Intro to Linear Algebra: A Summary

William Boyles

Updated January 6, 2022

Contents

Chapter 0

Background & Review

Everything mentioned in this chapter should already be familiar to you from other math classes. These topics span two main areas: algebra/pre-calculus and single variable calculus. Ideas from algebra and pre-calculus will be used either implicitly or with only a passing reference. Topics from single variable calculus will mostly be used in setting up problems.
If you are unfamiliar with anything mentioned, you can use many of the great only resources like Khan Academy, to familiarize yourself before moving forward.

0.1 Algebra and Pre-Calculus

0.1.1 Sets

Definition. *A set A is a collection of distinct elements. Those elements can be anything, like numbers, functions, and even other sets.*

We can define a set by giving its elements, like $A = \{-2, 5, 3\}$ or by describing its properties, like $A = \{x \mid x > 0\}$ where the vertical bar means "such that". If an object x is a member of the set A, we write $x \in A$.

A set A is called a subset of a set B if every element of A is also an element of B. We can write this as $A \subseteq B$. For example, $\{7, 10, 16\} \subseteq \{5, 6, 7, 9, 10, 11, 16\}$. Note that this relation can be strict if there exists at least one element in B that is not also an element of A. Some common sets and their informal definitions are given below:

Set Name	Symbol	Informal Definition
Natural numbers	$\mathbb{N}$	$\{1, 2, 3, \dots\}$
Integers	$\mathbb{Z}$	$\{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$
Rational numbers	$\mathbb{Q}$	$\{\frac{m}{n} \mid m, n \in \mathbb{Z}$ and $n \neq 0\}$
Real numbers	$\mathbb{R}$	Any number on the number line[1]

This means that $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$.

There are several common operations that can be performed on sets. The union $A \cup B$ of two sets A and B is the set of all elements that are elements of A or of B. Similarly, the intersection $A \cap B$ of two sets A and B is the set of all elements that are also elements of both A and B.

Example. *If $A = \{\sqrt{2}, 2, 5, 8\}$ and $B = \{-9, 8, 2.3\}$, what are $A \cup B$ and $A \cap B$?*

To find the union, we combine the sets, making sure to include any repeated element only once:

$$A \cup B = \{-9, \sqrt{2}, 2, 2.3, 5, 8\}.$$

Then, since the only element both sets share is 8, we also have

$$A \cap B = \{8\}.$$

0.1.2 Intervals

Definition. *We call a subset I of $\mathbb{R}$ an interval if, for any $a, b \in I$ and $x \in \mathbb{R}$ such that $a \leq x \leq b$, then $x \in I$.*

We can write an interval more simply using the notation $[a, b]$, which is equivalent to $\{x \in \mathbb{R} \mid a \leq x \leq b\}$. This is called a closed interval, and to make the inequalities strict, we can also define an open interval by using parantheses instead of square brackets.

In addition, we can mix the two to create half-open intervals, where one inequality is strict and the other isn't. For instance, $(2, 5]$ refers to the set $\{x \in \mathbb{R} \mid x < 2 \leq 5\}$ Finally, if the interval is unbounded in either direction, we use the notations $-\infty$ and ∞ to indicate that there is no minimum or maximum, respectively.

Example. *Is $8 \in (-\infty, 4) \cup [8, 100)$?*

Since $8 \leq 8 < 100$ is a true statement, $8 \in [8, 100)$. Since we are taking the union with another set, all of the members of the right interval will also be members of the union of intervals. Therefore, the statement is true.

0.1.3 Functions

Definition. *A function f is a rule between a pair of sets, denoted $f : D \to C$, that assigns values from the first set, the domain D, to the second set, the codomain C.*

We call the subset of the codomain C that constitutes all values f can actually attain the range $R \subseteq C$. Note that when we draw a graph of a function, all we are doing is drawing all ordered pairs $\{(x, f(x)) \mid x \in D\}$.

Example. *Find the domain of the following function:*

$$f(x) = \frac{1}{(1-x)\sqrt{5-x^2}}$$

We know that $\frac{n}{0}$ is undefined for all $n \in \mathbb{R}$ and $\sqrt{x}$ is only defined for $x \geq 0$. The first condition applies to the first term in the denominator and both conditions apply to the second, giving us

$$(1-x) \neq 0 \text{ and } 5 - x^2 > 0$$

The first condition implies $x \neq 1$ while the second implies $|x| < \sqrt{5}$. Putting these together, we find that the domain is

$$\{x \mid x \neq 1, |x| < \sqrt{5}\} \text{ or } (-\sqrt{5}, 1) \cup (1, \sqrt{5})$$

We can also compose two functions, such that the ouput of one function is the input of another:

$$(f \circ g)(x) = f(g(x)).$$

Definition. *A function g is called an inverse function of f if $f(g(x)) = x$ for all x in the domain of g and $g(f(x))$ for all x in the domain of f. We write this as $g = f^{-1}$.*

One common algorithm for finding an inverse function is to set $y = f(x)$, substitute all x's for y's, and then solve for y.

Example. *Find the inverse function of*

$$f(x) = \frac{5x+2}{4x-3}.$$

We first make the substitutions to set up the algorithm:

$$y = \frac{5x+2}{4x-3} \text{ followed by } x = \frac{5y+2}{4y-3}$$

After multiplying both sides by the denominator and simplifying, we have

$$\implies 4xy - 3x = -5y - 2 \implies y = f^{-1}(x) = \frac{3x-2}{4x+5}.$$

We say that a function f is even if it satisfies $f(-x) = f(x)$ for all $x \in D$. Likewise, we say that a function f is odd if it satisfies $f(-x) = -f(x)$ for all $x \in D$. Geometrically, we can see that the graph of an even function is symmetric with respect to the y-axis, while the graph of an odd function is symmetric with respect to the origin.

Example. *Is $f(x) = 2x - x^2$ even, odd, or neither?*

We see that

$$f(-x) = 2(-x) - (-x)^2 = -2x - x^2.$$

Since $f(-x) \neq f(x)$ and $f(-x) \neq -f(x)$, the function is neither even nor odd.

0.1.4 Complex Numbers

Definition. *i is called the imaginary unit. It's defined by $i^2 = -1$.*

The set of complex numbers ($\mathbb{C}$) is an extension of the real numbers. Complex numbers have the form $z = \alpha + \beta i$, where α and β are real numbers. The α part of z is called the real part, so $\Re(z) = \alpha$. The β part of z is called the imaginary part, so $\Im(z) = \beta i$.

Often, complex numbers are visualized as points or vectors in a 2D plane, called the complex plane, where α is the x-component, and β is the y-component. Thinking of complex numbers like points helps us define the magnitude of complex numbers and compare them. Since a point (x, y) has a distance $\sqrt{x^2 + y^2}$ from the origin, we can say the magnitude of z, $|z|$ is $\sqrt{\alpha^2 + \beta^2}$. Thinking of complex numbers like vectors helps us understand adding two complex numbers, since you just add the components like vectors.

A common operation on complex numbers is the complex conjugate. The complex conjugate of $z = \alpha + \beta i$ is $\overline{z} = \alpha - \beta i$. z and $\overline{z}$ are called a conjugate pair.

Conjugate pairs have the following properties. Let $z, w \in \mathbb{C}$ and $n \in \mathbb{Z}$.

$$\begin{aligned}
\overline{z \pm w} &= \overline{z} \pm \overline{w} \\
\overline{zw} &= \overline{z} \cdot \overline{w} \\
\overline{z} &= z \Leftrightarrow z \in \mathbb{R} \\
z\overline{z} &= |z|^2 = |\overline{z}|^2 \\
\overline{\overline{z}} &= z \\
\overline{z}^n &= \overline{z^n} \\
z^{-1} &= \frac{\overline{z}}{|z|^2}
\end{aligned}$$

0.1.5 Factoring Polynomials

We want to break up a polynomial like $f(x) = a_0 + a_1x^1 + \ldots a_nx^n$ into linear factors so that $f(x) = c(x - b_1) \cdot \ldots \cdot (x - b_n)$. This form makes it simple to see that the roots of f, solutions to $f(x) = 0$, are $x = b_1 \ldots b_n$.

For quadratics, $f(x) = ax^2 + bx + c$, there exists a simple formula that will give us both roots, the quadratic formula

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

We can see that when $b^2 - 4ac < 0$, like for $f(x) = x^2 + 5x + 10$, we will get complex roots $\alpha \pm \beta i$. For any polynomial, these roots come in pairs, so if $\alpha + \beta i$ is a root, then so is $\alpha - \beta i$. This means that every conjugate pair $\alpha \pm \beta i$ has a quadratic equation with those roots. Sometimes we will not factor quadratics with complex roots into linear terms.

Although there do exist explicit formulas for finding roots for cubic (degree 3) and quartic (degree 4) equations, they are too long and not useful enough to memorize. When working by hand, we instead use other tricks to find roots.

There are a few useful tricks that can help. If the polynomial doesn't have a constant term, then 0 is a root. If all the coefficients sum to 0, then 1 is a root. For certain polynomials with an even number of terms, like all cubics of the form $ax^3 + bx^2 + cax + cb$ we can factor out a term from the first two and last two terms to get $x^2(ax+b)+c(ax+b) = (ax+b)(x^2+c)$. For other polynomials, we might just try guessing and checking values. However, we need a more efficient way that works in general.

Since we are looking to find linear factors $f(x) = (x-b_1)\cdot\ldots\cdot(x-b_n)$, we can see that the constant term in the polynomial is the product of the roots $b_1 \ldots b_n$. In fact, since the coefficients of polynomials are completely determined by the roots and the leading coefficient, all the coefficients are sums and products of roots. You might remember when factoring quadratics that the coefficient of x term is the sum of the two roots. These rules are called Vieta's formulas.

So, if we have the constant term, we can check all of its integer factors to see if any are roots. For each root, we can divide, using a technique like synthetic division, to continue finding the rest of the roots. This method is especially useful on tests because the roots tend to be integers.

Example. *Factor the polynomial* $x^5 + x^4 - 2x^3 + 4x^2 - 24x$.

We can immediately see that there is no constant term, so $x = 0$ is a root. Now we need to work on factoring $x^4 + x^3 - 2x^2 + 4x - 24$.
The factors of -24 are: -24, -12, -8, -6, -4, -3, -2, -1, 1, 2, 3, 4, 6, 8, 12, and 24. Starting from roots close to 0 and working outwards, we find that $x = 2$ is a root. So, we synthetic divide like so

$$\begin{array}{r|ccccc} x=2 & 1 & 1 & -2 & 4 & -24 \\ & \downarrow & 2 & 6 & 8 & 24 \\ \hline & 1 & 3 & 4 & 12 & |\,0 \end{array}$$

to see that now we need to work on factoring $x^3 + 3x^2 + 4x + 12$. $x^3 + 3x^2 + 4x + 12 = x^2(x+3) + 4(x+3) = (x+3)(x^2+4)$, so $x = -3$ is a root, and we need to work on factoring $x^2 + 4$. $x^2 + 4$ has two complex roots $\pm 2i$, so we'll leave it as a quadratic.

$$x^5 + x^4 - 2x^3 + 4x^2 - 24x = x(x-2)(x-3)(x^2+4)$$

0.1.6 Trig Functions & The Unit Circle

Imagine a circle of radius 1 centered at the origin that we'll call the unit circle. The x and y coordinates of a point on the unit circle are completely determined by the angle θ in radians between the

x-axis and a line from the origin to the point.

The function $\cos\theta$ tells us x-coordinate of the point, while $\sin\theta$ tells us the y-coordinate of the point. The function $\tan\theta = \frac{\sin\theta}{\cos\theta}$ tells us the slope of the line from the origin to the point. Most of the trig functions have geometric interpretations as shown below. The most used ones are sin, cos, $\tan = \frac{\sin}{\cos}$, $\cot = \frac{\cos}{\sin}$, $\csc = \frac{1}{\sin}$, and $\sec = \frac{1}{\cos}$.

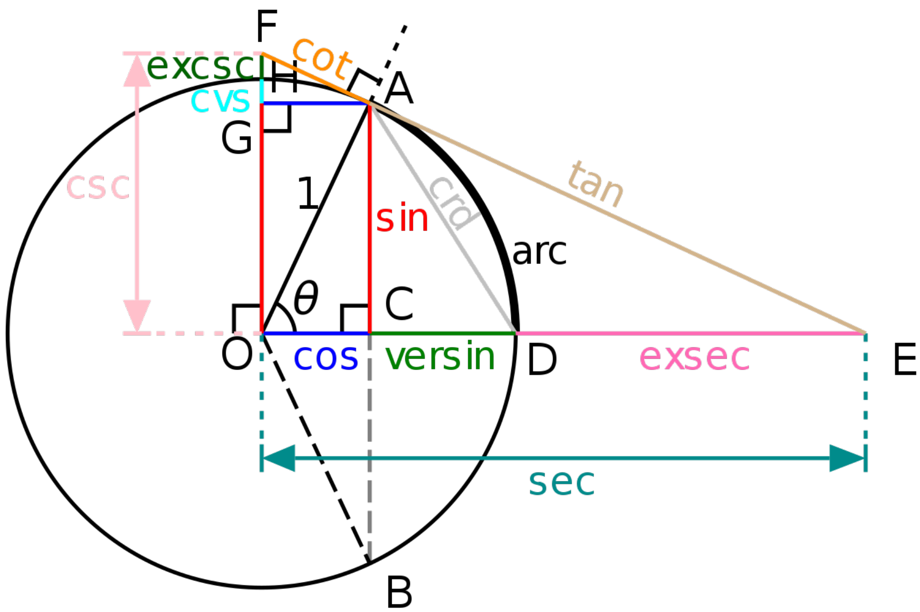

Figure 1: Wikipedia - Unit circle

We can also think about the inverses of these trig functions. These are either notated with a -1 exponent on the function, or the prefix arc in front of the function name. Many of these functions are only defined on a part of the domain $[0, 2\pi]$. Below is a table of the inverse trig functions and their domains.

Function	Domain
arcsin	$[-1, 1]$
arccos	$[-1, 1]$
arctan	$(-\infty, \infty)$
arccot	$(-\infty, \infty)$
arccsc	$(-\infty, -1] \cup [1, \infty)$
arcsec	$(-\infty, -1] \cup [1, \infty)$

0.1.7 Trig Identities

As we could see in Figure 0.1.6, sin and cos form a right triangle with hypotenuse 1. So, using the Pythagorean Theorem,

$$\sin^2\theta + \cos^2\theta = 1.$$

By dividing by $\sin^2$ or $\cos^2$, we can also get

$$1 + \cot^2\theta = \csc^2\theta \text{ and } \tan^2\theta + 1 = \sec^2\theta.$$

Together, these 3 identities are called the Pythagorean Identities.

We can also relate functions and co-functions.

$$\text{xxx}(\theta) = \text{coxxx}\left(\frac{\pi}{2} - \theta\right).$$

Some of the most useful and used identities are the sum and difference.

$$\begin{aligned}
\sin(\alpha \pm \beta) &= \sin\alpha\cos\beta \pm \cos\alpha\sin\beta \\
\cos(\alpha \pm \beta) &= \cos\alpha\cos\beta \mp \sin\alpha\sin\beta \\
\tan(\alpha \pm \beta) &= \frac{\tan\alpha \pm \tan\beta}{1 \mp \tan\alpha\tan\beta} \\
\sin\alpha \pm \sin\beta &= 2\sin\left(\frac{\alpha \pm \beta}{2}\right)\cos\left(\frac{\alpha \mp \beta}{2}\right) \\
\cos\alpha + \cos\beta &= 2\cos\left(\frac{\alpha + \beta}{2}\right)\cos\left(\frac{\alpha - \beta}{2}\right) \\
\cos\alpha - \cos\beta &= -2\sin\left(\frac{\alpha + \beta}{2}\right)\sin\left(\frac{\alpha - \beta}{2}\right)
\end{aligned}$$

0.1.8 Exponentials & Logarithms

Definition. *e is the base of the natural logarithm. It's defined by the limit*

$$e = \lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n.$$

The functions $\exp x = e^x$ and $\ln x$ are inverses of each other:

$$e^{\ln x} = x \text{ and } \ln e^x = x.$$

Just like other exponentials, the normal rules for adding, subtracting, and multiplying exponents apply:

$$e^x e^y = e^{x+y},\ \frac{e^x}{e^y} = e^{x-y}, \text{ and } (e^x)^k = e^{xk}.$$

Similar rules apply for logarithms:

$$\ln x + \ln y = \ln xy,\ \ln x - \ln y = \ln\left(\frac{x}{y}\right), \text{ and } \ln\left(a^b\right) = b\ln a.$$

We can also write a logarithm of any base using natural logarithms:

$$\log_b a = \frac{\ln a}{\ln b}.$$

The number e is also unique in that it is the only real number a satisfying the equation

$$\frac{\mathrm{d}}{\mathrm{d}x} a^x = a^x,$$

meaning e^x is its own derivative.

0.1.9 Partial Fractions

If we have a function of two polynomials $f(x) = \frac{P(x)}{Q(x)}$, it's often easier to break this quotient into a sum of parts where the denominator is a linear or quadratic factor and the numerator is always a smaller degree than the denominator.

Example.

$$\frac{2x-1}{x^3-6x^2+11x-6} = \frac{1/2}{x-1} + \frac{-3}{x-2} + \frac{5/2}{x-3}.$$

One natural way to find these small denominators comes from the linear factors of the denominator where we keep quadratics with complex roots. This way, when making a common denominator, we get back the original big denominator. However, there are a few special cases we have to take care of.

Linear Factors

This is the the most basic type where the degree of the numerator is less than the degree of the denominator and the denominator factors into all linear factors with no repeated roots. In this case we can write

$$\frac{P(x)}{Q(x)} = \frac{A_1}{(x-a_1)} + \ldots + \frac{A_n}{(x-a_n)}.$$

Multiplying each side by $Q(x)$,

$$P(x) = A_1(x-a_2)\ldots(x-a_n) + \ldots + A_n(x-a_1)\ldots(x-a_{n-1}).$$

We can then find each A_i by evaluating both sides at $x = a_i$, since every term except the ith has an $(x-a_i)$ factor that will go to 0. So,

$$A_i = \frac{P(a_i)}{(x-a_i)\ldots(x-a_{i-1})(x-a_{i+1})\ldots(x-a_n)}.$$

Example. *Find the partial fraction decomposition of the following expression:*

$$\frac{2x-1}{x^3-6x^2+11x-6}.$$

Factoring,

$$x^3 - 6x^2 + 11x - 6 = (x-1)(x-2)(x-3).$$

So,

$$\frac{2x-1}{x^3 - 6x^2 + 11x - 6} = \frac{A_1}{x-1} + \frac{A_2}{x-2} + \frac{A_3}{x-3}.$$

Multiplying each side by the denominator,

$$2x - 1 = A_1(x-2)(x-3) + A_2(x-1)(x-3) + A_3(x-1)(x-2).$$

At $x = 1$,

$$1 = A_1(1-2)(1-3) \implies A_1 = \frac{1}{2}.$$

At $x = 2$,

$$3 = A_2(2-1)(2-3) \implies A_2 = -3.$$

At $x = 3$,

$$5 = A_3(3-1)(3-2) \implies A_3 = \frac{5}{2}.$$

So,

$$\frac{2x-1}{x^3 - 6x^2 + 11x - 6} = \frac{1/2}{x-1} + \frac{-3}{x-2} + \frac{5/2}{x-3},$$

just as was shown in the previous example.

Repeated Linear Factors

If $Q(x)$ has repeated roots, it factors into

$$Q(x) = R(x)(x-a)^k,\ k \geq 2 \text{ and } R(a) \neq 0.$$

When making the common denominator for each repeated root of multiplicity k, we do

$$\frac{P(x)}{R(x)(x-a)^k} = (\text{Decomposition of } R(x)) + \frac{A_1}{x-a} + \ldots + \frac{A_k}{(x-a)^k}.$$

You would then multiply each side by the denominator like in the linear factors case and solve for the coefficients. The only additional difficulty is that you might have to use previous results or solve a system of linear equations to get some of the constants.

Example. *Find the partial fraction of the following expression:*

$$\frac{x^2 + 5x - 6}{x^3 - 7x^2 + 16x - 12}.$$

Factoring,

$$x^3 - 7x^2 + 16x - 12 = (x-3)(x-2)^2.$$

So,

$$\frac{x^2 + 5x - 6}{x^3 - 7x^2 + 16x - 12} = \frac{A_1}{x-3} + \frac{A_2}{x-2} + \frac{A_3}{(x-2)^2}.$$

Multiplying each side by the denominator,

$$x^2 + 5x - 6 = A_1(x-2)^2 + A_2(x-2)(x-3) + A_3(x-3).$$

At $x = 2$,

$$8 = A_3(2-3) \implies A_3 = -8.$$

At $x = 3$,

$$18 = A_1(3-2)^2 \implies A_1 = 18.$$

Now we'll use our results for A_1 and A_3 to find A_2 using a value for x that isn't 2 or 3 so the A_2 term doesn't become 0. A good choice is $x = 0$.
At $x = 0$,

$$-6 = 18(0-2)^2 + A_2(0-2)(0-3) + -8(0-3) \implies A_2 = -17.$$

So,

$$\frac{x^2 + 5x - 6}{x^3 - 7x^2 + 16x - 12} = \frac{18}{x-3} - \frac{17}{x-2} - \frac{8}{(x-2)^2}.$$

Quadratic Factors

If a quadratic doesn't have real roots, then we have a quadratic factor. Here, we'll assume that the quadratic factor isn't repeated. So, $Q(x) = R(x)(ax^2 + bx + c)$, $b^2 - 4ac < 0$, and $R(x)$ is not evenly divisible by $ax^2 + bx + c$. In this case, we say

$$\frac{P(x)}{R(x)(ax^2 + bx + c)} = (\text{Decomposition of } R(x)) + \frac{A_1 x + B_1}{ax^2 + bx + c}.$$

We then solve for the constants in the numerator, possibly having to solve a system of equations or using previous results and less convenient values for x.

Example. *Find the partial fraction decomposition of the following expression:*

$$\frac{6x^2 + 21x + 11}{x^3 + 5x^2 + 3x + 15}.$$

Factoring,

$$x^2 + 5x^2 + 3x + 15 = (x+5)(x^2+3).$$

So,

$$\frac{6x^2 + 21x + 11}{x^3 + 5x^2 + 3x + 15} = \frac{A_1}{x+5} + \frac{A_2 x + B_2}{x^2 + 3}.$$

Multiplying each side by the denominator,

$$6x^2 + 21x + 11 = A_1(x^2 + 3) + (A_2x + B_2)(x + 5).$$

At $x = -5$,

$$56 = 28A_1 \implies A_1 = 2.$$

Now we'll use the previous result and another value for x. We can use $x = 0$ to not have to worry about the A_2 term. At $x = 0$,

$$11 = 2(3) + (B_2)(5) \implies B_2 = 1.$$

Now we'll use the previous 2 results to find A_2. $x = 1$ is a good choice to keep the numbers small. At $x = 1$,

$$38 = 2(1 + 3) + (A_2 + 1)(6) \implies A_2 = 4.$$

So,

$$\frac{6x^2 + 21x + 11}{x^3 + 5x^2 + 3x + 15} = \frac{2}{x + 5} + \frac{4x + 1}{x^2 + 3}.$$

Repeated Quadratic Factors

If a quadratic factor that can't be broken into linear factors is repeated, then we can write $Q(x) = R(x)(ax^2 + bx + c)^k$, $k \geq 0$, and $R(x)$ is not divisible by $(ax^2 + bx + c)^k$. Now we have to do a combination of what we did for repeated linear factors and quadratic factors. We say

$$\frac{P(x)}{R(x)(ax^2 + bx + c)^k} = (\text{Decomposition of } R(x)) + \frac{A_1x + B_1}{ax^2 + bx + c} + \ldots + \frac{A_kx + B_k}{(ax^2 + bx + c)^k}.$$

We then solve for the coefficients in the numerator.

Example. *Find the partial fraction decomposition of the following expression:*

$$\frac{3x^4 - 2x^3 + 6x^2 - 3x + 3}{x^5 + 3x^4 + 4x^3 + 12x^2 + 4x + 12}.$$

Factoring,

$$x^5 + 3x^4 + 4x^3 + 12x^2 + 4x + 12 = (x + 3)(x^2 + 2)^2.$$

So,

$$\frac{3x^4 - 2x^3 + 6x^2 - 3x + 3}{x^5 + 3x^4 + 4x^3 + 12x^2 + 4x + 12} = \frac{A_1}{x + 3} + \frac{A_2x + B_2}{x^2 + 2} + \frac{A_3x + B_3}{(x^2 + 2)^2}.$$

Multiplying each side by the denominator,

$$3x^4 - 2x^3 + 6x^2 - 3x + 3 = A_1(x^2 + 2)^2 + (A_2x + B_2)(x^2 + 2)(x + 3) + (A_3x + B_3)(x + 3).$$

At $x = -3$,

$$363 = 121A_1 \implies A_1 = 3.$$

Now, we'll use our result for A_1 and pick a value for x that minimizes the number of things we need to solve for. We'll have to solve a linear system with 4 unknowns, so we'll need up to 4 values. At $x = 0$,

$$3 = 3(2)^2 + B_2(2)(3) + B_3(3) \implies 2B_2 + B_3 = -3.$$

At $x = 1$,

$$7 = 3(3)^2 + (A_2 + B_2)(3)(4) + (A_3 + B_3)(4) \implies 3A_2 + A_3 + 3B_2 + B_3 = -5.$$

At $x = -1$,

$$17 = 3(3)^2 + (-A_2 + B_2)(3)(2) + (-A_3 + B_3)(2) \implies -3A_2 - A_3 + 3B_2 + B_3 = -5.$$

At $x = 2$,

$$53 = 3(6)^2 + (2A_2 + B_2)(6)(5) + (2A_3 + B_3)(5) \implies 12A_2 + 2A_3 + 6B_2 + B_3 = -11.$$

Now we have the following system of equations:

$$\begin{cases} 0A_2 + 0A_3 + 2B_2 + B_3 & = -3 \\ 3A_2 + A_3 + 3B_2 + B_3 & = -5 \\ -3A_2 - A_3 + 3B_2 + B_3 & = -5 \\ 12A_2 + 2A_3 + 6B_2 + B_3 & = -11 \end{cases}.$$

Solving,

$$A_2 = 0, A_3 = 0, B_2 = -2, \text{ and } B_3 = 1.$$

So,

$$\frac{3x^4 - 2x^3 + 6x^2 - 3x + 3}{x^5 + 3x^4 + 4x^3 + 12x^2 + 4x + 12} = \frac{3}{x+3} - \frac{2}{x^2+2} + \frac{1}{(x^2+2)^2}.$$

Improper Fractions

If the degree of the numerator is greater than or equal to the degree of the denominator, we have a case of improper fractions. In this case, we have to do polynomial long division to get a quotient and remainder and then decompose the remainder if necessary. So,

$$\frac{P(x)}{Q(x)} = R(x) + \frac{S(x)}{Q(x)}.$$

Example. *Find the partial fraction decomposition of the following expression:*

$$\frac{x^3 + 3}{x^2 - 2x - 3}.$$

First we do polynomial long division to find that

$$\frac{x^3+3}{x^2-2x-3} = x+2+\frac{7x+9}{x^2-2x-3}.$$

Now that the numerator is of a lesser degree than the denominator, we can decompose it normally.

$$x^2-2x-3 = (x-3)(x+1).$$

So,

$$\frac{7x+9}{x^2-2x-3} = \frac{A_1}{x-3}+\frac{A_2}{x+1}.$$

Multiplying each side by the denominator,

$$7x+9 = A_1(x+1)+A_2(x-3).$$

At $x=-1$,

$$2 = -4A_2 \implies A_2 = \frac{-1}{2}.$$

At $x=3$,

$$30 = 4A_1 \implies A_1 = \frac{15}{2}.$$

So,

$$\frac{x^3+3}{x^2-2x-3} = x+2+\frac{15/2}{x-3}+\frac{-1/2}{x+1}.$$

0.2 Single Variable Calculus

0.2.1 Derivatives and Integrals

Derivatives

The derivative of a function $y=f(x)$, notated $f'(x)$, gives the slope of the tangent line to f at x.

Definition.

$$f'(x) = \lim_{h\to 0}\frac{f(x+h)-f(x)}{h}.$$

Below are some properties of the derivative. Let f and g be functions of x and p a scalar.

- **Linearity**

$$(pf\pm g)' = pf'\pm g'$$

- **Product Rule**

$$(fg)' = f'g+fg'$$

- **Quotient Rule**

$$\left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}$$

- **Chain Rule**

$$(f \circ g)' = (f' \circ g) \cdot g'$$

- **Power Rule**

$$\frac{\mathrm{d}}{\mathrm{d}x} p^x = px^{p-1}, p \neq 0$$

- **Exponent Rule**

$$\frac{\mathrm{d}}{\mathrm{d}x} p^x = p^x \ln p, p > 0$$

The Power Rule and Exponent Rule are two cases of the same rule

$$\frac{\mathrm{d}}{\mathrm{d}x} f^g = gf^{g-1}f' + f^g \ln(f)g'.$$

Using the definition of the derivative and these rules, we can find the derivatives to some common functions.

$$\begin{array}{r|l}
\frac{\mathrm{d}}{\mathrm{d}x} p = 0 & \frac{\mathrm{d}}{\mathrm{d}x} e^x = e^x \\
\frac{\mathrm{d}}{\mathrm{d}x} \ln x = \frac{1}{x} & \frac{\mathrm{d}}{\mathrm{d}x} \sin x = \cos x \\
\frac{\mathrm{d}}{\mathrm{d}x} \cos x = -\sin x & \frac{\mathrm{d}}{\mathrm{d}x} \tan x = \sec^2 x
\end{array}$$

Integrals

The definite integral of a function $f(x)$ from $x = a$ to $x = b$ where $a \leq b$ is the area between $f(x)$ and the x-axis bounded by the lines $x = a$ and $x = b$ where area above the x-axis is positive, and area below the x-axis is negative.

Definition.

$$\int_a^b f(x)\mathrm{d}x = \lim_{h \to 0} \sum_{n=1}^{\frac{b-a}{h}} f(a + (n-1)h) \cdot h.$$

We also define an indefinite integral, or antiderivative of $f(x)$, notated $F(x)$ where

$$F'(x) = f(x) \implies \int f(x)\mathrm{d}x = F(x).$$

Note that there are infinitely many such functions F, since adding a constant to F does not affect its derivative. To notate this, we add a constant C to the indefinite integral. Given an initial condition for f, we can solve for C.

Below are some properties of the integral. Let f and g be functions of x and p, a, b, and c where $a < b < c$, and f and g are continuous on the closed interval $[a, c]$.

- **Linearity**

$$\int (pf \pm g)\mathrm{d}x = p \int f\mathrm{d}x \pm \int g\mathrm{d}x$$

- **Flipped Bounds**

$$\int_a^b f\mathrm{d}x = -\int_b^a f\mathrm{d}x$$

- **Union of Intervals**

$$\int_a^b f\mathrm{d}x + \int_b^c f\mathrm{d}x = \int_a^c f\mathrm{d}x$$

- **Power Rule**

$$\int x^n\mathrm{d}x = \frac{x^{n+1}}{n+1} + C, n \neq -1$$

- **U-Substitution**

$$\int (f' \circ g)\, g'\mathrm{d}x = f \circ g + C$$

- **Integration by Parts**

$$\int f'g\mathrm{d}x = fg - \int fg'\mathrm{d}x$$

- **Fundamental Theorem of Calculus**

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^x f(s)\mathrm{d}s = f(x)$$

Using the definition of the integral and the above rules, we can find the indefinite integral of some common functions.

$$\int \frac{1}{x}\mathrm{d}x = \ln\left|x\right| + C$$
$$\int \sin x\mathrm{d}x = -\cos x + C$$
$$\int \cos x\mathrm{d}x = \sin x + C$$
$$\int \tan x\mathrm{d}x = -\ln\left|\cos x\right| + C$$

0.2.2 Taylor Series

A Taylor series as a way of approximating a function about a point $x = a$ using polynomials. The first approximation just keeps the same value at $x = a$, the second approximation keeps the same value and first derivative at $x = a$, etc.

Definition.

$$f(x) = f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots$$

If we approximate a function about $x = 0$, we call this a Maclaurin series. Below are some common Maclaurin series, and their radii of convergence if applicable.

$$\begin{aligned}
e^x &= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots \\
\sin x &= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots \\
\cos x &= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots \\
\frac{1}{1+x} &= 1 - x + x^2 - \ldots, \text{ where } |x| < 1 \\
\ln(1+x) &= x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots, \text{ where } |x| < 1
\end{aligned}$$

Euler's Identity

Let's see what happens when we look at the Maclaurin series for e^{ix}.

$$\begin{aligned}
e^{ix} &= 1 + (ix) + \frac{(ix)^2}{2!} + \frac{(ix)^3}{3!} + \frac{(ix)^4}{4!} + \frac{(ix)^5}{5!} \ldots \\
&= 1 + ix - \frac{x^2}{2!} - i\frac{x^3}{3!} + \frac{x^4}{4!} + i\frac{x^5}{5!} - \ldots \\
&= \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \ldots\right) + i\left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ldots\right).
\end{aligned}$$

The two expressions in parenthesis are exactly the Maclaurin series for $\cos x$ and $\sin x$. So,

$$e^{ix} = \cos x + i \sin x.$$

In the case that $x = \pi$,

$$e^{i\pi} = \cos \pi + i \sin \pi = -1 + 0.$$

So,

$$e^{i\pi} + 1 = 0.$$

Chapter 1

Vector Spaces

Linear algebra deals with vector spaces and special maps between them. Here we'll define a vector space and provide its most important properties.

1.1 Binary Operations

1.1.1 Definition & Notations

Definition. *For some domain D, A binary operation is a function $f : D \times D \to D$. That is, it combines two elements of D to produce another element of D.*

You've already been working with binary operations since elementary school.

Example. *The following are binary operations with domain $\mathbb{R}$.*

- *Addition $(+)$*
- *Multiplication $(\times)$*

Subtraction $(-)$ is also a binary operation on $\mathbb{R}$. It is just a special case of addition, since for all $a, b \in \mathbb{R}$, $a - b = a + (-b)$, and $-b \in \mathbb{R}$.

One needs to specify over which domain a function is a binary operation. Functions can be binary operations over some domains but not others.

Example.

- *Division $(\div)$ is not a binary operation on $\mathbb{R}$ since $0 \in \mathbb{R}$, but dividing by 0 is not defined. However if we consider the domain $\mathbb{R} \setminus \{0\}$, then division is a binary operation, since we no longer have to worry about dividing by 0.*
- *Although we saw that subtraction was a binary operation over $\mathbb{R}$, it is not over $D = \{x \in \mathbb{R} \mid x \geq 0\}$ because $1, 2 \in D$, but $1 - 2 = -1 \notin D$.*

For many binary operations, we tend to abandon function notation in favor of an "infix" notation. For example, we write $a+b$ or $a \times b$ rather than $+(a,b)$ or $\times(a,b)$. Sometimes, if the operation is clear from context, we may drop the symbol for the operation entirely and just write the inputs next to each other. For example, if $a,b \in \mathbb{R}$, then ab is understood to mean $a \times b$.

1.1.2 Properties

A binary operation may have certain useful properties. Let $\circ$ and $\square$ be two binary operations with the same domain. Let a, b, and c be elements of the domain.

- **Commutative**: $a \circ b = b \circ a$
- **Associative**: $(a \circ b) \circ c = a \circ (b \circ c)$
- **Distributive over** $\square$: $a \circ (b \square c) = (a \circ b) \square (a \circ c)$
- **Identity**: There exists I in the domain such that $a \circ I = a = I \circ a$
- **Inverse**: There exists a^{-1} in the domain such that $a \circ a^{-1} = I = a^{-1} \circ a$

Many of the binary operations you're already familiar with satisfy some of these properties.

Example.

- *Addition over $\mathbb{R}$ is commutative and associative. Its identity element is* 0*, and the inverse of a is $-a$.*
- *Multiplication over $\mathbb{R}$ is commutative and associative. It is distributive over addition. Its identity element is* 1*. It is not invertable, since 0 doesn't have an inverse.*
- *Division over $\mathbb{R} \setminus \{0\}$ is not commutative, since* $1/2 \neq 2/1$*. It is not associative, since* $(1 \div 2) \div 3 = 1/6 \neq 1 \div (2 \div 3) = 3/2$*.*

1.2 Euclidean Vectors

1.2.1 Definition & Notations

Definition. *A Euclidean vector is a finite, ordered list of real numbers. Each entry in the vector is called a component. The number of components in a vector is called its dimension. The set of all Euclidean vectors with dimension n is notated $\mathbb{R}^n$ and called "Euclidean n-space".*

Often, vectors are written in square brackets with their components ordered vertically from top to bottom or in angled brackets with their components ordered horizontally from left to right. We tend to symbolically represent some vector by a symbol with an arrow above it. We can represent a certain component of a vector with a subscript.

Example. *The vector* $\vec{v} = \begin{bmatrix} 4 \\ 5 \end{bmatrix} = \langle 4, 5 \rangle$ *is a vector in* $\mathbb{R}^2$*. The components are* $\vec{v}_1 = 4$ *and* $\vec{v}_2 = 5$*. Since* $\vec{v}$ *can be visualized as a straight arrow in the 2D plane with its tail at the origin and its tip at* $(4, 5)$*, we might also notate the x and y components as* $\vec{v}_x = 4$ *and* $\vec{v}_y = 5$*.*

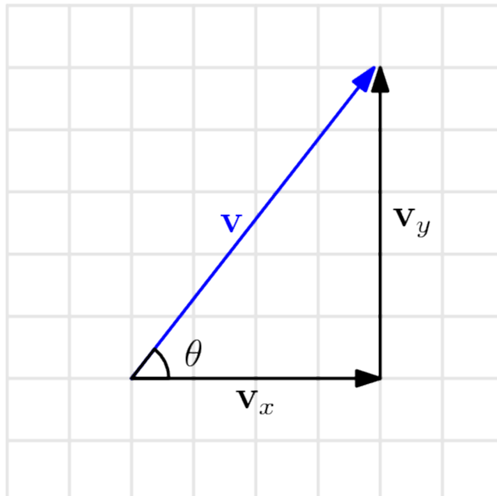

Figure 1.1: The x and y components of a vector v

1.2.2 Properties

We can define addition and scalar multiplication of Euclidean vectors.

Definition. *Let* $\vec{a}, \vec{b} \in \mathbb{R}^n$ *and* $c \in \mathbb{R}$ *then*

1. $$\vec{a} + \vec{b} = [a_i + b_i \mid i \in \{1, 2, \dots, n\}],$$

2. $$c \cdot \vec{a} = [c \cdot a_i \mid i \in \{1, 2, \dots, n\}].$$

Example. *Let* $\vec{u} = \begin{bmatrix} 1 \\ -2 \\ 3 \end{bmatrix}$, $\vec{v} = \begin{bmatrix} -1 \\ 4 \\ 3 \end{bmatrix}$, *and* $\vec{w} = \begin{bmatrix} 4 \\ 2 \\ 6 \end{bmatrix}$*. Find* $(2\vec{u} + \vec{v}) - 3\vec{w}$*.*

We find

$$\begin{aligned}
(2\vec{u}+\vec{v})-3\vec{w} &= \left(2\begin{bmatrix}1\\-2\\3\end{bmatrix}+\begin{bmatrix}-1\\4\\3\end{bmatrix}\right)-3\begin{bmatrix}4\\2\\6\end{bmatrix}\\
&= \left(\begin{bmatrix}2\\-4\\6\end{bmatrix}+\begin{bmatrix}-1\\4\\3\end{bmatrix}\right)+\begin{bmatrix}-12\\-6\\-18\end{bmatrix}\\
&= \begin{bmatrix}1\\0\\9\end{bmatrix}+\begin{bmatrix}-12\\-6\\-18\end{bmatrix}\\
&= \begin{bmatrix}-11\\-6\\-9\end{bmatrix}.
\end{aligned}$$

Notice that addition of vectors is a binary operation over $\mathbb{R}^n$. Also, many of the properties we said binary operations could have are satisfied.

Theorem. *The addition of two vectors in $\mathbb{R}^n$ is commutative and associative. The identity vector is the one containing all 0's. The inverse of any vector $\vec{v}$ is $-1\cdot\vec{v}=-\vec{v}$. Vector-scalar multiplication*[1] *is distributive over vector addition.*

1.3 Abstract Vector Spaces

1.3.1 Definition

Definition. *A set V is a vector space of a field F (for us, $\mathbb{R}$ or $\mathbb{C}$) provided that the operations $\oplus : V\times V\to V$, called "addition", and $\odot : V\times F\to V$, called "scalar multiplication", are defined and satisfy the following for all $\vec{u},\vec{v},\vec{w}\in V$ and $a,b\in F$:*

Addition Properties

1. ***Commutative:*** $\vec{u}\oplus\vec{v}=\vec{v}\oplus\vec{u}$
2. ***Associative:*** $\vec{u}\oplus(\vec{v}\oplus\vec{w})=(\vec{u}\oplus\vec{v})\oplus\vec{w}$
3. ***Additive Identity**: There exists $\vec{0}\in V$ such that for all $\vec{u}\in V$, $\vec{u}\oplus\vec{0}=\vec{u}$.*
4. ***Additive Inverse**: For each $\vec{u}\in V$, there exists $-\vec{u}\in V$ such that $\vec{u}\oplus-\vec{u}=\vec{0}$.*

Multiplication Properties

1. ***Multiplicatative Identity**: There exists $1\in F$ such that for all $\vec{u}\in V$, $1\odot\vec{u}=\vec{u}$.*

[1] Vector-scalar multiplication is not a binary operation, but the distributive property still holds

Distributive Properties

1. ***Left Distributes Over*** $\oplus$*:* $a \odot (\vec{u} \oplus \vec{v}) = (a \odot \vec{u}) \oplus (a \odot \vec{v})$
2. ***Right Distributes Over*** $\oplus$*:* $(\vec{u} \oplus \vec{v}) \odot a = (\vec{u} \odot a) \oplus (\vec{v} \odot a)$
3. ***Distributes Over Field Multiplication****:* $(ab) \odot \vec{u} = a \odot (b \odot \vec{u})$

Example. *Here are a few examples of vector spaces:*

- $\mathbb{R}^n$ *with operations of vector addition and vector-scalar multiplication over the field* $\mathbb{R}$
- $\mathbb{R}^2$ *where*
$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \oplus \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} u_1 + v_1 + 1 \\ u_2 + v_2 + 1 \end{bmatrix} \text{ and } a \odot \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} = \begin{bmatrix} au_1 + 1 \\ au_2 - 1 \end{bmatrix}$$
- *All polynomials of degree n with real coefficients*
- $\mathbb{Q}$ *over the field* $\mathbb{Q}$
- *All real functions that are continuous on* $[0, 1]$

Not all operations for $\oplus$ and $\odot$ will result in vector spaces.

Example. $\mathbb{R}$ *where*
$$a \oplus b = 2a + 2b \text{ and } k \odot a = ka$$
is not a vector space because
$$2 \oplus (3 \oplus 4) = 2 \oplus 14 = 32,$$
$$(2 \oplus 3) \oplus 4 = 10 \oplus 4 = 28,$$
so $\oplus$ *isn't associative.*

1.3.2 Properties

Here are several useful properties of vector spaces. Most of these likely seem intuitively true.

Let $(V, +, \cdot)$ be a vector space.

Theorem. *The additive identity is unique.*

Proof. We will show that any two elements of V that behave like the additive identity must be equal. Suppose V has two additive identities $\vec{0}$ and $\vec{0}'$. Since both are additive identities,
$$\vec{0} = \vec{0} + \vec{0}'$$
$$\vec{0}' = \vec{0} + \vec{0}'.$$
Since the expressions on the right of each line are equal, the expressions on the left must also be equal. So $\vec{0} = \vec{0}'$, as desired. ■

Theorem. *The additive inverse of an element is unique.*

Proof. Suppose $\vec{v}$ and $\vec{v'}$ are both additive inverses of $\vec{u}$. Then

$$\begin{aligned}\vec{u}+\vec{v'} = \vec{0} \text{ and } \vec{u}+\vec{v} &= \vec{0} \\ \vec{v} &= \vec{v}+\vec{0} \\ &= \vec{v}+\left(\vec{u}+\vec{v'}\right) \\ &= (\vec{v}+\vec{u})+\vec{v'} \\ &= \vec{0}+\vec{v'} \\ &= \vec{v'},\end{aligned}$$

as desired. ■

Theorem. *For all $\vec{v} \in V$, $0 \cdot \vec{v} = \vec{0}$.*

Proof. Notice that

$$\begin{aligned}0 \cdot \vec{v} &= (0+0) \cdot \vec{v} \\ &= 0 \cdot \vec{v} + 0 \cdot \vec{v}.\end{aligned}$$

Let $\vec{w}$ be the additive inverse of $0 \cdot \vec{v}$. Then

$$\begin{aligned}0 \cdot \vec{v} + \vec{w} &= 0 \cdot \vec{v} + 0 \cdot \vec{v} + \vec{w} \\ \vec{0} &= 0 \cdot \vec{v} + \vec{0} \\ \vec{0} &= 0 \cdot \vec{v},\end{aligned}$$

as desired. ■

Theorem. *For all $a \in \mathbb{R}$, $a \cdot \vec{0} = \vec{0}$.*

Proof. Notice that

$$\begin{aligned}a \cdot \vec{0} &= a \cdot \left(\vec{0}+\vec{0}\right) \\ &= a \cdot \vec{0} + a \cdot \vec{0}.\end{aligned}$$

Let $\vec{w}$ be the additive inverse of $a \cdot \vec{0}$. Then

$$\begin{aligned}a \cdot \vec{0} + \vec{w} &= a \cdot \vec{0} + a \cdot \vec{0} + \vec{w} \\ \vec{0} &= a \cdot \vec{0} + \vec{0} \\ \vec{0} &= a \cdot \vec{0},\end{aligned}$$

as desired. ■

Theorem. *For all $\vec{v} \in V$, $-1 \cdot \vec{v} = -\vec{v}$.*

Proof. Notice,

$$\begin{aligned}
-1 \cdot \vec{v} &= -1 \cdot \vec{v} + \vec{0} \\
&= -1 \cdot \vec{v} + (\vec{v} - \vec{v}) \\
&= (-1 \cdot \vec{v} + \vec{v}) - \vec{v} \\
&= (-1 + 1)\, \vec{v} - \vec{v} \\
&= \vec{0} - \vec{v} \\
&= -\vec{v},
\end{aligned}$$

as desired. ■

1.3.3 Subspaces

Definition. *A non-empty subset U of a vector space V is a subspace of V if U is a vector space with the operations inherited from V.*

Example. *Here are a few examples of subspaces:*

- *The set of vectors* $\left\{ \begin{bmatrix} x \\ y \\ 0 \end{bmatrix} \mid x, y \in \mathbb{R} \right\}$ *is a subspace of* $\mathbb{R}^3$.
- *The set of real-valued functions* $\{f \in \mathbb{R}^{[0,1]} \mid \int_0^1 f(x)\mathrm{d}x = 0\}$ *is a subspace of all real-valued functions that are integrable over* $[0, 1]$, $\mathbb{R}^{[0,1]}$.
- *The set of functions* $\{f : [0, 1] \to \mathbb{R} \mid f'(1/2) = 0\}$ *is a subspace of* $\mathbb{R}^{[0,1]}$.

Example. *Here are a few non-examples of subspaces:*

- *The set of vectors* $W = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x, y, z \in \mathbb{R}, 2x + y - z = -4 \right\}$ *is not a subspace of* $\mathbb{R}^3$ *because* $\vec{0} \notin W$*, meaning there is no additive identity.*
- *The set of real-valued functions* $\{f \in \mathbb{R}^{[0,1]} \mid \int_0^1 f(x)\mathrm{d}x = 1\}$ *is not a subspace of* $\mathbb{R}^{[0,1]}$ *because it is not closed under addition.*
- *The set of functions* $\{f : [0, 1] \to \mathbb{R} \mid f'(1/2) = 1\}$ *is not a subspace of* $\mathbb{R}^{[0,1]}$ *because it is not closed under scalar multiplication.*

It would be tedious to recheck all the vector space conditions when seeing if a set is a subspace. Thankfully, we have an easier subspace test

Theorem. *Let U be a non-empty subset of a vector space V. U is a subspace of V if and only if U is closed under addition and scalar multiplication.*

Proof. Assume that U is closed under addition and scalar multiplication. Since $U \subseteq V$, the addition properties of commutativity and associativity hold, the multiplicative identity is in U, and the distributive properties all hold. Thus, all that remains to check is whether the additive inverse exists, and whether each element of U has an additive inverse.

Let $\vec{u} \in U$. Since $U \subset V$ and V is a vector space, there exists $-\vec{u} \in V$. By previous result, we know that $-\vec{u} = -1 \cdot \vec{u}$. Since U is closed under scalar multiplication, $-\vec{u} \in U$. Further, since $\vec{u} + -\vec{u} = \vec{0}$, and U is closed under addition, $\vec{0} \in U$. ■

Chapter 2

Finite Dimensional Vector Spaces

If you've worked with vectors in $\mathbb{R}^2$ before, you might have noticed that we can write any vector using only two vectors and scalar multiplication: one to represent the x component and another to represent the y component.

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = v_x \begin{bmatrix} 1 \\ 0 \end{bmatrix} + v_y \begin{bmatrix} 0 \\ 1 \end{bmatrix}.$$

Although $\langle 1, 0 \rangle$ and $\langle 0, 1 \rangle$ are probably the most convenient vectors to use to break up any vector, other representations are possible.

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \frac{v_y + v_x}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + \frac{v_y - v_x}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix}.$$

We'll formalize for vector spaces like $\mathbb{R}^2$ when one can or can't write a vectors in this way, and how many vectors are needed to do so.

2.1 Linear Span

2.1.1 Definition

Definition. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors from a vector space V. A linear combination of vectors from S is any vector $\vec{v}$ such that*

$$\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n} = \sum_{i=1}^{n} a_i\vec{v_i},$$

where each a_i is from the field associated with V.

From above we see that any vectors in $\mathbb{R}^2$ can be written as a linear combination of the vectors $\langle 1, 0 \rangle$ and $\langle 0, 1 \rangle$.

Definition. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors from a vector space V. The span of S, written* $\operatorname{span} S$*, is the set of all vectors from V that can be written as a linear combination of vectors from S.*

Example.

$$\begin{bmatrix}1\\3\\5\end{bmatrix} \in \operatorname{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\} = \mathbb{R}^3\text{; } 2x^2 - 3x + 3 \in \operatorname{span}\{2, 1-x, 1+x^2\}.$$

Example.

$$\operatorname{span}\left\{\begin{bmatrix}1\\2\\3\end{bmatrix}, \begin{bmatrix}2\\0\\1\end{bmatrix}\right\} = \left\{\begin{bmatrix}x\\y\\z\end{bmatrix} \in \mathbb{R}^3 \mid 2x+5y-4z=0\right\}$$

2.1.2 Span Is a Subspace

Theorem. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors from a vector space V. Then* $\operatorname{span} S$ *is the smallest subspace of V containing all vectors in S.*

Proof. Since $S \subseteq V$, and V is closed under additional and scalar multiplication by virtue of being a vector space, $\operatorname{span} S \subseteq V$.
First we'll show closure under addition. Let $\vec{a}, \vec{b} \in \operatorname{span} S$. Then

$$\begin{aligned}\vec{a} &= a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n}\\ \vec{b} &= b_1\vec{v_1} + b_2\vec{v_2} + \cdots + b_n\vec{v_n}.\end{aligned}$$

So,

$$\vec{a} + \vec{b} = (a_1+b_1)\vec{v_1} + (a_2+b_2)\vec{v_2} + \cdots + (a_n+b_n)\vec{v_n}.$$

Since each $(a_i + b_i)$ is in the associated field, we see that we can write $\vec{a} + \vec{b}$ as a linear combination of vectors from S. Thus, $\operatorname{span} S$ is closed under addition.
Now we'll show scalar multiplication. Let $\vec{a} \in \operatorname{span} S$ and k be an element from the associated field.

$$k\vec{a} = (ka_1)\vec{v_1} + (ka_2)\vec{v_2} + \cdots + (ka_n)\vec{v_n}.$$

Since each ka_i is in the associated field, we see that we can write $k\vec{a}$ as a linear combination of vectors from S. Thus, $\operatorname{span} S$ is closed under multiplication.
Now that we've show $\operatorname{span} S$ is a subspace of V, we'll show its the smallest one containing all elements from S. Suppose U is a subspace of V containing all elements from S. Let $\vec{a} \in \operatorname{span} S$. Since U contains all vectors from S and is a subspace, $\vec{a} \in U$. Thus, $\operatorname{span} S \subseteq U$. ■

2.1.3 Finite & Infinite Dimension

Definition. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors. If* $\operatorname{span} S = V$*, a vector space, then we say*

- *S spans V*
- *S is a spanning set of V*

Definition. *A vector space is finite dimensional if it has a finite spanning set. Otherwise, it is infinite dimensional.*

Example. *We see from the previous examples that*

$$\operatorname{span}\left\{\begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}, \begin{bmatrix}0\\0\\1\end{bmatrix}\right\} = \mathbb{R}^3.$$

Thus, $\mathbb{R}^3$ is a finite dimensional vector space.

Example. *The set of all polynomials with real coefficients is infinite dimensional because any finite set could only represent polynomials up to some finite degree.*

2.2 Linear Independence

2.2.1 Definition

Definition. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors from a vector space V. The set S is said to be linearly independent if the only solution to*

$$a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n} = \vec{0}$$

is $a_1 = a_2 = \cdots = a_n = 0$. Otherwise, S is said to be linearly dependent.

In other words, if we can write a vector from S as a linear combination of the other vectors from S, then S is linearly dependent. Linearly independent sets don't have such redundancies. Equivalently, S is linearly independent if and only if every element of $\operatorname{span} S$ can be written in one and only way as a linear combination of elements from S. If there were two unique representations, we could subtract them to get a non-zero solution to our equation. Notice too that this equation itself is asking for a representation of $\vec{0}$ other than all coefficients set to 0.

Example. *Determine whether or not the following sets are linearly independent or dependent*

1.

$$\{1, x, x^2, 2x^2 - x\}$$

2.

$$\left\{\begin{bmatrix}1\\0\\1\\2\end{bmatrix}, \begin{bmatrix}0\\1\\1\\2\end{bmatrix}, \begin{bmatrix}1\\1\\1\\3\end{bmatrix}\right\}$$

1. This set is not linearly independent because

$$0(1) + -1(x) + 2(x^2) + 1(2x^2 - x) = 0.$$

2. Let the three vectors be called $\vec{v_1}$, $\vec{v_2}$, and $\vec{v_3}$ respectively. Let's consider solutions ot the equation

$$a\vec{v_1} + b\vec{v_2} + c\vec{v_3} = \vec{0}.$$

Since $\vec{v_1}$ and $\vec{v_3}$, both have 1 in their first component, $a+c=0$. Similarly, looking at $\vec{v_2}$ and $\vec{v_3}$, we find that $b+c=0$. Thus, $a=b$ and $c=-a$. So, our equation becomes

$$\begin{aligned} a\vec{v_1} + a\vec{v_2} - a\vec{v_3} &= \vec{0} \\ a\left(\vec{v_1} + \vec{v_2} - \vec{v_3}\right) &= \vec{0} \\ a\begin{bmatrix} 0 \\ 0 \\ 1 \\ 1 \end{bmatrix} &= \vec{0} \\ a &= 0. \end{aligned}$$

We see that the only solution is $a=b=c=0$, so the set is linearly independent.

2.2.2 Properties

Theorem. *Let V be a vector space. Let $S \subseteq T \subseteq V$. If T is linearly independent, then so is S.*

Notice that since all sets that are not linearly independent are linearly dependent, the contrapositive would be "If S is linearly dependent, then so is T".

Theorem. *Let $S = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ be a set of n vectors from a vector space V. The set S is linearly dependent if and only if there exists some $\vec{v_j} \in S$ such that $\vec{v_j} \in \operatorname{span}(S \setminus \{\vec{v_j}\})$. If this is the case, then* $\operatorname{span} S = \operatorname{span}(S \setminus \{\vec{v_j}\})$.

Theorem. *Let V be finite dimensional vector space. The size of any linearly independent set in V is at most the size of a spanning set of V.*

Proof. Suppose $U = \{\vec{u_1}, \ldots, \vec{v_m}\}$ is linearly independent in V. Suppose that $W = \{\vec{w_1}, \ldots, \vec{w_n}\}$ spans V.
Initialize $B = W$, then add $\vec{u_1}$ to B Since W spans V, adding $\vec{u_1}$ to B ensures that B is linearly dependent. By previous result, we can remove some vector other than $\vec{u_1}$ from B while still ensuring that B spans V. Remove one such vector.

Continue in this way, adding each vector from U to B one at a time. By previous result, we know we can remove one vector from B while still being sure it will span S. Since all the vectors from U are linearly independent, we can be sure that this vector will be from W.

After doing this for all vectors in U, we see that at each step we were sure there was a vector from W to remove. Thus, there must at least as many elements of W than of U. ∎

Theorem. *Every subspace of a finite dimensional vector space is finite dimensional.*

Proof. Suppose U is a subspace of V where V is finite dimensional. If $U = \{\vec{0}\}$, then we are done; otherwise, select a vector $\vec{v_1} \in U$. Continue in the following way: If $U = \text{span}\{\vec{v_1}, \vec{v_2}, ...\vec{v_{j-1}}\}$ then we are done; otherwise, select a vector $\vec{v_j}$ not in this span. Since no vector we selected is in the span of the previous, the set of selected vectors is independent. By previous result, the set of selected vectors cannot be longer than a spanning set of V, so we know this process terminates. Thus, the set of selected vectors spans U and is finite, meaning that U is finite dimensional. ∎

2.3 Bases

2.3.1 Definition

Definition. *Let V be a vector space. A non-empty set S of vectors from V is a basis for V if*

1. *S is a spanning set of V.*
2. *S is linearly independent.*

Example. *We saw before that the set $\{\langle 1,0,0\rangle, \langle 0,1,0\rangle, \langle 0,0,1\rangle\}$ is linearly independent and spans $\mathbb{R}^3$. Thus, it is a basis for $\mathbb{R}^3$. This is called the "standard basis" for $\mathbb{R}^3$, and the same pattern holds for $\mathbb{R}^n$.*

Example. *The set $\{1, x, x^2, \dots\}$ is a basis for the infinite dimensional vector space of polynomials with real coefficients.*

Example. *Although $\{\langle 1,0,1,2\rangle, \langle 0,1,1,2\rangle, \langle 1,1,1,3\rangle\}$ is linearly independent, it does not span $\mathbb{R}^4$ because $\langle 1,1,1,1\rangle$ is in $\mathbb{R}^4$ but is not in the span of these three vectors.*

2.3.2 Properties

A basis extends the idea of linear independence from just considering the zero vector to all vectors in a vector space.

Theorem. *A set of vectors $\{\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}\}$ from a vector space V is a basis for V if and only every vector $\vec{v} \in V$ can be uniquely written in the form*

$$\vec{v} = a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n},$$

where $a_1, a_2, \dots, a_n$ are from the field associated with V.

Theorem. *Every spanning set of a vector space V can be made into a basis for V by removing some elements.*

Proof. To make a spanning set into a basis, we need to remove vectors from the set to make it linearly independent, while still maintaining a spanning set of V. Start by removing all zero vectors from the set. Next, continue by removing vector $\vec{v_j}$ from the set if it's in $\text{span}\{\vec{v_1}, \dots, \vec{v_{j-1}}\}$. The resulting set is linearly independent because all vectors are not in the span of the previous. The resulting set also still spans V because we only removed vectors that were already in the span of previous vectors. Thus, the resulting set is a basis for V. ∎

Corollary. *Every finite dimensional vector space has a basis.*

Proof. By definition, a finite dimensional vector space has a finite spanning set. So, we simply apply the process described previously to make a basis from the spanning set. ∎

Just like we could remove vectors from a spanning set to form a basis, we can also add vectors to a linearly independent set.

Theorem. *Every linearly independent set of vectors from a finite dimensional vector space V can be extend to make a basis for V.*

Proof. Let $U = \{\vec{u_1}, \vec{u_2}, \ldots, \vec{v_m}\}$ be a linearly independent set of vectors from V. Let $W = \{\vec{w_1}, \vec{w_2}, \ldots, \vec{w_n}\}$ be a basis of V. Then the set $U \cup W = \{\vec{u_1}, \ldots, \vec{u_m}, \vec{w_1}, \ldots \vec{w_n}\}$ certainly spans V. When we apply the procedure to turn this spanning set into a basis for V, none of the vectors from U get removed, since they are linearly independent. Thus, our resulting basis is an extension of U with some vectors from W, as desired. ∎

2.4 Dimension

We saw that a basis is both a linearly independent and spanning set. We shall see that we can't have two bases for the same vector space of different sizes. Thus, the length of a basis for a vector space is a special number, which we'll call "dimension".

2.4.1 Definition

Theorem. *Any two bases of the same finite dimensional vector space V have the same size.*

Proof. Let B_1 and B_2 be two bases for V. Since B_1 is linearly independent in V, and B_2 spans V,

$$|B_1| \leq |B_2|,$$

by previous result. Similarly, since B_2 is linearly independent in V, and B_1 spans V,

$$|B_1| \leq |B_2|,$$

by previous result. Thus,

$$|B_1| = |B_2|,$$

as desired. ∎

Definition. *Let V be a finite dimensional vector space. The dimension of V, denoted* $\dim V$*, is the size of any basis for V.*

Example. *Here are the dimension of some vector spaces:*

- $\dim \mathbb{R}^n = n$*, since the standard basis has n elements.*
- $\mathcal{P}_m$*, the set of a polynomials with real coefficients of degree at most m, has dimension* $m+1$ *since* $\{1, x, x^2, \ldots, x^m\}$ *is a basis.*
- *The set of all $n \times m$ matrices with real coefficients has dimension nm (imagine "unwrapping" the matrix to look like a vector in $\mathbb{R}^{nm}$).*

2.4.2 Properties

Our results about the length of linearly independent sets and subspaces extend to dimension.

Theorem. *If U is a subspace of a finite dimensional vector space V, then* $\dim U \leq \dim V$.

Proof. A basis for U is linearly independent in V, and a basis for V spans V. So, applying a previous result and the definition of dimension, $\dim U \leq \dim V$. ∎

Any linearly independent set of vectors with size equal to the dimension is a basis.

Theorem. *Let V be a finite dimensional vector space. Then every linearly independent set of vectors from V with size* $\dim V$ *is a basis for V.*

Proof. Let $\dim V = n$, and suppose $U = \{\vec{v_1}, \vec{v_2}, \ldots, \vec{v_n}\}$ is linearly independent in V. We know by previous result that we can extend U to form a basis. This basis must have n elements. Thus, the extension is trivial and adds no vectors, meaning U is a basis for V. ∎

Similarly, any spanning set of vectors with size equal to the dimension is a basis.

Theorem. *Let V be a finite dimensional vector space. Then every spanning set of vectors from V with size* $\dim V$ *is a basis for V*

Proof. Let $\dim V = n$, and suppose $S = \{\vec{v_1}, \ldots, \vec{v_n}\}$ is a spanning set in V. We know by previous result that we can remove vectors from S to form a basis. Thus basis must have n elements. Thus, the removal is trivial and removes no vectors, meaning S is a basis for V. ∎

Example. *Show that* $\{\langle 1, 1\rangle, \langle -1, 1\rangle\}$ *is a basis for* $\mathbb{R}^2$.

> We showed previously that any vector from $\mathbb{R}^2$ can be written as a linear combination of these vectors. Thus, the set is spanning. Since $\dim \mathbb{R}^2 = 2$, and our spanning set has 2 elements, it must be a basis.

Chapter 3

Linear Maps

Now that we have an understanding of finite dimensional vector spaces, we can investigate the mappings between vector spaces that preserve these properties.

3.1 Linear Transformations

3.1.1 Definition

Definition. *A linear map (linear transformation) from a vector space V to a vector space W is a function $T : V \to W$ that preserves the linearity properties of V and W. That is, for all $\vec{u}, \vec{v} \in V$ and $k \in F$, where F is the field associated with V and W,*

$$\begin{aligned} T(\vec{u} + \vec{v}) &= T(\vec{u}) + T(\vec{v}) \\ T(k\vec{u}) &= kT(\vec{u}). \end{aligned}$$

Often, we'll drop the parentheses and write $T\vec{u}$.

Example. *Show that $T : \mathbb{R}^2 \to \mathbb{R}^2$ defined by $T(\langle x, y\rangle) = \langle x + y, x - y\rangle$ is a linear transformation.*

We'll show each of the properties hold. For the first property:

$$\begin{aligned} T(\langle x_1, y_1\rangle + \langle x_2, y_2\rangle) &= T(\langle x_1 + x_2, y_1 + y_2\rangle) \\ &= \langle x_1 + x_2 + y_1 + y_2, x_1 - x_2 + y_1 - y_2\rangle \\ &= \langle x_1 + y_1, x_1 - y_1\rangle + \langle x_2 + y_2, x_2 - y_2\rangle \\ &= T(\langle x_1, y_1\rangle) + T(\langle x_2, y_2\rangle). \end{aligned}$$

For the second property:

$$\begin{aligned} T(k\langle x, y\rangle) &= T(\langle kx, ky\rangle) \\ &= \langle kx + ky, kx = ky\rangle \\ &= k\langle x + y, x - y\rangle \\ &= kT(\langle x, y\rangle). \end{aligned}$$

Set of All Linear Maps

Definition. *The set of all linear maps from V to W is denoted $\mathcal{L}(V, W)$.*

There are two important elements of this set, which have special notations.

- **Zero** This is the transformation that sends all elements to W's zero vector. It is notated as 0, or as **0**.
- **Identity** This map exists when $V = W$ and sends all elements to themselves. It is notated as I, or as **1**.

3.1.2 Properties

Theorem. *A linear transformation $T : V \to W$ is uniquely determined by how is transforms the vectors in a basis for V.*

Proof Sketch. Let $B = \{\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}\}$ be a basis for V. Let $\vec{v} \in V$. Since B is a basis, $\vec{v}$ has a unique representation as a linear combination of vectors from B. So,

$$\begin{aligned}\vec{v} &= a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n}\\ T\vec{v} &= T(a_1\vec{v_1} + a_2\vec{v_2} + \cdots + a_n\vec{v_n})\\ &= T(a_1\vec{v_1}) + T(a_2\vec{v_2}) + \cdots + T(a_n\vec{v_n})\\ &= a_1T\vec{v_1} + a_2T\vec{v_2} + \cdots + a_nT\vec{v_n}.\end{aligned}$$

■

Corollary. *For any linear transformation $T : V \to W$, $T\vec{0} = \vec{0}$.*

Theorem. *Let $B = \{\vec{v_1}, \vec{v_2}, \dots, \vec{v_n}\}$ be a basis for V and $\vec{w}_1, \dots, \vec{w}_n \in W$. Then there exists a unique linear map $T : V \to W$ such that*

$$T\vec{v_j} = \vec{w_j}$$

for $j = 1, 2, \dots, n$. In particular, the transformation is

$$T(a_1\vec{v_1} + \cdots + a_n\vec{v_n}) = a_1\vec{w_1} + \cdots + a_n\vec{w_n}.$$

Operations on Set of Linear Transformations

We can define addition and scalar multiplication on $\mathcal{L}(V, W)$ and show that we in fact have a vector space.

Definition. *Let $S, T \in \mathcal{L}(V, W)$ and k be a scalar from the associated field. Then for all $\vec{v} \in V$,*

$$(S + T)(\vec{v}) = S\vec{v} + T\vec{v}$$

and

$$(kT)(\vec{v}) = k(T\vec{v}).$$

Theorem. *The set $\mathcal{L}(V, W)$ with the operations defined above is a vector space.*

We can also define the multiplication of two elements as function composition.

Definition. *Let $S \in \mathcal{L}(V, U)$ and $T \in \mathcal{L}(U, W)$. Then for all $\vec{v} \in V$,*

$$(ST)(\vec{v}) = S(T\vec{v}).$$

This is normally notated $ST\vec{v}$. Notice that even if the domains and codomains allow $TS\vec{v}$ to be defined, multiplication of linear transformations is not in general commutative: $ST\vec{v} \neq TS\vec{v}$.
There are however some properties of multiplication that hold

Theorem. *The following properties of multiplication of linear transformations holds (assuming the domains and codomains allow the operations to be defined):*

- ***Associativity**: $(T_1T_2)T_3 = T_1(T_2T_3)$.*
- ***Identity**: $TI = IT = T$.*
- ***Left & Right Distributive**: $(T_1 + T_2)T_3 = T_1T_3 + T_2T_3$ and $T_1(T_2 + T_3) = T_1T_2 + T_1T_3$.*

3.2 Null Space & Range

3.2.1 Null Space

Definition

Definition. *For $T \in \mathcal{L}(V, W)$, the null space of T, notated* null T*, is the set of all vectors from V that are mapped to the zero vector of W.*

$$\text{null}\, T = \{\vec{v} \in V \mid T\vec{v} = \vec{0}\}.$$

You may also hear the null space referred to as the "kernel".

Example. *Find the null space of the following linear map from $\mathbb{R}^3$ to $\mathbb{R}^2$:*

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x - z \\ y - z \end{bmatrix}.$$

Since we are looking for when the output of the transformation is the zero vector, we want when $x - z = 0$ and $y - z = 0$. Thus, $x = z$ and $y = z$, so $x = y = z$. So,

$$\text{null}\, T = \left\{ x \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \mid x \in \mathbb{R} \right\}.$$

Properties

Theorem. *For $T \in \mathcal{L}(V, W)$,* null T *is a subspace of* V.

Proof. By previous result, we know that $\vec{0} \in \operatorname{null} T$. Suppose $\vec{u}, \vec{v} \in \operatorname{null} T$. Then

$$T(\vec{u} + \vec{v}) = T\vec{u} + T\vec{v} = \vec{0} + \vec{0} = \vec{0}.$$

So, $\vec{u} + \vec{v} \in \operatorname{null} T$. Let k be in the field associated with V and W. Then

$$T(k\vec{u}) = kT\vec{u} = k\vec{0} = \vec{0}.$$

So, $k\vec{u} \in \operatorname{null} T$. Since null T is non-empty and closed under addition and scalar multiplication, it is a subspace of V. ■

Certain linear maps only contain the zero vector in their null space. These are in fact exactly the injective linear maps.

Definition. *A function $T : V \to W$ is injective if $T\vec{u} = T\vec{v} \implies \vec{u} = \vec{v}$.*

That is, injective functions are those that don't map two different elements in the domain to the same element in the codomain. You may also hear injective functions called "one-to-one".

Theorem. *For $T \in \mathcal{L}(V, W)$, T is injective if and only of* null $T = \{\vec{0}\}$.

Proof. $\implies$: Suppose T is injective. By previous result $\vec{0} \in \operatorname{null} T$. Suppose $\vec{v} \in \operatorname{null} T$. Then $T(\vec{v}) = \vec{0} = T(\vec{0})$. Since T is injective, we must have $\vec{v} = \vec{0}$. Thus, $\vec{0}$ is the only element of null T.

$\impliedby$: Suppose null $T = \{\vec{0}\}$. Let $\vec{u}, \vec{v} \in V$ with $T\vec{u} = T\vec{v}$. Then $\vec{0} = T\vec{u} - T\vec{v} = T(\vec{u} - \vec{v})$. Since $\vec{0}$ is the only element of null T, we must have $\vec{u} - \vec{v} = 0$, or $\vec{u} = \vec{v}$. Thus, T is injective. ■

3.2.2 Range

Definition

Definition. *For $T : V \to W$, the range of T is the subset of W containing all elements of the form $T\vec{v}$ for some $\vec{v} \in V$.*

$$\operatorname{range} T = \{T\vec{v} \mid \vec{v} \in V\}.$$

You may also hear the range referred to as the "image" of V, while the image of $\vec{v} \in V$ is $T\vec{v}$.

Example. *Find the range of the following linear map from $\mathbb{R}^2$ to $\mathbb{R}^3$:*

$$T \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x \\ y \\ x + y \end{bmatrix}.$$

We see from the definition of the transformation that the range is all vectors in $\mathbb{R}^3$ with a z components equal to the sum of the x and y components.

$$\operatorname{range} T = \left\{ \begin{bmatrix} x \\ y \\ z \end{bmatrix} \mid x, y, z \in \mathbb{R}, x + y = z \right\}.$$

Properties

Theorem. *For $T \in \mathcal{L}(V, W)$,* range T *is a subspace of* W.

Proof. Since $T\vec{0} = \vec{0}$, $\vec{0} \in \text{range}\, T$. Suppose $\vec{w_1}, \vec{w_2} \in \text{range}\, T$. Then there exists $\vec{v_1}, \vec{v_2} \in V$ such that $T\vec{v_1} = \vec{w_1}$ and $T\vec{v_2} = \vec{w_2}$. So,

$$T(\vec{v_1} + \vec{v_2}) = T\vec{v_1} + T\vec{v_2} = \vec{w_1} + \vec{w_2}.$$

So, $\vec{w_1} + \vec{w_2} \in \text{range}\, T$. Let k be an element of the associated field. Then

$$T(k\vec{v_1}) = kT\vec{v_1} = k\vec{w_1}.$$

So, $k\vec{w_1} \in \text{range}\, T$. Since range T is non-empty and closed under addition and scalar multiplication, it is a subspace of W. ■

Certain linear maps have range $T = W$. These are in fact exactly the surjective linear maps.

Definition. *A function $T : V \to W$ is surjective if* range $T = W$.

You may also hear surjective functions called "onto".

3.3 Fundamental Theorem of Linear Maps

For finite-dimensional domains, the dimension of the null space and range of and dimension of the range of any linear transformation sum to give the dimension of the domain.

Theorem. *Let V be a finite dimensional vector space and $T \in \mathcal{L}(V, W)$. Then* range T *is finite dimensional, and*

$$\dim V = \dim \text{null}\, T + \dim \text{range}\, T.$$

Proof. Let $U = \{\vec{u_1}, \ldots, \vec{u_m}\}$ be a basis for null T, so $\dim \text{null}\, T = m$. This linearly independent list containing elements from V can be extend to be a basis $U' = \{\vec{u_1}, \ldots, \vec{u_m}, \vec{v_1}, \ldots \vec{v_n}\}$, which means $\dim V = m + n$.

Now we just need to show that $\dim \text{range}\, T = n$. We will do show by showing that $T\vec{v_1}, \ldots, T\vec{v_n}$ is a basis for range T. Let $\vec{v} \in V$. Since U' spans V, we can write

$$\begin{aligned}
\vec{v} &= a_1\vec{u_1} + \cdots + a_m\vec{u_m} + b_1\vec{v_1} + \cdots + b_n\vec{v_n} \\
T\vec{v} &= T(a_1\vec{u_1}) + \cdots + T(a_m\vec{u_m}) + T(b_1\vec{v_1}) + \cdots + T(b_n\vec{v_n}) \\
&= a_1\vec{0} + \cdots + a_m\vec{0} + b_1T\vec{v_1} + \cdots + b_n\vec{v_n} \\
&= b_1T\vec{v_1} + \cdots + b_n\vec{v_n}
\end{aligned}$$

So, we see that $X = \{T\vec{v}_1, \ldots, T\vec{v}_n\}$ spans range T. Now we just need to show that X is linearly independent. Let $c_1, \ldots, c_n$ be from the associated field where

$$\begin{aligned} c_1 T\vec{v}_1 + \cdots + c_n T\vec{v}_n &= \vec{0} \\ T(c_1\vec{v}_1 + \cdots + c_n\vec{v}_n) &= \vec{0} \\ c_1\vec{v}_1 + \cdots + c_n\vec{v}_n &\in \operatorname{null} T \\ c_1\vec{v}_1 + \cdots + c_n\vec{v}_n &= d_1\vec{u}_1 + \cdots + d_m\vec{u}_m. \end{aligned}$$

Since each of the $\vec{v}_i$'s and $\vec{u}_i$'s are linearly independent, the only solution to this equation must be all c_i's and d_i's are 0. Thus, X is a linearly indepepdent set, meaning $\dim \operatorname{range} T = n$. ■

Corollary. *Let V and W be finite dimensional vector spaces with* $\dim V > \dim W$. *Then no linear map from V to W is injective.*

Proof. Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned} \dim V &= \dim \operatorname{range} T + \dim T \\ \dim T &= \dim V - \dim \operatorname{range} T \\ &\geq \dim V - \dim W \\ &> 0. \end{aligned}$$

Thus, the null space contains something other than $\vec{0}$, meaning T cannot be injective. ■

Corollary. *Let V and W be finite dimensional vector spaces with* $\dim V < \dim W$. *Then no linear map from V to W is surjective.*

Proof. Let $T \in \mathcal{L}(V, W)$. Then

$$\begin{aligned} \dim V &= \dim \operatorname{range} T + \dim T \\ \dim \operatorname{range} T &= \dim V - \dim T \\ &\leq \dim V \\ &< \dim W. \end{aligned}$$

Since $\dim \operatorname{range} T < \dim W$, range $T \neq W$, so T is not surjective. ■

3.4 Isomorphisms

3.4.1 Invertible Functions

Generally

Definition. *A function $T : V \to W$ is invertible if there exists an inverse function $T^{-1} : W \to V$ such that $T \circ T^{-1}$ is the identity on W and $T^{-1} \circ T$ is the identity on V.*

That is, invertible functions are those that are reversible both ways. Notice too that T is the inverse of T^{-1}. That is, $(T^{-1})^{-1} = T$.

Theorem. *The inverse of function is unique.*

Proof. Let $T : V \to W$ be an invertible function. Let $S_1, S_2 : W \to V$ be inverses of T. Then

$$S_1 = S_1 \circ I = S_1 \circ (T \circ S_2) = (S_1 \circ T) \circ S_2 = I \circ S_2 = S_2.$$

Thus, all inverses of T are in fact the same function, as desired. ■

This uniqueness property is what allows the notation T^{-1} be be well-defined.

Example. *Find the inverse function of the $T : \mathbb{R}^3 \to \mathbb{R}^3$ defined by*

$$T \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \frac{1}{2} \begin{bmatrix} x+y \\ y+z \\ z+x \end{bmatrix}.$$

We'll check that the function $S : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$S \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} x-y+z \\ x+y-z \\ -x+y+z \end{bmatrix}$$

is the inverse.

$$(T \circ S) \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = T \left(\begin{bmatrix} x-y+z \\ x+y-z \\ -x+y+z \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} (x-y+z)+(x+y-z) \\ (x+y-z)+(-x+y+z) \\ (-x+y+z)+(x-y+z) \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

$$(S \circ T) \left(\begin{bmatrix} x \\ y \\ z \end{bmatrix} \right) = S \left(\frac{1}{2} \begin{bmatrix} x+y \\ y+z \\ z+x \end{bmatrix} \right) = \frac{1}{2} \begin{bmatrix} (x+y)-(y+z)+(z+x) \\ (x+y)+(y+z)-(z+x) \\ -(x+y)+(y+z)+(z+x) \end{bmatrix} = \begin{bmatrix} x \\ y \\ z \end{bmatrix}.$$

Since $T \circ S$ and $S \circ T$ are both the identity, S is indeed the inverse of T.

Definition. *If a function $T : V \to W$ is both injective and surjective, then it is bijective. We say then that T is a bijection.*

Theorem. *A function is invertible if and only if it is bijective.*

Proof. Let $T : V \to W$ be a function.

$\Longrightarrow$: Suppose T is invertible. We'll show that T is a bijection by showing it's both injective in surjective. Let $u, v \in V$ such taht $T(u) = T(v)$. Then

$$u = T^{-1}(T(u)) = T^{-1}(T(v)) = v.$$

So, T is injective. Let $w \in W$. Since T is invertible, $w = T(T^{-1}(w))$. Hence, $w \in \text{range } T$, meaning $\text{range } T = W$, so T is surjective.

$\Longleftarrow$: Suppose T is a bijection. For each element $w \in W$, define $S(w)$ to be the unique element of V such that $(T \circ S)(w) = w$. The existence of S and uniqueness of its output are implied by T being a bijection. We see that by construction, $T \circ S$ is the identity on W. So, all that remains is to show that $S \circ T$ is the identity on V. Let $v \in V$. Then

$$T((S \circ T)(v)) = (T \circ S)(T(v)) = I(T(v)) = T(v).$$

Since T is injective, this result implies that $(S \circ T)(v) = v$, meaning $S \circ T$ is the identity on V. ∎

For Linear Maps

Inverse functions behave nicely when looking just at linear maps. In particular, the inverse of a linear map is also a linear map.

Theorem. *The inverse of a linear map is a linear map.*

Proof. Let $T : U \to V$ be an invertible linear map. Let $x, y \in V$. Then

$$\begin{aligned} T^{-1}(x+y) &= T^{-1}\left(T(T^{-1}(x)) + T(T^{-1}(y))\right) \\ &= T^{-1}(T(T^{-1}(x) + T^{-1}(y))) \\ &= T^{-1}(x) + T^{-1}(y). \end{aligned}$$

Let k be from the field associated with U and V. Then

$$\begin{aligned} T^{-1}(kx) &= T^{-1}(kT(T^{-1}(T(x)))) \\ &= T^{-1}(T(kT^{-1}(x))) \\ &= kT^{-1}(x). \end{aligned}$$

Thus, T^{-1} is a linear map. ∎

3.4.2 Isomorphic Vector Spaces

In working on examples, you might have noticed that certain vector spaces seem nearly the same. For example, when working with elements as just members of a vector space, there seems to be no difference between $\mathbb{R}^4$ and $\mathcal{M}_{2\times 2}$, the vector space of 2×2 matrices with real entries. We'll formalize this idea of vector spaces being the same and find that all finite dimensional vector spaces of the same dimension and field are the same.

Definition. *An isomorphism between two vector spaces is an invertible linear map. If such a function exists, then the two vector spaces are said to be isomorphic to each other.*

Theorem. *Two finite dimensional vector spaces over the same field are isomorphic if and only if they have the same dimension.*

Proof. $\implies$: Let $T : V \to W$ be the isomorphism between finite dimensional vector spaces V and W. So, null $T = \{\vec{0}\}$ and range $T = W$, meaning dim null $T = 0$ and dim range $T =$ dim W. So, by the Fundamental Theorem of Linear Maps, dim $V =$ dim W.

$\impliedby$: Let V and W be finite dimensional vector spaces with the same dimension. Let $\{\vec{v_1}, \ldots, \vec{v_n}\}$ be a basis for V and $\{\vec{w_1}, \ldots, \vec{w_n}\}$ be a basis for W. Define $T : V \to W$ by

$$T(c_1\vec{v_1} + \cdots + c_2\vec{v_n}) = c_1\vec{w_1} + \ldots c_n\vec{w_n}.$$

The function T is surjective because the $\vec{w}$'s span W, meaning we can adjust the scalars to get any vector in W as the output. Further, T is injective because the $\vec{w}$'s are linearly independent, meaning $T = \{0\}$. Since T is injective and surjective, it is an isomorphism, meaning V and W are isomorphic. ■

Corollary. *Let F be a field. Then all vector spaces with field F and dimension n are isomorphic to F^n.*

Example. *Show that $\mathcal{M}_{2\times 2}$ and $\mathbb{R}^4$ are isomorphic.*

Both these vector spaces have dimension 4, so they are isomorphic by previous result. One simple isomorphism between them is

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

Theorem. *Let V and W be finite dimensional vector spaces. Then $\mathcal{L}(V, W)$, the vector space of all linear transformations from V to W is finite dimensional with*

$$\dim \mathcal{L}(V, W) = \dim V \cdot \dim W.$$

Chapter 4

Matrices

We'll introduce matrices, how they can represent linear maps and systems of linear equations, and useful operations we can perform on them.

4.1 Definition

Definition. *An $m \times n$ matrix is an array of objects (usually field elements) arranged in m rows and n columns.*

Matrices are usually written inside square brackets. We tend to use uppercase letters like M to represent matrices as variables. The notation $M_{a,b}$ represents the element in row a and column b.

Example. *Matrix M is* 3×4.

$$M = \begin{bmatrix} 1 & 4 & 0 \\ 8 & -1 & -2 \\ 3 & 7 & 4 \end{bmatrix}$$

We see that $M_{2,2} = -1$ and $M_{3,1} = 3$.

4.2 Basic Operations

4.2.1 Vector Space Operations

Two matrices are considered equal if they have the same number of rows and columns and all entries are equal. Scalar multiplication works by multiplying each element by the scalar. Addition works element by element.

Example.

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 2 \\ -1 & 3 \end{bmatrix}, B = \begin{bmatrix} 4 & 0 \\ -2 & 5 \end{bmatrix}. \\
A+B &= \begin{bmatrix} 1+4 & 2+0 \\ -1+-2 & 3+5 \end{bmatrix} = \begin{bmatrix} 5 & 2 \\ -3 & 8 \end{bmatrix} \\
3A &= \begin{bmatrix} 3\cdot 1 & 3\cdot 2 \\ 3\cdot -1 & 3\cdot 3 \end{bmatrix} = \begin{bmatrix} 3 & 6 \\ -3 & 9 \end{bmatrix}.
\end{aligned}$$

Notice addition is commutative, associative, has an additive identity (the all 0's matrix), and has an additive inverse (scalar multiply by -1). Further, scalar multiplication has a multiplicative identity (1), is distributive over both addition and field multiplication. Thus, the set of all matrices of the same size form a vector space. We denote the vector space of all $m \times n$ matrices with real entries as $\mathcal{M}_{m\times n}$.

4.2.2 Multiplication

We can also define an operation for multiplying two matrices of compatible sizes that outputs another matrix.

Definition. *Let A be an $m \times n$ matrix, and let B be and $n \times k$ matrix. Then $C = AB$ is an $m \times k$ matrix where*

$$C_{i,j} = \sum_{d=1}^{n} A_{i,d}B_{d,j}.$$

If you're familiar with the concept of dot products, then $C_{i,j}$ is the dot product of the ith row of A with the jth column of B.

Example.

$$\begin{aligned}
A &= \begin{bmatrix} 1 & 2 & 3 \\ 0 & -1 & 2 \end{bmatrix}, B = \begin{bmatrix} 1 & -1 & 0 & 2 \\ 2 & 0 & 3 & -1 \\ 0 & 1 & 1 & 5 \end{bmatrix}. \\
AB &= \begin{bmatrix} 5 & 2 & 9 & 15 \\ -2 & 2 & -1 & 11 \end{bmatrix}.
\end{aligned}$$

Similar to scalar multiplication, there exists a multiplicative identity matrix. However, this matrix only behaves like an identity when the matrix it's being multiplied is $n \times n$ (i.e. a square matrix).

Definition. *The $n \times n$ matrix I_n is called the identity matrix and is defined by*

$$I_{i,j} = \begin{cases} 1 & i = j \\ 0 & \text{otherwise} \end{cases}.$$

Theorem. *Let A be an $n \times n$ matrix. Then $AI_n = I_nA = A$.*

Proof. Let $C = AI_n$. Notice,

$$\begin{aligned}
C_{i,j} &= \sum_{d=1}^{n} A_{i,d}(I_n)_{d,j} \\
&= \sum_{d=1}^{n} A_{i,d} \begin{cases} 1 & d = j \\ 0 & \text{otherwise} \end{cases} \\
&= \sum_{d=1}^{n} \begin{cases} A_{i,j} & d = j \\ 0 & \text{otherwise} \end{cases} \\
&= A_{i,j}.
\end{aligned}$$

This same line of reasoning works to also show that $I_nA = A$. Since all entries of A and C are equal, $C = AI_n = A$, as desired. ■

Also similar to scalar multiplication, matrix multiplication is associative and distributive.

Theorem. *Let A, B, and C be matrices. Let k be a scalar. Then the following properties hold (assuming the matrices have the correct dimensions):*

- ***Associative**: $A(BC) = (AB)C$.*
- ***Distributive Over Matrix Multiplication**: $k(AB) = (kA)B = A(kB)$.*
- ***Left Distributive Over Addition**: $A(B+C) = AB + AC$.*
- ***Right Distributive Over Addition**: $(A+B)C = AC + BC$.*

Unlike scalar multiplication, matrix multiplication is not commutative. For one, if AB is defined, BA won't also be defined unless A and B are both square matrices Even if this is the case, $AB \neq BA$ in general.

Theorem. *An $n \times n$ matrix A commutes only and all matrices in the vector space* $\operatorname{span}(\{I_n, A\})$.

4.3 As Linear Maps

4.3.1 Definition

Matrices of size $m \times n$ can represent a linear map from a space with dimension m to one with dimension n.

Definition. *Let $T \in \mathcal{L}(V, W)$. Let $\{\vec{v_1}, \ldots, \vec{v_n}\}$ be a basis for V and $\{\vec{w_1}, \ldots, \vec{w_m}\}$ be a basis for W. The matrix of T with respect to these bases is the $m \times n$ matrix $\mathcal{M}(T)$ whose entries $\mathcal{M}(T)_{j,k}$ are defined by*

$$T\vec{v_k} = \mathcal{M}(T)_{1,k}\vec{w_1} + \cdots + \mathcal{M}(T)_{m,k}\vec{w_m}.$$

That is, the kth column of A gives the scalars to write $T\vec{v_k}$ as a linear combination of the basis vectors for W.

Example. *Let $T : \mathbb{R}^2 \to \mathbb{R}^3$ be defined by*

$$T\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} x+3y \\ 2x+5y \\ 7x+9y \end{bmatrix}.$$

Find $\mathcal{M}(T)$ assuming standard bases.

To find the kth column of $\mathcal{M}(T)$, we simply need to calculate T of the kth basis vector.

$$T\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 7 \end{bmatrix}$$
$$T\begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 3 \\ 5 \\ 9 \end{bmatrix} .\mathcal{M}(T) = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}.$$

We can check that this gives the correct result whether we go through T or $\mathcal{M}(T)$.

$$T\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 17 \\ 34 \end{bmatrix}$$
$$\mathcal{M}(T)\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 & 3 \\ 2 & 5 \\ 7 & 9 \end{bmatrix}\begin{bmatrix} 1 \\ 3 \end{bmatrix} = \begin{bmatrix} 10 \\ 17 \\ 34 \end{bmatrix}.$$

Since we can now put any linear transformation between finite dimensional vector spaces int a matrix, we can determine $\dim \mathcal{L}(V, W)$.

Corollary. *The V and W be finite dimensional vector spaces with $\dim V = m$ and $\dim W = n$. Then $\dim \mathcal{L}(V, W) = \dim M_{m\times n} = mn$.*

4.3.2 Properties

Just like how we can add, scalar multiply, and compose linear maps to make new ones, adding and multiplying these matrices gives the matrix of these maps.

Theorem. *Let $S, T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(S+T) = \mathcal{M}(S) + \mathcal{M}(T)$.*

Theorem. *Let $T \in \mathcal{L}(V, W)$ and k be from the associated field. Then $\mathcal{M}(kT) = k\mathcal{M}(T)$.*

Theorem. *Let $S \in \mathcal{L}(U, V)$ and $T \in \mathcal{L}(V, W)$. Then $\mathcal{M}(ST) = \mathcal{M}(S)\mathcal{M}(T)$.*

4.4 As Systems of Linear Equations

4.4.1 Invertiblity

Definition

We can represent a system of linear equations using matrices.

Example.

$$\begin{cases} x+2y+3z &= 8 \\ -x+4z &= 12 \\ 2x-y+z &= 0 \end{cases} \to \begin{bmatrix} 1 & 2 & 3 \\ -1 & 0 & 4 \\ 2 & -1 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 8 \\ 12 \\ 0 \end{bmatrix}.$$

To solve such a system using matrices, we'd like a matrix that we could multiply both sides of the equation on the left by that multiplies with the existing matrix to give an identity matrix. Then we'd be left with an equation like $\langle x, y, z\rangle = \langle \dots \rangle$, which would give us a solution.

Definition. *A matrix A is invertible if there exists a matrix A^{-1} such that $AA^{-1} = A^{-1}A = I$. Such a matrix is called the inverse of A.*

Notice that the definition implies that the inverse of A^{-1} is just A, so $(A^{-1})^{-1} = A$.

Theorem. *The inverse of a matrix, if it exists, is unique.*

Proof. Let A be a matrix with B and C as inverses. Then

$$\begin{aligned} I &= AC \\ BI &= B(AC) \\ &= (BA)C \\ &= IC \\ B &= C. \end{aligned}$$

Thus, the two inverses are the same, as desired. ∎

This uniqueness result is what allows the notation A^{-1} to make sense.
We can compute the inverse of a matrix A by creating a matrix of the form $[A \mid I]$ and then performing operations like adding linear combinations of rows together to put this matrix in the form $[I \mid A^{-1}]$. Each one of these operations (scalar multiplying at row, adding linear combinations of rows, swapping two rows) is a linear transformation. So, when simplifying to the form where we have an identity matrix on the left, we've found a composition of operations that "undoes" the linear transformation represented by A.

Example. *Find the inverse of the following matrix:*

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

Setting up $[A \mid I_3]$ and simplifying,

$$\left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ -1 & 2 & 0 & 0 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \overset{R_2=R_2+R_1}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 0 & 1 \end{array}\right] \overset{R_3=R_2+3R_3}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 3 & -2 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array}\right]$$

$$\overset{R_2=R_2+2R_3}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 3 & 0 & 3 & 3 & 6 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array}\right] \overset{R_2=R_2/3}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & -2 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array}\right] \overset{R_1=R_1+2R_3}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 1 & 0 & 3 & 2 & 6 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array}\right]$$

$$\overset{R_1=R_1-R_2}{\rightarrow} \left[\begin{array}{ccc|ccc} 1 & 0 & 0 & 2 & 1 & 4 \\ 0 & 1 & 0 & 1 & 1 & 2 \\ 0 & 0 & 1 & 1 & 1 & 3 \end{array}\right].$$

So, we see that

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

Properties

Now that we know how to invert matrices, we might want to know whether we can invert sums and products of invertible matrices. Let A and B be invertible matrices of the same size, and let k be a scalar.

- $A+B$ may not be invertible. For example, consider $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, which are both invertible. We see that $A+B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, which is not invertible.
- kA is invertible, and $(kA)^{-1} = \frac{1}{k}A^{-1}$.
- AB is invertible, and $(AB)^{-1} = B^{-1}A^{-1}$.

Invertible matrices are those such that there exists and inverse matrix. Since matrices represent linear transformations, it should come as no surprise that a linear transformation is invertible if and only if its matrix is invertible. Since invertible linear transformations are bijections, we can determine when a system of equations has a unique solution.

Theorem. *If the $n \times n$ matrix A is invertible, then the system $A\vec{x} = \vec{b}$ has a unique solution for any vector $\vec{b} \in \mathbb{R}^n$, namely $\vec{x} = A^{-1}\vec{b}$.*

4.4.2 Column Space & Null Space

Definitions

Since matrices represent linear transformations, we can define a null space.

Definition. *Let A be an $m \times n$ matrix. The null spaces of A, notated,* $\operatorname{null} A$*, is the set of vectors $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{0}$.*

We can also define a concept like the range of a linear transformation.

Definition. *Let A be an $m \times n$ matrix. The column space of A, notated* $\operatorname{col} A$*, is the set of all linear combinations of the columns of A. That is,* $\operatorname{col} A = \operatorname{span}\{$*columns of A*$\}$.

Properties

Just like the null space and range of linear transformations, the null space and column space of matrices are subspaces.

Theorem. *Let A be an $m \times n$ matrix. Then A is a subspace of $\mathbb{R}^n$, and* $\operatorname{col} A$ *is a subspace of $\mathbb{R}^m$ (or whichever field).*

Proof. First we'll show closure under addition. Let $\vec{x}, \vec{y} \in \operatorname{null} A$. Then

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y} = \vec{0} + \vec{0} = \vec{0}.$$

So, $\vec{x} + \vec{y} \in \operatorname{null} A$.
Now let $\vec{x}, \vec{y} \in \operatorname{col} A$. Let $\{\vec{v_1}, \ldots, \vec{v_n}\}$ be the columns of A. Then

$$\begin{aligned} \vec{x} &= a_1\vec{v_1} + \cdots + a_n\vec{v_n} \\ \vec{y} &= b_1\vec{v_1} + \cdots + b_n\vec{v_n} \\ \vec{x} + \vec{y} &= (a_1 + b_1)\vec{v_1} + \cdots + (a_n + b_n)\vec{v_n}. \end{aligned}$$

So, $\vec{x} + \vec{y} \in \operatorname{col} A$.

Now we'll show closure under scalar multiplication. Let $x \in \operatorname{null} A$ and k be a scalar. Then

$$A(k\vec{x}) = k(A\vec{x}) = k\vec{0} = \vec{0}.$$

So, $k\vec{x} \in \operatorname{null} A$. Now let $\vec{x} \in \operatorname{col} A$. Then

$$\begin{aligned} \vec{x} &= a_1\vec{v_1} + \cdots + a_n\vec{v_n} \\ k\vec{x} &= ka_1\vec{v_1} + \cdots + ka_n\vec{v_n}. \end{aligned}$$

So, $k\vec{x} \in \operatorname{col} A$. ■

Theorem. *Let A be an $m \times n$ matrix. Then the system $A\vec{x} = \vec{b}$ is consistent (i.e. has a solution) if and only if $\vec{b} \in$* $\operatorname{col} A$*.*

Theorem. *The rows/columns of the $n \times n$ matrix A are a basis for $\mathbb{R}^n$ if and only if A is invertible.*

We also have a version of the Fundamental Theorem of Linear Maps but for matrices.

Definition. *Let A be an $m \times n$ matrix. The rank of A, notated* $\operatorname{rank} A$*, is the dimension of the column space of A. The nullity of A, notated* $\operatorname{nullity} A$*, is the dimension of the null space of A.*

Theorem. *Let A be an $m \times n$ matrix. Then*

$$\operatorname{rank} A + \operatorname{nullity} A = n.$$

4.5 Determinants

TODO

Chapter 5

Additional Resources

5.1 Contributors

Special thanks to everyone who made contributions to this project on Github. They are listed in order of number of commits as `name (GitHub username)`.

William Boyles (wmboyles)	Nate (aneziac)	Ashwin Murali (Suzukazole)
आदित्य देव (dev-aditya)	Calvin McPhail-Snyder (esselltwo)	github-actions[bot] (github-actions[bot])
Jared Geller (jaredgeller)	Robert Washbourne (rawsh)	