Linear Algebra

Math 314 Fall 2025

Eloísa Grifo

September 21, 2025

Contents

Chapter 1

Systems of Equations

1.1 What is Linear Algebra?

Linear algebra is the study of linear equations.

Definition 1.1. A **linear equation** in the variables $x_1, x_2, \ldots, x_n$ is an equation that can be written in the form:

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, \ldots, a_n, b$ are constants (real numbers). The constant a_i is the **coefficient** of x_i, and b is the **constant term**.

Example 1.2.

a) The equation

$$2x_1 - 5x_2 + 2 = -x_1$$

is a linear equation, as it is equivalent to the equation

$$3x_1 - 5x_2 = -2.$$

b) The equation

$$x_2 = 2(\sqrt{6} - x_1) + x_3$$

is also a linear equation: note that

$$x_2 = 2(\sqrt{6} - x_1) + x_3 \iff 2x_1 + x_2 - x_3 = 2\sqrt{6}.$$

c) The equation

$$x_1x_2 = 6$$

is **not** a linear equation.

d) The equation

$$x_1 + \log x_2 - x_3 = 2$$

is **not** a linear equation.

e) The equation

$$x_1^2 = 7$$

is **not** a linear equation.

In this class, we will study systems of linear equations:

Definition 1.3. A **system of linear equations** or **linear system** is a collection of one or more linear equations. A **solution** to a system of equations in the variables $x_1, \ldots, x_n$ is a list $s = (s_1, \ldots, s_n)$ of numbers that satisfy every equation in the system, meaning that if we replace x_1 by s_1, x_2 by s_2, and so on, then we obtain a true equality.

The **solution set** of a system is the set of all possible solutions.

Example 1.4.

a) The system of linear equations

$$\begin{cases} x_1 = 4 \\ 2x_1 + x_2 = 0 \end{cases}$$

has one solution, the point $(4, -8)$. The solution set is $\{(4, -8)\}$, which is how we denote the set that has only one element $(4, -8)$.

b) The system of linear equations

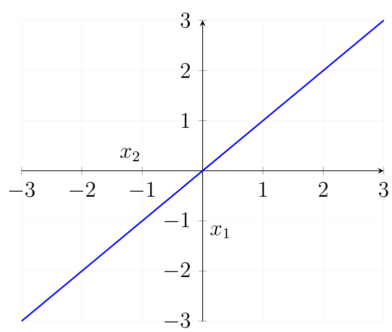
$$\begin{cases} x_1 = 4 \\ x_1 = 7 \end{cases}$$

is impossible, and it has no solutions. The solution set is the **empty set** $\varnothing$.

c) The solution set of the equation

$$x_1 - x_2 = 0$$

is a line:



We will later explain why the following holds:

Important

In general, a system of linear equations may have:

- No solutions,
- Exactly one solution, or
- Infinitely many solutions.

But it can never have a finite number of solutions greater than one.

Definition 1.5. Two systems of linear equations in the same variables $x_1, \ldots, x_n$ are **equivalent** if they have the same solution set.

To study linear systems of equations, we keep replacing our system by an equivalent system, until the solution set becomes easy to find. To do this, we will use matrices.

Definition 1.6. An $m \times n$ (read m *by* n) matrix is a rectangular array of numbers with m rows and n columns. The (i, j) entry of A is the value on the ith row and jth column.

Example 1.7. The following is a 2×3 matrix:

$$\begin{bmatrix} 2 & 5 & 0 \\ 7 & -3 & 13 \end{bmatrix}.$$

Definition 1.8. A system of linear equations

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

has **coefficient matrix** A and **constant vector** $\mathbf{b}$ below:

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{and} \quad \mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}.$$

The **augmented matrix** of the system is

$$\begin{bmatrix} A | \mathbf{b} \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{bmatrix} \quad \text{also written} \quad \left[\begin{array}{ccc:c} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & \ddots & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{array}\right].$$

Sometimes we will write $Ax = \mathbf{b}$ to refer to the system in a more compact way.

Remark 1.9. In the coefficient matrix for a system of linear equations,

$$\text{number of rows} = \text{number of equations}$$
$$\text{number of columns} = \text{number of variables}.$$

In contrast, the augmented matrix always has exactly one extra column.

Example 1.10. Given the system

$$\begin{cases} 3x_1 + x_2 = 5 \\ 2x_1 - x_3 = 6 \end{cases}$$

has coefficient matrix $\begin{bmatrix} 3 & 1 & 0 \\ 2 & 0 & -1 \end{bmatrix}$ and augmented matrix $\begin{bmatrix} 3 & 1 & 0 & 5 \\ 2 & 0 & -1 & 6 \end{bmatrix}$.

How do we solve systems of linear equations?

Theorem 1.11. *Any system of linear equations can be solved using the following* **elementary row operations** *on the augmented matrix:*

1. *Replace: Replace one row by the sum of itself and a multiple of another row.*
2. *Swap: Swap two rows.*
3. *Scale: Multiply all entries of a row by a nonzero constant.*

How does this work in practice?

Example 1.12. Let us take the first step in resolving the following system of linear equations:

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 5x_1 - 5x_3 = 10. \end{cases} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

Replace
$R_3 \to R_3 - 5R_1$

$$\begin{cases} x_1 - 2x_2 + x_3 = 0 \\ 2x_2 - 8x_3 = 8 \\ 10x_2 - 10x_3 = 10. \end{cases} \qquad \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix}$$

Old R_3		5	0	-5	10
$-5R_1$	+	-5	10	-5	0
New R_3		0	10	-10	10

Definition 1.13. We say two $n \times m$ matrices A and B are **row equivalent** if there exists a finite sequence of row operations that converts A into B. We will write $A \sim B$ to say that A and B are equivalent.

Example 1.14. The calculation we did in Example 1.12 shows that

$$\begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix} \sim \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 0 & 10 & -10 & 10 \end{bmatrix}.$$

Remark 1.15. Row operations are always reversible. If matrix A is row equivalent to B, then B is also row equivalent to A. So if we write $A \sim B$, it is also true that $B \sim A$.

Theorem 1.16. *If the augmented matrices of two linear systems are row equivalent, then the systems have the same solution set.*

In other words, if the augmented matrices are row equivalent, then the corresponding linear systems are equivalent. We use this idea to solve systems of linear equations: we keep performing row operations until we have a simpler system we can solve.

Example 1.17. Consider the linear system below, and its augmented matrix:

$$\begin{cases} 2x_2 - 8x_3 = 8 \\ x_1 - 2x_2 = 0 \\ 5x_1 - 5x_3 = 10 \end{cases} \qquad \left[\begin{array}{ccc|c} 0 & 2 & -8 & 8 \\ 1 & -2 & 0 & 0 \\ 5 & 0 & -5 & 10 \end{array}\right].$$

Let us row reduce step by step:

$$\left[\begin{array}{ccc|c} 0 & 2 & -8 & 8 \\ 1 & -2 & 0 & 0 \\ 5 & 0 & -5 & 10 \end{array}\right] \xrightarrow{R_1 \leftrightarrow R_2} \left[\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{array}\right]$$

$$\xrightarrow[R_3 \to \frac{1}{5}R_3]{R_2 \to \frac{1}{2}R_2} \left[\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 1 & -4 & 4 \\ 1 & 0 & -1 & 10 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - R_1} \left[\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 2 & -1 & 10 \end{array}\right]$$

$$\xrightarrow{R_3 \to R_3 - 2R_2} \left[\begin{array}{ccc|c} 1 & -2 & 0 & 0 \\ 0 & 1 & -4 & 4 \\ 0 & 0 & 7 & 2 \end{array}\right].$$

This system is now in **triangular form**, which is sufficient to allow us to find the solutions by **back substitution**: we see that

$$7x_3 = 2,$$

so

$$x_3 = \frac{2}{7}.$$

We can now substitute this back in the second equation to obtain

$$x_2 = 4x_3 + 4 = \frac{8}{7} + 4 = \frac{36}{7}$$

and substituting into the first equation gives us

$$x_1 = 2x_2 = \frac{72}{7}.$$

The solution set is $\{(\frac{72}{7}, \frac{36}{7}, \frac{2}{7})\}$.

The big question is how to apply elementary row operations efficiently. This is where Gauss Elimination will come in.

1.2 Gaussian Elimination and Row Echelon form

Definition 1.18. Given a matrix, the **leading entry** of a particular row is the first nonzero entry in that row (from the left).

Definition 1.19. A rectangular matrix is in **row echelon form** if:

- Any rows consisting entirely of zeros are at the bottom.
- The leading entry of each nonzero row is to the right of the leading entry of the row above.
- All entries below a leading entry (in the same column) are zero.

Example 1.20. In each of the matrices below, we circled the leading entries.

1) The matrix

$$\begin{bmatrix} 2 & -3 & 0 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

is in echelon form.

2) The matrix

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

is not in echelon form. (The leading entries are below each other!)

3) The matrix

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

is not echelon form. (The rows should be switched!)

4) The matrix

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

is not in echelon form. (The second row should be at the bottom!)

Definition 1.21. A matrix is in **reduced row echelon form (RREF)** if it is in row echelon form and:

- The leading entry in each nonzero row is 1.
- Each leading 1 is the only nonzero entry in its column.

Remark 1.22. A typical matrix in RREF has the following format:

$$\begin{bmatrix} \cdots & 0 & 1 & \star & 0 & \star & 0 & \cdots & 0 & \star & \star \\ \cdots & 0 & 0 & 0 & 1 & \star & 0 & \cdots & 0 & \star & \star \\ \cdots & 0 & 0 & 0 & 0 & 0 & 1 & \cdots & 0 & \star & \star \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots \\ 0 & \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 1 & \star & \star \\ \multicolumn{11}{c}{\text{if there are zero rows}} \\ \multicolumn{11}{c}{\text{they are at the bottom}} \end{bmatrix}$$

Example 1.23. In each of the matrices below, we circled the leading entries.

a) The matrix
$$\begin{bmatrix} 2 & -3 & 0 & 1 \\ 0 & 1 & -4 & 8 \\ 0 & 0 & 0 & \frac{1}{3} \\ 0 & 0 & 0 & 0 \end{bmatrix}$$
is in row echelon form, but not in reduced row echelon form.

b) The matrix
$$\begin{bmatrix} 1 & -3 & 0 & 1 \\ 0 & 1 & 4 & 8 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$
is in row echelon form, but not reduced.

c) The matrix
$$\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & -4 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$
is in echelon form, not reduced.

d) The matrix
$$\begin{bmatrix} 1 & 0 & 0 & 29 \\ 0 & 1 & 0 & 36 \\ 0 & 0 & 1 & 7 \end{bmatrix}$$
is in reduced row echelon form.

Definition 1.24. A **pivot position** in a matrix, often shortened to **pivot**, is a position that corresponds to a leading 1 in the reduced echelon form of the matrix. A **pivot column** is a column that contains a pivot position.

Remark 1.25. Important note: a pivot is a *position*, not a value.

Example 1.26. The matrix

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}$$

has RREF

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

We can see the pivots easily from the RREF B:

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

Thus its pivot columns are the first three columns, and we can easily mark the pivots in A:

$$A = \begin{bmatrix} 1 & -2 & 1 & 0 \\ 0 & 2 & -8 & 8 \\ 5 & 0 & -5 & 10 \end{bmatrix}.$$

Theorem 1.27. *Each matrix is row equivalent to one and only one matrix in reduced echelon form.*

However, while the reduced echelon form is unique, note that there are many different paths to the reduced echelon form.

Important

To solve a linear system of equations, we are going to:

1. Write out the augmented matrix corresponding to the system.
2. Get the augmented matrix in row reduced echelon form.

   Remember: there is only <u>one</u> possible row reduced echelon form.
3. Read the solution to the system from the row reduced echelon form.

To get the augmented matrix in row reduced echelon form, we will use an algorithm known as Gauss Elimination, or sometimes also called Gauss–Jordan Elimination.

Algorithm 1.28 (Gaussian Elimination)**.** To get a matrix into reduced row echelon form:

0. Start with the leftmost nonzero column. This will be the first pivot column, with a pivot at the very top.
1. Choose a nonzero entry in this pivot column; swap rows if needed to move it into the top position, and do nothing if the pivot is already in place. From this point on, we will not switch this row with another ever again.
2. Use row operations to eliminate all other entries in this pivot column.
3. Move (right) to the next pivot column and repeat.
4. Scale pivot rows so that each pivot is 1. This can be done together with the previous steps, or all together at the end.
5. Eliminate all entries *above* each pivot.

Example 1.29. Let us solve the following system of linear equations:

$$\begin{cases} x_1 - x_2 + x_3 = 2 \\ 2x_1 - 2x_2 + 3x_3 = 5 \\ -x_1 + x_2 - 2x_3 = -3. \end{cases}$$

First, we write the augmented matrix of the system:

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 2 & -2 & 3 & 5 \\ -1 & 1 & -2 & -3 \end{array}\right].$$

Since the first column is nonzero, that will be our first pivot column, with a pivot on the first row. Luckily, the top entry of the first column is already nonzero, so the first row is not going anywhere.

The next step is to eliminate the rest of the first column using the first row:

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 2 & -2 & 3 & 5 \\ -1 & 1 & -2 & -3 \end{array}\right] \xrightarrow{R_2 \to R_2 - 2R_1} \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ -1 & 1 & -2 & -3 \end{array}\right] \xrightarrow{R_3 \to R_3 + R_1} \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \end{array}\right].$$

Where is the next pivot column? To identify it, we need to now ignore the first row and find the next column with nonzero elements *in another row*. Since the second column has all zeroes outside of the first row, the next pivot column is actually the third column. So now we use the second row to zero out everything below the pivot in the third column.

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & -1 \end{array}\right] \xrightarrow{R_3 \to R_3 + R_2} \left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

This matrix is now in row echelon form, and the pivot positions already have all 1s, but the matrix is not in RREF yet. To achieve that, we need to clear the entries above the pivots too.

$$\left[\begin{array}{ccc:c} 1 & -1 & 1 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \to R_1 - R_2} \left[\begin{array}{ccc:c} 1 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the RREF is

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

In fact, we have circled the pivots below:

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

Now that we know how to apply Gauss Elimination, how will we read the solutions from the row reduced echelon form of the augmented matrix?

Example 1.30. The system

$$\begin{cases} x_1 = 4 \\ 2x_1 + x_2 = 0 \end{cases}$$

has augmented matrix

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

Applying the elementary row operation $R_2 \mapsto R_2 - 2R_1$, we see that its reduced row echelon form is

$$\begin{bmatrix} 1 & 0 & 4 \\ 0 & 1 & -8 \end{bmatrix}.$$

This means that

$$\begin{cases} x_1 = 4 \\ x_2 = -8 \end{cases}$$

and so $(4, -8)$ is only solution.

The key point that made the previous example easy is that every column corresponding to one of the variables $x_1, \ldots, x_n$ has a pivot. But this will not happen in general.

Discussion 1.31 (How to read the solutions from the RREF?)**.** Consider the columns of the augmented matrix corresponding to each of the variables $x_1, \ldots, x_n$, and ignore the last column (corresponding to the constant vector). The columns without pivots give us **free variables**, meaning that these are variables that can take any value. Each choice of values for the free variables will correspond to one solution to the system, because they impose conditions on the variables that are not free. We might call the variables that are not free **leading variables**. We then write an expression for the remaining variables (leading variables) depending on the free variables.

Important

Once we obtain the RREF of a system:

- Columns without pivots among $x_1, \ldots, x_n$ correspond to free variables.
- Free variables can take arbitrary values.
- Each choice of free variables gives one solution to the system.

Example 1.32. We saw in Example 1.29 that the augmented matrix of the system

$$\begin{cases} x_1 - x_2 + x_3 = 2 \\ 2x_1 - 2x_2 + 3x_3 = 5 \\ -x_1 + x_2 - 2x_3 = -3. \end{cases}$$

has RREF

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

The second column has no pivot, so x_2 is a free variable. The columns corresponding to the variables x_1 and x_3 have pivots, so they are not free. This means we can write x_1 and x_3 in terms of the free variable x_2: looking at our system, which has now been reduced to

$$\begin{cases} x_1 - x_2 = 1 \\ x_3 = 1 \\ 0 = 0 \end{cases}$$

we get

$$x_1 = 1 + x_2 \quad \text{and} \quad x_3 = 1.$$

The variable x_2 can take any value, say $x_2 = t$, where t is a parameter that varies. The solutions to the system are all the points of the form

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

where t can take any value. The solution set is

$$\{(1 + t, t, 1) \mid t \text{ any real value}\}.$$

Example 1.33. Suppose the RREF of the augmented matrix of a system is

$$\left[\begin{array}{cccc:c} 1 & 0 & 2 & -1 & 4 \\ 0 & 1 & -3 & 2 & -7 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right].$$

Then x_3 and x_4 are free variables, while x_1 and x_2 are not. To write down all solutions, we need to let the free variables take any values possible. Setting $x_3 = s$ and $x_4 = t$, where s and t are now parameters that will vary over all real numbers, we get

$$\begin{aligned} x_1 &= 4 - 2s + t, \\ x_2 &= -7 + 3s - 2t. \end{aligned}$$

So the solution set is

$$\{(4{-}2s{+}t,\ -7{+}3s{-}2t,\ s,\ t) \mid s, t \in \mathbb{R}\}.$$

Given a system of linear equations, rather than finding the solution set we might just want to know the answers to the following questions:

- Does the system have at least one solution?
- If a solution exists, is it unique? Meaning, does the system have only one solution, or infinitely many?

Definition 1.34. A system of linear equations is:

- **Consistent** if it has at least one solution.
- **Inconsistent** if it has no solutions.

Remark 1.35. A consistent linear system might have one solution or infinitely many solutions.

Theorem 1.36 (Consistency Criterion)**.** *A linear system of equations is inconsistent if and only if the reduced echelon form of its augmented matrix has a pivot in the last column.*

Remark 1.37. A simpler way to say this: a system is inconsistent if the RREF has a row of the form

$$[0\ 0\ \cdots\ 0 \mid 1].$$

Example 1.38. The system with augmented matrix

$$\begin{bmatrix} 1 & 3 & 5 & 7 \\ 0 & 0 & 0 & 42 \end{bmatrix}$$

is inconsistent: we can see that the second row corresponds to the impossible equation $0 = 42$. We can also check the system is inconsistent by seeing that the reduced echelon form is

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

which has a pivot on the last column.

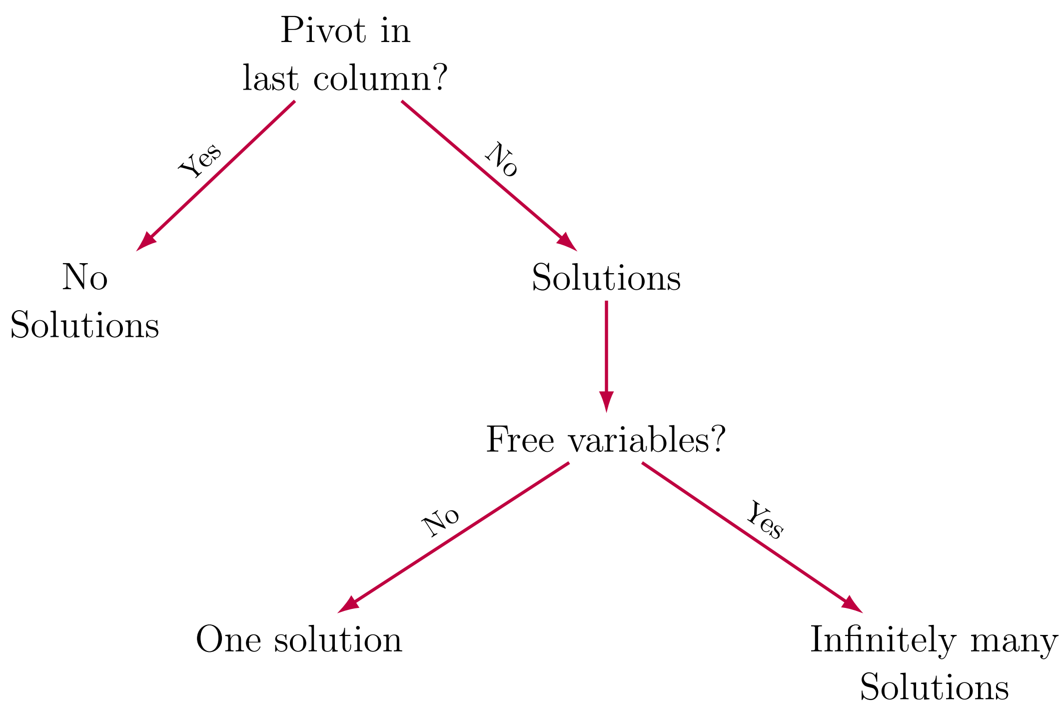
In summary:

Important

To determine how many solutions a system has, look at the reduced row echelon form of the augmented matrix:

- Pivot in the last column $\implies$ inconsistent system, no solutions.
- No pivot in the last column, no free variables $\implies$ exactly one solution.
- No pivot in the last column, some free variables $\implies$ infinitely many solutions.

Pivot in last column	Yes	Yes	No	No
Free variables	Yes	No	No	Yes
Number of solutions	0	0	1	∞



Example 1.39.

a) The system whose augmented matrix has reduced row echelon form

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

has a free variable (x_2) and no pivot in the last column, so it has infinitely many solutions. In fact, the solution set is

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

b) The system whose augmented matrix has reduced row echelon form

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

has a free variable (x_2) and a pivot in the last column, so it no solutions.

c) The system whose augmented matrix has reduced row echelon form

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

has no free variables and no pivot in the last column, so it has exactlty one solution. In fact, the unique solution is $(0, 3)$, so the solution set is $\{(0, 3)\}$.

1.3 The geometry of the solution set of a linear system of equations

Discussion 1.40 (One equation in two variables)**.** The solution set of one linear equation in two variables

$$a_1x_1 + a_2x_2 = b$$

is typically a line, except:

- If $a_1 = a_2 = 0$ and $b \neq 0$, the system is **inconsistent**, as it is equivalent to the equation

  $$0 = b,$$

  which is false. The solution set is the empty set $\varnothing$.

- If $a_1 = a_2 = b = 0$, the system is equivalent to

  $$0 = 0$$

  and the solution set is the entire plane $\mathbb{R}^2$.

Discussion 1.41 (Two equations in two variables)**.** What is the solution set of a system of two linear equations in two variables? Consider the system

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases}$$

where a_{11} and a_{12} are not both zero, and a_{21} and a_{22} are not both zero. Each equation determines a line, so the solution set to this system of equations is the intersection of two lines. This can be:

- A point (one solution),
- A line (infinitely many solutions),
- The empty set (no solution, i.e. if the two lines are parallel).

Example 1.42.

a) The system of equations

$$\begin{cases} x_1 = x_2 \\ x_1 + x_2 = 2 \end{cases}$$

has one solution: the solution set is $\{(1,1)\}$. If we were to represent this geometrically, we only draw one point.

b) The system of equations

$$\begin{cases} x_1 - x_2 = 2 \\ x_1 - x_2 = 0 \end{cases}$$

has no solutions: the solution set is the empty set $\varnothing$. (The two lines corresponding to each equation are parallel!)

c) The system of equations

$$\begin{cases} x_1 = x_2 \\ x_1 + x_2 = 2 \end{cases}$$

has infinitely many solution: the solution set is a whole line. A fancy mathematical way to indicate that line is

$$\{(x_1, x_2) \mid x_1 - 2x_2 = -1\}.$$

Example 1.43 (Planes in three dimensions)**.** A linear equation in three variables, such as

$$x + y + z = 0$$

determines a plane in three-dimensional space. The solution to a system of linear equations in three variables such as

$$\begin{cases} 3x - y + z = 0 \\ 2x + y + 2z = 2 \\ x + 4y - 2z = 11 \end{cases}$$

is the intersection of the three planes corresponding to each equation.

Chapter 2

Vectors

2.1 Introduction to vectors

Definition 2.1. A **vector** is a matrix with only one column, that is, an $n \times 1$ matrix

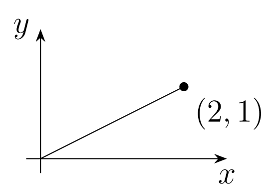
$$v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}.$$

The real number v_1 is the **first component** of v, and v_i is the ith component of v.

We write $\mathbb{R}^n$ for the set of all vectors with n components in the real numbers. The **zero vector** in $\mathbb{R}^n$ is the vector whose entries are all zero:

$$0 = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix} \quad \text{in } \mathbb{R}^n.$$

Discussion 2.2. $\mathbb{R}^2$ is a two-dimensional plane, when we think of the point (a, b) in the plane as corresponding to the vector $\begin{bmatrix} a \\ b \end{bmatrix}$:



$$\rightsquigarrow \quad \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

When we represent our vector with its tail at the origin, we say the vector is in **standard position**. We might also represent a vector with its head at point $A = (a_1, \dots, a_n)$ and its tail at point $B = (b_1, \dots, b_n)$, in which case the vector is

$$v = \begin{bmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{bmatrix}.$$

Definition 2.3. We can sum vectors:

$$\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ \vdots \\ u_n + v_n \end{bmatrix}.$$

We can also multiply vectors by **scalars** (real numbers):

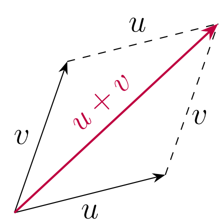
$$c \cdot \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix} = \begin{bmatrix} c\,u_1 \\ \vdots \\ c\,u_n \end{bmatrix}.$$

Example 2.4.

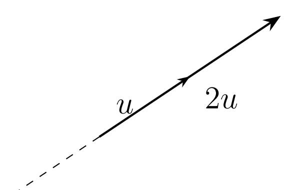
a) $\begin{bmatrix} 2 \\ 3 \end{bmatrix} + \begin{bmatrix} 8 \\ 7 \end{bmatrix} = \begin{bmatrix} 10 \\ 10 \end{bmatrix}$.

b) $2 \cdot \begin{bmatrix} 1 \\ 2 \end{bmatrix} = \begin{bmatrix} 2 \\ 4 \end{bmatrix}$.

Remark 2.5. Here is a geometric visualization of sums (parallelogram rule):



Here is a geometric visualization of scalar multiples:



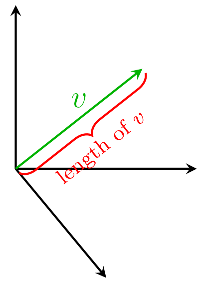
Theorem 2.6 (Properties of vector operations)**.** *For all vectors $u, v, w \in \mathbb{R}^n$ and scalars c, d:*

1. $u + v = v + u$
2. $(u + v) + w = u + (v + w)$
3. $u + 0 = 0 + u = u$
4. $u + (-u) = -u + u = 0$
5. $c(u + v) = cu + cv$
6. $(c + d)u = cu + du$
7. $c(du) = (cd)u$
8. $1u = u$

Definition 2.7. Let v be a vector in $\mathbb{R}^n$. The **length** or **norm** of v is the nonnegative real number

$$\|v\| := \sqrt{v_1^2 + \cdots + v_n^2}.$$

Remark 2.8. If we identify the vector $v = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$ with the point $(v_1, \ldots, v_n)$ in n-dimensional space, the norm of v is the length of the line segment between that point and the origin.



Theorem 2.9. *If v is a vector in $\mathbb{R}^n$ and c is any scalar, $\|cv\| = |c| \cdot \|v\|$.*

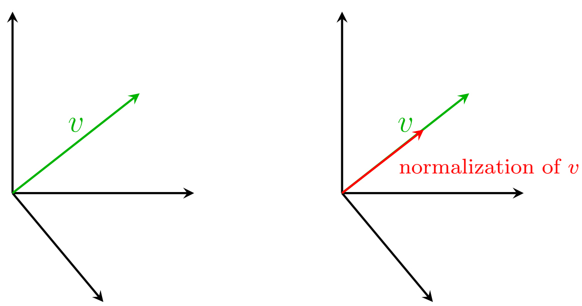
Definition 2.10. A vector in $\mathbb{R}^n$ whose length is 1 is called a **unit** vector.

We can always find a unit vector with the same direction as a given vector v by *normalizing* v:

Definition 2.11. Let $v \neq \mathbf{0}$ be a vector in $\mathbb{R}^n$. The **normalization** of v is the unit vector

$$\frac{v}{\|v\|}$$

which has the same direction as v.



The most important unit vectors are the standard unit vectors:

Definition 2.12. The ith standard basis vector in $\mathbb{R}^n$ is the unit vector

$$\mathbf{e}_i = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \begin{matrix} \\ \\ \\ \longrightarrow \text{position } i \\ \\ \\ \\ \end{matrix}$$

Notation 2.13. In $\mathbb{R}^3$, one sometimes writes $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ for the standard basis vectors:

$$\mathbf{i} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} = \mathbf{e}_1 \qquad \mathbf{j} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} = \mathbf{e}_2 \qquad \mathbf{k} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = \mathbf{e}_3.$$

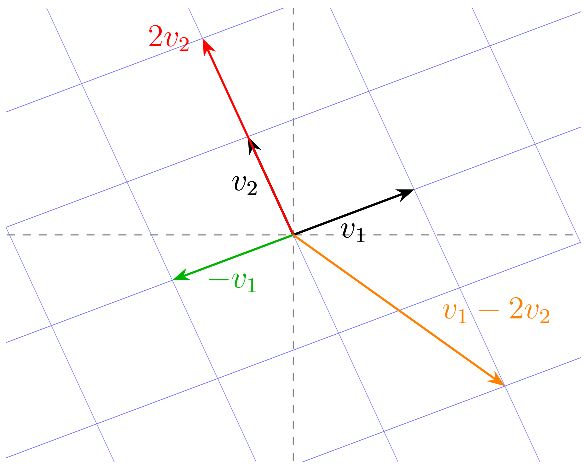
2.2 Linear combinations

Definition 2.14. Given vectors $v_1, \ldots, v_p$ and scalars $c_1, \ldots, c_p$, the vector

$$c_1 v_1 + \cdots + c_p v_p$$

is a **linear combination** of $v_1, \ldots, v_p$ with coefficients $c_1, \ldots, c_p$.

Remark 2.15. What does this look like geometrically? Here is a depiction of the linear combinations of v_1 and v_2:



Any point on the plane determined by v_1 and v_2 is a linear combination of v_1 and v_2.

Remark 2.16. Note that any vector in $\mathbb{R}^n$ can be written as a linear combination of the standard vectors $\mathbf{e}_1, \ldots, \mathbf{e}_n$: the vector $v \in \mathbb{R}^n$ is

$$v = v_1 \mathbf{e}_1 + \cdots + v_n \mathbf{e}_n.$$

A typical question we would like to answer is the following: given vectors $v_1, \ldots, v_p, b \in \mathbb{R}^n$, is b a linear combination of $v_1, \ldots, v_p$?

Discussion 2.17. Given vectors $v_1, \ldots, v_p, b \in \mathbb{R}^n$, b is a linear combination of $v_1, \ldots, v_p$ if and only if the vector equation

$$x_1 v_1 + \cdots + x_p v_p = b$$

has solutions. This vector equation has the same solutions as the linear system with augmented matrix

$$\left[\begin{array}{ccc:c} v_1 & \cdots & v_p & b \end{array}\right].$$

Example 2.18. Is $\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$ a linear combination of $\begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix}$?

Solution: We are asking if the system

$$x_1 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + x_2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$$

has a solution; equivalently, whether the linear system with augmented matrix

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

is consistent (= has solutions). We see that

$$\begin{bmatrix} 1 & 2 & 7 \\ -2 & 5 & 4 \\ -5 & 6 & -3 \end{bmatrix} \xrightarrow[R_3 \to R_3 + 5R_1]{R_2 \to R_2 + 2R_1} \begin{bmatrix} 1 & 2 & 7 \\ 0 & 9 & 18 \\ 0 & 16 & 32 \end{bmatrix} \xrightarrow{R_2 \to \frac{1}{9} R_2} \begin{bmatrix} 1 & 2 & 7 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix}$$

$$\xrightarrow{R_1 \to R_1 - 2R_2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 16 & 32 \end{bmatrix} \xrightarrow{R_3 \to R_3 - 16R_2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & 2 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{(reduced row echelon form).}$$

No pivots in the last column $\Rightarrow$ the system is consistent.

Answer: yes.

In fact, if we wanted to find an explicit way of writing our vector as a linear combination of the other two, all we need is a solution to the system. From the RREF, we see that there is a unique solution (no free variables!), given by $(3, 2)$. Thus

$$3 \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix} + 2 \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} = \begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}.$$

Definition 2.19 (Span)**.** Let $v_1, \ldots, v_p$ be vectors in $\mathbb{R}^n$. The set of all linear combinations of $v_1, \ldots, v_p$ is the **span** of $v_1, \ldots, v_p$, written

$$\operatorname{span}(\{v_1, \ldots, v_p\}) = \{c_1 v_1 + \cdots + c_p v_p \mid c_i \in \mathbb{R}\}.$$

Example 2.20.

a) $\operatorname{span}\left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\}$.

b) $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\} = \left\{ \begin{bmatrix} a \\ 0 \end{bmatrix} : a \text{ any value} \right\}$.

Remark 2.21. Note that for any vector $v \in \mathbb{R}^n$,

$$\operatorname{span}\{v\} = \{\lambda v \mid \lambda \in \mathbb{R}\}$$

is the set of all scalar multiples of v.

Example 2.22. Is $\begin{bmatrix} 7 \\ 4 \\ -3 \end{bmatrix}$ in $\operatorname{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ -5 \end{bmatrix}, \begin{bmatrix} 2 \\ 5 \\ 6 \end{bmatrix} \right\}$? According to Example 2.18, yes.

Example 2.23. Let u, v be vectors in $\mathbb{R}^3$, both nonzero. If u is a scalar multiple of v, then $\operatorname{span}\{u, v\} = \operatorname{span}\{u\}$ is a line. Otherwise, $\operatorname{span}\{u, v\}$ is a plane!

2.3 Matrix Equations

Definition 2.24 (Matrix-vector multiplication)**.** Let A be an $m \times n$ matrix and consider a vector $x \in \mathbb{R}^n$. The product Ax is the vector in $\mathbb{R}^m$ given by

$$\begin{aligned} Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} &= x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} \\ &= \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{bmatrix}. \end{aligned}$$

We might shorten this by setting the columns of A to be $a_1, \ldots, a_n$, so that we can write

$$Ax = x_1a_1 + \cdots + x_na_n.$$

This indicates a linear combination of the columns of A with coefficients $x_1, \ldots, x_n$.

Remark 2.25. For the product Ax of a matrix A with a vector x to be defined, we need the number of columns of A to match the number of rows of the vector x.

Notation 2.26. Given a system of linear equations

$$\begin{cases} a_{11}x_1 + \cdots + a_{1n}x_n = b_1 \\ \qquad\qquad \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n = b_m \end{cases}$$

with coefficient matrix A and constant vector b, we can write our system in matrix notation

$$Ax = b$$

where

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

a vector of *variables.*

Remark 2.27. The matrix equation $Ax = b$ has the *exact same solution set* as the vector equation $x_1a_1 + \cdots + x_na_n = b$ and as the linear system with augmented matrix

$$\begin{bmatrix} A \mid b \end{bmatrix} = \begin{bmatrix} a_1 & \cdots & a_n & | & b \end{bmatrix}.$$

Example 2.28. The linear system

$$\begin{cases} x_1 + 3x_2 = 4 \\ -x_1 + x_2 = 1 \end{cases}$$

has augmented matrix

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

and can be written as a matrix equation as follows:

$$\begin{bmatrix} 1 & 3 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix} \qquad \text{or equivalently} \qquad x_1 \begin{bmatrix} 1 \\ -1 \end{bmatrix} + x_2 \begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \end{bmatrix}.$$

Remark 2.29. The system $Ax = b$ has a solution if and only if b is a linear combination of the columns of A.

Theorem 2.30. *Fix an $m \times n$ matrix A. The following are equivalent:*

a) The system $Ax = b$ has a solution for every *vector $b \in \mathbb{R}^m$.*

b) Every vector $b \in \mathbb{R}^m$ is a linear combination of the columns of A.

c) The columns of A span $\mathbb{R}^m$.

d) The coefficient matrix A has a pivot in every row.

Remark 2.31. The last statement is about A itself, the coefficient matrix of the system, and *not* an augmented matrix.

Example 2.32. The matrix

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

has a pivot in every row. Hence the equation $Ax = b$ has solutions for every $b \in \mathbb{R}^2$. Indeed,

$$\operatorname{span}\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\} = \mathbb{R}^2.$$

Theorem 2.33 (Properties of matrix-vector products)**.** *Let c be a scalar, let $u, v \in \mathbb{R}^n$, and let A be an $m \times n$ matrix. Then*

$$A(u + v) = Au + Av \qquad \textit{and} \qquad A(cu) = c(Au).$$

2.4 Homogeneous linear systems of equations

Definition 2.34. A linear system is **homogeneous** if we can write it as

$$Ax = 0.$$

Remark 2.35. A homogeneous system always has a solution, $x = 0$. This is called the **trivial solution**. A solution $x \neq 0$ is called **nontrivial**.

Remark 2.36. Given a homogeneous system $Ax = 0$, the system has a nontrivial solution if and only if the system has at least one free variable.

Example 2.37. Consider the homogeneous linear system

$$\begin{cases} 3x_1 + 5x_2 - 4x_3 = 0 \\ -3x_1 - 2x_2 + 4x_3 = 0 \\ 6x_1 + x_2 - 8x_3 = 0. \end{cases}$$

This can be written in matrix notation as

$$\underbrace{\begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & -8 \end{bmatrix}}_{A} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}}_{x} = \underbrace{\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}}_{b} \qquad \text{(this is a homogeneous system).}$$

We can solve this system by finding the reduced row echelon form of the augmented matrix $[A\,|b]$. Since $b = 0$, row operations on A and on $[A\,|\,0]$ are the equivalent, and the augmented matrix does not add any new information. So it is sufficient to find the by finding the reduced row echelon form of A.

$$\left[\begin{array}{ccc|c} 3 & 5 & -4 & 0 \\ -3 & -2 & 4 & 0 \\ 6 & 1 & -8 & 0 \end{array}\right] \xrightarrow{R_2 \to R_2 + R_1} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 6 & 1 & -8 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 - 2R_1} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R_2 \to \frac{1}{3}R_2} \left[\begin{array}{ccc|c} 3 & 5 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{array}\right] \xrightarrow{R_1 \to R_1 - 5R_2} \left[\begin{array}{ccc|c} 3 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & -9 & 0 & 0 \end{array}\right] \xrightarrow{R_3 \to R_3 + 9R_2} \left[\begin{array}{ccc|c} 3 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

$$\xrightarrow{R1 \to \frac{1}{3}R1} \left[\begin{array}{ccc|c} 1 & 0 & -\frac{4}{3} & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

From the reduced row echelon form we see that

$$x_3 \text{ is free,} \qquad x_1 = \tfrac{4}{3}x_3, \qquad \text{and } x_2 = 0.$$

Hence the solution set is

$$x = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} \qquad \text{(a one parameter family; nontrivial solutions occur when } x_3 \neq 0\text{).}$$

The general solution to our system is

$$x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} \qquad \text{or} \qquad t \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}$$

where t is a parameter that can take any real value. The trivial solution comes from choosing $t = 0$. Each choice of $t \neq 0$ gives a nontrivial particular solution: for example, taking $t = 1$ gives the solution

$$\begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix}.$$

Remark 2.38. In summary, the general solution to a homogeneous system is a linear combination of vectors, with the free variables as coefficients. The general solution, when written in this format, is said to be in **parametric vector form**.

Definition 2.39. A **nonhomogeneous linear system** is a linear system of the form

$$Ax = b \qquad \text{for some } b \neq 0.$$

Theorem 2.40. *The general solution to the nonhomogeneous system $Ax = b$ is*

$$x = \textit{ one particular solution } + \textit{ general solution to the homogeneous system } Ax = 0.$$

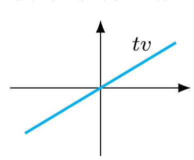
Remark 2.41. Theorem 2.40 says that the solution set of the nonhomogeneous system $Ax = b$ is obtained by translating the solution set for $Ax = 0$ by a vector corresponding to one particular solution to $Ax = b$.

For example, suppose that the general solution to $Ax = 0$ is $x = tv$, where the parameter t can take the value of any real number, and $v \in \mathbb{R}^n$ is any nonzero vector; note that $x = tv$ is a line with direction v. Then the general solution to $Ax = b$ is

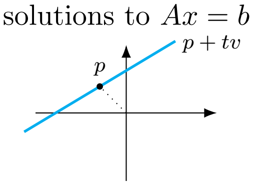
$$x = tv + p \qquad \text{for some vector } p.$$

Geometrically, this corresponds to a line parallel to v, but that goes through the point corresponding to p.

solutions to $Ax = 0$



solutions to $Ax = b$



Remark 2.42. Here are some useful geometric rules: given $u, v \in \mathbb{R}^n$,

- Parametric equation of the line through u parallel to v:

$$x = u + tv, \qquad t \in \mathbb{R}.$$

- Parametric equation of the line through u and v:

$$x = u + t(v - u), \qquad t \in \mathbb{R}.$$

Example 2.43. Consider the system

$$\begin{bmatrix} 3 & 5 & -4 \\ -3 & -2 & 4 \\ 6 & 1 & 8 \end{bmatrix} x = \begin{bmatrix} 7 \\ 1 \\ -4 \end{bmatrix}.$$

The reduced row echelon form of the augmented matrix is

$$\left[\begin{array}{ccc:c} 1 & 0 & -\frac{4}{3} & -1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 0 & 0 \end{array}\right].$$

Thus the general solution to the system is

$$x = x_3 \begin{bmatrix} \frac{4}{3} \\ 0 \\ 1 \end{bmatrix} + \begin{bmatrix} -1 \\ 2 \\ 0 \end{bmatrix}, \qquad \text{which is in parametric vector form.}$$

Important

To write the solution set of a consistent system:

1) Row-reduce the augmented matrix into reduced echelon form.
2) Write each non-free variable in terms of the free ones.
3) Write the general solution x as a vector whose entries depend on the free variables (if there are free variables).
4) Decompose this as a linear combination of vectors where each coefficient is a free variable (plus possibly one term with coefficient 1 for a particular solution).

Example 2.44. Let us find the general solution to the linear system with augmented matrix

$$\left[\begin{array}{ccccc:c} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{array}\right].$$

and write the solution in parametric vector form.

Solution. First, one uses Gauss Elimination to see that the augmented matrix has reduced echelon form

$$\left[\begin{array}{ccccc:c} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{array}\right].$$

From the RREF, we see that the free variables are x_3 and x_4. So the general solution is

$$\begin{cases} x_3, x_4 \ \text{ are free variables} \\ \quad x_1 = 2x_3 - 3x_4 - 24, \\ \quad x_2 = 2x_3 - 2x_4 - 7, \\ \quad x_5 = 4. \end{cases}$$

Let $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{bmatrix}$. Setting $x_3 = x_4 = 0$ gives us the particular solution

$$\mathbf{x} = \begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix}.$$

Now we can write the general solution in parametric vector form:

$$\mathbf{x} = \underbrace{x_3 \begin{bmatrix} 2 \\ 2 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -3 \\ -2 \\ 0 \\ 1 \\ 0 \end{bmatrix}}_{\text{general solution of the homogeneous system}} + \underbrace{\begin{bmatrix} -24 \\ -7 \\ 0 \\ 0 \\ 4 \end{bmatrix}}_{\text{particular solution}}.$$

We close this section with an important warning:

Important

Caution! Given a linear system $Ax = b$, there is a big difference between

the coefficient matrix A and the augmented matrix $\begin{bmatrix} A & b \end{bmatrix}$.

- Is the system $Ax = b$ consistent? $\Longrightarrow$ look at the augmented matrix.
- The system $Ax = 0$ is always consistent.

  We can solve the system by focusing only on A and then finding a particular solution, but if we do so we must remember A is *not* the augmented matrix of the system.

2.5 Linear Independence

Definition 2.45. A set of vectors $\{v_1, \dots, v_p\}$ in $\mathbb{R}^n$ is **linearly independent** if the vector equation

$$x_1 v_1 + \cdots + x_p v_p = 0$$

has only the trivial solution $x_1 = \cdots = x_p = 0$. We say that $v_1, \dots, v_p$ are linearly independent or that the set $\{v_1, \dots, v_p\}$ is linearly independent.

A set of vectors $\{v_1, \dots, v_p\}$ in $\mathbb{R}^n$ is **linearly dependent** if there exist scalars $c_1, \dots, c_p$, not all zero, such that

$$c_1 v_1 + \cdots + c_p v_p = 0.$$

Given such $c_1, \dots, c_p$, the equation

$$c_1 v_1 + \cdots + c_p v_p = 0.$$

is called a **relation of linear dependence** among $v_1, \dots, v_p$. We say that $v_1, \dots, v_p$ are linearly dependent or that the set $\{v_1, \dots, v_p\}$ is linearly dependent.

Remark 2.46. Equivalently, $\{v_1, \dots, v_p\}$ is linearly independent if and only if

$$c_1 v_1 + \cdots + c_p v_p = 0 \quad \Longrightarrow \quad c_1 = \cdots = c_p = 0.$$

Remark 2.47. The singleton set $\{v\}$ is linearly independent $\iff$ $v \neq 0$.

Example 2.48. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ are linearly dependent: for example, we can take

$$0 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

More generally:

Theorem 2.49. *Any set of vectors in $\mathbb{R}^n$ that contains the zero vector is linearly dependent.*

Why? Because we can always take any nonzero coefficient for the zero vector and 0 for the coefficients of all the (nonzero) vectors.

Example 2.50. $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ and $\begin{bmatrix} 2 \\ 0 \end{bmatrix}$ are linearly dependent, since one is a scalar multiple of the other. Indeed,

$$2 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + (-1) \begin{bmatrix} 2 \\ 0 \end{bmatrix} = 0$$

is a relation of linear dependence.

In fact, more generally, any two nonzero vectors that are scalar multiples of each other form a linearly dependent set.

Example 2.51. If v is any nonzero vector and $t \neq 1$, then v and tv are linearly dependent, since

$$t \cdot v + (-1) \cdot (tv) = 0,$$

and the coefficients t and -1 are not both zero. Thus any two nonzero vectors that are scalar multiples of each other form a linearly dependent set.

Remark 2.52. A set $\{v_1, \ldots, v_p\}$ of two or more vectors is linearly dependent if and only if one of the vectors is a linear combination of the others. However, note that this does *not* say that *every* v_i is a linear combination of the rest.

Example 2.53. $\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 2 \end{bmatrix} \right\}$ is linearly dependent, since

$$\begin{bmatrix} 2 \\ 2 \end{bmatrix} = 2 \begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

However, note that $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is not a linear combination of the other two vectors.

Remark 2.54. An equation of linear dependence among the vectors $v_1, \ldots, v_n$ is a nontrivial solution to the homogeneous system

$$x_1 v_1 + \cdots + x_n v_n = 0.$$

Thus to decide if the vectors $v_1, \ldots, v_n$ are linearly independent, we consider the matrix

$$A = \begin{bmatrix} | & & | \\ v_1 & \cdots & v_n \\ | & & | \end{bmatrix}$$

whose columns are the vectors $v_1, \ldots, v_n$, and ask whether the system $Ax = 0$ has a nontrivial solution. The vertical lines above are just for visual effect, as a reminder that each v_i is a vector; the correct way to write this is

$$A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}.$$

In summary:

Theorem 2.55. *The columns of a matrix A are linearly independent if and only if the homogeneous system $Ax = 0$ has only the trivial solution.*

The homogeneous system $Ax = 0$ has only the trivial solution if and only if the augmented matrix $[A \mid 0]$ of the homogeneous system $Ax = 0$ has no free variables. Thus we can check whether a set of vectors is linearly independent by looking at the RREF of a matrix with those vectors as columns:

Theorem 2.56. *Consider the matrix*

$$A = \begin{bmatrix} v_1 & \cdots & v_n \end{bmatrix}.$$

The column vectors $v_1, \ldots, v_n$ are linearly independent if and only if A has a pivot in every column.

Example 2.57. Consider the vectors

$$v_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \qquad v_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}, \qquad v_3 = \begin{bmatrix}2\\1\\0\end{bmatrix}.$$

Question. Are v_1, v_2, and v_3 linearly independent? linearly independent?
Solution. Consider the homogeneous system whose coefficient matrix has columns v_1, v_2, and v_3:

$$\left[\begin{array}{ccc:c} 1 & 4 & 2 & 0\\ 2 & 5 & 1 & 0\\ 3 & 6 & 0 & 0\end{array}\right] \sim \left[\begin{array}{ccc:c} 1 & 0 & -2 & 0\\ 0 & 1 & 1 & 0\\ 0 & 0 & 0 & 0\end{array}\right].$$

There is one free variable, so the system has nontrivial solutions. Hence the vectors are linearly dependent.

Question. How do we find a linear dependence relation among v_1, v_2, and v_3?
Solution. From the reduced row echelon form, we see that

$$\text{general solution: } \begin{cases} x_3 \text{ free} \\ x_1 = 2x_3 \\ x_2 = -x_3 \end{cases} \qquad \text{In parametric vector form: } \; x = t\begin{bmatrix}2\\-1\\1\end{bmatrix}.$$

To get a particular solution, we can take for example $t = 1$, giving us

$$x = \begin{bmatrix}2\\-1\\1\end{bmatrix}.$$

Remember: these are the coefficients in our relation of linear dependence. This gives us the following relation of linear dependence:

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

Theorem 2.58. *Any set of more than n vectors in $\mathbb{R}^n$ is linearly dependent.*

Note that if we have more than n vectors in $\mathbb{R}^n$, it is not possible for the matrix with those vectors as columns to have a pivot in every column.

Example 2.59. Using Theorem 2.58, we can see immediately that

$$v_1 = \begin{bmatrix}1\\2\\3\end{bmatrix}, \qquad v_2 = \begin{bmatrix}4\\5\\6\end{bmatrix}, \qquad v_3 = \begin{bmatrix}7\\8\\9\end{bmatrix}, \qquad \text{and} \qquad v_3 = \begin{bmatrix}10\\11\\12\end{bmatrix}$$

are linearly dependent, since we have 4 vectors in $\mathbb{R}^3$.

2.6 Matrix Transformations

Definition 2.60 (Matrix transformation)**.** Any $m \times n$ matrix A determines a function $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ as follows: for each vector $x \in \mathbb{R}^n$,

$$\mathrm{T}(x) = Ax.$$

Such a function is called a **matrix transformation**.

Remark 2.61. Helpful visual aid: the matrix A gives a function $\mathrm{T}\colon \mathbb{R}^{\#\text{ columns}} \longrightarrow \mathbb{R}^{\#\text{ rows}}$.

Example 2.62. The matrix

$$A = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix}$$

determines the function $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^3$ given by $\mathrm{T}(x) = Ax$. For example,

$$\mathrm{T}\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 3 \\ 9 \\ -3 \end{bmatrix}.$$

and

$$\mathrm{T}\left(\begin{bmatrix} 2 \\ -1 \end{bmatrix}\right) = \begin{bmatrix} 1 & -3 \\ 3 & 5 \\ -1 & 7 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 2+3 \\ 6-5 \\ -2-7 \end{bmatrix} = \begin{bmatrix} 5 \\ 1 \\ -9 \end{bmatrix}.$$

Example 2.63. Consider the transformation $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $\mathrm{T}(x) = Ax$, where

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

Note that for all values of x_1 and x_2,

$$\mathrm{T}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This is the *identity map*! And in fact, the matrix $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ is called the *identity matrix*.

Notation 2.64 (Identity matrix)**.** The $n \times n$ matrix

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

is the $n \times n$ **identity matrix**.

Example 2.65. The 3×3 identity matrix is

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

2.7 Linear transformations

Definition 2.66 (Linear transformation)**.** A function $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ is a **linear transformation** if for all vectors $u, v \in \mathbb{R}^n$ and all scalars c,

$$\mathrm{T}(u+v) = \mathrm{T}(u) + \mathrm{T}(v) \quad \text{and} \quad \mathrm{T}(c\,u) = c\,\mathrm{T}(u).$$

Theorem 2.67 (Properties of linear transformations)**.** *If* T *is a linear transformation, then*

a) $\mathrm{T}(0) = 0$.

b) $\mathrm{T}(c_1 u + c_2 v) = c_1\,\mathrm{T}(u) + c_2\,\mathrm{T}(v)$ *for all scalars* c_1 *and* c_2 *and all vectors* u *and* v*.*

Let us first see some examples of functions that are *not* linear transformations.

Example 2.68.

a) The function $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathrm{T}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1^2 \\ x_2 \end{bmatrix}$$

is not a linear transformation: it fails to preserve addition, since for example,

$$\mathrm{T}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = T\left(\begin{bmatrix} 3 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 9 \\ 0 \end{bmatrix}$$

while

$$\mathrm{T}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) + \mathrm{T}\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \end{bmatrix} = \begin{bmatrix} 5 \\ 0 \end{bmatrix}.$$

One can also see that this function does not preserve scaling:

$$\mathrm{T}\left(2\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = \mathrm{T}\left(\begin{bmatrix} 2 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 4 \\ 0 \end{bmatrix},$$

but

$$2\,\mathrm{T}\left(\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right) = 2\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ 0 \end{bmatrix}.$$

b) Let $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^3$ be the function given by

$$\mathrm{T}\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \begin{bmatrix} x_1 + x_2 + 1 \\ -x_2 \\ x_1 \end{bmatrix}.$$

This is not a linear function, since it fails to preserve addition and scaling. But an even easier way to see that it fails to preserve addition is to note that

$$\mathrm{T}\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} \neq \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

It turns out that every linear transformation is actually a matrix transformation.

Theorem 2.69. *A function* $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ *is a linear transformation if and only if it is a matrix transformation, meaning that there exists a matrix* A *such that*

$$\mathrm{T}(x) = Ax \quad \text{for all } x \in \mathbb{R}^n.$$

To find this matrix A, we do the following:

Definition 2.70 (Standard matrix of a linear transformation)**.** Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be the standard basis of $\mathbb{R}^n$. Given a linear transformation $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$, consider the matrix

$$A = \begin{bmatrix} \mathrm{T}(\mathbf{e}_1) & \cdots & \mathrm{T}(\mathbf{e}_n) \end{bmatrix}.$$

The matrix A is called the **standard matrix** of T and it satisfies

$$\mathrm{T}(x) = Ax \qquad \text{for all } x \in \mathbb{R}^n.$$

Remark 2.71. Let us check that the standard matrix of a linear transformation does what we claim it does. Suppose that T is a linear transformation with standard matrix A. Given any vector $x \in \mathbb{R}^n$, we saw earlier that we can decompose x into its components and write it as a linear combination of the standard basis elements:

$$x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n.$$

Then

$$\begin{aligned} \mathrm{T}(x) &= \mathrm{T}(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) \\ & x_1\,\mathrm{T}(\mathbf{e}_1) + \cdots + x_n\,\mathrm{T}(\mathbf{e}_n) && \text{since } T \text{ is a linear transformation} \\ &= Ax && \text{since the } \mathrm{T}(\mathbf{e}_i) \text{ are the columns of } A. \end{aligned}$$

This shows that T is in fact a matrix transformation, with associated matrix A.

Let us see some examples.

Example 2.72 (Dilation in $\mathbb{R}^2$)**.** Let us find the standard matrix for the dilation $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by

$$\mathrm{T}(x) = 2x.$$

It is not hard to check that this is indeed a linear transformation. Since

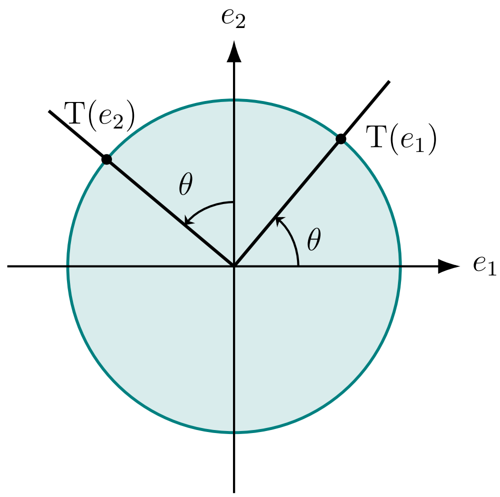
$$\mathrm{T}(\mathbf{e}_1) = \begin{bmatrix} 2 \\ 0 \end{bmatrix} \qquad \text{and} \qquad \mathrm{T}(\mathbf{e}_2) = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

we conclude that the standard matrix for this linear transformation is

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

Example 2.73. Generalizing what we saw in Example 2.63, the standard matrix for the identity function $\mathbb{R}^n \to \mathbb{R}^n$ is the identity matrix.

Example 2.74 (Rotation in the plane)**.** Consider the function $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ that rotates each point counterclockwise by an angle θ (in radians).



Then using trigonometry, one can show that

$$\mathrm{T}\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}\cos\theta\\ \sin\theta\end{bmatrix} \qquad \text{and} \qquad \mathrm{T}\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}-\sin\theta\\ \cos\theta\end{bmatrix}.$$

and thus the standard matrix for this linear transformation is

$$\begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}.$$

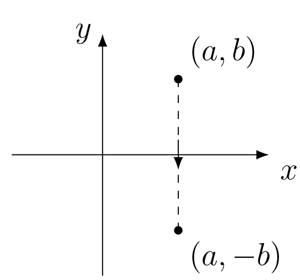
We conclude that

$$\mathrm{T}\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = \begin{bmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{bmatrix}\begin{bmatrix}x_1\\x_2\end{bmatrix}.$$

Example 2.75. Consider the matrix transformation $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ given by $\mathrm{T}(x) = Ax$, where

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

Note that $\mathrm{T}\left(\begin{bmatrix}a\\b\end{bmatrix}\right) = \begin{bmatrix}a\\-b\end{bmatrix}$. Geometrically,

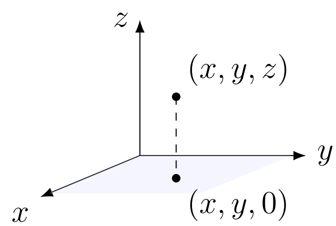


This is the reflection across the x-axis.

Example 2.76 (Geometric description in $\mathbb{R}^3$)**.** Let us give a geometric description of the matrix transformation with standard matrix

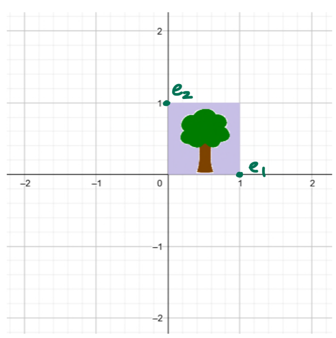
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

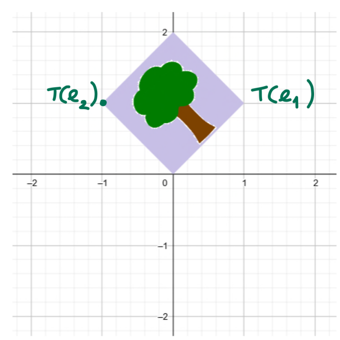
Note that $\quad \mathrm{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ 0 \end{bmatrix}. \quad$ Geometrically,



This is the orthogonal projection of $\mathbb{R}^3$ onto the xy-plane.

Example 2.77. Consider the linear transformation $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ that does the following:





Note how we explicitly marked the images of $\mathbf{e}_1$ and $\mathbf{e}_2$. How can we tell? We know that $\mathrm{T}(0) = 0$, so looking for the other bottom corner of the tree we find the image of $\mathrm{T}(e_1)$. Moreover, the $\mathrm{T}(e_2)$ must be the opposite corner. While a linear transformation might stretch things, it will not change the fact that e_1 and e_2 are on opposite corners of the tree.

This is sufficient for us to find the standard matrix, and thus to completely describe the linear transformation: the matrix is

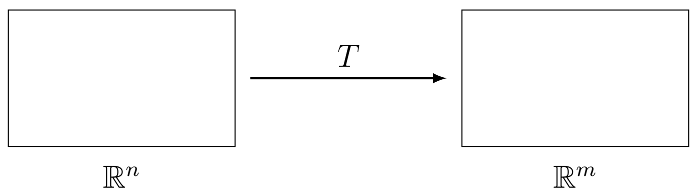
$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}.$$

2.8 Injective and surjective maps

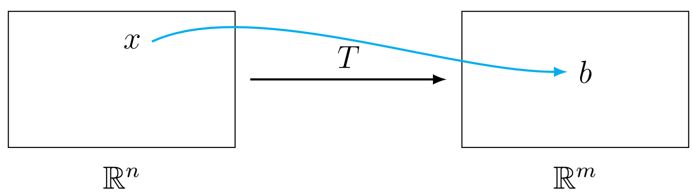
Definition 2.78. A function $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ has **domain** $\mathbb{R}^n$ and **codomain** $\mathbb{R}^m$.

Informally, the domain is the set of all inputs, and the codomain is the set of all *possible* outputs, whether or not they are *actual* outputs. Saying the codomain is $\mathbb{R}^m$ means that all the outputs are vectors in $\mathbb{R}^m$, but not that every vector in $\mathbb{R}^m$ can be obtained as a specific output.

Example 2.79 (In a picture)**.** We can visualize T as mapping the "domain box" to the "codomain box":



Here is a visual depiction of an input x going to an output b:

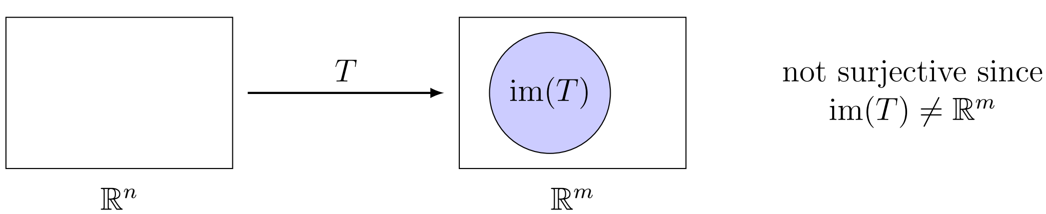


Definition 2.80 (Image or range)**.** The **image** or **range** of a function $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ is

$$\operatorname{im}(\mathrm{T}) := \{\, \mathrm{T}(x) \mid x \in \mathbb{R}^n \,\} \subseteq \mathbb{R}^m.$$

Definition 2.81 (Surjective function)**.** Let $\mathrm{T} : \mathbb{R}^n \to \mathbb{R}^m$ be a function. We say T is **surjective** if for every $b \in \mathbb{R}^m$ there exists *at least one* $x \in \mathbb{R}^n$ such that $\mathrm{T}(x) = b$. Equivalently, T is surjective if $\operatorname{im}(\mathrm{T}) = \mathbb{R}^m$, meaning the image is the entire codomain. Some authors also use the word **onto**.

Example 2.82. A function is <u>not</u> surjective if $\operatorname{im}(T)$ is a proper subset of $\mathbb{R}^m$.

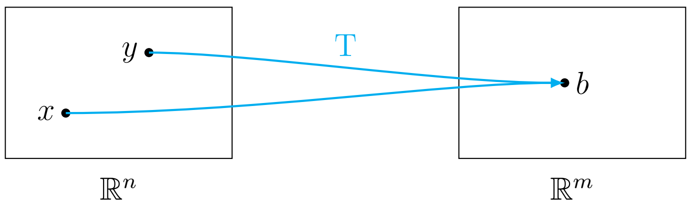


Definition 2.83 (Injective function)**.** Let $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ be a function. We say T is **injective** if for each $b \in \mathbb{R}^m$ there exists *at most one* $x \in \mathbb{R}^n$ such that $\mathrm{T}(x) = b$. Equivalently,

$$\mathrm{T}(x_1) = \mathrm{T}(x_2) \implies x_1 = x_2.$$

Remark 2.84. Some authors use the word **one-to-one** to refer to injective functions, but that can lead to some ambiguity, so we will avoid those words.

Remark 2.85 (In a picture)**.** A function is not injective if two different inputs map to the same output.



not injective since
$T(x) = b = T(y)$ and $x \neq y$

Definition 2.86. A function $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ is **bijective** if it is both injective and surjective.

Example 2.87. Consider the (nonlinear) function $\mathrm{T}: \mathbb{R} \to \mathbb{R}$ given by $\mathrm{T}(x) = x^2$. This function is not injective: for example, $\mathrm{T}(1) = 1 = \mathrm{T}(-1)$. It is also not surjective: $\operatorname{im} \mathrm{T}$ is just the set of nonnegative real numbers:

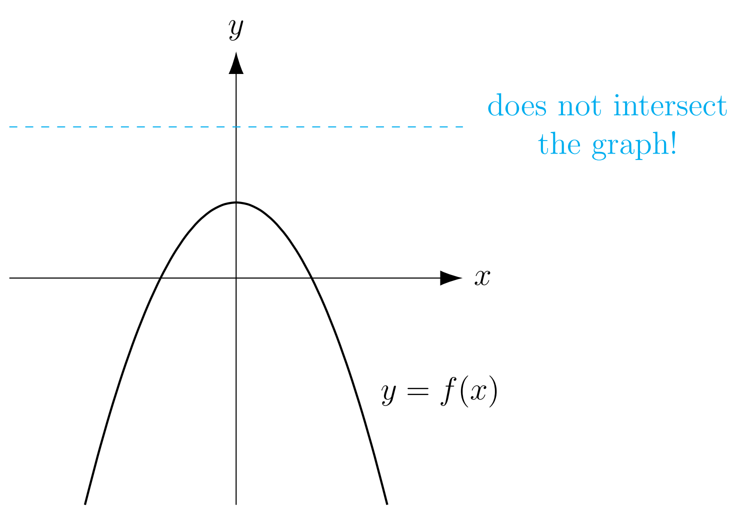
$$\operatorname{im} \mathrm{T} = \{x \in \mathbb{R} \mid x \geqslant 0\}.$$

Discussion 2.88. Let us focus on the more familiar case of functions $f: \mathbb{R} \to \mathbb{R}$, and consider the graph of such a function f. We can describe the injective and surjective properties visually:

- Surjective: f is surjective if and only if every horizontal line crosses the graph of f at least once.

  Note that if the horizontal line $y = b$ crosses the graph at (a, b), then that means that

$$f(a) = b.$$

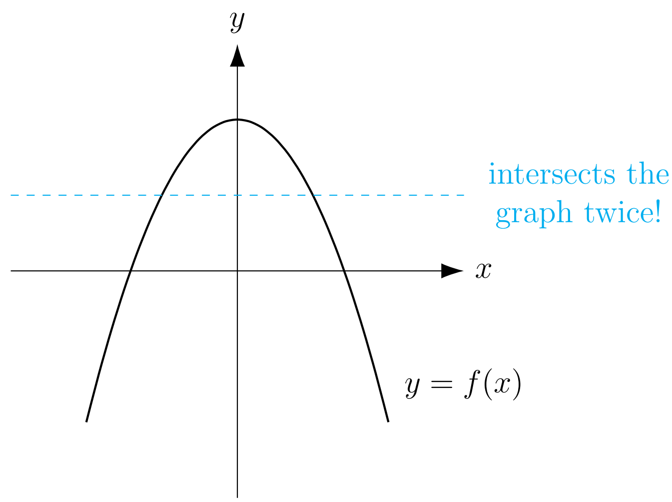


Example: f not surjective

<u>Injective:</u> f is injective if and only if every horizontal line crosses the graph of f at most once.

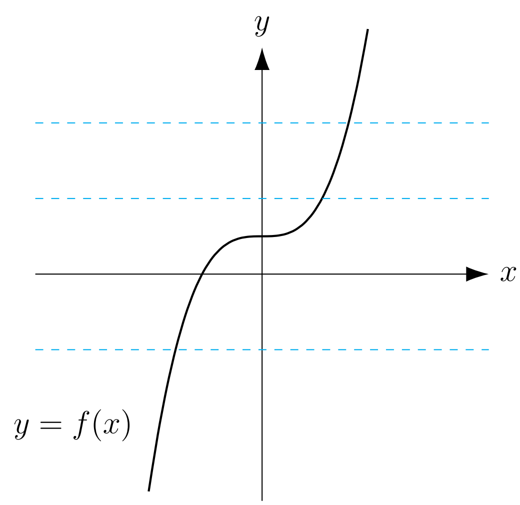
Note that if the horizontal line $y = b$ crosses the graph twice, say at (a, b) and (c, b), then that means that

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



Example: f not injective

Putting these ideas together: f is bijective if and only if every horizontal line intersects the graph of f exactly once.



Example: f bijective

Remark 2.89. Injectivity and surjectivity are <u>different</u> properties. A function $\mathrm{T} : \mathbb{R}^n \to \mathbb{R}^m$ can be

- injective but not surjective,
- surjective but not injective,
- bijective (both injective and surjective),
- or neither injective nor surjective.

We are of course interested specifically in the case where our function is a linear transformation.

Example 2.90 (Identity on $\mathbb{R}^2$ is bijective)**.** Let $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the identity map, so that

$$\mathrm{T}\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ b \end{bmatrix}.$$

- T is surjective: for any $\begin{bmatrix} a \\ b \end{bmatrix} \in \mathbb{R}^2$ we have $T\left(\begin{bmatrix} a \\ b \end{bmatrix}\right) = \begin{bmatrix} a \\ b \end{bmatrix}$.
- T is injective: each vector maps to itself, so equal outputs force equal inputs.

Example 2.91 (injective but not surjective)**.** Define $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^3$ by

$$T\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \begin{bmatrix} x \\ y \\ x+y \end{bmatrix}.$$

- This map T is not surjective because, for instance, $\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \notin \operatorname{im}(T)$.
- But T is injective: if $\mathrm{T}\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = T\left(\begin{bmatrix} u \\ v \end{bmatrix}\right)$, comparing the first two coordinates gives $x = u$ and $y = v$, so
$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} u \\ v \end{bmatrix}.$$

Example 2.92 (surjective but not injective)**.** Define $\mathrm{T}\colon \mathbb{R}^3 \to \mathbb{R}^2$ by

$$\mathrm{T}\left(\begin{bmatrix} x \\ y \\ z \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}.$$

- This function T is surjective because for any $\begin{bmatrix} x \\ y \end{bmatrix} \in \mathbb{R}^2$ we have $T\left(\begin{bmatrix} x \\ y \\ 0 \end{bmatrix}\right) = \begin{bmatrix} x \\ y \end{bmatrix}$.

- But T is not injective since, for example,

$$T\left(\begin{bmatrix}1\\0\\0\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix} = T\left(\begin{bmatrix}1\\0\\1\end{bmatrix}\right).$$

We can characterize injectivity via the kernel:

Definition 2.93. The **kernel** of a linear transformation $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ is the set

$$\ker(\mathrm{T}) := \{x \in \mathbb{R}^n \mid \mathrm{T}(x) = 0\}.$$

Remark 2.94. Note that the kernel of any linear transformation always contains the zero vector.

Theorem 2.95. *Let $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. Then T is injective if and only if the equation $\mathrm{T}(x) = 0$ has only the trivial solution $x = 0$. Equivalently, T is injective if and only if $\ker(T) = \{0\}$.*

We can also decide if a linear transformation is injective or surjective by looking at the RREF of the corresponding standard matrix.

Theorem 2.96. *Let $\mathrm{T}\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation with standard matrix A.*

a) The linear transformation T is surjective if and only if the columns of A span $\mathbb{R}^m$.

Equivalently: T is surjective if and only if A has a pivot in every row.

b) The linear transformation T is injective if and only if the columns of A are linearly independent.

Equivalently: T is surjective if and only if A has no free variables, meaning it has a pivot in every column.

Example 2.97 (Identity map is bijective)**.** The identity map $\mathrm{T}\colon \mathbb{R}^3 \to \mathbb{R}^3$ is both surjective and injective. And in fact, this is the linear transformation with standard matrix

$$I_3 = \begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}$$

which has a pivot in every column and every row.

Example 2.98 (surjective but not injective)**.** Let $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}$ be given by

$$\mathrm{T}\left(\begin{bmatrix}x_1\\x_2\end{bmatrix}\right) = x_1.$$

Its standard matrix is $A = \begin{bmatrix}1 & 0\end{bmatrix}$. Thus T is surjective, as A has a pivot in every row. In fact, we can see that for each $b \in \mathbb{R}$ we can take

$$\mathrm{T}\left(\begin{bmatrix}b\\0\end{bmatrix}\right) = b.$$

On the other hand, T is not injective since there is no pivot on the second column. In fact, we can see that for example

$$T\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = T\left(\begin{bmatrix}1\\1\end{bmatrix}\right) = 1.$$

Example 2.99 (injective but not surjective)**.** Let us again consider the map $\mathrm{T}\colon \mathbb{R}^2 \to \mathbb{R}^3$ from Example 2.91, given by

$$T\left(\begin{bmatrix}x\\y\end{bmatrix}\right) = \begin{bmatrix}x\\y\\x+y\end{bmatrix}.$$

Its standard matrix is

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

Thus T is injective but not surjective.

Example 2.100. Consider the linear transformation with standard matrix

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

Row-reducing gives

$$A \sim \begin{bmatrix}1 & 0 & -\frac{4}{3}\\0 & 1 & 0\\0 & 0 & 0\end{bmatrix}.$$

There is a pivot in each of the first two rows but none in the third, so T is not surjective. There is a missing pivot in the third column (a free variable), so T is not injective.

Chapter 3

Matrix operations

3.1 Adding and multiplying matrices

Let A be a matrix. Recall that the (i, j)-th entry of A is the value on the ith row and jth column. In what follows, we will write $A = [a_{ij}]$ to indicate that the matrix A has a_{ij} in the (i, j)-th entry. More precisely,

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

Definition 3.1 (Sum of matrices)**.** Let A and B be two $m \times n$ matrices. The **sum** of A and B is the $m \times n$ matrix $A + B$ whose (i, j)-th entry is the sum of the (i, j)-th entries of A and B. More precisely, if $A = [a_{ij}]$ and $B = [b_{ij}]$, then

$$A + B := [a_{ij} + b_{ij}].$$

Note that the sum of two matrices is only defined if they have the same size.

Example 3.2. We have $\begin{bmatrix} 3 & -1 \\ 2 & -11 \end{bmatrix} + \begin{bmatrix} 1 & 3 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 3+1 & -1+3 \\ 2+4 & -11+5 \end{bmatrix} = \begin{bmatrix} 4 & 2 \\ 6 & -6 \end{bmatrix}$.

Definition 3.3 (Multiplication by Scalars)**.** If c is a scalar and A is an $m \times n$ matrix, then cA is the $m \times n$ matrix whose entries are obtained by multiplying all the entries of A by c. If $A = [a_{ij}]$, then $cA = [ca_{ij}]$.

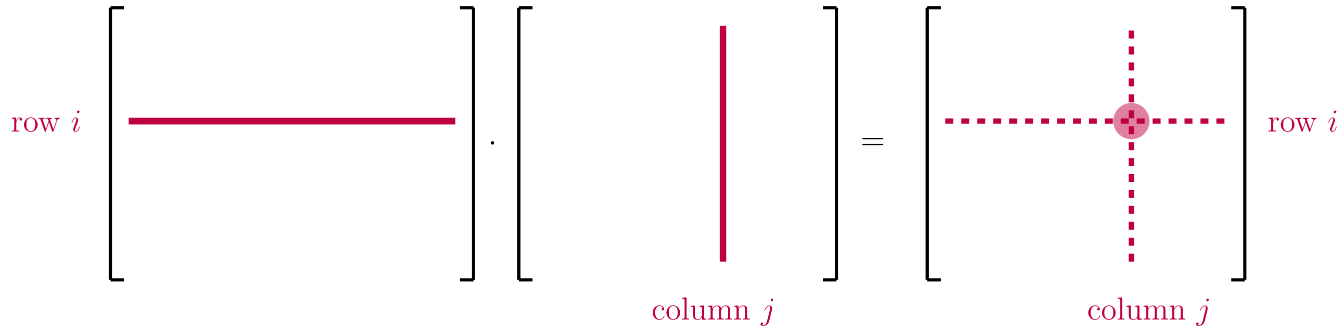
Example 3.4. We have $3 \cdot \begin{bmatrix} -1 & 1 \\ -1 & 5 \end{bmatrix} = \begin{bmatrix} -3 & 3 \\ -3 & 15 \end{bmatrix}$.

Definition 3.5 (Matrix Multiplication)**.** If A is an $m \times n$ matrix and B is an $n \times p$ matrix, then AB is the $m \times p$ matrix obtained by multiplying rows of A with columns of B: if b_j is the j-th column of B, then

$$AB = \begin{bmatrix} Ab_1 & Ab_2 & \cdots & Ab_p \end{bmatrix}$$

Note that Ab_i is a vector with m rows.

In a picture: to find the (i, j)th entry of AB, we focus on row i of A and column j of B



and the (i, j)th entry of AB is obtained by summing up the products of the successive elements in this row and this column, as follows:

$$(i, j)\text{th entry of } AB = a_{i1}b_{1j} + \cdots + a_{in}b_{nj}.$$

Remark 3.6. The multiplication AB is only defined if the number of columns of A matches the number of rows of B.

Example 3.7. The product

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

is not defined.

Example 3.8. We have

$$\begin{aligned} \begin{bmatrix} 2 & 3 \\ 1 & -5 \end{bmatrix} \begin{bmatrix} 4 & 3 & 6 \\ 1 & -2 & 3 \end{bmatrix} &= \begin{bmatrix} 2 \times 4 + 3 \times 1 & 2 \times 3 + 3 \times (-2) & 2 \times 6 + 3 \times 3 \\ 1 \times 4 - 5 \times 1 & 1 \times 3 + (-5) \times (-2) & 1 \times 6 + (-5) \times 3 \end{bmatrix} \\ &= \begin{bmatrix} 11 & 0 & 21 \\ -1 & 13 & -9 \end{bmatrix} \end{aligned}$$

Theorem 3.9 (Properties of Matrix Multiplication)**.** *Let A, B, and C be matrices, and assume that their sizes are such that the products AB and BC make sense.*

1) Associativity: $(AB)C = A(BC)$.

2) Left distributivity: $A(B + C) = AB + AC$.

3) Right distributivity: $(B + C)A = BA + CA$.

4) For any scalar α, $\alpha(AB) = (\alpha A)B = A(\alpha B)$.

5) Let I_m denote the $m \times m$ identity matrix. Then $I_m A = A = A I_n$.

Remark 3.10. Warning: the order of the matrices in a product matters! In fact, it could even be that one product is defined and the other one is not. But even if the two matrices are square, we may have $AB \neq BA$.

Here is an example:

Example 3.11. Consider
$$A = \begin{bmatrix} 5 & 1 \\ -1 & 3 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 2 & 0 \\ 4 & 3 \end{bmatrix}.$$
Then
$$AB = \begin{bmatrix} 5 \cdot 2 + 1 \cdot 4 & 5 \cdot 0 + 1 \cdot 3 \\ (-1) \cdot 2 + 3 \cdot 4 & (-1) \cdot 0 + 3 \cdot 3 \end{bmatrix} = \begin{bmatrix} 14 & 3 \\ 10 & 9 \end{bmatrix}$$
while
$$BA = \begin{bmatrix} 2 \cdot 5 + 0 \cdot (-1) & 2 \cdot 1 + 0 \cdot 3 \\ 4 \cdot 5 + 3 \cdot (-1) & 4 \cdot 1 + 3 \cdot 3 \end{bmatrix} = \begin{bmatrix} 10 & 2 \\ 17 & 13 \end{bmatrix}.$$
Hence $AB \neq BA$.

Definition 3.12 (Zero matrix)**.** The zero $m \times n$ matrix is the $m \times n$ matrix whose entries are all 0. We sometimes denote it simply by 0, if the size is clear from context.

Remark 3.13. Warning: cancellation fails. If $AB = AC$, it does *not* follow that $B = C$. Similarly, $BA = CA$ does *not* imply that $B = C$ Moreover, $AB = 0$ does *not* imply $A = 0$ or $B = 0$.

Example 3.14. Let
$$A = \begin{bmatrix} 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$
Then $AB = \begin{bmatrix} 0 \end{bmatrix}$, but $A \neq 0$ and $B \neq 0$.

Definition 3.15 (Powers of a matrix)**.** If A is an $n \times n$ (square) matrix, we define the powers of A by
$$A^2 = AA, \quad A^3 = AAA, \ \ldots, \quad A^k = \underbrace{A \cdot A \cdots A}_{k \text{ times}}.$$
We also set
$$A^1 = A \quad \text{and} \quad A^0 = I_n.$$

Definition 3.16. If A is an $m \times n$ matrix, then its **transpose** A^T is the $n \times m$ matrix whose rows are the columns of A.

Example 3.17. The transpose of
$$A = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{bmatrix} \quad \text{is} \quad A^\mathsf{T} = \begin{bmatrix} 1 & 3 \\ 2 & 4 \\ 5 & 6 \end{bmatrix}.$$

Theorem 3.18 (Facts about Transposes)**.** *Let A and B be matrices whose sizes make the following make sense.*

1) $(A^\mathsf{T})^\mathsf{T} = A$.

2) $(A + B)^\mathsf{T} = A^\mathsf{T} + B^\mathsf{T}$.

3) $(\alpha A)^\mathsf{T} = \alpha A^\mathsf{T}$ *for any scalar* α.

4) $(AB)^\mathsf{T} = B^\mathsf{T} A^\mathsf{T}$.

Definition 3.19. A square matrix A is **symmetric** if $A = A^\mathsf{T}$.

Index