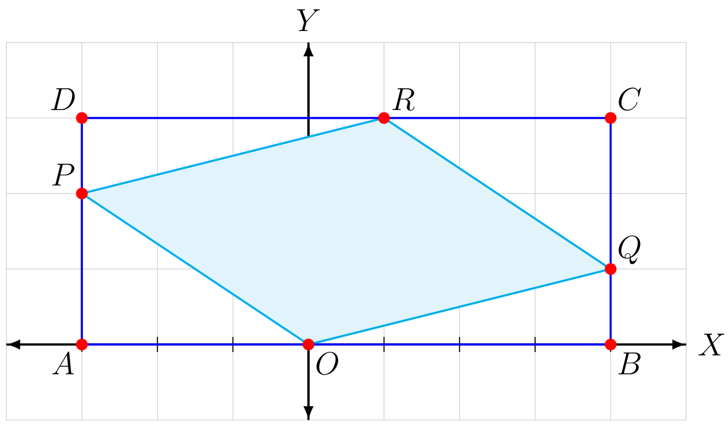

FIGURE 7. Diagram for Problem 6.

Solution to Problem 1. The Euclidean inner product between u and v is

$$\langle u, v\rangle = 1 \cdot (-1) + (-1) \cdot 0 + 0 \cdot 1 = -1.$$

The norms of u and v are $\|u\| = \|v\| = \sqrt{2}$. Therefore, the angle between u and v is

$$\theta = \arccos\left(\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}\right) = \arccos\left(-\frac{1}{2}\right) = 120°.$$

$\square$

Solution to Problem 2. (a) Using the formula given in class, we find that the distance is

$$\frac{|8 \cdot 0 + (-4) \cdot 2 + (-1) \cdot (-1) - 20|}{\sqrt{8^2 + (-4)^2 + (-1)^2}} = 3.$$

(b) Using the formula given in class, we find that the distance is

$$\frac{|8 \cdot 4 + (-4) \cdot 1 + (-1) \cdot 8 - 20|}{\sqrt{8^2 + (-4)^2 + (-1)^2}} = 0.$$

$\square$

Solution to Problem 3. The volume is

$$\mathrm{vol}(\mathcal{P}) = \left|\det \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \\ 3 & 1 & 2 \end{bmatrix}\right| = |6 + 6 + 6 - 27 - 1 - 8| = 18.$$

$\square$

Solution to Problem 4. We have

$$\langle \overrightarrow{OX}, \overrightarrow{OY}\rangle = 2t^2 - t - 1 = (t - 1)(2t + 1).$$

Hence, $\langle \overrightarrow{OX}, \overrightarrow{OY} \rangle = 0$ if and only if $t \in \{-1/2, 1\}$. Note that both $\overrightarrow{OX}$ and $\overrightarrow{OY}$ are nonzero vectors.

The set of all such real numbers t is $\{-\frac{1}{2}, 1\}$. $\square$

Solution to Problem 5. (a) The norm of u is $\|u\| = \sqrt{5^2 + 7^2} = \sqrt{74}$. The formula from class yields the distance

$$\frac{|5 \cdot 5 + 7 \cdot 4 - 24|}{\sqrt{74}} = \frac{29}{\sqrt{74}}.$$

(b) We have

$$\frac{29}{\sqrt{74}} \approx 3.37$$

(c) The answer to this part might be different depending on how you made your measurement. I got $x = 127$.

(d) The answer to this part might be different depending on how you made your measurement. I got $y = 61$.

(e) We have

$$d = \frac{y}{x} \cdot 7 = \frac{61}{127} \cdot 7 \approx 3.36,$$

which is very close (less than 1% away!) from the answer in part (b).

Indeed, this is as expected. The distance between A and B is 7 unit cells. In the measurement, one unit cell becomes (approximately)

$$\frac{x}{7} \text{ millimeters.}$$

The distance from P to L theoretically is $29/\sqrt{74}$ unit cells, which should be (approximately)

$$\frac{29}{\sqrt{74}} \cdot \frac{x}{7} \text{ millimeters.}$$

Hence,

$$y = \frac{29}{\sqrt{74}} \cdot \frac{x}{7},$$

which means

$$\frac{29}{\sqrt{74}} = \frac{y}{x} \cdot 7 = d.$$

$\square$

Solution to Problem 6. (a) The area is

$$\left| \det \begin{bmatrix} -3 & 4 \\ 2 & 1 \end{bmatrix} \right| = |-3 - 8| = 11.$$

(b) We have $a_1 = 3 \times 7 = 21$, $a_2 = \frac{1}{2} \times 3 \times 2 = 3$, $a_3 = \frac{1}{2} \times 1 \times 4 = 2$, $a_4 = \frac{1}{2} \times 3 \times 2 = 3$, and $a_5 = \frac{1}{2} \times 4 \times 1 = 2$.

(c) We have $a_1 - a_2 - a_3 - a_4 - a_5 = 21 - 3 - 2 - 3 - 2 = 11$, which is the same as the answer in part (a). $\square$

11. The Eleventh Week

11.1. **Lines through the origin in two dimensions.** We recall from the last lecture that a *line* in $\mathbb{R}^2$ is a set of the form

$$L := \left\{ X \in \mathbb{R}^2 \, : \left\langle u, \overrightarrow{OX} \right\rangle = t \right\},$$

for a certain nonzero vector $u \in \mathbb{R}^2$ and a certain real number t. Here, $O = (0,0) \in \mathbb{R}^2$ denotes the origin in two dimensions. The equation above says that L is the set of all points $X \in \mathbb{R}^2$ with the property that the Euclidean inner product between a fixed vector $u \in \mathbb{R}^2$ and the vector $\overrightarrow{OX}$ equals a fixed real number t.

The reader may have been familiar with a seemingly different *definition* of a line in $\mathbb{R}^2$ which says that a *line* is the set of all points $(x, y) \in \mathbb{R}^2$ such that

$$ax + by = c,$$

for certain real numbers a, b, c with $(a, b) \neq (0, 0)$. A moment's thought reveals that the two definitions are indeed the same. When we label our point $X \in \mathbb{R}^2$ by their Cartesian coordinates $X =: (x, y) \in \mathbb{R}^2$ and suppose that the vector u is written as

$$u = \begin{bmatrix} a \\ b \end{bmatrix}$$

and let $t = c$, we then see that the equation

$$\left\langle u, \overrightarrow{OX} \right\rangle = t$$

is the same as

$$\left\langle \begin{bmatrix} a \\ b \end{bmatrix}, \begin{bmatrix} x \\ y \end{bmatrix} \right\rangle = c,$$

which is simply

$$ax + by = c.$$

Now suppose we have a line L in $\mathbb{R}^2$. We say that L is a line **through the origin** if $O = (0, 0) \in L$; i.e., the origin lies on the line L.

Proposition 11.1. *The line $L \subseteq \mathbb{R}^2$ given by*

$$L := \left\{ X \in \mathbb{R}^2 \, : \left\langle u, \overrightarrow{OX} \right\rangle = t \right\},$$

for a certain nonzero vector $u \in \mathbb{R}^2$ and a certain real number t, is a line through the origin if and only if $t = 0$.

Proof. ($\Rightarrow$) Suppose that L is a line through the origin. Then $O \in L$, which implies

$$\left\langle u, \overrightarrow{OO} \right\rangle = t.$$

Note that $\overrightarrow{OO}$ is the zero vector. Therefore, $t = 0$.

($\Leftarrow$) Suppose that $t = 0$. Then

$$\left\langle u, \overrightarrow{OO} \right\rangle = 0 = t,$$

which implies that O is an element of L. Hence, L is a line through the origin. $\square$

We have the following corollary of the previous proposition.

Corollary 11.2. *Let $a, b, c \in \mathbb{R}$ be such that $(a, b) \neq (0, 0)$. Then the following are equivalent.*

- *The line defined by the equation $ax + by = c$ (in the variables x and y) is a line through the origin in $\mathbb{R}^2$.*
- *The equation $ax + by = c$ (in the variables x and y) is a homogeneous linear equation.*

11.2. **Reflections across lines through the origin.** In this subsection, let us fix a line

$$L := \left\{ X \in \mathbb{R}^2 : \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\}$$

through the origin, where $u \in \mathbb{R}^2$ is a fixed nonzero vector.

For any vector $v \in \mathbb{R}^2$ in the same Euclidean space as L, we can consider the *reflection* of v across L. Algebraically, this is a function $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(v) := v - \frac{2}{\|u\|^2} \cdot \langle u, v \rangle \cdot u.$$

Example 11.1. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ denote the reflection across the line

$$L := \left\{ X \in \mathbb{R}^2 : \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\},$$

where

$$u := \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad T\left(\begin{bmatrix} 5 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} -1 \\ 5 \end{bmatrix}, \quad \text{and} \quad T\left(\begin{bmatrix} 7 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 7 \\ 7 \end{bmatrix}.$$

Example 11.2. Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ denote the reflection across the line

$$L := \left\{X \in \mathbb{R}^2 \,:\, \left\langle w, \overrightarrow{OX} \right\rangle = 0\right\},$$

where

$$w := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Then

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \quad T\left(\begin{bmatrix} 5 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \quad \text{and} \quad T\left(\begin{bmatrix} 7 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 7 \\ -7 \end{bmatrix}.$$

For every nonzero vector $u \in \mathbb{R}^2$, the reflection $T : \mathbb{R}^2 \to \mathbb{R}^2$ across the line

$$L := \left\{X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 0\right\}$$

is a linear transformation. This implies that there exists a matrix $M \in \mathbb{R}^{2\times 2}$ such that

$$M \cdot v = T(v)$$

holds for every $v \in \mathbb{R}^2$. This matrix is called the **Householder matrix** of the linear transformation. The linear transformation itself is called a **Householder transformation**.

Example 11.3. Let us compute the Householder matrix of the Householder transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that is the reflection across the line

$$L := \left\{X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 0\right\},$$

where

$$u := \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

From our computation above, we find

$$M = [T(e_1) \,|\, T(e_2)] = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

Example 11.4. Let us compute the Householder matrix of the Householder transformation $T : \mathbb{R}^2 \to \mathbb{R}^2$ that is the reflection across the line

$$L := \left\{X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 0\right\},$$

where

$$u := \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

From our computation above, we find

$$M = [T(e_1) \,|\, T(e_2)] = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

11.3. **Orthogonal projections onto lines through the origin.** In this subsection, we continue to fix a line

$$L := \left\{ X \in \mathbb{R}^2 \, : \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\}$$

through the origin, where $u \in \mathbb{R}^2$ is a fixed nonzero vector.

For any vector $v \in \mathbb{R}^2$, we can consider the *orthogonal projection* of v onto L. This is a function $P : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$P(v) := v - \frac{1}{\|u\|^2} \cdot \langle u, v \rangle \cdot u.$$

Example 11.5. Let $P : \mathbb{R}^2 \to \mathbb{R}^2$ denote the orthogonal projection onto the line

$$L := \left\{ X \in \mathbb{R}^2 \, : \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\},$$

where

$$u := \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

Then

$$P\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \quad P\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} 1/2 \\ 1/2 \end{bmatrix}, \quad P\left(\begin{bmatrix} 5 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \text{and} \quad P\left(\begin{bmatrix} 7 \\ 7 \end{bmatrix}\right) = \begin{bmatrix} 7 \\ 7 \end{bmatrix}.$$

For every line L through the origin in $\mathbb{R}^2$, the orthogonal projection P onto L is a linear transformation. Therefore, there exists a matrix $M \in \mathbb{R}^{2\times 2}$ for which

$$M \cdot v = P(v)$$

holds for every vector $v \in \mathbb{R}^2$.

Example 11.6. Let P be the orthogonal projection from the example above. Then the matrix M such that $M \cdot v = P(v)$ for every $v \in \mathbb{R}^2$ is

$$M = \begin{bmatrix} 1/2 & 1/2 \\ 1/2 & 1/2 \end{bmatrix}.$$

11.4. **Rotations about the origin.** Let θ be a real number. We consider the function $T : \mathbb{R}^2 \to \mathbb{R}^2$ which takes any vector $v \in \mathbb{R}^2$ as input and which outputs the vector obtained from rotating v about the origin by angle θ counterclockwise.

The function T is a linear transformation. By trigonometry, it is not hard to see that

$$T\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \qquad \text{and} \qquad T\left(\begin{bmatrix} 0 \\ 1 \end{bmatrix}\right) = \begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix},$$

and therefore we obtain the matrix

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

with the property that

$$R(\theta) \cdot v = T(v),$$

for every $v \in \mathbb{R}^2$. This is called the **rotation matrix**.

Example 11.7. The rotation matrix of rotating by 90° is

$$R(\pi/2) = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

For example, if we start with the vector

$$v := \begin{bmatrix} 3 \\ 2 \end{bmatrix},$$

and we rotate it 90° counterclockwise, then we obtain

$$R(\pi/2) \cdot v = \begin{bmatrix} -2 \\ 3 \end{bmatrix}.$$

Now suppose that α and β are real numbers. Since the composition of rotating by α and rotating by β is the same as rotating by $\alpha + \beta$, we obtain the following identity

$$R(\alpha) \cdot R(\beta) = R(\alpha + \beta),$$

for every pair α, β of real numbers.

By directly multiplying matrices, we obtain

$$\begin{bmatrix} \cos\alpha\cos\beta - \sin\alpha\sin\beta & -\cos\alpha\sin\beta - \sin\alpha\cos\beta \\ \sin\alpha\cos\beta + \cos\alpha\sin\beta & -\sin\alpha\sin\beta + \cos\alpha\cos\beta \end{bmatrix} = \begin{bmatrix} \cos(\alpha+\beta) & -\sin(\alpha+\beta) \\ \sin(\alpha+\beta) & \cos(\alpha+\beta) \end{bmatrix}.$$

Thus, we have recovered the famous trigonometric formulas:

$$\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta,$$
$$\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta.$$

11.5. **Rotations in three dimensions and Euler angles.** In three dimensions, the rotations about the X-, Y-, and Z- axes are given by

$$R_X(\theta) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

$$R_Y(\theta) := \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix},$$

and

$$R_Z(\theta) := \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

These are rotations about the standard basis vectors e_1, e_2, e_3 by angle θ counterclockwise, using the right-hand rule.

Example 11.8. Let $P = (4, 0, 0) \in \mathbb{R}^3$. If we rotate the vector $\overrightarrow{OP}$ by $135°$ about the Z-axis and then rotate the new vector by $60°$ about the Y-axis, what is the resulting vector?

The final vector is

$$R_Y(\pi/3) \cdot R_Z(3\pi/4) \cdot \overrightarrow{OP} = \begin{bmatrix} 1/2 & 0 & \sqrt{3}/2 \\ 0 & 1 & 0 \\ -\sqrt{3}/2 & 0 & 1/2 \end{bmatrix} \cdot \begin{bmatrix} -1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 1/\sqrt{2} & -1/\sqrt{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} -\sqrt{2} \\ 2\sqrt{2} \\ \sqrt{6} \end{bmatrix}.$$

Every rotation in three dimensions can be obtained by first rotating about the Z-axis, second rotating about the X-axis, and third rotating about the Z-axis again. That is, every rotation is three dimensions can be represented as

$$R_Z(\gamma) \cdot R_X(\beta) \cdot R_Z(\alpha)$$

for certain $\alpha, \beta, \gamma \in \mathbb{R}$. These angles are called the **Euler angles**.

11.6. **Orthogonal matrices.** Let n be a positive integer. A matrix $Q \in \mathbb{R}^{n\times n}$ is said to be an **orthogonal matrix** if

$$Q \cdot Q^{\mathsf{T}} = Q^{\mathsf{T}} \cdot Q = I_n.$$

When $n = 2$, a matrix $M \in \mathbb{R}^{2\times 2}$ is an orthogonal matrix if and only if M is either a Householder matrix or a rotation matrix.

From the defining equation of orthogonal matrices, we find that

$$\det(Q)^2 = \det(Q \cdot Q^{\mathsf{T}}) = \det(I_n) = 1.$$

Since the entries of Q are real, we have that $\det(Q)$ is either -1 or 1.

In two dimensions (when $n = 2$), an orthogonal matrix has determinant 1 if and only if it is a rotation matrix, and it has determinant -1 if and only if it is a Householder matrix.

Proposition 11.3. *Every 2×2 orthogonal matrix with determinant -1 is a symmetric matrix. In other words, every 2×2 Householder matrix is symmetric.*

Proof. Suppose that

$$H := \begin{bmatrix} a & b \\ c & d \end{bmatrix} \in \mathbb{R}^{2\times 2}$$

is a 2×2 orthogonal matrix with $\det(H) = -1$. By orthogonality, we have $H \cdot H^{\mathsf{T}} = I_2$, which implies

$$\begin{bmatrix} a^2 + b^2 & ac + bd \\ ac + bd & c^2 + d^2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Since $a^2 + b^2 = 1$, we know from trigonometry that there exists $u \in \mathbb{R}$ for which

$$a = \cos(u) \qquad \text{and} \qquad b = \sin(u).$$

Similarly, since $c^2 + d^2 = 1$, there exists $v \in \mathbb{R}$ for which

$$c = \cos(v) \qquad \text{and} \qquad d = \sin(v).$$

From $\det(H) = -1$, we find that $ad - bc = -1$, which means

$$\cos(u)\sin(v) - \sin(u)\cos(v) = -1.$$

The left-hand side of the above equation is $\sin(v - u)$, and thus

$$\sin(v - u) = -1,$$

which implies that there exists an integer $k \in \mathbb{Z}$ for which

$$v - u = \frac{3\pi}{2} + 2k\pi.$$

This shows that

$$c = \cos(v) = \cos\left(u + \frac{3\pi}{2} + 2k\pi\right) = \cos\left(u + \frac{3\pi}{2}\right) = \sin(u) = b,$$

whence H is symmetric, as desired. $\square$

Example 11.9. The matrix

$$X := \begin{bmatrix} 4/5 & -3/5 & 0 \\ 9/25 & 12/25 & -4/5 \\ 12/25 & 16/25 & 3/5 \end{bmatrix}$$

is an orthogonal matrix. To see this, we can simply directly multiply $X \cdot X^{\mathsf{T}}$ and observe that the product is the identity I_3.

11.7. **Problems and Solutions.**

Problem 1. Let
$$u := \begin{bmatrix} -2 \\ 1 \end{bmatrix} \in \mathbb{R}^2.$$
Consider the line
$$L := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\}.$$

(a) Compute the coordinates of the reflection of the vector
$$v := \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$
across L.

(b) Let H denote the Householder matrix of the reflection across L. Compute the entries of H explicitly. Verify that H is an orthogonal matrix, and that $\det(H) = -1$.

(c) Compute $H \cdot v$. Is the answer the same as your answer in part (a)?

Problem 2. Let
$$u_1 := \begin{bmatrix} 5 \\ 2 \end{bmatrix} \qquad \text{and} \qquad u_2 := \begin{bmatrix} 2 \\ -5 \end{bmatrix}.$$
Consider the lines
$$L_1 := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle u_1, \overrightarrow{OX} \right\rangle = 0 \right\},$$
and
$$L_2 := \left\{ X \in \mathbb{R}^2 \,:\, \left\langle u_2, \overrightarrow{OX} \right\rangle = 0 \right\}.$$

(a) Determine the projection of
$$e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$
onto L_1.

(b) Determine the projection of
$$e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$
onto L_1.

(c) Determine the matrix $P_1 \in \mathbb{R}^{2\times 2}$ with the property that for every vector $v \in \mathbb{R}^2$, we have that $P_1 \cdot v$ is the projection of v onto L_1.

(d) Determine the matrix $P_2 \in \mathbb{R}^{2\times 2}$ with the property that for every vector $v \in \mathbb{R}^2$, we have that $P_2 \cdot v$ is the projection of v onto L_2.

(e) Compute $P_1 \cdot P_2$.

Problem 3. Suppose we start with the vector

$$w_0 := \begin{bmatrix} 8 \\ 4 \\ -8 \end{bmatrix}$$

in the three-dimensional Euclidean space $\mathbb{R}^3$. Let us perform the following steps.

(i) Rotate w_0 about the Y-axis by $30° = \pi/6$ counterclockwise (using the right-hand rule with the +Y direction). Call the result of this rotation w_1.

(ii) Rotate w_1 about the Z-axis by $90° = \pi/2$ counterclockwise (using the right-hand rule with the +Z direction). Call the result of this rotation w_2.

(iii) Rotate w_2 about the Y-axis by $315° = 7\pi/4$ counterclockwise (using the right-hand rule with the +Y direction). Call the result of this rotation w_3.

(a) Compute w_3 explicitly.

(b) Compute $\|w_0\|$ and $\|w_3\|$. Are the two norms the same?

Problem 4. Let a, b, x, y, z be real numbers such that the matrix

$$Q := \begin{bmatrix} a^2/9 & 2/\sqrt{b} & 5/\sqrt{42} \\ 1/\sqrt{3} & 2/\sqrt{14} & y \\ 1/\sqrt{3} & x & z \end{bmatrix} \in \mathbb{R}^{3\times 3}$$

is an orthogonal matrix with real entries. Evaluate

$$\frac{bz}{y} + 7a^2xy.$$

Hint 1. The answer should be an integer.

Hint 2. Recall that a^2 is a nonnegative real number. (The square of every real number is nonnegative.) Furthermore, since Q is a matrix with real entries, its $(1, 2)$-entry, $2/\sqrt{b}$, must be a real number. This implies that b is a positive real number. Also, recall that the square root of a positive real number is also a positive real number.

Problem 5. Let $\theta \in \mathbb{R}$ be a real number. Consider the matrix

$$A := \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Compute A^2. Is A an orthogonal matrix? Is A a Householder matrix? Is A a rotation matrix?

Solution to Problem 1. (a) The reflection is

$$v - \frac{2}{\|u\|^2} \cdot \langle u, v \rangle \cdot u = \begin{bmatrix} 4 \\ 3 \end{bmatrix} - \frac{2}{5} \cdot ((-2) \cdot 4 + 1 \cdot 3) \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 3 \end{bmatrix} + \begin{bmatrix} -4 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

(b) The reflection of e_1 across L is

$$e_1 - \frac{2}{\|u\|^2} \cdot \langle u, e_1 \rangle \cdot u = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{2}{5} \cdot ((-2) \cdot 1 + 1 \cdot 0) \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -8/5 \\ 4/5 \end{bmatrix} = \begin{bmatrix} -3/5 \\ 4/5 \end{bmatrix}.$$

The reflection of e_2 across L is

$$e_2 - \frac{2}{\|u\|^2} \cdot \langle u, e_2 \rangle \cdot u = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{2}{5} \cdot ((-2) \cdot 0 + 1 \cdot 1) \cdot \begin{bmatrix} -2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 4/5 \\ -2/5 \end{bmatrix} = \begin{bmatrix} 4/5 \\ 3/5 \end{bmatrix}.$$

This yields the Householder matrix

$$H = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix}.$$

Note that

$$H \cdot H^{\mathsf{T}} = H^2 = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix} \cdot \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$H^{\mathsf{T}} \cdot H = H^2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

This shows that

$$H \cdot H^{\mathsf{T}} = H^{\mathsf{T}} \cdot H = I_2,$$

implying that H is an orthogonal matrix.

The determinant of H is

$$\det(H) = (-3/5)(3/5) - (4/5)(4/5) = -\frac{9}{25} - \frac{16}{25} = -1,$$

as desired.

(c) The product is

$$H \cdot v = \begin{bmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{bmatrix} \cdot \begin{bmatrix} 4 \\ 3 \end{bmatrix} = \begin{bmatrix} (-3/5) \cdot 4 + (4/5) \cdot 3 \\ (4/5) \cdot 4 + (3/5) \cdot 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 5 \end{bmatrix}.$$

Yes, the answer is the same as the answer in part (a). $\square$

Solution to Problem 2. (a) The projection of e_1 onto L_1 is

$$e_1 - \frac{1}{\|u_1\|^2} \cdot \langle u_1, e_1 \rangle \cdot u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{29} \cdot (5 \cdot 1 + 2 \cdot 0) \cdot \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -25/29 \\ -10/29 \end{bmatrix} = \begin{bmatrix} 4/29 \\ -10/29 \end{bmatrix}.$$

(b) The projection of e_2 onto L_1 is

$$e_2 - \frac{1}{\|u_1\|^2} \cdot \langle u_1, e_2 \rangle \cdot u_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{1}{29} \cdot (5 \cdot 0 + 2 \cdot 1) \cdot \begin{bmatrix} 5 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -10/29 \\ -4/29 \end{bmatrix} = \begin{bmatrix} -10/29 \\ 25/29 \end{bmatrix}.$$

(c) Combining the results from parts (a) and (b), we find that

$$P_1 = \begin{bmatrix} 4/29 & -10/29 \\ -10/29 & 25/29 \end{bmatrix}.$$

(d) The projection of e_1 onto L_2 is

$$e_1 - \frac{1}{\|u_2\|^2} \cdot \langle u_2, e_1 \rangle \cdot u_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} - \frac{1}{29} \cdot (2 \cdot 1 + (-5) \cdot 0) \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -4/29 \\ 10/29 \end{bmatrix} = \begin{bmatrix} 25/29 \\ 10/29 \end{bmatrix}.$$

The projection of e_2 onto L_2 is

$$e_2 - \frac{1}{\|u_2\|^2} \cdot \langle u_2, e_2 \rangle \cdot u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} - \frac{1}{29} \cdot (2 \cdot 0 + (-5) \cdot 1) \cdot \begin{bmatrix} 2 \\ -5 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} + \begin{bmatrix} 10/29 \\ -25/29 \end{bmatrix} = \begin{bmatrix} 10/29 \\ 4/29 \end{bmatrix}.$$

Therefore,

$$P_2 = \begin{bmatrix} 25/29 & 10/29 \\ 10/29 & 4/29 \end{bmatrix}.$$

(e) We have

$$P_1 \cdot P_2 = \begin{bmatrix} (4/29)(25/29) + (-10/29)(10/29) & (4/29)(10/29) + (-10/29)(4/29) \\ (-10/29)(25/29) + (25/29)(10/29) & (-10/29)(10/29) + (25/29)(4/29) \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Remark. The result in part (e) represents the fact that if we start with any vector v, project it onto L_1, and then project the result onto L_2, then we obtain the zero vector, regardless of what v is. This is not surprising since the two lines L_1 and L_2 are perpendicular. □

Solution to Problem 3. (a) We have

$$w_1 = \begin{bmatrix} \sqrt{3}/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & \sqrt{3}/2 \end{bmatrix} \cdot w_0 = \begin{bmatrix} \sqrt{3}/2 & 0 & 1/2 \\ 0 & 1 & 0 \\ -1/2 & 0 & \sqrt{3}/2 \end{bmatrix} \cdot \begin{bmatrix} 8 \\ 4 \\ -8 \end{bmatrix} = \begin{bmatrix} -4 + 4\sqrt{3} \\ 4 \\ -4 - 4\sqrt{3} \end{bmatrix}.$$

Next,

$$w_2 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot w_1 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} -4+4\sqrt{3} \\ 4 \\ -4-4\sqrt{3} \end{bmatrix} = \begin{bmatrix} -4 \\ -4+4\sqrt{3} \\ -4-4\sqrt{3} \end{bmatrix}.$$

Finally,

$$w_3 = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \cdot w_2 = \begin{bmatrix} 1/\sqrt{2} & 0 & -1/\sqrt{2} \\ 0 & 1 & 0 \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} -4 \\ -4+4\sqrt{3} \\ -4-4\sqrt{3} \end{bmatrix} = \begin{bmatrix} 2\sqrt{6} \\ -4+4\sqrt{3} \\ -4\sqrt{2}-2\sqrt{6} \end{bmatrix}.$$

(b) We have

$$\|w_0\| = \sqrt{8^2+4^2+(-8)^2} = \sqrt{64+16+64} = \sqrt{144} = 12,$$

and

$$\begin{aligned} \|w_3\| &= \sqrt{(2\sqrt{6})^2+(-4+4\sqrt{3})^2+(-4\sqrt{2}-2\sqrt{6})^2} \\ &= \sqrt{24+16-32\sqrt{3}+48+32+32\sqrt{3}+24} \\ &= \sqrt{144} = 12. \end{aligned}$$

The two norms are the same. □

Solution to Problem 4. Note that

$$Q^{\mathrm{T}} = \begin{bmatrix} a^2/9 & 1/\sqrt{3} & 1/\sqrt{3} \\ 2/\sqrt{b} & 2/\sqrt{14} & x \\ 5/\sqrt{42} & y & z \end{bmatrix}.$$

Since Q is an orthogonal matrix, we have

$$Q \cdot Q^{\mathrm{T}} = Q^{\mathrm{T}} \cdot Q = I_3.$$

Note that the $(1,1)$-entry of $Q^{\mathrm{T}} \cdot Q$ is

$$(a^2/9)^2 + (1/\sqrt{3})^2 + (1/\sqrt{3})^2.$$

On the other hand, since $Q^{\mathrm{T}} \cdot Q = I_3$, this entry has to be 1. So we have

$$(a^2/9)^2 + (1/\sqrt{3})^2 + (1/\sqrt{3})^2 = 1.$$

Thus, $(a^2/9)^2 = 1/3$. Since $a^2/9$ is a nonnegative real number, we have that

$$\frac{a^2}{9} = \frac{1}{\sqrt{3}},$$

which implies $a^2 = 3\sqrt{3}$.

Next, by considering the $(1,1)$-entry of $Q \cdot Q^{\mathsf{T}}$, we find

$$(a^2/9)^2 + (2/\sqrt{b})^2 + (5/\sqrt{42})^2 = 1,$$

and so

$$\frac{1}{3} + \frac{4}{b} + \frac{25}{42} = 1.$$

This implies $b = 56$.

Now, we have

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{14} & 5/\sqrt{42} \\ 1/\sqrt{3} & 2/\sqrt{14} & y \\ 1/\sqrt{3} & x & z \end{bmatrix} \qquad \text{and} \qquad Q^{\mathsf{T}} = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{14} & 2/\sqrt{14} & x \\ 5/\sqrt{42} & y & z \end{bmatrix}.$$

Next, consider the $(1,2)$-entry of $Q \cdot Q^{\mathsf{T}}$. We have

$$(1/\sqrt{3})(1/\sqrt{3}) + (1/\sqrt{14})(2/\sqrt{14}) + (5/\sqrt{42}) \cdot y = 0,$$

which gives

$$\frac{1}{3} + \frac{1}{7} + \frac{5}{\sqrt{42}} \cdot y = 0.$$

Solving this yields

$$y = -\frac{4}{\sqrt{42}}.$$

Next, consider the $(2,1)$-entry of $Q^{\mathsf{T}} \cdot Q$. We have

$$(1/\sqrt{14})(1/\sqrt{3}) + (2/\sqrt{14})(1/\sqrt{3}) + x \cdot (1/\sqrt{3}) = 0.$$

This gives

$$x = -\frac{3}{\sqrt{14}}.$$

Now, consider the $(3,1)$-entry of $Q^{\mathsf{T}} \cdot Q$. We have

$$(5/\sqrt{42})(1/\sqrt{3}) + (-4/\sqrt{42})(1/\sqrt{3}) + z \cdot (1/\sqrt{3}) = 0.$$

This gives

$$z = -\frac{1}{\sqrt{42}}.$$

Our work above establishes that

$$Q = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{14} & 5/\sqrt{42} \\ 1/\sqrt{3} & 2/\sqrt{14} & -4/\sqrt{42} \\ 1/\sqrt{3} & -3/\sqrt{14} & -1/\sqrt{42} \end{bmatrix}.$$

Conversely, it is not difficult to check that this matrix is an orthogonal matrix.

The quantity we would like to compute is

$$\frac{bz}{y} + 7a^2xy = 56 \cdot \frac{-1/\sqrt{42}}{-4/\sqrt{42}} + 7 \cdot 3\sqrt{3} \cdot (-3/\sqrt{14})(-4/\sqrt{42}) = 14 + 18 = 32.$$

□

Solution to Problem 5. We have

$$\begin{aligned} A^2 &= \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \cdot \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \\ &= \begin{bmatrix} (\cos\theta)^2 + (\sin\theta)^2 & (\cos\theta)(\sin\theta) + (\sin\theta)(-\cos\theta) \\ (\sin\theta)(\cos\theta) + (-\cos\theta)(\sin\theta) & (\sin\theta)^2 + (-\cos\theta)^2 \end{bmatrix} \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I_2. \end{aligned}$$

Note that since A is symmetric, we have $A^{\mathsf{T}} = A$. Our calculation above then implies that

$$A \cdot A^{\mathsf{T}} = A^{\mathsf{T}} \cdot A = I_2,$$

which shows that A is an orthogonal matrix.

Now, let us compute the determinant of A:

$$\det(A) = (\cos\theta)(-\cos\theta) - (\sin\theta)(\sin\theta) = -1.$$

Since the matrix A is a 2×2-orthogonal matrix with determinant -1, we conclude that A is a Householder matrix. Note that A is *not* a rotation matrix (even though its entries involve trigonometric functions!) □

12. The Twelfth Week

We give a brief summary of this class.

12.1. **Week 1.**

In the first week, we talked about linear equations. A *linear equation* in the n variables $x_1, x_2, \ldots, x_n$ is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b, \tag{$\heartsuit$}$$

where $a_1, a_2, \ldots, a_n, b$ are constants: we might think of them as *given* or *known*. Equation ($\heartsuit$) is said to be *homogeneous*[31] if $b = 0$.

A *system of linear equations* (or a *linear system*) in the variables $x_1, x_2, \ldots, x_n$ is a collection of linear equations in the same (ordered) set of variables $x_1, x_2, \ldots, x_n$. A *solution* to a linear system in the variables $x_1, x_2, \ldots, x_n$ is an n-tuple $(t_1, t_2, \ldots, t_n)$ such that when we plug in $x_1 = t_1, x_2 = t_2, \ldots, x_n = t_n$, all the equations in the linear system hold true simultaneously.

A linear system is said to be *consistent* if it has at least one solution.

Example 12.1. Consider the following system

$$\begin{cases} t = 4, \\ t = 6. \end{cases}$$

What is the solution set of the system?

The solution set is $\varnothing$.

Example 12.2. Consider the following system

$$\begin{cases} t = 2, \\ t = 2, \\ t + 1 = 3. \end{cases}$$

What is the solution set of the system?

The solution set is $\{2\}$.

Example 12.3. Consider the following system

$$\begin{cases} t = 3, \\ 0 = 0. \end{cases}$$

What is the solution set of the system?

The solution set is $\{3\}$.

---

[31] pronounced "homoGEneous"

Example 12.4. Consider the following system

$$\begin{cases} t = 1, \\ 0 = 1. \end{cases}$$

What is the solution set of the system?

The solution set is $\varnothing$.

Example 12.5. Consider the following system

$$\begin{cases} 0 = 0, \\ 0 = 0. \end{cases}$$

What is the solution set of the system?

The solution set is $\mathbb{R}$.

Recall once again that if the problem asks to describe the solution set in this case, the answer should be $\mathbb{R}$, the set of real numbers. It would not be a complete description of the solution set, if one simply says that the solution set is “infinite”.

Example 12.6. Consider the following system

$$\begin{cases} t = 5, \\ 0 = 0, \\ 0 = 1. \end{cases}$$

What is the solution set of the system?

The solution set is $\varnothing$.

12.2. **Week 2.** In the second week, we talked about Gaussian elimination. Suppose we would like to solve a linear system. We write the augmented matrix from the system, and then perform Gaussian elimination on the augmented matrix to obtain a matrix in reduced row echelon form. Let us call this final matrix M.

To *solve* the linear system, we consider the matrix M.

Step 1. If M has a leading 1 in the rightmost column (or, equivalently, M has a row in which all entries except the last one are zero and the last entry is 1), then

$$\text{the solution set is } \varnothing.$$

In this case, we finish solving the linear system.

Step 2. If M does not have a leading 1 in the rightmost column, we introduce a new independent variable to each column that is (i) not the rightmost column and (ii) does not have a leading 1.

Replace each variable in each such column with the new variable. Write the linear system associated with M and move all the terms with the new variables to the right-hand side to obtain a parametric description of the solution set in terms of constants and the new variables.

Example 12.7. Suppose we have a linear system in the variables $x_1, x_2, \ldots, x_6$. After performing Gaussian elimination on the linear system, we obtain

$$M = \begin{array}{c} \begin{array}{cccccc} x_1 & x_2 & x_3 & x_4 & x_5 & x_6 \end{array} \\ \begin{bmatrix} 1 & 2 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & -2 \\ 0 & 0 & 0 & 1 & -1 & 0 & -3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array}.$$

There is no leading 1 in the rightmost column. (The top right 1 in the matrix is *not* a leading 1.)

There are three columns that (i) are not the rightmost column and (ii) do not have leading 1's: the columns corresponding to x_2, x_5, x_6. We introduce new variables t_2, t_5, t_6 for these columns.

Plug in $x_2 = t_2$, $x_5 = t_5$, $x_6 = t_6$. We find

$$\begin{cases} x_1 + 2t_2 + 3t_6 = 1, \\ x_3 = -2, \\ x_4 - t_5 = -3. \end{cases}$$

Move the new variables to the right-hand sides to obtain

$$\begin{cases} x_1 = -2t_2 - 3t_6 + 1, \\ x_3 = -2, \\ x_4 = t_5 - 3. \end{cases}$$

Therefore, the solution set of the linear system can be described parametrically as

$$\left\{(-2t_2 - 3t_6 + 1,\ t_2,\ -2,\ t_5 - 3,\ t_5,\ t_6) \,|\, t_2, t_5, t_6 \in \mathbb{R}\right\}.$$

12.3. **Week 3.** In the third week, we talked about matrices. We discussed how we can add, subtract, and multiply matrices.

If $A \in \mathbb{R}^{m \times n}$, then the *transpose* $A^{\intercal}$ of A is the $n \times m$-matrix $A^{\intercal} \in \mathbb{R}^{n \times m}$ whose (i,j)-entry is the (j,i)-entry of A. For example, if

$$A = \begin{bmatrix} 0 & 1 & -1 \\ 5 & -7 & 0 \end{bmatrix},$$

then

$$A^{\mathsf{T}} = \begin{bmatrix} 0 & 5 \\ 1 & -7 \\ -1 & 0 \end{bmatrix}.$$

If $B \in \mathbb{R}^{m\times m}$ is a square matrix, then the *trace* $\operatorname{tr}(B)$ of B is the sum of diagonal entries of B. For example, if

$$B = \begin{bmatrix} 0 & 1 & -2 \\ 3 & -4 & 5 \\ -6 & 7 & 8 \end{bmatrix},$$

then

$$\operatorname{tr}(B) = 0 + (-4) + 8 = 4.$$

Because the diagonal entries do not change after taking transpose, we find that for any square matrix C, we have

$$\operatorname{tr}(C) = \operatorname{tr}(C^{\mathsf{T}}).$$

12.4. **Week 4.** In the fourth week, we talked about some algebraic properties of matrices and inverses.

Two matrices A and B are said to *commute* if $A \cdot B = B \cdot A$. If A and B commute, they must be square matrices of the same dimension.

Suppose we have a 2×2-matrix

$$D = \begin{bmatrix} a & b \\ c & d \end{bmatrix},$$

where a, b, c, d are real numbers. The matrix D is *invertible* if and only if $ad - bc \neq 0$. When D is invertible, its inverse is

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

12.5. **Week 5.** In the fifth week, we talked about elementary matrices and how we use them to compute inverses and determinants.

Example 12.8. Suppose that

$$G = \begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ 2 & -3 & 1 \end{bmatrix}.$$

Let us perform Gaussian elimination on G:

$$G = \begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ 2 & -3 & 1 \end{bmatrix} \xmapsto[\det(E_1)=1]{R_3 \mapsto R_3 + R_2} \begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 0 & 1 & 0 \\ -2 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_2)=1]{R_2 \mapsto R_2 - 3R_1} \begin{bmatrix} 0 & 1 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} 0 & 1 & 0 \\ -2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_3)=-1]{R_1 \rightleftharpoons R_2} \begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

$$\begin{bmatrix} -2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xmapsto[\det(E_4)=-\frac{1}{2}]{R_1 \mapsto (-1/2)\cdot R_1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = I_3.$$

After a sequence of row operations on G, we transform G into I_3. This shows that G is invertible, and thus $\det(G) \neq 0$. We also have elementary matrices E_1, E_2, E_3, E_4 such that

$$(\spadesuit) \qquad E_4 \cdot E_3 \cdot E_2 \cdot E_1 \cdot G = I_3.$$

This shows that

$$\frac{1}{\det(G)} = \det(E_1) \cdot \det(E_2) \cdot \det(E_3) \cdot \det(E_4) = 1 \cdot 1 \cdot (-1) \cdot (-1/2) = 1/2.$$

Hence, $\det(G) = 2$.

Recall that we can write out the elementary matrices E_1, E_2, E_3, E_4 explicitly by doing the same operations to the identity matrix. For example, E_1 is the result of adding R_2 to R_3 ($R_3 \mapsto R_3 + R_2$) on I_3:

$$E_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{bmatrix}.$$

Similarly, we can write out E_2, E_3, and E_4.

From Equation ($\spadesuit$), we find that the inverse of G is

$$G^{-1} = E_4 \cdot E_3 \cdot E_2 \cdot E_1,$$

and we can also express G as

$$G = E_1^{-1} \cdot E_2^{-1} \cdot E_3^{-1} \cdot E_4^{-1},$$

where the right-hand side is a product of elementary matrices.

12.6. **Week 6.** In the sixth week, we introduced a method of finding the inverse of a matrix (if it exists) based on the idea from Week 5.

Suppose we start with a square matrix $H \in \mathbb{R}^{n\times n}$. We form the $n \times 2n$-matrix $[H|I_n]$, and perform Gaussian elimination on it to obtain

$$[K|L] \in \mathbb{R}^{n\times 2n}.$$

If K has an all-zero row, then H is non-invertible and $\det(H) = 0$.

On the other hand, if K is the identity I_n, then we have that H is invertible, and

$$H^{-1} = L.$$

We also talked about the *rule of Sarrus* to compute the determinant of a 3×3-matrix.

12.7. **Week 7.** In the seventh week, we talked about linear transformations.

A function $T : \mathbb{R}^n \to \mathbb{R}^m$ is a *linear transformation* if it satisfies the following two properties:

(i) for every $u, v \in \mathbb{R}^n$, we have $T(u + v) = T(u) + T(v)$, and

(ii) for every $w \in \mathbb{R}^n$ and for every $a \in \mathbb{R}$, we have $T(a \cdot w) = a \cdot T(w)$.

If $T : \mathbb{R}^n \to \mathbb{R}^m$ is a linear transformation, then the matrix

$$M := [\,T(e_1)\,|\,T(e_2)\,|\,\cdots\,|\,T(e_n)\,] \in \mathbb{R}^{m\times n}$$

is the matrix associated to the linear transformation T: it has the property that for every vector $v \in \mathbb{R}^n$,

$$M \cdot v = T(v),$$

where in the equation above, on the left-hand side, the matrix M is multiplied to the left of v (as matrices), while on the right-hand side, the function T is applied to v.

12.8. **Week 8.** In the eighth week, we talked about the Laplace expansion, minors, cofactors, and adjugates.

Recall that the (i, j)-*minor* m_{ij} of a square matrix N is the determinant of the matrix obtained by removing the i^{th} row and the j^{th} column of N. The (i, j)-*cofactor* is $c_{ij} := (-1)^{i+j} \cdot m_{ij}$. The *cofactor matrix* $\operatorname{Cof}(N)$ of N is the matrix of cofactors: its (i, j)-entry is the (i, j)-cofactor of N. Finally, the *adjugate* $\operatorname{adj}(N)$ of N is the transpose of the cofactor matrix of N.

The adjugate has the following important property. For every square matrix $N \in \mathbb{R}^{n\times n}$, we have

$$\operatorname{adj}(N) \cdot N = N \cdot \operatorname{adj}(N) = \det(N) \cdot I_n.$$

The cofactors can be used to compute the determinant of a square matrix by Laplace expansion.

12.9. **Week 9.** In the ninth week, we talked about column replacements, Cramer's rule, and Euclidean geometry.

We introduced the notation $A \xleftarrow{\mathrm{col}_i} b$, which denotes the matrix obtained by replacing the i^{th} column of A by the column vector b.

If P is a square matrix, then the function

$$u \mapsto \det\left(P \xleftarrow{\mathrm{col}_i} u\right)$$

is a linear transformation.

Cramer's rule is a method to solve a linear system, when the coefficient matrix of the linear system is invertible.

Example 12.9. Consider the system

$$\begin{cases} 3x - 2y = 4, \\ x + 3y = 5, \end{cases}$$

of two linear equations in the two variables x and y.

The coefficient matrix is

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

Note that $\det(C) = 3 \cdot 3 - (-2) \cdot 1 = 11$, so C is invertible. Cramer's rule then says that the system has a unique solution with

$$x = \frac{1}{\det(C)} \det \begin{bmatrix} 4 & -2 \\ 5 & 3 \end{bmatrix} = \frac{1}{11} \cdot (12 - (-10)) = 2,$$

and

$$y = \frac{1}{\det(C)} \det \begin{bmatrix} 3 & 4 \\ 1 & 5 \end{bmatrix} = \frac{1}{11} \cdot (15 - 4) = 1.$$

Therefore, the solution set is $\{(2, 1)\}$.

In Euclidean geometry, we might think of elements of $\mathbb{R}^n$ as *points* or *vectors*. For example, in $\mathbb{R}^3$, we have points like

$$P = (0, 1, 2) \qquad \text{and} \qquad Q = (2, 5, 7).$$

When we have two points in the Euclidean space $\mathbb{R}^n$, we can define the vector from P to Q, denoted $\overrightarrow{PQ}$. For example, for the two points above,

$$\overrightarrow{PQ} = \begin{bmatrix} 2 \\ 4 \\ 5 \end{bmatrix} \in \mathbb{R}^3.$$

In $\mathbb{R}^n$, the point O denotes the *origin*:

$$O = (0, 0, \ldots, 0) \in \mathbb{R}^n.$$

Thus, if $R = (r_1, r_2, \ldots, r_n) \in \mathbb{R}^n$ is a point in $\mathbb{R}^n$, then

$$\overrightarrow{OR} = \begin{bmatrix} r_1 \\ r_2 \\ \vdots \\ r_n \end{bmatrix} \in \mathbb{R}^n.$$

12.10. **Week 10.** In the tenth week, we continued to discuss geometry.

We defined the *Euclidean inner product* between two vectors in $\mathbb{R}^n$.

When u and v are two nonzero vectors in $\mathbb{R}^n$, the *angle* between u and v is given by

$$\arccos\left(\frac{\langle u, v\rangle}{\|u\| \cdot \|v\|}\right).$$

Recall the convention that the output of the function arccos is always in the interval $[0, \pi]$; that is, it is a non-negative angle between 0° and 180°, inclusive.

If u and v are nonzero vectors in $\mathbb{R}^n$ and $\langle u, v\rangle = 0$, then we say that u and v are *perpendicular*.

A hyperplane in $\mathbb{R}^n$ is a set of the form

$$H = \left\{X \in \mathbb{R}^n \mid \left\langle u, \overrightarrow{OX}\right\rangle = t\right\},$$

for a fixed nonzero vector $u \in \mathbb{R}^n$ and a fixed real number $t \in \mathbb{R}$. Here, O denotes the origin in $\mathbb{R}^n$.

The distance from a point $P \in \mathbb{R}^n$ to the hyperplane H given above can be computed by the formula

$$\frac{\left|\left\langle u, \overrightarrow{OP}\right\rangle - t\right|}{\|u\|}.$$

The volume of a parallelotope generated by a set of vectors $u_1, u_2, \ldots, u_n \in \mathbb{R}^n$ is given by the absolute value of the determinant

$$\det[\, u_1 \mid u_2 \mid \cdots \mid u_n \,].$$

12.11. **Week 11.** In the eleventh week, we talked about linear transformations in Euclidean spaces.

Some linear transformations we discussed include:

- Householder transformations in two dimensions,
- orthogonal projections onto lines through the origin in two dimensions,

- rotations in two dimensions, and
- rotations in three dimensions.

Suppose that

$$L := \left\{ X \in \mathbb{R}^2 \, : \left\langle u, \overrightarrow{OX} \right\rangle = 0 \right\},$$

where $u \in \mathbb{R}^2$ is a fixed nonzero vector.

Then the Householder transformation of reflecting across L is $T : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$T(v) := v - \frac{2}{\|u\|^2} \cdot \langle u, v \rangle \cdot u.$$

The orthogonal projection onto L is $P : \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$P(v) := v - \frac{1}{\|u\|^2} \cdot \langle u, v \rangle \cdot u.$$

The rotation by angle $\theta \in \mathbb{R}$ counterclockwise is a linear transformation that corresponds to the matrix

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

Every rotation in three dimensions can be obtained by composing rotations in three dimensions about the X-, the Y-, and the Z- axes:

$$R_X(\theta) := \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix},$$

$$R_Y(\theta) := \begin{bmatrix} \cos\theta & 0 & \sin\theta \\ 0 & 1 & 0 \\ -\sin\theta & 0 & \cos\theta \end{bmatrix},$$

and

$$R_Z(\theta) := \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In particular,

$$\{R_Z(\gamma) \cdot R_X(\beta) \cdot R_Z(\alpha) \,|\, \alpha, \beta, \gamma \in \mathbb{R}\}$$

is the set of *all* rotations in three dimensions. These angles α, β, γ used to describe a rotation in three dimensions are called *Euler angles*.

A matrix $Q \in \mathbb{R}^{n \times n}$ is said to be an *orthogonal matrix* if

$$Q \cdot Q^{\mathsf{T}} = Q^{\mathsf{T}} \cdot Q = I_n.$$

References

[AR14] Howard Anton and Chris Rorres. *Elementary linear algebra with supplemental applications, international student version*. Hoboken, NJ: John Wiley & Sons, 11th ed. edition, 2014.