Math 309: Introduction to Linear Algebra

Mark Sullivan

November 30, 2022

Contents

1 Systems of linear equations

1.1 Geometric view of systems of equations

First, let's introduce some notation. The set of real numbers is denoted by $\mathbb{R}$:

$$\mathbb{R} = \left\{ x \middle| x \text{ is a real number} \right\}. \tag{1}$$

You have dealt with this set before; you probably referred to it by the name "the real line."

Definition 1.1 *A linear equation (over $\mathbb{R}$) is any equation of the form*

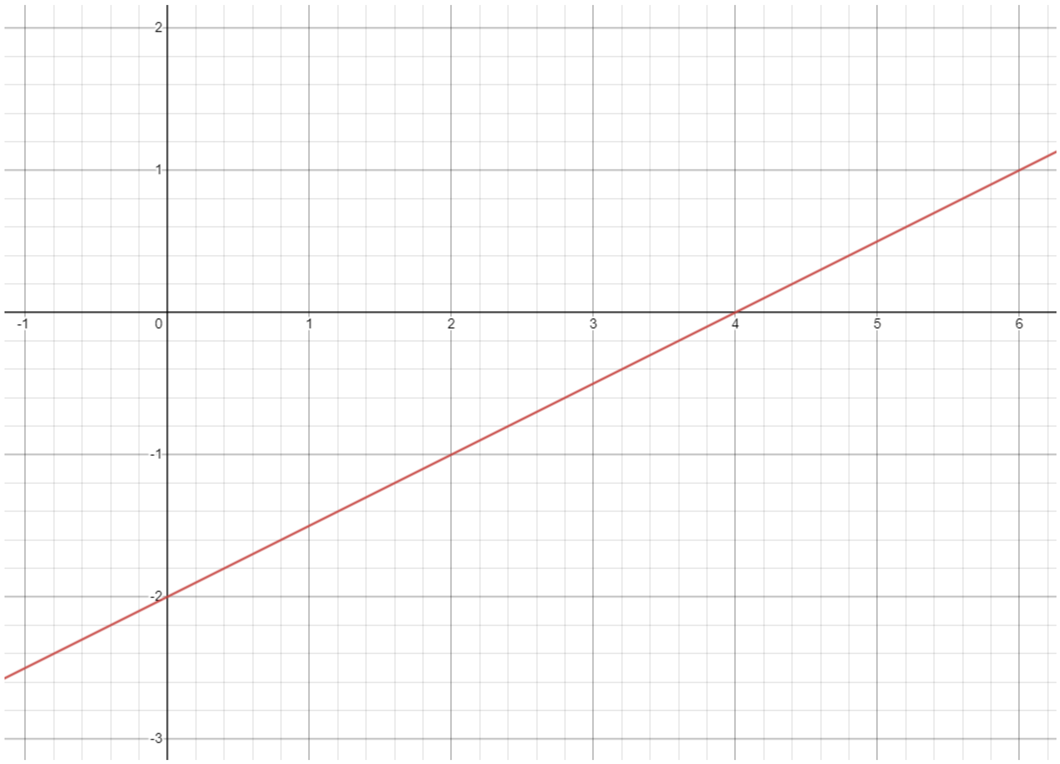
$$a_1x_1 + a_2x_2 + ... + a_nx_n = b, \tag{2}$$

where x_1, x_2, ..., x_n are real-valued variables, and $a_1, a_2, ..., a_n, b \in \mathbb{R}$ are constants. (Here the symbol "$\in$" means "is an element of.") The constants $a_1, a_2, ..., a_n$ are called the coefficients of the equation, while the constant b is called the constant term of the equation.

We can understand a linear equation as corresponding to a set of points (known as the "graph" of the equation) whose coordinates satisfy that equation. For example, given the line L described by an equation $y = \frac{1}{2}x - 2$, we can understand L as the set of points whose coordinates are real values and obey the equation:

$$L = \left\{ (x, y) \middle| y = \frac{1}{2}x - 2 \right\}. \tag{3}$$

This set can be drawn in the xy-plane as



In general, linear equations represent certain shapes living in space:

$$\begin{array}{ccc} a_1x_1 + a_2x_2 = b & \text{represents} & \text{a line living in the } x_1x_2\text{-plane} \\ a_1x_1 + a_2x_2 + a_3x_3 = b & \text{represents} & \text{a plane living in 3-space} \\ a_1x_1 + a_2x_2 + a_3x_3 + a_4x_4 = b & \text{represents} & \text{a hyperplane (?) living in 4-space} \end{array} \tag{4}$$

As you can see, once the number of variables involved in the equation exceeds 3, the equation becomes much harder to visualize as a shape in space. This is simply a limitation of the human intellect.

The problem becomes even more pronounced when we consider *systems* of linear equations. The general setup of a system of linear equations is

$$\begin{array}{c} a_{11}x_1 + a_{12}x_2 + ... + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + ... + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + ... + a_{mn}x_n = b_m \end{array} . \tag{5}$$

In these situations, we want to describe the set of points $(x_1, x_2, ..., x_n)$ that satisfy *all* of these equations *at the same time*. Given a system of two lines, we want to find the point of intersection, if any. Given a system of two planes, we can to find the line of intersection, if any. But when given a system of two hyperplanes, the problem of finding their plane of intersection becomes unreasonable from a purely geometric point of view.

Linear algebra is a set of tools that allows us to study linear equations and their systems without necessarily having a visual representation of the geometric objects that they represent.

1.2 Algebraic view of systems of equations

Question: how many solutions does a system of linear equations have?
Answer: It depends on the system.

Example 1.2 *Solve the system of linear equations*

$$\begin{aligned} x_1 + x_2 &= 1 \\ x_1 - x_2 &= 1 \end{aligned}. \tag{6}$$

There are many different ways of handling this; most of them you probably learned in high school. Here's just one: add the two equations together to get

$$2x_1 = 2. \tag{7}$$

This makes it clear that $x_1 = 1$*. Therefore, either equation can allow us to deduce that* $x_2 = 0$*. Therefore, there exists exactly one solution:* $\boxed{(1, 0)}$*. Geometrically, we can graph the two equations as lines in 2-dimensional space. The solution is the point of intersection of the lines.* □

Example 1.3 *Solve the system of linear equations*

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 3x_1 + 2x_2 &= 2 \end{aligned}. \tag{8}$$

Any solution of this system would be a point (x_1, x_2) that satisfies both equations. This would lead to a contradiction, since then $3x_1 + 2x_2$ would be both 1 and 2, leading us to conclude that $1 = 2$. As this is an unacceptable conclusion, we must admit that there exists $\boxed{\textit{no solution}}$ *of this system. If we were to graph these equations, we would find two lines that have no intersection point.* □

Definition 1.4 *We say that a system of linear equations is* <u>*inconsistent*</u> *provided that it has no solution.*

Example 1.5 *Solve the system of linear equations*

$$\begin{aligned} x_1 + x_2 &= 1 \\ 2x_1 + 2x_2 &= 2 \end{aligned}. \tag{9}$$

The fact of the matter is that the two equations are describing the same set of points: points (x_1, x_2) in 2-dimensional space whose coordinates sum to 1. There are infinitely many such points, so there are infinitely many solutions. Specifically, for any value of x_1, we can write $(x_1, 1 - x_1)$ as a solution. To describe the "solution set," we write

$$\boxed{\left\{(t, 1-t) \,\middle|\, t \in \mathbb{R}\right\}}. \tag{10}$$

□

As these examples illustrate, a system of linear equations could have one solution, no solution, or an infinite set of solutions. In fact, these are the *only* possibilities, in any number of variables.

1.3 Elementary operations

Definition 1.6 *Two systems of linear equations are called* <u>*equivalent*</u> *provided that they have the same set of solutions.*

Our main task for right now is as follows. Given a system of linear equations, we would like to find an equivalent system whose solutions are more readily obvious. It is most convenient to do this by performing "elementary row operations" on an augmented matrix.

Every system of linear equations can be re-written as an "augmented matrix." For example, the linear system

$$\begin{aligned} x_1 + 2x_2 - x_3 &= 1 \\ 5x_1 + 2x_2 + x_3 &= 2 \\ x_2 + 3x_3 &= 9 \end{aligned} \tag{11}$$

can be re-written as

$$\left(\begin{array}{ccc|c} 1 & 2 & -1 & 1 \\ 5 & 2 & 1 & 2 \\ 0 & 1 & 3 & 9 \end{array}\right) \tag{12}$$

Definition 1.7 *Let A be a matrix. The* <u>*elementary row operations on A*</u> *are as follows.*
(i) Exchanging rows of A.
(ii) Multiplying a row of A by a nonzero number.
(iii) Adding a multiple of one row of A to another row of A.

Definition 1.8 *Let A and B be matrices. We say that* <u>*A and B are row-equivalent*</u> *provided that B can be obtained from A by a finite sequence of elementary row operations.*

Theorem 1.9 *Let S_1 and S_2 be two systems of linear equations, and let A_1 and A_2 be their augmented matrices, respectively. If A_2 is row-equivalent to A_1, then S_1 and S_2 are equivalent systems.*

Example 1.10 *Find all points of intersection of the following planes in 3-space:*

$$\begin{aligned} P_1 &= \big\{(x,y,z)\big| x+y+z=6\big\} \\ P_2 &= \big\{(x,y,z)\big| 3x-2y+3z=3\big\}. \\ P_3 &= \big\{(x,y,z)\big| 2x-y+z=4\big\} \end{aligned} \tag{13}$$

This geometric question can be solved algebraically. Specifically, we need only find those points (x, y, z) with real-valued coordinates which satisfy all three equations *simultaneously. In other words, we seek to solve the following system of linear equations:*

$$\begin{aligned} x + y + z &= 6 \\ 3x - 2y + 3z &= 3. \\ 2x - y + z &= 4 \end{aligned} \tag{14}$$

To this end, we set up an augmented matrix:

$$\begin{pmatrix} 1 & 1 & 1 & | & 6 \\ 3 & -2 & 3 & | & 3 \\ 2 & -1 & 1 & | & 4 \end{pmatrix}. \tag{15}$$

We can do some elementary row operations to the augmented matrix without changing the solution set:

$$\begin{aligned}
&\begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ \boxed{3} & -2 & 3 & | & 3 \\ \boxed{2} & -1 & 1 & | & 4 \end{pmatrix} \xrightarrow{R_2 - 3R_1 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{-5} & 0 & | & -15 \\ \boxed{2} & -1 & 1 & | & 4 \end{pmatrix} \\
&\xrightarrow{R_3 - 2R_1 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{-5} & 0 & | & -15 \\ 0 & \boxed{-3} & -1 & | & -8 \end{pmatrix} \xrightarrow{\frac{1}{5}R_2 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{-1} & 0 & | & -3 \\ 0 & \boxed{-3} & -1 & | & -8 \end{pmatrix} \\
&\xrightarrow{R_3 - 3R_2 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{-1} & 0 & | & -3 \\ 0 & 0 & \boxed{-1} & | & 1 \end{pmatrix} \xrightarrow{-1R_2 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{1} & 0 & | & 3 \\ 0 & 0 & \boxed{-1} & | & 1 \end{pmatrix} \\
&\xrightarrow{-1R_3 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 6 \\ 0 & \boxed{1} & 0 & | & 3 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix} \xrightarrow{R_1 - R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 3 \\ 0 & \boxed{1} & 0 & | & 3 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix} \\
&\xrightarrow{R_1 - R_3 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 4 \\ 0 & \boxed{1} & 0 & | & 3 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix}
\end{aligned} \tag{16}$$

This final augmented matrix corresponds to the following system:

$$\begin{aligned} 1x + 0y + 0z &= 4 \\ 0x + 1y + 0z &= 3 \\ 0x + 0y + 1z &= -1 \end{aligned} \;, \tag{17}$$

from which we deduce that the point $\boxed{(4,3,-1)}$ *is the only point of intersection of the three planes.* □

1.4 Gaussian elimination

The previous example was our first example of the use of a method called "Gauss-Jordan elimination," or "row reduction." This procedure has two steps: "Gaussian elimination," and then "Jordan elimination." We'll start with Gaussian elimination.

The goal of Gaussian elimination is to find a "row echelon form" of the original matrix. This is defined as follows:

Definition 1.11 *Let A be a matrix. Given a row i of A, a* <u>*pivot point of the ith row of A*</u> *is a nonzero entry of row i that has no nonzero entries to its left. A column containing a pivot point of A is called a* <u>*pivot column of A*</u>.

Definition 1.12 *Let A be a matrix. Given a matrix B,* <u>*B is a row echelon form of A*</u> *provided that the following statements are true.*
(i) B is row-equivalent to A.
(ii) All zero-rows of B are below all nonzero-rows of B.
(iii) Each pivot point of B is in a column that is right of the pivot point of any row above it.

Example 1.13 *The following matrices are row echelon forms. The pivot points of each matrix are highlighted.*

$$\begin{pmatrix} \boxed{2} & 0 & 0 & 1 \\ 0 & \boxed{-1} & 1 & 1 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}, \quad \begin{pmatrix} \boxed{-5} & 1 & 2 \\ 0 & \boxed{-1} & 4 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{pmatrix}. \tag{18}$$

*The following matrices are **not** row echelon forms. The pivot points of each matrix are highlighted.*

$$\begin{pmatrix} \boxed{1} & 0 & 0 \\ \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \end{pmatrix}, \quad \begin{pmatrix} \boxed{3} & \frac{1}{2} & \pi & 0 \\ 0 & 0 & 0 & 0 \\ 0 & \boxed{2} & 3 & 4 \end{pmatrix}, \quad \begin{pmatrix} 0 & \boxed{1} & 8 & 9 & 0 \\ \boxed{9} & 0 & 0 & 1 & 0 \end{pmatrix} \tag{19}$$

□

In Gaussian elimination, our only goal to to find a row echelon form of our given (augmented) matrix.

Example 1.14 *Find a row echelon form for the augmented matrix*

$$\left(\begin{array}{ccc|c} 1 & -1 & 1 & 0 \\ 2 & -1 & 1 & 1 \\ 3 & 1 & -3 & 2 \end{array}\right). \tag{20}$$

We do some elementary row operations, with the intent to eventually get a row echelon form:

$$\left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ \boxed{2} & -1 & 1 & 1 \\ \boxed{3} & 1 & -3 & 2 \end{array}\right) \xrightarrow{R_2-2R_1\to R_2} \left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ 0 & \boxed{1} & -1 & 1 \\ \boxed{3} & 1 & -3 & 2 \end{array}\right)$$
$$\xrightarrow{R_3-3R_1\to R_3} \left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ 0 & \boxed{1} & -1 & 1 \\ 0 & \boxed{4} & -6 & 2 \end{array}\right) \xrightarrow{R_3-4R_2\to R_3} \boxed{\left(\begin{array}{ccc|c} \boxed{1} & -1 & 1 & 0 \\ 0 & \boxed{1} & -1 & 1 \\ 0 & 0 & \boxed{-2} & -2 \end{array}\right)}. \tag{21}$$

*Note: this is **not the only** row echelon form associated to this augmented matrix.* □

1.5 Gauss-Jordan elimination

The second step, Jordan elimination, involves taking an echelon form matrix and finding an equivalent "reduced row echelon form:"

Definition 1.15 *Let A be a matrix. Given a matrix B, B is a reduced row echelon form of A provided that the following statements are true.*
(i) B is a row echelon form of A.
(ii) Each pivot point of B is 1.
(iii) All entries above a pivot point are 0.

Example 1.16 *The following matrices are reduced row echelon forms. The pivot points of each matrix are highlighted.*

$$\begin{pmatrix} \boxed{1} & 0 & 0 \\ 0 & \boxed{1} & 0 \\ 0 & 0 & \boxed{1} \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & 0 & 5 & 2 \\ 0 & \boxed{1} & 1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & \boxed{1} & 0 & 0 & 5 \\ 0 & 0 & \boxed{1} & 0 & -1 \\ 0 & 0 & 0 & \boxed{1} & 2 \end{pmatrix}. \tag{22}$$

The following matrices are row echelon forms that are ***not*** *reduced row echelon forms. The pivot points of each matrix are highlighted.*

$$\begin{pmatrix} \boxed{1} & 0 \\ 0 & \boxed{6} \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & 2 & 0 \\ 0 & \boxed{1} & 0 \end{pmatrix}, \quad \begin{pmatrix} \boxed{1} & 0 & 0 & 1 \\ 0 & \boxed{1} & 0 & 0 \\ 0 & 0 & 0 & \boxed{1} \end{pmatrix}. \tag{23}$$

Example 1.17 *Find a reduced row echelon form for the following augmented matrix:*

$$\begin{pmatrix} 1 & 1 & -1 & | & 0 \\ 1 & 1 & 1 & | & 2 \\ 2 & -1 & 0 & | & 1 \end{pmatrix} \tag{24}$$

First, we implement Gaussian elimination, seeking a row echelon form:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ \boxed{1} & 1 & 1 & | & 2 \\ \boxed{2} & -1 & 0 & | & 1 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & 0 & \boxed{2} & | & 2 \\ \boxed{2} & -1 & 0 & | & 1 \end{pmatrix} \\ &\xrightarrow{R_3 - 2R_1 \to R_3} \begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & 0 & \boxed{2} & | & 2 \\ 0 & \boxed{-3} & 2 & | & 1 \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & \boxed{-3} & 2 & | & 1 \\ 0 & 0 & \boxed{2} & | & 2 \end{pmatrix}. \end{aligned} \tag{25}$$

Now we do Jordan elimination to get a reduced row echelon form:

$$\begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & \boxed{-3} & 2 & | & 1 \\ 0 & 0 & \boxed{2} & | & 2 \end{pmatrix} \xrightarrow{R_2 - R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & \boxed{-3} & 0 & | & -1 \\ 0 & 0 & \boxed{2} & | & 2 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}R_3 \to R_3} \begin{pmatrix} \boxed{1} & 1 & -1 & | & 0 \\ 0 & \boxed{-3} & 0 & | & -1 \\ 0 & 0 & \boxed{1} & | & 1 \end{pmatrix} \xrightarrow{R_1 + R_3 \to R_1} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 1 \\ 0 & \boxed{-3} & 0 & | & -1 \\ 0 & 0 & \boxed{1} & | & 1 \end{pmatrix}$$

$$\xrightarrow{-\frac{1}{3}R_2 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 1 \\ 0 & \boxed{1} & 0 & | & \frac{1}{3} \\ 0 & 0 & \boxed{1} & | & 1 \end{pmatrix} \xrightarrow{R_1 - R_2 \to R_1} \boxed{\begin{pmatrix} \boxed{1} & 0 & 0 & | & \frac{2}{3} \\ 0 & \boxed{1} & 0 & | & \frac{1}{3} \\ 0 & 0 & \boxed{1} & | & 1 \end{pmatrix}}. \quad (26)$$

□

Finding a reduced row echelon form of an augmented matrix allows one to easily read off the solutions of the corresponding system. For this reason, Gauss-Jordan elimination is a valuable computational tool.

1.7 Uniqueness of the reduced echelon form

Theorem 1.18 *Let A be a matrix. If B_1 and B_2 are reduced row echelon forms of A, then $B_1 = B_2$.*

In other words, every matrix corresponds to a *unique* reduced row echelon form.

Definition 1.19 *Let S be a system of m equations in the variables $x_1, x_2, ..., x_n$, and let A be its augmented matrix. (Note that A has m rows and $n+1$ columns.) Given $i \in \{1, 2, ..., n\}$, we say that x_i is a pivot variable of S provided that* ***the*** *reduced row echelon form of A has a pivot point in the ith column. We say that x_i is a free variable of S provided that it is not a pivot variable.*

Example 1.20 *Describe the set of solutions of the following system of linear equa-*

tions:

$$S : \begin{cases} 3x - y + 5z = 8 \\ y - 10z = 1 \\ 6x - y = 17 \end{cases} . \tag{27}$$

First, we set up an augmented matrix for S:

$$A = \begin{pmatrix} 3 & -1 & 5 & | & 8 \\ 0 & 1 & -10 & | & 1 \\ 6 & -1 & 0 & | & 17 \end{pmatrix} \tag{28}$$

We proceed with Gauss-Jordan elimination:

$$\begin{aligned} &\begin{pmatrix} \boxed{3} & -1 & 5 & | & 8 \\ 0 & \boxed{1} & -10 & | & 1 \\ \boxed{6} & -1 & 0 & | & 17 \end{pmatrix} \xrightarrow{R_3 - 2R_1 \to R_3} \begin{pmatrix} \boxed{3} & -1 & 5 & | & 8 \\ 0 & \boxed{1} & -10 & | & 1 \\ 0 & \boxed{1} & -10 & | & 1 \end{pmatrix} \\ &\xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} \boxed{3} & -1 & 5 & | & 8 \\ 0 & \boxed{1} & -10 & | & 1 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \xrightarrow{R_1 + R_2 \to R_1} \begin{pmatrix} \boxed{3} & 0 & -5 & | & 9 \\ 0 & \boxed{1} & -10 & | & 1 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \\ &\xrightarrow{\frac{1}{3}R_1 \to R_1} \begin{pmatrix} \boxed{1} & 0 & -\frac{5}{3} & | & 3 \\ 0 & \boxed{1} & -10 & | & 1 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} . \end{aligned} \tag{29}$$

According to our definitions, x and y are pivot variables, while z is a free variable. The system corresponding to our reduced row echelon form can be written as

$$\begin{matrix} x - \frac{5}{3}z = 3 \\ y - 10z = 1 \end{matrix} . \tag{30}$$

Any (x, y, z) with coordinates in $\mathbb{R}$ that satisfies these two descriptions is a solution. We consider z as a "parameter," calling it t, so that

$$\begin{matrix} x = 3 + \frac{5}{3}t \\ y = 1 + 10t \end{matrix} . \tag{31}$$

Now, if t is any real number whatsoever, then the x and y described above will determine the solution point (x, y, t). In other words, the set of solutions is the infinite set

$$\boxed{\left\{\left(3+\tfrac{5}{3}t, 1+10t, t\right) \,\middle|\, t \in \mathbb{R}\right\}}. \tag{32}$$

For example, $(3,1,0)$ is a solution (letting $t = 0$), as is $(8,31,3)$ (letting $t = 3$), as well as $\left(3+\frac{5}{3}, 11, 1\right)$ (letting $t = 1$), and so on; there is a different solution for each real value of t. $\square$

Definition 1.21 *Let A be a matrix, and let B be* ***the*** *reduced row echelon form of A. The rank of A is the number of pivot points in B.*

Example 1.22 *Determine the rank of the following matrix:*

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 5 & 9 \\ 2 & 4 & 6 \end{pmatrix} \tag{33}$$

To determine $\operatorname{rank}(A)$*, we need to know the reduced row echelon form of A. We proceed with Gauss-Jordan elimination:*

$$\begin{pmatrix} \boxed{1} & 2 & 3 \\ \boxed{1} & 5 & 9 \\ \boxed{2} & 4 & 6 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{3} & 6 \\ \boxed{2} & 4 & 6 \end{pmatrix} \xrightarrow{R_3 - 2R_1 \to R_3} \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{3} & 6 \\ 0 & 0 & 0 \end{pmatrix}$$
$$\xrightarrow{\frac{1}{3}R_2 \to R_3} \begin{pmatrix} \boxed{1} & 2 & 3 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{R_1 - 2R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & -1 \\ 0 & \boxed{1} & 2 \\ 0 & 0 & 0 \end{pmatrix}. \tag{34}$$

This tells us that $\boxed{\operatorname{rank}(A) = 2}$*, the number of pivot points in the reduced row echelon form of A.*

By definition, the rank of a matrix must be less than or equal to the number of columns of that matrix. This motivates the following definition.

Definition 1.23 *Let A be a matrix with n many columns. We say that A has full rank provided that* $\operatorname{rank}(A) = n$.

Theorem 1.24 *Let S be a system of m equations in n variables, and let A be* ***the left side*** *of its augmented matrix. The number of free variables of S is* $n - \operatorname{rank}(A)$.

A system has a unique solution (that is, only one point as a solution) if and only if it has no free variables. Therefore, the above theorem tells us that a system will have a unique solution if and only if the left side of its augmented matrix has full rank.

For applications of the techniques that we discussed in this section, read sections 1.9, 1.10, and/or 1.11 of the textbook.

2 Vectors in $\mathbb{R}^n$

2.1 Points and vectors

Definition 2.1 *A* <u>*vector [over $\mathbb{R}$]*</u> *is a line segment, together with a specific direction.*

Just as a point has coordinates, which indicate how far it is from each of the axes, a vector has "components," which indicate how far along each axis it goes. We denote a vector using a matrix with a single column. For instance, a vector that extends from the point $(1, 1)$ to $(6, -2)$ has the representation

$$\overrightarrow{v} = \begin{pmatrix} 5 \\ -3 \end{pmatrix}. \tag{35}$$

Two vectors should be regarded as the same (equal) if they have the same length and direction. This means that translating a vector from one location in space to a different location in space without altering its length or direction has no effect. So, for instance, the vector from $(1, 1)$ to $(6, -2)$ is the same as the vector from $(0, 0)$ to $(5, -3)$. To summarize, two vectors are equal if and only if they have equal corresponding components:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \quad \text{if and only if} \quad \begin{matrix} a_1 = b_1, \text{ and} \\ a_2 = b_2, \text{ and} \\ \vdots \\ a_n = b_n. \end{matrix} \tag{36}$$

Definition 2.2 *The* <u>*zero vector in $\mathbb{R}^n$*</u> *is the vector in n-space whose components are all zero:*

$$\overrightarrow{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{37}$$

Definition 2.3 *Let $p = (x_1, x_2, ..., x_n)$ be a point in n-space. The* <u>*position vector*</u>

of p is the vector that goes from the origin to p. In other words, the position vector $\overrightarrow{p}$ is given by

$$\overrightarrow{p} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}. \tag{38}$$

Real-valued quantities that are not vectors will sometimes be called "scalars," in order to distinguish them.

Let's introduce some notation. The set of vectors of real numbers with two components is denoted $\mathbb{R}^2$:

$$\mathbb{R}^2 = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \middle| x, y \in \mathbb{R} \right\} \tag{39}$$

Similarly, the set of vectors of real numbers with three components is denoted $\mathbb{R}^3$:

$$\mathbb{R}^3 = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| x, y, z \in \mathbb{R} \right\}. \tag{40}$$

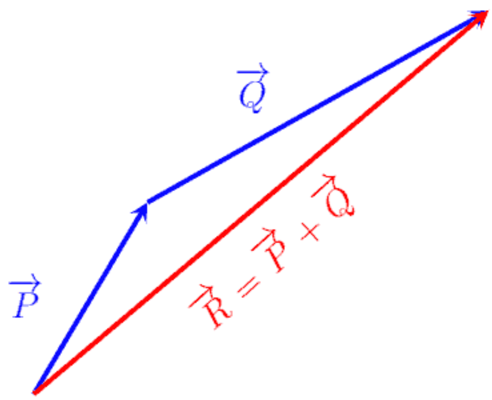
We can extend this to any positive integer n:

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| x_1, x_2, ..., x_n \in \mathbb{R} \right\}. \tag{41}$$

These sets are typically called "n-dimensional Euclidean space."

2.2 Addition

Vectors in $\mathbb{R}^n$ can be added together. Geometrically, adding the vectors $\overrightarrow{P}$ and $\overrightarrow{Q}$ looks like this:



Algebraically, vector addition just requires adding together the corresponding components of each vector:

$$\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix} \tag{42}$$

2.3 Scalar multiplication

Vectors in $\mathbb{R}^n$ can be multiplied by scalars. Geometrically, if the scalar is $r > 0$, then $r\overrightarrow{v}$ refers to a vector in the same direction as $\overrightarrow{v}$ with a length that is r times the length of $\overrightarrow{v}$. If the scalar is $r < 0$, then $r\overrightarrow{v}$ refers to a vector in the opposite direction as $\overrightarrow{v}$ with a length that is $|r|$ times the length of $\overrightarrow{v}$. If the scalar is $r = 0$, then $r\overrightarrow{v} = \overrightarrow{0}$, no matter what.

Algebraically, we can multiply a vector by a scalar by just multiplying each component by that scalar:

$$r\begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} ra_1 \\ ra_2 \\ \vdots \\ ra_n \end{pmatrix}. \tag{43}$$

Notice that the following properties are true for vectors in $\mathbb{R}^n$:

Theorem 2.4 *The following statements are true.*
(i) Given $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3} \in \mathbb{R}^n$*,*

$$\overrightarrow{v_1} + (\overrightarrow{v_2} + \overrightarrow{v_3}) = (\overrightarrow{v_1} + \overrightarrow{v_2}) + \overrightarrow{v_3} \tag{44}$$

(vector addition is associative).
(ii) There exists a vector $\overrightarrow{0} \in \mathbb{R}^n$ *such that for any vector* $\overrightarrow{v} \in \mathbb{R}^n$*,*

$$\overrightarrow{v} + \overrightarrow{0} = \overrightarrow{0} + \overrightarrow{v} = \overrightarrow{v} \tag{45}$$

(an additive identity exists for vector addition).
(iii) For each vector $\overrightarrow{v} \in \mathbb{R}^n$*, there exists a vector* $\overrightarrow{u} \in \mathbb{R}^n$ *such that*

$$\overrightarrow{v} + \overrightarrow{u} = \overrightarrow{u} + \overrightarrow{v} = \overrightarrow{0} \tag{46}$$

(each vector has an additive inverse).
(iv) Given vectors $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathbb{R}^n$*,*

$$\overrightarrow{v_1} + \overrightarrow{v_2} = \overrightarrow{v_2} + \overrightarrow{v_1} \tag{47}$$

(vector addition is commutative).
(v) Given a vector $\overrightarrow{v} \in \mathbb{R}^n$ *and scalars* $s, t \in \mathbb{R}$*,*

$$s\left(t\overrightarrow{v}\right) = (st)\overrightarrow{v} \tag{48}$$

(scalar multiplication is associative).
(vi) Given vectors $\overrightarrow{v_1}\overrightarrow{v_2} \in \mathbb{R}^n$ *and a scalar* $s \in \mathbb{R}$*,*

$$s\left(\overrightarrow{v_1} + \overrightarrow{v_2}\right) = s\overrightarrow{v_1} + s\overrightarrow{v_2} \tag{49}$$

(scalar multiplication distributes over vector addition)
(vii) Given a vector $\overrightarrow{v} \in \mathbb{R}^n$ *and scalars* $s, t \in \mathbb{R}$*,*

$$(s + t)\overrightarrow{v} = s\overrightarrow{v} + t\overrightarrow{v} \tag{50}$$

(scalar multiplication distributes over scalar addition).
(viii) There exists a scalar $1 \in \mathbb{R}$ *such that for any vector* $\overrightarrow{v} \in \mathbb{R}^n$*,*

$$1\overrightarrow{v} = \overrightarrow{v} \tag{51}$$

(a multiplicative identity exists for scalar multiplication).

Later we will see that any set that satisfies the above properties is called a "vector space over $\mathbb{R}$," so for now, we'll just say that $\mathbb{R}^n$ is a vector space over $\mathbb{R}$ (whatever that means).

2.4 Linear combinations

Definition 2.5 *Let* $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ *be some vectors in* $\mathbb{R}^n$ *for some positive integer* n*. Given any scalars* $s_1, s_2, ..., s_k \in \mathbb{R}$*, the vector*

$$\overrightarrow{u} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} \tag{52}$$

is called a <u>*linear combination of* $v_1, v_2, ..., v_k$ *over* $\mathbb{R}$</u>*.*

Example 2.6 *The vector*

$$\overrightarrow{u} = \begin{pmatrix} -1 \\ -2 \\ -2 \end{pmatrix} \tag{53}$$

is a linear combination of the vectors

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \overrightarrow{v_3} = \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}, \tag{54}$$

because $\overrightarrow{u} = 3\overrightarrow{v_1} - 2\overrightarrow{v_2} + 5\overrightarrow{v_3}$. □

Example 2.7 *The vector*

$$\overrightarrow{u} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{55}$$

*is **not** a linear combination of the vectors*

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \tag{56}$$

because for any scalars $s_1, s_2 \in \mathbb{R}$, $\overrightarrow{u} \neq s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2}$. $\square$

Question: given vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ and $\overrightarrow{u}$ in $\mathbb{R}^n$, how can we tell whether $\overrightarrow{u}$ can be expressed as a linear combination of the vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ over $\mathbb{R}$?

This same question can be understood as a problem about systems of equations. Suppose that

$$\overrightarrow{v_1} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix}, \quad ..., \quad \overrightarrow{v_k} = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}, \quad \overrightarrow{u} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{57}$$

We are asking whether there exist scalars $s_1, s_2, ..., s_k \in \mathbb{R}$ such that

$$s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} = \overrightarrow{u}. \tag{58}$$

In other words, we want to know whether there exist scalars $s_1, s_2, ..., s_k \in \mathbb{R}$ such that

$$s_1 \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} + s_2 \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} + ... + s_k \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{59}$$

This equation can be written as

$$\begin{pmatrix} s_1a_{11} + s_2a_{12} + ... + s_ka_{1k} \\ s_1a_{21} + s_2a_{22} + ... + s_ka_{2k} \\ \vdots \\ s_1a_{n1} + s_2a_{n2} + ... + s_ka_{nk} \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}. \tag{60}$$

As mentioned, these two vectors are equal if and only if their corresponding components are equal. Thus, the scalars exist if and only if there exists a solution to the system of equations

$$S : \begin{cases} s_1 a_{11} + s_2 a_{12} + ... + s_k a_{1k} = b_1 \\ s_1 a_{21} + s_2 a_{22} + ... + s_k a_{2k} = b_2 \\ \vdots \\ s_1 a_{n1} + s_2 a_{n2} + ... + s_k a_{nk} = b_n \end{cases} . \tag{61}$$

Example 2.8 *Express the vector*

$$\overrightarrow{u} = \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} \tag{62}$$

as a linear combination of the vectors

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} \quad \overrightarrow{v_3} = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \tag{63}$$

or prove that no such expression is possible.

We seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ *such that*

$$s_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + s_3 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix}. \tag{64}$$

This is tantamount to asking for a solution to the system

$$S : \begin{cases} s_1 + s_2 = 3 \\ s_1 + 2s_2 + s_3 = 4 \\ s_1 + s_2 + s_3 = 3 \end{cases} . \tag{65}$$

To determine whether such solutions exist, we set up an augmented matrix:

$$\begin{pmatrix} 1 & 1 & 0 & | & 3 \\ 1 & 2 & 1 & | & 4 \\ 1 & 1 & 1 & | & 3 \end{pmatrix}. \tag{66}$$

We now apply Gauss-Jordan elimination to find the reduced row echelon form:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 1 & 0 & | & 3 \\ \boxed{1} & 2 & 1 & | & 4 \\ \boxed{1} & 1 & 1 & | & 3 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 3 \\ 0 & \boxed{1} & 1 & | & 1 \\ \boxed{1} & 1 & 1 & | & 3 \end{pmatrix} \\ &\xrightarrow{R_3 - R_1 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 3 \\ 0 & \boxed{1} & 1 & | & 1 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \xrightarrow{R_2 - R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 3 \\ 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \\ &\xrightarrow{R_1 - R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 2 \\ 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \end{aligned} \tag{67}$$

The reduced row echelon form corresponds to the system

$$\begin{cases} 1s_1 + 0s_2 + 0s_3 = 2 \\ 0s_1 + 1s_2 + 0s_3 = 1 \\ 0s_1 + 0s_2 + 1s_3 = 0 \end{cases}. \tag{68}$$

In other words, if we let $s_1 = 2$, $s_2 = 1$, and $s_3 = 0$, then we can write the linear combination $\overrightarrow{u} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + s_3\overrightarrow{v_3}$:

$$\boxed{\begin{pmatrix} 3 \\ 4 \\ 3 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + 1\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + 0\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}}. \tag{69}$$

□

Example 2.9 *Express the vector*

$$\overrightarrow{u} = \begin{pmatrix} -1 \\ -2 \\ -4 \\ -10 \end{pmatrix} \tag{70}$$

as a linear combination of the vectors

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix}, \quad \overrightarrow{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \tag{71}$$

or prove that no such expression is possible.

We seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ *such that*

$$s_1 \begin{pmatrix} 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \\ 2 \\ 3 \end{pmatrix} + s_3 \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ -2 \\ -4 \\ -10 \end{pmatrix}. \tag{72}$$

This is the same as looking for solutions to the system

$$S : \begin{cases} s_1 + 0s_1 + s_3 = -1 \\ s_1 + s_2 + s_3 = -2 \\ 0s_1 + 2s_2 + s_3 = -4 \\ s_1 + 3s_2 + s_3 = -10 \end{cases}, \tag{73}$$

which can be rephrased using an augmented matrix:

$$\begin{pmatrix} 1 & 0 & 1 & | & -1 \\ 1 & 1 & 1 & | & -2 \\ 0 & 2 & 1 & | & -4 \\ 1 & 3 & 1 & | & -10 \end{pmatrix}. \tag{74}$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{1} & 0 & 1 & | & -1 \\ \boxed{1} & 1 & 1 & | & -2 \\ 0 & \boxed{2} & 1 & | & -4 \\ \boxed{1} & 3 & 1 & | & -10 \end{pmatrix} \xrightarrow{R_2-R_1\to R_2} \begin{pmatrix} \boxed{1} & 0 & 1 & | & -1 \\ 0 & \boxed{1} & 0 & | & -1 \\ 0 & \boxed{2} & 1 & | & -4 \\ \boxed{1} & 3 & 1 & | & -10 \end{pmatrix}$$
$$\xrightarrow{R_3-2R_2\to R_3} \begin{pmatrix} \boxed{1} & 0 & 1 & | & -1 \\ 0 & \boxed{1} & 0 & | & -1 \\ 0 & 0 & \boxed{1} & | & -2 \\ \boxed{1} & 3 & 1 & | & -10 \end{pmatrix} \xrightarrow{R_4-R_1\to R_4} \begin{pmatrix} \boxed{1} & 0 & 1 & | & -1 \\ 0 & \boxed{1} & 0 & | & -1 \\ 0 & 0 & \boxed{1} & | & -2 \\ 0 & 3 & 0 & | & -9 \end{pmatrix}$$
$$\xrightarrow{R_4-3R_2\to R_4} \begin{pmatrix} \boxed{1} & 0 & 1 & | & -1 \\ 0 & \boxed{1} & 0 & | & -1 \\ 0 & 0 & \boxed{1} & | & -2 \\ 0 & 0 & 0 & | & -6 \end{pmatrix}. \tag{75}$$

This is not the reduced row echelon form, but we need go no further. It is already clear that this system has no solution; the last equation of the last form would imply that $0s_1 + 0s_2 + 0s_3 = -6$*. Regardless of the values of* s_1, s_2 *and* s_3*, this is impossible, so no solution exists. Therefore:*

$$\boxed{\overrightarrow{u} \text{ cannot be expressed as a linear combination of } \overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3} \text{ over } \mathbb{R}.}$$

□

2.5 Length of a vector

Definition 2.10 *Let*

$$\overrightarrow{v} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \tag{76}$$

be a vector in $\mathbb{R}^n$*. The* *<u>magnitude</u>* *(also called the* *<u>norm</u>*) *of* $\overrightarrow{v}$ *is the value*

$$||\overrightarrow{v}|| = \sqrt{{a_1}^2 + {a_2}^2 + ... + {a_n}^2}. \tag{77}$$

Geometrically, the magnitude of a vector corresponds exactly to its length as a line segment.

Example 2.11 *Find the magnitude of the following vector in* $\mathbb{R}^4$*:*

$$\overrightarrow{v} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 2 \end{pmatrix} \tag{78}$$

This is just an application of the definition of magnitude:

$$||\overrightarrow{v}|| = \sqrt{1^2 + 2^2 + 0^2 + 2^2} = \sqrt{9} = \boxed{3}. \tag{79}$$

□

The fact that a notion of magnitude exists in $\mathbb{R}^n$ allows us to define the distance between two points (or two vectors) as the magnitude of the vector that goes between them.

2.6 The dot product

Definition 2.12 *Let*

$$\overrightarrow{u} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \quad \textit{and} \quad \overrightarrow{v} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \tag{80}$$

be vectors in $\mathbb{R}^n$*. The* <u>*dot product*</u> *of* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *is the scalar value*

$$\overrightarrow{u} \cdot \overrightarrow{v} = a_1 b_1 + a_2 b_2 + ... + a_n b_n. \tag{81}$$

Notice that for any vector $\overrightarrow{v}$ in $\mathbb{R}^n$, $\overrightarrow{v} \cdot \overrightarrow{v} = \left|\left|\overrightarrow{v}\right|\right|^2$.

The dot product is significant because it allows us to define a notion of perpendicularity in $\mathbb{R}^n$.

Definition 2.13 *(Different from the textbook) Let* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *be nonzero vectors in* $\mathbb{R}^n$*. We say that* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *are* <u>*orthogonal*</u> *provided that* $\overrightarrow{u} \cdot \overrightarrow{v} = 0$*.*

Geometrically, orthogonality is like this: there exists a plane P which contains both $\overrightarrow{u}$ and $\overrightarrow{v}$ (later, we will say that this plane is "spanned" by the vectors). The vectors $\overrightarrow{u}$ and $\overrightarrow{v}$ are orthogonal if and only if the angle between them in P is $\frac{\pi}{2}$ radians. In other words, "orthogonal" is just a generalization of "perpendicular."

Theorem 2.14 *(The Cauchy-Schwarz inequality for* $\mathbb{R}^n$*) Let* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *be vectors in* $\mathbb{R}^n$*. In that case,*

$$\left|\overrightarrow{u} \cdot \overrightarrow{v}\right| \leq \left|\left|\overrightarrow{u}\right|\right| \left|\left|\overrightarrow{v}\right|\right|. \tag{82}$$

This theorem implies another theorem:

Theorem 2.15 *(The triangle inequality for* $\mathbb{R}^n$*) Let* $\overrightarrow{u}$ *and* $\overrightarrow{v}$ *be vectors in* $\mathbb{R}^n$*. In that case,*

$$\left|\left|\overrightarrow{u} + \overrightarrow{v}\right|\right| \leq \left|\left|\overrightarrow{u}\right|\right| + \left|\left|\overrightarrow{v}\right|\right| \tag{83}$$

The proofs of these theorems can be found in the textbook. You will not be expected to be able to reproduce them in this class. However, we will use the techniques of this chapter to showcase a proof of the Pythagorean theorem:

Theorem 2.16 *(The Pythagorean theorem for $\mathbb{R}^2$) Let $\triangle ABC$ be a right triangle in $\mathbb{R}^2$. Let a be the length of the side opposite to vertex A, let b be the length of the side opposite to vertex B, and let c be the side opposite to vertex C. If the right angle of $\triangle ABC$ exists at the vertex c, then*

$$a^2 + b^2 = c^2. \tag{84}$$

Proof Define the vectors $\overrightarrow{w} = \overrightarrow{BA}$, $\overrightarrow{u} = \overrightarrow{BC}$ and $\overrightarrow{v} = \overrightarrow{CA}$. We notice that $||\overrightarrow{w}|| = c$, $||\overrightarrow{u}|| = a$, and $||\overrightarrow{v}|| = b$. Additionally,

$$\overrightarrow{w} = \overrightarrow{u} + \overrightarrow{v}. \tag{85}$$

We compute the dot product of each side of the above equation with itself:

$$\overrightarrow{w} \cdot \overrightarrow{w} = (\overrightarrow{u} + \overrightarrow{v}) \cdot (\overrightarrow{u} + \overrightarrow{v}). \tag{86}$$

The dot product distributes over vector addition, so we can re-write this as

$$\overrightarrow{w} \cdot \overrightarrow{w} = (\overrightarrow{u} \cdot \overrightarrow{u}) + (\overrightarrow{v} \cdot \overrightarrow{u}) + (\overrightarrow{u} \cdot \overrightarrow{v}) + (\overrightarrow{v} \cdot \overrightarrow{v}). \tag{87}$$

However, $\overrightarrow{u}$ and $\overrightarrow{v}$ are perpendicular, so $\overrightarrow{u} \cdot \overrightarrow{v} = \overrightarrow{v} \cdot \overrightarrow{u} = 0$. Thus, the above equation becomes

$$||\overrightarrow{w}||^2 = \overrightarrow{w} \cdot \overrightarrow{w} = (\overrightarrow{u} \cdot \overrightarrow{u}) + (\overrightarrow{v} \cdot \overrightarrow{v}) = ||\overrightarrow{u}||^2 + ||\overrightarrow{v}||^2. \tag{88}$$

In other words, $c^2 = a^2 + b^2$. $\square$

4 Matrices

We would like to be able to write any linear system of equations as a single equation consisting of some operations on some mathematical objects. Then, we could study the system by studying the mathematical objects involved in that single equation. This brings us to the topic of matrices.

4.1 Definition and equality

Definition 4.1 *An $m \times n$ matrix is a rectangular array of mathematical objects with m rows and n columns:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}. \tag{89}$$

Two $m \times n$ matrices are equal provided that they have equal corresponding entries.

Definition 4.2 *Let A be an $m \times n$ matrix. We say that A is a square matrix provided that $m = n$.*

Definition 4.3 *A matrix of size $1 \times n$ is called a row vector. A matrix of size $m \times 1$ is called a column vector.*

4.2 Addition

To add two matrices (of the same size) together, just add their corresponding entries:

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1n}+b_{1n} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2n}+b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}+b_{m1} & a_{m2}+b_{m2} & \dots & a_{mn}+b_{mn} \end{pmatrix}. \quad (90)$$

Definition 4.4 *The* $\underline{m \times n \text{ zero matrix}}$ *is the matrix of size* $m \times n$ *whose entries are all* 0*.*

4.3 Scalar multiplication

To multiply a scalar by a matrix, just multiply each entry by that scalar:

$$s \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} = \begin{pmatrix} sa_{11} & sa_{12} & \dots & sa_{1n} \\ sa_{21} & sa_{22} & \dots & sa_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ sa_{m1} & sa_{m2} & \dots & sa_{mn} \end{pmatrix}. \quad (91)$$

4.4 Matrix multiplication

Multiplying two matrices together is a bit more complicated. The motivation is as follows: we would like to define multiplication of matrices so that an augmented matrix like

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} & | & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & | & b_2 \\ \vdots & \vdots & \ddots & \vdots & | & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & | & b_m \end{pmatrix} \tag{92}$$

can be written as an equation involving a matrix and vectors, like

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}. \tag{93}$$

To accomplish this, we need to define matrix multiplication in a very particular way:

Definition 4.5 *Let M be an $m \times n$ matrix, and let N be an $n \times p$ matrix. The* <u>*matrix product MN*</u> *is the $m \times p$ matrix*

$$MN = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1p} \\ c_{21} & c_{22} & \dots & c_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ c_{m1} & c_{m2} & \dots & c_{mp} \end{pmatrix}, \tag{94}$$

where for each $i \in \{1, 2, ..., m\}$ and for each $j \in \{1, 2, ..., p\}$, c_{ij} is the dot product of the ith row of M with the jth column of N.

Example 4.6 *Given the matrices*

$$A = \begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 2 & 4 & -1 & 2 \\ 1 & 0 & 2 & 1 \\ -1 & 1 & -2 & 1 \end{pmatrix}, \tag{95}$$

compute the matrix product AB.

We apply the definition of the matrix product to get

$$\begin{pmatrix} 1 & 1 & 0 \\ -1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 & 4 & -1 & 2 \\ 1 & 0 & 2 & 1 \\ -1 & 1 & -2 & 1 \end{pmatrix} = \boxed{\begin{pmatrix} 3 & 4 & 1 & 3 \\ -1 & -5 & 3 & -3 \end{pmatrix}}. \tag{96}$$

□

Definition 4.7 *The* <u>$n \times n$ *identity matrix*</u> *is the* $n \times n$ *matrix* I_n *whose entries along its main diagonal are* 1*, and whose other entries are* 0.

Example 4.8 *The* 2×2*,* 3×3 *and* 4×4 *identity matrices are shown below.*

$$I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{97}$$

□

The identity matrix I_n has the property that for any $m \times n$ matrix M, $MI_n = M$, and for any $n \times m$ matrix N, $I_n N = N$.

4.5 Matrix inverses

Definition 4.9 *Let* A *be an* $n \times n$ *matrix. An* <u>*inverse matrix of* A</u> *is an* $n \times n$ *matrix* B *such that* $AB = BA = I_n$*. If such a matrix exists, we say that* A *is an* <u>*invertible matrix*</u>*.*

Theorem 4.10 *If* A *is an invertible* $n \times n$ *matrix, then there exists exactly one inverse matrix, called* A^{-1}.

Given a square matrix A, how can we either find A^{-1}, or prove that A is not invertible?

Theorem 4.11 *Let* A *be an* $n \times n$ *matrix. The matrix* A *is invertible if and only if* A *is row-equivalent to* I_n.

The procedure for finding the inverse is as follows. Augment the original matrix with the identity matrix on the right. Then, do Gauss-Jordan elimination. If the left side's reduced row echelon form is the identity matrix, then the inverse matrix is on the right side. If it is not the identity matrix, then the matrix is not invertible.

Example 4.12 *Find the inverse matrix of*

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}, \tag{98}$$

or prove that A is not invertible.

We set up an augmented matrix with I_2 on the right:

$$\begin{pmatrix} 1 & 2 & | & 1 & 0 \\ 2 & 3 & | & 0 & 1 \end{pmatrix}, \tag{99}$$

and proceed with Gauss-Jordan elimination:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 2 & | & 1 & 0 \\ \boxed{2} & 3 & | & 0 & 1 \end{pmatrix} \xrightarrow{R_2 - 2R_1 \to R_2} \begin{pmatrix} \boxed{1} & 2 & | & 1 & 0 \\ 0 & \boxed{-1} & | & -2 & 1 \end{pmatrix} \\ &\xrightarrow{-R_2 \to R_2} \begin{pmatrix} \boxed{1} & 2 & | & 1 & 0 \\ 0 & \boxed{1} & | & 2 & -1 \end{pmatrix} \xrightarrow{R_1 - 2R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & | & -3 & 2 \\ 0 & \boxed{1} & | & 2 & -1 \end{pmatrix}. \end{aligned} \tag{100}$$

Therefore, the inverse matrix is

$$\boxed{A^{-1} = \begin{pmatrix} -3 & 2 \\ 2 & -1 \end{pmatrix}}. \tag{101}$$

□

Example 4.13 *Find the inverse matrix of*

$$A = \begin{pmatrix} -4 & -2 \\ 2 & 1 \end{pmatrix}, \tag{102}$$

or prove that A is not invertible.

We set up an augmented matrix with I_2 on the right:

$$\begin{pmatrix} -4 & -2 & | & 1 & 0 \\ 2 & 1 & | & 0 & 1 \end{pmatrix}, \tag{103}$$

and proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{-4} & -2 & | & 1 & 0 \\ \boxed{2} & 1 & | & 0 & 1 \end{pmatrix} \xrightarrow{R_1+2R_2 \to R_1} \begin{pmatrix} 0 & 0 & | & 1 & 2 \\ \boxed{2} & 1 & | & 0 & 1 \end{pmatrix}$$
$$\xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} \boxed{2} & 1 & | & 0 & 1 \\ 0 & 0 & | & 1 & 2 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1 \to R_1} \begin{pmatrix} \boxed{1} & \frac{1}{2} & | & 0 & \frac{1}{2} \\ 0 & 0 & | & 1 & 2 \end{pmatrix}. \tag{104}$$

This is the reduced row echelon form. However, the left side is not the identity matrix I_2. Therefore, the matrix A is not invertible. □

4.7 The transpose

Definition 4.14 *Let A be an $m \times n$ matrix. The transpose of A is the $n \times m$ matrix A^T such that for all $i \in \{1, 2, ..., n\}$ and for all $j \in \{1, 2, ..., m\}$, the (i, j)-entry of A^T is the (j, i) entry of A.*

Example 4.15 *Find the transpose of the following matrix:*

$$A = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \end{pmatrix}. \tag{105}$$

The transpose is

$$\boxed{A^T = \begin{pmatrix} 1 & 5 \\ 2 & 6 \\ 3 & 7 \\ 4 & 8 \end{pmatrix}}. \tag{106}$$

□

Definition 4.16 *Let A be an $n \times n$ matrix. We say that A is a symmetric matrix provided that $A^T = A$.*

5 Spans, linear independence and bases in $\mathbb{R}^n$

5.1 Spans

Definition 5.1 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. The span of $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is the set of linear combinations of $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$:*

$$\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}) = \left\{ s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} \middle| s_1, s_2, ..., s_k \in \mathbb{R} \right\}. \quad (107)$$

In light of this definition, $\overrightarrow{u}$ is a linear combination of $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ if and only if $\overrightarrow{u} \in \operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$.

Example 5.2 *Describe* $\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2})$*, where*

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} \quad \textit{and} \quad \overrightarrow{v_2} = \begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix}. \quad (108)$$

By definition,

$$\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}) = \left\{ s\begin{pmatrix} 1 \\ 2 \\ 0 \\ 1 \end{pmatrix} + t\begin{pmatrix} -1 \\ 1 \\ -1 \\ 1 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} s-t \\ 2s+t \\ -t \\ s+t \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}. \quad (109)$$

□

Definition 5.3 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. We say that $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is redundant provided that there exists some $i \in \{1, 2, ..., k\}$ such that*

$$\operatorname{span}(v_1, v_2, ..., v_{i-1}, v_{i+1}, ..., v_k) = \operatorname{span}(v_1, v_2, ..., v_{i-1}, v_i, v_{i+1}, ..., v_k). \quad (110)$$

Example 5.4 *Determine whether $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ is redundant, where*

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_3} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}. \tag{111}$$

We notice that

$$\begin{aligned} \text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\right) &= \left\{ r \begin{pmatrix} 1 \\ 1 \end{pmatrix} + s \begin{pmatrix} 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \end{pmatrix} \middle| r, s, t \in \mathbb{R} \right\} \\ &= \left\{ \begin{pmatrix} r+t \\ r+s \end{pmatrix} \middle| r, s, t \in \mathbb{R} \right\} \end{aligned} \tag{112}$$

However, this is the same set as

$$\left\{ \begin{pmatrix} u \\ v \end{pmatrix} \middle| u, v \in \mathbb{R} \right\} = \left\{ u \begin{pmatrix} 1 \\ 0 \end{pmatrix} + v \begin{pmatrix} 0 \\ 1 \end{pmatrix} \middle| u, v \in \mathbb{R} \right\} = \text{span}\left(\overrightarrow{v_2}, \overrightarrow{v_3}\right). \tag{113}$$

Therefore $\text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}\right) = \text{span}\left(\overrightarrow{v_2}, \overrightarrow{v_3}\right)$*, and so* $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ *is* $\boxed{\textit{redundant}}$. □

5.2 Linear independence

Under what circumstances is a finite sequence of vectors redundant? Suppose that $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is a redundant finite sequence of vectors. This means that

$$\text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}\right) = \text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_{i-1}}, \overrightarrow{v_{i+1}}, \overrightarrow{v_k}\right), \tag{114}$$

for some $i \in \{1, 2, ..., k\}$. We notice that

$$\overrightarrow{v_i} \in \text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}\right), \tag{115}$$

so this means that

$$\overrightarrow{v_i} \in \text{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_{i-1}}, \overrightarrow{v_{i+1}}, \overrightarrow{v_k}\right). \tag{116}$$

In other words, $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is redundant if and only if some one of the vectors is a linear combination of the others.

Definition 5.5 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. We say that this finite sequence of vectors is linearly dependent (over $\mathbb{R}$) provided that some one of them is a linear combination of the others. Otherwise, we say that the finite sequence is linearly independent (over $\mathbb{R}$).*

Suppose that $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$ is a linearly dependent finite sequence of vectors. This is the case if and only if, for some $i \in \{1, 2, ..., k\}$, $\overrightarrow{v_i}$ is a linear combination of $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_{i-1}}, \overrightarrow{v_{i+1}}, ..., \overrightarrow{v_k}$. This means that there exists some finite sequence of scalars $s_1, s_2, ..., s_{i-1}, s_{i+1} ..., s_k \in \mathbb{R}$ such that

$$\overrightarrow{v_i} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_{i-1}\overrightarrow{v_{i-1}} + s_{i+1}\overrightarrow{v_{i+1}} + ... + s_k\overrightarrow{v_k}. \tag{117}$$

In other words,

$$s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_{i-1}\overrightarrow{v_{i-1}} + (-1)\,\overrightarrow{v_i} + s_{i+1}\overrightarrow{v_{i+1}} + ... + s_k\overrightarrow{v_k} = \overrightarrow{0}. \tag{118}$$

Thus, there exists a finite sequence of scalars $s_1, s_2, ..., s_k \in \mathbb{R}$ which are not all zero such that $s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} = \overrightarrow{0}$. In fact, we can rephrase linear dependence to mean exactly this.

Theorem 5.6 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. The finite sequence $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is linearly independent if and only if the equation*

$$s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} = \overrightarrow{0} \tag{119}$$

has only the trivial solution: $s_1 = s_2 = ... = s_k = 0$.

Example 5.7 *The vectors*

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \textit{and} \quad \overrightarrow{v_2} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \tag{120}$$

are linearly independent. We know this because if

$$s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} = \overrightarrow{0}, \tag{121}$$

then

$$s_1 \begin{pmatrix}1\\0\\0\end{pmatrix} + s_2 \begin{pmatrix}0\\0\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{122}$$

which is equivalent to saying that

$$\begin{pmatrix}s_1\\0\\s_2\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{123}$$

or in other words, $s_1 = s_2 = 0$. □

Example 5.8 *Determine whether the vectors*

$$\overrightarrow{v_1} = \begin{pmatrix}1\\0\\1\end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix}0\\1\\1\end{pmatrix} \quad \overrightarrow{v_3} = \begin{pmatrix}-2\\3\\1\end{pmatrix} \tag{124}$$

are linearly independent.

We want to know whether there exist scalars $s_1, s_2, s_3 \in \mathbb{R}$ *such that*

$$s_1 \begin{pmatrix}1\\0\\1\end{pmatrix} + s_2 \begin{pmatrix}0\\1\\1\end{pmatrix} + s_3 \begin{pmatrix}-2\\3\\1\end{pmatrix} = \begin{pmatrix}0\\0\\0\end{pmatrix}, \tag{125}$$

and not all of s_1, s_2, s_3 *are zero. In other words, we want to learn about the solutions of the following system of equations:*

$$\begin{gathered} s_1 + 0s_2 - 2s_3 = 0 \\ 0s_1 + s_2 + 3s_3 = 0 \\ s_1 + s_2 + s_3 = 0 \end{gathered} \tag{126}$$

To this end, we set up an augmented matrix:

$$\begin{pmatrix} 1 & 0 & -2 & | & 0 \\ 0 & 1 & 3 & | & 0 \\ 1 & 1 & 1 & | & 0 \end{pmatrix}, \tag{127}$$

and proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{1} & 0 & -2 & | & 0 \\ 0 & \boxed{1} & 3 & | & 0 \\ \boxed{1} & 1 & 1 & | & 0 \end{pmatrix} \xrightarrow{R_3 - R_1 \to R_3} \begin{pmatrix} \boxed{1} & 0 & -2 & | & 0 \\ 0 & \boxed{1} & 3 & | & 0 \\ 0 & \boxed{1} & 3 & | & 0 \end{pmatrix}$$
$$\xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} \boxed{1} & 0 & -2 & | & 0 \\ 0 & \boxed{1} & 3 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}. \tag{128}$$

Since a non-pivot column exists, we will have a free variable, s_3. This indicates that there will be infinitely many solutions to the system. Ergo, more solutions exist than just the trivial solution, and so the vectors $\overrightarrow{v_1}$, $\overrightarrow{v_2}$ and $\overrightarrow{v_3}$ are $\boxed{\textit{linearly dependent}}$.
□

The following theorem summarizes a few of our results so far.

Theorem 5.9 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. Let A be the matrix whose columns are $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$. The following statements are equivalent.*
(i) A has full rank.
(ii) The system $A\overrightarrow{x} = \overrightarrow{0}$ has no free variables.
(iii) The system $A\overrightarrow{x} = \overrightarrow{0}$ has a unique solution: $\overrightarrow{x} = \overrightarrow{0}$.
(iv) $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ are linearly independent.
(v) The columns of A are linearly independent.

We observe the following consequence of this theorem: given a finite sequence of vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$, if $k > n$, then the matrix A whose columns are $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ cannot possibly have full rank. Therefore:

Corollary 5.10 *Let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. If $k > n$, then $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ are linearly dependent.*

5.3 Subspaces of $\mathbb{R}^n$

Definition 5.11 *Let S_1 and S_2 be sets. We say that $\underline{S_1 \text{ is a subset of } S_2}$ provided that every element of S_1 is also an element of S_2.*

Notation The symbol $\subseteq$ means "is a subset of."

Definition 5.12 *Let $V \subseteq \mathbb{R}^n$. We say that $\underline{V \text{ is a vector subspace of } \mathbb{R}^n}$ provided that the following statements are true.*
(i) $\overrightarrow{0} \in V$.
(ii) For any $\overrightarrow{u}, \overrightarrow{v} \in V$, $\overrightarrow{u} + \overrightarrow{v} \in V$. ("$V$ is closed under addition.")
(iii) For any $\overrightarrow{v} \in V$ and any $s \in \mathbb{R}$, $s\overrightarrow{v} \in V$. ("$V$ is closed under scalar multiplication.")

Example 5.13 *The following sets are vector subspaces of $\mathbb{R}^3$:*

$$V_1 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| x, y \in \mathbb{R} \right\}, \quad V_2 = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \middle| z \in \mathbb{R} \right\}, \quad V_3 = \left\{ \overrightarrow{0} \right\}, \quad V_4 = \mathbb{R}^3. \tag{129}$$

*The following subsets of $\mathbb{R}^3$ are **not** vector subspaces of $\mathbb{R}^3$:*

$$\left\{ \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \middle| x, y \in \mathbb{R} \right\}, \quad \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \middle| xy = 0 \right\},. \tag{130}$$

□

Question: let $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in \mathbb{R}^n$. Is $\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$ always a vector subspace of $\mathbb{R}^n$?

Yes. Check each condition of the definition of a vector subspace of $\mathbb{R}^n$.

5.4 Basis and dimension

As previously mentioned, all spans are vector subspaces of $\mathbb{R}^n$. Conversely, every vector subspace of $\mathbb{R}^n$ is the span of some finite sequence of vectors. In fact, by deleting any redundancies, we can find a *linearly independent* set of vectors whose span is the given vector subspace.

Theorem 5.14 *Let V be a vector subspace of $\mathbb{R}^n$. In that case, there exist linearly independent vectors $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in V$ such that* $\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}) = V$.

Definition 5.15 *Let V be a vector subspace of $\mathbb{R}^n$. A basis of V is a finite sequence $(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$ of vectors in V such that the following statements are true.*
(i) $\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}$ is linearly independent.
(ii) $\operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k}) = V$.

Example 5.16 *Determine whether the vectors*

$$\overrightarrow{v_1} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \quad \overrightarrow{v_3} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \tag{131}$$

form a basis for $\mathbb{R}^3$.

We need to determine whether the vectors are linearly independent and span $\mathbb{R}^3$. Let

$$\overrightarrow{u} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3 \tag{132}$$

be arbitrary. We want to know whether there is a solution to the equation

$$s_1 \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \tag{133}$$

To this end, we set up an augmented matrix and proceed with Gauss-Jordan elimi-

nation:

$$\begin{pmatrix} \boxed{2} & 1 & 1 & | & x \\ \boxed{2} & 2 & 1 & | & y \\ \boxed{1} & 0 & 1 & | & z \end{pmatrix} \xrightarrow{R_1-R_3\to R_1} \begin{pmatrix} \boxed{1} & 1 & 0 & | & x-z \\ \boxed{2} & 2 & 1 & | & y \\ \boxed{1} & 0 & 1 & | & z \end{pmatrix}$$
$$\xrightarrow{R_2-2R_1\to R_2} \begin{pmatrix} \boxed{1} & 1 & 0 & | & x-z \\ 0 & 0 & \boxed{1} & | & -2x+y+2z \\ \boxed{1} & 0 & 1 & | & z \end{pmatrix}$$
$$\xrightarrow{R_3-R_2\to R_3} \begin{pmatrix} \boxed{1} & 1 & 0 & | & x-z \\ 0 & 0 & \boxed{1} & | & -2x+y+2z \\ \boxed{1} & 0 & 0 & | & 2x-y-z \end{pmatrix}$$
$$\xrightarrow{R_1-R_3\to R_1} \begin{pmatrix} 0 & \boxed{1} & 0 & | & -x+y \\ 0 & 0 & \boxed{1} & | & -2x+y+2z \\ \boxed{1} & 0 & 0 & | & 2x-y-z \end{pmatrix}$$
$$\xrightarrow{R_1\leftrightarrow R_2} \begin{pmatrix} 0 & 0 & \boxed{1} & | & -2x+y+2z \\ 0 & \boxed{1} & 0 & | & -x+y \\ \boxed{1} & 0 & 0 & | & 2x-y-z \end{pmatrix}$$
$$\xrightarrow{R_1\leftrightarrow R_3} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 2x-y-z \\ 0 & \boxed{1} & 0 & | & -x+y \\ 0 & 0 & \boxed{1} & | & -2x+y+2z \end{pmatrix}. \quad (134)$$

This shows that a solution exists: $(2x-y-z, -x+y, -2x+y+2z)$. *Therefore,* $\overrightarrow{u} \in \text{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3})$. *Since* $\overrightarrow{u}$ *is arbitrary, this shows that* $\mathbb{R}^3 \subseteq \text{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3})$, *hence* $\mathbb{R}^3 = \text{span}(\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3})$.

Additionally, the left side of the augmented matrix has full rank, so according to Theorem 5.9, $\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}$ *must be linearly independent.*

Now, by definition, $\boxed{\textit{the vectors form a basis for } \mathbb{R}^3}$. □

The previous example illustrates a general fact about bases.

Theorem 5.17 *Let* A *be an* $n \times n$ *matrix. The columns of* A *form a basis for* $\mathbb{R}^n$ *if and only if* A *is row-equivalent to* I_n.

Definition 5.18 *The* <u>*standard basis for* $\mathbb{R}^n$</u> *is the finite sequence* $(\overrightarrow{e_1}, \overrightarrow{e_2}, ..., \overrightarrow{e_n})$, *where*

$$\overrightarrow{e_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \overrightarrow{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \overrightarrow{e_3} = \begin{pmatrix} 0 \\ 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \quad ..., \quad \overrightarrow{e_n} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}. \tag{135}$$

What if the matrix is not row-equivalent to the identity matrix? In that case, just delete any vectors corresponding to non-pivot columns of the reduced row echelon form. The remaining vectors will have the same span as the original sequence, and will be linearly independent.

Example 5.19 *Find a basis for the following vector subspace of* $\mathbb{R}^3$*:*

$$V = \operatorname{span}\left(\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}, \begin{pmatrix} 3 \\ 5 \\ 7 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}\right) \tag{136}$$

We consider the matrix A*, whose columns are the vectors that span* V*:*

$$A = \begin{pmatrix} 2 & -1 & 1 & 3 & -1 \\ 0 & 0 & 3 & 5 & 1 \\ -2 & 1 & 5 & 7 & 3 \end{pmatrix}. \tag{137}$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{2} & -1 & 1 & 3 & -1 \\ 0 & 0 & \boxed{3} & 5 & 1 \\ \boxed{-2} & 1 & 5 & 7 & 3 \end{pmatrix} \xrightarrow{R_3+R_1\to R_3} \begin{pmatrix} \boxed{2} & -1 & 1 & 3 & -1 \\ 0 & 0 & \boxed{3} & 5 & 1 \\ 0 & 0 & \boxed{6} & 10 & 2 \end{pmatrix}$$
$$\xrightarrow{R_3-2R_2} \begin{pmatrix} \boxed{2} & -1 & 1 & 3 & -1 \\ 0 & 0 & \boxed{3} & 5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\frac{1}{3}R_2\to R_2} \begin{pmatrix} \boxed{2} & -1 & 1 & 3 & -1 \\ 0 & 0 & \boxed{1} & \frac{5}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$
$$\xrightarrow{R_1-R_2\to R_1} \begin{pmatrix} \boxed{2} & -1 & 0 & \frac{4}{3} & -\frac{4}{3} \\ 0 & 0 & \boxed{1} & \frac{5}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1\to R_1} \begin{pmatrix} \boxed{1} & -\frac{1}{2} & 0 & \frac{2}{3} & -\frac{2}{3} \\ 0 & 0 & \boxed{1} & \frac{5}{3} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \tag{138}$$

Deleting the second, fourth and fifth vectors gives us

$$V = \text{span}\left(\begin{pmatrix} 2 \\ 0 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}\right). \tag{139}$$

Moreover, these vectors are linearly independent, and so the following is a basis for V:

$$\boxed{\left(\begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 5 \end{pmatrix}\right)}. \tag{140}$$

□

Theorem 5.20 *Let V be a vector subspace of $\mathbb{R}^n$. A finite sequence $(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$ of vectors in V is a basis for V if and only if every vector $\overrightarrow{u} \in V$ can be* ***uniquely*** *expressed as a linear combination*

$$\overrightarrow{u} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k}. \tag{141}$$

In other words, the theorem says that once a vector is chosen, the scalars that form it as a linear combination of the basis are uniquely determined. This motivates the following definition.

Definition 5.21 *Let V be a vector subspace of $\mathbb{R}^n$. Given a basis $B = (\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$ of V and a vector $\overrightarrow{u} \in V$, the* <u>*coordinates*</u> *of $\overrightarrow{u}$ with respect to B is the finite sequence $(s_1, s_2, ..., s_k)$, where $\overrightarrow{u} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k}$.*

Example 5.22 *Consider the basis*

$$B = \left(\begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right) \tag{142}$$

of $\mathbb{R}^3$. Suppose that $\overrightarrow{u} \in \mathbb{R}^3$ can be represented with respect to the standard basis of $\mathbb{R}^3$ as

$$\overrightarrow{u} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}. \tag{143}$$

Find the coordinates of $\overrightarrow{u}$ with respect to the basis B.

We seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ such that

$$s_1 \begin{pmatrix} 1 \\ 2 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \\ 0 \end{pmatrix}. \tag{144}$$

To this end, we set up an augmented matrix:

$$\begin{pmatrix} 1 & 0 & -1 & | & 4 \\ 2 & 1 & 0 & | & 2 \\ 1 & 0 & 1 & | & 0 \end{pmatrix}, \tag{145}$$

and proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{1} & 0 & -1 & | & 4 \\ \boxed{2} & 1 & 0 & | & 2 \\ \boxed{1} & 0 & 1 & | & 0 \end{pmatrix} \xrightarrow{R_2-2R_1\to R_2} \begin{pmatrix} \boxed{1} & 0 & -1 & | & 4 \\ 0 & \boxed{1} & 2 & | & -6 \\ \boxed{1} & 0 & 1 & | & 0 \end{pmatrix}$$
$$\xrightarrow{R_3-R_1\to R_3} \begin{pmatrix} \boxed{1} & 0 & -1 & | & 4 \\ 0 & \boxed{1} & 2 & | & -6 \\ 0 & 0 & \boxed{2} & | & -4 \end{pmatrix} \xrightarrow{R_2-R_3\to R_2} \begin{pmatrix} \boxed{1} & 0 & -1 & | & 4 \\ 0 & \boxed{1} & 0 & | & -2 \\ 0 & 0 & \boxed{2} & | & -4 \end{pmatrix}$$
$$\xrightarrow{\frac{1}{2}R_3\to R_3} \begin{pmatrix} \boxed{1} & 0 & -1 & | & 4 \\ 0 & \boxed{1} & 0 & | & -2 \\ 0 & 0 & \boxed{1} & | & -2 \end{pmatrix} \xrightarrow{R_1+R_3\to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 2 \\ 0 & \boxed{1} & 0 & | & -2 \\ 0 & 0 & \boxed{1} & | & -2 \end{pmatrix} \quad (146)$$

This gives us the coordinates $s_1 = 2$, $s_2 = -2$, and $s_3 = -2$. Our notation for this is

$$\boxed{[\overrightarrow{v}]_B = \begin{pmatrix} 2 \\ -2 \\ -2 \end{pmatrix}}. \quad (147)$$

□

Theorem 5.23 *Let V be a vector subspace of $\mathbb{R}^n$. Given bases B_1 and B_2 of $\mathbb{R}^n$, B_1 and B_2 have the same number of distinct vectors.*

Definition 5.24 *Let V be a vector subspace of $\mathbb{R}^n$.*
(i) If $V = \left\{\overrightarrow{0}\right\}$, we say that the dimension of V is zero.
(ii) Assume that $V \neq \left\{\overrightarrow{0}\right\}$. The dimension of V is the number of vectors in any basis of V.

5.5 Column space, row space and null space of a matrix

Definition 5.25 *Let A be an $m \times n$ matrix. The columnspace of A is the vector subspace of $\mathbb{R}^m$ spanned by the columns of A. The nullspace of A is the vector subspace*

$$\text{null}\,(A) = \left\{\overrightarrow{v} \in \mathbb{R}^n \middle| A\overrightarrow{v} = \overrightarrow{0}\right\} \quad (148)$$

of $\mathbb{R}^n$.

Example 5.26 *Find bases for the columnspace and nullspace of the following matrix:*

$$A = \begin{pmatrix} 1 & -1 & -1 & 3 \\ 2 & -2 & 0 & 4 \end{pmatrix}. \tag{149}$$

By definition, we know that

$$\mathrm{col}\,(A) = \mathrm{span}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ -2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}\right). \tag{150}$$

However, this does not tell us a basis for $\mathrm{col}\,(A)$*, because the vectors above are not linearly independent. In order to find a linearly independent sequence with the same span, we proceed with Gauss-Jordan elimination:*

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & -1 & -1 & 3 \\ \boxed{2} & -2 & 0 & 4 \end{pmatrix} \xrightarrow{R_2 - 2R_1} \begin{pmatrix} \boxed{1} & -1 & -1 & 3 \\ 0 & 0 & 2 & -2 \end{pmatrix} \\ &\xrightarrow{\frac{1}{2}R_2 \to R_2} \begin{pmatrix} \boxed{1} & -1 & -1 & 3 \\ 0 & 0 & \boxed{1} & -1 \end{pmatrix} \xrightarrow{R_1 + R_2 \to R_1} \begin{pmatrix} \boxed{1} & -1 & 0 & 2 \\ 0 & 0 & \boxed{1} & -1 \end{pmatrix} \end{aligned} \tag{151}$$

A basis for $\mathrm{col}\,(A)$ *can be found by taking the vectors corresponding to the pivot columns of the reduced row echelon form:*

$$\boxed{\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix}\right)} \tag{152}$$

To find a basis for $\mathrm{null}\,(A)$*, we first recognize that by definition,* $\mathrm{null}\,(A)$ *is the* *set*

$$\text{null}(A) = \left\{ \vec{v} \in \mathbb{R}^4 \middle| A\vec{v} = \vec{0} \right\}$$
$$= \left\{ \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} \in \mathbb{R}^4 \middle| \begin{pmatrix} 1 & -1 & -1 & 3 \\ 2 & -2 & 0 & 4 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}. \quad (153)$$

In other words, we need to solve the matrix equation

$$\begin{pmatrix} 1 & -1 & 0 & 2 \\ 0 & 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} w \\ x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \quad (154)$$

This corresponds to the system

$$\begin{matrix} w - x + 2z = 0 \\ y - z = 0 \end{matrix}. \quad (155)$$

Here x and z are free variables. We give them the names s and t to get

$$\text{null}(A) = \left\{ \begin{pmatrix} s - 2t \\ s \\ t \\ t \end{pmatrix} \middle| s, t \in \mathbb{R} \right\}$$
$$= \left\{ s \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} \middle| s, t \in \mathbb{R} \right\} = \text{span}\left(\begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -2 \\ 0 \\ 1 \\ 1 \end{pmatrix} \right). \quad (156)$$

in fact, the vectors above are linearly independent, so they form a basis for $\operatorname{null}(A)$*:*

$$\boxed{\left(\begin{pmatrix}1\\1\\0\\0\end{pmatrix},\begin{pmatrix}-2\\0\\1\\1\end{pmatrix}\right)}. \tag{157}$$

□

Theorem 5.27 *If A and B are row-equivalent matrices, then* $\operatorname{col}(A) = \operatorname{col}(B)$.

Theorem 5.28 *(Rank-nullity theorem) Let A be an $m \times n$ matrix. In that case,*

$$\begin{aligned} \dim(\operatorname{col}(A)) &= \operatorname{rank}(A) \\ \dim(\operatorname{row}(A)) &= \operatorname{rank}(A) \\ \dim(\operatorname{null}(A)) &= n - \operatorname{rank}(A) \end{aligned} \quad . \tag{158}$$

6 Linear transformations in $\mathbb{R}^n$

6.1 Linear transformations

Definition 6.1 *Let X and Y be sets. A function from X to Y is a rule which associates each element $x \in X$ to some element $f(x) \in Y$.*

Notation $f : X \to Y$.

In single variable calculus, you always dealt with functions $f : \mathbb{R} \to \mathbb{R}$, like the following:

$$\begin{array}{lll} f : \mathbb{R} \to \mathbb{R} & g : \mathbb{R} \to \mathbb{R} & h : \mathbb{R} \to \mathbb{R} \\ f(x) = x^2 & g(x) = \sin x & h(x) = e^{2x} \end{array}. \tag{159}$$

In multivariable calculus, you deal with more general notions of functions, sometimes considering $f : \mathbb{R}^n \to \mathbb{R}^m$, such as the following:

$$\begin{array}{ccc} f : \mathbb{R} \to \mathbb{R}^3 & g : \mathbb{R}^2 \to \mathbb{R} & h : \mathbb{R}^2 \to \mathbb{R}^3 \\ f(t) = \begin{pmatrix} t^2 - 1 \\ t + 1 \\ e^t \end{pmatrix} & g\begin{pmatrix} x \\ y \end{pmatrix} = x^2 + y^2 & h\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ y \\ \cos(xy) \end{pmatrix} \end{array}. \tag{160}$$

In general, though, a function can be defined between any two sets.

In this class, we will be particularly interested in a certain kind of function.

Definition 6.2 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a function. We say that T is a linear transformation or linear map provided that the following two conditions are true.*
(i) For any $\overrightarrow{u}, \overrightarrow{v} \in \mathbb{R}^n$, $T(\overrightarrow{u} + \overrightarrow{v}) = T(\overrightarrow{u}) + T(\overrightarrow{v})$. ("T preserves addition.")
(ii) For any $s \in \mathbb{R}$ and any $\overrightarrow{v} \in \mathbb{R}^n$, $T(s\overrightarrow{v}) = sT(\overrightarrow{v})$. ("T preserves scalar multiplication.")

Example 6.3 *The following functions are linear transformations.*

$$\begin{array}{ccc} f : \mathbb{R}^3 \to \mathbb{R}^3 & g : \mathbb{R}^3 \to \mathbb{R}^3 & h : \mathbb{R}^2 \to \mathbb{R} \\ f\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + y \\ x + y + z \\ 0 \end{pmatrix}, & g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + y \\ x - 2y \\ -x - y \end{pmatrix}, & h\begin{pmatrix} x \\ y \end{pmatrix} = x \end{array}. \tag{161}$$

*The following functions are **not** linear transformations.*

$$f : \mathbb{R} \to \mathbb{R} \qquad g : \mathbb{R}^3 \to \mathbb{R}^3 \qquad h : \mathbb{R} \to \mathbb{R}^2$$
$$f(x) = x^2, \quad g\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+1 \\ y+1 \\ z+1 \end{pmatrix}, \quad h(x) = \begin{pmatrix} \cos x \\ \sin x \end{pmatrix}. \tag{162}$$

6.2 The matrix of a linear transformation

Example 6.4 *Let*

$$A = \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix}. \tag{163}$$

Define a function $T_A : \mathbb{R}^3 \to \mathbb{R}^2$ via $T_A(\overrightarrow{v}) = A\overrightarrow{v}$. Determine whether T_A is a linear transformation.

First, let $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathbb{R}^3$ be arbitrary. In that case, for some appropriate scalars $x_1, y_1, z_1 \in \mathbb{R}$ and $x_2, y_2, z_2 \in \mathbb{R}$,

$$\overrightarrow{v_1} = \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix}. \tag{164}$$

Now,

$$
\begin{aligned}
T_A\left(\overrightarrow{v_1}+\overrightarrow{v_2}\right) = T_A \begin{pmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} &= \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} \\
&= \begin{pmatrix} 3\left(x_1+x_2\right)+\left(y_1+y_2\right)+4\left(z_1+z_2\right) \\ 2\left(x_1+x_2\right)+\left(z_1+z_2\right) \end{pmatrix} \\
&= \begin{pmatrix} 3x_1+y_1+4z_1 \\ 2x_1+z_1 \end{pmatrix} + \begin{pmatrix} 3x_2+y_2+4z_2 \\ 2x_2+z_2 \end{pmatrix} \\
&= \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} \\
&= T_A \begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + T_A \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = T_A\left(\overrightarrow{v_1}\right)+T_A\left(\overrightarrow{v_2}\right). \qquad (165)
\end{aligned}
$$

This shows that T_A preserves addition. Next, let $s \in \mathbb{R}$ and $\overrightarrow{v} \in \mathbb{R}^3$ be arbitrary. In that case, for some appropriate scalars $x, y, z \in \mathbb{R}$,

$$
\overrightarrow{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}. \qquad (166)
$$

Now,

$$T_A\left(s\overrightarrow{v}\right) = T_A\begin{pmatrix} sx \\ sy \\ sz \end{pmatrix} = \begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} sx \\ sy \\ sz \end{pmatrix} = \begin{pmatrix} 3sx + sy + 4sz \\ 2sx + sz \end{pmatrix}$$
$$= \begin{pmatrix} s\left(3x + y + 4z\right) \\ s\left(2x + z\right) \end{pmatrix} = s\begin{pmatrix} 3x + y + 4z \\ 2x + z \end{pmatrix}$$
$$= s\begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = sT_A\begin{pmatrix} x \\ y \\ z \end{pmatrix} = sT_A\left(\overrightarrow{v}\right). \quad (167)$$

This shows that T_A preserves scalar multiplication. Thus, T_A is a linear transformation. □

The previous example illustrates a more general fact about matrices and linear transformations.

Proposition 6.5 *Let A be an $m \times n$ matrix. Define a function $T_A : \mathbb{R}^n \to \mathbb{R}^m$ via $T_A\left(\overrightarrow{v}\right) = A\overrightarrow{v}$. The function T_A is a linear transformation.*

A more surprising and interesting fact is that every linear transformation actually corresponds to some matrix multiplication:

Theorem 6.6 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. There exists an $m \times n$ matrix A such that for all $\overrightarrow{v} \in \mathbb{R}^n$, $T\left(\overrightarrow{v}\right) = A\overrightarrow{v}$.*

To construct the matrix corresponding to a linear transformation, find the images of the standard basis vectors under T: $T\left(\overrightarrow{e_1}\right), T\left(\overrightarrow{e_2}\right), ..., T\left(\overrightarrow{e_n}\right)$. Construct a matrix whose columns are these vectors.

Example 6.7 *Find the matrix corresponding to the linear map $T : \mathbb{R}^3 \to \mathbb{R}^2$ defined by*

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x + 2y + 3z \\ 2y - 3x + z \end{pmatrix}. \quad (168)$$

To do this, we notice that

$$T\begin{pmatrix}1\\0\\0\end{pmatrix}=\begin{pmatrix}1\\-3\end{pmatrix},\quad T\begin{pmatrix}0\\1\\0\end{pmatrix}=\begin{pmatrix}2\\2\end{pmatrix},\quad T\begin{pmatrix}0\\0\\1\end{pmatrix}=\begin{pmatrix}3\\1\end{pmatrix}. \tag{169}$$

The matrix corresponding to this linear transformation is:

$$\boxed{M_T=\begin{pmatrix}1&2&3\\-3&2&1\end{pmatrix}}. \tag{170}$$

To check our work, we examine the multiplication of this matrix by an arbitrary vector in $\mathbb{R}^3$:

$$M_T\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}1&2&3\\-3&2&1\end{pmatrix}\begin{pmatrix}x\\y\\z\end{pmatrix}=\begin{pmatrix}x+2y+3z\\-3x+2y+z\end{pmatrix}=T\begin{pmatrix}x\\y\\z\end{pmatrix} \tag{171}$$

□

6.4 Properties of linear transformations

Definition 6.8 *Let $f: X\to Y$ and $g: Y\to Z$ be functions. The composition of f and g is the function $g\circ f: X\to Z$ defined such that $(g\circ f)(x)=g(f(x))$.*

Suppose that $T_1:\mathbb{R}^k\to\mathbb{R}^n$ and $T_2:\mathbb{R}^n\to\mathbb{R}^m$ are linear transformations. In that case, they correspond to some matrices; call these M_1 and M_2, respectively. We should notice that for any vector $\overrightarrow{v}\in\mathbb{R}^k$,

$$(T_2\circ T_1)(\overrightarrow{v})=T_2(T_1(\overrightarrow{v}))=T_2(M_1\overrightarrow{v})=M_2(M_1\overrightarrow{v})=(M_2M_1)\overrightarrow{v}. \tag{172}$$

Thus, we have the following theorem.

Theorem 6.9 *Let $T_1:\mathbb{R}^k\to\mathbb{R}^n$ and $T_2:\mathbb{R}^n\to\mathbb{R}^m$ be linear transformations with matrix representations M_1 and M_2, respectively. In that case, $T_2\circ T_1$ is also a linear transformation, and its matrix representation is the product M_2M_1.*

Definition 6.10 *Let $f : X \to Y$ and $g : Y \to X$ be functions. We say that* *f and g are inverse functions* *provided that for all $x \in X$, $g(f(x)) = x$ and for all $y \in Y$, $f(g(y)) = y$.*

Example 6.11 *The following pair of functions are inverse functions.*

$$\begin{array}{ll} f : \mathbb{R} \to \mathbb{R} & g : \mathbb{R} \to \mathbb{R} \\ f(x) = e^x, & g(x) = \ln x \end{array}. \tag{173}$$

□

Theorem 6.12 *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation, and let M be the matrix representation of T. The function T has an inverse function if and only if M is an invertible matrix. In that case, the matrix representation of the inverse function is M^{-1}.*

9 Vector spaces

9.1 Definition of vector spaces

Definition 9.1 *The field of complex numbers is the set*

$$\mathbb{C} = \left\{ a + bi \middle| a, b \in \mathbb{R} \right\}, \tag{174}$$

where $i^2 = -1$.

Definition 9.2 *Let* K *be a "field." (For this class, usually* $K = \mathbb{R}$ *or* $K = \mathbb{C}$*.) A vector space over* K *is a nonempty set* V *together with two operations called* $+$ *and* $\cdot$ *satisfying the following conditions.*
(i) For any $a, b \in V$, $a + b \in V$*. ("*V *is closed under addition.")*
(ii) For any $a, b, c \in V$,

$$a + (b + c) = (a + b) + c. \tag{175}$$

("Addition is associative.")
(iii) There exists an element $0 \in V$ *such that for any* $a \in V$,

$$a + 0 = 0 + a = a. \tag{176}$$

("V *contains an additive identity.")*
(iv) For any $a \in V$*, there exists a* $b \in V$ *such that*

$$a + b = b + a = 0. \tag{177}$$

("Additive inverses exist in V*.")*
(v) For any $a, b \in V$,

$$a + b = b + a. \tag{178}$$

("Addition is commutative.")
(vi) For any $s \in K$ *and* $a \in V$, $s \cdot a \in V$*. ("*V *is closed under scalar multiplication.")*

(vii) For any $r, s \in K$ *and* $a \in V$*,*

$$r \cdot (s \cdot a) = (rs) \cdot a. \tag{179}$$

("Scalar multiplication is associative.")

(viii) For any $s \in K$ *and* $a, b \in V$*,*

$$s \cdot (a + b) = (s \cdot a) + (s \cdot b). \tag{180}$$

("Scalar multiplication distributes over vector addition.")

(ix) For any $r, s \in K$ *and* $a \in V$*,*

$$(r + s) \cdot a = (r \cdot a) + (s \cdot a). \tag{181}$$

("Scalar multiplication distributes over scalar addition.")

(x) For any $a \in V$*,*

$$1 \cdot a = a. \tag{182}$$

Example 9.3 *The following sets are vector spaces over* $\mathbb{R}$*.*

1. Given any positive integer n*,*

$$\mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \middle| x_1, x_2, ..., x_n \in \mathbb{R} \right\}. \tag{183}$$

2. Given any positive integer n*,*

$$P_n(\mathbb{R}) = \left\{ a_0 + a_1 x + a_2 x^2 + ... + a_n x^n \middle| a_0, a_1, a_2, ..., a_n \in \mathbb{R} \right\}. \tag{184}$$

3.

$$\mathbb{R}[X] = \left\{ f : \mathbb{R} \to \mathbb{R} \middle| f \text{ is a polynomial with real coefficients} \right\} \tag{185}$$

4. Given any positive integers m and n,

$$M_{m\times n}(\mathbb{R}) = \left\{ \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \middle| a_{11}, a_{12}, \dots, a_{mn} \in \mathbb{R} \right\}. \tag{186}$$

5. The set

$$\operatorname{Func}(\mathbb{R}, \mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \middle| f \textit{ is a function} \right\} \tag{187}$$

is a vector space.

6. Given any non-negative integer n,

$$\mathcal{C}^n(\mathbb{R}, \mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \middle| f^{(n)} \textit{ is continuous} \right\}. \tag{188}$$

7.

$$\mathcal{C}^\infty(\mathbb{R}, \mathbb{R}) = \left\{ f : \mathbb{R} \to \mathbb{R} \middle| \textit{for every non-negative integer } n,\ f^{(n)} \textit{ is continuous} \right\}. \tag{189}$$

8.

$$\operatorname{Seq}(\mathbb{R}) = \left\{ (a_1, a_2, a_3, \dots) \middle| a_1, a_2, a_3, \dots \in \mathbb{R} \right\}. \tag{190}$$

9.2 Linear combinations, span, and linear independence

Definition 9.4 *Let V be a vector space over a field K. Let $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_k} \in V$. A vector $\overrightarrow{u} \in V$ is a linear combination of $\overrightarrow{v_1}, \overrightarrow{v_2}, \dots, \overrightarrow{v_k}$ provided that there exist scalars $s_1, s_2, \dots, s_k \in K$ such that*

$$\overrightarrow{u} = s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + \dots + s_k\overrightarrow{v_k}. \tag{191}$$

Example 9.5 *Determine whether the matrix*

$$A = \begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} \tag{192}$$

can be written as a linear combination of the matrices

$$\overrightarrow{v_1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \overrightarrow{v_3} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \overrightarrow{v_4} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \tag{193}$$

over $\mathbb{R}$.

We seek scalars $s_1, s_2, s_3, s_4 \in \mathbb{R}$ *such that*

$$\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = s_1 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + s_2 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + s_3 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + s_4 \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} s_1 + s_2 & s_3 - s_4 \\ s_3 + s_4 & s_1 - s_2 \end{pmatrix} \tag{194}$$

This gives us a system of four linear equations:

$$\begin{aligned} s_1 + s_2 &= 1 \\ s_3 - s_4 &= 3 \\ s_3 + s_4 &= -1 \\ s_1 - s_2 &= 2 \end{aligned}. \tag{195}$$

We set up an augmented matrix corresponding to the system:

$$\begin{pmatrix} 1 & 1 & 0 & 0 & | & 1 \\ 0 & 0 & 1 & -1 & | & 3 \\ 0 & 0 & 1 & 1 & | & -1 \\ 1 & -1 & 0 & 0 & | & 2 \end{pmatrix}. \tag{196}$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & -1 & | & 3 \\ 0 & 0 & \boxed{1} & 1 & | & -1 \\ \boxed{1} & -1 & 0 & 0 & | & 2 \end{pmatrix} \xrightarrow{R_4 - R_1 \to R_4} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & -1 & | & 3 \\ 0 & 0 & \boxed{1} & 1 & | & -1 \\ 0 & \boxed{-2} & 0 & 0 & | & 1 \end{pmatrix}$$

$$\xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & -1 & | & 3 \\ 0 & 0 & 0 & \boxed{2} & | & -4 \\ 0 & \boxed{-2} & 0 & 0 & | & 1 \end{pmatrix}$$

$$\xrightarrow{\frac{1}{2}R_3 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & -1 & | & 3 \\ 0 & 0 & 0 & \boxed{1} & | & -2 \\ 0 & \boxed{-2} & 0 & 0 & | & 1 \end{pmatrix}$$

$$\xrightarrow{R_2 + R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & | & -2 \\ 0 & \boxed{-2} & 0 & 0 & | & 1 \end{pmatrix}$$

$$\xrightarrow{-\frac{1}{2}R_4 \to R_4} \begin{pmatrix} \boxed{1} & 1 & 0 & 0 & | & 1 \\ 0 & 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & | & -2 \\ 0 & \boxed{1} & 0 & 0 & | & -\frac{1}{2} \end{pmatrix}$$

$$\xrightarrow{R_1 - R_4 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & | & \frac{3}{2} \\ 0 & 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & | & -2 \\ 0 & \boxed{1} & 0 & 0 & | & -\frac{1}{2} \end{pmatrix} \xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & | & \frac{3}{2} \\ 0 & 0 & 0 & \boxed{1} & | & -2 \\ 0 & 0 & \boxed{1} & 0 & | & 1 \\ 0 & \boxed{1} & 0 & 0 & | & -\frac{1}{2} \end{pmatrix}$$

$$\xrightarrow{R_2 \leftrightarrow R_4} \begin{pmatrix} \boxed{1} & 0 & 0 & 0 & | & \frac{3}{2} \\ 0 & \boxed{1} & 0 & 0 & | & -\frac{1}{2} \\ 0 & 0 & \boxed{1} & 0 & | & 1 \\ 0 & 0 & 0 & \boxed{1} & | & -2 \end{pmatrix} \quad (197)$$

From this we deduce that a single solution exists: $s_1 = \frac{3}{2}$, $s_2 = -\frac{1}{2}$, $s_3 = 1$ *and* $s_4 = -2$. *Thus,*

$$\boxed{\begin{pmatrix} 1 & 3 \\ -1 & 2 \end{pmatrix} = \frac{3}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} + 1\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} - 2\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}. \tag{198}$$

□

Example 9.6 *Determine whether* $p\left(x\right) = 7x^2 + 4x - 3$ *can be written as a linear combination of*

$$q_1\left(x\right) = x^2, \quad q_2\left(x\right) = \left(x+1\right)^2, \quad q_3\left(x\right) = \left(x+2\right)^2. \tag{199}$$

We seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ *which satisfy the equation*

$$s_1 q_1 + s_2 q_2 + s_3 q_3 = p. \tag{200}$$

In other words, we need scalars such that

$$s_1\left(x^2\right) + s_2(x+1)^2 + s_3(x+2)^2 = 7x^2 + 4x - 3 \tag{201}$$

for all values of x. *To find them, we first distribute:*

$$s_1 x^2 + s_2 x^2 + 2s_2 x + s_2 + s_3 x^2 + 4s_3 x + 4s_3 = 7x^2 + 4x - 3. \tag{202}$$

Now collect like terms in powers of x *to get*

$$\left(s_1 + s_2 + s_3\right)x^2 + \left(2s_2 + 4s_3\right)x + \left(s_2 + 4s_3\right) = 7x^2 + 4x - 3. \tag{203}$$

We have that these two polynomials must be equal for all values of x. *Therefore, they must be the same polynomial, with the same coefficients. This gives us the* *following system of equations:*

$$\begin{aligned} s_1 + s_2 + s_3 &= 7 \\ 2s_2 + 4s_3 &= 4 \\ s_2 + 4s_3 &= -3 \end{aligned} \quad . \tag{204}$$

We convert this to an augmented matrix:

$$\begin{pmatrix} 1 & 1 & 1 & | & 7 \\ 0 & 2 & 4 & | & 4 \\ 0 & 1 & 4 & | & -3 \end{pmatrix}, \tag{205}$$

and proceed with Gauss-Jordan elimination:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 1 & 1 & | & 7 \\ 0 & \boxed{2} & 4 & | & 4 \\ 0 & \boxed{1} & 4 & | & -3 \end{pmatrix} \xrightarrow{R_2 - R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 7 \\ 0 & \boxed{1} & 0 & | & 7 \\ 0 & \boxed{1} & 4 & | & -3 \end{pmatrix} \\ &\xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 7 \\ 0 & \boxed{1} & 0 & | & 7 \\ 0 & 0 & \boxed{4} & | & -10 \end{pmatrix} \xrightarrow{\frac{1}{4}R_3 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 7 \\ 0 & \boxed{1} & 0 & | & 7 \\ 0 & 0 & \boxed{1} & | & -\frac{5}{2} \end{pmatrix} \\ &\xrightarrow{R_1 - R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ 0 & \boxed{1} & 0 & | & 7 \\ 0 & 0 & \boxed{1} & | & -\frac{5}{2} \end{pmatrix} \xrightarrow{R_1 - R_3 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & | & \frac{5}{2} \\ 0 & \boxed{1} & 0 & | & 7 \\ 0 & 0 & \boxed{1} & | & -\frac{5}{2} \end{pmatrix}. \end{aligned} \tag{206}$$

This reveals that $s_1 = \frac{5}{2}$, $s_2 = 7$ *and* $s_3 = -\frac{5}{2}$. *Thus,*

$$\boxed{7x^2 + 4x - 3 = \frac{5}{2}x^2 + 7(x+1)^2 - \frac{5}{2}(x+2)^2}. \tag{207}$$

□

Definition 9.7 *Let* V *be a vector space over a field* K. *Given a subset* $S \subseteq V$, *the* *span of S is the set of linear combinations of elements of S:*

$$\operatorname{span}(S) = \left\{ s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} \middle| s_1, s_2, ..., s_+k \in K \textit{ and } \overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k} \in S \right\}. \tag{208}$$

Definition 9.8 *Let V be a vector space over a field K. A finite sequence of vectors $(\overrightarrow{v_1}, \overrightarrow{v_2}, ..., \overrightarrow{v_k})$ in V is linearly independent (over K) provided that the equation*

$$s_1\overrightarrow{v_1} + s_2\overrightarrow{v_2} + ... + s_k\overrightarrow{v_k} = 0 \tag{209}$$

has only the trivial solution: $s_1 = s_2 = ... = s_k = 0$. An infinite sequence of vectors is called linearly independent (over $\mathbb{R}$) provided that every finite subsequence is linearly independent.

Example 9.9 *Determine whether the following polynomials are linearly independent over $\mathbb{R}$:*

$$p_1(x) = x^2, \quad p_2(x) = x^2 + 2x - 1, \quad p_3(x) = 2x^2 - x + 3. \tag{210}$$

We seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ satisfying the equation

$$s_1p_1 + s_2p_2 + s_3p_3 = 0, \tag{211}$$

or in other words,

$$s_1x^2 + s_2\left(x^2 + 2x - 1\right) + s_3\left(2x^2 - x + 3\right) = 0 \tag{212}$$

for all values of x. We distribute:

$$s_1x^2 + s_2x^2 + 2s_2x - s_2 + 2s_3x^2 - s_3x + 3s_3 = 0, \tag{213}$$

and collect like terms in powers of x:

$$(s_1 + s_2 + 2s_3)\,x^2 + (2s_2 - s_3)\,x + (3s_3 - s_2) = 0x^2 + 0x + 0. \tag{214}$$

Since these two polynomials are equal for all values of x, they must be the same *polynomial, with the same coefficients. This gives us the system of equations*

$$\begin{aligned} s_1 + s_2 + 2s_3 &= 0 \\ 2s_2 - s_3 &= 0 \\ -s_2 + 3s_3 &= 0 \end{aligned} \quad . \tag{215}$$

We convert to an augmented matrix:

$$\begin{pmatrix} 1 & 1 & 2 & | & 0 \\ 0 & 2 & -1 & | & 0 \\ 0 & -1 & 3 & | & 0 \end{pmatrix}, \tag{216}$$

and proceed with Gauss-Jordan elimination:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & \boxed{2} & -1 & | & 0 \\ 0 & \boxed{-1} & 3 & | & 0 \end{pmatrix} \xrightarrow{R_2+2R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & 0 & \boxed{5} & | & 0 \\ 0 & \boxed{-1} & 3 & | & 0 \end{pmatrix} \\ &\xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & \boxed{-1} & 3 & | & 0 \\ 0 & 0 & \boxed{5} & | & 0 \end{pmatrix} \xrightarrow{\frac{1}{5}R_3 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & \boxed{-1} & 3 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \\ &\xrightarrow{-R_2 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & \boxed{1} & -3 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \xrightarrow{R_2+3R_3 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 2 & | & 0 \\ 0 & \boxed{1} & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \\ &\xrightarrow{R_1-R_2 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 2 & | & 0 \\ 0 & \boxed{1} & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \xrightarrow{R_1-2R_3 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 0 \\ 0 & \boxed{1} & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix}. \end{aligned} \tag{217}$$

This gives us the unique solution $s_1 = s_2 = s_3 = 0$. By definition of linear independence, this shows that p_1, p_2 and p_3 are $\boxed{\textit{linearly independent}}$. □

Example 9.10 *In the vector space* $\mathrm{Func}(\mathbb{R}, \mathbb{R})$*, determine whether the following vectors are linearly independent:*

$$f(x) = \sin x \quad g(x) = \cos x. \tag{218}$$

We seek scalars $s_1, s_2 \in \mathbb{R}$ such that

$$s_1 f + s_2 g = 0, \tag{219}$$

or in other words,

$$s_1 \sin x + s_2 \cos x = 0 \tag{220}$$

for all values of x. In particular, if $x = \frac{\pi}{2}$, then the equation must hold:

$$\begin{array}{c} s_1 \sin\left(\frac{\pi}{2}\right) + s_2 \cos\left(\frac{\pi}{2}\right) = 0 \\ s_1 (1) + s_2 (0) = 0 \\ s_1 = 0 \end{array} . \tag{221}$$

On the other hand, if $x = 0$, then the equation must hold:

$$\begin{array}{c} s_1 \sin(0) + s_2 \cos(0) = 0 \\ s_1 (0) + s_2 (1) = 0 \\ s_2 = 0 \end{array} . \tag{222}$$

In this way, we've shown that both $s_1 = s_2 = 0$. By definition, f and g are $\boxed{\text{linearly independent}}$. □

9.3 Subspaces

Definition 9.11 *Let V be a vector space over a field K, and let $W \subseteq V$. We say that W is a vector subspace of V provided that the following statements are true.*
(i) $\overrightarrow{0} \in W$.
(ii) For any $\overrightarrow{u}, \overrightarrow{v} \in W$, $\overrightarrow{u} + \overrightarrow{v} \in W$.
(iii) For any $s \in K$ and $\overrightarrow{v} \in W$, $s\overrightarrow{v} \in W$.

Example 9.12 *The following are examples of vector subspaces.*
1. Given any non-negative integer n, $C^n(\mathbb{R}, \mathbb{R})$ is a vector subspace of $\operatorname{Func}(\mathbb{R}, \mathbb{R})$. Further, $C^\infty(\mathbb{R}, \mathbb{R})$ is a vector subspace of $C^n(\mathbb{R}, \mathbb{R})$.

2. The set

$$S = \left\{ \begin{pmatrix} a & b \\ 0 & d \end{pmatrix} \middle| a, b, d \in \mathbb{R} \right\} \tag{223}$$

is a vector subspace of $M_{2\times 2}(\mathbb{R})$.

3. Any vector space is a vector subspace of itself.

4. Given a vector space V *whose additive identity is denoted* $\overrightarrow{0}$, *the set* $\left\{\overrightarrow{0}\right\}$ *is a vector subspace of* V.

9.4 Basis and dimension

Definition 9.13 *Let* V *be a vector space over a field* K, *and let* $B = (\overrightarrow{v_1}, \overrightarrow{v_2}, \overrightarrow{v_3}, ...)$ *be a sequence of vectors in* V. *We say that* B *is a* basis for V *provided that the following statements are true.*

(i) $\mathrm{span}(B) = V$.

(ii) B *is linearly independent.*

Theorem 9.14 *Given any vector space* V *over a field* K, *there exists a basis for* V.

Theorem 9.15 *Let* V *be a vector space over a field* K. *Given bases* B_1 *and* B_2 *of* V, B_1 *and* B_2 *have the same number of elements.*

Definition 9.16 *Let* V *be a vector space over a field* K. *The* dimension of V *is the number of elements in a basis of* V.

10 Linear transformations of vector spaces

10.1 Definition and examples

Definition 10.1 *Let V and W be vector spaces over a field of scalars K. A function $T : V \to W$ is called a* <u>*linear transformation from V to W*</u> *provided that the following two statements are true.*
(i) For all $v_1, v_2 \in V$, $T(v_1 + v_2) = T(v_1) + T(v_2)$.
(ii) For all $v \in V$ and for all $s \in K$, $sT(v) = T(sv)$.
In the case that $V = W$, we say that T is a <u>*linear operator on V*</u>*.*

Examples (and/or nonexamples) of linear transformations can be found on your Homework 03.

10.3 Linear transformations defined on a basis

Proposition 10.2 *Let V and W be vector spaces over a field K, and suppose that $T_1, T_2 : V \to W$ are linear transformations. Given a set $S \subseteq V$, suppose that* $\operatorname{span}(S) = V$. *If, for all $v \in S$, $T_1(v) = T_2(v)$, then $T_1 = T_2$.*

Theorem 10.3 *Let V and W be vector spaces over a field K. Let $(v_1, v_2, ..., v_n)$ be a basis for V. Given any vectors $w_1, w_2, ..., w_n \in W$, there exists a unique linear transformation $T : V \to W$ such that $T(v_i) = w_i$ for each $i \in \{1, 2, ..., n\}$.*

To summarize, every linear transformation is completely defined by what it does to basis elements.

10.4 The matrix of a linear transformation

Just as with Euclidean spaces, every linear transformation can be fully described by a matrix, provided that some bases are chosen for the two vector spaces in question.

Proposition 10.4 *Let V and W be finite-dimensional vector spaces over a field K. Suppose that B is a basis for V and C is a basis for W. Given a linear* *transformation $T : V \to W$, there exists a unique matrix A such that for any $v \in V$, $A[v]_B = [T(v)]_C$.*

To write the matrix representation of a linear transformation, find $T(v_i)$ for each i, and write its coordinates in terms of C. This will become the ith column of the matrix representation.

Example 10.5 *Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be the linear transformation defined via*

$$T\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} a+b \\ b \\ a \end{pmatrix}. \tag{224}$$

Consider the bases

$$\begin{gathered} B = \left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \end{pmatrix} \right) \\ C = \left(\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right) \end{gathered}. \tag{225}$$

Find the matrix representation of T with respect to the bases B and C.

First, we need to find the values of T on the elements of B:

$$T\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad T\begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}. \tag{226}$$

Next, we need to find the coordinates of these vectors with respect to the basis C. In other words, we seek scalars $s_1, s_2, s_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = s_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + s_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + s_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{227}$$

and scalars $t_1, t_2, t_3 \in \mathbb{R}$ such that

$$\begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = t_1 \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} + t_3 \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}. \tag{228}$$

For the first problem, we can take $s_1 = s_2 = 0$ and $s_3 = 1$. Thus, we get the coordinate representation

$$\left[T \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right]_C = \left[\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right]_C = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}. \tag{229}$$

The second problem is less obvious; for this we set up an augmented matrix

$$\begin{pmatrix} 1 & 0 & 1 & | & 0 \\ 1 & 1 & 0 & | & 1 \\ 0 & 1 & 1 & | & -1 \end{pmatrix} \tag{230}$$

and proceed with Gauss-Jordan elimination:

$$\begin{aligned} &\begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ \boxed{1} & 1 & 0 & | & 1 \\ 0 & \boxed{1} & 1 & | & -1 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ 0 & \boxed{1} & -1 & | & 1 \\ 0 & \boxed{1} & 1 & | & -1 \end{pmatrix} \\ &\xrightarrow{R_3 - R_2 \to R_3} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ 0 & \boxed{1} & -1 & | & 1 \\ 0 & 0 & \boxed{2} & | & -2 \end{pmatrix} \xrightarrow{\frac{1}{2} R_3 \to R_3} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ 0 & \boxed{1} & -1 & | & 1 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix} \\ &\xrightarrow{R_2 + R_3 \to R_2} \begin{pmatrix} \boxed{1} & 0 & 1 & | & 0 \\ 0 & \boxed{1} & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix} \xrightarrow{R_1 - R_3 \to R_1} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 1 \\ 0 & \boxed{1} & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & -1 \end{pmatrix} \end{aligned} \tag{231}$$

This gives us $t_1 = 1$, $t_2 = 0$, and $t_3 = -1$. Thus,

$$\left[T\begin{pmatrix}-1\\1\end{pmatrix}\right]_C = \left[\begin{pmatrix}0\\1\\-1\end{pmatrix}\right]_C = \begin{pmatrix}1\\0\\-1\end{pmatrix}. \tag{232}$$

Now the matrix representation of T with respect to the bases B and C is

$$\boxed{M_T = \begin{pmatrix}0 & 1\\0 & 0\\1 & -1\end{pmatrix}}. \tag{233}$$

□

Example 10.6 *Let $T : P_3(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ be the linear transformation defined via*

$$T\left(ax^3 + bx^2 + cx + d\right) = \begin{pmatrix}a+d & b-c\\b+c & a-d\end{pmatrix}. \tag{234}$$

Consider the bases $B = (1, x, x^2, x^3)$ and

$$C = \left(\begin{pmatrix}1 & 0\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 1\\0 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\1 & 0\end{pmatrix}, \begin{pmatrix}0 & 0\\0 & 1\end{pmatrix}\right) \tag{235}$$

of $P_3(\mathbb{R})$ and $M_{2\times 2}(\mathbb{R})$, respectively. Find the matrix representation of T with respect to the bases B and C.

First, we need to know what T does to the elements of B:

$$\begin{aligned} T(1) &= \begin{pmatrix}1 & 0\\0 & -1\end{pmatrix}, \quad T(x) = \begin{pmatrix}0 & -1\\1 & 0\end{pmatrix},\\ T(x^2) &= \begin{pmatrix}0 & 1\\1 & 0\end{pmatrix}, \quad T(x^3) = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} \end{aligned}. \tag{236}$$

Next, we need to write each of these as a linear combination of the elements of the *basis C. In this case, the basis C is simple enough that this is easy to do:*

$$\begin{array}{c} T(1) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} = 1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} - 1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ T(x) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} - 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ T(x^2) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = 0 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \\ T(x^3) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + 0 \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + 1 \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \end{array}. \tag{237}$$

Now the coordinates are

$$\begin{array}{cc} [T(1)]_C = \begin{pmatrix} 1 \\ 0 \\ 0 \\ -1 \end{pmatrix}, & [T(x)]_C = \begin{pmatrix} 0 \\ -1 \\ 1 \\ 0 \end{pmatrix}, \\ [T(x^2)]_C = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 0 \end{pmatrix}, & [T(x^3)]_C = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \end{array}. \tag{238}$$

Thus, the matrix representation of T with respect to the bases B and C is

$$\boxed{M_T = \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & -1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix}}. \tag{239}$$

□

Example 10.7 *Let*

$$M = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}, \tag{240}$$

and consider the linear operator $T : M_{2\times 2}(\mathbb{R}) \to M_{2\times 2}(\mathbb{R})$ defined via the relationship $T(A) = MAM$. Find the matrix representation of T with respect to the following basis:

$$B = \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right). \tag{241}$$

As usual, we need to find what T does to the elements of B:

$$\begin{gathered} T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & -4 \end{pmatrix} \\ T\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -2 & 1 \\ 4 & -2 \end{pmatrix} \\ T\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} \\ T\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} = \begin{pmatrix} -4 & 2 \\ -2 & 1 \end{pmatrix} \end{gathered}. \tag{242}$$

Next, we need to write the coordinates of each of these with respect to the basis B:

$$\begin{gathered} \left[T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]_C = \begin{pmatrix} 1 \\ 2 \\ -2 \\ -4 \end{pmatrix} \quad \left[T\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}\right]_C = \begin{pmatrix} -2 \\ 1 \\ 4 \\ -2 \end{pmatrix} \\ \left[T\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}\right]_C = \begin{pmatrix} 2 \\ 4 \\ 1 \\ 2 \end{pmatrix} \quad \left[T\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right]_C = \begin{pmatrix} -4 \\ 2 \\ -2 \\ 1 \end{pmatrix} \end{gathered}. \tag{243}$$

Now the matrix representation of the linear transformation with respect to the basis

B is

$$\boxed{M_T = \begin{pmatrix} 1 & -2 & 2 & -4 \\ 2 & 1 & 4 & 2 \\ -2 & 4 & 1 & -2 \\ -4 & -2 & 2 & 1 \end{pmatrix}}. \tag{244}$$

$\square$

Example 10.8 *Let $D : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}^2$ be the linear transformation defined via*

$$D\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix}. \tag{245}$$

Consider the bases

$$\begin{aligned} B &= \left(\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \right) \\ C &= \left(\begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right) \end{aligned} \tag{246}$$

of $M_{2\times 2}(\mathbb{R})$ and $\mathbb{R}^2$, respectively. Find the matrix representation of D with respect to the bases B and C.

Again, we need to know what D does to each element of B:

$$\begin{aligned} D\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \\ D\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ D\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \\ D\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} &= \begin{pmatrix} 1 \\ 1 \end{pmatrix} \end{aligned}. \tag{247}$$

Next, we write these vectors as coordinates with respect to the basis C:

$$\begin{aligned} \left[D\begin{pmatrix}1 & 0\\ 0 & 0\end{pmatrix}\right]_C &= \left[\begin{pmatrix}1\\ 0\end{pmatrix}\right]_C = \begin{pmatrix}-1\\ 1\end{pmatrix}, \\ \left[D\begin{pmatrix}1 & 1\\ 0 & 0\end{pmatrix}\right]_C &= \left[\begin{pmatrix}1\\ 0\end{pmatrix}\right]_C = \begin{pmatrix}-1\\ 1\end{pmatrix} \\ \left[D\begin{pmatrix}1 & 1\\ 1 & 0\end{pmatrix}\right]_C &= \left[\begin{pmatrix}1\\ 1\end{pmatrix}\right]_C = \begin{pmatrix}0\\ 1\end{pmatrix} \\ \left[D\begin{pmatrix}1 & 1\\ 1 & 1\end{pmatrix}\right]_C &= \left[\begin{pmatrix}1\\ 1\end{pmatrix}\right]_C = \begin{pmatrix}0\\ 1\end{pmatrix} \end{aligned}. \tag{248}$$

The matrix representation is constructed by putting these vectors as columns in a matrix:

$$\boxed{M_D = \begin{pmatrix}-1 & -1 & 0 & 0\\ 1 & 1 & 1 & 1\end{pmatrix}}. \tag{249}$$

□

7 Determinants

We now turn our attention to figuring out under what conditions a matrix (and therefore, a linear transformation) is invertible. To this end, we define a value called the "determinant" of a matrix. Since no non-square matrix can have a two-sided inverse matrix, we will restrict our attention to square matrices in this chapter.

7.1 Determinants of 2×2- and 3×3-matrices

Definition 7.1 *Let*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{250}$$

be a 2×2 *matrix. The* *determinant of* A *is the value* $\det(A) = ad - bc$.

Example 7.2 *Find the determinant of*

$$A = \begin{pmatrix} 1 & 7 \\ 2 & 6 \end{pmatrix}. \tag{251}$$

Directly from the definition,

$$\det(A) = (1)(6) - (7)(2) = 6 - 14 = \boxed{-8}. \tag{252}$$

□

7.2 Minors and cofactors

Determinants of larger matrices can all be expressed in terms of determinants of small matrices.

Definition 7.3 *Let* A *be an* $n \times n$ *matrix. Given* $i, j \in \{1, 2, ..., n\}$, *the* *$(i, j)$-minor of* A *is the determinant* M_{ij} *of the* $(n-1) \times (n-1)$*-submatrix obtained by deleting the* i*th row and* j*th column of* A.

Example 7.4 *Let*

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix}. \tag{253}$$

Determine the minors M_{12} and M_{33}.

By definition,

$$M_{12} = \begin{vmatrix} 0 & 2 \\ 3 & 1 \end{vmatrix} = \boxed{-6} \tag{254}$$

and

$$M_{33} = \begin{vmatrix} 1 & 2 \\ 0 & 0 \end{vmatrix} = \boxed{0}. \tag{255}$$

□

Definition 7.5 *Let A be an $n \times n$ matrix. Given $i, j \in \{1, 2, ..., n\}$, the* <u>*$(i, j)$-cofactor of A*</u> *is the value*

$$C_{ij} = (-1)^{i+j} M_{ij}, \tag{256}$$

where M_{ij} is the (i, j)-minor of A.

Example 7.6 *Let*

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix}. \tag{257}$$

Determine the cofactors C_{12} and C_{33}.

This is the same matrix as in the previous example. Thus,

$$C_{12} = (-1)^{1+2} M_{12} = (-1)(-6) = \boxed{6} \tag{258}$$

and

$$C_{33} = (-1)^{3+3} M_{33} = (1)(0) = \boxed{0}. \tag{259}$$

□

Definition 7.7 *Let A be the following $n \times n$ matrix:*

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix}. \tag{260}$$

(i) Given $i \in \{1, 2, ..., n\}$, the *cofactor expansion of A along row i* *is the value*

$$D_i = a_{i1}C_{i1} + a_{i2}C_{i2} + ... + a_{in}C_{in}. \tag{261}$$

(ii) Given $j \in \{1, 2, ..., n\}$, the *cofactor expansion of A along row j* *is the value*

$$E_j = a_{1j}C_{1j} + a_{2j}C_{2j} + ... + a_{nj}C_{nj}. \tag{262}$$

Example 7.8 *Let*

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 0 & 2 \\ 3 & 2 & 1 \end{pmatrix}. \tag{263}$$

Find the cofactor expansions D_2 and E_1.

By definition of the cofactor expansion along row 2,

$$D_2 = 0C_{21} + 0C_{22} + 2C_{23}. \tag{264}$$

Now,

$$C_{23} = (-1)^{2+3} M_{23} = - \begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} = 4. \tag{265}$$

Thus,

$$D_2 = 2C_{23} = 2\,(4) = \boxed{8}. \tag{266}$$

By definition of cofactor expansion along column 1,

$$E_1 = 1C_{11} + 0C_{21} + 3C_{31}. \tag{267}$$

Now,

$$\begin{aligned} C_{11} = (-1)^{1+1} M_{11} = \begin{pmatrix} 0 & 2 \\ 2 & 1 \end{pmatrix} = -4 \\ C_{31} = (-1)^{3+1} M_{31} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 4 \end{aligned}. \tag{268}$$

Thus,

$$E_1 = 1\,(-4) + 3\,(4) = \boxed{8}. \tag{269}$$

□

The fact that the cofactor expansions in the previous example are equal is not unusual. In fact, this is always the case.

Theorem 7.9 *Let A be an $n \times n$ matrix. Every cofactor expansion of A along a given row is equal to every cofactor expansion of A along a given column.*

Definition 7.10 *Let A be an $n \times n$ matrix. The* <u>*determinant of A*</u> *is the cofactor expansion of A along any row or column.*

Example 7.11 *Find the determinant of the following matrix:*

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 2 & 1 & -1 & 0 \\ 1 & 2 & 0 & 1 \\ -1 & 0 & 2 & 1 \end{pmatrix}. \tag{270}$$

We compute $\det(A)$ *by doing a cofactor expansion along the first row:*

$$\det(A) = 1C_{11} + 0C_{12} + 0C_{13} + 0C_{14} = C_{11}. \tag{271}$$

Now,

$$C_{11} = (-1)^{1+1} M_{11} = \begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \end{vmatrix}. \tag{272}$$

This reduces the problem to finding the determinant of the above 3×3 matrix. In turn, we can do cofactor expansion along the third column of the above matrix:

$$\begin{vmatrix} 1 & -1 & 0 \\ 2 & 0 & 1 \\ 0 & 2 & 1 \end{vmatrix} = 0C_{13} + 1C_{23} + 1C_{33}$$
$$= (-1)^{2+3} M_{23} + (-1)^{3+3} M_{33} = - \begin{vmatrix} 1 & -1 \\ 0 & 2 \end{vmatrix} + \begin{vmatrix} 1 & -1 \\ 2 & 0 \end{vmatrix}$$
$$= -(2) + (2) = \boxed{0}. \quad (273)$$

□

7.3 The determinant of a triangular matrix

Definition 7.12 *Let A be the following $n \times n$ matrix:*

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{pmatrix}. \quad (274)$$

(i) A is called an upper triangular matrix provided that for all $i, j \in \{1, 2, ..., n\}$, if $i > j$, then $a_{ij} = 0$.
(ii) A is called a lower triangular matrix provided that for all $i, j \in \{1, 2, ..., n\}$, if $i < j$, then $a_{ij} = 0$.

Example 7.13 *The following matrices are upper triangular:*

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 2 & 0 & 1 & 0 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 4 \end{pmatrix}. \quad (275)$$

The following matrices are lower triangular:

$$A = \begin{pmatrix} 2 & 0 \\ 1 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 0 & 1 & 2 \end{pmatrix} \quad C = \begin{pmatrix} 5 & 0 & 0 & 0 \\ 1 & 2 & 0 & 0 \\ -1 & 1 & 6 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix}. \tag{276}$$

The following matrices are both upper triangular and lower triangular:

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} \quad B = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \tag{277}$$

The following square matrices are neither upper triangular nor lower triangular:

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \quad B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \quad C = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 0 & 0 \end{pmatrix}. \tag{278}$$

Theorem 7.14 *Let A be an $n \times n$ matrix:*

$$A = \begin{pmatrix} a_{11} & a_{12} & ... & a_{1n} \\ a_{21} & a_{22} & ... & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & ... & a_{nn} \end{pmatrix}. \tag{279}$$

If A is upper triangular or lower triangular, then $\det(A) = a_{11}a_{22}...a_{nn}$.

7.4 Determinants and row operations

Theorem 7.15 *Let A be an $n \times n$ matrix.*
(i) If B is obtained by switching two rows of A, then $\det(B) = -\det(A)$.
(ii) If B is obtained by multiplying a nonzero scalar s by one row of A, then

$\det(B) = s\det(A)$.
(iii) If B is obtained by adding a multiple of one row of A to another row of A, then $\det(B) = \det(A)$.

This theorem gives us the following strategy: to compute the determinant of a large matrix, use elementary row operations to find a row-equivalent triangular matrix. Then, the determinant will be easy to compute.

Example 7.16 *Find the determinant of the following matrix:*

$$A = \begin{pmatrix} 0 & 0 & 0 & 5 \\ 1 & 7 & -1 & 2 \\ 0 & 0 & 3 & 4 \\ 0 & 2 & -2 & -10 \end{pmatrix}. \tag{280}$$

We'll use the theorem to predict how the determinant will change when we use elementary row operations. Let's apply elementary row operations until we get an upper triangular matrix:

$$\begin{vmatrix} 0 & 0 & 0 & 5 \\ 1 & 7 & -1 & 2 \\ 0 & 0 & 3 & 4 \\ 0 & 2 & -2 & -10 \end{vmatrix} = - \begin{vmatrix} 1 & 7 & -1 & 2 \\ 0 & 0 & 0 & 5 \\ 0 & 0 & 3 & 4 \\ 0 & 2 & -2 & -10 \end{vmatrix} = \begin{vmatrix} 1 & 7 & -1 & 2 \\ 0 & 2 & -2 & -10 \\ 0 & 0 & 3 & 4 \\ 0 & 0 & 0 & 5 \end{vmatrix} = \boxed{30}. \tag{281}$$

□

7.5 Properties of determinants

Theorem 7.17 *Let A be a square matrix. The matrix A is invertible if and only if* $\det(A) \neq 0$.

To summarize all of our results about square matrices, we state the following very big theorem.

Theorem 7.18 *Let A be an $n \times n$ matrix. The following statements are equivalent.*
(i) $\det A \neq 0$.

(ii) A is invertible.

(iii) Any linear transformation that has A as its matrix representation is invertible.

(iv) The equation $A\overrightarrow{v} = \overrightarrow{0}$ has only the trivial solution.

(v) The columns of A are linearly independent.

(vi) $\dim\left(\mathrm{Col}\left(A\right)\right) = n$.

(vii) The columns of A form a basis for $\mathbb{R}^n$.

(viii) A has full rank.

(ix) $\mathrm{rank}\left(A\right) = n$.

(x) A is row-equivalent to I_n.

(xi) Any augmented matrix whose left side is A has no free variables.

(xii) $\mathrm{null}\left(A\right) = \left\{\overrightarrow{0}\right\}$.

Theorem 7.19 *Let A and B be $n \times n$ matrices. In that case,*

$$\det\left(AB\right) = \det\left(A\right)\det\left(B\right). \tag{282}$$

In particular, if A and B are inverse matrices, then $AB = I_n$. The theorem above now indicates that

$$1 = \det\left(I_n\right) = \det\left(AB\right) = \det\left(A\right)\det\left(B\right). \tag{283}$$

From this, we deduce that $\det B = \frac{1}{\det A}$.

8 Eigenvalues, eigenvectors and diagonalization

8.1 Eigenvalues and eigenvectors

Definition 8.1 *Let A be an $n \times n$ matrix with entries in $\mathbb{C}$. An eigenvalue of A is a scalar $\lambda \in \mathbb{C}$ such that there exists a nonzero vector $\overrightarrow{v} \in \mathbb{C}^n$ satisfying the equation $A\overrightarrow{v} = \lambda\overrightarrow{v}$. In that case, $\overrightarrow{v}$ is called an eigenvector of A corresponding to λ.*

Do the eigenvectors corresponding to a particular eigenvalue of a matrix form a vector space? Almost. Strictly speaking, the zero vector is *not* an eigenvector. However, we have the following.

Definition 8.2 *Let A be an $n \times n$ matrix with entries in $\mathbb{C}$, and let $\lambda \in \mathbb{C}$ be an eigenvalue of A. The eigenspace of A associated to λ is the set*

$$E_\lambda = \left\{ \overrightarrow{v} \in \mathbb{C}^n \middle| A\overrightarrow{v} = \lambda\overrightarrow{v} \right\}. \tag{284}$$

8.2 Finding eigenvalues

Suppose that A is an $n \times n$ matrix that has an eigenvalue λ, with an associated eigenvector $\overrightarrow{v}$. In that case, $A\overrightarrow{v} = \lambda\overrightarrow{v}$. This can also be written as

$$(A - \lambda I_n)\overrightarrow{v} = \overrightarrow{0}. \tag{285}$$

By definition, in order for $\overrightarrow{v}$ to be an eigenvector of A corresponding to λ, we need that $\overrightarrow{v} \neq \overrightarrow{0}$. Now, the equation above has a non-trivial solution if and only if

$$\det(A - \lambda I_n) = 0. \tag{286}$$

(The left side of the equation above is called the "characteristic polynomial" of the matrix, and the entire equation is called the "characteristic equation" of the matrix.) Using this fact, we can find the eigenvalues of any square matrix. From there, finding the eigenvectors associated to each eigenvalue is possible. These vectors form an actual vector space associated to each eigenvalue.

Example 8.3 *Find the eigenvalues and associated eigenspaces of the following matrix:*

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{287}$$

We begin by writing the characteristic polynomial:

$$\begin{aligned} \det(A - \lambda I_2) &= \det\left(\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\right) \\ &= \begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1 = \lambda^2 - 4\lambda + 3. \end{aligned} \tag{288}$$

We now solve the characteristic equation by setting this to 0*:*

$$0 = \lambda^2 - 4\lambda + 3 = (\lambda - 3)(\lambda - 1). \tag{289}$$

This gives us the eigenvalues $\boxed{\lambda_1 = 1}$ *and* $\boxed{\lambda_2 = 3}$.

To find the eigenvectors associated to $\lambda_1 = 1$*, we now apply the definition of eigenvector; we seek* $\overrightarrow{v} \in \mathbb{R}^2$ *which are nonzero and satisfy*

$$A\overrightarrow{v} = 1\overrightarrow{v}. \tag{290}$$

This can be re-written as

$$(A - 1I_2)\overrightarrow{v} = \overrightarrow{0}. \tag{291}$$

In other words, we seek vectors that solve the equation

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{292}$$

This gives us the redundant system of equations

$$\begin{aligned} a + b &= 0 \\ a + b &= 0 \end{aligned}. \tag{293}$$

It is easy to see that b *is a free variable of this system, so we give it the name* s*. Now* *all eigenvectors can be described as*

$$\boxed{E_1 = \left\{ \begin{pmatrix} -s \\ s \end{pmatrix} \middle| s \in \mathbb{C} \right\}}. \tag{294}$$

In the same way, to find the eigenvectors associated to $\lambda_2 = 3$, we see $\overrightarrow{v} \in \mathbb{R}^2$ which are nonzero and satisfy

$$(A - 3I_2)\overrightarrow{v} = \overrightarrow{0}. \tag{295}$$

In other words, we want to solve the equation

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{296}$$

Again, we get a system of equations that does not have full rank:

$$\begin{aligned} -c + d &= 0 \\ c - d &= 0 \end{aligned}. \tag{297}$$

Consider d as a free variable, and call it t. In that case, all eigenvectors can be described as

$$\boxed{E_3 = \left\{ \begin{pmatrix} t \\ t \end{pmatrix} \middle| t \in \mathbb{C} \right\}}. \tag{298}$$

□

Example 8.4 *Find the eigenvalues and associated eigenspaces of the following matrix:*

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}. \tag{299}$$

We begin by writing the characteristic equation:

$$0 = \det(A - \lambda I_2) = \begin{vmatrix} 1-\lambda & 1 \\ -1 & 1-\lambda \end{vmatrix} = (\lambda - 1)^2 + 1 = \lambda^2 - 2\lambda + 2. \tag{300}$$

We solve this by using the quadratic formula:

$$\lambda = \frac{-(-2) \pm \sqrt{(-2)^2 - 4(1)(2)}}{2(1)} = \boxed{1 \pm i}. \tag{301}$$

To find the eigenspace associated to $1+i$*, we must solve the equation*

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{302}$$

This gives us the system of equations

$$\begin{aligned} -ia + b = 0 \\ -a - ib = 0 \end{aligned}. \tag{303}$$

This system of equations has b *as a free variable, so we name it* s*. The eigenspace can now be written as*

$$\boxed{E_{1+i} = \left\{ \begin{pmatrix} -is \\ s \end{pmatrix} \middle| s \in \mathbb{C} \right\}}. \tag{304}$$

To find the eigenspace associated to $1-i$*, we must solve*

$$\begin{pmatrix} i & 1 \\ -1 & i \end{pmatrix} \begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{305}$$

This gives us the system of equations

$$\begin{aligned} ic + d = 0 \\ -c + id = 0 \end{aligned}, \tag{306}$$

with d *as a free variable. We call* $d = t$*. Now the eigenspace associated to* $1-i$ *is*

$$\boxed{E_{1-i} = \left\{ \begin{pmatrix} it \\ t \end{pmatrix} \middle| t \in \mathbb{C} \right\}}. \tag{307}$$

□

Example 8.5 *Find the eigenvalues and associated eigenspaces of the following matrix:*

$$A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}. \tag{308}$$

We write the characteristic equation:

$$0 = \det(A - \lambda I_2) = \begin{vmatrix} 3-\lambda & 1 \\ 0 & 3-\lambda \end{vmatrix} = (\lambda - 3)^2. \tag{309}$$

This gives the single, repeated eigenvalue $\lambda = 3$.

To find the eigenspace associated to $\lambda = 3$, we must solve

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{310}$$

This gives us the single requirement that $b = 0$. Now a is a free variable, and so the eigenspace can be written as

$$\boxed{E_3 = \left\{ \begin{pmatrix} t \\ 0 \end{pmatrix} \middle| t \in \mathbb{C} \right\}}. \tag{311}$$

□

Example 8.6 *Find the eigenvalues and associated eigenspaces of the following matrix:*

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{312}$$

We write the characteristic equation:

$$0 = \det(A - \lambda I_3) = \begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 0 \\ 0 & 0 & 2-\lambda \end{vmatrix} = (2-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix}$$
$$= (2-\lambda)\left((\lambda-2)^2 - 1\right) = (2-\lambda)\left(\lambda^2 - 4\lambda + 3\right)$$
$$= (2-\lambda)(\lambda-1)(\lambda-3). \quad (313)$$

This gives us three eigenvalues: $\lambda_1 = 1$, $\lambda_2 = 2$ *and* $\lambda_3 = 3$.

To find the eigenspace associated to $\lambda_1 = 1$, *we solve the equation*

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \quad (314)$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{1} & 1 & 1 & | & 0 \\ \boxed{1} & 1 & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 0 \\ 0 & 0 & \boxed{-1} & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \end{pmatrix}$$
$$\xrightarrow{R_2 \leftrightarrow R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \\ 0 & 0 & \boxed{-1} & | & 0 \end{pmatrix} \xrightarrow{R_3 + R_2 \to R_3} \begin{pmatrix} \boxed{1} & 1 & 1 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}$$
$$\xrightarrow{R_1 - R_2 \to R_1} \begin{pmatrix} \boxed{1} & 1 & 0 & | & 0 \\ 0 & 0 & \boxed{1} & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}. \quad (315)$$

This gives us the system

$$\begin{aligned} a + b &= 0 \\ c &= 0 \end{aligned}, \quad (316)$$

where b is a free variable that we call r. Now the eigenspace is

$$\boxed{E_1 = \left\{ \begin{pmatrix} -r \\ r \\ 0 \end{pmatrix} \middle| r \in \mathbb{C} \right\}}. \tag{317}$$

To find the eigenspace associated to $\lambda_2 = 2$, we solve the equation

$$\begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{318}$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} 0 & \boxed{1} & 1 & | & 0 \\ \boxed{1} & 0 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \xrightarrow{R_1 \leftrightarrow R_2} \begin{pmatrix} \boxed{1} & 0 & 0 & | & 0 \\ 0 & \boxed{1} & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}. \tag{319}$$

This gives us the system

$$\begin{aligned} a &= 0 \\ b + c &= 0 \end{aligned}. \tag{320}$$

With c as the free variable s, the eigenspace is

$$\boxed{E_2 = \left\{ \begin{pmatrix} 0 \\ -s \\ s \end{pmatrix} \middle| s \in \mathbb{C} \right\}}. \tag{321}$$

To find the eigenspace associated to $\lambda_3 = 3$, we solve the equation

$$\begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{322}$$

We proceed with Gauss-Jordan elimination:

$$\begin{pmatrix} \boxed{-1} & 1 & 1 & | & 0 \\ \boxed{1} & -1 & 0 & | & 0 \\ 0 & 0 & \boxed{-1} & | & 0 \end{pmatrix} \xrightarrow{R_1+R_2\to R_1} \begin{pmatrix} 0 & 0 & 1 & | & 0 \\ \boxed{1} & -1 & 0 & | & 0 \\ 0 & 0 & \boxed{-1} & | & 0 \end{pmatrix}$$
$$\xrightarrow{R_3+R_1\to R_3} \begin{pmatrix} 0 & 0 & 1 & | & 0 \\ \boxed{1} & -1 & 0 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix} \xrightarrow{R_1\leftrightarrow R_2} \begin{pmatrix} \boxed{1} & -1 & 0 & | & 0 \\ 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & | & 0 \end{pmatrix}. \tag{323}$$

This gives us the system

$$\begin{aligned} a - b &= 0 \\ c &= 0 \end{aligned} \tag{324}$$

with b as a free variable, which we will call t. Now the eigenspace is

$$\boxed{E_3 = \left\{ \begin{pmatrix} t \\ t \\ 0 \end{pmatrix} \middle| t \in \mathbb{C} \right\}} \tag{325}$$

□

Proposition 8.7 *Let A be an upper or lower triangular matrix. The eigenvalues of A are the entries along the main diagonal.*

8.4 Diagonalization

Definition 8.8 *Let A be an $n \times n$ matrix. We say that A is a* <u>*diagonal matrix*</u> *provided that all entries of A that are not on the main diagonal are 0. (In other words, A is both upper and lower triangular.)*

Proposition 8.9 *Let*

$$D = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix} \tag{326}$$

be a diagonal matrix. Given any integer k,

$$D^k = \begin{pmatrix} a_{11} & 0 & 0 & \ldots & 0 \\ 0 & a_{22} & 0 & \ldots & 0 \\ 0 & 0 & a_{33} & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & a_{nn} \end{pmatrix}^k = \begin{pmatrix} {a_{11}}^k & 0 & 0 & \ldots & 0 \\ 0 & {a_{22}}^k & 0 & \ldots & 0 \\ 0 & 0 & {a_{33}}^k & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & {a_{nn}}^k \end{pmatrix}. \tag{327}$$

This proposition allows us to easily compute a kth power of any diagonal matrix. What if a matrix is not diagonal, though? Well, even in that case, there may be something we can do.

Definition 8.10 *Let A and B be $n \times n$ matrices. We say that A and B are similar matrices provided that there exists an invertible $n \times n$ matrix P such that $P^{-1}AP = B$.*

Definition 8.11 *Let A be an $n \times n$ matrix. We say that A is a diagonalizable matrix provided that there exists a diagonal matrix D such that A is similar to D.*

Theorem 8.12 *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors. In that case, $P^{-1}AP = D$, where P is the matrix whose columns are linearly independent eigenvectors and D is the diagonal matrix whose diagonal entries are the corresponding eigenvalues.*

To "diagonalize" a matrix A means to find the matrices P and D such that $P^{-1}AP = D$ and D is diagonal.

Example 8.13 *Diagonalize the following matrix.*

$$A = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}. \tag{328}$$

We first need the eigenvalues and eigenvectors:

$$\begin{aligned} 0 = \det\left(A - \lambda I_2\right) &= \begin{vmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{vmatrix} \\ &= \left(\lambda - 3\right)^2 - 1 = \lambda^2 - 6\lambda + 8 = \left(\lambda - 2\right)\left(\lambda - 4\right). \end{aligned} \tag{329}$$

This gives us the distinct, real eigenvalues $\lambda_1 = 4$ and $\lambda_2 = 2$.

To find the eigenspace corresponding to $\lambda_1 = 4$, we solve the following equation:

$$\left(A - 4I_2\right)\overrightarrow{v} = \overrightarrow{0}, \tag{330}$$

which we write as

$$\begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{331}$$

This gives us the system $a - b = 0$, with b as a free variable. Call $b = s$. Now $a - s = 0$, so $a = s$. Thus, the eigenvectors are

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} s \\ s \end{pmatrix} = s\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \tag{332}$$

where $s \neq 0$.

To find the eigenspace associated to $\lambda_2 = 2$, we solve

$$\left(A - 2I_2\right)\overrightarrow{v} = \overrightarrow{0}, \tag{333}$$

which we write as

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \tag{334}$$

This gives us $c + d = 0$, with d as a free variable. Call $d = t$. Now $c + t = 0$, so

$c = -t$. Thus, the eigenvectors take the form

$$\begin{pmatrix} c \\ d \end{pmatrix} = \begin{pmatrix} -t \\ t \end{pmatrix} = t \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \tag{335}$$

Based on our eigenvectors, we can construct the matrix P:

$$P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}. \tag{336}$$

We now find the inverse matrix:

$$\begin{pmatrix} 1 & -1 & | & 1 & 0 \\ 1 & 1 & | & 0 & 1 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} 1 & -1 & | & 1 & 0 \\ 0 & 2 & | & -1 & 1 \end{pmatrix}$$
$$\xrightarrow{\frac{1}{2}R_2 \to R_2} \begin{pmatrix} 1 & -1 & | & 1 & 0 \\ 0 & 1 & | & -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \xrightarrow{R_1 + R_2 \to R_1} \begin{pmatrix} 1 & 0 & | & \frac{1}{2} & \frac{1}{2} \\ 0 & 1 & | & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \tag{337}$$

giving us

$$P^{-1} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix}. \tag{338}$$

According to our theorem, $P^{-1}AP$ should be a diagonal matrix whose diagonal entries are 4 and 2. In fact, this is the case:

$$\boxed{\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}}. \tag{339}$$

□

Example 8.14 *Find A^{100}, where*

$$A = \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}. \tag{340}$$

Our strategy is as follows. If we can diagonalize A, then we can find a matrix

P such that $P^{-1}AP = D$, a diagonal matrix. Then, $A = PDP^{-1}$. We note that

$$
\begin{aligned}
A^{100} = \left(PDP^{-1}\right)^{100} &= \underbrace{\left(PDP^{-1}\right)\left(PDP^{-1}\right)...\left(PDP^{-1}\right)}_{100\,factors} \\
&= P\underbrace{DD...D}_{100\,factors}P^{-1} = PD^{100}P^{-1}. \qquad (341)
\end{aligned}
$$

So, we begin by finding the eigenvalues of A. We compute the characteristic polynomial:

$$\det\left(A - \lambda I_2\right) = \begin{vmatrix} 4-\lambda & -2 \\ 1 & 1-\lambda \end{vmatrix} = (\lambda - 3)(\lambda - 2). \qquad (342)$$

This gives us the distinct real eigenvalues $\lambda_1 = 3$ and $\lambda_2 = 2$.

To find the eigenvectors corresponding to $\lambda_1 = 3$, we need solutions of

$$
\begin{aligned}
\left(A - 3I_2\right)\overrightarrow{v} &= \overrightarrow{0} \\
\begin{pmatrix} 1 & -2 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned} \qquad (343)
$$

Doing Gauss-Jordan elimination makes it clear that our solution can be read from the augmented matrix

$$\left(\begin{array}{cc|c} 1 & -2 & 0 \\ 1 & -2 & 0 \end{array}\right) \xrightarrow{R_2 - R_1 \to R_2} \left(\begin{array}{cc|c} 1 & -2 & 0 \\ 0 & 0 & 0 \end{array}\right), \qquad (344)$$

giving us that $a - 2b = 0$ with b as a free variable. Call $b = s$. Now all eigenvectors take the form

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2s \\ s \end{pmatrix} = s\begin{pmatrix} 2 \\ 1 \end{pmatrix}. \qquad (345)$$

To find the eigenvectors corresponding to $\lambda_2 = 2$, we need solutions of

$$
\begin{aligned}
\left(A - 2I_2\right)\overrightarrow{v} &= \overrightarrow{0} \\
\begin{pmatrix} 2 & -2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}.
\end{aligned} \qquad (346)
$$

Doing Gauss-Jordan elimination makes it clear that our solution can be read from the augmented matrix

$$\begin{pmatrix} 2 & -2 & | & 0 \\ 1 & -1 & | & 0 \end{pmatrix} \xrightarrow{\frac{1}{2}R_1 \to R_1} \begin{pmatrix} 1 & -1 & | & 0 \\ 1 & -1 & | & 0 \end{pmatrix} \xrightarrow{R_2 + R_1 \to R_2} \begin{pmatrix} 1 & -1 & | & 0 \\ 0 & 0 & | & 0 \end{pmatrix}, \quad (347)$$

giving us that $a - b = 0$ with b as a free variable. Call $b = t$. Now all eigenvectors take the form

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} t \\ t \end{pmatrix} = t \begin{pmatrix} 1 \\ 1 \end{pmatrix}. \quad (348)$$

We construct the matrix P using two linearly independent eigenvectors as columns. For no particular reason, we make the choice

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \quad (349)$$

We compute the inverse matrix by Gauss-Jordan elimination:

$$\begin{pmatrix} 2 & 1| & 1 & 0 \\ 1 & 1| & 0 & 1 \end{pmatrix} \xrightarrow{R_1 - R_2 \to R_1} \begin{pmatrix} 1 & 0| & 1 & -1 \\ 1 & 1| & 0 & 1 \end{pmatrix} \xrightarrow{R_2 - R_1 \to R_2} \begin{pmatrix} 1 & 0| & 1 & -1 \\ 0 & 1| & -1 & 2 \end{pmatrix}. \quad (350)$$

Now $P^{-1}AP = D$, a diagonal matrix:

$$\begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix}. \quad (351)$$

From here,

$$\begin{aligned} \begin{pmatrix} 4 & -2 \\ 1 & 1 \end{pmatrix}^{100} = PD^{100}P^{-1} &= \boxed{\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 3^{100} & 0 \\ 0 & 2^{100} \end{pmatrix} \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}} \\ &= \boxed{\begin{pmatrix} 2\left(3^{100}\right) - 2^{100} & 2^{101} - 2\left(3^{100}\right) \\ 3^{100} - 2^{100} & 2^{101} - 3^{100} \end{pmatrix}}. \end{aligned} \quad (352)$$

□

Example 8.15 *Find A^{25}, where*

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{pmatrix}. \tag{353}$$

Our strategy is to diagonalize A (if possible) and then use the properties of the diagonal matrix to complete the computation. First, we compute the eigenvalues by factoring the characteristic polynomial:

$$\det(A - \lambda I_3) = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 1 & 4-\lambda & -1 \\ -2 & -4 & 4-\lambda \end{vmatrix}$$
$$= (2-\lambda)\begin{vmatrix} 4-\lambda & -1 \\ -4 & 4-\lambda \end{vmatrix} = (2-\lambda)(\lambda-2)(\lambda-6). \tag{354}$$

This gives us two distinct eigenvalues: $\lambda_1 = 6$, with a multiplicity of 1, and $\lambda_2 = 2$, with a multiplicity of 2.

To find the eigenvectors associated to $\lambda_1 = 6$, we need to solve the following equation:

$$(A - 6I_3)\overrightarrow{v} = \overrightarrow{0}, \tag{355}$$

which we write as

$$\begin{pmatrix} -4 & 0 & 0 \\ 1 & -2 & -1 \\ -2 & -4 & -2 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{356}$$

Gauss-Jordan elimination gives

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{357}$$

This gives us the system

$$\begin{aligned} a &= 0 \\ b + \tfrac{1}{2}c &= 0 \end{aligned}, \tag{358}$$

with c as a free variable. Call $c = r$. Now the eigenvectors all take the form

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ -\frac{1}{2}r \\ r \end{pmatrix} = r \begin{pmatrix} 0 \\ -\frac{1}{2} \\ 1 \end{pmatrix}, \tag{359}$$

where $r \neq 0$.

To find the eigenvectors associated to $\lambda_2 = 2$ we need to solve

$$(A - 2I_3)\overrightarrow{v} = \overrightarrow{0}, \tag{360}$$

which we write as

$$\begin{pmatrix} 0 & 0 & 0 \\ 1 & 2 & -1 \\ -2 & -4 & 2 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{361}$$

Gauss-Jordan elimination gives

$$\begin{pmatrix} 1 & 2 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}. \tag{362}$$

This gives us $d + 2e - f = 0$, with e and f being free variables. Call $e = s$ and $f = t$. Now $d + 2s - t = 0$, so $d = -2s + t$. The eigenvectors take the form

$$\begin{pmatrix} d \\ e \\ f \end{pmatrix} = \begin{pmatrix} -2s + t \\ s \\ t \end{pmatrix} = s \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \tag{363}$$

where s and t are not both 0.

To construct the matrix P, we take linearly independent eigenvectors as the *columns:*

$$P = \begin{pmatrix} 0 & -2 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}. \tag{364}$$

We take the inverse matrix:

$$P^{-1} = \begin{pmatrix} -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \end{pmatrix}. \tag{365}$$

According to our theorem, $P^{-1}AP$ should be a diagonal matrix, and in fact, it is:

$$\begin{pmatrix} -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \end{pmatrix} \begin{pmatrix} 2 & 0 & 0 \\ 1 & 4 & -1 \\ -2 & -4 & 4 \end{pmatrix} \begin{pmatrix} 0 & -2 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}. \tag{366}$$

Now, to find A^{25}, we note that

$$P^{-1}A^{25}P = \left(P^{-1}AP\right)^{25} = D^{25} = \begin{pmatrix} 6 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}^{25} = \begin{pmatrix} 6^{25} & 0 & 0 \\ 0 & 2^{25} & 0 \\ 0 & 0 & 2^{25} \end{pmatrix} \tag{367}$$

Therefore,

$$\begin{aligned} A^{25} &= P \begin{pmatrix} 6^{25} & 0 & 0 \\ 0 & 2^{25} & 0 \\ 0 & 0 & 2^{25} \end{pmatrix} P^{-1} \\ &= \boxed{\begin{pmatrix} 0 & -2 & 1 \\ -1 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix} \begin{pmatrix} 6^{25} & 0 & 0 \\ 0 & 2^{25} & 0 \\ 0 & 0 & 2^{25} \end{pmatrix} \begin{pmatrix} -\frac{1}{4} & -\frac{1}{2} & \frac{1}{4} \\ -\frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{2} & 1 & \frac{1}{2} \end{pmatrix}} \end{aligned} \tag{368}$$

□

8.9 Properties of eigenvectors and eigenvalues

Definition 8.16 *Let A be an $n \times n$ matrix, and let λ be an eigenvalue of A.*
(i) The algebraic multiplicity of λ is the largest integer k such that $(x - \lambda)^k$ divides the characteristic polynomial of A.
(ii) The geometric multiplicity of λ is the dimension of the eigenspace of λ.
(iii) The defect of λ is the algebraic multiplicity of λ minus the geometric multiplicity of λ.
(iv) If the defect of an eigenvalue is zero, we say that the eigenvalue is complete.
(v) If the defect of an eigenvalue is greater than zero, we say that the eigenvalue is defective.
(vi) If a matrix has a defective eigenvalue, we say that the matrix is defective.

Proposition 8.17 *Given any $n \times n$ matrix A, the defect of any given eigenvalue of A is non-negative.*

Theorem 8.18 *Let A be an $n \times n$ matrix. The matrix A is diagonalizable if and only if A is not defective.*

Example 8.19 *Determine whether the following matrix is diagonalizable.*

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 \\ 0 & 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{pmatrix}. \tag{369}$$

We need to know whether any eigenvalue of A is defective. This is an upper triangular matrix, so we already know that the eigenvalues are $\lambda_1 = 5$, having an algebraic multiplicity of 1, $\lambda_2 = 4$, having an algebraic multiplicity of 2, and $\lambda_3 = 3$, having an algebraic multiplicity of 2. Since the geometric multiplicity of $\lambda_1 = 5$ is at least 1, we must have that $\lambda_1 = 5$ has a defect of $1 - 1 = 0$. This leaves $\lambda_2 = 4$ and $\lambda_3 = 3$.

Consider $\lambda_3 = 3$. We will compute its geometric multiplicity by finding the *eigenspace E_3. We must solve*

$$
\begin{aligned}
(A - 3I_5)\,\vec{v} &= \vec{0} \\
\begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 2 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.
\end{aligned} \tag{370}
$$

The reduced row echelon form of this system is

$$
\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & | & 0 \\ 0 & 0 & 1 & 0 & 0 & | & 0 \\ 0 & 0 & 0 & 1 & 0 & | & 0 \\ 0 & 0 & 0 & 0 & 1 & | & 0 \\ 0 & 0 & 0 & 0 & 0 & | & 0 \end{pmatrix}. \tag{371}
$$

This gives us that $b = c = d = e = 0$, and a is a free variable. We give the name $a = t$. Now all eigenvectors take the form

$$
\begin{pmatrix} a \\ b \\ c \\ d \\ e \end{pmatrix} = \begin{pmatrix} t \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix} = t \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}. \tag{372}
$$

This shows that the eigenspace E_3 is spanned by a single vector, and so the geometric multiplicity of $\lambda_3 = 3$ is 1. Therefore, the defect of $\lambda_3 = 3$ is $2 - 1 = 1$. This shows that $\lambda_3 = 3$ is defective, and therefore, A is defective. Our theorem indicates that A $\boxed{\text{cannot be diagonalized}}$. $\square$

11 Inner product spaces

11.1 Real inner product spaces

Definition 11.1 *Let V be a vector space over $\mathbb{R}$. A* <u>*real inner product on V*</u> *is an operation $\langle , \rangle$ that takes two input vectors from V and associates an element of $\mathbb{R}$ as the output, satisfying the following condtions:*
(i) For any $u, v \in V$, $\langle u, v \rangle = \langle v, u \rangle$ ("symmetry")
(ii) For any $u, v, w \in V$ and $s, t \in \mathbb{R}$, $\langle u, sv + tw \rangle = s \langle u, v \rangle + t \langle u, w \rangle$. ("linearity")
(iii) For any $u \in V$, $\langle u, u \rangle \geq 0$, and $\langle u, u \rangle = 0$ if and only if $u = 0$ ("positive definite")
If $\langle , \rangle$ is a real inner product on V, we say that V is a <u>*real inner product space*</u>*.*

Example 11.2 *The following are real inner product spaces.*
(i) $\mathbb{R}^n$, equipped with the dot product.
(ii) Let $V = \mathbb{R}^2$, and let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}. \tag{373}$$

Define $\langle u, v \rangle = u^T A v$.
(iii) Let $a, b \in \mathbb{R}$ such that $a < b$. Let

$$\mathcal{C}\left(\left[a, b\right], \mathbb{R}\right) = \left\{ f : \left[a, b\right] \to \mathbb{R} \middle| f \text{ is continuous} \right\}. \tag{374}$$

Define

$$\langle f, g \rangle = \int_a^b f\left(x\right) g\left(x\right) \, \mathrm{d}x. \tag{375}$$

Definition 11.3 *Let V be an inner product space. Given $v \in V$, the* <u>*norm, or magnitude of v*</u> *is the value $||v|| = \sqrt{\langle v, v \rangle}$.*

Example 11.4 *Compute the norm of 1, x, x^2, and x^3 in $\mathcal{C}\left(\left[-1, 1\right], \mathbb{R}\right)$.*

By definition,

$$\langle 1, 1 \rangle = \int_{-1}^{1} \left(1\right)\left(1\right) \, \mathrm{d}x = x \Big|_{-1}^{1} = 2. \tag{376}$$

Therefore, $\boxed{||1|| = \sqrt{2}}$*. Similarly,*

$$\langle x, x\rangle = \int_{-1}^{1} (x)\,(x)\ \mathrm{d}x = \frac{1}{3}x^3\Big|_{-1}^{1} = \frac{2}{3}, \tag{377}$$

so $\boxed{||x|| = \sqrt{\frac{2}{3}}}$*. Next,*

$$\left\langle x^2, x^2\right\rangle = \int_{-1}^{1} \left(x^2\right)\left(x^2\right)\ \mathrm{d}x = \frac{1}{5}x^5\Big|_{-1}^{1} = \frac{2}{5}, \tag{378}$$

so $\boxed{||x^2|| = \sqrt{\frac{2}{5}}}$*. Finally,*

$$\left\langle x^3, x^3\right\rangle = \int_{-1}^{1} \left(x^3\right)\left(x^3\right)\ \mathrm{d}x = \frac{1}{7}x^7\Big|_{-1}^{1} = \frac{2}{7}, \tag{379}$$

so $\boxed{||x^3|| = \sqrt{\frac{2}{7}}}$. $\square$

Definition 11.5 *let V be an inner product space. Given $v \in V$, we say that v is a unit vector [in V] provided that $||v|| = 1$.*

Theorem 11.6 *(Cauchy-Schwarz inequality) Let V be an inner product space. Given $u, v \in V$,*

$$\langle u, v\rangle \leq ||u||||v||. \tag{380}$$

Theorem 11.7 *(Triangle inequality) Let V be an inner product space. Given $u, v \in V$,*

$$||u + v|| \leq ||u|| + ||v||. \tag{381}$$

11.2 Orthogonality

Definition 11.8 *Let V be an inner product space, and let $u, v \in V$. We say that u and v are orthogonal [in V] provided that $\langle u, v\rangle = 0$.*

Notation $u \perp v$

Definition 11.9 *Let V be an inner product space, and let $S \subseteq V$. The* *<u>orthogonal complement of S in V</u>* *is the set*

$$S^{\perp} = \left\{ v \in V \middle| \text{For all } u \in S,\ \langle u, v \rangle = 0 \right\}. \tag{382}$$

Proposition 11.10 *Let V be an inner product space. Given any $S \subseteq V$, $S^{\perp}$ is a vector subspace of V.*

Example 11.11 *Let $V = \mathbb{R}^3$, with an inner product defined by the usual dot product. Let*

$$S = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}. \tag{383}$$

Find $S^{\perp}$, the orthogonal complement of S in V.

Let

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \overrightarrow{v_2} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \tag{384}$$

By definition,

$$S^{\perp} = \left\{ \overrightarrow{u} \in \mathbb{R}^3 \middle| \overrightarrow{u} \cdot \overrightarrow{v_1} = 0 \text{ and } \overrightarrow{u} \cdot \overrightarrow{v_2} = 0 \right\}. \tag{385}$$

Every element $\overrightarrow{u} \in \mathbb{R}^3$ takes the form

$$\overrightarrow{u} = \begin{pmatrix} a \\ b \\ c \end{pmatrix}. \tag{386}$$

We must have

$$\begin{aligned} 0 = \vec{u} \cdot \vec{v_1} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = a \\ 0 = \vec{u} \cdot \vec{v_2} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = a + c. \end{aligned} \tag{387}$$

This indicates that $a = 0$ *and* $c = 0$*. Thus,*

$$\boxed{S^{\perp} = \left\{ \begin{pmatrix} a \\ b \\ c \end{pmatrix} \in \mathbb{R}^3 \middle| a = c = 0 \right\} = \left\{ \begin{pmatrix} 0 \\ b \\ 0 \end{pmatrix} \middle| b \in \mathbb{R} \right\}}. \tag{388}$$

□

Example 11.12 *Let* $V = \mathbb{R}^2$ *with an inner product defined by*

$$\langle \vec{u}, \vec{v} \rangle = \vec{u}^T \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \vec{v}. \tag{389}$$

Let $S = \{\vec{w}\}$*, where*

$$\vec{w} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}. \tag{390}$$

Find $S^{\perp}$*, the orthogonal complement of* S *in* V*.*

By definition,

$$S^{\perp} = \left\{ \vec{v} \in \mathbb{R}^2 \middle| \vec{w}^T \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \vec{v} = 0 \right\}. \tag{391}$$

Every vector of $\mathbb{R}^2$ *takes the form*

$$\vec{v} = \begin{pmatrix} a \\ b \end{pmatrix}, \tag{392}$$

So this implies that $\overrightarrow{v} \in S^{\perp}$ *if and only if*

$$0 = \begin{pmatrix} 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 3 & -2 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = 3a - 2b. \quad (393)$$

We deduce that

$$\boxed{S^{\perp} = \left\{ \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \middle| 3a - 2b = 0 \right\}}. \quad (394)$$

□

Example 11.13 *Let* $V = P_3(\mathbb{R})$ *with an inner product defined by*

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\, g(x)\ \mathrm{d}x \quad (395)$$

Let $S = \{x^2\}$. *Find* $S^{\perp}$, *the orthogonal complement of* S *in* V.

By definition,

$$S^{\perp} = \left\{ f \in P_3(\mathbb{R}) \middle| \int_{-1}^{1} x^2 f(x)\ \mathrm{d}x = 0 \right\}. \quad (396)$$

Each element $f \in P_3(\mathbb{R})$ *can be written as*

$$f(x) = a_3x^3 + a_2x^2 + a_1x + a_0, \quad (397)$$

for some $a_0, a_1, a_2, a_3 \in \mathbb{R}$. *Thus,* $f \in S^{\perp}$ *if and only if*

$$\begin{aligned} 0 &= \int_{-1}^{1} x^2 \left(a_3x^3 + a_2x^2 + a_1x + a_0\right)\ \mathrm{d}x \\ &= \frac{1}{6}a_3x^6 + \frac{1}{5}a_2x^5 + \frac{1}{4}a_1x^4 + \frac{1}{3}a_0x^3 \bigg|_{-1}^{1} = \frac{2}{5}a_1 + \frac{2}{3}a_0. \end{aligned} \quad (398)$$

In other words,

$$\boxed{S^{\perp} = \left\{ a_3x^3 + a_2x^2 + a_1x + a_0 \middle| \frac{2}{5}a_2 + \frac{2}{3}a_0 = 0 \right\}}. \tag{399}$$

□

Definition 11.14 *Let V be an inner product space. Given a set of nonzero vectors $S \subseteq V$, we say that S is an orthogonal set in V provided that for all $u, v \in S$, if $u \neq v$, then $u \perp v$. We say that S is an orthonormal set in V provided that S is an orthogonal set in V and each element of S is a unit vector in V.*

Proposition 11.15 *Let V be an inner product space. Given a nonempty $S \subseteq V$, if S is orthogonal, then S is linearly independent.*

In particular, the previous proposition indicates that any orthogonal set that spans an entire subspace is a basis for that subspace.

Proposition 11.16 *Let V be an inner product space, and let $B = (u_1, u_2, ..., u_k)$ be an orthogonal basis for V. Given any $v \in V$, the coordinates of v with respect to B are*

$$[v]_B = \begin{pmatrix} \frac{\langle u_1, v \rangle}{||u_1||^2} \\ \frac{\langle u_2, v \rangle}{||u_2||^2} \\ \vdots \\ \frac{\langle u_k, v \rangle}{||u_k||^2} \end{pmatrix}. \tag{400}$$

The proposition indicates that if we have an orthogonal basis, then the coordinates of any vector in the space are easy to compute. Further, these coordinates have a name.

Definition 11.17 *Let V be a real inner product space, and let B be an orthogonal basis of V. Given $v \in V$, the Fourier coefficients of v with respect to B are the coordinates of v with respect to B.*

11.3 The Gram-Schmidt orthogonalization procedure

We note that orthogonal bases (and especially orthonormal bases) for inner product spaces are often much easier to use than general bases. To see why, consider $\mathbb{R}^3$ with the usual dot product. Which of the following two bases for $\mathbb{R}^3$ is easier to work with?

$$B_1 = \left(\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} \right) \quad B_2 = \left(\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \tag{401}$$

A skeptic may say that it would make no real difference, since both are bases. In some ways, this is true. However, consider the problem of finding the coordinates of the following vector:

$$\overrightarrow{v} = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix} \tag{402}$$

In order to find the coordinates with respect to B_1, we need to find the unique scalars $s_1, s_2, s_3 \in \mathbb{R}$ such that

$$s_1 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + s_2 \begin{pmatrix} 2 \\ 2 \\ 0 \end{pmatrix} + s_3 \begin{pmatrix} 3 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix}. \tag{403}$$

This comes down to solving the system of equations

$$\begin{pmatrix} 1 & 2 & 3 & | & -2 \\ 1 & 2 & 0 & | & 4 \\ 1 & 0 & 0 & | & 3 \end{pmatrix}. \tag{404}$$

This is certainly not impossible, but it would require an amount of effort. On the other hand, if we want to find the coordinates of $\overrightarrow{v}$ with respect to B_2, then we need scalars $t_1, t_2, t_3 \in \mathbb{R}$ such that

$$t_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + t_2 \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + t_3 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 4 \\ 3 \end{pmatrix}. \tag{405}$$

This is a system that already corresponds to a reduced row echelon form:

$$\begin{pmatrix} 1 & 0 & 0 & | & -2 \\ 0 & 1 & 0 & | & 4 \\ 0 & 0 & 1 & | & 3 \end{pmatrix}. \tag{406}$$

Therefore, at least with regards to finding coordinates of vectors, an orthonormal basis is highly desirable.

Suppose that v_1 and v_2 are linearly independent, but not orthogonal. To construct a new vector u_2 that is orthogonal to v_1, we consider $u_2 = v_2 - tv_1$ for some $t \in \mathbb{R}$. To find t, we use the fact that $v_1 \perp u_2$:

$$0 = \langle v_1, u_2 \rangle = \langle v_1, v_2 - tv_1 \rangle = \langle v_1, v_2 \rangle - t \langle v_1, v_1 \rangle. \tag{407}$$

Now $t = \frac{\langle v_1, v_2 \rangle}{\langle v_1, v_1 \rangle} = \frac{\langle v_1, v_2 \rangle}{||v_1||^2}$. We can extend this to find an orthogonal basis for any real inner product space.

Proposition 11.18 *(Gram-Schmidt orthogonalization) Let V be an inner product space, and let $(v_1, v_2, ..., v_n)$ be a basis for V. Define*

$$\begin{array}{c} u_1 = v_1 \\ u_2 = v_2 - \frac{\langle u_1, v_2 \rangle}{||u_1||^2} u_1 \\ u_3 = v_3 - \frac{\langle u_1, v_3 \rangle}{||u_1||^2} u_1 - \frac{\langle u_2, v_3 \rangle}{||u_2||^2} u_2 \\ \vdots \\ u_n = v_n - \frac{\langle u_1, v_n \rangle}{||u_1||^2} u_1 - \frac{\langle u_2, v_n \rangle}{||u_2||^2} u_2 - \dots - \frac{\langle u_{n-1}, v_n \rangle}{||u_{n-1}||^2} u_{n-1} \end{array}. \tag{408}$$

In that case, $(u_1, u_2, ..., u_n)$ is an orthogonal basis for V.

Example 11.19 *Find an orthogonal basis for*

$$V = \operatorname{span}\left(\begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \begin{pmatrix}1\\1\\0\\0\end{pmatrix}\right) \tag{409}$$

in $\mathbb{R}^4$ with the dot product.

Define

$$v_1 = \begin{pmatrix}1\\1\\1\\1\end{pmatrix}, \quad v_2 = \begin{pmatrix}1\\1\\1\\0\end{pmatrix}, \quad v_3 = \begin{pmatrix}1\\1\\0\\0\end{pmatrix}. \tag{410}$$

We note that (v_1, v_2, v_3) is linearly independent, and so it is a basis for V. We follow the Gram-Schmidt orthogonalization procedure:

$$\begin{aligned} u_1 = v_1 &= \begin{pmatrix}1\\1\\1\\1\end{pmatrix} \\ u_2 = v_2 - \tfrac{u_1 \cdot v_2}{u_1 \cdot u_1} u_1 &= \begin{pmatrix}1\\1\\1\\0\end{pmatrix} - \tfrac{3}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} = \begin{pmatrix}\frac{1}{4}\\\frac{1}{4}\\\frac{1}{4}\\-\frac{3}{4}\end{pmatrix} \\ u_3 = v_3 - \tfrac{u_1 \cdot v_3}{u_1 \cdot u_1} u_1 - \tfrac{u_2 \cdot v_3}{u_2 \cdot u_2} u_2 &= \begin{pmatrix}1\\1\\0\\0\end{pmatrix} - \tfrac{2}{4}\begin{pmatrix}1\\1\\1\\1\end{pmatrix} - \tfrac{\left(\frac{1}{2}\right)}{\left(\frac{3}{4}\right)}\begin{pmatrix}\frac{1}{4}\\\frac{1}{4}\\\frac{1}{4}\\-\frac{3}{4}\end{pmatrix} = \begin{pmatrix}\frac{1}{3}\\\frac{1}{3}\\-\frac{2}{3}\\0\end{pmatrix} \end{aligned} \tag{411}$$

Now, according to the proposition, $\boxed{(u_1, u_2, u_3)}$ is an orthogonal basis for V. □

Example 11.20 *Consider the vector space $V = P_2(\mathbb{R})$ of polynomials with degree*

2 *or less with coefficients in* $\mathbb{R}$, *with an inner product defined by*

$$\langle f, g\rangle = \int_0^1 f(x)\, g(x)\ \mathrm{d}x. \tag{412}$$

Find an orthogonal basis for V.

We note that $V = \operatorname{span}(1, x, x^2)$, *and in fact that* $(1, x, x^2)$ *is a basis for* V. *We use the Gram-Schmidt orthogonalization procedure to find an orthogonal basis. First,* $u_1 = v_1 = 1$. *Next,*

$$u_2 = v_2 - \frac{\langle v_2, u_1\rangle}{||u_1||^2} u_1 = x - \frac{\langle x, 1\rangle}{||1||^2}(1) = x - \frac{1}{2} \tag{413}$$

Finally,

$$\begin{aligned} u_3 = v_3 - \frac{\langle v_3, u_1\rangle}{||u_1||^2} u_1 - \frac{\langle v_3, u_2\rangle}{||u_2||^2} u_2 & \\ = x^2 - \frac{\langle x^2, 1\rangle}{||1||^2}(1) - \frac{\left\langle x^2, x - \frac{1}{2}\right\rangle}{||x - \frac{1}{2}||^2}\left(x - \frac{1}{2}\right) & \\ = x^2 - \frac{\left(\frac{1}{3}\right)}{1^2}(1) - \frac{\left(\frac{1}{12}\right)}{\left(\frac{1}{12}\right)^2}\left(x - \frac{1}{2}\right) = x^2 - 12x + \frac{17}{3}. & \end{aligned} \tag{414}$$

Thus, our orthogonal basis is $\boxed{\left(1, x - \frac{1}{2}, x^2 - 12x + \frac{17}{3}\right)}$. □

11.4 Orthogonal projections and Fourier series

Let V be an inner product space containing a vector subspace W and a given vector $v \in V$. Suppose that we want to find the vector $w \in W$ which is as "close" to v as possible. Now that we have a notion of distance, we can define this rigorously. Specifically, we desire to minimize the distance between w and v, which is the real-valued quantity $||w - v||$. We have a name for the vector that does this.

Definition 11.21 *Let* V *be an inner product space containing a vector subspace* W *and a given vector* $v \in V$. *The orthogonal projection of* v *onto* w *(also called the best approximation of* v *in* W*) is the vector* $\operatorname{proj}_W(v) \in W$ *such that for any vector*

$w \in W$,

$$||\text{proj}_W(v) - v|| \leq ||w - v||. \tag{415}$$

Proposition 11.22 *Let V be an inner product space, $v \in V$, and let W be a vector subspace of V. Assume that $B = (u_1, u_2, ..., u_k)$ is an orthogonal basis for W. In that case,*

$$\text{proj}_W(v) = \frac{\langle u_1, v\rangle}{||u_1||^2}u_1 + \frac{\langle u_2, v\rangle}{||u_2||^2}u_2 + ... + \frac{\langle u_k, v\rangle}{||u_k||^2}u_k. \tag{416}$$

In other words, the coordinates of $\text{proj}_W(v)$ with respect to the basis B are

$$[\text{proj}_W(v)]_B = \begin{pmatrix} \frac{\langle u_1,v\rangle}{||u_1||^2} \\ \frac{\langle u_2,v\rangle}{||u_2||^2} \\ \vdots \\ \frac{\langle u_k,v\rangle}{||u_k||^2} \end{pmatrix}. \tag{417}$$

Example 11.23 *Let $V = \mathbb{R}^4$ with the usual dot product, and let*

$$W = \text{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} \right\}. \tag{418}$$

Let

$$\overrightarrow{v} = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}. \tag{419}$$

Find $\text{proj}_W(v)$, the orthogonal projection of v onto W.

First, we need an orthogonal basis for W. To do this, we enact the Gram-

Schmidt orthogonalization procedure:

$$\begin{array}{c} \overrightarrow{u_1} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \\ \overrightarrow{u_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} \end{array}. \tag{420}$$

We name this basis B. By the previous proposition,

$$\operatorname{proj}_W\left(\overrightarrow{v}\right) = \frac{\overrightarrow{u_1}\cdot\overrightarrow{v}}{\overrightarrow{u_1}\cdot\overrightarrow{u_1}}\overrightarrow{u_1} + \frac{\overrightarrow{u_2}\cdot\overrightarrow{v}}{\overrightarrow{u_2}\cdot\overrightarrow{u_2}}\overrightarrow{u_2} = \frac{5}{2}\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} + \frac{\left(\frac{3}{2}\right)}{\left(\frac{3}{2}\right)}\begin{pmatrix} -\frac{1}{2} \\ 1 \\ 0 \\ \frac{1}{2} \end{pmatrix} = \boxed{\begin{pmatrix} 2 \\ 1 \\ 0 \\ 3 \end{pmatrix}}. \tag{421}$$

□

Definition 11.24 *Let V be an infinite-dimensional inner product space containing an orthogonal set $S = \{u_1, u_2, u_3, ...\}$. Let $W = \operatorname{span}(S)$. Given $v \in V$, a generalized Fourier approximation of v in W is a vector of the form*

$$v_k = \frac{\langle u_1, v\rangle}{||u_1||^2} + \frac{\langle u_2, v\rangle}{||u_2||^2} + ... + \frac{\langle u_k, v\rangle}{||u_k||^2} \tag{422}$$

for some positive integer k. The limit $\lim_{k\to\infty} v_k$ is called the generalized Fourier series of v in W.

Based on the name, we should hope that the orthogonal projection of a vector onto a subspace W is orthogonal to every vector of W. Indeed, that is the case, as the following proposition indicates.

Proposition 11.25 *Let V be an inner product space containing a vector subspace W. Given $v \in V$, $\operatorname{proj}_W(v) \in W^{\perp}$.*

11.6 Orthogonal functions and orthogonal matrices

Definition 11.26 *Let V and W be inner product spaces, and let $T : V \to W$ be a linear transformation. We say that T is an isometry from V to W provided that for all $v_1, v_2 \in V$,*

$$\langle T(v_1), T(v_2) \rangle = \langle v_1, v_2 \rangle . \tag{423}$$

We say that T is an orthogonal transformation provided that T is an invertible isometry.

Example 11.27 *Let $T : \mathbb{R}^2 \to \mathbb{R}^2$ via*

$$T \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} -y \\ x \end{pmatrix} . \tag{424}$$

Determine whether T is an isometry.

Let $\overrightarrow{v_1}, \overrightarrow{v_2} \in \mathbb{R}^2$ be arbitrary. First, we note that

$$\overrightarrow{v_1} = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \quad \textit{and} \quad \overrightarrow{v_2} = \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \tag{425}$$

for some appropriate $a_1, a_2, b_1, b_2 \in \mathbb{R}$. Now,

$$\begin{aligned} \langle T(\overrightarrow{v_1}), T(\overrightarrow{v_2}) \rangle &= \left\langle T \begin{pmatrix} a_1 \\ b_1 \end{pmatrix}, T \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} \right\rangle = \begin{pmatrix} -b_1 \\ a_1 \end{pmatrix} \cdot \begin{pmatrix} -b_2 \\ a_2 \end{pmatrix} \\ &= a_1 a_2 + b_1 b_2 = \begin{pmatrix} a_1 \\ b_1 \end{pmatrix} \cdot \begin{pmatrix} a_2 \\ b_2 \end{pmatrix} = \langle \overrightarrow{v_1}, \overrightarrow{v_2} \rangle . \end{aligned} \tag{426}$$

Thus, T is an isometry. □

Proposition 11.28 *Let V and W be finite-dimensional inner product spaces, and let $T : V \to W$ be a linear transformation. Let B be a basis for V, and let C be a basis for W. Suppose that M is the matrix representation of T with respect to the bases B and C. The following statements are true.*
(i) T is an isometry if and only if $M^T M$ is an identity matrix.

(ii) T is an orthogonal transformation if and only if $M^T M$ is an identity matrix and $\dim V = \dim W$.

Definition 11.29 *Let A be an $n \times n$ matrix. We say that A is an orthogonal matrix provided that $A^T A = I_n$.*

Proposition 11.30 *Let $A \in M_{n\times n}(\mathbb{R})$. The following statements are equivalent.*
(i) A is orthogonal.
(ii) The columns of A form an orthonormal set.
(iii) A is invertible, and $A^T = A^{-1}$.

11.7 Diagonalization of symmetric matrices

Proposition 11.31 *Let $A \in M_{n\times n}(\mathbb{R})$. If A is symmetric, then all eigenvalues of A are real numbers. Further, eigenvectors from different eigenspaces are orthogonal.*

Definition 11.32 *Let A be an $n \times n$ matrix. We say that A is orthogonally diagonalizable provided that there exists an orthogonal matrix P and a diagonal matrix D such that $P^{-1}AP = D$.*

Theorem 11.33 *Let $A \in M_{n\times n}(\mathbb{R})$. The matrix A is orthogonally diagonalizable if and only if A is symmetric.*

To find an orthogonal diagonalization, follow the same procedure as for ordinary diagonalization, but also ensure that the vectors chosen for the matrix P are orthonormal.

Example 11.34 *Orthogonally diagonalize the following matrix:*

$$A = \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix}. \tag{427}$$

(In other words, find matrices P and D such that P is orthogonal, D is diagonal, and $P^{-1}AP = D$.)

We first need the eigenvalues of A, so we write the characteristic polynomial:

$$\det(A - \lambda I_2) = \begin{vmatrix} 3-\lambda & 2 \\ 2 & 6-\lambda \end{vmatrix} = \lambda^2 - 9\lambda + 14 = (\lambda - 2)(\lambda - 7). \qquad (428)$$

This gives us the eigenvalues $\lambda_1 = 7$ and $\lambda_2 = 2$, each with an algebraic multiplicity of 1.

To find the eigenspace corresponding to $\lambda_1 = 7$, we solve

$$\begin{gathered} (A - 7I_2)\overrightarrow{v} = \overrightarrow{0} \\ \begin{pmatrix} -4 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{gathered} \qquad (429)$$

This gives us that $2a - b = 0$, with b as a free variable. Call $b = s$. Now $a = \frac{1}{2}s$, and so each element of the eigenspace E_7 takes the form

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{1}{2}s \\ s \end{pmatrix} = s \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}. \qquad (430)$$

To find the eigenspace corresponding to $\lambda_2 = 2$, we solve

$$\begin{gathered} (A - 2I_2)\overrightarrow{v} = \overrightarrow{0} \\ \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}. \end{gathered} \qquad (431)$$

This gives us that $a + 2b = 0$, with b as a free variable. Call $b = t$. Now $a = -2t$, so each element of the eigenspace E_2 takes the form

$$\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -2t \\ t \end{pmatrix} = t \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \qquad (432)$$

As usual, to construct P, we need to select two linearly independent eigenvec-

tors. This is equivalent to selecting values for s *and* t*:*

$$\overrightarrow{v_1} = s \begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}, \quad \overrightarrow{v_2} = t \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \tag{433}$$

Unlike before, though, we must also select these eigenvectors so that they are orthogonal unit vectors. Therefore, we need that $||\overrightarrow{v_1}|| = ||\overrightarrow{v_2}|| = 1$*. This gives us the equations*

$$\begin{aligned} 1 = ||\overrightarrow{v_1}|| = \sqrt{\left(\tfrac{1}{2}s\right)^2 + (s)^2} \\ 1 = ||\overrightarrow{v_2}|| = \sqrt{(-2t)^2 + (t)^2} \end{aligned}, \tag{434}$$

from which we can get the solutions $s = \frac{2}{\sqrt{5}}$ *and* $t = \frac{1}{\sqrt{5}}$*. Further,*

$$\overrightarrow{v_1} \cdot \overrightarrow{v_2} = \begin{pmatrix} \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{pmatrix} \cdot \begin{pmatrix} -\frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{pmatrix} = 0, \tag{435}$$

so these eigenvectors are also orthogonal. We use these to construct P*:*

$$P = \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}. \tag{436}$$

Since this matrix is orthogonal, its inverse is equal to its transpose:

$$P^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}. \tag{437}$$

This gives us the diagonalization

$$P^{-1}AP = \boxed{\begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -\frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} 7 & 0 \\ 0 & 2 \end{pmatrix}} = D. \tag{438}$$

□

11.8 Positive semidefinite and positive definite matrices

In an earlier example in this chapter, we defined an inner product on $\mathbb{R}^2$ via the relationship

$$\langle \overrightarrow{u}, \overrightarrow{v} \rangle = \overrightarrow{u}^T A \overrightarrow{v}, \tag{439}$$

where

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}. \tag{440}$$

Which matrices A will cause this operation to define an inner product? We explore the answer to this question in this subsection.

Definition 11.35 *Let $A \in M_{n\times n}(\mathbb{R})$.*
(i) We say that A is a <u>*positive semidefinite matrix*</u> *provided that A is symmetric and for any $\overrightarrow{v} \in \mathbb{R}^n$, $\overrightarrow{v}^T A \overrightarrow{v} \geq 0$.*
(ii) We say that A is a <u>*positive definite matrix*</u> *provided that A is positive semidefinite and for any $\overrightarrow{v} \in \mathbb{R}^n$, if $\overrightarrow{v}^T A \overrightarrow{v} = 0$, then $\overrightarrow{v} = \overrightarrow{0}$.*

A matrix is positive definite if and only if it can be used to define an inner product of the form $\langle \overrightarrow{u}, \overrightarrow{v} \rangle = \overrightarrow{u}^T A \overrightarrow{v}$, as the following proposition shows.

Proposition 11.36 *Let $A \in M_{n\times n}(\mathbb{R})$. The formula*

$$\langle \overrightarrow{u}, \overrightarrow{v} \rangle = \overrightarrow{u}^T A \overrightarrow{v} \tag{441}$$

defines an inner product on $\mathbb{R}^n$ if and only if A is positive definite. Conversely, every inner product on $\mathbb{R}^n$ follows the above formula for some positive definite matrix A.

We have a characterization for all positive definite matrices, as seen in the following proposition.

Proposition 11.37 *Let $A \in M_{n\times n}(\mathbb{R})$ be a symmetric matrix. The following statements are true.*
(i) A is positive semidefinite if and only if all eigenvalues of A are non-negative.
(ii) A is positive definite if and only if all eigenvalues of A are positive.

Example 11.38 *The following matrices are positive definite.*

$$A_1 = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} \tag{442}$$

The following matrices are positive semidefinite, but not positive definite.

$$A_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}, \quad A_6 = \begin{pmatrix} 3 & -3 \\ -3 & 3 \end{pmatrix} \tag{443}$$

The following matrices are not positive semidefinite (and therefore, cannot be positive definite).

$$A_7 = \begin{pmatrix} 1 & 2 \\ 0 & 3 \end{pmatrix}, \quad A_8 = \begin{pmatrix} 3 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_9 = \begin{pmatrix} \frac{-1+\sqrt{5}}{2} & 1 \\ 1 & \frac{-1-\sqrt{5}}{2} \end{pmatrix} \tag{444}$$

□

11.10 Complex inner product spaces

When we work over the complex numbers $\mathbb{C}$, we need to alter the definition of inner product a bit. The reason for this is that the square of a complex number need not be a positive real number, or even a real number at all. In order to get a condition on complex inner product spaces that mimics condition (iii) of the definition of a real inner product space, we need to consider complex conjugates.

Definition 11.39 *Let $z = x + yi \in \mathbb{C}$. The complex conjugate of z is the complex number $\overline{z} = x - yi$.*

What happens when you multiply a complex number by its conjugate? Let's see: suppose $z = x + iy \in \mathbb{C}$. Now

$$z\overline{z} = (x + iy)(x - iy) = x^2 + y^2. \tag{445}$$

This is always a real number, and in fact, it is positive if and only if $z \neq 0$. This allows us to get a positivity condition on complex inner product spaces.

We start by introducing the complex dot product.

Definition 11.40 *Let*

$$\overrightarrow{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix} \quad \textit{and } \overrightarrow{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} \tag{446}$$

be vectors in $\mathbb{C}^n$. The complex dot product of u and v is the value

$$\overrightarrow{u} \cdot \overrightarrow{v} = \overline{u_1} v_1 + \overline{u_2} v_2 + ... + \overline{u_n} v_n. \tag{447}$$

This is the most common example of a complex inner product, which we now define.

Definition 11.41 *Let V be a vector space over $\mathbb{C}$. A complex inner product on V is an operation $\langle , \rangle$ that takes two input vectors from V and associates an element of $\mathbb{C}$ as the output, satisfying the following condtions:*
(i) For any $u, v \in V$, $\langle u, v \rangle = \overline{\langle v, u \rangle}$. ("conjugate symmetry")
(ii) For any $u, v, w \in V$ and $s, t \in \mathbb{C}$, $\langle u, sv + tw \rangle = s \langle u, v \rangle + t \langle u, w \rangle$. ("linearity on the right")
(iii) For any $u, v, w \in V$ and $s, t \in \mathbb{C}$, $\langle su + tv, w \rangle = \overline{s} \langle u, w \rangle + \overline{t} \langle v, w \rangle$. ("anti-linearity on the left")
(iii) For any $u \in V$, $\langle u, u \rangle \geq 0$, and $\langle u, u \rangle = 0$ if and only if $u = 0$. ("positive definite")
If $\langle , \rangle$ is a complex inner product on V, we say that V is a complex inner product space.

Example 11.42 *The following are examples of complex inner product spaces.*
(i) $\mathbb{C}^n$, with the complex dot product.

(ii) Let $a, b \in \mathbb{R}$ such that $a < b$. Let

$$\mathcal{C}\left(\left[a,b\right],\mathbb{C}\right) = \left\{f : [a,b] \to \mathbb{C} \middle| f \text{ is continuous}\right\}. \tag{448}$$

Define

$$\langle f, g\rangle = \int_a^b \overline{f\left(x\right)} g\left(x\right)\, \mathrm{d}x. \tag{449}$$

All of the theorems that we previously learned about real inner products also apply to complex inner products. However, there is one idiosyncrasy that should not be ignored: in a complex inner product space, $\langle u, v\rangle$ is, in general, **not** the same as $\langle v, u\rangle$. This means that **ordering matters** if you want to state a theorem about a complex inner product.

Example 11.43 *In the complex inner product space $\mathbb{C}^3$ with the (complex) dot product, let*

$$\overrightarrow{v_1} = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \quad \text{and} \quad \overrightarrow{v_2} = \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}. \tag{450}$$

Use the Gram-Schmidt procedure to find an orthogonal basis for $W = \operatorname{span}\left(\overrightarrow{v_1}, \overrightarrow{v_2}\right)$.

The first two steps of the Gram-Schmidt procedure are

$$\begin{aligned} \overrightarrow{u_1} &= \overrightarrow{v_1} = \begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} \\ \overrightarrow{u_2} &= \overrightarrow{v_2} - \frac{\langle \overrightarrow{u_1}, \overrightarrow{v_2}\rangle}{||\overrightarrow{u_1}||^2}\overrightarrow{u_1} = \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix} - \frac{1}{2}\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix} = \begin{pmatrix} -\frac{1}{2}i \\ \frac{1}{2}i \\ 1 \end{pmatrix} \end{aligned}. \tag{451}$$

Thus, our orthogonal basis is

$$\boxed{\left(\begin{pmatrix} 1 \\ i \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}\right)}. \tag{452}$$

□

Example 11.44 *In the complex inner product space $\mathbb{C}^3$ with the (complex) inner product, let*

$$\overrightarrow{v_1} = \begin{pmatrix} 1-i \\ 1+i \\ i \end{pmatrix}, \quad \textit{and} \quad \overrightarrow{v_2} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}. \tag{453}$$

Use the Gram-Schmidt procedure to find an orthogonal basis for $W = \operatorname{span}(\overrightarrow{v_1}, \overrightarrow{v_2})$.

The first two steps of the Gram-Schmidt procedure are

$$\begin{aligned} \overrightarrow{u_1} = \overrightarrow{v_1} &= \begin{pmatrix} 1-i \\ 1+i \\ i \end{pmatrix} \\ \overrightarrow{u_2} = \overrightarrow{v_2} - \frac{\langle \overrightarrow{u_1}, \overrightarrow{v_2} \rangle}{||\overrightarrow{u_1}||^2} \overrightarrow{u_1} &= \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{2-i}{5} \begin{pmatrix} 1-i \\ 1+i \\ i \end{pmatrix} = \begin{pmatrix} \frac{4+3i}{5} \\ \frac{2-i}{5} \\ \frac{4-2i}{5} \end{pmatrix} \end{aligned}. \tag{454}$$

Thus, our orthogonal basis is

$$\boxed{\left(\begin{pmatrix} 1-i \\ 1+i \\ i \end{pmatrix}, \begin{pmatrix} \frac{4+3i}{5} \\ \frac{2-i}{5} \\ \frac{4-2i}{5} \end{pmatrix} \right)}. \tag{455}$$

□

When working in complex inner product spaces, it's often necessary to take both the transpose and complex conjugate of a matrix. Thus, we have a name for this matrix.

Definition 11.45 *Let $A \in M_{m \times n}(\mathbb{C})$. The <u>adjoint matrix of A</u> is the transpose of the complex conjugate of A.*

Notation The symbols A^* mean "adjoint matrix of A." The definition indicates that $A^* = \left(\overline{A}\right)^T$.

We also have names for the situations that $A = A^*$, and for $A^{-1} = A^*$.

Definition 11.46 *Let $A \in M_{n\times n}(\mathbb{C})$. We say that A is a Hermitian matrix provided that $A^* = A$.*

Definition 11.47 *Let $A \in M_{n\times n}(\mathbb{C})$. We say that A is a unitary matrix provided that $A^* A = I_n$.*