Algebra including linear algebra

Adam Boult (www.bou.lt)

April 29, 2024

Contents

Part I

Elementary number theory

Chapter 1

The integers

1.1 The integers

1.1.1 Integers

Defining integers

To extend the number line to negative numbers, we define:

$\forall ab \in \mathbb{N} \exists c(a + c = b)$

For any pair of numbers there exists a terms which can be added to one to get the other.

For $1 + x = 3$ this is another natural number, however for $3 + x = 1$ there is no such number.

Integers are defined as the solutions for any pair of natural numbers.

There are an infinite number of ways to write any integer. -1 can be written as $0 - 1$, $1 - 2$ etc.

The class of these terms form an equivalence class.

Integers as ordered pairs

Integers can be defined as an ordered pair of natural numbers, where the integer is valued at: $a - b$.

For example -1 could be shown as:

$-1 = \{\{0\}, \{0, 1\}\}$

$-1 = \{\{5\}, \{5, 6\}\}$

$(a, b) = a - b$

Converting natural numbers to integers

Natural numbers can be shown as integers by using:

$(n, 0)$

Natural numbers can be converted to integers:

$\{\{a\}, \{a, 0\}\}$

Cardinality of integers

1.1.2 Ordering of the integers

Ordering integers

Integers are an ordered pair of naturals.

$\{\{x\}, \{x, y\}\}$

For example -4 can be:

$\{\{4\}, \{4, 8\}\}$

$\{\{0\}, \{0, 8\}\}$

We extend the ordering to say:

$\{\{x\}, \{x, y\}\} \leq \{\{s(x)\}, \{s(x), y\}\}$

$\{\{x\}, \{x, s(y)\}\} \leq \{\{x\}, \{x, y\}\}$

So can we define this on an arbitrary pair:

$\{\{a\}, \{a, b\}\} \leq \{\{c\}, \{c, d\}\}$

We know that:

$\{\{a\}, \{a, b\}\} = \{\{s(a)\}, \{s(a), s(b)\}\}$

And either of:

$\{\{a\}, \{a, b\}\} = \{\{0\}, \{0, A\}\}$

$\{\{a\}, \{a, b\}\} = \{\{B\}, \{B, 0\}\}$

$\{\{a\}, \{a, b\}\} = \{\{0\}, \{0, 0\}\}$

As the latter is a case of either of the other 2, we consider only the first 2.

So we can define:

$\{\{a\}, \{a, b\}\} \leq \{\{c\}, \{c, d\}\}$

As any of:

$1 : \{\{0\}, \{0, A\}\} \leq \{\{0\}, \{0, C\}\}$

$2 : \{\{0\}, \{0, A\}\} \leq \{\{D\}, \{D, 0\}\}$

$3 : \{\{B\}, \{B, 0\}\} \leq \{\{0\}, \{0, C\}\}$

$4 : \{\{B\}, \{B, 0\}\} \leq \{\{D\}, \{D, 0\}\}$

Case 1:

$\{\{0\}, \{0, A\}\} \leq \{\{0\}, \{0, C\}\}$

Trivial, depends on relative size of A and C.

Case 2:

$\{\{0\}, \{0, A\}\} \leq \{\{D\}, \{D, 0\}\}$

We can see that:

$\{\{D\}, \{D, A\}\} \leq \{\{D\}, \{D, 0\}\}$

And therefore this holds.

Case 3:

$\{\{B\}, \{B, 0\}\} \leq \{\{0\}, \{0, C\}\}$

We can see that:

$\{\{B\}, \{B, 0\}\} \leq \{\{B\}, \{B, C\}\}$

And therefore this does not hold.

Case 4:

$\{\{B\}, \{B, 0\}\} \leq \{\{D\}, \{D, 0\}\}$

Trivial, like case 1.

1.1.3 Functions of integers

Addition

Then we can define addition as:

$(a, b) + (c, d) = (a + c, b + d)$

Integer addition can then be defined:

$a + b = \{\{a_1\}, \{a_1, a_2\}\} + \{\{b_1\}, \{b_1, b_2\}\}$

$a + b = \{\{a_1 + b_1\}, \{a_1 + b_1, a_2 + b_2\}\}$

Or:

$a + b = c$

$c_1 = a_1 + b_1$

$c_2 = a_2 + b_2$

Multiplication

Similarly, multiplication can be defined as:

$(a, b).(c, d) = (ac + bd, ad + bc)$

$ab = c$

$c_1 = a_1b_1 + a_2b_2$

$c_2 = a_2b_1 + a_1b_2$

Subtraction

$a - b = c$

$c_1 = a_1 + b_2$

$c_2 = a_2 + b_1$

1.1.4 Cardinality of the integers

Cardinality of integers

Chapter 2

The rational numbers

2.1 Rational numbers

2.1.1 Rational numbers

Defining rational numbers

We previously defined integers in terms of natural numbers. Similarly we can define rational numbers in terms of integers.

$\forall ab \in \mathbb{I}(\neg(b = 0) \to \exists c(b.c = a))$

A rational is an ordered pair of integers.

$\{\{a\}, \{a, b\}\}$

So that:

$\{\{a\}, \{a, b\}\} = \frac{a}{b}$

Converting integers to rational numbers

Integers can be shown as rational numbers using:

$(i, 1)$

Integers can then be turned into rational numbers:

$\mathbb{Q} = \frac{a}{1}$

$a = \frac{a_1}{a_2}$

$b = \frac{b_1}{b_2}$

$c = \frac{c_1}{c_2}$

Equivalence classes of rationals

There are an infinite number of ways to write any rational number, as with integers. $\frac{1}{2}$ can be written as $\frac{1}{2}$, $\frac{-2}{-4}$ etc.

The class of these terms form an equivalence class.

We can show these are equal:

$\frac{a}{b} = \{\{a\}, \{a, b\}\}$

$\frac{ca}{cb} = \{\{a\}, \{a, b\}\}$

$\frac{ca}{cb} = \{\{ca\}, \{ca, cb\}\}$

$\{\{a\}, \{a, b\}\} = \{\{ca\}, \{ca, cb\}\}$

2.1.2 Ordering of rationals

2.1.3 Functions of rational numbers

Rational addition

Then we can define addition as:

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

$a + b = c$

$c_1 = a_1 b_2 + a_2 b_1$

$c_1 = a_2 b_2$

Rational subtraction

$a - b = c$

$c_1 = a_1 b_2 - a_2 b_1$

$c_1 = a_2 b_2$

Rational multiplication

Similarly, multiplication can be defined as:

$(a, b).(c, d) = (a.c, b.d)$

$ab = c$

$c_1 = a_1 b_1$

$c_1 = a_2 b_2$

Rational division

$\frac{a}{b} = c$

$c_1 = a_1 b_2$

$c_1 = a_2 b_1$

2.1.4 Cardinality of the rationals

Cardinality of rational numbers

We can see rational numbers as cartesian products of integers. That is:

$\mathbb{Q} = Z.Z$

We can order the rational numbers like so:

$\{\frac{1}{1}, \frac{2}{1}, \frac{1}{2}, \frac{1}{3}, \frac{2}{2}\frac{3}{1}...\}$

These can be mapped from natural numbers, so there is a bijunctive function.

So:

$|\mathbb{Q}| = |\mathbb{Z}.\mathbb{Z}| = |\mathbb{N}| = \aleph_0$

As: $|\mathbb{Z}.\mathbb{Z}| = |\mathbb{Z}|^2$

$|\mathbb{N}|^n = \mathbb{N}$

2.1.5 Fraction rules

Addition

$\frac{A}{B} + \frac{C}{D} = \frac{AD + BC}{BD}$

Multiplication

$\frac{A}{B}\frac{C}{D} = \frac{AC}{BD}$

b Scaler addition

$C + \frac{A}{B} = \frac{BC + A}{B}$

Scaler multiplication

$C\frac{A}{B} = \frac{AC}{B}$

Other

$$\frac{A+B}{C} = \frac{A}{C} + \frac{B}{C}$$

$$\frac{A}{B} = \frac{AC}{BC}$$

2.1.6 Partial fraction decomposition

We have: $\frac{1}{A.B}$

We want this in the form of:

$$\frac{a}{A} + \frac{b}{B}$$

First, lets define M as the mean of these two numbers, and define $\delta = M - B$. Then:

$$\frac{1}{AB} = \frac{1}{(M+\delta)(M-\delta)} = \frac{a}{M+\delta} + \frac{b}{M-\delta}$$

We can rearrange the latter two to find:

$$1 = a(M-\delta) + b(M+\delta)$$

Now we need to find values of a and b to choose.

Let's examine a.

$$a = \frac{1 - b(M+\delta)}{M-\delta}$$

$$a = -\frac{bM + b\delta - 1}{M-\delta}$$

$$a = -\frac{bM + b\delta - 1}{M-\delta}$$

For this to divide neatly we need both the numerator to be a constant multiplier of the denominator. This means the ratio the multiplier for the left hand side of the denominator is equal to the right:

$$\frac{bM}{M} = \frac{b\delta - 1}{-\delta}$$

$$b = \frac{b\delta - 1}{-\delta}$$

$$b = \frac{1}{2\delta}$$

We can do the same for a.

$$a = -\frac{1}{2\delta}$$

We can plug these back into our original formula:

$$\frac{1}{(M+\delta)(M-\delta)} = \frac{-\frac{1}{2\delta}}{M+\delta} + \frac{\frac{1}{2\delta}}{M-\delta}$$

$$\frac{1}{(M+\delta)(M-\delta)} = \frac{1}{2\delta}[\frac{1}{M-\delta} - \frac{1}{M+\delta}]$$

2.1.7 Density of the rationals

Rationals are dense in rationals

For any pair of rationals, there is another rational between them:

$a = \frac{p}{q}$

$b = \frac{m}{n}$

Where $b > a$.

We define a new rational:

$c = \frac{a+b}{2}$

$c = \frac{pn+qm}{2qn}$

This is a rational number.

We can write:

$a = \frac{2pn}{2qn}$

$b = \frac{2qm}{2qn}$

As $b > a$ we know $2qm > 2pn$

So: $a < c < b$

Chapter 3

Algebraic numbers

Chapter 4

Complex numbers

4.1 Introducing complex numbers

4.1.1 Defining complex numbers

Define as an ordered pair of reals

We have a complete set of real numbers. Do we need any more?

For the real numbers, we showed there were functions on the rational numbers which did not have rational solutions. We can similarly show that there are functions on real numbers which do not have real solutions.

Consider:

$f(x) = \sqrt{x}$

This has no real solution for $x < 0$.

We define:

$i := \sqrt{-1}$

i and $-i$ can be used interchangeably.

$(-i)^2 = (-1)^2 i^2 = i^2 = -1$

Complex numbers can be shown more generally as:

$a + bi$

We define the complex conjugate of

$x = a + bi$

As

$\bar{x} = a - bi$

Note that

$x\bar{x} = (a + bi)(a - bi) = a^2 - b^2$

We can take exponents of imaginary numbers

$c^{i\theta} = a + bi$

We know the opposite is true.

$c^{-i\theta} = a - bi$

So

$c^{i\theta}c^{-i\theta} = (a + bi)(a - bi)$

$1 = a^2 - b^2$

The case where $c = e$ is of particular note. We explore this later.

4.1.2 Real numbers aren't closed

Define as an ordered pair of reals

We have a complete set of real numbers. Do we need any more?

For the real numbers, we showed there were functions on the rational numbers which did not have rational solutions. We can similarly show that there are functions on real numbers which do not have real solutions.

Consider:

$f(x) = \sqrt{x}$

This has no real solution for $x < 0$.

We define:

$i := \sqrt{-1}$

i and $-i$ can be used interchangeably.

$(-i)^2 = (-1)^2 i^2 = i^2 = -1$

Complex numbers can be shown more generally as:

$a + bi$

We define the complex conjugate of

$x = a + bi$

As

$\bar{x} = a - bi$

Note that

$x\bar{x} = (a + bi)(a - bi) = a^2 - b^2$

We can take exponents of imaginary numbers

$c^{i\theta} = a + bi$

We know the opposite is true.

$c^{-i\theta} = a - bi$

So

$c^{i\theta}c^{-i\theta} = (a + bi)(a - bi)$

$1 = a^2 - b^2$

The case where $c = e$ is of particular note. We explore this later.

4.2 Operators on complex numbers

4.2.1 Arithmetic on complex numbers

For each of these we have:

$x = a + bi$

$y = c + di$

Addition is defined as:

$x + y = a + bi + c + di$

$x + y = (a + c) + (b + d)i$

Subtraction is defined as:

$x - y = a + bi - c - di$

$x - y = (a - c) + (b - d)i$

Multiplication is defined as:

$xy = (a + bi)(c + di)$

$xy = ac - bd + adi + bci$

$xy = (ac - bd) + (ad + bc)i$

Division is defined as:

$$\frac{x}{y} = \frac{a + bi}{c + di}$$

$$\frac{x}{y} = \frac{(a + bi)(c - di)}{(c + di)(c - di)}$$

$$\frac{x}{y} = \frac{(ac + bd) + (bc - ad)i}{c^2 + d^2}$$

4.2.2 Complex conjugate

We have $z = a + bi$.

The complex conjugate is:

$\bar{z} = a - bi$

4.2.3 Absolute value

$|z| = \sqrt{z\bar{z}}$

$|z| = \sqrt{(a + bi)(a - bi)}$

$|z| = \sqrt{a^2 + b^2}$

4.3 Results

4.3.1 Roots of unity

4.3.2 Complex logarithms

4.3.3 Disks

A disk is the area contained by a circle.

An open disk at (a, b) of radius r is:

$\{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 < r^2\}$

For a closed disk it is:

$\{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 \leq r^2\}$

4.3.4 Disks

We defined an open disk at (a, b) of radius r as:

$\{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 < r^2\}$

For a closed disk it is:

$\{(x, y) \in \mathbb{R}^2 : (x - a)^2 + (y - b)^2 \leq r^2\}$

4.3.5 Annulus

An annulus is a disk, which excludes a smaller disk inside the disk

4.3.6 Punctured disk

If the interior disk is just a point, it is a punctured disk.

Chapter 5

Infinite sequences and limits

5.1 More on sequences

5.1.1 Limit of a sequence

A sequence converges to a limit if

Can converge to a number (1/x)

Can converge to +/- infinity (x)

Otherwise, does not converge (1,-1,1,-1. . .)

Superior and inferior limits

A bounded increasing sequence converges to least upper bound

Identifying the limit of a sequence

Direct comparison test

Root test

5.2 Divergent series

5.2.1 Partial sum

Take a series. We can define the partial sum as:

$s_k = \sum_{i=1}^{k} a_i$

5.2.2 Cesàro sum

The Cesàro sum is the limit of the average of the first n partial sums.

That is:

$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} s_k$

Consider the sequence $\{1, -1, 1, -1, ...\}$

The partial sum is:

$s_k = \sum_{i=1}^{k} a_i$

$s_k = k \mod (2)$

The Cesàro sum is: $\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} s_k$

$\lim_{n\to\infty} \frac{1}{n} \sum_{k=1}^{n} k \mod (2)$

$\frac{1}{2}$

5.2.3 Abel summation

Part II

Elementary algebra

Chapter 6

Solving quadratic polynomials

6.1 Single-variable polynomials

6.1.1 Introduction

A single-variable polynomial is an equation of the form:

$\sum_{i=0}^{n} a_i x^i = 0$

For example:

- $x = 1$
- $x^2 = 4$
- $x^2 - 3x + 2 = 0$

6.1.2 Degrees

The degree of a polynomial is the highest-order term.

For example $x^3 + x = 0$ has degree 3.

6.1.3 Roots of single-variable polynomials

A solution to a polynomial is a root.

For example 1 and 2 are roots of $x^2 - 3x + 2 = 0$

6.2 Solving quadratic polynomials

6.2.1 Quadratic polynomials

Quadratic polynomials are of the form $ax^2 + bx + c = 0$.

6.2.2 Solving quadratic polynomials

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

6.2.3 Proof

We can get the two solutions to a quadratic equation from the following manipulation.

$$ax^2 + bx + c = 0$$

$$a[x^2 + \frac{b}{a}x] = -c$$

$$a[(x + \frac{b}{2a})^2 - \frac{b^2}{4a^2}] = -c$$

$$a[(x + \frac{b}{2a})^2] = \frac{b^2}{4a} - c$$

$$(x + \frac{b}{2a})^2 = \frac{b^2 - 4ac}{4a^2}$$

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

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

Chapter 7

Solving cubic polynomials: Solving depressed cubic equations and reducing to depressed cubic form

7.1 Solving cubic polynomials

7.1.1 Cubic polynomials

Cubic polynomials are of the form $ax^3 + bx^2 + cx + d = 0$.

7.1.2 Solving specific cases

We start by solving when $b = 0$, that is:

$aX^3 + bx + c = 0$

7.1.3 Solving the general case

Chapter 8

Elliptic curves and finite-field elliptic curves

8.1 Elliptic curves

8.1.1 Elliptic curves

Of the form $y^2 = x^3 + ax + b$.

group law: line on graph which hits two points will hit 3rd and no more

defines an abelian group "addition" operation for points on a curve, if we use the one on the opposite (ie flip y value)a

8.1.2 Finite-field elliptic curves

not a curve. collection of points in int mod p

addition operation still exists

Chapter 9

Generating functions

9.1 Generating functions

9.1.1 Generating functions

Definition

A series can be described as:

$\sum_{i=0}^{\infty} s_i x^i$

If we know the function equal to this series, we can identify the ith number.

9.1.2 Fibonacci sequence

The generating function

Let's use a generating function to create a function for the Fibonacci sequence's cth digit. $F(c) = \sum_{i=c} x^i s_i$

Let's look at it for other starts:

$F(c+k) = \sum_{i=c} x^{i+k} s_{i+k}$

$F(c+k) = \sum_{i=c+k} x^i s_i$

$F(c+1) = \sum_{i=c} x^{i+1} s_{i+1}$

$F(c+2) = \sum_{i=c} x^{i+2} s_{i+2}$

This means

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^i s_i x^2 + \sum_{i=c} x^{i+1} s_{i+1} x$

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2} s_i + \sum_{i=c} x^{i+2} s_{i+1}$

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2}(s_i + s_{i+1})$

Using the definiton of the Fibonacci sequence

From the definition of the fibonacci sequence, $s_i + s_{i+1} = s_{i+2}$.

$F(c)x^2 + F(c+1)x = \sum_{i=c} x^{i+2}(s_{i+2})$

$F(c)x^2 + F(c+1)x = F(c+2)$

Reducing the functions

Next, we expand out $F(c+1)$ and $F(c+2)$.

$F(c) - F(c+k) = \sum_{i=c} x^i s_i - \sum_{i=c+k} x^i s_i$

$F(c) - F(c+k) = \sum_{i=c}^{c+k} x^i s_i$

$F(c+k) = F(c) - \sum_{i=c}^{c+k} x^i s_i$

So:

$F(c+1) = F(c) - \sum_{i=c}^{c+1} x^i s_i$

$F(c+1) = F(c) - x^c s_c$

$F(c+2) = F(c) - \sum_{i=c}^{c+2} x^i s_i$

$F(c+2) = F(c) - x^{c+1} s_{c+1} - x^c s_c$

Let's take our previous equation

$F(c)x^2 + F(c+1)x = F(c+2)$

$F(c)x^2 + [F(c) - x^c s_c]x = F(c) - x^{c+1} s_{c+1} - x^c s_c$

$F(c)x^2 + F(c)x - x^{c+1} s_c = F(c) - x^{c+1} s_{c+1} - x^c s_c$

$F(c)[x^2 + x - 1] = x^{c+1} s_c - x^{c+1} s_{c+1} - x^c s_c$

$F(c) = \dfrac{x^c s_c + x^{c+1} s_{c+1} - x^{c+1} s_c}{1 - x - x^2}$

Using the first element in the sequence

For the start of the sequence, $c = 0$, $s_0 = s_1 = 1$.

$F(0) = \dfrac{x^0 1 + x - x}{1 - x - x^2}$

$F(0) = \dfrac{1}{1 - x - x^2}$

Let's factorise this:

$F(0) = \dfrac{-1}{(x + \frac{1}{2} + \frac{\sqrt{5}}{2})(x + \frac{1}{2} - \frac{\sqrt{5}}{2})}$

We can then use partial fraction decomposition

$$\frac{1}{(M+\delta)(M-\delta)} = \frac{1}{2\delta}[\frac{1}{M-\delta} - \frac{1}{M+\delta}]$$

To show that

$$F(0) = \frac{-1}{\sqrt{5}}[\frac{1}{x+\frac{1}{2}-\frac{\sqrt{5}}{2}} - \frac{1}{x+\frac{1}{2}+\frac{\sqrt{5}}{2}}]$$

$$F(0) = \frac{-1}{\sqrt{5}}[\frac{\frac{1}{2}+\frac{\sqrt{5}}{2}}{(x+\frac{1}{2}-\frac{\sqrt{5}}{2})(\frac{1}{2}+\frac{\sqrt{5}}{2})} - \frac{\frac{1}{2}-\frac{\sqrt{5}}{2}}{(x+\frac{1}{2}+\frac{\sqrt{5}}{2})(\frac{1}{2}-\frac{\sqrt{5}}{2})}]$$

$$F(0) = \frac{-1}{\sqrt{5}}[\frac{\frac{1}{2}+\frac{\sqrt{5}}{2}}{x(\frac{1}{2}+\frac{\sqrt{5}}{2})-1} - \frac{\frac{1}{2}-\frac{\sqrt{5}}{2}}{x(\frac{1}{2}-\frac{\sqrt{5}}{2})-1}]$$

$$F(0) = \frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})\frac{1}{1-x(\frac{1}{2}+\frac{\sqrt{5}}{2})} - (\frac{1}{2}-\frac{\sqrt{5}}{2})\frac{1}{1-x(\frac{1}{2}-\frac{\sqrt{5}}{2})}]$$

Finishing off

As we know

$$\frac{1}{1-x} = \sum_{i=0} x^i$$

So

$$F(0) = \frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})\sum_{i=0} x^i(\frac{1}{2}+\frac{\sqrt{5}}{2})^i - (\frac{1}{2}-\frac{\sqrt{5}}{2})\sum_{i=0} x^i(\frac{1}{2}-\frac{\sqrt{5}}{2})^i]$$

$$F(0) = \frac{1}{\sqrt{5}}[\sum_{i=0} x^i(\frac{1}{2}+\frac{\sqrt{5}}{2})^{i+1} - \sum_{i=0} x^i(\frac{1}{2}-\frac{\sqrt{5}}{2})^{i+1}]$$

$$F(0) = \frac{1}{\sqrt{5}}\sum_{i=0} x^i[(\frac{1}{2}+\frac{\sqrt{5}}{2})^{i+1} - (\frac{1}{2}-\frac{\sqrt{5}}{2})^{i+1}]$$

So the nth number in the sequence (treating $n = 1$ as the first number) is:

$$\frac{1}{\sqrt{5}}[(\frac{1}{2}+\frac{\sqrt{5}}{2})^n - (\frac{1}{2}-\frac{\sqrt{5}}{2})^n]$$

Chapter 10

Diophantine equations

Part III

Systems of linear equations

Chapter 11

Solving systems of linear equations

11.1 Introduction

11.1.1 Introduction

$m_{11}x + m_{12}y + m_{13}z = v_1$

$m_{21}x + m_{22}y + m_{23}z = v_2$

$m_{31}x + m_{32}y + m_{33}z = v_3$

11.1.2 Matrix and vector notation

We can write the above as:

$\mathbf{M}x = \mathbf{v}$

What are the properties of $\mathbf{M}$ and $\mathbf{v}$?

They are linear in addition and scalar multiplication.

11.2 Rank

11.2.1 Matrix rank

Rank function

The rank of a matrix is the dimension of the span of its component columns.

$rank(M) = span(m_1, m_2, ..., m_n)$

Column and row span

The span of the rows is the same as the span of the columns.

11.2.2 Types of matrices

Empty matrix

A matrix where every element is 0. There is one for each dimension of matrix.

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

11.2.3 Triangular matrix

A matrix where $a_{ij} = 0$ where $i < j$ is upper triangular.

A matrix where $a_{ij} = 0$ where $i > j$ is lower triangular.

A matrix which is either upper or lower triangular is a triangular matrix.

11.2.4 Symmetric matrices

All symmetric matrices are square.

The identity matrix is an example.

A matrix where $a_{ij} = a_{ji}$ is symmetric.

11.2.5 Diagonal matrix

A matrix where $a_ij = 0$ where $i \neq j$ is diagonal.

All diagonal matrices are symmetric.

The identity matrix is an example.

11.3 Other

11.3.1 Basis of an endomorphism

11.3.2 Changing the basis

For any two bases, there is a unique linear mapping from of the element vectors to the other.

Chapter 12

Inverting matrices

12.1 Inversion

12.1.1 Inverse matrices

An invertible matrix implies that if the matrix is multiplied by another matrix, the original matrix can be recovered.

That is, if we have matrix A, there exists matrix A^{-1} such that $AA^{-1} = I$.

Consider a linear map on a vector space.

$Ax = y$

If A is invertible we can have:

$A^{-1}Ax = A^{-1}y$

$x = A^{-1}y$

If we set $y = \mathbf{0}$ then:

$x = \mathbf{0}$

So if there is a non-zero vector x such that:

$Ax = \mathbf{0}$ then A is not invertible.

12.1.2 Left and right inverses

That is, for all matrices A, the left and right inverses of B, B_L^{-1} and B_R^{-1}, are defined such that:

$A(BB_R^{-1}) = A$

$A(B_L^{-1}B) = A$

Left and right inversions are equal

Note that if the left inverse exists then:

$B_L^{-1}B = I$

And if the right inverse exists:

$BB_R^{-1} = I$

Let's take the first:

$B_L^{-1}B = I$

$B_L^{-1}BB_L^{-1} = B_L^{-1}$

$B_L^{-1}BB_L^{-1} - B_L^{-1} = 0$

$B_L^{-1}(BB_L^{-1} - I) = 0$

12.1.3 Inversion of products

$(AB)(AB)^{-1} = I$

$A^{-1}AB(AB)^{-1} = A^{-1}$

$B^{-1}B(AB)^{-1} = B^{-1}A^{-1}$

$(AB)^{-1} = B^{-1}A^{-1}$

12.1.4 Inversion of a diagonal matrix

$DD^{-1} = I$

$D_{ii}D_{ii}^{-1} = 1$

$D_{ii}^{-1} = \frac{1}{D_{ii}}$

12.1.5 Degenerate (singular) matrices

12.1.6 Elementary row operations

Some operations to a matrix can be reversed to arrive at the original matrix. Trivially, multiplying by the identity matrix is reversible.

Similarly, some operations are not reversible. Such as multiplying by the empty matrix.

All matrix operations which can be reversed are combinations of 3 elementary row operations. These are: Swapping rows

$$T_{12} = \begin{bmatrix} 0 & 1 & ... & 0 \\ 1 & 0 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

Multiplying rows by a vector

$$D_2(m) = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & m & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

Adding rows to other rows

$$L_{12}(m) = \begin{bmatrix} 1 & 0 & ... & 0 \\ m & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix}$$

12.1.7 Gaussian elimination

Simultaneous equations

Matricies can be used to solve simultaneous equations. Condsider the following set of equations.

- $2x + y - z = 8$
- $-3x - y + 2z = -11$
- $-2x + y + 2z = -3$

We can write this in matrix form.

$Ax = y$

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

$$x = \begin{bmatrix} x \\ y \\ z \end{bmatrix}$$

$$y = \begin{bmatrix} 8 \\ -11 \\ -3 \end{bmatrix}$$

Augmented matrix

Consider a form for summarising these equations. This is the augmented matrix.

$$(A|y) = \begin{bmatrix} 2 & 1 & -1 & | & -8 \\ -3 & -1 & 2 & | & -11 \\ -2 & 1 & 2 & | & -3 \end{bmatrix}$$

We can take this and recovery our original A and y.

However we can also do things to this augmented matrix which preserve solutions to the set of equations. These are:

Undertaking combinations of these can make it easier to solve the equation. In particular, if we can arrive at the form:

$$(A|y) = \begin{bmatrix} 1 & 0 & 0 & | & a \\ 0 & 1 & 0 & | & b \\ 0 & 0 & 1 & | & c \end{bmatrix}$$

The solutions for x, y, z are a, b, c.

Echeleon / triangular form

We first aim for:

$$(A|y) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & | & a \\ 0 & a_{22} & a_{23} & | & b \\ 0 & 0 & a_{33} & | & c \end{bmatrix}$$

If this cannot be reached there is no single solution. There may be infinite or no solutions.

Solving

Once we have the triangular form, we can easily solve.

$$(A|y) = \begin{bmatrix} a_{11} & a_{12} & a_{13} & | & a \\ 0 & a_{22} & a_{23} & | & b \\ 0 & 0 & a_{33} & | & c \end{bmatrix}$$

$$(A|y) = \begin{bmatrix} 1 & 0 & 0 & | & a \\ 0 & 1 & 0 & | & b \\ 0 & 0 & 1 & | & c \end{bmatrix}$$

This process is back substitution (or forward substitution if the matrix is triangular the other way).

Matrix inversion

We can think of the inverse of a matrix as one which which takes a series of reverible operations and does these to a matrix then arriving at the identity matrix.

That is, only the three elementary row operations, and combinations of them, can transform a matrix in a way in which it can be reversed. As such All reversible matricies are combinations of the identity matrix and a series of elementary row operations. The inverse matrix is then those series of row operations, in reverse.

We can find identify an inversion by undertaking gaussian elimination. Each step done on the matrix is done to the identify matrix, reversing the process. The end result is the inverted matrix.

Instead of:

$$(A|y) = \begin{bmatrix} 2 & 1 & -1 & | & -8 \\ -3 & -1 & 2 & | & -11 \\ -2 & 1 & 2 & | & -3 \end{bmatrix}$$

Take:

$$(A|I) = \begin{bmatrix} 2 & 1 & -1 & | & 1 & 0 & 0 \\ -3 & -1 & 2 & | & 0 & 1 & 0 \\ -2 & 1 & 2 & | & 0 & 0 & 1 \end{bmatrix}$$

When we solve this we get:

$$(I|A^{-1}) = \begin{bmatrix} 1 & 0 & 0 & | & \frac{3}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 1 & 0 & | & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & | & \frac{1}{4} & \frac{1}{2} & \frac{3}{4} \end{bmatrix}$$

Chapter 13

Eigenvalues, Eigenvectors, decomposition and operations

13.1 Eigenvalues and eigenvectors

13.1.1 Eigenvalues and eigenvectors

Which vectors remain unchanged in direction after a transformation?

That is, for a matrix A, what vectors v are equal to scalar multiplication by λ following the operation of the matrix.

$Av = \lambda v$

13.1.2 Spectrum

The spectrum of a matrix is the set of its eigenvalues.

13.1.3 Eigenvectors as a basis

If eigen vectors space space, we can write

$v = \sum_i \alpha_i |\lambda_i\rangle$

Under what circumstances do they span the entirity?

13.1.4 Calculating eigenvalues and eigenvectors using the characteristic polynomial

The characteristic polynomial of a matrix is a polynomial whose roots are the eigenvalues of the matrix.

We know from the definition of eigenvalues and eigenvectors that:

$Av = \lambda v$

Note that

$Av - \lambda v = 0$

$Av - \lambda I v = 0$

$(A - \lambda I)v = 0$

Trivially we see that $v = 0$ is a solution.

Otherwise matrix $A - \lambda I$ must be non-invertible. That is:

$Det(A - \lambda I) = 0$

13.1.5 Calculating eigenvalues

For example

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

$A - \lambda I = \begin{bmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{bmatrix}$

$Det(A - \lambda I) = (2-\lambda)(2-\lambda) - 1$

When this is 0.

$(2-\lambda)(2-\lambda) - 1 = 0$

$\lambda = 1, 3$

13.1.6 Calculating eigenvectors

You can plug this into the original problem.

For example

$Av = 3v$

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

As vectors can be defined at any point on the line, we normalise $x_1 = 1$.

$\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 3 \\ 3x_2 \end{bmatrix}$

Here $x_2 = 1$ and so the eigenvector corresponding to eigenvalue 3 is:

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

13.1.7 Traces

The trace of a matrix is the sum of its diagonal components.

$Tr(M) = \sum_i^n m_{ii}$

The trace of a matrix is equal to the sum of its eigenvectors.

Traces can be shown as the sum of inner products.

$Tr(M) = \sum_i^n e_i M e^i$

13.1.8 Properties of traces

Traces commute

$Tr(AB) = Tr(BA)$

Traces of 1×1 matrices are equal to their component.

$Tr(M) = m_{11}$

13.1.9 Trace trick

If we want to manipulate the scalar:

$v^T M v$

We can use properties of the trace.

$v^T M v = Tr(v^T M v)$

$v^T M v = Tr([v^T][Mv])$

$v^T M v = Tr([Mv][v^T])$

$v^T M v = Tr(M v v^T)$

13.2 Matrix operations

13.2.1 Matrix powers

For a square matrix M we can calculate $MMMM...$, or M^n where $n \in \mathbb{N}$.

13.2.2 Powers of diagonal matrices

Generally, calculating a matrix to an integer power can be complicated. For diagonal matrices it is trivial.

For a diagonal matrix $M = D^n$, $m_{ij} = d_{ij}^n$.

13.2.3 Matrix exponentials

The exponential of a complex number is defined as:

$e^x = \sum \frac{1}{j!} x^j$

We can extend this definition to matrices.

$e^X := \sum \frac{1}{j!} X^j$

The dimension of a matrix and its exponential are the same.

13.2.4 Matrix logarithms

If we have $e^A = B$ where A and B are matrices then we can say that A is matrix logarithm of B.

That is:

$\log B = A$

The dimensions of a matrix and its logarithm are the same.

13.2.5 Matrix square roots

For a matrix M, the square root $M^{\frac{1}{2}}$ is A where $AA = M$.

This does not necessarily exist.

Square roots may not be unique.

Real matrices may have no real square root.

13.3 Matrix decomposition

13.3.1 Similar matrices

In hermitian, show all symmtric matrices are hermitian

For a diagonal matrix, eigenvalues are the diagonal entries?

Similar matrix:

$M = P^{-1}AP$

M and A have the same eigenvalues. If A diagonal, then entries are eigenvalues.

13.3.2 Defective and diagonalisable matrices

13.3.3 Diagonalisable matrices and eigendecomposition

If matrix M is diagonalisable if there exists matrix P and diagonal matrix A such that:

$M = P^{-1}AP$

Diagonalisiable matrices and powers

If these exist then we can more easily work out matrix powers.

$M^n = (P^{-1}AP)^n = P^{-1}A^nP$

A^n is easy to calculate, as each entry in the diagonal taken to the power of n.

Defective matrices

Defective matrices are those which cannot be diagonalised.

Non-singular matries can be defective or not defective, for example the identiy matrix.

Singular matrices can also be defective or not defective, for example the empty matrix.

Eigen-decomposition

Consider an eigenvector v and eigenvalue λ of matrix M.

We known that $Mv = \lambda v$.

If M is full rank then we can generalise for all eigenvectors and eigenvalues:

$MQ = Q\Lambda$

Where Q is the eigenvectors as columns, and Λ is a diagonal matrix with the corresponding eigenvalues. We can then show that:

$M = Q\Lambda Q^{-1}$

This is only possible to calculate if the matrix of eigenvectors is non-singular. Otherwise the matrix is defective.

If there are linearly dependent eigenvectors then we cannot use eigen-decomposition.

13.3.4 Using the eigen-decomposition to invert a matrix

This can be used to invert M.

We know that:

$M^{-1} = (Q\Lambda Q^{-1})^{-1}$

$M^{-1} = Q^{-1}\Lambda^{-1}Q$

We know Λ can be easily inverted by taking the reciprocal of each diagonal element. We already know both Q and its inverse from the decomposition.

If any eigenvalues are 0 then Λ cannot be inverted. These are singular matrices.

13.3.5 Spectral theorem for finite-dimensional vector spaces

13.4 Other

13.4.1 Commutation

We define a function, the commuter, between two objects a and b as:

$[a, b] = ab - ba$

For numbers, $ab - ba = 0$, however for matrices this is not generally true.

13.4.2 Commutators and eigenvectors

Consider two matrices which share an eigenvector v.

$Av = \lambda_A v$

$Bv = \lambda_B v$

Now consider:

$ABv = A\lambda_B v$

$ABv = \lambda_A \lambda_B v$

$BAv = \lambda_A \lambda_B v$

If the matrices share all the same eigenvectors, then the matrices commute, and $AB = BA$.

13.4.3 Identity matrix and the Kronecker delta

13.4.4 Matrix additon and multiplication

Matrix multiplication

$A = A^{mn}$

$B = B^{no}$

$C = C^{mo} = A.B$

$c_{ij} = \sum_{r=1}^{n} a_{ir}b_{rj}$

Matrix multiplication depends on the order. Unlike for real numbers,

$AB \neq BA$

Matrix multiplication is not defined unless the condition above on dimensions is met.

A matrix multiplied by the identity matrix returns the original matrix.

For matrix $M = M^{mn}$

$M = MI^m = I^n M$

Matrix addition

2 matricies of the same size, that is with idental dimensions, can be added together.

If we have 2 matrices A^{mn} and B^{mn}

$C = A + B$

$c_{ij} = a_{ij} + b_{ij}$

An empty matrix with 0s of the same size as the other matrix is the identity matrix for addition.

Scalar multiplication

A matrix can be multiplied by a scalar. Every element in the matrix is multiplied by this.

$B = cA$

$b_{ij} = ca_{ij}$

The scalar 1 is the identity scalar.

13.4.5 Transposition and conjugation

Transposition

A matrix of dimensions $m * n$ can be transformed into a matrix $n * m$ by transposition.

$B = A^T$

$b_{ij} = aji$

Transpose rules

$(M^T)^T = M$

$(AB)^T = B^T A^T$

$(A + B)^T = A^T + B^T$

$(zM)^T = zM^T$

Conjugation

With conjugation we take the complex conjugate of each element.

$B = \overline{A}$

$b_{ij} = \overline{a}_{ij}$

Conjugation rules

$\overline{(\overline{A})} = A$

$\overline{(AB)} = (\overline{A})(\overline{B})$

$\overline{(A + B)} = \overline{A} + \overline{B}$

$\overline{(zM)} = \overline{z}\overline{M}$

Conjugate transposition

Like transposition, but with conjucate.

$B = A^*$

$b_{ij} = \bar{a_{ji}}$

Alternatively, and particularly in physics, the following symbol is often used instead.

$(A^*)^T = A^\dagger$

13.4.6 Matrix rank

Rank function

The rank of a matrix is the dimension of the span of its component columns.

$rank(M) = span(m_1, m_2, ..., m_n)$

Column and row span

The span of the rows is the same as the span of the columns.

13.4.7 Types of matrices

Empty matrix

A matrix where every element is 0. There is one for each dimension of matrix.

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

13.4.8 Triangular matrix

A matrix where $a_{ij} = 0$ where $i < j$ is upper triangular.

A matrix where $a_{ij} = 0$ where $i > j$ is lower triangular.

A matrix which is either upper or lower triangular is a triangular matrix.

13.4.9 Symmetric matrices

All symmetric matrices are square.

The identity matrix is an example.

A matrix where $a_{ij} = a_{ji}$ is symmetric.

13.4.10 Diagonal matrix

A matrix where $a_i j = 0$ where $i \neq j$ is diagonal.

All diagonal matrices are symmetric.

The identity matrix is an example.

13.5 Multilinear forms and determinants

13.5.1 Multilinear forms

13.5.2 Determinants

From invertible matrix section in endo

A matrix can only be inverted if it can be created from a combination of elementary row operations.

How can we identify if a matrix is invertible? We want to create a scalar from the matrix which tells us if this possible. We can this scalar the determinant.

For a matrix A we label the determinant $|A|$, or $\det A$

We propose $|A| = 0$ when the matrix is not invertible.

So how can we identify the function we need to undertake on the matrix?

New 1

We know that linear dependence results in determinants of 0.

We can model this as a function on the columns of the matrix.

$\det M = \det([M_1, ..., M_n)$

If there is linear depednence, for example if two columns are the same then:

$\det([M_1, ..., M_i, ..., M_i, ..., M_n]) = 0$

Similarly, if there is a column of 0 then the determinant is 0.

$\det([M_1, ..., 0, ..., M_n]) = 0$

New 2

Show linear in addition

How can we identify the determinant of less simple matrices? We can use the multilinear form.

$\sum c_i \mathbf{M}_i = \mathbf{0}$

Where $\mathbf{c} \neq \mathbf{0}$

Or:

$M\mathbf{c} = \mathbf{0}$

Rule 1: Columns of matrices can be the input to a multilinear form

A matrix can be shown in terms of its columns. $A = [v_1, ..., v_n]$

$\det A = \det[v_1, ..., v_n]$

$\det A = \sum_{k_1=1}^m ... \sum_{k_n=1} \prod_{i=1}^m a_{ik_i} \det([e_{k_1}, ..., e_{k_n}])$

Multiplying a matrix by a constant multiplies the determinant by the same amount

If a whole row or columns is 0 then:

$\det A = \det[v_1, ..., v_i, ..., v_n]$

$\det A' = \det[v_1, ..., cv_i, ..., v_n]$

$\det A = \det[v_1, ..., v_i, ..., v_n]$

$\det A' = \det[v_1, ..., cv_i, ..., v_n]$

$\det A' = c \det[v_1, ..., v_i, ..., v_n]$

$\det A' = c \det A$

As a result, multiplying a column by 0 makes the determinant 0.

A matrix with a column of 0 therefore has determinant 0

Rule 2: A matrix with equal columns has a determinant of 0.

$A = [a_1, ..., a_i, ..., a_i, ..., a_n]$

$D(A) = D([a_1, ..., a_i, ..., a_i, ..., a_n])$

We know from Result 3 that swapping columns reverses the sign. Reversing columns results in the same matrix, so the determinant must be unchanged.

$D(A) = -D(A)$

$D(A) = 0$

Linear dependence

If a column is a linear combination of other columns, then the matrix cannot be inverted.

$A = [a_1, ..., \sum_{j \neq i}^n c_j a_j, ..., a_n]$

$\det A = \det([v_1, ..., \sum_{j \neq i}^n c_j v_j, ..., v_n])$

$\det A = \sum_{j \neq i}^n c_j \det([v_1, ..., v_j, ..., v_n])$

$\det A = \sum_{j \neq i}^n c_j \det([v_1, ..., v_j, , ..., v_j, ..., v_n])$

As there is a repeating vector:

$\det A = 0$

Swapping columns multiplies the determinant by -1

$A = [v_1, ..., v_i + v_j, ..., v_i + v_j, ..., v_n]$

We know.

$\det A = 0$

$\det A = \det([a_1, ..., a_i, ..., a_i, ..., a_n]) + \det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) + \det([a_1, ..., a_j, ..., a_j, ..., a_n])$

So:

$\det([a_1, ..., a_i, ..., a_i, ..., a_n]) + \det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) + \det([a_1, ..., a_j, ..., a_j, ..., a_n]) = 0$

As 2 of these have equal columns these are equal to 0.

$\det([a_1, ..., a_i, ..., a_j, ..., a_n]) + \det([a_1, ..., a_j, ..., a_i, ..., a_n]) = 0$

$\det([a_1, ..., a_i, ..., a_j, ..., a_n]) = -\det([a_1, ..., a_j, ..., a_i, ..., a_n])$

Calculating the determinant

We have

$\det A = \sum_{k_1=1}^m ... \sum_{k_n=1} \prod_{i=1}^m a_{ik_i} \det([e_{k_1}, ..., e_{k_n}])$

So what is the value of the determinant here?

We know that the determinant of the identity matrix is 1.

We know that the determinant of a matrix with identical columns is 0.

We know that swapping columns multiplies the determinant by -1.

Therefore the determinants where the values of k are not all unique are 0.

The determinants of the others are either -1 or 1 depending on how many swaps are required to restore to the identity matrix.

This is also shown as the Leibni formula.

$\det A = \sum_{\sigma \in S_n} sgn(\sigma) \prod_{i=1}^n a_{i,\sigma_i}$

13.5.3 Properties of determinants

Identity

$\det I = 1$

Multiplication

$\det(AB) = \det A \det B$

Inverse

$\det(M^{-1}) = \dfrac{1}{\det M}$

We know this because:

$\det(MM^{-1}) = \det I = 1$

$\det M \det M^{-1} = 1$

$\det(M^{-1}) = \dfrac{1}{\det M}$

Complex cojugate

$\det(M^*) = \overline{\det M}$

Transpose

$\det(M^T) = \det M$

Addition

$\det(A + B) = \det A + \det B$

Scalar multiplication

$\det cM = c^n \det M$

Determinants and eigenvalues

The determinant is equal to the product of the eigenvalues.

13.5.4 Determinants of 2x2 and 3x3 matrices

The determinant of a 2x2 matrix

$M = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

$|M| = ad - bc$

The determinant of a 3x3 matrix

$M = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$

$|M| = aei + bfg + cdh - ceg - dbi - afh$

Part IV

Inequalities

Chapter 14

Arthmetic/geometric mean inequality, Shur's inequality and Muirhead's inequality

14.1 Introduction

14.1.1 The arithmetic mean/geometric mean inequality

14.1.2 Shur's inequality

14.1.3 Muirhead's inequality