High School Mathematics Extensions

From Wikibooks, the open-content textbooks collection

Note: current version of this book can be found at http://en.wikibooks.org/wiki/High_school_extensions"

Contents

Authors

Zhuo Jia Dai, User:R3m0t, Martin Warmer, Tom Lam.

Primes and modular arithmetic

Primes

Introduction

A prime number (or prime for short) is a natural number that can only be wholly divided by 1 and itself. For theoretical reasons, the number 1 is not considered a prime (we shall see why later on in this chapter). For example, 2 is a prime, 3 is prime, and 5 is prime, but 4 is not a prime because 4 divided by 2 equals 2 without a remainder.

The first 20 primes are 2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, 47, 53, 59, 61, 67, 71.

Primes are an endless source of fascination for mathematicians. Some of the problems concerning primes are so difficult that even decades of work by some of the most brilliant mathematicians have failed to solve them. One such problem is Goldbach's conjecture, which states that all even numbers greater than 3 can be expressed as the sum of two primes.

Geometric meaning of primes

Let's start with an example. Given 12 pieces of square floor tiles, can we assemble them into a rectangular shape in more than one way? Of course we can, this is due to the fact that

$$\begin{aligned} 12 &= 12 \times 1 \\ &= 6 \times 2 \\ &= 4 \times 3 \end{aligned}$$

We do not distinguish between 2×6 and 6×2 because they are essentially equivalent arrangements.

But what about the number 7? Can you arrange 7 square floor tiles into rectangular shapes in more than one way? The answer is no, because 7 is a prime number.

Fundamental Theorem of Arithmetic

A **theorem** is a non-obvious mathematical fact. A theorem must be proven; a proposition that is generally believed to be true, but without a proof, is called a **conjecture**.

With those definitions out of the way the fundamental theorem of arithmetic simply states that:

Any natural number (except for 1) can be expressed as the product of primes in one and only one way.

For example

$$12 = 2 \times 2 \times 3$$

Rearranging the multiplication order is not considered a different representation of the number, so there is no other way of expressing 12 as the product of primes.

A few more examples

$$
\begin{aligned}
99 &= 3 \times 3 \times 11 \\
52 &= 2 \times 2 \times 13 \\
17 &= 17
\end{aligned}
$$

A number that can be factorised into more than 1 prime factor is a called a composite number (or composite for short). Composite is the opposite of prime.

Think about it

Bearing in mind the definition of the fundamental theorem of arithmetic, why isn't the number 1 considered a prime?

Factorisation

We know from the fundamental theorem of arithmetic that any integer can be expressed as the product of primes. The million dollar question is: given a number *x*, is there an *easy* way to find all prime factors of *x*?

If *x* is a small number that is easy. For example 90 = 2 × 3 × 3 × 5. But what if *x* is large? For example *x = 4539*? Most people can't factorise 4539 into primes in their heads. But can a computer do it? Yes, the computer can factorise 4539 in no time. In fact 4539 = 3 × 17 × 89.

There is, indeed, an *easy* way to factorise a number into prime factors. Just apply the method to be described below (using a computer). However, that method is too slow for large numbers: trying to factorise a number with thousands of digits would take more time than the current age of the universe. But is there a *fast* way? Or more precisely, is there an *efficient* way? There may be, but no one has found one yet. Some of the most widely used encryption schemes today (such as RSA) make use of the fact that we can't factorise large numbers into prime factors quickly. If such a method is found, a lot of internet transactions will be rendered unsafe. So if you happen to be the discoverer of such a method, don't be too forthcoming in publishing it, consult your national security agency first!

Since computers are very good at doing arithmetic, we can work out all the factors of *x* by simply instructing the computer to divide *x* by 2 and then 3 then by 5 then by 7 then by 11 ... and so on, and check whether at any point the result is a whole number. Consider the following three examples of the dividing method in action.

Example 1

x = 21

x / 2 = 10.5 not a whole number

$x / 3 = 7$ hence 3 and 7 are the factors of 21.

Example 2

$x = 153$

$x / 2 = 76.5$ hence 2 is not a factor of 153

$x / 3 = 51$ hence 3 and 51 are factors of 153

51 / 3 = 17 hence 3 and 17 are factors of 153

It is clear that 3, 9, 17 and 51 are the factors of 153. The prime factors of 153 are 3, 3 and 17 ($3 \times 3 \times 17 = 153$)

Example 3

2057 / 2 = 1028.5

...

2057 / 11 = 187

187 / 11 = 17

hence 11, 11 and 17 are the prime factors of 2057.

Exercise

Factor the following numbers:

1. 13
2. 26
3. 59
4. 82
5. 101
6. 121
7. 2187 Give up if it takes too long. There is a quick way.

Fun Fact -- Is this prime?

Interestingly, due to recent developments, we can tell quickly, with the help of a computer program, *whether* a number is prime with 100% accuracy.

2, 5 and 3

The primes 2, 5, and 3 hold a special place in factorisation. Firstly, all even numbers have 2 as one of their prime factors. Secondly, all numbers whose last digit is 0 or 5 can be divided wholly by 5.

The third case, where 3 is a prime factor, is the focus of this section. The underlying question is: is there a simple way to decide whether a number has 3 as one of its prime factors? Yes. See the following theorem

Theorem - Divisibility by 3

A number is divisible by 3 if and only if the ***sum of its digits*** *is divisible by 3*

E.g. 272 is not divisible by 3, because 2+7+2=11. And 11 isn't divisible by 3.

945 is divisible by 3, because 9+4+5 = 18. And 18 is divisible by 3. In fact 945 / 3 = 315

Is 123456789 divisible by 3?

$1 + 2 + 3 + 4 + 5 + 6 + 7 + 8 + 9 = (1 + 9) \times 9 / 2 = 45$

$4 + 5 = 9$

Nine is divisible by 3, therefore 45 is divisible by 3, therefore 123456789 is divisible by 3!

(That this method works is a theorem, though the proof is not given here.)

The beauty of the theorem lies in its recursive nature. A number is divisible by 3 if and only if the sum of its digits is divisible by 3. How do we know whether the sum of its digits is divisible by 3? Apply the theorem again! It's too true that "... to recurse, [is] divine".

Try a few more numbers yourself.

info -- Recursion

A prominent computer scientist once said "To iterate is human, to recurse, divine." But what does it mean *to recurse*? Before that, what is *to iterate*? "To iterate" simply means doing the same thing over and over again, and computers are very good at that. An example of iteration in mathematics is the exponential operation, e.g. x^n means doing x times x times x times x...n times. That is an example of iteration. *Thinking* about iteration *economically* (in terms of mental resources), by defining a problem in terms of itself, is "to recurse". To recursively represent x^n, we write:

$x^n = 1$ if n equals 0.

$x^n = x \times x^{n-1}$ if $n > 0$

What is 9^9? That is 9 times 9^8. But what is 9^8, it is 9 times 9^7. Repeating this way is an example of recursion.

Exercises

1. Factorise

1. 45
2. 4050
3. 2187

2. Show that the divisible-by-3 theorem works for any 3 digits numbers (Hint: Express a 3 digit number as 100a + 10b + c, where $0 \le a$, b and $c \le 9$)

3. "A number is divisible by 9 if and only if the sum of its digits is divisible by 9." True or false? Determine whether 89, 558, 51858, and 41857 are divisible by 9. Check your answers.

Finding primes

The prime sieve is a relatively efficient method of finding all primes less than or equal to a specified number. Let say we want to find all primes less than or equal to 50.

First we write out all numbers between 0 and 51 in a table as below

1	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50

Cross out 1, because it's not a prime.

X	2	3	4	5	6	7	8	9	10
11	12	13	14	15	16	17	18	19	20
21	22	23	24	25	26	27	28	29	30
31	32	33	34	35	36	37	38	39	40
41	42	43	44	45	46	47	48	49	50

Now 2 is the smallest number not crossed out in the table. We mark 2 as a prime and cross out all multiples of 2 i.e. 4, 6, 8, 10 ...

X	2_p	3	X	5	X	7	X	9	X
11	X	13	X	15	X	17	X	19	X
21	X	23	X	25	X	27	X	29	X
31	X	33	X	35	X	37	X	39	X
41	X	43	X	45	X	47	X	49	X

Now 3 is the smallest number not marked in anyway. We mark 3 as a prime and cross out all multiples of 3 i.e. 6, 9, 12, 15 ...

X	2_p	3_p	X	5	X	7	X	X	X
11	X	13	X	X	X	17	X	19	X
X	X	23	X	25	X	X	X	29	X
31	X	X	X	35	X	37	X	X	X
41	X	43	X	X	X	47	X	49	X

Continue is this way to find all the primes. When do you know you have found all the primes under 50?

Exercise

1.

X	2	3	X	5	X	7	X	X	X
11	X	13	X	X	X	17	X	19	X
X	X	23	X	X	X	X	X	29	X
31	X	X	X	X	X	37	X	X	X
41	X	43	X	X	X	47	X	X	X

The prime sieve has been applied to the table above. Notice that every number situated directly below 2 and 5 are crossed out. Construct a rectangular grid of numbers running from 1 to 60 so that after the prime sieve has been performed on it, all numbers situated directly below 3 and 5 are crossed out. What is the width of the grid?

2. Find all primes below 200

Infinitely many primes

We know some numbers can be factorised into primes. Some have only themselves as a prime factor, because they are prime. So how many primes are there? There are infinitely many! Here is a classical proof of the infinitude of primes dating back 2000 years to the ancient Greek mathematician Euclid:

Proof of infinitude of primes

Let us first assume that

there are a finite number of primes

therefore

there must be one prime that is greater than all others,

let this prime be referred to as *n*. We now proceed to show the two assumptions made above will lead to non-sense, and so there are infinitely many primes.

Take the product of all prime numbers to yield a number *x*. Thus:

$$x = 2 \times 3 \times 5 \times \ldots \times n$$

Then, let *y* equal one more than *x*:

$$y = x + 1$$

One may now conclude that *y* is not divisible by any of the primes up to *n*, since *y* differs from a multiple of each such prime by exactly 1. Since *y* is not divisible by any prime number, *y* must either be prime, or its prime factors must all be greater than *n*, a contradiction of the original assumption that *n* is the largest prime! Therefore, one must declare the original assumption incorrect, and that there exists an infinite number of primes.

Fun Fact -- Largest known prime

The largest prime known to human is $2^{30402457}-1$. It has a whopping 9152052 digits! Primes of the form 2^n-1 are called Mersenne primes named after the French monk/amateur mathematician.

Primes in arithmetic progression

Consider the arithmetic progression

a, *a* + *b*, *a* + 2*b*, *a* + 3*b* ....

if *a* and *b* share a common factor greater than 1, then *a* + *kb* for any k is not prime. But if *a* and *b* are coprimes, then there are infinitely many *k*'s such that *a* + *kb* is prime! For example, consider a = 3, b = 4,

3, 7, 11, 15, 19, 23, 27, 31...

in this rather short list, 3, 7, 11, 19, 23 and 31 are prime and they are all equal to 4*k* + 3 for some *k*. And there are infinitely many primes in this progression (see: proof by contradiction exercise Mathematical Proofs).

Modular Arithmetic

Introduction

Modular arithmetic connects with primes in an interesting way. Firstly recall clock-arithmetic. It is a system by which all numbers up to some positive integer, *n* say, are used. So if you were to start counting you would go 0, 1, 2, 3, ... , *n* - 1 but instead of counting *n* you would start over at 0. And what would have been *n* + 1 would be 1 and what would have been *n* + 2 would be 2. Once 2*n* had been reach the number is reset to 0 again, and so on. Very much like the clocks we have which starts at 1 and continues to 12 then back to 1 again.

The sequence also continues into what would be the negitive numbers. What would have been -1 is now *n* - 1. For example, let's start with modulo 7 arithmetic, it's just like ordinary arithmetic except the only numbers we use are 0, 1, 2, 3, 4, 5 and 6. If we see a number outside of this range we add 7 to (or subtract 7 from) it, until it lies within that range.

As mentioned above, modular arithmetic is not much different to ordinary arithmetic. For example, consider modulo 7 arithmetic

$$3 + 2 = 5$$

$$5 + 6 = 11 = 4$$

$$5 - 6 = -1 = 6$$

The same deal for multiplication

$$\begin{aligned} 3 \times 5 = 15 = 1 \\ 5 \times -6 = -30 = 5 \end{aligned}$$

We have done some calculation with negative numbers. Consider 5 × -6. Since -6 does not lie in the range 0 to 6, we need to add 7 to it until it does. And -6 + 7 = 1. So in modular 7 arithmetic, -6 = 1. In the above example we showed that 5 × -6 = -30 = 5, but 5 × 1 = 5. So we didn't do ourselves any harm by using -6 instead of 1. *Why?*

Note - Negatives: The preferred representation of -3 is 4, as -3 + 7 = 4, but using either -3 and 4 in a calculation will give us the same answer as long as we convert the final answer to a number between 0 and 6 (inclusive).

Exercise

Find in modulo 11

1.

-1 × -5

2.

3×7

3. Compute all the powers of 2

$2^1, 2^2, 2^3, \ldots, 2^{10}$

What do you notice? Using the powers of 2 find

$6^1, 6^2, 6^3, \ldots, 6^{10}$

What do you notice again?

4.

$$\sqrt{4}$$

i.e. find, by trial and error (or otherwise), all numbers x such that $x^2 = 4 \pmod{11}$. There are two solutions, find both .

5.

$$\sqrt{9}$$

i.e. find all numbers x such that $x^2 = 9 \pmod{11}$. There are two solutions, find both.

Inverses

Let's consider a number n, the inverse of n is the number that when multiplied by n will give 1. Let's consider a simple example, we want to solve the following equation in modulo 7,

$$5x = 3 \pmod{7}$$

the (mod 7) is used to make clear that we doing arithemetic modulo 7. We want to get rid of the 5 somehow, notice that

$$3 \times 5 = 15 = 1 \pmod{7}$$

because 3 multiplied by 5 gives 1, so we say 3 is the inverse of 5 in modulo 7. Now we multiply both sides by 3

$$\begin{aligned} 3 \times 5 \quad x &= 3 \times 3 \pmod{7} \\ x &= 9 \pmod{7} \\ &= 2 \pmod{7} \end{aligned}$$

So $x = 2$ modulo 7 is the required solution.

Inverse is unique

From above, we know the inverse of 5 is 3, but does 5 have another inverse? The answer is no. In fact, in any reasonable number system, a number can have one and only one inverse. We can see that from the following proof

Suppose *n* has two inverses *b* and *c*

$$b = b \times 1 = b(nc) = (bn)c = 1 \times c = c$$

From the above argument, all inverses of *n* must be equal. As a result, if the number *n* has an inverse, the inverse must be unique.

From now on, we will use x^{-1} to denote the inverse of *x* if it exists.

An interesting property of any modulo *n* arithmetic is that the number *n* - 1 has itself as an inverse. That is, $(n - 1) \times (n - 1) = 1 \pmod{n}$, or we can write $(n - 1)^2 = (-1)^2 = 1 \pmod{n}$. The proof is left as an exercise at the end of the section.

Existence of inverse

Not every number has an inverse in every modulo arithmetic. For example, 3 doesn't have an inverse mod 6, i.e., we can't find a number *x* such that $3x = 1$ mod 6 (the reader can easily check).

Let's consider modulo 15 arithmetic and note that 15 is composite. We know the inverse of 1 is 1 and of 14 is 14. But what about 3, 6, 9, 12, 5 and 10? None of them has an inverse! Note that each of them shares a common factor with 15!

Let's look at 3, we want to use a *proof by contradiction* argument to show that 3 does not have an inverse modulo 15. Suppose 3 has an inverse, which is denoted by *x*.

$$3x = 1 \pmod{15}$$

We make the *jump* from modular arithemetic into rational number arithmetic. If $3x = 1$ in modulo 15 arithmetic, then

$$3x = 15k + 1$$

for some integer *k*. Now we divide both sides by 3, we get

$$x = 5k + \frac{1}{3}$$

But this is can't be true, because we know that *x* is an integer, not a fraction. Therefore 3 doesn't have an inverse in mod 15 arithmetic. To show that 10 doesn't have an inverse is harder and is left as an exercise.

We will now state the theorem regarding the existence of inverses in modular arithmetic.

Theorem

If *n* is prime then every number (except 0) has an inverse in modulo *n* arithmetic.

Similarly

If *n* is composite then every number that doesn't share a common factor with *n* has an inverse.

It is interesting to note that *division* is closely related to the concept of inverses. Consider the following expresion

$$6 \times 3^{-1} \pmod 7$$

the conventional way to calculate the above would be to find the inverse of 3 (being 5). So

$$6 \times 3^{-1} = 6 \times 5 = 30 = 2 \pmod 7$$

Let's write the inverse of 3 as 1/3, so we think of multiplying 3^{-1} as *dividing* by 3, we get

$$6 \times \frac{1}{3} = \frac{6}{3} = 2 \pmod 7$$

Notice that we got the same answer! In fact, the division method will always work if the inverse exists.

Be very careful though, the expression in a different modulo system will produce the wrong answer, for example

$$6 \times 3^{-1} \pmod 9$$

we don't get 2, as 3^{-1} does not exist in modulo 9, so we can't use the division method.

Exercise

1. Does 8 have an inverse in mod 16 arithemetic? If not, why not?

2. Find *x* mod 7 if *x* exists:

$$x = 2^{-1}$$

$$x = 3^{-1}$$

$$x = 4^{-1}$$

$$x = 5^{-1}$$

$$x = 6^{-1}$$

$$x = 7^{-1}$$

3. Calculate *x* in two ways, *finding inverse* and *division*

$$x = 28 \cdot 7^{-1} \pmod{29}$$

4. (Trick) Find *x*

$$x = 5^{99} \times (40 + 3^{-1}) \pmod{11}$$

5. Find all inverses mod *n* ($n \leq 19$)
This exercise may seem tedious, but it will increase your understanding of the topic by tenfold

Coprime and greatest common divisor

Two numbers are said to be coprimes if their greatest common divisor is 1. The greatest common divisor (gcd) is just what its name says it is. And there is a quick and elegant way to compute the gcd of two numbers, called Euclid's algorithm. Let's illustrate with a few examples:

Example 1:

Find the gcd of 21 and 49.

We set up a 2-column table where the bigger of the two numbers is on the right hand side as follows

smaller	**larger**
21	49

We now compute 49(mod 21) which is 7 and put it in the second row *smaller* column, and put 21 into the *larger* column.

smaller	**larger**
21	49
7	21

Perform the same action on the second row to produce the third row.

smaller	**larger**
21	49
7	21
0	**7**

Whenever we see the number 0 appear on the *smaller* column, we know the corresponding *larger* number is the the gcd of the two numbers we started with, i.e. 7 is the gcd of 21 and 49. This *algorithm* is called Euclid's algorithm.

Example 2

Find the gcd of 31 and 101

smaller	**larger**
31	101
8	31
7	8
1	7

0	**1**

Example 3

Find the gcd of 132 and 200

smaller	**larger**
132	200
68	132
64	68
4	64
0	**4**

Remember

1. The gcd need not be a prime number.
2. The gcd of two primes is 1

Why does the Euclid's algorithm work? It is more fun for the questioner to discover.

Exercise

1. Determine whether the following sets of numbers are coprimes

1. 5050 5051
2. 59 78
3. 111 369
4. 2021 4032

2. Find the gcd of the numbers 15, 510 and 375

info -- Algorithm

An algorithm is a step-by-step description of a series of actions when performed correctly can accomplish a task. There are algorithms for finding primes, deciding whether 2 numbers are coprimes, finding inverses and many other purposes. You'll learn how to implement some of the algorithms we have seen using a computer in the chapter Mathematical Programming.

Finding Inverses

Let's look at the idea of inverse again, but from a different angle. In fact we will provide a sure-fire method to find the inverse of any number. Let's consider:

$5x = 1 \pmod 7$

We know x is the inverse of 5 and we can work out it's 3 pretty quickly. But x = 10 is also a solution, so is x = 17, 24, 31, ... 7n + 3. So there are infinitely many solutions; therefore we say 3 is equivalent to 10, 17, 24, 31 and so on. This is a crucial observation

Now let's consider

$$216x \equiv 1 \pmod{811}$$

A new notation is introduced here, it is the equal sign with three strokes instead of two. It is the "equivalent" sign; the above statement should read "216x is EQUIVALENT to 1" instead of "216x is EQUAL to 1". From now on, we will use the equivalent sign for modulo arithmetic and the equal sign for ordinary arithmetic.

Back to the example, we know that x exists, as gcd(811,216) = 1. The problem with the above question is that there is no quick way to decide the value of x! The best way we know is to multiply 216 by 1, 2, 3, 4... until we get the answer, there are at most 816 calculations, way too tedious for humans. But there is a better way, and we have touched on it quite a few times!

We notice that we could make the *jump* just like before into rational mathematics:

$$\begin{aligned} 216a &= 1 + 811b \\ 0 &\equiv 1 + 163b \pmod{216} \end{aligned}$$

We jump into rational maths again

$$\begin{aligned} 216c &= 1 + 163b \\ 53c &\equiv 1 \pmod{163} \end{aligned}$$

We jump once more

$$\begin{aligned} 53c &= 1 + 163d \\ 0 &\equiv 1 + 4d \pmod{53} \end{aligned}$$

Now the pattern is clear, we shall start from the beginning so that the process is not broken:

$$\begin{aligned}
216a &= 1 + 811b \\
216c &= 1 + 163b \\
53c &= 1 + 163d \\
53e &= 1 + 4d \\
e &= 1 + 4f
\end{aligned}$$

Now all we have to do is choose a value for *f* and substitute it back to find *a*! Remember *a* is the inverse of 216 mod 811. We choose f = 0, therefore e = 1, d = 13, c = 40, b = 53 and finally a = 199! If choose *f* to be 1 we will get a different value for *a*.

The very perceptive reader should have noticed that this is just Euclid's gcd algorithm in reverse.

Here are a few more examples of this ingenious method in action:

Example 1

Find the smallest positive value of *a*:

$$\begin{aligned}
33a &\equiv 1 \pmod{101} \\
33a &= 1 + 101b \\
33c &= 1 + 2b \\
c &= 1 + 2d
\end{aligned}$$

Choose d = 0, therefore a = 49.

Example 2 Find the smallest positive value of *a*:

$$\begin{aligned}
27a &\equiv 1 \pmod{821} \\
27a &= 1 + 821b \\
27c &= 1 + 11b \\
5c &= 1 + 11d \\
5e &= 1 + d
\end{aligned}$$

Choose e = 0, therefore a = -152 = 669

Example 3 Find the smallest positive value of *a*:

$$\begin{aligned}
34a &\equiv 1 \pmod{55} \\
34a &= 1 + 55b \\
34c &= 1 + 21b \\
13c &= 1 + 21d \\
13e &= 1 + 8d \\
5e &= 1 + 8f \\
5g &= 1 + 3f \\
2g &= 1 + 3h \\
2i &= 1 + h
\end{aligned}$$

Set i = 0, then a = -21 = 34. *Why is this so slow for two numbers that are so small? What can you say about the coefficients?*

Example 4 Find the smallest positive value of *a*:

$$\begin{aligned} 21a &\equiv 1 \pmod{102} \\ 21a &= 1 + 102b \\ 21c &= 1 + 18b \\ 3c &= 1 + 18d \\ 3d &= 1 \end{aligned}$$

Now *d* is not an integer, therefore 21 does not have an inverse mod 102.

What we have discussed so far is the method of finding integer solutions to equations of the form:

ax + by = 1

where *x* and *y* are the unknowns and *a* and *b* are two given constants, these equations are called linear *Diophantine equations*. It is interesting to note that sometimes there is no solution, but if a solution exists, it implies that infinitely many solutions exist.

Diophantine equation

In the Modular Arithmetic section, we stated a theorem that says if gcd(*a*,*m*) = 1 then a^{-1} (the inverse of *a*) exists in mod *m*. It is not difficult to see that if *p* is prime then gcd(*b*,*p*) = 1 for all *b* less than *p*, therefore we can say that in mod *p*, every number except 0 has an inverse.

We also showed a way to find the inverse of any element mod *p*. In fact, finding the inverse of a number in modular arithmetic amounts to solving a type of equations called Diophantine equations. A Diophantine equation is an equation of the form

ax + *by* = *d*

where *x* and *y* are unknown.

As an example, we should try to find the inverse of 216 in mod 811. Let the inverse of 216 be *x*, we can write

$$216x \equiv 1 \pmod{811}$$

we can rewrite the above in every day arithmetic

$$216x + 811y = 1$$

which is in the form of a Diophantine equation.

Now we are going to do the inelegant method of solving the above problem, and then the elegant method (using Magic Tables).

Both methods mentioned above uses the Euclid's algorithm for finding the gcd of two numbers. In fact, the gcd is closely related to the idea of an inverse. Let's apply the Euclid's algorithm on the two numbers 216 and 811. This time, however, we should store more details; more specifically, we want to set up an additional column called PQ which stands for partial quotient. The partial quotient is just a technical term for "how many *n* goes into *m*" e.g. The partial quotient of 3 and 19 is 6, the partial quotient of 4 and 21 is 5 and one last example the partial quotient of 7 and 49 is 7.

smaller	**larger**	**PQ**
216	811	**3**
163		

The tables says three 216s goes into 811 with remainder 163, or symbollically:

811 = 3×216 + 163.

Let's continue:

smaller	**larger**	**PQ**
216	811	**3**
163	216	**1**
53	163	**3**
4	53	**13**
1	4	**4**
0	1	

Reading off the table, we can form the following expressions

811 = 3× 216 + 163

216 = 1× 163 + 53

163 = 3× 53 + 4

53 =13× 4 + 1

Now that we can work out the inverse of 216 by working the results backwards

1 = 53 - 13×4

1 = 53 - 13×(163 - 3×53)

$1 = 40 \times 53 - 13 \times 163$

$1 = 40 \times (216 - 163) - 13 \times 163$

$1 = 40 \times 216 - 53 \times 163$

$1 = 40 \times 216 - 53 \times (811 - 3 \times 216)$

$1 = 199 \times 216 - 53 \times 811$

Now look at the equation mod 811, we will see the inverse of 216 is 199.

Magic Table

The Magic Table is a more elegant way to do the above calculations, let us use the table we form from Euclid's algorithm

smaller	**larger**	**PQ**
216	811	**3**
163	216	**1**
53	163	**3**
4	53	**13**
1	4	**4**
0	1	

Now we set up the so-called "magic table" which looks like this initially

0	1
1	0

Now we write the partial quotient on the first row:

		3	1	3	1 3	4
0	1					
1	0					

We produce the table according to the following rule:

Multiply a partial quotient one space to the left of it in a different row, add the product to the number two space to the left on the same row and put the sum in the corresponding row.

It sounds more complicated then it should. Let's illustrate by producing a column:

		3	1	3	1 3	4
0	1	3				
1	0	1				

We put a 3 in the second row because 3 = 3×1 + 0. We put a 1 in the third row because 1 = 3×0 + 1.

We shall now produce the whole table without disruption:

		3	1	3	13	4
0	1	3	4	1 5	19 9	81 1
1	0	1	1	4	53	21 6

I claim

|199×216 - 811×53| = 1

In fact, if you have done the magic table properly and *cross multiplied and subtracted* the last two column correctly, then you will always get 1 or -1, provided the two numbers you started with are coprimes. The magic table is just a cleaner way of doing the mathematics.

Exercises

1. Find the smallest positive *x*:

$$216x \equiv 1 \pmod{816}$$

2. Find the smallest positive *x*:

$$42x \equiv 7 \pmod{217}$$

3.

(a) Produce the magic table for 33*a* = 1 (mod 101)

(b) Evaluate and express in the form p/q

$$3 + \frac{1}{16 + \frac{1}{2}}$$

What do yo notice?

4.

(a) Produce the magic table for 17a = 1 (mod 317)

(b) Evaluate and express in the form p/q

$$18 + \frac{1}{1 + \frac{1}{1 + \frac{1}{1 + \frac{1}{5}}}}$$

What do yo notice?

Chinese remainder theorem

The Chinese remainder theorem is known in China as *Han Xing Dian Bing*, which in its most naive translation means *Han Xing counts his soldiers*. The original problem goes like this:

There exists a number *x*, when divided by 3 leaves remainder 2, when divided by 5 leaves remainder 3 and when divided by 7 leaves remaider 2. Find the smallest *x*.

We translate the question into symbolic form:

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &\equiv 3 \pmod 5 \\ x &\equiv 2 \pmod 7 \end{aligned}$$

How do we go about finding such a *x*? We shall use a familar method and it is best illustrated by example:

Looking at x = 2 (mod 3), we make the *jump* into ordinary mathematics

$$\begin{aligned} x &\equiv 2 \pmod 3 \\ x &= 2 + 3a \qquad (1) \end{aligned}$$

Now we look at the equation modulo 5

$$\begin{aligned} 2 + 3a &\equiv 3 \pmod 5 \\ 3a &\equiv 1 \pmod 5 \\ a &\equiv 2 \pmod 5 \\ a &= 2 + 5b \end{aligned}$$

Substitute into (1) to get the following

$$\begin{aligned} x &= 2 + 3(2 + 5b) \\ &= 8 + 15b \end{aligned}$$

Now look at the above modulo 7

$$x = 8 + 15b \equiv 2 \pmod 7$$

we get

$$b \equiv 1 \pmod 7$$

We choose $b = 1$ to minimise x, therefore $x = 23$. And a simple check (to be performed by the reader) should confirm that $x = 23$ is a solution. A good question to ask is what is the next smallest x that satisfies the three congruences? The answer is $x = 128$, and the next is 233 and the next is 338, and they differ by 105, the product of 3, 5 and 7.

We will illustrate the method of solving a system of congruencies further by the following examples:

Example 1 Find the smallest x that satifies:

$$x \equiv 1 \pmod 3$$

$$x \equiv 2 \pmod 5$$

$$x \equiv 3 \pmod 7$$

Solution

$$x = \quad 1 + 3 \qquad a \equiv 2 \pmod 5$$
$$a = 2 + 5b$$

now substitute back into the first equation, we get

$$\begin{aligned} x &= 1 + 3(2 + 5b) \\ &= 7 + 15b \\ &\equiv 3 \pmod 7 \end{aligned}$$

we obtain

$$b \equiv 3 \pmod 7$$
$$b = 3 + 7c$$

again substituting back

$$\begin{aligned} x &= 7 + 15(3 + 7c) \\ &= 52 + 15 \times 7c \end{aligned}$$

Therefore 52 is the smallest *x* that satifies the congruencies.

Example 2

Find the smallest *x* that satisfies:

$$x \equiv 5 \pmod{11}$$
$$x \equiv 3 \pmod 7$$
$$x \equiv 8 \pmod 9$$

Solution

$$x \quad = 5 + 11 \quad a \equiv 3 \pmod 7$$
$$a \equiv 3 \pmod 7$$
$$a = 3 + 7b$$

substituting back

$$\begin{aligned} x = & \quad 5 + 11(3 + 7b) \\ = & \quad 38 + 11 \times 7b \\ \equiv & \quad 8 \pmod 9 \end{aligned}$$

now solve for *b*

$$\begin{aligned} 2 + 2 \times 7b & \equiv 8 \pmod 9 \\ b & \equiv 3 \pmod 9 \\ b & = 3 + 9c \end{aligned}$$

again, substitue back

$$\begin{aligned} x & = 38 + 11 \times 7(3 + 9c) \\ & = 269 + 11 \times 7 \times 9c \end{aligned}$$

Therefore 269 is the smallest *x* that satifies the Congruences.

Excercises

1. Solve for *x*

$$\begin{aligned} 3x & \equiv 5 \pmod{14} \\ 2x & \equiv -3 \pmod{17} \\ x & \equiv 6 \pmod{15} \end{aligned}$$

2. Solve for *x*

$$\begin{aligned} 3x & \equiv 5 \pmod{19} \\ 7x & \equiv -3 \pmod{17} \\ x & \equiv 6 \pmod{11} \end{aligned}$$

Existence of a solution

The exercises above all have a solution. So does there exist a system of congruences such that no solution could be found? It certainly is possible, consider:

$x \equiv 5 \pmod{15}$

$x \equiv 10 \pmod{21}$

a cheekier example is:

$x \equiv 1 \pmod 2$

$x \equiv 0 \pmod 2$

but we won't consider silly examples like that.

Back to the first example, we can try to solve it by doing:

$$\begin{aligned} x = 5 + 15k &\equiv 10 \pmod{21} \\ 15k &\equiv 5 \\ 3k &\equiv 1 \end{aligned}$$

the above equation has no solution because 3 does not have an inverse modulo 21!

One may be quick to conclude that if two modulo systems share a common factor then there is no solution. But this is not true! Consider:

$$x \equiv 4 \pmod{15}$$

$$x \equiv 7 \pmod{21}$$

we can find a solution

$$\begin{aligned} x = 4 + 15k &\equiv 7 \pmod{21} \\ 15k &\equiv 3 \pmod{21} \\ 5 \times 3k &\equiv 3 \pmod{21} \end{aligned}$$

we now multiply both sides by the inverse of 5 (which is 17), we obtain

$$3k \equiv 9$$

obviously, $k = 3$ is a solution, and the two modulo systems are the same as the first example (i.e. 15 and 21).

So what determines whether a system of congruences has a solution or not? Let's consider the general case:

$$x \equiv a \pmod m$$

$$x \equiv b \pmod n$$

we have

$$x = a + km$$

$$x = b + ln$$

essentially, the problem asks us to find *k* and *l* such that the above equations are satisfied.

We can approach the problem as follows

$$0 = (a - b) + (km - ln)$$

$$(ln - km) = (a - b)$$

now suppose m and n have gcd(m,n) = d, and m = dm_o, n = dn_o. We have

$$dln_o - dkm_o = (a - b)$$

$$ln_o - km_o = (a - b)/d$$

if $(a - b)/d$ is an integer then we can read the equation mod m_o, we have:

$$ln_o \equiv (a - b)/d \pmod{m_o}$$

Again, the above only makes sense if $(a - b)/d$ is integeral. Also if $(a - b)/d$ is an integer, then there is a solution, as m_o and n_o are coprimes!

In summary: for a system of two congruent equations

$$x \equiv a \pmod{m}$$

$$x \equiv b \pmod{n}$$

there is a solution if and only if

d = gcd(m,n) divides $(a - b)$

And the above generalises well into more than 2 congruences. For a system of *n* congruences:

$$x \equiv a_1 \pmod{m_1}$$

$$x \equiv a_2 \pmod{m_2}$$

...

$$x \equiv a_n \pmod{m_n}$$

for a solution to exist, we require that if $i \neq j$

gcd(m_i,m_j) divides $(a_i - a_j)$

Exercises

Decide whether a solution exists for each of the congruencies. Explain why.

1.

$x \equiv 7 \pmod{25}$

$x \equiv 22 \pmod{45}$

2.

$x \equiv 7 \pmod{23}$

$x \equiv 3 \pmod{11}$

$x \equiv 3 \pmod{13}$

3.

$x \equiv 7 \pmod{25}$

$x \equiv 22 \pmod{45}$

$x \equiv 7 \pmod{11}$

4.

$x \equiv 4 \pmod{28}$

$x \equiv 28 \pmod{52}$

$x \equiv 24 \pmod{32}$

To go further

This chapter has been a gentle introduction to number theory, a profoundly beautiful branch of mathematics. It is gentle in the sense that it is mathematically light and overall quite easy. If you enjoyed the material in this chapter, you would also enjoy Further Modular Arithmetic, which is a harder and more rigorous treatment of the subject.

Also, if you feel like a challenge you may like to try out the Probelem Set we have prepared for you. On the other hand, the project asks you to take a more investigative approach to work through some of the finer implications of the Chinese Remainder Theorem.

Acknowledgement

Acknowledgement: This chapter of the textbook owes much of its inspiration to Terry Gagen, Emeritus Associate Professor of Mathematics at the University of Sydney, and his lecture notes on "Number Theory and Algebra". Terry is a much loved figure among his students and is

renowned for his entertaining style of teaching.

Reference

1. Largest Known Primes--A Summary

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Problem Set

1. Is there a rule to determine whether a 3-digit number is divisible by 11? If yes, derive that rule.

2. Show that p, $p + 2$ and $p + 4$ cannot all be primes. (p a positive integer)

3. Find x

$$x \equiv 3^7 + 1^7 + 2^7 + 4^7 + 5^7 + 6^7 + 7^7 \pmod 7$$

4. Show that there are no integers x and y such that

$$x^2 - 5y^2 = 3$$

5. In modular arithmetic, if

$$x^2 \equiv y \pmod m$$

for some m, then we can write

$$x \equiv \sqrt{y} \pmod m$$

we say, x is the square root of y mod m.

Note that if x satisfies $x^2 \equiv y$, then $m - x \equiv -x$ when squared is also equivalent to y. We consider both x and $-x$ to be square roots of y.

Let p be a prime number. Show that

(a)

$$(p-1)! \equiv -1 \ (\text{mod p})$$

where

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$$

E.g. 3! = 1*2*3 = 6

(b)

Hence, show that

$$\sqrt{-1} \equiv \frac{p-1}{2}! \pmod{p}$$

for $p \equiv 1 \pmod{4}$, i.e., show that the above when squared gives one.

Square root of minus 1

Project -- The Square Root of -1

Notation: In modular arithmetic, if

$$x^2 \equiv y \pmod{m}$$

for some m, then we can write

$$x \equiv \sqrt{y} \pmod{m}$$

we say, x is the square root of y mod m.

Note that if x satisfies $x^2 \equiv y$, then $m - x \equiv -x$ when squared is also equivalent to y. We consider both x and $-x$ to be square roots of y.

1. Question 5 of the Problem Set showed that

$$x \equiv \sqrt{-1} \equiv \sqrt{p-1} \pmod{p}$$

exists for $p \equiv 1$ (mod 4) prime. Explain why no square root of -1 exist if $p \equiv 3$ (mod 4) prime.

2. Show that for $p \equiv 1$ (mod 4) prime, there are exactly 2 solutions to

$$x \equiv \sqrt{-1} \pmod{p}$$

3. Suppose m and n are integers with gcd(n,m) = 1. Show that for each of the numbers 0, 1, 2, 3, ... , $nm - 1$ there is a unique pair of numbers a and b such that the smallest number x that satisfies:

$x \equiv$ a (mod m)

$x \equiv$ b (mod n)

is that number. E.g. Suppose m = 2, n = 3, then 4 is uniquely represented by

$x \equiv$ 0 (mod 2)

$x \equiv$ 1 (mod 3)

as the smallest x that satisfies the above two congruencies is *4*. In this case the unique pair of numbers are 0 and 1.

4. If $p \equiv 1$ (mod 4) prime and $q \equiv 3$ (mod 4) prime. Does

$$x \equiv \sqrt{-1} \pmod{pq}$$

have a solution? Why?

5. If $p \equiv 1 \pmod 4$ prime and $q \equiv 1 \pmod 4$ prime and $p \neq q$. Show that

$$x \equiv \sqrt{-1} \pmod{pq}$$

has 4 solutions.

6. Find the 4 solutions to

$$x \equiv \sqrt{-1} \pmod{493}$$

note that $493 = 17 \times 29$.

7. Take an integer n with more than 2 prime factors. Consider:

$$x \equiv \sqrt{-1} \pmod{n}$$

Under what condition is there a solution? Explain thoroughly.

Solutions to exercises

HSE Primes|Primes and Modular Arithmetic

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

Factorisation Exercises

Factorise the following numbers. (note: I know you didn't have to, this is just for those who are curious)

1. 13 is prime
2. $26 = 13 \cdot 2$
3. 59 is prime
4. $82 = 41 \cdot 2$
5. 101 is prime
6. $121 = 11 \cdot 11$
7. $2187 = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$

Recursive Factorisation Exercises

Factorise using recursion.

1. $45 = 3 \cdot 3 \cdot 5$
2. $4050 = 2 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 5 \cdot 5$
3. $2187 = 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3 \cdot 3$

Prime Sieve Exercises

1. Use the above result to quickly work out the numbers that still need to be crossed out in the table below, knowing 5 is the next prime:

X	2_p	3_p	X	5	X	7	X	X	X
11	X	13	X	X	X	17	X	19	X
X	X	23	X	25	X	X	X	29	X
31	X	X	X	35	X	37	X	X	X
41	X	43	X	X	X	47	X	49	X

The next prime number is 5. Because 5 is an unmarked prime number, and 5 * 5 = 25, cross out 25. Also, 7 is an unmarked prime number, and 5 * 7 = 35, so cross off 35. However, 5 * 11 = 55, which is too high, so mark 5 as prime ad move on to 7. The only number low enough to be marked off is 7 * 7, which equals 35. You can go no higher.

2. Find all primes below 200.

The method will not be outlined here, as it is too long. However, all primes below 200 are:

2 3 5 7 11 13 17 19 23 29 31 37 41 43 47 53 59 61 67 71 73 79 83 89 97 101 103 107 109 113 127 131 137 139 149 151 157 163 167 173 179 181 191 193 197 199

Modular Arithmetic Exercises

1. $(-1)\cdot(-5) \mod 11 = 5$ alternatively, -1 = 10, -5 = 6: 10 × 6 = 60 = 5&times 11 + 5 = 5

2. $3 \cdot 7 \mod 11 = 21 = 10$

3. $2^1 = 2, 2^2 = 4, 2^3 = 8, 2^4 = 16 = 5$
$2^5 = 32 = 10, 2^6 = 64 = 9, 2^7 = 128 = 7$
$2^8 = 256 = 3, 2^9 = 512 = 6, 2^{10} = 1024 = 1$
An easier list: 2, 4, 8, 5, 10, 9, 7, 3, 6, 1
Notice that it is not necessary to acutally
compute 2^{10} to find 2^{10} mod 11.
If you know 2^9 mod 11 = 6.
You can find 2^{10} mod 11 = (2*(2^9 mod 11)) mod 11 = 2*6 mod 11 = 12 mod 11 = 1.
We can note that $2^9 = 6$ and $2^{10} = 1$, we can calculate 6^2 easily: $6^2 = 2^{18}$ = 2^8 = 3. OR by the above method
$6^1 = 6, 6^2 = 36 = 3, 6^3 = 6 * 3 = 18 = 7$,
$6^4 = 6 * 7 = 42 = 9, 6^5 = 6 * 9 = 54 = 10, 6^6 = 6 * 10 = 60 = 5$,
$6^7 = 6 * 5 = 30 = 8, 6^8 = 6 * 8 = 48 = 4, 6^9 = 6 * 4 = 24 = 2, 6^{10} = 6 * 2 = 12 = 1$.
An easier list: 6, 3, 7, 9, 10, 5, 8, 4, 2, 1.

4. $0^2 = 0$, $1^2 = 1$, $2^2 = 4$, $3^2 = 9$,
$4^2 = 16 = 5$, $5^2 = 25 = 5$, $6^2 = 36 = 3$, $7^2 = 49 = 3$,
$8^2 = 64 = 5$, $9^2 = 81 = 4$, $10^2 = 100 = 1$
An easier list: 0, 1, 4, 9, 5, 3, 3, 5, 9, 4, 1
Thus $\sqrt{4} = 2 \text{ and } \sqrt{4} = 9$

5. $x^2 = -2 = 9$
Just look at the list above and you'll see that $\sqrt{-2} = 8 \text{ and } \sqrt{-2} = 3$

Division and Inverses Exercises

1.

$x = 2^{-1} = 4$

$x = 3^{-1} = 5$

$x = 4^{-1} = 2$

$x = 5^{-1} = 3$

$x = 6^{-1} = 6$

$x = 7^{-1} = 0^{-1}$ therefore the inverse does not exist

2. $x = \frac{28}{7} = 4 \pmod{29}$

$7^{-1} = 25 \pmod{29}$

$x = 28 \cdot 25 = 4 \pmod{29}$

3.

$x = 5^{99} \times (40 + \frac{1}{3}) \pmod{11}$

$x = 5^{99} \times (40 + 4) \pmod{11}$

$x = 5^{99} \times 0 \pmod{11}$

$x = 0 \pmod{11}$

4.

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	
	1																		mod 2
	1	2																	mod 3
	1		3																mod 4
	1	3	2	4															mod 5
	1				5														mod 6
	1	4	5	2	3	6													mod 7
	1		3		5		7												mod 8
	1	5		7	2		4	8											mod 9
	1		7				3		9										mod 10
	1	6	4	3	9	2	8	7	5	10									mod 11
	1				5		7				11								mod 12
	1	7	9	10	8	11	2	5	3	4	6	12							mod 13
	1		5		3				11		9		13						mod 14
	1	8		4			13	2			11		7	14					mod 15
	1		11		13		7		9		3		5		15				mod 16
	1	9	6	13	7	3	5	15	2	12	14	10	4	11	8	16			mod 17
	1				11		13				5		7				17		mod 18
	1	10	13	5	4	16	11	12	17	2	7	8	3	15	14	6	9	18	mod 19

Coprime and greatest common divisor Exercises

1.

1.

smaller	**larger**
5050	5051
1	5050
0	**1**

5050 and 5051 are coprime

2.

smaller	**larger**
59	78
19	59
2	19
1	2
0	**1**

59 and 79 are coprime

3.

smaller	larger
111	369
36	111
3	36
0	**3**

111 and 369 are not coprime

4.

smaller	larger
2021	4032
2011	2021
10	2011
1	10
0	**1**

2021 and 4032 are coprime

2.We first calculate the gcd for all combinations

smaller	larger
15	510
0	**15**

smaller	larger
15	375
0	**15**

smaller	larger
375	510
135	375
105	135
30	105
15	30
0	**15**

The gcd for any combination of the numbers is 15 so the gcd is 15 for the three numbers.

Diophantine equation Exercises

1.

$$\begin{aligned} 216x &= 1 + 816b \\ 216c &= 1 + 168b \\ 48c &= 1 + 168d \\ 48e &= 1 + 24d \\ 24e &= 1 + 24f \end{aligned}$$

There is no solution, because can never become an integer.

2.

$$\begin{aligned} 42x &= 7 + 217b \\ 42c &= 7 + 7b \\ 7c &= 0 + 7d \end{aligned}$$

We choose d=1, then x=26.

3.

(a)

smaller	**larger**	**PQ**
33	101	**3**
2	33	**16**
1	2	**2**
0	1	

		3	16	2
0	1	3	49	101
1	0	1	16	33

(b) *To be added*

4.

(a)

smaller	larger	PQ
17	317	18
11	17	1
6	11	1
5	6	1
1	5	5
0	1	

		18	1	1	1	5
0	1	18	19	37	56	317
1	0	1	1	2	3	17

(b) To be added

Chinese remainder theorem exercises

1.

$$\begin{array}{rcccl} 3x & \equiv & 5 \pmod{14} & & \\ x & \equiv & 11 \pmod{14} & & \\ x & = & 11 + 14a & & \\ 2x & = & 2(11 + 14a) & \equiv & -3 \pmod{17} \\ & & 22 + 28a & \equiv & -3 \pmod{17} \\ & & 11a & \equiv & -8 \pmod{17} \\ & & a & = & 7 + 17b \\ x & = & 11 + 14(7 + 17b) & \equiv & 6 \pmod{15} \\ & = & 109 + 238b & \equiv & 6 \pmod{15} \\ & = & 4 + 13b & \equiv & 6 \pmod{15} \\ & = & 13b & \equiv & 2 \pmod{15} \\ & & b & \equiv & 14 \pmod{15} \\ & & b & = & 14 + 15c \\ x & = & 109 + 238(14 + 15c) & & \\ x & = & 3441 + 3570c & & \end{array}$$

Question 1

Show that the divisible-by-3 theorem works for any 3 digits numbers (Hint: Express a 3 digit number as 100a + 10b + c, where a, b and c are ≥ 0 and < 10)

Solution 1 Any 3 digits integer *x* can be expressed as follows

x = 100a + 10b + c

where a, b and c are positive integer between 0 and 9 inclusive. Now

$$x \equiv 100a + 10b + c \equiv a + b + c \pmod{3}$$

$$x \equiv 0 \pmod{3}$$

if and only if a + b + c = 3k for some k. But a, b and c are the digits of x.

Question 2

"A number is divisible by 9 if and only if the sum of its digits is divisible by 9." True or false? Determine whether 89, 558, 51858, and 41857 are divisible by 9. Check your answers.

Solution 2 The statement is true and can be proven as in question 1.

Question 4

The prime sieve has been applied to the table of numbers above. Notice that every number situated directly below 2 and 5 are crossed out. Construct a rectangular grid of numbers running from 1 to 60 so that after the prime sieve has been performed on it, all numbers situated directly below 3 and 5 are crossed out. What is the width of the grid?

Solution 4 The width of the grid should be 15 or a multiple of it.

Question 6

Show that n - 1 has itself as an inverse modulo n.

Solution 6

$(n - 1)^2 = n^2 - 2n + 1 = 1 \pmod n$

Alternatively

$(n - 1)^2 = (-1)^2 = 1 \pmod n$

Question 7

Show that 10 does not have an inverse modulo 15.

Solution 7 Suppose 10 does have an inverse *x* mod 15,

10x = 1 (mod 15)

2Ã-5x = 1 (mod 15)

5x = 8 (mod 15)

5x = 8 + 15k

for some integer k

x = 1.6 + 3k

but now *x* is not an integer, therefore 10 does not have an inverse

Problem set solutions

Question 1

Is there a rule to determine whether a 3-digit number is divisible by 11? If yes, derive that rule.

Solution

Let *x* be a 3-digit number We have

$$x = 100a + 10b + c$$

now

$$x \equiv a + 10b + c \equiv a - b + c \pmod{11}$$

We can conclude a 3-digit number is divisible by 11 if and only if the sum of first and last digit minus the second is divisible by 11.

Question 2

Show that p, $p + 2$ and $p + 4$ cannot all be primes. (p a positive integer)

Solution

We look at the arithmetic mod 3, then p slotted into one of three categories

1st category

$$p \equiv 0 \pmod{3}$$

we deduce p is not prime, as it's a multiple of 3

2nd category

$$p \equiv 1 \pmod{3}$$

$$p + 2 \equiv 0 \pmod{3}$$

so p + 2 is not prime

3rd category

$$p \equiv 2 \pmod{3}$$

$$p + 4 \equiv 0 \pmod{3}$$

therefore $p + 4$ is not prime

Therefore *p*, *p* + 2 and *p* + 4 cannot all be primes.

Question 3

Find *x*

$$x \equiv 1^7 + 2^7 + 3^7 + 4^7 + 5^7 + 6^7 + 7^7 \pmod 7$$

Solution

Notice that

$$-a \equiv 7 - a \pmod 7.$$

Then

$$1^7 \equiv (7-6)^7 \equiv (-6)^7 \equiv -(6^7) \pmod 7.$$

Likewise,

$$2^7 \equiv -5^7 \pmod 7$$

and

$$3^7 \equiv -4^7 \pmod 7.$$

Then

$$\begin{aligned} x &\equiv 1^7 + 2^7 + 3^7 + 4^7 + 5^7 + 6^7 + 7^7 \\ &\equiv 1^7 + 2^7 + 3^7 - 3^7 - 2^7 - 1^7 + 7^7 \\ &\equiv 0 \pmod 7 \end{aligned}$$

Question 4

9. Show that there are no integers *x* and *y* such that

$$x^2 - 5y^2 = 3$$

Solution

Look at the equation mod 5, we have

$$x^2 = 3 \pmod 5$$

but

$1^2 \equiv 1$
$2^2 \equiv 4$
$3^2 \equiv 4$
$4^2 \equiv 1$

therefore there does not exist *a* x such that

$$x^2 \equiv 3 \pmod 5$$

Question 5

Let *p* be a prime number. Show that

(a)

$$(p-1)! \equiv -1 \pmod p$$

where

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$$

E.g. 3! = 1Ã-2Ã-3 = 6

(b) Hence, show that

$$\sqrt{-1} \equiv \frac{p-1}{2}! \pmod p$$

for *p* ≡1 (mod 4)

Solution

a) If *p* = 2, then it's obvious. So we suppose *p* is an odd prime. Since *p* is prime, some deep thought will reveal that every distinct element multiplied by some other element will give 1. Since

$$(p-1)! = (p-1)(p-2)(p-3)\cdots 2$$

we can pair up the inverses (two numbers that multiply to give one), and (p - 1) has itself as an inverse, therefore it's the only element not "eliminated"

$$\boxed{(p-1)! \equiv (p-1) \equiv -1}$$

as required.

b) From part a)

$$-1 \equiv (p-1)!$$

since *p* = 4*k* + 1 for some positive integer *k*, (p - 1)! has 4*k* terms

$$-1 = 1 \times 2 \times 3 \times \cdots 2k \times (-2k) \cdots \times (-3) \times (-2) \times (-1)$$

there are an even number of minuses on the right hand side, so

$$-1 = (1 \times 2 \times 3 \times \cdots 2k)^2$$

it follows

$$\sqrt{-1} = 1 \times 2 \times 3 \times ...2k$$

and finally we note that p = 4k + 1, we can conclude

$$\sqrt{-1} = \frac{p-1}{2}!$$

Definitions

Logic

Introduction

Logic is the study of the way we humans reason. In this chapter, we focus on the *methods* of logical reasoning, i.e. digital logic, predicate calculus, application to proofs and the (insanely) fun logical puzzles.

Boolean algebra

In the black and white world of ideals, there is absolute truth. That is to say *everything* is either **true** or **false**. With this philosophical backdrop, we consider the following examples:

"One plus one equals two." True or false?

That is (without a doubt) true!

"1 + 1 = 2 AND 2 + 2 = 4." True or false?

That is also true.

But what about:

"1 + 1 = 3 OR Sydney is in Australia" True or false?

It is true! Although 1 + 1 = 3 is not true, the OR in the statement made it so that if either part of the statement is true then the whole statement is true.

Now let's consider a more puzzling example

"2 + 2 = 4 OR 1 + 1 = 3 AND 1 - 3 = -1" True or false?

The truth or falsity of the statements depends on the *order* in which you evaluate the statement. If you evaluate "2 + 2 = 4 OR 1 + 1 = 3" first, the statement is false, and otherwise true. As in ordinary algebra, it is necessary that we define some rules to govern the order of evaluation, so we don't have to deal with ambiguity.

Before we decide which order to evaluate the statements in, we do what most mathematician love to do -- replace sentences with symbols.
Let x represent the truth or falsity of the statement 2 + 2 = 4.
Let y represent the truth or falsity of the statement 1 + 1 = 3.
Let z represent the truth or falsity of the statement 1 - 3 = -1.

Then the above example can be rewritten in a more compact way:

x OR y AND z

To go one step further, mathematicians also replace OR by + and AND by Ã-, the statement becomes:

$$x + y \times z$$

Now that the order of precedence is clear. We evaluate (y AND z) first and then OR it with x. The statement "x + yz" is true, or symbolically

x + yz = 1

where the number 1 represents "true".

There is a good reason why we choose the multiplicative sign for the AND operation. As we shall see later, we can draw some parallels between the AND operation and multiplication.

The Boolean algebra we are about to investigate is named after the British mathematician George Boole. Boolean algebra is about two things -- "true" or "false" which are often represented by the numbers 1 and 0 respectively. Alternative, T and F are also used.

Boolean algebra has operations (AND and OR) analogous to the ordinary algebra that we know and love.

Basic Truth tables

We have all had to memorise the 9 by 9 multiplication table and now we know it off by heart. In Boolean algebra, the idea of a truth table is somewhat similar.

Let's consider the AND operation which is analogous to the multiplication. We want to consider:

x AND y

where and *x* and *y* each represent a true or false statement (e.g. It is raining today). It is true if and only if both *x* and *y* are true, in table form:

The AND function		
x	**y**	**x AND y**
F	F	F
F	T	F
T	F	F
T	T	T

We shall use 1 instead of T and 0 instead of F from now on.

The AND function		
x	**y**	**x AND y**
0	0	0
0	1	0
1	0	0
1	1	1

Now you should be able to see why we say AND is analogous to multiplication, we shall replace the AND by Ã-, so **x AND y** becomes **xÃ-y** (or just *xy*). From the AND truth table, we have:

0 Ã- 0 = 0

0 Ã- 1 = 0

1 Ã- 0 = 0

1 Ã- 1 = 1

To the OR operation. ***x* OR *y*** is FALSE if and only if both *x* and *y* are false. In table form:

The OR function		
x	**y**	**x OR y**
0	0	0
0	1	1
1	0	1
1	1	1

We say OR is almost analoguous to addition. We shall illustrate this by replacing OR with +:

0 + 0 = 0

0 + 1 = 1

1 + 0 = 1

1 + 1 = 1 (like 1 OR 1 is 1)

The NOT operation is not a *binary operation*, like AND and OR, but a *unary operation*, meaning it works with one argument. **NOT *x*** is true if *x* is false and *false* if *x* is true. In table form:

The NOT function		
x	**NOT *x***	
0	1	
1	0	

In symbolic form, **NOT** x is denoted x' or ~x (or by a bar over the top of x).

Alternative notations:

$$x \times y = x \wedge y$$

and

$$x + y = x \vee y$$

Compound truth tables

The three truth tables presented above are the most basic of truth tables and they serve as the building blocks for more complex ones. Suppose we want to construct a truth table for xy + z (i.e. x AND y OR z). Notice this table involves three variables (x, y and z), so we would expect it to be bigger then the previous ones.

To construct a truth table, firstly we write down all the possible combinations of the three variables:

x	y	z
0	0	0
0	0	1
0	1	0
0	1	1
1	0	0
1	0	1
1	1	0

1	1	1

There is a pattern to the way the combinations are written down. We always start with 000 and end with 111. As to the middle part, it is up to the reader to figure out.

We then complete the table by hand computing what value each combination is going to produce using the expression xy + z. For example:

000

x = 0, y = 0 and z = 0

xy + z = 0

001

x = 0, y = 0 and z = 1

xy + z = 1

We continue in this way until we fill up the whole table

x	y	z	xy OR z
0	0	0	0
0	0	1	1
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	1
1	1	1	1

The procedure we follow to produce truth tables are now clear. Here are a few more examples of truth tables.

Example 1 -- x + y + z

x	y	z	x+y+z
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	1
1	0	0	1
1	0	1	1
1	1	0	1
1	1	1	1

Example 2 -- (x + yz)'

x	y	z	x+yz	(x+yz)'
0	0	0	0	1
0	0	1	0	1
0	1	0	0	1
0	1	1	1	0
1	0	0	1	0
1	0	1	1	0
1	1	0	1	0
1	1	1	1	0

When an expression is hard to compute, we can first compute intermediate results and then the final result.

Example 3 -- (x + yz')w

x	y	z	w	(x+yz')w
0	0	0	0	0
0	0	0	1	0
0	0	1	0	0
0	0	1	1	0
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	0
1	0	0	0	0
1	0	0	1	1
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	1
1	1	1	0	0
1	1	1	1	1

Exercise

Produce the truth tables for the following operations:

8. NAND: x NAND y = NOT (x AND y)

9. NOR: x NOR y = NOT (x OR y)

10. XOR: x XOR y is true if and ONLY if one of x or y is true.

Produce truth tables for:

4. xyz

5. x'y'z'

6. xyz + xy'z

7. xz

8. (x + y)'

9. x'y'

10. (xy)'

11. x' + y'

Laws of Boolean algebra

In ordinary algebra, two expressions may be equivalent to each other, e.g. xz + yz = (x + y)z. The same can be said of Boolean algebra. Let's construct truth tables for:

xz + yz

(x + y)z

xz + yz

x	y	z	xz+yz
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

(x + y)z

x	y	z	(x+y)z
0	0	0	0
0	0	1	0
0	1	0	0
0	1	1	1
1	0	0	0
1	0	1	1
1	1	0	0
1	1	1	1

By comparing the two tables, you will have noticed that the outputs (i.e. the last column) of the two tables are the same!

Definition

We say two Boolean expressions are **equivalent** if the output of their truth tables are the same.

We list a few expressions that are equivalent to each other

x + 0 = x

x Ã- 1 = x

xz + yz = (x + y)z

x + x' = 1

x Ã- x' = 0

x Ã- x = x

x + yz = (x + y)(x + z)

Take a few moments to think about why each of those laws might be true.

The last law is not obvious but we can prove that it's true using the other laws:

$$\begin{aligned}(x+y)(x+z) &= x(x+z)+y(x+z)\\ &= xx+xz+xy+yz\\ &= x+xz+xy+yz\\ &= x(1+z+y)+yz\\ &= x+yz\end{aligned}$$

As Dr Kuo Tzee-Char, Honorary reader of mathematics at the University of Sydney, is so fond of saying: "The only thing to remember in mathematics is that there is nothing to remember. Remember that!". You should not try to commit to memory the laws as they are stated, because some of them are so deadly obvious once you are familiar with the AND, OR and NOT operations. You should only try to remember those things that are most basic, once a high level of familiarity is developed, you will agree there really isn't anything to remember.

Simplification

Once we have those laws, we will want to simplify Boolean expressiosn just like we do in ordinary algebra. We can all simplify the following example with ease:

$$\begin{aligned} xyzw' + xyzw &= xyz(w + w') \\ &= xyz \end{aligned}$$

the same can be said about:

$$\begin{aligned} (x + y)(x' + y') &= x(x' + y') + y(x' + y') \\ &= xx' + xy' + yx' + yy' \\ &= 0 + xy' + yx' + 0 \\ &= xy' + yx' \end{aligned}$$

From those two examples we can see that complex-looking expressions can be reduced very significantly. Of particular interest are expressions of the form of a *sum-of-product*, for example:

xyz + xyz' + xy'z + x'yz + x'y'z' + x'y'z

We can factorise and simplify the expression as follows

xyz + *xyz*' + *xy*'z + *x*'yz + *x*'*y*'z' + *x*'*y*'z

$$= xy(z + z') + xy'z + x'yz + x'y'(z' + z)$$

$$= xy + xy'z + x'yz + x'y'$$

$$= x(y + y'z) + x'(yz + y')$$

It is only hard to go any further, although we can. We use the identity:

x + yz = (x + y)(x + z)

If the next step is unclear, try constructing truth tables as an aid to understanding.

$$\begin{aligned} &= x(y + z) + x'(z + y') \\ &= xy + xz + x'z + x'y' \\ &= xy + (x + x')z + x'y' \\ &= xy + z + x'y' \end{aligned}$$

And this is as far as we can go using the algebraic approach (or any other approach). The algebraic approach to simplication relies on the priniciple of elimination. Consider, in ordinary algebra:

x + y - x

We simplify by rearranging the expression as follows

(x - x) + y = y

Although we only go through the process in our head, the idea is clear: we *bring* together terms that cancel themselves out and so the expression is simplified.

De Morgan's theorems

So far we have only dealt with expressions in the form of a *sum of products* e.g. xyz + x'z + y'z'. De Morgan's theorems help us to deal with another type of Boolean expressions. We revisit the AND and OR truth tables:

x	y	x Ã- y	x + y
0	0	0	
0			
0	1	0	
1			
1	0	0	
1			
1	1	1	
1			

You would be correct to suspect that the two operations are connected somehow due to the similarities between the two tables. In fact, if you invert the AND operation, i.e. you perform the NOT operations on x AND y. The outputs of the two operations are almost the same:

x	y	(x Ã- y)'	x + y
0	0	1	
0			
0	1	1	
1			
1	0	1	
1			
1	1	0	
1			

The connection between AND, OR and NOT is revealed by *reversing* the output of x + y by replacing it with x' + y'.

x	y	(x Ã- y)'	x' + y'
0	0	1	
1			
0	1	1	
1			
1	0	1	
1			
1	1	0	
0			

Now the two outputs match and so we can equate them:

(xy)' = x' + y'

this is one of de Morgan's laws. The other which can be derived using a similar process is:

(x + y)' = x'y'

We can apply those two laws to simplify equations:

Example **1**
Express *x* in *sum of product* form

$$\begin{aligned} x &= (ab' + c)' \\ &= (ab')'c' \\ &= (a' + b)c' \\ &= a'c' + bc' \end{aligned}$$

Example **2**
Express *x* in *sum of product* form

$$\begin{aligned} x &= (a + b + c)' \\ &= (a + b)'c' \\ &= a'b'c' \end{aligned}$$

This points to a possible extension of De Morgan's laws to 3 or more variables.

Example **3**

Express *x* in *sum of product* form

$$\begin{aligned} x &= [(a'+c)\cdot(b+d')]' \\ &= (a'+c)'+(b+d')' \\ &= ac'+b'd \end{aligned}$$

Example **4**

Express *x* in *sum of product* form

$$\begin{aligned} x &= [(a+bc)\cdot(d+ef)]' \\ &= (a+bc)'+(d+ef)' \\ &= a'(bc)'+d'(ef)' \\ &= a'(b'+c')+d'(e'+f') \\ &= a'b'+a'c'+d'e'+d'f' \end{aligned}$$

Another thing of interest we learnt is that we can *reverse* the truth table of any expression by replacing each of its variables by their opposites, i.e. replace x by x' and y' by y etc. This result shouldn't have been a suprise at all, try a few examples yourself.

De Morgan's laws

$(x + y)' = x'y'$

$(xy)' = x' + y'$

Exercise

3. Express in simplified sum-of-product form:

1. z = ab'c' + ab'c + abc

2. z = ab(c + d)

3. z = (a + b)(c + d + f)

4. z = a'c(a'bd)' + a'bc'd' + ab'c

5. z = (a' + b)(a + b + d)d'

4. Show that x + yz is equivalent to (x + y)(x + z)

Propositions

We have been dealing with propositions since the start of this chapter, although we are not told they are propositions. A proposition is simply a statement (or sentence) that is either TRUE or FALSE. Hence, we can use Boolean algebra to handle propositions.

There are two special types of propositions -- tautology and contradiction. A tautology is a proposition that is always TRUE, e.g. "1 + 1 = 2". A contradiction is the opposite of a tautology, it is a proposition that is always FALSE, e.g. 1 + 1 = 3. As usual, we use 1 to represent TRUE and 0 to represent FALSE. Please note that opinions are not propositions, e.g. "George W. Bush started the war on Iraq for its oil." is just an opinion, its truth or falsity is not universal, meaning some think it's true, some do not.

Examples

- "It is raining today" is a proposition.
- "Sydney is in Australia" is a proposition.
- "1 + 2 + 3 + 4 + 5 = 16" is a proposition.
- "Earth is a perfect sphere" is a proposition.
- "How do you do?" is *not* a proposition - it's a question.
- "Go clean your room!" is *not* a proposition - it's a command.
- "Martians exist" is a proposition.

Since each proposition can only take two values (TRUE or FALSE), we can represent each by a *variable* and decide whether compound propositions are true by using Boolean algebra, just like we have been doing. For example "It is always hot in Antarctica OR 1 + 1 = 2" will be evaluated as true.

Implications

Propositions of the type if *something something* then *something something* are called implications. The logic of implications are widely applicable in mathematics, computer science and general everyday common sense reasoning! Let's start with a simple example

"**If** 1 + 1 = 2 **then** 2 - 1 = 1"

is an example of implication, it simply says that 2 - 1 = 1 is a consequence of 1 + 1 = 2. It's like a cause and effect relationship. Consider this example:

John says: "*If* I become a millionaire, *then* I will donate $500,000 to the Red Cross."

There are four situations:

5. John becomes a millionaire and donates $500,000 to the Red Cross

6. John becomes a millionaire and does not donate $500,000 to the Red Cross

7. John does not become a millionaire and donates $500,000 to the Red Cross

8. John does not become a millionaire and does not donate $500,000 to the Red Cross

In which of the four situations did John NOT fulfill his promise? Clearly, if and only if the second situation occured. So, we say the proposition is FALSE if and only if John becomes a millionaire and does not donate. If John did not become a millionaire then he can't break his promise, because his promise is now claiming nothing, therefore it must be evaluated TRUE.

If *x* and *y* are two propositions, *x* implies *y* (if *x* then *y*), or symbolically

$$x \Rightarrow y$$

has the following truth table:

x	y	$x \Rightarrow y$
0	0	1
0	1	1
1	0	0
1	1	1

For emphasis, $x \Rightarrow y$ is FALSE if and only if *x* is true and *y* false. If *x* is FALSE, it does not matter what value *y* takes, the proposition is automatically TRUE. On a side note, the two propositions *x* and *y* need not have anything to do with each other, e.g. "1 + 1 = 2 implies Australia is in the southern hemisphere" evaluates to TRUE!

If

$$(x \Rightarrow y) \text{ AND } (y \Rightarrow x)$$

then we express it symbolically as

$$x \Leftrightarrow y.$$

It is a two way implication which translates to *x* is TRUE if and only if *y* is true. The *if and only if* operation has the following truth table:

x	y	$x \Leftrightarrow y$
0	0	1
0	1	0
1	0	0
1	1	1

The two new operations we have introduced are not really new, they are just combinations of AND, OR and NOT. For example:

$$x \Rightarrow y = x' + y$$

Check it with a truth table. Because we can express the *implication* operations in terms of AND, OR and NOT, we have open them to manipulation by Boolean algebra and de Morgan's laws.

Example 1

Is the following proposition a tautology (a proposition that's always true)

$$[(x \Rightarrow y)(y \Rightarrow z)] \Rightarrow (x \Rightarrow z)$$

Solution 1

$$= [(x \Rightarrow y)(y \Rightarrow z)] \Rightarrow (x \Rightarrow z)$$

$$= [(x' + y)(y' + z)]' + (x' + z)$$

$$= (x' + y)' + (y' + z)' + x' + z$$

$$= xy' + yz' + x' + z$$

$$= 1 = y' + y + x' + z$$

Therefore it's a tautology.

Solution 2

A somewhat easier solution is to draw up a truth table of the proposition, and note that the output column are all 1s. Therefore the proposition is a tautology, because the output is 1 regardless of the *inputs* (i.e. x, y and z).

Example 2

Show that the proposition z is a contradiction (a proposition that is always false):

z = xy(x + y)'

Solution

$$\begin{aligned} z &= xy(x+y)' \\ &= xy(x'y') \\ &= 0 \end{aligned}$$

Therefore it's a contradiction.

Back to Example 1, : $[(x \Rightarrow y)(y \Rightarrow z)] \Rightarrow (x \Rightarrow z)$. This isn't just a slab of symbols, you should be able translate it into everyday language and understand intuitively why it's true.

Exercises

8. Decide whether the following propositions are true or false:

1. If 1 + 2 = 3, then 2 + 2 = 5

2. If 1 + 1 = 3, then fish can't swim

9. Show that the following pair of propositions are equivalent

1. $x \Rightarrow y$: $y' \Rightarrow x'$

Logic Puzzles

Puzzle is an all-encompassing word, it refers to anything trivial that requires solving. Here is a collection of logic puzzles that we can solve using Boolean algebra.

Example 1

We have two type of people -- knights or knaves. A knight always tell the truth but the knaves always lie.

Two people, Alex and Barbara, are chatting. Alex says :"We are both knaves"

Who is who?

We can probably work out that *Alex* is a knave in our heads, but the algebraic approach to determine *Alex* 's identity is as follows:

Let *A* be TRUE if Alex is a knight

Let *B* be TRUE if Barbara is a knight

There are *two* situations, either:

Alex is a knight and what he says is TRUE, OR

he is NOT a knight and what he says it FALSE.

There we have it, we only need to translate it into symbols:

A(A'B') + A'[(A'B')'] = 1

we simplify:

(AA')B' + A'[A + B] = 1

A'A + A'B = 1

A'B = 1

Therefore *A* is FALSE and *B* is TRUE. Therefore Alex is a knave and Barbara a knight.

Example 2

There are three businessmen, conveniently named **A**bner, **B**ill and **C**harley, who order martinis together every weekend according to the following rules:

4. If A orders a martini, so does B.
5. Either B or C always order a martini, but never at the same lunch.
6. Either A or C always order a martini (or both)
7. If C orders a martini, so does A.

2. $A \Rightarrow B$ or $AB + A'B' = 1$
3. $B'C + BC' = 1$
4. $A + C = 1$
5. $C \Rightarrow A$ or $CA + C'A' = 1$

Putting all these into one formula and simplifying:

$$
\begin{aligned}
1 &= (AB+A'B')(B'C+BC')(A+C)(CA+C'A) \\
&= (AB+A'B')(B'C+BC')(A+C)(C+C')A \\
&= (AB+A'B')(B'C+BC')(A+C)1A \\
&= (AB+A'B')(B'C+BC')(A+C)A \\
& \text{Now that we know that } A=1 \text{ we can substitute that in:} \\
&= (1B+0B')(B'C+BC')(1+C)1 \\
&= (B)(B'C+BC') \\
& \text{Now that we know that } B=1 \text{ we can substitute that in:} \\
&= (1)(0C+1C') \\
&= C' \\
& \text{If } 1=C' \text{ then } C=0 \\
& ABC'=1
\end{aligned}
$$

Exercises

Please go to Logic puzzles.

Problem Set

1. Decide whether the following propositions are equivalent:

$x' \Rightarrow y'$

$y \Rightarrow x$

2. Express in simplest sum-of-product form the following proposition:

$(x \Leftrightarrow y) \Rightarrow z$

3. Translate the following sentences into symbolic form and decide if it's true:

a. For all x, if $x^2 = 9$ then $x^2 - 6x - 3 = 0$

b. We can find a x, such that $x^2 = 9$ and $x^2 - 6x - 3 = 0$ are both true.

4. NAND is a binary operation:

x NAND y = (xy)'

Find a proposition that consists of only NAND operators, equivalent to:

(x + y)w + z

5. Do the same with NOR operators. Recall that x NOR y = (x + y)'

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Mathematical proofs

"It is by logic that we prove, but by intuition that we discover."

Introduction

Mathematicians have been, for the past five hundred years or so, obsessed with proofs. They want to prove everything, and in the process proved that they can't prove everything (see this). This chapter will introduce the techniques of mathematical induction, proof by contradiction and the axiomatic approach to mathematics.

Mathematical induction

Deductive reasoning is the process of reaching a conclusion that is guaranteed to follow. For example, if we know

- All ravens are black birds, and
- For every action, there is an equal and opposite reaction

then we can conclude:

- This bird is a raven, therefore it is black.
- This billiard ball will move when struck with a cue.

Induction is the opposite of deduction. To induce, we observe how things behave in specific cases and from that we draw conclusions as to how things behave in the general case. For example:

$$1 + 2 + 3 + ... + n = \frac{(n+1)n}{2}$$

We know it is true for all numbers, because Gauss told us. But how do we show that it's true for all positive integers? Even if we can show the identity holds for numbers from one to a trillion or any larger number we can think of, we still haven't proved that it's true for all positive integers. This is where mathematical induction comes in, it works somewhat like the dominoes.

If we can show that the identity holds for some number k, and that mere fact implies that the identity also holds for $k + 1$, then we have effectively shown that it works for all integers.

Example 1 Show that the identity

$$1 + 2 + 3 + ... + n = \frac{(n+1)n}{2}$$

holds for all positive integers.

Solution Firstly, we show that it holds for integers 1, 2 and 3

1 = 2Ã-1/2

1 + 2 = 3Ã-2/2

1 + 2 + 3 = 4Ã-3/2 = 6

Suppose the identity holds for some number *k*:

$$1 + 2 + 3 + ... + k = \frac{1}{2}(k+1)k$$

This supposition is known as the induction hypothesis. We assume it is true, and aim to show that,

$$1 + 2 + 3 + ... + k + (k+1) = \frac{1}{2}(k+2)(k+1)$$

is also true.

We proceed

$$\begin{aligned} 1 + 2 + 3 + ... + k \quad &= \quad \tfrac{1}{2}(k+1)k \\ 1 + 2 + 3 + ... + k \quad +(k+1) &= \tfrac{1}{2}(k+1)k + (k+1) \\ &= (k+1)(\tfrac{k}{2} + 1) \\ &= \tfrac{1}{2}(k+1)(k+2) \end{aligned}$$

which is what we have set out to show. Since the identity holds for 3, it also holds for 4, and since it holds for 4 it also holds for 5, and 6, and 7, and so on.

There are two types of mathematical induction: strong and weak. In weak induction, you assume the identity holds for certain value k, and prove it for k+1. In strong induction, the identity must be true for any value lesser or equal to k, and then prove it for k+1.

Example 2 Show that $n! > 2^n$ for $n \geq 4$.

Solution The claim is true for n = 4. As $4! > 2^4$, i.e. 24 > 16. Now suppose it's true for n = k, k ≥ 4, i.e.

$k! > 2^k$

it follows that

$(k+1)k! > (k+1)2^k > 2^{k+1}$

$(k+1)! > 2^{k+1}$

We have shown that if for n = k then it's also true for n = k + 1. Since it's true for n = 4, it's true for n = 5, 6, 7, 8 and so on for all n.

Example 3 Show that

$$1^3 + 2^3 + ... + n^3 = \frac{(n+1)^2 n^2}{4}$$

Solution Suppose it's true for n = k, i.e.

$$1^3 + 2^3 + ... + k^3 = \frac{(k+1)^2 k^2}{4}$$

it follows that

$$\begin{aligned} 1^3 + 2^3 + ... + k^3 + (k+1)^3 &= \frac{(k+1)^2 k^2}{4} + (k+1)^3 \\ &= (k+1)^2 (\frac{k^2}{4} + (k+1)) \\ &= \frac{1}{4}(k+1)^2 (k^2 + 4k + 4) \\ &= \frac{1}{4}(k+1)^2 (k+2)^2 \end{aligned}$$

We have shown that if it's true for n = k then it's also true for n = k + 1. Now it's true for n = 1 (clear). Therefore it's true for all integers.

Exercises

1. Prove that $1^2 + 2^2 + ... + n^2 = \frac{n(2n^2 + 3n + 1)}{6}$

2. Prove that for $n \geq 1$,

$$(1 + \sqrt{5})^n = x_n + y_n \sqrt{5}$$

where x_n and y_n are integers.

3. Note that

$$\sum_{i=1}^{n}[i^k - (i-1)^k] = n^k$$

Prove that there exists an explicit formula for

$$\sum_{i=1}^{n} i^m$$ for all integer *m*. E.g.

$$1^3 + 2^3 + \ldots + n^3 = \frac{n^2(n+1)^2}{4}$$

4. The sum of all of the interior angles of a triangle is 180° ; the sum of all the angles of a rectangle is 360° . Prove that the sum of all the angles of a polygon with *n* sides, is $(n-2)\cdot 180^\circ$.

Proof by contradiction

"When you have eliminated the impossible, what ever remains, however improbable must be the truth." Sir Arthur Conan Doyle

The idea of a proof by contradiction is to:

6. First, we assume that the *opposite* of what we wish to prove is true.
7. Then, we show that the logical consequences of the assumption include a contradiction.
8. Finally, we conclude that the assumption must have been false.

Root 2 is irrational

As an example, we shall prove that $\sqrt{2}$ is not a rational number. Recall that a rational number is a number which can be expressed in the form of p/q, where p and q are integers and q does not equal 0 (see the 'categorizing numbers' section here).

First, assume that $\sqrt{2}$ is *rational*:

$$\sqrt{2} = \frac{a}{b}$$

where *a* and *b* are coprimes (i.e. both integers with no common factors). If *a* and *b* are not coprimes, we remove all common factors. In other words, *a/b* is in simplest form. Now, continuing:

$$\begin{aligned} \sqrt{2} &= a/b \\ 2 &= a^2/b^2 \\ 2b^2 &= a^2 \end{aligned}$$

We have now found that a^2 is some integer multiplied by 2. Therefore, a^2 must be divisible by two. If a^2 is even, then *a* must also be even, for an odd number squared yields an odd number. Therefore we can write $a = 2c$, where *c* is another integer.

$$\begin{aligned} 2b^2 &= a^2 \\ 2b^2 &= (2c)^2 \\ 2b^2 &= 4c^2 \\ b^2 &= 2c^2 \end{aligned}$$

We have discovered that b^2 is also an integer multiplied by two. It follows that *b* must be even. We have a contradiction! Both *a* and *b* are even integers. In other words, both have the common factor of 2. But we already said that *a/b* is in simplest form. Since such a contradiction has been established, we must conclude that our original assumption was false. Therefore, âˆš2 is irrational.

Contrapositive

Some propositions that take the form of *if xxx then yyy* can be hard to prove. It is sometimes useful to consider the *contrapositive* of the statement. Before I explain what contrapositive is let us see an example

"If x^2 is odd then *x* is is also odd"

is harder to prove than

"if *x* even then x^2 is also even"

although they mean the same thing. So instead of proving the first proposition directly, we prove the second proposition instead.

If *A* and *B* are two propositions, and we aim to prove

If *A* is true then *B* is true

we may prove the equivalent statement

If *B* is false then *A* is false

instead. This technique is called proof by contrapositive.

To see why those two statements are equivalent, we show the following boolean algebra expressions is true (see Logic)

$$p \Rightarrow q \equiv q' \Rightarrow p'$$

(to be done by the reader).

Exercises

1.Prove that there is no perfect square number for 11,111,1111,11111......

2. Prove that there are infinitely number of k's such that, $4k + 3$, is prime. (Hint: consider $N = p_1 p_2 ... p_m + 3$)

Reading higher mathematics

This is some basic information to help with reading other higher mathematical literature. ... *to be expanded*

Quantifiers

Sometimes we need propositions that involve some description of rough quantity, e.g. "For *all* odd integers x, x^2 is also odd". The word *all* is a description of quantity. The word "some" is also used to describe quantity.

Two special symbols are used to describe the quanties "all" and "some"

$\forall$ means "for all" or "for any"

$\exists$ means "there are some" or "there exists"

Example 1
The proposition:

For all even integers x, *x^2 is also even.*

can be expressed symbolically as:

$$(\forall x)(x \text{ is even} \Rightarrow x^2 \text{ is even})$$

Example 2
The proposition:

There are some odd integers x, *such that x^2 is even.*

can be expressed symbolically as:

$$(\exists x)(x \text{ is odd} \Rightarrow x^2 \text{ is even})$$

This proposition is false.

Example 3
Consider the proposition concerning (z = x'y' + xy):

For any value of x, *there exist a value for* y, *such that z = 1.*

can be expressed symbolically as:

$$(\forall x)(\exists y)(z = 1)$$

This proposition is true. Note that the order of the quantifiers is important. While the above statement is true, the statement

$$(\exists y)(\forall x)(z = 1)$$

is false. It asserts that there is one value of y which is the same for all x for which z=1. The first statement only asserts that there is a y for each x, but different values of x may have different values of y.

Negation

Negation is just a fancy word for the opposite, e.g. The *negation* of "All named Britney can sing" is "Some named Britney can't sing". What this says is that to disprove that all people named Britney can sing, we only need to find one named Britney who can't sing. To express symbolically:

Let *p* represent a person named Britney

$$[(\forall p)(p \text{ can sing})]' = (\exists p)(p \text{ cannot sing})$$

Similarly, to disprove

$$(\forall x)(x \text{ is odd} \Rightarrow x^2 \text{ is even})$$

we only need to find one odd number that doesn't satisfy the condition. Three is odd, but 3Ã-3 = 9 is also odd, therefore the proposition is FALSE and

$$(\exists x)(x \text{ is odd} \Rightarrow x^2 \text{ is odd})$$

is TRUE

In summary, to obtain the *negation* of a proposition involving a quantifier, you replace the quantifier by its opposite (e.g. $\forall$ with $\exists$) and the *quantified proposition* (e.g. "x is even") by its negation (e.g. "x is odd").

Example 1

$$(\forall x)(\exists y)(x(x+1)(x+2)(x+3)+1=y^2)$$

is a true statement. Its negation is

$$(\exists x)(\forall y)(x(x+1)(x+2)(x+3)+1\neq y^2)$$

Axioms and Inference

If today's mathematicians were to describe the greatest achievement in mathematics in the 20th century in one word, that word will be **abstraction**. True to its name, abstraction is a very abstract concept (see Abstraction).

In this chapter we shall discuss the *essence* of some of the number systems we are familiar with. For example, the real numbers and the rational numbers. We look at the most fundamental properties that, in some sense, *define* those number systems.

We begin our discussion by looking at some of the more obscure results we were told to be true

- 0 times any number gives you 0
- a negative number multiplied by a negative number gives you a positive number

Most people simply accept that they are true (and they are), but the two results above are simple consequences of what we believe to be true in a number system like the real numbers!

To understand this we introduce the idea of axiomatic mathematics (mathematics with simple assumptions). An axiom is a statement about a number system that we assume to be true. Each number system has a few axioms, from these axioms we can draw conclusions (inferences).

Let's consider the Real numbers, it has axioms Let *a*, *b* and *c* be real numbers

For *a*, *b*, and *c* taken from the real numbers

A1: *a*+*b* is a real number also (*closure*)

A2: There exist 0, such that 0 + *a* = *a* for all *a* (existence of zero - an *identity*)

A3: For every *a*, there exist *b* (written -*a*), such that a + b = 0 (existence of an additive inverse)

A4: (*a* + *b*) + *c* = *a* + (*b* + *c*) (associativity of addition)

A5: *a* + *b* = *b* + *a* (commutativity of addition)

For *a*, *b*, and *c* taken from the real numbers excluding zero

M1: *ab* (*closure*)

M2: There exist an element, 1, such that 1*a* = *a* for all *a* (existence of one - an *identity*)

M3: For every a there exists a b such that $ab = 1$

M4: $(ab)c = a(bc)$ (associativity of multiplication)

M5: $ab = ba$ (commutativity of multiplication)

D1: $a(b + c) = ab + ac$ (distributivity)

These are the *minimums* we assume to be true in this system. These are *minimum* in the sense that everything else that is true about this number system can be derived from those axioms!

Let's consider the following true identity

(x + y)z = xz + yz

which is not included in the axioms, but we can prove it using the axioms. We proceed:

$$\begin{aligned}(x+y)z &= z(x+y) \text{ by M5}\\ &= zx + zy \text{ by D1}\\ &= xz + yz \text{ by M5}\end{aligned}$$

Before we proceed any further, you will have notice that the real numbers are not the only numbers that satifies those axioms! For example the rational numbers also satify all the axioms. This leads to the abstract concept of a *field*. In simple terms, a *field* is a number system that satisfies all those axiom. Let's define a *field* more carefully:

A number system, F, is a *field* if it supports + and Ã- operations such that:

For a, b, and c taken from F

A1: $a + b$ is in F also (*closure*)

A2: There exist 0, such that $0 + a = a$ for all a (existence of zero - an *identity*)

A3: For every a, there exist b (written $-a$), such that a + b = 0 (existence of an additive inverse)

A4: $(a + b) + c = a + (b + c)$ (associativity of addition)

A5: $a + b = b + a$ (commutativity of addition)

For a, b, and c taken from F with the zero removed (sometimes written F^*)

M1: ab is in F (*closure*)

M2: There exist an element, 1, such that $1a = a$ for all a (existence of one - the *identity*)

M3: For every a there exists a b such that $ab = 1$ (inverses)

M4: $(ab)c = a(bc)$ (associativity of multiplication)

M5: $ab = ba$ (commutativity of multiplication)

D1: $a(b + c) = ab + ac$ (distributivity)

Now, for **M3**, we do not let b be zero, since 1/0 has no meaning. However for the *M* axioms, we have excluded zero anyway.

For interested students, the requirements of *closure*, *identity*, having *inverses* and *associativity* on an operation and a set are known as a group. If F is a group with addition and F^* is a group with multiplication, plus the distributivity requirement, F is a field. The above axioms merely state this fact in full.

Note that the natural numbers are not a field, as **M3** is general not satified, i.e. not every natural number has an inverse that is also a natural number.

Please note also that $(-a)$ denotes the additive inverse of a, it doesn't say that (-a) = (-1)(a), although we can prove that they are equivalent.

Example 1

Prove using only the axioms that 0 = -0, where -0 is the additive inverse of 0.

Solution 1

0 = 0 + (-0) by **A3:** existence of inverse

0 = (-0) by **A2:** 0 + a = a

Example 2

Let F be a field and a an element of F. Prove using nothing more than the axioms that $0a = 0$ for all a.

Solution

0 = 0a + (-0a) by **A3** existence of inverse

0 = (0 + 0)a + (-0a) by Example 1

0 = (0a + 0a) + (-0a) by distributivity and commutativity of multiplication

0 = 0a + (0a + (-0a)) by associativity of addition

0 = 0a + 0 by **A3**

0 = 0a by **A2**.

Example 3

Prove that $(-a) = (-1)a$.

Solution 3

(-a) = (-a) + 0

(-a) = (-a) + 0a by Example 2

(-a) = (-a) + (1 + (-1))a

(-a) = (-a) + (1a + (-1)a)

(-a) = (-a) + (a + (-1)a)

(-a) = ((-a) + a) + (-1)a

(-a) = 0 + (-1)a

(-a) = (-1)a

One wonders why we need to prove such obvious things (obvious since primary school). But the idea is not to prove that they are true, but to practise inferencing, how to logically join up arguments to prove a point. That is a vital skill in mathematics.

Exercises

1. Describe a field in which 1 = 0

2. Prove using only the axioms if u + v = u + w then v = w (subtracting u from both sides is not accepted as a solution)

3. Prove that if xy = 0 then either x = 0 or y = 0

4. In F_-, the operation + is defined to be the difference of two numbers and the Ã- operation is defined to be the ratio of two numbers. E.g. 1 + 2 = -1, 5 + 3 = 2 and 9Ã-3 = 3, 5Ã-2; = 2.5. Is F_- a field?

5. Explain why Z_6 (modular arithmetic modular 6) is not a field.

Problem Set

1. Prove

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + ... + \frac{1}{\sqrt{n}} \geq \sqrt{n}$$

for $n \geq 1$

2. Prove by induction that $2n^3 - 3n^2 + n + 31 \geq 0$

3. Prove by induction

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n$$

where

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}$$ and $$n! = n \cdot (n-1) \cdot (n-2) \cdot ... \cdot 2 \cdot 1$$

and 0! = 1 by definition.

4. Prove by induction $\binom{n}{0} + 2\binom{n}{1} + 2^2\binom{n}{2} + ... + 2^n\binom{n}{n} = 3^n$

5. Prove that if x and y are integers and n an odd integer then $\frac{x^n + y^n}{x + y}$ is an integer.

6. Prove that (n~m) = n!/((n-m)!m!) is an integer. Where n! = n(n-1)(n-2)...1. E.g 3! = 3Ã-2Ã-1 = 6, and (5~3) = (5!/3!)/2! = 10.

Many questions in other chapters require you to prove things. Be sure to try the techniques discussed in this chapter.

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Exercises

Mathematical proofs

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

Mathematical induction exercises

1.

Prove that $1^2 + 2^2 + ... + n^2 = n(n+1)(2n+1)/6$

When n=1,

L.H.S. = $1^2 = 1$

R.H.S. = 1*2*3/6 = 6/6 = 1

Therefore L.H.S. = R.H.S.

Therefore this is true when n=1.

Assume that this is true for some positive integer k,

i.e. $1^2 + 2^2 + ... + k^2 = k(k+1)(2k+1)/6$

$$\begin{aligned} 1^2+2^2+3^2+...+k^2 &= \tfrac{k(k+1)(2k+1)}{6} \\ 1^2+2^2+3^2+...+k^2+(k+1)^2 &= \tfrac{k(k+1)(2k+1)}{6}+(k+1)^2 \\ &= \tfrac{1}{6}(k+1)\left[k(2k+1)+6(k+1)\right] \\ &= \tfrac{1}{6}(k+1)\left[2k^2+7k+6\right] \\ &= \tfrac{(k+1)(k+2)(2k+3)}{6} \end{aligned}$$

Therefore this is also true for k+1.

Therefore, by the principle of mathematical induction, this holds for all positive integer n.

2. Prove that for $n \geq 1$,

$$(1+\sqrt{5})^n = x_n + y_n\sqrt{5}$$

where x_n and y_n are integers.

When n=1,

$$1+\sqrt{5}=x_1+y_1\sqrt{5}$$

Therefore x_1=1 and y_1=1, which are both integers.

Therefore this is true when n=1.

Assume that this is true for some positive integer k,

i.e. $(1+\sqrt{5})^k=x_k+y_k\sqrt{5}$ where x_k and y_k are integers.

$$\begin{aligned}(1+\sqrt{5})^k &= x_k+y_k\sqrt{5}\\ (1+\sqrt{5})^{k+1} &= (x_k+y_k\sqrt{5})(1+\sqrt{5})\\ &= x_k+y_k\sqrt{5}+x_k\sqrt{5}+5y_k\\ &= (x_k+5y_k)+(x_k+y_k)\sqrt{5}\end{aligned}$$

Because x_k and y_k are both integers, therefore $x_k + 5y_k$ and $x_k + y_k$ are integers also.

Therefore this is true for k+1 also.

Therefore, by the principle of mathematical induction, this holds for all positive integer n.

3. (The solution assume knowledge in binomial expansion and summation notation)

Note that

$$\sum_{i=1}^{n}[i^k-(i-1)^k]=n^k$$

Prove that there exists an explicit formula for

$$\sum_{i=1}^{n} i^m$$

for all integer *m*. E.g.

$$1^3+2^3+\ldots+n^3=\frac{n^2(n+1)^2}{4}$$

It's clear that $1^1 + 2^1 + \ldots = (n+1)n/2$. So the proposition is true for m=1.

Suppose that

$$\sum_{i=1}^{n} i^j$$

has an explicit formula in terms of *n* for all *j* < *k* (**), we aim to prove that

$$\sum_{i=1}^{n} i^k$$

also has an explicit formula.

Starting from the property given, i.e.

$$\sum_{i=1}^{n}[i^{k+1} - (i-1)^{k+1}] = n^{k+1}$$

$$\sum_{i=1}^{n}[i^{k+1} - \sum_{j=0}^{k+1} \binom{k+1}{j} i^j] = n^{k+1}$$

$$\sum_{i=1}^{n}[i^{k+1} - \binom{k+1}{k+1} i^{k+1} - \sum_{j=0}^{k} \binom{k+1}{j} i^j] = n^{k+1}$$

$$\sum_{i=1}^{n}[\sum_{j=0}^{k} \binom{k+1}{j} i^j] = n^{k+1}$$

$$\sum_{j=0}^{k}[\sum_{i=1}^{n} \binom{k+1}{j} i^j] = n^{k+1}$$

$$\sum_{j=0}^{k}[\binom{k+1}{j} \sum_{i=1}^{n} i^j] = n^{k+1}$$

Since we know the formula for power sum of any power less then *k* (**), we can solve the above equation and find out the formula for the *k*-th power directly.

Hence, by the principle of strong mathematical induction, this proposition is true.

Additional info for question 3

The method employed in question 3 to find out the general formula for power sum is called **the method of difference**, as shown by that we consider the sum of all difference of adjacant terms. Aside from the method above, which lead to a recursive solution for finding the general formula, there're also other methods, such as that of using generating function. Refer to the last question in the generating function project page for detail.

Problem set

Mathematical Proofs Problem Set

1.

For all

$$
\begin{array}{ccc}
a & > & 0 \\
n+a & > & n \\
n & > & n-a \\
\sqrt{n} & > & \sqrt{n-a} \\
1 & > & \frac{\sqrt{n-a}}{\sqrt{n}} \\
\frac{1}{\sqrt{n-a}} & > & \frac{1}{\sqrt{n}}
\end{array}
$$

Therefore

$$\frac{1}{\sqrt{1}} \quad , \quad \frac{1}{\sqrt{2}} \quad , \quad \frac{1}{\sqrt{3}} \quad \cdots\cdots\cdots \quad > \frac{1}{\sqrt{n}}$$

When a>b and c>d, a+c>b+d (See also Replace it if you find a better one).

Therefore we have:

$$\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \cdots\cdots + \frac{1}{\sqrt{n}} > n \times \frac{1}{\sqrt{n}}$$

$$
\begin{array}{l}
\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \cdots\cdots + \frac{1}{\sqrt{n}} > \frac{n}{\sqrt{n}} \times \frac{\sqrt{n}}{\sqrt{n}} \\
\frac{1}{\sqrt{1}} + \frac{1}{\sqrt{2}} + \frac{1}{\sqrt{3}} \cdots\cdots + \frac{1}{\sqrt{n}} > \sqrt{n}\;\overline{n} \\
\frac{}{\sqrt{1}} + \frac{}{\sqrt{2}} + \frac{}{\sqrt{3}} \cdots\cdots + \frac{}{\sqrt{n}} > \frac{}{n}
\end{array}
$$

3.Let us call the proposition

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n$$

be P(n)

Assume this is true for some n, then

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{n} = 2^n$$

$$2 \times \left\{ \binom{n}{0} + \binom{n}{1} + \binom{n}{2} + ... + \binom{n}{2} \right\} = 2^{n+1}$$

$$\left\{ \binom{n}{0} + \binom{n}{n} \right\} + \left\{ \binom{n}{0} + 2\binom{n}{1} + 2\binom{n}{2} + ... + 2\binom{n}{n-1} + \binom{n}{n} \right\} = 2^{n+1}$$

$$\left\{ \binom{n}{0} + \binom{n}{n} \right\} + \left\{ \binom{n}{0} + \binom{n}{1} \right\} + \left\{ \binom{n}{1} + \binom{n}{2} \right\} + \left\{ \binom{n}{2} + \binom{n}{3} \right\} + ... + \left\{ \binom{n}{n-1} + \binom{n}{n} \right\} = 2^{n+1}$$

$\binom{n}{a} + \binom{n}{a+1} = \binom{n+1}{a+1}$ function:(Note:If anyone find wikibooks ever mentioned

$$\left\{ \binom{n}{0} + \binom{n}{n} \right\} + \binom{n+1}{1} + \binom{n+1}{2} + \binom{n+1}{3} + ... + \binom{n+1}{n} = 2^{n+1}$$

$$\binom{n}{0} = \binom{n}{n} = 1$$

$$\binom{n+1}{0} + \binom{n+1}{n+1} + \binom{n+1}{1} + \binom{n+1}{2} + \binom{n+1}{3} + ... + \binom{n+1}{n} = 2^{n+1}$$

$$\binom{n+1}{0} + \binom{n+1}{1} + \binom{n+1}{2} + \binom{n+1}{3} + ... + \binom{n+1}{n} + \binom{n+1}{n+1} = 2^{n+1}$$

Therefore P(n) implies P(n+1), and by simple substitution P(0) is true.

Therefore by the principal of mathematical induction, P(n) is true for all n.

Alternate solution

Notice that

$$(a+b)^n = \binom{n}{0}a^n + \binom{n}{1}a^{n-1}b + \cdots + \binom{n}{n}b^n$$

letting *a* = *b* = 1, we get

$$(1+1)^n = 2^n = \binom{n}{0} + \binom{n}{1} + \cdots + \binom{n}{n}$$

as required.

5.

Let $P(x) = x^n + y^n$ be a polynomial with x as the variable, y and n as constants.

$$\begin{aligned} P(-y) &= (-y)^n + y^n \\ &= -y^n + y^n \text{(When n is an odd integer)} \\ &= 0 \end{aligned}$$

Therefore by factor theorem(link here please), (x-(-y))=(x+y) is a factor of P(x).

Since the other factor, which is also a polynomial, has integer value for all integer x,y and n(I've skipped the part about making sure all coeifficients are of integer value for this moment), it's now obivous that

$\frac{x^n + y^n}{x + y}$ is an integer for all integer value of x,y and n when n is odd.

Infinity and infinite processes

Introduction

As soon as a child first learns about numbers, they become interested in big ones, a million, a billion, a trillion. They even make up their own, a zillion etc. One of the first mathematical questions a child asks is "what is the largest number?" This will often lead to a short explanation that there are infinitely many numbers.

But there are many different *types* of infinity - in fact, there are infinite types of infinity! This chapter will try to explain what some of these types mean and the differences between them.

Infinite Sets

How big is infinity?

Infinity is unlike a normal number because, by definition, it is not **finite**. Dividing infinity by any positive number (except infinity) gives us infinity. You can also multiply it by anything except zero (or infinity) and it will not get bigger. So let's look more carefully at the different types of infinity.

There was once a mathematician called Georg Cantor who created a new branch of mathematics called **set theory** in the late 19th century. Set theory involves collections of numbers or objects. Here's a set:

{1,2,3,4,5}

Is it the same size as this one?

{6,7,8,9,10}

Cantor's notion of sets being "the same size" does not consider whether the numbers are bigger, but whether there are the same amount of objects in it. You can easily see here that they are the same size, because you can simply count the number of members in each set. But with an infinite number of members you cannot, in a finite amount of time, count all the members of one set to see if they are the same number of them as there are in another set.

In order to decide if two infinite sets have the same number of members we need to think carefully about what we do when we count. Think of a small child sharing out sweets, between her and her brother.

"One for you, and one for me, two for you and two for me"

and so on. She knows that they both get the same number of sweets because of the way the sharing out was done. Even if she runs out of numbers (like if she can only count up to ten) she can still distribute the candy with the even process of "another one for you and another one for me".

We can use the same idea to compare infinite sets. If we can find a way to pair up one member of set A with one member of set B, and if there are no members of A without a partner in B and vice versa then we can say that set A and set B have the same number of members.

Example

Let Set N be all counting numbers. N is called the set of natural numbers. 1,2,3,4,5,6, ... and so to infinity. Let Set B be the negative numbers -1,-2,-3, ... and so on to -infinity. Can the members of N and B be paired up? The formal way of saying this is "Can A and B be put into a one to one correspondance"?

Obviously the answer is yes. 1 in set N corresponds with -1 in B. Likewise:

N	B
1	-1
2	-2
3	-3

and so on.

So useful is the set of counting numbers that any set that can be put into a one to one correspondence with it is said to be *countably infinite*.

Let's look at some more examples. Is the set of integers countably infinite? Integers are set N, set B and 0.

... -3,-2,-1, 0, 1, 2, 3, ...

Historically this set is usually called Z. Note that N the set of natural numbers is a subset of Z. All members of N are in Z, but not all members of Z are in N.

What we need to find out is if Z can be put into a one to one correspondence with N. Your first answer, given that N is a subset of Z, may be no but you would be wrong! In set theory, the *order* of the elements is unimportant. There is no reason why we can't rearrange the elements into any order we please as long as we don't leave any out. Z as presented above doesn't look countable, but if we rewrite it as 0, -1, 1, -2, 2, -3, 3 ..... and so on we can see that it is countable.

Z	N
0	1
-1	2
1	3

-2 4

and so on. Strange indeed! A subset of Z (namely the natural numbers) has the same number of members as Z itself? Infinite sets are not like ordinary finite sets. In fact this is sometimes used as a definition of an infinite set. **An infinite set is any set which can be put into a one to one correspondence with at least one of its subsets**. Rather than saying "The number of members" of a set, people sometimes use the word **cardinality** or **cardinal value**. Z and N are said to have the same cardinality.

Exercises

1. Is the number of even numbers the same as the natural numbers?

2. What about the number of square numbers?

3. Is the cardinality of positive even numbers less than 100 equal to the cardinality of natural numbers less than 100? Which set is bigger? How do you know? In what ways do finite sets differ from infinite ones?

4. Using the idea of one to one correspondance prove that infinity + 1 = infinity, what about infinity + A where A is a *finite* set? What about infinity plus C where C is a countably infinite set?

Is the set of rational numbers bigger than N?

In this section we will look to see if we can find a set that is **bigger** than the countable infinity we have looked at so far. To illustrate the idea we can imagine a story.

There was once a criminal who went to prison. The prison was not a nice place so the poor criminal went to the prison master and pleaded to be let out. She replied:

"Oh all right - I'm thinking of a number, every day you can have a go at guessing it. If you get it correct, you can leave."

Now the question is - can the criminal get himself out of jail? (Think about if for a while before you read on)

Obviously it depends on the number. If the prison master chooses a natural number, then the criminal guesses 1, to first day, 2,the second day and so on until he reaches the correct number. Likewise for the integers 0 on the first day, -1 on the second day. 1 on the third day and so on. If the number is very large then it may take a long time to get out of prison but get out he will.

What the prison master needs to do is choose a set that is not countable in this way. Think of a number line. The integers are widely spaced out. There are plenty of numbers inbetween the integers 0&5 for example. So we need to look at *denser* sets. The first set that springs to most peoples mind are the fractions. There are an infinite number of fractions between 0 and 1 so surely there are more fractions than integers? Is it possible to count fractions? Let's think about that possibility for a while. If we try to use the approach of counting all the fractions between 0 & 1 then go on to 1 - 2 and so on we will come unstuck becuase we will never finish counting the ones up to 1 (there are an infinite number of them). But does this mean that they are uncountable ? Think of the situation with the integers. Ordering them ...-2, -1, 0, 1, 2, ... renders them impossible to count, but *reordering* them 0, -1, 1, -2, 2, ... allows them to be counted.

There is in fact a way of ordering fractions to allow them to be counted. Before we go on to it let's revert to the normal mathematical language. Mathematicians use the term *rational number* to define what we have been calling fractions. A rational number is any number that can be written in the form p/q where p and q are integers. So 3/4 is rational, as is -1/2. The set of all rational numbers is usally called Q. Note that Z is a subset of Q becuase any integer can be divided by 1 to make it into a rational. E.g. the number 3 can be written in the form p/q as 3/1.

Now as all the numbers in Q are defined by two numbers p and q it makes sense to write Q out in the form of a table.

$$\begin{matrix} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \cdots \\ \\ \frac{2}{1} & \frac{2}{2} & \frac{2}{3} & \cdots \\ \\ \frac{3}{1} & \frac{3}{2} & \frac{3}{3} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{matrix}$$

Note that this table isn't an exact representation of Q. It only has the positive members of Q and has a number of multiple entries.(e.g. 1/1 and 2/2 are the same number) We shall call this set Q'. It is simple enough to see that if Q' is countable then so is Q.

So how do we go about counting Q'? If we try counting the first row then the second and so on we will fail because the rows are infinite in length. Likewise if we try to count columns. But look at the diagonals. In one direction they are infinite (e.g. 1/1, 2/2, 3/3, ...) but in the other direction they are finite. So this set is countable. We count them along the finite diagonals, 1/1, 1/2, 2/1, 1/3, 2/2, 3/1....

Exercises

1. Adapt the method of counting the set Q' to show that thet Q is also countable. How will you include 0 and the negative rationals? How will you solve the problem of multiple entries representing the same number ?
2. Show that $\infty \times \infty = \infty$ (provided that the infinites are both countable)

Can we find any sets that are bigger than N?

So far we have looked at N, Z, and Q and found them all to be the same size, even though N is a subset of Z which is a subset of Q. You might be beginning to think "Is that it? Are all infinities the same size?" In this section we will look at an set that is *bigger* than N. A set that *cannot* be put into a one to one correspondence with N no matter how it is arranged.

The set in question is R the real numbers. A real number is any number on the number line that is not in Q. Remember that the set Q contains all the numbers that can be written in the form p/q with p and q rational. Most numbers can never be put in this form. Examples of irrational numbers include Ï€,*e*, and $\sqrt{2}$.

The set R is huge! Much bigger than Q. To get a feel for the different sizes of these two infinite sets consider the decimal expansions of a real number and a rational number. Rational numbers always either terminate:

- 1/8 = 0.125

or repeat:

- 1/9 = 0.1111111......

Imagine measuring an object such as a book. If you use a ruler you might get 10cm. If you take a bit more care to and read the mm you might get 10.2cm. You'd then have to go on to more accurate measing devices such as vernier micrometers and find that you get 10.235cm. Going onto a travelling microscope you may find its 10.235823cm and so on. In general the decimal expansion of any *real* measurement will be a list of digits that look completely random.

Now imagine you measure a book and found it to be 10.101010101010cm. You'd be pretty surprised wouldn't you? But this is exactly the sort of result you would need to get if the book's length were rational. Rational numbers are dense (you find them all over the number line), infinite, yet much much rarer than real numbers.

How we can prove that R is bigger than Q

It's good to get a feel for the size of infinities as in the previous section. But to be really sure we have to come up with some form of proof. In order to prove that R is bigger than Q we use a classic method. We assume that R is the same size as Q and come up with a contradiction. For the sake of clarity we will restrict our proof to the real numbers between 0 and 1.We will call this set R' Clearly if we can prove that R' is bigger than Q then R must be bigger than Q also.

If R' was the same size as Q it would mean that it is countable. This means that we would be able to write out some form of list of all the members of R (This is what countable means, so far we have managed to write out all our infinite sets in the form of an infinitely long list). Let's consider this list.

R_1

R_2

R_3

R_4

.

.

.

Where R_1 is the first number in our list, R_2 is the second, and so on. Note that we haven't said what order the list is to be written. For this proof we don't need to say what the order of the list needs to be, only that the mermbers of R are listable (hence countable).

Now lets write out the decimal expansion of each of the numbers in the list.

$0.r_{11}r_{12}r_{13}r_{14}...$

$0.r_{21}r_{22}r_{23}r_{24}...$

$0.r_{31}r_{32}r_{33}r_{44}...$

$0.r_{41}r_{42}r_{43}r_{44}...$

.

.

.

Here r_{11} means the first digit after the decimal point of the first number in the list. So if our first number happened to be 0.36921... r_{11} would be 3, r_{12} would be 6 and so on. Remember that this list is meant to be complete. By that we mean that it contains *every* member of R'. What we are going to do in order to prove that R is not countable is to construct a number that is not already on the list. Since the list is supposed to contain *every* member of R', this will cause a contradiction and therfore show that R' is unlistable.

In order to construct this unlisted number we choose a decimal representation:

$0.a_1a_2a_3a_4...$

Where a_1 is the first digit after the point etc.

We let a_1 take any value from 0 - 9 inclusive *except* the digit r_{11}. So if $r_{11} = 3$ then a_1 can be 0, 1, 2, 4, 5, 6, 7, 8, or 9. Then we let a_2 be any digit except r_{22} (the second digit of the second number on the list). Then a_3 be any digit except r_{33} and so on.

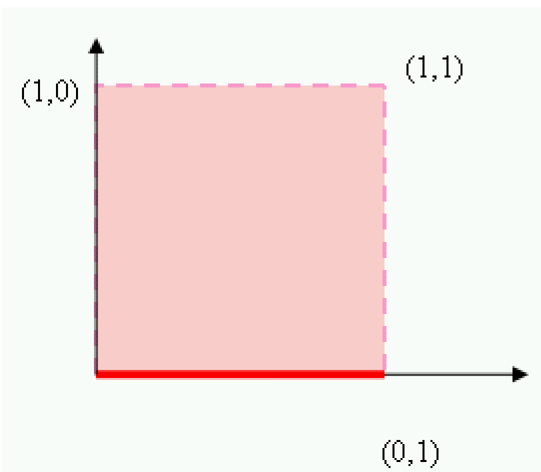
Now if this number, that we have just constructed *were* on the list somewhere then it would have to be equal to $R_{something}$. Let's see what $R_{something}$ it might be equal to. It can't be equal to R_1 because it has a different first digit (r_{11} and a_1. Nor can it be equal to R_2 because it has a different second digit, and so on. In fact it can't be equal to *any* number on the list becuse it differs by at least one digit from *all* of them.

We have done what we set out to do. We have constructed a number that is in R but is not on the "list of all members of R". This means that R is bigger than any list. It is not listable. It is not countable. It is a bigger infinity than Q.

Are there even bigger infinities?

There are but they are difficult to describe. The set of all the possible combinations of any number of real numbers is a bigger infinity than R. However trying to imagine such a set is mind boggling. Let's look instead at a set that looks like it should be bigger than R but turns out not to be.

Remember R', which we defined earlier on as the set of all numbers on the number line between 0 and 1. Let us now consider the set of all numbers in the plane from [0,0] to [1,1]. At first sight it would seem obvious that there must be more points on the whole plane than there are in a line. But in transfinite mathematics the "obvious" is not always true and proof is the only way to go. Cantor spent three years trying to prove it true but failed. His reason for failure was the best possible. It's false.



Each point in this plane is specified by two numbers, the x coordinate and the y coordinate; x and y both belong to R. Lets consider one point in the line. $0.a_1a_2a_3a_4$.... Can you think of a way of using this one number to specify a point in the plane ? Likewise can you think of a way of combining the two numbers $x= 0.x_1x_2x_3x_4$.... and $y= 0.y_1y_2y_3y_4$.... to specify a point on the line? (think about it before you read on)

One way is to do it is to take

$a_1 = x_1$

$a_2 = y_1$

$a_3 = x_2$

$a_4 = y_2$

.

.

.

This defines a one to one correspondence between the points in the plane and the points in the line. (Actually, for the sharp amongst you, not quite one to one. Can you spot the problem and how to cure it?)

Exercises

1. Prove that the number of points in a cube is the same as the number of points on one of its sides.

Continuum hypothesis

We shall end the section on infinite sets by looking at the Continuum hypothesis. This hypothesis states that there are no infinities between the natural numbers and the real numbers. Cantor came up with a number system for transfinite numbers. He called the smallest infinity $\aleph_0$ with the next biggest one $\aleph_1$ and so on. It is easy to prove that the cardinality of N is $\aleph_0$ (Write any smaller infinity out as a list. Either the list terminates, in which case it's finite, or it goes on forever, in which case it's the same size as N) but is the cardinality of the reals = $\aleph_1$?

To put it another way, the hypotheses states that:

There are no infinite sets larger than the set of natural numbers but smaller than the set of real numbers.

The hypothesis is interesting because it has been proved that "It is not possible to prove the hypothesis true or false, using the normal axioms of set theory"

Further reading

If you want to learn more about set theory or infinite sets try one of the many interesting pages on our sister project en:wikipedia.

- ordinal numbers
- Aleph numbers
- Set theory
- Hilbert's Hotel

Limits *Infinity got rid of*

The theory of infinite sets seems weird to us in the 21st century, but in Cantor's day it was downright unpalatable for most mathematicians. In those days the idea of infinity was too troublesome, they tried to avoid it wherever possible.

Unfortunately the mathematical topic called **analysis** was found to be highly useful in mathematics, physics, engineering. It was far too useful a field to simply drop yet analysis relies on infinity or at least infinite processes. To get around this problem the idea of a *limit* was invented.

Consider the series

$$\frac{1}{1}, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n} \ldots$$

This series is called the harmonic series.

Note that the terms of the series get smaller and smaller as you go further and further along the series. What happens if we let n become infinite? The term would become $\frac{1}{\infty}$

But this doesn't make sense. (Mathematicians consider it sloppy to divide by infinity. Infinity is not a normal number, you can't divide by it). A better way to think about it (The way you probably already do think about it, if you've ever considered the matter) is to take this approach: Infinity is very big, bigger than any number you care to think about. So let's let *n* become bigger and bigger and see if 1/*n* approaches some fixed number. In this case as *n* gets bigger and bigger 1/*n* gets smaller and smaller. So it is reasonable to say that the *limit* is 0.

In mathematics we write this as

$$\lim_{n \to \infty} \frac{1}{n} = 0$$

and it reads:

the limit of 1/*n* as *n* approaches infinity is zero

Note that we are not dividing 1 by infinity and getting the answer 0. We are letting the number *n* get bigger and bigger and so the reciprocal gets closer and closer to zero. Those 18th Century mathematicians loved this idea because it got rid of the pesky idea of *dividing by infinity*. At all times *n* remains finite. Of course, no matter how huge *n* is, 1/*n* will not be *exactly* equal to zero, there is always a small difference. This difference (or error) is usually denoted by Îµ (epsilon).

info -- infinitely small

When we talk about infinity, we think of it as something big. But there is also the infinitely small, denoted by Îµ (epsilon). This animal is closer to zero than any other number. Mathematicians also use the character Îµ to represent anything small. For example, the famous Hungarian mathematician Paul Erdos used to refer to small children as epsilons.

Examples

Lets look at the function

$$\frac{x^2+x}{x^2}$$

What is the limit as x approaches infinity ?

This is where the idea of limits really come into its own. Just replacing *x* with infinity gives us very little:

$$\frac{\infty^2+\infty}{\infty^2}=?$$

But by using limits we can solve it

$$\lim_{x\to\infty}\frac{x^2+x}{x^2}=1+\lim_{x\to\infty}\frac{1}{x}=1$$

For our second example consider this limit as x approaches infinity of x^3 - x^2

Again lets look at the *wrong* way to do it. Substituting $x=\infty$ into the expression gives $\infty^3-\infty^2$. Note that you cannot say that these two infinities just cancel out to give the answer zero.

Now lets look at doing it the *correct* way, using limits

$$\lim_{x\to\infty}x^3-x^2=\lim_{x\to\infty}x^2(x-1)=\infty$$

The last expression is two functions multiplied together. Both of these functions approach infinity as *x* approaches infinity, so the product is infinity also. This means that the *limit* does not exist, i.e. there is no finite number that the function approaches as *x* gets bigger and bigger.

One more just to get you really familer with how it works. Calculate:

$$\lim_{x\to\infty}\frac{\sin x}{x}$$

To make things very clear we shall rewrite it as

$$\lim_{x\to\infty}\frac{1}{x}(\sin x)$$

Now to calculate this limit we need to look at the properties of sin(x). Sin(x)is a function that you should already be familiar with (or you soon will be) its value oscillates between 1 and -1 depending on x. This means that the absolute value of sin(x) (the value ignoring the plus or minus sign) is always less than or equal to 1:

$$|x| \leq 1$$

So we have 1/x which we already know goes to zero as x goes to infinity multiplied by sin(x) which always remains finite no matter how big x gets. This gives us

$$\lim_{x\to\infty} \frac{1}{x}(\sin x) = 0$$

Exercises

Evaluate the following limits;

1. $\lim_{x\to\infty} \frac{3x^2-4}{2x^2+x}$
2. $\lim_{x\to\infty} \frac{x^2-1}{2x^3+3}$
3. $\lim_{x\to\infty} \frac{cosx}{x^2}$
4. $\lim_{x\to\infty} (2x^2 - x^4)$

Infinite series

Consider the infinite sum 1/1 + 1/2 + 1/4 + 1/8 + 1/16 + ... Do you think that this sum will equal infinity once all the terms have been added ? Let's sum the first few terms.

$$\begin{array}{lclcl} S_1 & = & \frac{1}{1} & = & 1 \\ S_2 & = & \frac{1}{1}+\frac{1}{2} & = & 1.5 \\ S_3 & = & \frac{1}{1}+\frac{1}{2}+\frac{1}{4} & = & 1.75 \\ S_4 & = & \frac{1}{1}+\frac{1}{2}+\frac{1}{4}+\frac{1}{8} & = & 1.825 \end{array}$$

Can you guess what S_∞ is ?

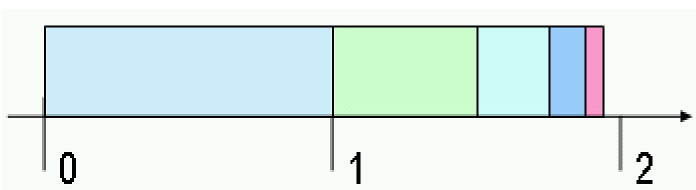
Here is another way of looking at it. Imagine a point on a number line moving along as the sum progresses. In the first term the point jumps to the position 1. This is half way from 0 to 2. In the second stage the point jumps to position 1.5 - half way from 1 to 2. At each stage in the process (shown in a different colour on the diagram) the distance to 2 is halved. The point can get as close to the point 2 as you like. You just need to do the appropriate number of jumps, but the point will never actually reach 2 in a finite number of steps. We say that in the limit as n approaches infinity, S_n approaches 2.

Zeno's Paradox

The ancient Greeks had a big problem with summing infinite series. A famous paradox from the philosopher Zeno is as follows:

In the paradox of Achilles and the tortoise, we imagine the Greek hero Achilles in a footrace with the plodding reptile. Because he is so fast a runner, Achilles graciously allows the tortoise a head start of a hundred feet. If we suppose that each racer starts running at some constant speed (one very fast and one very slow), then after some finite time, Achilles will have run a hundred feet, bringing him to the tortoise's starting point.

During this time, the tortoise has "run" a (much shorter) distance, say one foot. It will then take Achilles some further period of time to run that distance, during which the tortoise will advance farther; and then another period of time to reach this third point, while the tortoise moves ahead. Thus, whenever Achilles reaches somewhere the tortoise has been, he still has farther to go. Therefore, Zeno says, swift Achilles can never overtake the tortoise.



Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Exercises

Infinity and infinite processes

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

These solutions were not written by the author of the rest of the book. They are simply the answers I thought were correct while doing the exercises. I hope these answers are usefull for someone and that people will correct my work if I made some mistakes

How big is infinity? exercises

1. The number of even numbers is the same as the number of natural numbers because both are countably infinite. You can clearly see the one to one correspondence. (E means even numbers and is not an official set like N)

E	**N**
2	1
4	2
6	3
8	4

2. The number of square numbers is also equal to the number of natural numbers. They are both countably infinite and can be put in one to one correspondence. (S means square numbers and is not an official set like N)

S	**N**
1	1
4	2
9	3
16	4

3. The cardinality of even numbers less than 100 is not equal to the cardinality of natural numbers less than 100. You can simply write out both of them and count the numbers. Then you will see that cardinality of even numbers less than 100 is 49 and the cardinality of natural numbers less than 100 is 99. Thus the set of natural numbers less than 100 is bigger than the set

of even numbers less than 100. The big difference between infinite and finite sets thus is that a finite set can not be put into one to one correspondence with any of it's subsets. While an infinite set can be put into one to one correspondence with at least one of it's subsets.
4. Each part of the sum is answered below

infinity + 1 = infinity

You can prove this by taking a set with a cardinality of 1, for example a set consisting only of the number 0. You simply add this set in front of the countably infinite set to put the infinite set and the inifinite+1 set into one to one correspondence.

N	N+1
1	0
2	1
3	2
4	3

infinity + A = infinity (where A is a finite set)

You simply add the finite set in front of the infinite set like above, only the finite set doesn't need to have a cardinality of one anymore.

infinity + C = infinity (where C is a countably infinite set)

You take one item of each set (infinity or C) in turns, this will make the new list also countably infinite.

Is the set of rational numbers bigger than N? exercises

1. To change the matrix from Q' to Q the first step you need to take is to remove the multiple entries for the same number. You can do this by leaving an empty space in the table when gcd(topnr,bottomnr) ≠ 1 because when the gcd isn't 1 the fraction can be simplified by dividing the top and bottom number by the gcd. This will leave you with the following table.

$$\begin{array}{cccc} \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \cdots \\ \frac{2}{1} & & \frac{2}{3} & \cdots \\ \frac{3}{1} & \frac{3}{2} & & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{array}$$

Now we only need to add zero to the matrix and we're finished. So we add a vertical row for zero and only write the topmost element in it (0/1) (taking gcd doesn't work here because gcd(0,a)=a) This leaves us with the following table where we have to count all fractions in the diagonal rows to see that Q is countably infinite.

$$\begin{array}{ccccc} \frac{0}{1} & \frac{1}{1} & \frac{1}{2} & \frac{1}{3} & \cdots \\ & \frac{2}{1} & & \frac{2}{3} & \cdots \\ & \frac{3}{1} & \frac{3}{2} & & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{array}$$

2. To show that $\infty \times \infty = \infty$ you have to make a table where you put one infinity in the horizontal row and one infinity in the vertical row. Now you can start counting the number of place in the table diagonally just like Q' was counted. This works because a table of size AxB contains A*B places.

Are there even bigger infinities? exercises

1. You have to use a method to map the coordinates in a plain onto a point on the line and the other way around, like the one described in the text. This method shows you that for every number on the line there is a place on the plain and for every place on the plain there is a place on the line. Thus the number of points on the line and the plain are the same. Limits Infinity got rid of exercises

1. $\lim_{x\to\infty} \frac{3x^2-4}{2x^2+x} = \lim_{x\to\infty} (\frac{3x^2}{2x^2+x} - \frac{4}{2x^2+x}) = \lim_{x\to\infty} \frac{3x^2}{2x^2+x} = \frac{3}{2}$
2. $\lim_{x\to\infty} \frac{x^2-1}{2x^3+3} = 0$
3. $\lim_{x\to\infty} \frac{cosx}{x^2} = \lim_{x\to\infty} \frac{1}{x^2} cosx = 0$
4. $\lim_{x\to\infty} (2x^2 - x^4) = \lim_{x\to\infty} x^2 \times (2 - x^2) = -\infty$

Problem set

HSE PS Infinity and infinite processes

Counting and Generating functions

Before we begin: This chapter assumes knowledge of

11. Ordered selection (permutation) and unordered selection (combination) covered in Basic counting,
12. Method of Partial Fractions and,
13. Competence in manipulating Summation Signs

Some Counting Problems

..more to come

Generating functions

..some motivation to be written

Generating functions, otherwise known as Formal Power Series, are useful for solving problems like:

$x_1 + x_2 + 2x_3 = m$

where

$x_n \geq 0$; n = 1, 2, 3

how many unique solutions are there if $m = 55$?

Before we tackle that problem, let's consider the infinite polynomial:

$S = 1 + x + x^2 + x^3 + \ldots + x^n + x^{n+1}\ldots$

We want to obtain a *closed form* of this infinite polynomial. The *closed form* is simply a way of expressing the polynomial so that it involves only a finite number of operations.

$$\begin{array}{ccll} S & = & 1+ & x + x^2 + x^3 + \ldots \\ xS & = & & x + x^2 + x^3 + \ldots \end{array}$$

$$\begin{array}{ccl} (1-x)S & = & 1 \\ S & = & \frac{1}{1-x} \end{array}$$

So the closed form of

$1 + x + x^2 + x^3 + \ldots$

is

$$\frac{1}{1-x}$$

We can equate them (actually, we can't. Refer to info).

$$1+x+x^2+x^3+...=\frac{1}{1-x}\ ;\ -1<x<1$$

info - Infinite sums

The two expressions are not *equal*. It's just that for certain values of *x* ($-1 < x < 1$), we can approximate the right hand side *as closely as possible* by adding up a large number of terms on the left hand side. For example, suppose $x = 1/2$, RHS = 2; we approximate the LHS using only 5 terms we get LHS equals 1 + 1/2 + 1/4 + 1/8 + 1/16 = 1.9375, which is close to 2, as you can imagine by adding more and more terms, we will get closer and closer to 2. For a more detailed discussion of the above, head to Infinity and infinite processes.

Anyway we really only care about its nice algebraic properties not its numerical value. From now on we will omit the condition for equality to be true when writing out generating functions.

Consider a more general case:

$S = A + ABx + AB^2x^2 + AB^3x^3 + ...$

where *A* and *B* are constants.

We can derive the *closed-form* as follows:

$$\begin{array}{rcll} S & = & A+ & ABx + AB^2x^2 + AB^3x^3 + ... \\ BxS & = & & ABx + AB^2x^2 + AB^3x^3 + ... \end{array}$$

$$\begin{array}{rcl} (1-Bx)S & = & A \\ S & = & \frac{A}{1-Bx} \end{array}$$

The following identity as derived above is worth investing time and effort memorising.

$$A+ABx+AB^2x^2+AB^3x^3+...=\frac{A}{1-Bx}$$

Exercises

1. Find the closed-form:

(a)$1 - z + z^2 - z^3 + z^4 - z^5 + ...$

(b)$1 + 2z + 4z^2 + 8z^3 + 16z^4 + 32z^5 + ...$

(c)$z + z^2 + z^3 + z^4 + z^5 + ...$

(d)$3 - 4z + 4z^2 - 4z^3 + 4z^4 - 4z^5 + ...$

(e)$1 - z^2 + z^4 - z^6 + z^8 - z^{10} + ...$

2. Given the closed-form, find a function f(n) for the coefficients of x^n:

(a) $\frac{1}{1+z}$ (Hint: note the plus sign in the denominator)

(b) $\frac{z^3}{1-z^2}$ (Hint: obtain the generating function for 1/(1 - z^2) first, then multiply by the appropriate expression)

(c) $\frac{z^2-1}{1+3z^3}$ (Hint: break into the sum of two distinct close forms)

Method of Substitution

We are given that:

$1 + z + z^2 + ... = 1/(1 - z)$

and we can obtain many other generating functions by substitution. For example: letting $z = x^2$ we have:

$1 + x^2 + x^4 + ... = 1/(1 - x^2)$

Similarly

$A + ABx + A(Bx)^2 + ... = A/(1 - Bx)$

is obtained by letting z = Bx then multiplying the whole expression by A.

Exercises

1. What are the coefficients of the powers of x:

$1/(1 - 2x^3)$

2. What are the coefficients of the powers of x (Hint: take out a factor of 1/2):

$1/(2 - x)$

Linear Recurrence Relations

The Fibonacci series

1, 1, 2, 3, 5, 8, 13, 21, 34, 55...

where each and every number, except the first two, is the sum of the two preceding numbers. We say the numbers are *related* if the value a number takes depends on the values that come before it in the sequence. The Fibonacci sequence is an example of a recurrence relation, it is expressed as:

$$\begin{aligned} x_n &= x_{n-1} + x_{n-2}; \text{ for n} \geq 2 \\ x_0 &= 1 \\ x_1 &= 1 \end{aligned}$$

where x_n is the $(n+1)$th number in the sequence. Note that the first number in the sequence is denoted x_0. Given this recurrence relation, the question we want to ask is "can we find a formula for the $(n+1)$th number in the sequence?". The answer is yes, but before we come to that, let's look at some examples.

Example 1

The expressions

$$\begin{aligned} x_n &= 2x_{n-1} + 3x_{n-2}; \text{ for n} \geq 2 \\ x_0 &= 1 \\ x_1 &= 1 \end{aligned}$$

define a recurrence relation. The sequence is: 1, 1, 5, 13, 41, 121, 365... Find a formula for the $(n+1)$th number in the sequence.

Solution Let G(z) be generating function of the sequence, meaning the coefficient of each power (in ascending order) is the corresponding number in the sequence. So the generating functions looks like this

$G(z) = 1 + z + 5z^2 + 13z^3 + 41z^4 + 121z^5 + \ldots$

Now, by a series of algebraic manipulations, we can find the closed form of the generating function and from that the formula for each coefficient

$$\begin{array}{llllllll} & G(z) & = & x_0 + & x_1 z + & x_2 z^2 & + & x_3 z^3 + x_4 z^4 + x_5 z^5 + \ldots \\ 2z & \times G(z) & = & & 2x_0 z + & 2x_1 z^2 & + & \ldots \\ 3z^2 & \times G(z) & = & & & 3x_0 z^2 & + & \ldots \end{array}$$

$$
\begin{aligned}
G(z) - 2zG(z) - 3z^2G(z) \quad = \quad & x_0 + (x_1 - 2x_0)z \quad + \\
& (x_2 - 2x_1 - 3x_0)z^2 \quad + \\
& (x_3 - 2x_2 - 3x_1)z^3 \quad + \quad ...
\end{aligned}
$$

by definition $x_n - 2x_{n-1} - 3x_{n-2} = 0$

$$
\begin{aligned}
(1 - 2z - 3z^2) \quad \times \quad G(z) \quad &= \quad x_0 + (x_1 - 2x_0)z \\
G(z) \quad &= \quad \tfrac{1-z}{1-2z-3z^2} \\
G(z) \quad &= \quad \tfrac{1-z}{(1-3z)(1+z)}
\end{aligned}
$$

by the method of partial fractions we get:

$$G(z) = \frac{1}{2} \times \frac{1}{1-3z} + \frac{1}{2} \times \frac{1}{1+z}$$

each part of the sum is in a recognisable closed-form. We can conclude that:

$$x_n = \frac{1}{2} \times 3^n + \frac{1}{2} \times (-1)^n$$

the reader can easily check the accuracy of the formula.

Example 2

$$
\begin{aligned}
x_n \quad &= \quad x_{n-1} \quad + \quad x_{n-2} \quad - \quad x_{n-3};\ \text{for n} \geq 3 \\
x_0 \quad &= \quad 1 \\
x_1 \quad &= \quad 1 \\
x_2 \quad &= \quad 1
\end{aligned}
$$

Find a non-recurrent formula for x_n.

Solution Let G(z) be the generating function of the sequence described above.

$G(z) = x_0 + x_1z + x_2z^2 + ...$

$$
\begin{aligned}
G(z)(1 - z - z^2 + z^3) \quad &= \quad x_0 \quad + \quad (x_1 - x_0)z + (x_2 - x_1 - x_0)z^2 \\
G(z)(1 - z - z^2 + z^3) \quad &= \quad 1 - z^2
\end{aligned}
$$

$$G(z) = \frac{1-z^2}{1-z-z^2+z^3}$$

$$G(z) = \frac{1-z}{1-2z+z^2}$$

$$G(z) = \frac{1}{1-z}$$

Therefore $x_n = 1$, for all n.

Example 3

A linear recurrence relation is defined by:

$$\begin{aligned} x_n &= x_{n-1} + 6x_{n-2} + 1; \text{ for n} \geq 2 \\ x_0 &= 1 \\ x_1 &= 1 \end{aligned}$$

Find the general formula for x_n.

Solution Let G(z) be the generating function of the recurrence relation.

$G(z)(1 - z - 6z^2) = x_0 + (x_1 - x_0)z + (x_2 - x_1 - 6x_0)z^2 + \ldots$

$G(z)(1 - z - 6z^2) = 1 + z^2 + z^3 + z^4 + \ldots$

$G(z)(1 - z - 6z^2) = 1 + z^2(1 + z + z^2 + \ldots)$

$$G(z)(1 - z - 6z^2) = 1 + \frac{z^2}{1-z}$$

$$G(z)(1 - z - 6z^2) = \frac{1 - z + z^2}{1-z}$$

$$\begin{aligned} G(z) &= \frac{1-z+z^2}{(1-z)(1+2z)(1-3z)} \\ G(z) &= -\frac{1}{6(1-z)} + \frac{7}{15(1+2z)} + \frac{7}{10(1-3z)} \end{aligned}$$

Therefore

$$x_n = -\frac{1}{6} + \frac{7}{15}(-2)^n + \frac{7}{10}3^n$$

Exercises

1. Derive the formula for the $(n+1)$th numbers in the sequence defined by the linear recurrence relations:

$$\begin{aligned} x_n &= 2x_{n-1} - 1; \text{ for n} \geq 1 \\ x_0 &= 1 \end{aligned}$$

2. Derive the formula for the (*n+1*)th numbers in the sequence defined by the linear recurrence relations:

$$\begin{aligned} 3x_n &= -4x_{n-1} + x_{n-2}; \text{ for n} \geq 2 \\ x_0 &= 1 \\ x_1 &= 1 \end{aligned}$$

3. (Optional) Derive the formula for the (*n+1*)th Fibonacci numbers.

Further Counting

Consider the equation

a + b = n; a, b $\geq$ 0 are integers

For a fixed positive integer n, how many solutions are there? We can count the number of solutions:

0 + n = n

1 + (n - 1) = n

2 + (n - 2) = n

...

n + 0 = n

As you can see there are (n + 1) solutions. Another way to solve the problem is to consider the generating function

$G(z) = 1 + z + z^2 + ... + z^n$

Let H(z) = G(z)G(z), i.e.

$H(z) = (1 + z + z^2 + ... + z^n)^2$

I claim that the coefficient of z^n in H(z) is the number of solutions to a + b = n, a, b > 0. The reason why lies in the fact that *when multiplying like terms, indices add*.

Consider

$A(z) = 1 + z + z^2 + z^3 + ...$

Let

$B(z) = A^2(z)$

it follows

$B(z) = (1 + z + z^2 + z^3 + ...) + z(1 + z + z^2 + z^3 + ...) + z^2(1 + z + z^2 + z^3 + ...) + z^3(1 + z + z^2 + z^3) + ...$

$B(z) = 1 + 2z + 3z^2 + ...$

Now the coefficient of z^n (for $n \geq 0$) is clearly the number of solutions to $a + b = n$ ($a, b > 0$).

We are ready now to derive a very important result: let t_k be the number solutions to $a + b = n$ ($a, b > 0$). Then the generating function for the sequence t_k is

$T(z) = (1 + z + z^2 + ... + z^n + ...)(1 + z + z^2 + ... + z^n + ...)$

$$T(z) = \frac{1}{(1-z)^2}$$

i.e.

$$\frac{1}{(1-z)^2} = 1 + 2z + 3z^2 + 4z^3 + ... + (n+1)z^n + ...$$

Counting Solutions to $a_1 + a_2 + ... + a_m = n$

Consider the number of solutions to the following equation:

$a_1 + a_2 + ... + a_m = n$

where $a_i \geq 0$; $i = 1, 2, ... m$. By applying the method discussed previously. If t_k is the number of solutions to the above equation when $n = k$. The generating function for t_k is

$$T(z) = \frac{1}{(1-z)^m}$$

but what is t_k? Unless you have learnt calculus, it's hard to derive a formula just by looking the equation of $T(z)$. Without assuming knowledge of calculus, we consider the following counting problem.

"You have three sisters, and you have n ($n \geq 3$) lollies. You decide to give each of your sisters at least one lolly. In how many ways can this be done?"

One way to solve the problem is to put all the lollies on the table in a straightline. Since there are *n* lollies there are (*n* - 1) gaps between them (just as you have 5 fingers on each hand and 4 gaps between them). Now get 2 dividers, from the (*n* - 1) gaps available, *choose* 2 and put a divider in each of the gaps you have chosen! There you have it, you have divided the *n* lollies into three parts, one for each sister. There are $\binom{n-1}{2}$ ways to do it! If you have 4 sisters, then there are $\binom{n-1}{3}$ ways to do it. If you have m sisters there are $\binom{n-1}{m-1}$ ways to do it.

Now consider: "You have three sisters, and you have n lollies. You decide to give each of your sisters some lollies (with no restriction as to how much you give to each sister). In how many ways can this be done?"

Please note that you are just solving:

$a_1 + a_2 + a_3 = n$

where $a_i \geq 0$; i = 1, 2, 3.

You can solve the problem by putting n + 3 lollies on the table in a straightline. Get two dividers and *choose* 2 gaps from the n + 2 gaps available. Now that you have divided n + 3 lollies into 3 parts, with each part having 1 or more lollies. Now take back 1 lollies from each part, and you have solved the problem! So the number of solutions is $\binom{n+2}{2}$. More generally, if you have m sisters and n lollies the number of ways to share the lollies is

$$\binom{n+m-1}{m-1} = \binom{n+m-1}{n}.$$

Now to the important result, as discussed above the number of solutions to

$a_1 + a_2 + ... + a_m = n$

where $a_i \geq 0$; i = 1, 2, 3 ... m is

$$\binom{n+m-1}{n}$$

i.e.

$$\frac{1}{(1-z)^m} = \sum_{i=0}^{\infty} \binom{m+i-1}{i} z^i$$

Example 1

The closed form of a generating function T(z) is

$$T(z) = \frac{z}{(1-z)^2}$$

and t_k in the coefficient of z^k is T(z). Find an explicit formula for t_k.

Solution

$$\frac{1}{(1-z)^2} = \sum_{i=0}^{\infty}(i+1)z^i$$

$$\frac{z}{(1-z)^2} = z\sum_{i=0}^{\infty}(i+1)z^i$$

$$= \sum_{i=0}^{\infty}(i+1)z^{i+1}$$

Therefore $t_k = k$

Example 2

Find the number of solutions to:

a + b + c + d = n

for all positive integers n (including zero) with the restriction a, b, c ,d ≥ 0.

Solution By the formula

$$\frac{1}{(1-z)^4} = \sum_{i=0}^{\infty}\binom{n+3}{3}z^i$$

so

the number of solutions is $\binom{n+3}{3}$

More Counting

We turn to a sligthly harder problem of the same kind. Suppose we are to count the number of solutions to:

$2a + 3b + c = n$

for some integer $n \geq 0$, with *a*, *b*, also *c* greater than or equal zero. We can write down the closed form straight away, we note the coefficient of x^n of:

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

is the required solution. This is due to, again, the fact that when multiplying powers, indices add.

To obtain the number of solutions, we break the expression into recognisable closed-forms by method of partial fraction.

Example 1

Let s_k be the number of solutions to the following equation:

$2a + 2b = n;\ a, b \geq 0$

Find the generating function for s_k, then find an explicit formula for s_n in terms of *n*.

Solution

Let T(z) be the generating functions of t_k

$T(z) = (1 + z^2 + z^4 + ... + z^{2n} + ...)^2$

$$T(z) = \frac{1}{(1 - z^2)^2}$$

It's not hard to see that

$$s_n = 0 \text{ if n is odd}$$

$$s_n = \binom{n/2 + 1}{n/2} = \binom{n/2 + 1}{1} = n/2 + 1 \text{ if n is even}$$

Example 2

Let t_k be the number of solutions to the following equation:

$a + 2b = n;\ a, b \geq 0$

Find the generating function for t_k, then find an explicit formula for t_n in terms of *n*.

Solution

Let T(z) be the generating functions of t_k

$T(z) = (1 + z + z^2 + ... + z^n + ...)(1 + z^2 + z^4 + ... + z^{2n} + ...)$

$$T(z) = \frac{1}{(1 - z)} \times \frac{1}{1 - z^2}$$

$$T(z) = \frac{1}{(1-z)^2} \times \frac{1}{1+z}$$

$$T(z) = \frac{Az+B}{(1-z)^2} + \frac{C}{1+z}$$

A = -1/4, B = 3/4, C = 1/4

$$T(z) = -\frac{1}{4}\sum_{i=0}^{\infty}(i+1)z^{i+1} + \frac{3}{4}\sum_{i=0}^{\infty}(i+1)z^{i} + \frac{1}{4}\sum_{i=0}^{\infty}(-1)^i z^i$$

$$t_k = -\frac{1}{4}k + \frac{3}{4}(k+1) + \frac{1}{4}(-1)^k$$

Exercises

1. Let

$$T(z) = \frac{1}{(1+z)^2}$$

be the generating functions for t_k (k = 0, 1, 2 ...). Find an explicit formula for t_k in terms of *k*.

2. How many solutions are there the following equations if *m* is a given constant

$a + b + 2c = m$

where a, b and c ≥ 0

Problem Set

1. A new Company has borrowed $250,000 initial capital. The monthly interest is 3%. The company plans to repay $*x* before the end of each month. Interest is added to the debt on the last day of the month (compounded monthly).

Let D_n be the remaining debt after *n* months.

a) Define D_n recursively.

b) Find the minimum values of *x*.

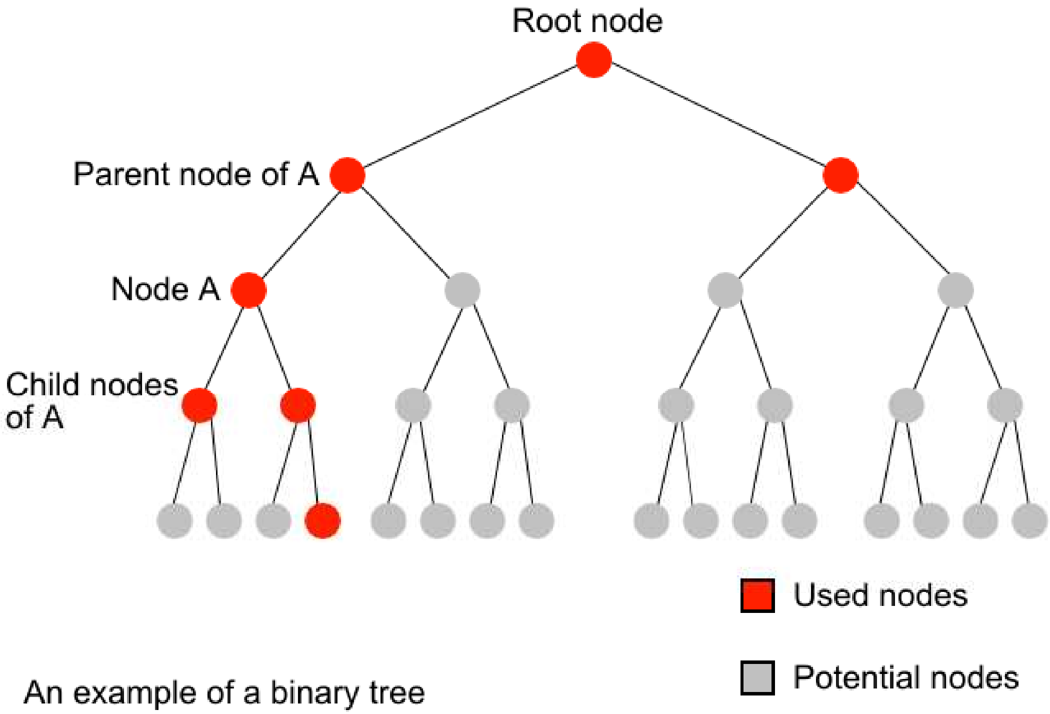
c) Find out the general formula for D_n.

d) Hence, determine how many months are need to repay the debt if *x* = 12,000.

2. A partion of *n* is a sequence of positive integers (λ1,λ1,..,λr) such that λ1 ≥ λ2 ≥ .. ≥ λr and

$\lambda 1 + \lambda 2 + .. + \lambda r = n$. For example, let $n = 5$, then (5), (4,1), (3,2), (3,1,1), (2,2,1), (2,1,1,1), (1,1,1,1,1) are all the partions of 5. So we say the number of partions of 5 is 7. Derive a formula for the number of partions of a general *n*.

3. A binary tree is a tree where each node can have up to two child nodes. The figure below is an example of a binary tree.



An example of a binary tree

a) Let c_n be the number of unique arrangements of a binary tree with totally n nodes. Let C(z) be a generating function of c_n.

(i) Define C(z) using recursion.

(ii) Hence find the closed form of C(z).

b) Let $P(x) = \sqrt{1+ax} = p_0 + p_1 x + p_2 x^2 + p_3 x^3 \ldots$ be a power series.

(i) By considering the n-th derivative of P(x), find a formula for p_n.

(ii) Using results from a) and b)(i) , or otherwise, derive a formula for c_n.

Hint: Instead of doing recursion of finding the change in c_n when adding nodes at the buttom, try to think in the opposite way, and direction.(And no, not deleting nodes)

Project - Exponential generating function

This project assumes knowledge of differentiation.

(Optional)0.

(a)

(i) Differentiate log x by first principle.

(ii)*** Show that the remaining limit in last part that can't be evaluated indeed converges. Hence finish the differentiation by assigning this number as a constant.

(b) Hence differentiate a^x.

1. Consider $E(x) = e^x$

(a) Find out the n-th derivative of E(x).

(b) By considering the value of the n-th derivative of E(x) at x = 0, express E(x) in power series/infinite polynomial form.

(Optional)2.

(a) Find out the condition for the geometric progression(that is the ordinary generating function introduced at the begining of this chapter) to converges. (Hint: Find out the partial sum)

(b) Hence show that E(x) in the last question converges for all real values of x. (Hint: For any fixed x, the numerator of the general term is exponential, while the denominator of the general term is factorial. Then what?)

3. The function E(x) is the most fundamental and important exponential generating function, it is similar to the ordinary generating function, but with some difference, most obviously having a fractorial fraction attached to each term.

(a) Similar to ordinary generating function, each term of the polynomial expansion of E(x) can have number attached to it as coefficient. Now consider

$$A(x) = a_1 + a_2\frac{x}{1!} + a_3\frac{x^2}{2!} + a_4\frac{x^3}{3!} + ...$$

Find $A'(x)$ and compare it with A(x). What do you discover?

(b) Substitute nz, where n is a real number and z is a free variable, into E(x), i.e. E(nz). What have you found?

4. Apart from A(x) defined in question 2, let $B(x) = b_1 + b_2\frac{x}{1!} + b_3\frac{x^2}{2!} + b_4\frac{x^3}{3!} + ...$

(a) What is A(x) multiplied by B(x)? Compare this with ordinary generating function, what is the difference?

(b) What if we blindly multiply A(x) with x(or x^n in general)? Will it shift coefficient like what happened in ordinary generating function?

Notes: Question with *** are difficult questions, although you're not expected to be able to answer those, feel free to try your best.

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Exercises

Counting and Generating functions

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

These solutions were not written by the author of the rest of the book. They are simply the answers I thought were correct while doing the exercises. I hope these answers are useful for someone and that people will correct my work if I made some mistakes

Generating functions exercises

1.

(a)$S = 1 - z + z^2 - z^3 + z^4 - z^5 + ...$

$zS = z - z^2 + z^3 - z^4 + z^5 - ...$

$(1 + z)S = 1$

$$S = \frac{1}{1+z}$$

(b)$S = 1 + 2z + 4z^2 + 8z^3 + 16z^4 + 32z^5 + ...$

$2zS = 2z + 4z^2 + 8z^3 + 16z^4 + 32z^5 + ...$

$(1 - 2z)S = 1$

$$S = \frac{1}{1-2z}$$

(c)$S = z + z^2 + z^3 + z^4 + z^5 + ...$

$zS = z^2 + z^3 + z^4 + z^5 + ...$

$(1 - z)S = z$

$$S = \frac{z}{1-z}$$

(d)$S = 3 - 4z + 4z^2 - 4z^3 + 4z^4 - 4z^5 + ...$

$z(S + 1) = 4z - 4z^2 + 4z^3 - 4z^4 + 4z^5 - ...$

$S + z(S + 1) = 3$

$S + zS + z = 3$

$(1 + z)S = 3 - z$

$$S = \frac{3 - z}{1 + z}$$

2.

(a) $S = \dfrac{1}{1 + z}$

$$S = \frac{1}{1 - -z}$$

$S = 1 - x + x^2 - x^3 + x^4 - x^5 + \dots$

$f(n) = (- 1)^n$

(b) $S = \dfrac{z^3}{1 - z^2}$

$(1 - z^2)S = z^3$

$S = z^3 + z^5 + z^7 + z^9 + \dots$

$$f(n) = 1; \text{for n} \geq 2 \text{ and even}$$

$f(n) = 0$;for n is odd

2c only contains the exercise and not the answer for the moment

(c) $\dfrac{z^2 - 1}{1 + 3z^3}$

Linear Recurrence Relations exercises

This section only contains the incomplete answers because I couldn't figure out where to go from here.

1.

$$\begin{aligned} x_n &= 2x_{n-1} - 1;\ \text{for n} \geq 1 \\ x_0 &= 1 \end{aligned}$$

Let G(z) be the generating function of the sequence described above.

$G(z) = x_0 + x_1z + x_2z^2 + ...$

$(1 - 2z)G(z) = x_0 + (x_1 - 2x_0)z + (x_2 - 2x_1)z^2 + ...$

$(1 - 2z)G(z) = 1 - z - z^2 - z^3 - z^4 - ...$

$(1 - 2z)G(z) = 1 - z(1 + z + z^2 + ...)$

$$(1-2z)G(z) = 1 - \frac{z}{1-z}$$

$$(1-2z)G(z) = \frac{1-2z}{1-z}$$

$$G(z) = \frac{1}{1-z}$$

$x_n = 1$

2.

$$\begin{array}{rcl} 3x_n & = & -4x_{n-1} + x_{n-2};\ \text{for n} \geq 2 \\ x_0 & = & 1 \\ x_1 & = & 1 \end{array}$$

Let G(z) be the generating function of the sequence described above.

$G(z) = x_0 + x_1z + x_2z^2 + ...$

$(3 + 4z - z^2)G(z) = 3x_0 + (3x_1 + 4x_0)z + (3x_2 + 4x_1 - x_0)z^2 + (3x_3 + 4x_2 - x_1)z^3 + ...$

$(3 + 4z - z^2)G(z) = 3x_0 + (3x_1 + 4x_0)z$

$(3 + 4z - z^2)G(z) = 3 + 7z$

$$G(z) = \frac{3+7z}{-z^2+4z+3}$$

3. Let G(z) be the generating function of the sequence described above.

$G(z) = x_0 + x_1z + x_2z^2 + ...$

$(1 - z - z^2)G(z) = x_0 + (x_1 - x_0)z + (x_2 - x_1 - x_0)z^2 + (x_3 - x_2 - x_1)z^2 + ...$

$(1 - z - z^2)G(z) = 1$

$$G(z) = \frac{1}{1 - z - z^2}$$

$$G(z) = \frac{-1}{z^2 + z - 1}$$

We want to factorize $f(z) = z^2 + z - 1$ into $(z - \alpha)(z - \beta)$, by the converse of factor theorem, if $(z - p)$ is a factor of f(z), f(p)=0.

Hence α and β are the roots of the quadratic equation $z^2 + z - 1 = 0$

Using the quadratic formula to find the roots:

$$\alpha = \frac{\sqrt{5} - 1}{2}, \beta = -\frac{\sqrt{5} + 1}{2}$$

In fact, these two numbers are the faomus golden ratio and to make things simple, we use the greek symbols for golden ratio from now on.

Note: $\frac{\sqrt{5} - 1}{2}$ is denoted φ and $\frac{\sqrt{5} + 1}{2}$ is denoted Φ

$$G(z) = \frac{-1}{(z - \phi)(z + \Phi)}$$

By the method of partial fraction:

$$G(z) = \frac{1}{\sqrt{5}(z + \Phi)} - \frac{1}{\sqrt{5}(z - \phi)}$$

$$G(z) = \frac{1}{\Phi\sqrt{5}(\frac{z}{\Phi} + 1)} - \frac{1}{\phi\sqrt{5}(\frac{z}{\phi} - 1)}$$

$$G(z) = \frac{1}{\Phi\sqrt{5}(1 - -\phi z)} + \frac{1}{\phi\sqrt{5}(1 - \Phi z)}$$

$$x_n = \frac{\phi}{\sqrt{5}} \times (-\phi)^n + \frac{\Phi}{\sqrt{5}} \times \Phi^n$$

$$x_n = \frac{\Phi^{n+1} - (-\phi)^{n+1}}{\sqrt{5}}$$

Further Counting exercises

1. We know that

$$T(z) = \frac{1}{(1-z)^2} = \sum_{i=0}^{\infty} \binom{i+1}{i} z^i = \sum_{i=0}^{\infty} (i+1) z^i$$

therefore

$$T(z) = \frac{1}{(1+z)^2} = \sum_{i=0}^{\infty} (i+1)(-1)^i z^i$$

Thus

$T_k = (-1)^k(k + 1)$

2. $a + b + c = m$

$$T(z) = \frac{1}{(1-z)^3} = \sum_{i=0}^{\infty} \binom{i+2}{i} z^i$$

Thus

$$T_k = \binom{i+2}{i}$$

Differentiate from first principle exercises

1.

$$f'(z) = \lim_{h \to 0} \frac{1}{(1-(z+h))^2} - \frac{1}{(1-z)^2} =$$

$$\lim_{h \to 0} \frac{1}{h} \frac{(1-z)^2 - (1-(z+h))^2}{(1-z-h)^2(1-z)^2} =$$

$$\lim_{h \to 0} \frac{1}{h} \frac{z^2 - 2z + 1 - (z+h)^2 + 2(z+h) - 1}{(1-z-h)^2(1-z)^2} =$$

$$\lim_{h \to 0} \frac{1}{h} \frac{z^2 - 2z + 1 - z^2 - h^2 - 2zh + 2z + 2h - 1}{(1-z-h)^2(1-z)^2} =$$

$$\lim_{h\to 0} \frac{1}{h}\frac{-h^2 - 2zh + 2h}{(1-z-h)^2(1-z)^2} =$$

$$\lim_{h\to 0} \frac{-h - 2z + 2}{(1-z-h)^2(1-z)^2} =$$

$$\frac{-2z+2}{(1-z)^4} =$$

$$\frac{-2}{(1-z)^3}$$

Discrete Probability

Introduction

Probability theory is one of the most widely applicable mathematical theories. It deals with uncertainty and teaches you how to manage it. It is simply one of the most useful theories you will ever learn.

Please do not misunderstand. We are not learning to predict things, rather we learn to utilise predicted *chances* and make them useful. Therefore, we don't care, *what is the probability it will rain tomorrow?*, but given the probability is 60% we can make deductions, the easiest of which is *the probability it will not rain tomorrow is 40%*.

As suggested above, a *probability* is a percentage and it's between 0% and 100% (inclusive). Mathematicians like to express a *probability* as a *proportion* i.e. as a number between 0 and 1.

info - Why discrete?

Probability comes in two flavours, discrete and continuous. The continuous case is considered to be far more difficult to understand, and much less intuitive, than discrete probability and it requires knowledge of calculus. But we will touch on a little bit of the continuous case later on in the chapter.

Event and Probability

Roughly, an *event* is something we can assign a *probability* to. For example *the probability it will rain tomorrow is 0.6*, in here the event is *it will rain tomorrow* the assigned probability is 0.6. We can write

P(it will rain tomorrow) = 0.6

as mathematicians like to do we can use abstract letters to represent events. In this case we choose *A* to represent the event *it will rain tomorrow*, so the above expression can be written as

$P(A) = 0.6$

Another example *a fair die will turn up 1, 2, 3, 4, 5 or 6 equally probably each time it is tossed*. Let *B* be the event that it turns up 1 in the next toss, we write

$P(B) = 1/6$

Misconception

Please note that the probability 1/6 does **not** mean that it will turns up 1 in at most six tries. Its precise meaning will be discussed later on in the chapter. Roughly, it just means that on the long run (i.e. the die being tossed a large number of times), the proportion of 1's will be very close to 1/6.

Impossible and Certain events

Two types of events are special. One type are the impossible events (e.g., the sum of digits of a two-digit number is greater than 18); the other type are certain to happen (e.g., a roll of a die will turn up 1, 2, 3, 4, 5 or 6). The probability of an impossible event is 0, while that of a certain event is 1. We write

P(Impossible event) = 0

P(Certain event) = 1

The above reinforces a very important principle concerning probability. Namely, the range of probability is between 0 and 1. You can **never** have a probability of 2.5! So remember the following

$$0 \leq P(E) \leq 1$$

for all events E.

Complement of an event

A most useful concept is the **complement** of an event. We use : $\overline{B}$ to represent the *event* that *the die will NOT turn up 1 in the next toss*. Generally, putting a bar over a variable (that represents an event) means the opposite of that event. In the above case of a die:

$$P(\overline{B}) = 5/6$$

it means *the die will turn up 2, 3, 4, 5 or 6 in the next toss has probability 5/6*. Please note that

$$P(\overline{E}) = 1 - P(E)$$

for any event E.

Combining independent probabilities

It is interesting how *independent* probabilities can be combined to yield probabilities for more complex events. I stress the word *independent* here, because the following demonstrations will not work without that requirement. The exact meaning of the word will be discussed a little later on in the chapter, and we will show why *independence* is important in Exercise 10 of this section.

Adding probabilities

Probabilities are added together whenever an event can occur in multiple "ways." As this is a rather loose concept, the following example may be helpful. Consider rolling a single die; if we want to calculate the probability for, say, rolling an odd number, we must add up the probabilities for all the "ways" in which this can happen -- rolling a 1, 3, or 5. Consequently, we come to the following calculation:

P(rolling an odd number) = P(rolling a 1) + P(rolling a 3) + P(rolling a 5) = 1/6 + 1/6 + 1/6 = 3/6 = 1/2 = 50%

Note that the addition of probabilities is often associated with the use of the word "or" -- whenever we say that some event E is equivalent to any of the events X, Y, **or** Z occurring, we use addition to combine their probabilities.

A general rule of thumb is that the probability of an event and the probability of its complement must add up to 1. This makes sense, since we intuitively believe that events, when well-defined, must either happen or not happen.

Multiplying probabilities

Probabilities are multiplied together whenever an event occurs in multiple "stages" or "steps." For example, consider rolling a single die twice; the probability of rolling a 6 both times is calculated by multiplying the probabilities for the individual steps involved. Intuitively, the first step is simply the first roll, and the second step is the second roll. Therefore, the final probability for rolling a 6 twice is as follows:

P(rolling a 6 twice) = P(rolling a 6 the first time) $\times$ P(rolling a 6 the second time) = $\frac{1}{6} \times \frac{1}{6}$= 1/36 $\approx$ 2.8%

Similarly, note that the multiplication of probabilities is often associated with the use of the word "and" -- whenever we say that some event E is equivalent to **all** of the events X, Y, **and** Z occurring, we use multiplication to combine their probabilities.

Also, it is important to recognize that the product of multiple probabilities must be less than or equal to each of the individual probabilities, since probabilities are restricted to the range 0 through 1. This agrees with our intuitive notion that relatively complex events are usually less likely to occur.

Combining addition and multiplication

It is often necessary to use both of these operations simultaneously. Once again, consider one die being rolled twice in succession. In contrast with the previous case, we will now consider the event of rolling two numbers that add up to 3. In this case, there are clearly two steps involved, and therefore multiplication will be used, but there are also multiple ways in which the event under consideration can occur, meaning addition must be involved as well. The die could turn up 1 on the first roll and 2 on the second roll, or 2 on the first and 1 on the second. This leads to the following calculation:

P(rolling a sum of 3) = P(1 on 1st roll) $\times$ P(2 on 2nd roll) + P(2 on 1st roll) $\times$ P(1 on 2nd roll)

$= \frac{1}{6} \times \frac{1}{6} + \frac{1}{6} \times \frac{1}{6} = 1/18 \approx 5.5\%$

This is only a simple example, and the addition and multiplication of probabilities can be used to calculate much more complex probabilities.

Exercises

Let *A* represent the number that turns up in a (fair) die roll, let *C* represent the number that turns up in a separate (fair) die roll, and let *B* represent a card randomly picked out of a deck:

1. A die is rolled. What is the probability of rolling a 3 i.e. calculate P(A = 3)?

2. A die is rolled. What is the probability of rolling a 2, 3, **or** 5 i.e. calculate P(A = 2) + P(A = 3) + P(A = 5)?

3. What is the probability of choosing a card of the suit Diamonds? 4. A die is rolled and a card is randomly picked from a deck of cards. What is the probability of rolling a 4 **and** picking the Ace of spades, i.e. calculate P(A = 4)×P(B = Ace of spades).

5. Two dice are rolled. What is the probability of getting a 1 followed by a 3?

6. Two dice are rolled. What is the probability of getting a 1 and a 3, regardless of order?

7. Calculate the probability of rolling two numbers that add up to 7.

8. (Optional) Show the probability of *C* is equal to *A* is 1/6.

9. What is the probability that *C* is greater than *A*?

10. Gareth was told that in his class 50% of the pupils play football, 30% play video games and 30% study mathematics. So if he was to choose a student from the class randomly, he calculated the probability that the student plays football, video games or studies mathematics is 50% + 30% + 30% = 1/2 + 3/10 + 3/10 = 11/10. But all probabilities should be between 0 and 1. What mistake did Gareth make?

Solutions

1. P(A = 3) = 1/6

2. P(A = 2) + P(A = 3) + P(A = 5) = 1/6 + 1/6 + 1/6 = 1/2

3. P(B = Ace of Diamonds) + ... + P(B = King of Diamonds) = 13 × 1/52 = 1/4

4. P(A = 4) × P(B = Ace of Spades) = 1/6 × 1/52 = 1/312

5. P(A = 1) × P(A = 3) = 1/36

6. P(A = 1) × P(A = 3) + P(A = 3) × P(A = 1) = 1/36 + 1/36 = 1/18

7. Here are the possible combinations: $1 + 6 = 2 + 5 = 3 + 4 = 7$. Probability of getting each of the combinations are 1/18 as in Q6. There are 3 such combinations, so the probability is $3 \times 1/18 = 1/6$.

9. Since both dice are fair, $C > A$ is just as likely as $C < A$. So

$P(C > A) = P(C < A)$

and

$P(C > A) + P(C < A) + P(A = C) = 1$

But

$P(A = C) = 1/6$

so $P(C > A) = 5/12$.

10. For example, some of those 50% who play football may also study mathematics. So we can not simply add them.

Random Variables

A *random experiment*, such as *throwing a die* or *tossing a coin*, is a process that produces some uncertain outcome. We also require that a random experiments can be repeated easily. In this section we shall start using a capital letter to represent the outcome of a random experiment. For example, let *D* be the outcome of a die roll, *D* could take the value 1, 2, 3, 4, 5 or 6, but it is uncertain. We say *D* is a *random variable*. Suppose now I throw a die, and it turns up 5, we say the *observed value* of *D* is 5.

A random variable is simply the outcome of a certain random experiment. It is usually denoted by a CAPITAL letter, but its observed value is not. For example let

$D_1,D_2,...,D_n$

denote the outcome of *n* die throws, then we usually use

$d_1,d_2,...,d_n$

to denoted the observed values of each of D_i's.

From here on, random variable may be abbreviated as simply rv (a common abbreviation in other probability literatures).

The Bernoulli

This section is optional and it assumes knowledge of binomial expansion.

A Bernoulli experiment is basically a "coin-toss". If we toss a coin, we will expect to get a head or a tail equally probably. A Bernoulli experiment is slightly more versatile than that, in that the two possible outcomes need not have the same probability.

In a Bernoulli experiment you will either get a

success, denoted by 1, with proability *p* (where *p* is a number between 0 and 1)

or a

failure, denoted by 0, with probaility 1 - *p*.

If the random variable *B* is the outcome of a Bernoulli experiment, and the probability of getting a 1 is *p*, we say *B* comes from a *Bernoulli distribution* with success probability *p* and we write:

$$B \sim Ber(p)$$

For example, if

$$C \sim Ber(0.65)$$

then

P(C = 1) = 0.65

and

P(C = 0) = 1 - 0.65 = 0.35

Binomial Distribution

Suppose we want to repeat the Bernoulli experiment *n* times, then we get a binomial distribution. For example:

$$C_i \sim Ber(p)$$

for i = 1, 2, ... , n. That is, there are *n* variables C_1, C_2, ... , C_n and they all come from the same Bernoulli distribution. We consider:

$B = C_1 + C_2 + ... + C_n$

, then *B* is simply the rv that counts the number of successes in *n* trials (experiments). Such a variables is called a binomial variable, and we write

$$B \sim B(n, p)$$

Example 1

Aditya, Gareth, and John are equally able. Their probability of scoring 100 in an exam follows a Bernoulli distribution with success probability 0.9. What is the probability of

i) One of them getting 100?

ii) Two of them getting 100?

iii) All 3 getting 100?

iv) None getting 100?

Solution

We are dealing with a binomial variable, which we will call *B*. And

$$B \sim Bin(3, 0.9)$$

i) We want to calculate

$P(B = 1)$

The probability of any of them getting 100 (success) and the other two getting below 100 (failure) is

$$0.9 \times 0.1 \times 0.1 = 0.009$$

but there are 3 possible candidates for getting 100 so

$$P(B = 1) = 3 \times 0.009 = 0.027$$

ii) We want to calculate

$P(B = 2)$

The probability is

$$0.9 \times 0.9 \times 0.1 = 0.081$$

but there are $\binom{3}{2}$ combinations of candidates for getting 100, so

$$P(B = 2) = \binom{3}{2} \times 0.081 = 0.243$$

iii) To calculate

$$P(B = 3) = 0.9 \times 0.9 \times 0.9 = 0.729$$

iv) The probability of "None getting 100" is getting 0 success, so

$$P(B = 0) = 0.1 \times 0.1 \times 0.1 = 0.001$$

The above example strongly hints at the fact the binomial distribution is connected with the binomial expansion. The following result regarding the binomial distribution is provided without proof, the reader is encouraged to check its correctness.

If

$$B \sim Bin(n, p)$$

then

$$P(B = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

This is the *k*th term of the binomial expansion of $(p + q)^n$, where $q = 1 - p$.

Exercises ...

Distribution

...

Events

In the previous sections, we have slightly abused the use of the word event. An event should be thought of as a collection of random outcomes of a certain rv.

Let us introduce some notations first. Let *A* and *B* be two events, we define

$$A \cap B$$

to be the event of *A and B*. We also define

$$A \cup B$$

to be the event of *A or B*. As demonstrated in exercise 10 above,

$$P(A \cup B) \neq P(A) + P(B)$$

in general.

Let's see some examples. Let *A* be the event of getting a number less than or equal to 4 when tossing a die, and let *B* be the event of getting an odd number. Now

P(*A*) = 2/3

and

P(*B*) = 1/2

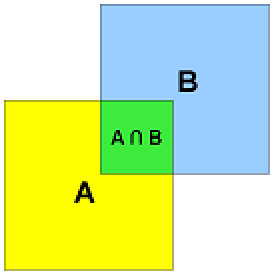
but the probability of *A or B* does not equal to the sum of the probabilities, as below

$$P(A \cup B) \neq P(A) + P(B) = \frac{1}{2} + \frac{2}{3} = \frac{7}{6}$$

as 7/6 is greater than 1.

It is not difficult to see that the event of throwing a 1 or 3 is included in both *A* and *B*. So if we simply add P(*A*) and P(*B*), some events' probabilities are being added twice!

The Venn diagram below should clarify the situation a little more,



think of the blue square as the probability of *B* and the yellow square as the probability of *A*. These two probabilities overlap, and where they do is the probability of *A and B*. So the probability of *A or B* should be:

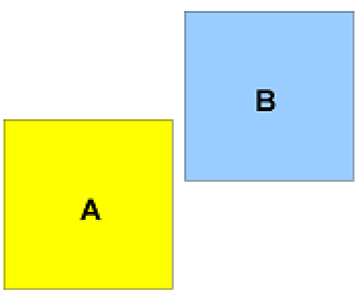
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

The above formula is called the *Simple Inclusion Exclusion Formula*.

If for events *A* and *B*, we have

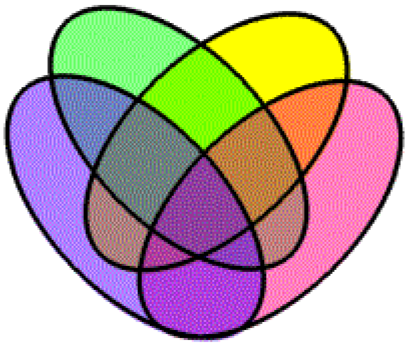
$$P(A \cap B) = 0$$

we say *A* and *B* are **disjoint**. The word means *to separate*. If two events are disjoint we have the following Venn diagram representing them:



info -- Venn Diagram

Traditionally, Venn Diagrams are used to illustrate sets graphically. A set being simply a collection of things, e.g. {1, 2, 3} is a set consisting of 1, 2 and 3. Note that Venn diagrams are usually drawn round. It is generally very difficult to draw Venn diagrams for more than 3 intersecting sets. E.g. below is a Venn diagram showing four intersecting sets:



Expectation

The expectation of a random variable can be roughly thought of as the long term average of the outcome of a certain repeatable random experiment. By long term average it is meant that if we perform the underlying experiment many times and average the outcomes. For example, let *D* be as above, the observed values of *D* (1,2 ... or 6) are equally likely to occur. So if you were to toss the die a large number of times, you would expect each of the numbers to turn up roughly an equal number of times. So the expectation is

$$\frac{1+2+3+4+5+6}{6} = 3.5$$

. We denote the expection of *D* by E(*D*), so

$E(D) = 3.5$

We should now properly define the expectation.

Consider a random variable *R*, and suppose the possible values it can take are $r_1, r_2, r_3, \ldots, r_n$. We define the expectation to be

$E(R) = r_1P(R = r_1) + r_2P(R = r_2) + \ldots + r_nP(R = r_n)$

Think about it: Taking into account the expectation is the long term average of the outcomes. Can you explain why is E(*R*) defined the way it is?

Example 1 In a fair coin toss, let 1 represent tossing a head and 0 a tail. The same coin is tossed 8 times. Let *C* be a random variable representing the number of heads in 8 tosses? What is the expectation of *C*, i.e. calculate E(*C*)?

Solution 1 ...

Solution 2 ...

Areas as probability

The uniform distributions. ... .

Order Statistics

Estimate the x in U[0, x]. ...

Addition of the Uniform distribution

Adding U[0,1]'s and introduce the CLT. ....

to be continued ...

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Financial Options

Binary tree option pricing

Introduction

We have all heard of at least one stock exchange. NASDAQ, Dow Jones, FTSE and Hang Sheng. Less well-known, but more useful to many people, are the future exchanges. A stock exchange allows stock brokers to trade company stocks, while the future exchanges allow more exotic derivatives to be traded. For example, financial options, which is also the focus of this chapter.

An option is a contract that gives the holder the choice to buy (or sell) a certain good in some time in the future for a certain price. What are options for? Initially, they are used to protect against risk. But they are also used to take advantage foreseeable opportunities, like what Thales has done[1].

Thales, the great Greek philosopher, was credited with the first recorded use of an option in the western world. A popular anecdote suggests, in one particular year while still in winter, he forecasted a great harvest of olives in the coming year. He had next to no money, so he purchased the option for the use of all the olive presses in his area. Naturally, when the time to harvest came everyone wanted to use the presses he had optioned! Needless to say he made a lot of money out of it.

Basics

An **option** is a contract of choice. You can choose whether to exercise the option or not.

If you own an option that states

You may purchase 1 kg of sugar from Shop A tomorrow for $2

suppose tomorrow the market price of sugar is $3, you would want to exercise the option i.e. buy the sugar for $2. Then you would sell it for $3 on the market and make $1 in the process. But if the market price for 1 kg of sugar is $1, then you would choose not to exercise the option, because it's cheaper on the market.

Let us be a little bit more formal about what an option is. In particular there are two types of option:

Call Option

A call option is a contract that gives the owner the option to *buy* an 'underlying stock' at the 'strike price', on the 'expiry date'.

Put Option

A put option is a contract that gives the owner the option to *sell* an 'underlying stock' at the 'strike price', on the 'expiry date'.

In the above example, the 'underlying stock' was sugar and the 'strike price' $2 and 'expiry date' is tomorrow.

We shall represent an option like below

{C or P, $amount, # periods to expiry}

. For example

{C,$3,1}

represents a call option with strike price for some unspecified underlying stock expiring in one time-unit's time. A time-unit here may be a year, a month, one day or one hour. The important point is the mathematics we will present later does not really depend on what this time-unit is. Also, we need not specify the underlying stock either. Another example

{P,$100,2}

represents a put option with strike price $100 for some unspecified underlying stock expiring in *2* time-units' time.

Now that we have a basic idea of what an option is, we can start to imagine a market place where options are traded. We assume that such a *market* exists. Also we assume that there is no fee of any kind to participate in a trade. Such a market is called a *frictionless* market. Of course, a market place where the underlying stock is traded is also assumed to exist.

Info -- American or European

Actually there are two major types of options: American or European. An European option allows you to exercise the option *only* on the 'expiry date'; while the American version allows you to exercise the option at any time prior to the 'expiry date'. We shall only discuss European options in this chapter.

Arbitrage

Another very important concept is *arbitrage*. In short, an arbitrage is a way to make money out of nothing. We assume that there is no *free-lunch* in this world, in other words our market is arbitrage-free. We will show an example of how to perform an arbitrage later on in the chapter.

The real meat of this chapter is the technique used to price the options. In simple terms, we have an option, how much should it be? From this angle, we will see that the arbitrage-free requirement is a very strong one, in that it basically dictates what the price of the option should be.

Option's value on expiry

Pricing the option is about how much is it worth *now*. Of course the present value of an option depends on its possible future values. Therefore it is vital to understand how much the option is worth at expiry, when it is time to choose whether to exercise the option or not. For example, consider the option

{C,$2,1}

it is the call (buy) option that expires in 1 week's time (or day or year or whatever time period it is suppose to be). How much should the option be if the market price of the underlying stock on expiry is $3? What if the market price is $1?

It is sensible to say the option has a value of $1 if the market price (for the underlying stock) is $3, and the option should be worthless ($0) if the market price is $1.

Why do we say it is *sensible* to price the option as above? It is because we assume the market is arbitrage-free. Also in a market, we assume

- there is a bank that's willing to lend you money
- if you repay the bank in the same day you borrowed, no fee will be charged.

With those assumptions, we show that if you price the option any differently, someone can make money without using any of his/her own money. For example, suppose on expiry, the market price for the underlying stock is $3 and you decide to sell the option for $0.7 (not $1 as is sensible). An intelligent buyer would do the following:

Action	**Money**	**Balance**
Borrow $2.7	+$2.7	$2.7
Purchase your option for $0.7	-$0.7	$2
Purchase sugar for $2 with option	-$2	$0
Sell 1kg of sugar for $3 in market	+$3	$3
Repay bank $2.7	-$2.7	$0.3

He/she made $0.3 and at no time did he/she use his/her own money (i.e. balance never less than zero)! This is a *free lunch*, which is contrary to the assumption of a arbitrage-free market!

Exercises

1. In an *arbitrage-free* market, consdier an option T = {C,$100,1}.

i) How much should the option be on expiry if the price of the underlying stock is $90.

ii) What if the underlying stock costs \$110 on expiry.

iii} \$100?

2. Consdier an option T = {C,\$10,1}.

i) On expiry, would you consider buying the option if it was for sell for \$2 if the underlying stock costs \$8?

ii) What if the underlying stock costs \$7.

3. Consider the *put* option T = {P,\$2,1}. On expiry the underlying stock costs \$1. Jenny owns T, she decides on the following actions

Borrow \$1

Purchase the underlying stock from the market for \$1

Exercise the option i.e. sell the stock for \$2

Repay \$1

Did she do the right thing?

4. In an *arbitrage-free* market, consider the *put* option T = {P,\$2,1}.

i) On expiry, how much should the option cost if the underlying stock costs \$1?

ii) \$3?

5. Consider the *put* option T = {P,\$2,1}. On expiry the underlying stock costs \$1. And the option T is on sale for \$0.5. Jenny immediately sees an arbitrage oppoturnity. Detail the actions she should take to capitalise on the arbitrage oppoturnity. (Hint: immitate the Action, Money, Balance table)

Pricing an option

Consider this hypothetical situation where a company, MassiveSoft, is in negatiation to merge with another company, Pears. The share price of MassiveSoft currently stands at \$7. If the negotiation is successful, the share price will rise to \$11; otherwise it will fall to \$5. Experts predict the probability of a success is 90%. Consider a call option that lets you buy 1000 shares of MassiveSoft at \$8 when the negotiation is finalised. How much should the option be?

Since the market is arbitrage-free, the value of the option at expiry is already determined. Of course

if the negotiation is successful, the option is valued at (11 - 8) × 1000 = \$3000

otherwise, the option should be worthless (\$0) the above are the only *correct* values of the option at expiry or people can "rip you off".

Let *x* be the price of the option at present, we can use the following diagrams to illustrate the situation,

	$\nearrow$	\$3000
$\$x$		
	$\searrow$	\$0

the diagram shows that the current price of the option should be $*x*, and if the negotiation is successful, it will be worth \$3000, otherwise it is worthless. In similar fashion, the following diagram shows the value of the company stock now, and in the future

	$\nearrow$	\$11
\$7		
	$\searrow$	\$5

You may have notice that we didn't put down the probability of success or failure. Interestingly (and counter-intuitively), they don't matter! Again, the arbitrage-free principle dictates that what we have in the two diagrams above are sufficient for us to price the option!

How?

What is the option? It is the contract that gives you the option to buy ... Wait, wait, wait. Think of it from another angle

it is a tradable object that is worth \$2000 if the negotiation is successful, and \$0 if otherwise

This is the main idea behind how to price the option. The option must be the same price as another *object* that goes up to \$2000 or down to \$0 depending on the success of the negotiation. Hopefully, this object is something we know the price of. This idea is called constructing a *replicating portfolio*.

A *portfolio* is a collection of tradable *things*. We want to construct a portfolio that behaves in the same way as the option. It turns out that we can construstt a portfolio that behaves in the way as the option by using only two things. They are

12. MassiveSoft shares

13. and *money*

let's assume that *money* is *tradable* in the sense that you can buy a dollar with a dollar. This concept may seem very unintuitive at first. However let's proceed with the mathematics, suppose this porfolio consists of *y* units of MassiveSoft shares and *z* units of money. If the negotiation is successful, then each share will be worth \$9, and the whole portolio should be worth \$2000, as it behaves in the same way as the option, so we have the following

$$11y + z = 3000$$

but if the negotiation is unsuccessful then the portfolio is worthless (\$0) and MassiveSoft share prices will fall to \$5, giving

$$5y + z = 0$$

we can easily solve the above simultaneous equations. We get

$$6y = 3000$$

and so

$y = 500$ and $z = -\$2500$

So this portolio consists of 500 MassiveSoft shares and -\$2500. But what is -\$2500? This can be understood as an *obligation* to pay back some money (e.g. from borrowings) on the expiry date of the option. So the porfolio we constructed can be thought of as

500 MassiveSoft shares and an obligation to pay \$2500

Now, 500 MassiveSoft shares costs \$7 × 500 = 3500, so the option should be priced as 3500 - 2500 = \$1000.

Let's price a few more options.

...

The famous mathematicain, John Nash, as portrayed in the movie "A beautiful mind", did some pioneering work in portfolio theory with equivalent functions.

Call-Put parity

...more to come

Reference

5. A Brief History of Options

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Matrices

Introduction

A matrix may be more popularly known as a giant computer simulation, but in mathematics it is a totally different thing. To be more precise, a matrix (plural matrices) is an array of numbers. For example, below is a typical way to write a matrix, with numbers arranged in rows and column and with round brackets around the numbers

$$\begin{pmatrix} 1 & 5 & 10 & 20 \\ 1 & -3 & -5 & 9 \\ 3 & -1 & -1 & -1 \\ 3 & 2 & 4 & -5 \end{pmatrix}$$

The above matrix has 4 rows and 4 columns, so we call it a 4 × 4 (4 by 4) matrix. Also, we can have matrices of many different shapes. The *shape* of a matrix is the name for the dimensions of matrix (m by n, where m is the number of rows and n the number of columns). Here are some more examples of matrices

This is an example of a 3 × 3 matrix:

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

This is an example of a 5 × 4 matrix:

$$\begin{pmatrix} a & b & c & d \\ h & g & f & e \\ i & j & k & l \\ p & o & n & m \\ q & r & s & t \end{pmatrix}$$

This is an example of a 1 × 6 matrix:

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

The theory of matrices is intimately connected with that of (linear) simultaneous equations. The ancient Chinese had established a systematic way to solve simultaneous equations. The theory of simultaneous equations is furthered in the east by the Japanese mathematician, Seki and a little later by Leibniz, Newton's greatest rival. Later, Gauss (1777 - 1855), one of the three pillars of modern mathematics popularised the use of Gaussian elimination, which is a simple step by step algorithm for solving any number of linear simultaneous equations. By then the use of matrices to represent simultaneous equation neatly on paper (as discussed above) has become quite common[1].

Consider the simultaneous equations:

$x + y = 10$

$x - y = 4$

it has the solution $x = 7$ and $y = 3$, and the usual way to solve it is to add the two equations together to eliminate the y. Matrix theory offers us another way to solve the above simultaneous equations via matrix *multiplication* (covered below). We will study the widely accepted way to *multiply* two matrices together. In theory with *matrix multiplication* we can solve any number of simultaneous equations, but we shall mainly restrict our attention to 2 × 2 matrices. But even with that restriction, we have opened up doors to topics simultaneous equations could never offer us. Two such examples are

9. using matrices to solve linear recurrence relations which can be used to model population growth, and

10. encrypting messages wtih matrices.

We shall commence our study by learning some of the more fundamental concepts of matrices. Once we have a firm grasp of the basics, we shall move on to study the real meat of this chapter, matrix multiplication.

Elements

An element of a matrix is a particular number inside the matrix, and it is uniquely located with a pair of numbers. E.g. let the following matrix be denoted by A, or symbolically:

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

the (2,2)th entry of A is 5; the (1,1)th entry of A is 1, the (3,3) entry of A is 9 and the (3,2)th entry of A is 8. The (i , j)th entry of A is usually denoted $a_{i,j}$ and the (i , j)th entry of a matrix B is usually denoted by $b_{i,j}$ and so on.

Summary

- A matrix is an array of numbers
- A $m \times n$ matrix has m rows and n columns

- The *shape* of a matrix is determined by its number of rows and columns
- The (i,j)th element of a matrix is located in ith row and jth column

Matrix addition & Multiplication by a scalar

Matrices can be added together. But only the matrices of the same *shape* can be added. This is very natural. E.g.

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

$$B = \begin{pmatrix} 2 & 9 & 8 \\ 0 & -1 & 8 \\ 4 & 6 & 7 \end{pmatrix}$$

then

$$A + B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} + \begin{pmatrix} 2 & 9 & 8 \\ 0 & -1 & 8 \\ 4 & 6 & 7 \end{pmatrix} = \begin{pmatrix} 1+2 & 2+9 & 3+8 \\ 4+0 & 5+(-1) & 6+8 \\ 7+4 & 8+6 & 9+7 \end{pmatrix}$$

$$= \begin{pmatrix} 3 & 11 & 11 \\ 4 & 4 & 14 \\ 11 & 14 & 16 \end{pmatrix}$$

Similarly matrices can be multiplied by a number, we call the number a *scalar*, this is to distinguish it from a matrix. The reader need not worry about the definition here, just remember that a *scalar* is just a number.

$$5A = A + A + A + A + A = 5\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 5 & 10 & 15 \\ 20 & 25 & 30 \\ 35 & 40 & 45 \end{pmatrix}$$

in this case the scalar value is 5. In general, when we do $s \times A$, where s is a scalar and A a matrix, we multiply each entry of A by s.

Matrix Multiplication

The widely accepted way to multiply two matrices together is definitely non-intuitive. As mentioned above, multiplication can help with solving simultaneous equations. We will now give a brief outline of how this can be done. Firstly, any system of linear simultaneous equations can be written as a matrix of coefficients multiplied by a matrix of unknowns equaling a matrix of results. This description may sound a little complicated, but in symbolic form it is quite clear. The previous statement simply says that if *A*, *x* and *b* are matrices, then $Ax = b$, can be used to represent some system of simultaneous equations. The beautiful thing about matrix multiplications is that some matrices can have multiplicative inverses, that is we can multiply both sides of the equation by A^{-1} to get $x = A^{-1}b$, which effectively solves the simultaneous equations.

The reader will surely come to understand matrix multiplication better as this chapter progresses. For now we should consider the simplest case of matrix multiplication, multiplying *vectors*. We will see a few examples and then we will explain process of multiplication

$$A_{2\times 1} = \begin{pmatrix} 2 \\ 9 \end{pmatrix} \quad , \quad B_{1\times 2} = \begin{pmatrix} 3 & 5 \end{pmatrix}$$

then

$$B_{1\times 2} \times A_{2\times 1} = \begin{pmatrix} 3 & 5 \end{pmatrix} \times \begin{pmatrix} 2 \\ 9 \end{pmatrix} = \begin{pmatrix} (3 \times 2) + (5 \times 9) \end{pmatrix} = \begin{pmatrix} 51 \end{pmatrix}$$

Similarly if:

$$A_{3\times 1} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad , \quad B_{1\times 3} = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix}$$

then

$$B_{1\times 3} \times A_{3\times 1} = \begin{pmatrix} 4 & 5 & 6 \end{pmatrix} \times \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} (4 \times 1) + (5 \times 2) + (6 \times 3) \end{pmatrix} = \begin{pmatrix} 32 \end{pmatrix}$$

A matrix with just one row is called a *row vector*, similarly a matrix with just one column is called a *column vector*. When we multiply a row vector *A*, with a column vector *B*, we multiply the element in the first column of *A* by the element in the first row of *B* and add to that the product of the second column of *A* and second row of *B* and so on. More generally we multiply $a_{1,i}$ by $b_{i,1}$ (where *i* ranges from 1 to n, the number of rows/columns) and sum up all of the products. Symbolically:

$$A_{1\times n} \times B_{n\times 1} = (\sum_{i=1}^{n} a_{1,i} \times b_{i,1})$$ (for information on the $\sum$ sign, see [Summation_Sign])

where *n* is the number of rows/columns.

In words: the product of a column vector and a row vector is the sum of the product of item 1,*i* from the row vector and *i*,1 from the column vector where *i* is from 1 to the width/height of these vectors.

Note: The product of matrices is also a matrix. The product of a row vector and column vector is a 1 by 1 matrix, not a scalar.

Exercises

Multiply:

$$\begin{pmatrix} 1 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix} \begin{pmatrix} 1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} \frac{1}{8} & 9 \end{pmatrix} \begin{pmatrix} 16 \\ 2 \end{pmatrix}$$

$$\begin{pmatrix} a & b \end{pmatrix} \begin{pmatrix} d \\ e \end{pmatrix}$$

$$\begin{pmatrix} 6+6b & 3-b \end{pmatrix} \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & abc \end{pmatrix} \begin{pmatrix} a \\ 0 \end{pmatrix}$$

Multiplication of non-vector matrices

Suppose $A_{m \times n} B_{n \times p} = C_{m \times p}$ where *A*, *B* and *C* are matrices. We multiply the *i*th row of *A* with the *j*th column of *B* as if they are vector-matrices. The resulting number is the (*i*,*j*)th element of *C*. Symbolically:

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} \times b_{k,j}$$

Example 1

Evaluate $AB = C$ and $BA' = \mathrm{D}$, *where*

$$A = \begin{pmatrix} 3 & 2 \\ 5 & 6 \end{pmatrix}$$

and

$$B = \begin{pmatrix} 2 & 6 \\ 8 & 7 \end{pmatrix}$$

Solution

$$c_{1,1} = \begin{pmatrix} 3 & 2 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = (3 \times 2 + 2 \times 8) = 22$$

$$c_{1,2} = \begin{pmatrix} 3 & 2 \end{pmatrix} \begin{pmatrix} 6 \\ 7 \end{pmatrix} = (3 \times 6 + 2 \times 7) = 32$$

$$c_{2,1} = \begin{pmatrix} 5 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ 8 \end{pmatrix} = (5 \times 2 + 6 \times 8) = 58$$

$$c_{2,2} = \begin{pmatrix} 5 & 6 \end{pmatrix} \begin{pmatrix} 6 \\ 7 \end{pmatrix} = (5 \times 6 + 6 \times 7) = 72$$

i.e.

$$C = \begin{pmatrix} 22 & 32 \\ 58 & 72 \end{pmatrix}$$

$$d_{1,1} = \begin{pmatrix} 2 & 6 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = (2 \times 3 + 6 \times 5) = 36$$

$$d_{1,2} = \begin{pmatrix} 2 & 6 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix} = (2 \times 2 + 6 \times 6) = 40$$

$$d_{2,1} = \begin{pmatrix} 8 & 7 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} = (8 \times 3 + 7 \times 5) = 59$$

$$d_{2,2} = \begin{pmatrix} 8 & 7 \end{pmatrix} \begin{pmatrix} 2 \\ 6 \end{pmatrix} = (8 \times 2 + 7 \times 6) = 58$$

i.e.

$$D = \begin{pmatrix} 36 & 40 \\ 59 & 58 \end{pmatrix}$$

Example 2 Evaluate *AB* and *BA* where

$$A = \begin{pmatrix} 5 & 17 \\ 2 & 7 \end{pmatrix}$$

$$B = \begin{pmatrix} 7 & -17 \\ -2 & 5 \end{pmatrix}$$

Solution

$$\begin{pmatrix} 5 & 17 \\ 2 & 7 \end{pmatrix}\begin{pmatrix} 7 & -17 \\ -2 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 7 & -17 \\ -2 & 5 \end{pmatrix}\begin{pmatrix} 5 & 17 \\ 2 & 7 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Example 3 Evaluate *AB* and *BA* where

$$A = \begin{pmatrix} 2 & 6 \\ 0 & 5 \end{pmatrix}$$

$$B = \begin{pmatrix} 5 & -6 \\ 0 & 2 \end{pmatrix}$$

Solution

$$\begin{pmatrix} 2 & 6 \\ 0 & 5 \end{pmatrix}\begin{pmatrix} 5 & -6 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}$$

$$\begin{pmatrix} 5 & -6 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 2 & 6 \\ 0 & 5 \end{pmatrix} = \begin{pmatrix} 10 & 0 \\ 0 & 10 \end{pmatrix}$$

Example 4 Evaluate the following multiplication:

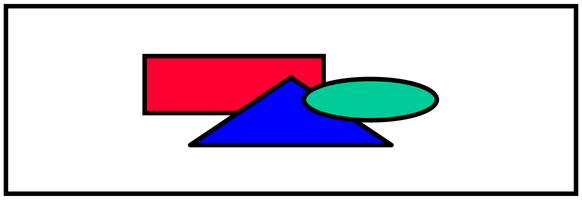
$$\begin{pmatrix} a \\ b \end{pmatrix}\begin{pmatrix} c & d \end{pmatrix}$$

Solution

Note that:

$$\begin{pmatrix} a \\ b \end{pmatrix}$$

is a 2 by 1 matrix and



is a 1 by 2 matrix. So the multiplication makes sense and the product should be a 2 by 2 matrix.

$$\begin{pmatrix} a \\ b \end{pmatrix}\begin{pmatrix} c & d \end{pmatrix} = \begin{pmatrix} ac & ad \\ bc & bd \end{pmatrix}$$

Example 5 Evaluate the following multiplication:

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix}$$

Solution

$$\begin{pmatrix} 1 \\ 2 \end{pmatrix}\begin{pmatrix} 3 & 4 \end{pmatrix} = \begin{pmatrix} 1 \times 3 & 1 \times 4 \\ 2 \times 3 & 2 \times 4 \end{pmatrix} = \begin{pmatrix} 3 & 4 \\ 6 & 8 \end{pmatrix}$$

Example 6 Evaluate the following multiplication:

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}$$

Solution $$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix} = \begin{pmatrix} ac & 0 \\ 0 & bd \end{pmatrix}$$

Example 7 Evaluate the following multiplication:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

Solution $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax+by \\ cx+dy \end{pmatrix}$

Note Multiplication of matrices is generally not commutative, i.e. generally $AB \neq BA$.

Diagonal matrices

A diagonal matrix is a matrix with zero entries everywhere except possibly down the diagonal. Multiplying diagonal matrices is really convenient, as you need only to multiply the diagonal entries together.

Examples

The following are all diagonal matrices

$$\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} a & 0 & 0 \\ 0 & c & 0 \\ 0 & 0 & 0 \end{pmatrix}$$

Example 1 $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} e & 0 \\ 0 & f \end{pmatrix}\begin{pmatrix} h & 0 \\ 0 & i \end{pmatrix} = \begin{pmatrix} aeh & 0 \\ 0 & bfi \end{pmatrix}$

Example 2 $\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}\begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix} = \begin{pmatrix} a^3 & 0 \\ 0 & b^3 \end{pmatrix}$

The above examples show that if D is a diagonal matrix then D^k is very easy to compute, all we need to do is to take the diagonal entries to the kth power. This will be an extremely useful fact later on, when we learn how to compute the nth Fibonacci number using matrices.

Exercises

1. State the dimensions of C

a) $C = A_{n\times p}B_{p\times m}$

b) $C = \begin{pmatrix} 10^{10} & 20 \\ 5000 & 0 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 5 & 6 & 6 \end{pmatrix}$

2. Evaluate. Please note that in matrix multiplication $(AB)C = A(BC)$ i.e. the order in which you do the multiplications does not matter (proved later).

a)

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

b)

$$\begin{pmatrix}3 & 1\\2 & 8\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}$$

3. Performing the following multiplications:

$$C = \begin{pmatrix}1 & 2\\4 & 5\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}$$

$$D = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\4 & 5\end{pmatrix}$$

What do you notice?

The Identity & multiplication laws

The exercise above showed us that the matrix:

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

is a very special. It is called the 2 by 2 identity matrix. An identity matrix is a *square* matrix, whose diagonal entries are 1's and all other entries are zero. The identity matrix, *I*, has the following very special properties

10. $$A \times I = A$$

11. $$I \times A = A$$

for all matrices *A*. We don't usually specify the *shape* of the identity because it's obvious from the context, and in this chapter we will only deal with the 2 by 2 identity matrix. In the real number system, the number 1 satisfies: r × 1 = r = 1 × r, so it's clear that the identity matrix is analogous to "1".

Associativity, distributivity and (non)-commutativity

Matrix multiplication is a great deal different to the multiplication we know from multiplying real numbers. So it is comforting to know that many of the laws the real numbers satisfy also carries over to the matrix world. But with one **big** exception, in general $AB \neq BA$.

Let *A*, *B*, and *C* be matrices. *Associativity* means

$(AB)C = A(BC)$

i.e. the order in which you multiply the matrices is unimportant, because the final result you get is the same regardless of the order which you do the multiplications.

On the other hand, *distributivity* means

$A(B + C) = AB + AC$

and

$(A + B)C = AC + BC$

Note: The commutative property of the real numbers (i.e. ab = ba), does not carry over to the matrix world.

Convince yourself

For all 2 by 2 matrices A, B and C. And I the identity matrix.

1. Convince yourself that in the 2 by 2 case:

$A(B + C) = AB + AC$

and

$(A + B)C = AC + BC$

2. Convince yourself that in the 2 by 2 case:

$A(BC) = (AB)C$

3. Convince yourself that:

$$AB \neq BA$$

in general. When does $AB = BA$? Name at least one case.

Note that all of the above are true for all matrices (of any dimension/shape).

Determinant and Inverses

We shall consider the simultaneous equations:

$ax + by = \alpha$ (1)

$cx + dy = \beta$ (2)

where a, b, c, d, α and β are constants. We want to determine the necessary conditions for (1) and (2) to have a *unique* solution for x and y. We proceed:

Let (1') = (1) × c

Let (2') = (2) × a

i.e.

$acx + bcy = c\alpha$ (1')

$acx + ady = a\beta$ (2')

Now

let (3) = (2') - (1')

$(ad - bc)y = a\beta - c\alpha$ (3)

Now y can be uniquely determined if and only if $(ad - bc) \neq 0$. So the necessary condition for (1) and (2) to have a unique solution depends on all four of the coefficients of x and y. We call this number $(ad - bc)$ the *determinant*, because it tells us whether there is a unique solutions to two simultaneous equations of 2 variables. In summary

if $(ad - bc) = 0$ then there is *no unique* solution

if $(ad - bc) \neq 0$ then there *is* a *unique* solution.

Note: *Unique*, we can not emphasise this word enough. If the determinant is zero, it doesn't necessarily mean that there is no solution to the simultaneous equations! Consider:

$x + y = 2$

$7x + 7y = 14$

the above set of equations has determinant zero, but there is obviously a solution, namely $x = y = 1$. In fact there are infiinitely many solutions! On the other hand consider also:

$x + y = 1$

$x + y = 2$

this set of equations has determinant zero, and there is no solution at all. So if determinant is zero then there is either no solution or infinitely many solutions.

Determinant of a matrix

We define the determinant of a 2 × 2 matrix

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

to be

$$\det(A) = ad - bc$$

Inverses

It is perhaps, at this stage, not very clear what's the use of the det(A). But it's intimately connected with the idea of an inverse. Consider in the real number system a number b, it has (multiplicative) inverse $1/b$, i.e. $b(1/b) = (1/b)b = 1$. We know that $1/b$ does not exist when $b = 0$.

In the world of matrices, a matrix *A* may or may not have an inverse depending on the value of the determinant det(A)! How is this so? Let's suppose *A* (known) does have an inverse *B* (i.e. $AB = I = BA$). So we aim to find *B*. Let's suppose further that

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

and

$$B = \begin{pmatrix} w & x \\ y & z \end{pmatrix}$$

we need to solve four simultaneous equations to get the values of *w*, *x*, *y* and *z* in terms of *a*, *b*, *c*, *d* and det(*A*).

$aw + by = 1$

$cw + dy = 0$

$ax + bz = 0$

$cx + dz = 1$

the reader can try to solve the above by him/herself. The required answer is

$$B = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

In here we assumed that *A* has an inverse, but this doesn't make sense if $\det(A) = 0$, as we can not divide by zero. So A^{-1} (the inverse of *A*) exists if and only if $\det(A) \neq 0$.

Summary

If $AB = BA = I$, then we say *B* is the inverse of *A*. We denote the inverse of *A* by A^{-1}. The inverse of a 2 × 2 matrix

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

is

$$A^{-1} = \frac{1}{\det(A)} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

provided the determinant of *A*, det(*A*) is zero.

Solving simultaneous equations

Suppose we are to solve:

$ax + by = \alpha$

$cx + dy = \beta$

We let

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

$$w = \begin{pmatrix} x \\ y \end{pmatrix}$$

$$\gamma = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

we can translate it into matrix form

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

i.e

$Aw = \gamma$

If A's determinant is not zero, then we can pre-multiply both sides by A^{-1} (the inverse of A)

$$\begin{aligned} A^{-1}Aw &= A^{-1}\gamma \\ Iw &= A^{-1}\gamma \\ w &= A^{-1}\gamma \end{aligned}$$

i.e.

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad - bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

which implies that x and y are unique.

Examples

Find the inverse of A, if it exists

a) $A = \begin{pmatrix} 1 & 5 \\ 2 & 3 \end{pmatrix}$

b) $A = \begin{pmatrix} 10 & 2 \\ 2 & 7 \end{pmatrix}$

c) $A = \begin{pmatrix} a & b \\ 3a & 3b \end{pmatrix}$

d) $A = \begin{pmatrix} 3 & 5 \\ 5 & 3 \end{pmatrix}$

Solutions

a) $A^{-1} = \frac{1}{-7}\begin{pmatrix} 3 & -5 \\ -2 & 1 \end{pmatrix}$

b) $A^{-1} = \frac{1}{66}\begin{pmatrix} 7 & -2 \\ -2 & 10 \end{pmatrix}$

c) No solution, as det(A) = 3ab - 3ab = 0

d) $A^{-1} = \frac{1}{-16}\begin{pmatrix} 3 & -5 \\ -5 & 3 \end{pmatrix}$

Exercises

1. Find the determinant of

$A = \begin{pmatrix} \frac{2}{5} & \frac{2}{3} \\ \frac{3}{2} & \frac{5}{2} \end{pmatrix}$. Using the determinant of A, decide whether there's a unique solution to the following simultaneous equations

$$\frac{2}{5}x + \frac{2}{3}y = 0$$
$$\frac{3}{2}x + \frac{5}{2}y = 0$$

2. Suppose

C = *AB*

show that

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

for the 2×2 case. Note: it's true for all cases.

3. Show that if you swap the rows of A to get A' , then $\det(A) = -\det(A')$

4. Using the result of 2

a) Prove that if:

$A = P^{-1}BP$

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

b) Prove that if:

$A^k = 0$

for some positive integer k, then $\det(A) = 0$.

5. a) Compute A^5, i.e. multiply A by itself 5 times, where

$$A = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$$

b) Find the inverse of P where

$$P = \begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}$$

c) Verify that

$$A = P^{-1}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}P$$

d) Compute A^5 by using part (b) and (c).

f) Compute A^{10}

Other Sections

Exercises

Matrices

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

Matrix Multiplication exercises

$$\begin{pmatrix}1 & 2\end{pmatrix}\begin{pmatrix}1\\2\end{pmatrix} = \begin{pmatrix}(1\times 1)+(2\times 2)\end{pmatrix} = \begin{pmatrix}5\end{pmatrix}$$

$$\begin{pmatrix}1\\2\end{pmatrix}\begin{pmatrix}1 & 2\end{pmatrix} = \begin{pmatrix}1\times 1 & 1\times 2\\ 2\times 1 & 2\times 2\end{pmatrix} = \begin{pmatrix}1 & 2\\ 2 & 4\end{pmatrix}$$

$$\begin{pmatrix}1/8 & 9\end{pmatrix}\begin{pmatrix}16\\2\end{pmatrix} = \begin{pmatrix}(1/8\times 16)+(9\times 2)\end{pmatrix} = \begin{pmatrix}20\end{pmatrix}$$

$$\begin{pmatrix}a & b\end{pmatrix}\begin{pmatrix}d\\e\end{pmatrix} = \begin{pmatrix}(a\times d)+(b\times e)\end{pmatrix} = \begin{pmatrix}a\times d+b\times e\end{pmatrix}$$

$$\begin{pmatrix}6+6b & 3-b\end{pmatrix}\begin{pmatrix}0\\0\end{pmatrix} = \begin{pmatrix}((6+6b)\times 0)+((3-b)\times 0)\end{pmatrix} = \begin{pmatrix}0\end{pmatrix}$$

$$\begin{pmatrix}0 & abc\end{pmatrix}\begin{pmatrix}a\\0\end{pmatrix} = \begin{pmatrix}(0\times a)+(abc\times 0)\end{pmatrix} = \begin{pmatrix}0\end{pmatrix}$$

Multiplication of non-vector matrices exercises

1.

a) $n \times m$

b) 2×4

2.

a)

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} =$$

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} =$$

$$\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}3\\1\end{pmatrix} =$$

$$\begin{pmatrix}4\\1\end{pmatrix}$$

b)

$$\begin{pmatrix}3 & 1\\2 & 8\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 2\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 1\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix} =$$

$$\begin{pmatrix}3 & 1\\2 & 8\end{pmatrix}\begin{pmatrix}1 & 1\\0 & 2\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix} =$$

$$\begin{pmatrix}3 & 1\\2 & 8\end{pmatrix}\begin{pmatrix}3\\2\end{pmatrix} =$$

$$\begin{pmatrix}11\\22\end{pmatrix}$$

3.

$$C = \begin{pmatrix}1 & 2\\4 & 5\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2\\4 & 5\end{pmatrix}$$

$$D = \begin{pmatrix}1 & 0\\0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\4 & 5\end{pmatrix} = \begin{pmatrix}1 & 2\\4 & 5\end{pmatrix}$$

The important thing to notice here is that the 2x2 matrix remains the same when multiplied with the other matrix. The matrix with only 1s on the diagonal and 0s elsewhere is known as the *identity* matrix, called *I*, and any matrix multiplied on either side of it stays the same. That is $A \times I = I \times A$

NB: The remaining exercises in this section are leftovers from previous exercises in the 'Multiplication of non-vector matrices' section

3.

$$C = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

$$D = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

The important thing to notice here is that the 1 to 9 matrix remains the same when multiplied with the other matrix. The matrix with only 1s on the diagonal and 0s elsewhere is known as the *identity* matrix, called *I*, and any matrix multiplied on either side of it stays the same. That is $A \times I = I \times A$

4. a)

$$A^5 = \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} =$$

$$\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -5 & 18 \\ -3 & 10 \end{pmatrix} =$$

$$\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -13 & 42 \\ -7 & 22 \end{pmatrix} =$$

$$\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} \begin{pmatrix} -29 & 90 \\ -15 & 46 \end{pmatrix} =$$

$$\begin{pmatrix} -61 & 186 \\ -31 & 94 \end{pmatrix}$$

b)

$$\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} (1 \times 3) + (-2 \times 1) & (1 \times 2) + (-2 \times 1) \\ (-1 \times 3) + (3 \times 1) & (-1 \times 2) + (3 \times 1) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

c) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} (a\times 1)+(b\times 0) & (a\times 0)+(b\times 1) \\ (c\times 1)+(d\times 0) & (c\times 0)+(d\times 1) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} (1\times a)+(0\times b) & (0\times a)+(1\times b) \\ (1\times c)+(0\times d) & (0\times c)+(1\times d) \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

d)

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

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

$$\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}$$

e) As an example I will first calculate A^2

$$A^2 = \begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix}\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}^2\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} =$$

$$\begin{pmatrix} 3 & 2 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 1^2 & 0 \\ 0 & 2^2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ -1 & 3 \end{pmatrix} =$$

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

$$\begin{pmatrix}3 & 8\\1 & 4\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}-5 & 18\\-3 & 10\end{pmatrix}$$

Now lets do the same simplifications I have done above with A^5-

$$A^5 = \begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}^5\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1^5 & 0\\0 & 2^5\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 32\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 64\\1 & 32\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}-61 & 186\\-31 & 94\end{pmatrix}$$

f)

$$A^{100} = \begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 2\end{pmatrix}^{100}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1^{100} & 0\\0 & 2^{100}\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2\\1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\0 & 1267650600228229401496703205376\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2535301200456458802993406410752\\1 & 1267650600228229401496703205376\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}-2535301200456458802993406410751 & 7605903601369376408980219232254\\-1267650600228229401496703205373 & 3802951800684688204490109616122\end{pmatrix}$$

Determinant and Inverses exercises

1.

$$\det(A) = \frac{2}{5} \times \frac{5}{2} - \frac{2}{3} \times \frac{3}{2} = 0$$

The simultaneous equation will be translated into the following matrices $\begin{pmatrix} \frac{2}{5} & \frac{2}{3} \\ \frac{3}{2} & \frac{5}{2} \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Because we already know that

$$\det(\begin{pmatrix} \frac{2}{5} & \frac{2}{3} \\ \frac{3}{2} & \frac{5}{2} \end{pmatrix}) = 0$$

We can say that there is no unique solution to these simultaneous equations.

2. First calculate the value when you multiply the determinants

$$\det(\begin{pmatrix} a & b \\ c & d \end{pmatrix}) \det(\begin{pmatrix} e & f \\ g & h \end{pmatrix}) =$$

$(ad - bc)(eh - fg) =$

$adeh - bceh - adfg + bcfg$

Now let's calculate C by doing the matrix multiplication first

$$\det(\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} e & f \\ g & h \end{pmatrix}) =$$

$$\det(\begin{pmatrix} ae + bg & af + bh \\ ce + dg & cf + dh \end{pmatrix}) =$$

$(ae + bg)(cf + dh) - (af + bh)(ce + dg) =$

$aecf + bgcf + aedh + bgdh - afce - bgce - afdg - bhdg =$

$bgcf + aedh - bgce - afdg$

Which is equal to the value we calculated when we multiplied the determinants, thus det(C) = det(A)det(B)

for the 2Ã-2 case.

3.

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

det(A) = ad - bc

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

det(A') = cb - da

- det(A') = - (bc - ad) = ad - bc

Thus det(A) = -det(A') is true.

4. a)

$A = P^{-1}BP$

det(A) = det(P^{-1})det(B)det(P) =

det(P^{-1})det(P)det(B) =

det($P^{-1}P$)det(B) =

det(I)det(B) =

det(B) as det(I) = 1.

thus det(A) = det(B) b) if $A^k = 0$ for some k it means that det(A^k) = 0. But we can write det(A^k) = det(A)^k, thus det(A)^k = 0. This means that det(A) = 0.

5. a)

A^5 =

$$\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} =$$

$$\left(\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\right)\left(\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix}\right)\begin{pmatrix} -1 & 6 \\ -1 & 4 \end{pmatrix} =$$

$$\begin{pmatrix}-5 & 18\\ -3 & 10\end{pmatrix}\begin{pmatrix}-5 & 18\\ -3 & 10\end{pmatrix}\begin{pmatrix}-1 & 6\\ -1 & 4\end{pmatrix} =$$

$$\begin{pmatrix}-5 & 18\\ -3 & 10\end{pmatrix}\begin{pmatrix}-13 & 42\\ -7 & 22\end{pmatrix} =$$

$$\begin{pmatrix}-61 & 186\\ -31 & 94\end{pmatrix}$$

b)

$$P^{-1} = \frac{1}{1}\begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix} = \begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix}$$

c)

$$\begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}1 & -2\\ -1 & 3\end{pmatrix} =$$

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

$$\begin{pmatrix}-1 & 6\\ -1 & 4\end{pmatrix}$$

d)

$A^5 =$

$$(P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}P)^5 =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}PP^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}PP^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}PP^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}PP^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}P =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}I\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}I\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}I\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}I\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}P =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}^5 P =$$

$$P^{-1}\begin{pmatrix}1^5 & 0\\ 0 & 2^5\end{pmatrix}P =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 32\end{pmatrix}P =$$

$$\begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 32\end{pmatrix}\begin{pmatrix}1 & -2\\ -1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 64\\ 1 & 32\end{pmatrix}\begin{pmatrix}1 & -2\\ -1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}-61 & 186\\ -31 & 94\end{pmatrix}$$

We see that P and it's inverse disappear when you raise the matrix to the fifth power. Thus you can see that we can calculate A^n very easily because you only have to raise the diagonal matrix to the n-th power. Raising diagonal matrices to a certain power is very easy because you only have to raise the numbers on the diagonal to that power.

f) We use the method derived in the exercise above.

$A^{100} =$

$$(P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}P)^{100} =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}^{100}P =$$

$$P^{-1}\begin{pmatrix}1^{100} & 0\\ 0 & 2^{100}\end{pmatrix}P =$$

$$P^{-1}\begin{pmatrix}1 & 0\\ 0 & 2^{100}\end{pmatrix}P =$$

$$\begin{pmatrix}3 & 2\\ 1 & 1\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 2^{100}\end{pmatrix}\begin{pmatrix}1 & -2\\ -1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3 & 2^{101}\\1 & 2^{100}\end{pmatrix}\begin{pmatrix}1 & -2\\-1 & 3\end{pmatrix} =$$

$$\begin{pmatrix}3-2^{101} & 3\times 2^{101}-6\\1-2^{100} & 3\times 2^{100}-2\end{pmatrix}$$

Problem set

Matrices Problem Set

At the moment, the main focus is on authoring the main content of each chapter. Therefore this exercise solutions section may be out of date and appear disorganised.

If you have a question please leave a comment in the "discussion section" or contact the author or any of the major contributors.

1.

$$\begin{pmatrix}2 & 3\\3 & 5\end{pmatrix}\begin{pmatrix}? & ? & ? & ?\\? & ? & ? & ?\end{pmatrix} = \begin{pmatrix}28 & 94 & 70 & 102\\44 & 153 & 112 & 163\end{pmatrix}$$

$$\begin{pmatrix}2 & 3\\3 & 5\end{pmatrix}^{-1}\begin{pmatrix}2 & 3\\3 & 5\end{pmatrix}\begin{pmatrix}? & ? & ? & ?\\? & ? & ? & ?\end{pmatrix} = \begin{pmatrix}2 & 3\\3 & 5\end{pmatrix}^{-1}\begin{pmatrix}28 & 94 & 70 & 102\\44 & 153 & 112 & 163\end{pmatrix}$$

$$\begin{pmatrix}? & ? & ? & ?\\? & ? & ? & ?\end{pmatrix} = \begin{pmatrix}2 & 3\\3 & 5\end{pmatrix}^{-1}\begin{pmatrix}28 & 94 & 70 & 102\\44 & 153 & 112 & 163\end{pmatrix}$$

$$= \frac{1}{2\times 5-3\times 3}\begin{pmatrix}5 & -3\\-3 & 2\end{pmatrix}\begin{pmatrix}28 & 94 & 70 & 102\\44 & 153 & 112 & 163\end{pmatrix}$$

$$= \begin{pmatrix}(5\times 28+(-3)\times 44) & (5\times 94+(-3)\times 153) & (5\times 70+(-3)\times 112) & (5\times 102+(-3)\times 163)\\((-3)\times 28+2\times 44) & ((-3)\times 94+2\times 153) & ((-3)\times 70+2\times 112) & ((-3)\times 102+2\times 163)\end{pmatrix}$$

$$= \begin{pmatrix}8 & 11 & 14 & 21\\4 & 24 & 14 & 20\end{pmatrix}$$

Therefore the message is "iloveyou"

2.

Combine the two matrices together, we have

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

$$A\begin{pmatrix}1 & 3\\ 2 & 4\end{pmatrix} = I$$

Therefore the inverse of $\boldsymbol{A}$ is

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

Further Modular Arithmetic

In this chapter we assume the reader can find inverses and be able to solve a system of congruences (Chinese Remainder Theorem) (see: Primes and Modular Arithmetic).

Introduction

Mathematics is the queen of the sciences and number theory is the queen of mathematics. -- Carl Friedrich Gauss 1777 - 1855

In the Primes and Modular Arithmetic section, we discussed the elementary properties of a prime and its connection to modular arithmetic. Our attention has been, for the most part, restricted to arithmetic mod p, where p is prime. But this need not be.

In this chapter, we will start by discussing some more elementary results in arithmetic modulo p, and then we will move on to discuss those results modulo m where m is composite. In particular, we will take a closer look at the Chinese Remainder Theorem, and how it allows us to break arithmetic modulo m into *components*. From that point of view, the CRT is an extremely powerful tool that can help us unlock the many secrets of modulo arithmetic (with relative ease).

Lastly, we will introduce the idea of an *abelian group* through multiplication in modular arithmetic and discuss the discrete log problem which underpins one of the most important cryptographic systems known today.

Wilson's Theorem

Wilson's theorem is a simple result that leads to a number of interesting observations in elementary number theory. It states that, if p is prime then

$$1 \cdot 2 \cdot 3 \cdots (p-1) \equiv p-1 \pmod p$$

We know the inverse of p - 1 is p - 1, so each other number can be paired up by its inverse and *eliminated*. For example, let p = 7, we consider

$$1 \times 2 \times .. \times 6 \equiv (2 \times 4) \times (3 \times 5) \times 1 \times 6 = 6$$

What we have done is that we paired up numbers that are inverses of each other, then we are left with numbers whose inverse is itself. It this case, they are 1 and 6.

But there is a technical difficulty. For a general prime number, p, how do we know that 1 and p - 1 are the only numbers in mod p which when squared give 1? For m not a prime, there are more than 2 solutions to $x^2 \equiv 1 \pmod m$, for example, let m = 15, then x = 1, 14, 4, 11 are solutions to $x^2 \equiv 1 \pmod m$.

We can not say Wilson's theorem is true, unless we show that there can only be (at most) two solutions to $x^2 \equiv 1 \pmod{p}$ when p is prime. We shall overcome this final hurdle by a simple *proof by contradiction* argument. You may want to skip the following proof and come up with your own justification of why Wilson's theorem is true.

Let p be a prime, and $x^2 \equiv 1 \pmod{p}$. We aim to prove that there can only be 2 solutions, namely x = 1, -1

$$\begin{aligned} x^2 - 1 &\equiv 0 \\ (x-1)(x+1) &\equiv 0 \end{aligned}$$

it obvious from the above that $x = 1, -1$ ($\equiv p - 1$) are solutions. Suppose there is another solution, $x = d$, and d not equal to 1 or -1. Since p is prime, we know $d - 1$ must have an inverse. We substitute x with d and multiply both sides by the inverse, i.e.

$$\begin{aligned} (d-1)(d+1) &= 0 \\ d+1 &= 0 \\ d &= -1 \end{aligned}$$

but we our initial assumption was that $d \neq -1$. This is a contradiction. Therefore there can only be 2 solutions to $x^2 \equiv 1 \pmod{p}$.

Fermat's little Theorem

There is a remarkable little theorem named after Fermat, the prince of amateur mathematicians. It states that if p is prime and given $a \neq 0$ then

$$a^{p-1} \equiv 1 \pmod{p}$$

This theorem hinges on the fact that p is prime. It won't work otherwise. How so? Recall that if p is prime then $a \neq 0$ has an inverse. So for any b and c we must have

$ab \equiv ac \pmod{p}$ if and only if $b \equiv c \pmod{p}$

A simple consequence of the above is that the following numbers must all be different mod p

$a, 2a, 3a, 4a, ..., (p-1)a$

and there are $p - 1$ of these numbers! Therefore the above list is just the numbers 1, 2, ... $p - 1$ in a different order. Let's see an example, take $p = 5$, and $a = 2$:

1, 2, 3, 4

multiply each of the above by 2 in mod 5, we get

2, 4, 1, 3

They are just the original numbers in a different order.

So for any p and using Wilson's Theorem (recall: $1 \times 2 \times ... \times (p-1) \equiv -1$),

, we get

$$\begin{aligned} a \cdot 2a \cdots (p-1)a &\equiv 1 \cdot 2 \cdots (p-1) \\ &\equiv -1 \end{aligned}$$

on the other hand we also get

$$\begin{aligned} a \cdot 2a \cdots (p-1)a &\equiv a^{p-1}(1 \cdot 2 \cdots (p-1)) \\ &\equiv -a^{p-1} \end{aligned}$$

Equating the two results, we get

$$-a^{p-1} \equiv -1$$

which is essentially Fermat's little theorem.

Modular Arithmetic with a general m

Chinese Remainder Theorem revisited

This section is rather theoretical, and is aimed at justifying the arithmetic we will cover in the next section. Therefore it is not necessary to fully understand the material here, and the reader may safely choose to skip the material below.

Recall the Chinese Remainder Theorem (CRT) we covered in the Modular Arithmetic section. In states that the following congruences

$$x \equiv b \pmod{n_1}$$

$$x \equiv c \pmod{n_2}$$

have a solution if and only if gcd(n_1,n_2) divides (b - c).

This deceptively simple theorem holds the key to arithmetic modulo m (not prime)! We shall consider the case where m has only two prime factors, and then the general case shall follow.

Suppose $m = p^i q^j$, where p and q are distinct primes, then every natural number below m (0, 1, 2, ..., m - 1) corresponds *uniquely* to a system of congruence mod p^i and mod q^j. This is due to the fact that gcd(p^i,q^j) = 1, so it divides all numbers.

Consider a number n, it corresponds to

$$n \equiv x_n \pmod{p^i}$$

$$n \equiv y_n \pmod{q^j}$$

for some x_n and y_n. If $r \neq n$ then r corresponds to

$$r \equiv x_r \pmod{p^i}$$

$$r \equiv y_r \pmod{q^j}$$

Now since r and n are different, we must have either $x_r \neq x_n$ and/or $y_r \neq y_n$

For example take $m = 12 = 2^2 \times 3$, then we can construct the following table showing the x_n, y_n for each n (0, 1, 2 ... 11)

n	***n* (mod 2^2)**	***n* (mod 3)**
0	0	0
1	1	1
2	2	2
3	3	0
4	0	1
5	1	2
6	2	0
7	3	1
8	0	2
9	1	0
10	2	1
11	3	2

Note that as predicted each number corresponds *uniquely* to two different systems of congruences mod 2^2 and mod 3.

Exercises

1. Consider $m = 45 = 3^2$ Ã- 5. Complete the table below and verify that any two numbers must differ in at least one place in the second and third column

n	***n* (mod 3^2)**	***n* (mod 5)**
0	0	0
1	1	1
2	2	2
...		
44	?	?

2. Suppose $m = p^i q^j$, n corresponds to

$$n \equiv x_n \pmod{p^i}$$

$$n \equiv y_n \pmod{q^j}$$

and r corresponds to

$$r \equiv x_r \pmod{p^i}$$

$$r \equiv y_r \pmod{q^j}$$

Is it true that

$$n + r \equiv x_n + x_r \pmod{p^i}$$

$$n + r \equiv y_n + y_r \pmod{q^j}$$

and that

$$nr \equiv x_n x_r \pmod{p^i}$$

$$nr \equiv y_n y_r \pmod{q^j}$$

Arithmetic with CRT

Exercise 2 above gave the biggest indication yet as to how the CRT can help with arithmetic modulo *m*. It is not essential for the reader to fully understand the above at this stage. We will proceed to describe how CRT can help with arithmetic modulo *m*. In simple terms, the CRT helps to break a modulo-*m* calculation into smaller calculations modulo prime factors of *m*. We will see what we mean very soon.

As always, let's consider a simple example first. Let $m = 63 = 3^2 \times 7$ and we see that m has two distinct prime factors. We should demonstrate multiplication of 51 and 13 modulo 63 in two ways. Firstly, the standard way

$$\begin{aligned} 51 \times 13 &= 663 \\ &= 10 \times 63 + 33 \\ &\equiv 33 \pmod{63} \end{aligned}$$

Alternatively, we notice that

$$51 \equiv 6 \pmod 9$$

and

$$51 \equiv 2 \pmod 7$$

We can represent the two expressions above as a two-tuple (6,2). We abuse the notation a little by writing 51 = (6,2). Similarly, we write 13 = (4,6). When we do multiplication with two-tuples, we multiply *component-wise*, i.e. (a,b) Ã- (c,d) = (ac,bd),

$$\begin{aligned} 51 \times 13 &= (6,2) \times (4,6) \\ &= (24,12) \\ &= (2 \times 9 + 6, 7 + 5) \\ &\equiv (6,5) \end{aligned}$$

Now let's solve

$$x \equiv 6 \pmod 9$$

and

$$x \equiv 5 \pmod 7$$

we write x = 6 + 9a, which is the first congruence equation, and then

$$\begin{aligned} 6 + 9a &\equiv 5 \pmod 7 \\ 2a &\equiv 6 \\ a &\equiv 3 \end{aligned}$$

therefore we have $a = 3 + 7b$, substitute back to get

$$x = 6 + 9(3 + 7b) = 33 + 63b \equiv 33 \pmod{63}$$

which is the same answer we got from multiplying 51 and 13 (mod 63) the standard way!

Let's summarise what we did. By representing the two numbers (51 and 13) as two two-tuples and multiplying *component-wise*, we ended up with another two-tuple. And this two-tuple corresponds to the product of the two numbers (mod m) via the Chinese Remainder Theorem.

We will do two more examples. Let $m = 88 = 2^3 \times 11$, and lets multiply 66 and 40 in two ways. Firstly, the standard way

$$\begin{aligned} 66 \times 40 &= 2640 \\ &= 30 \times 88 \\ &\equiv 0 \pmod{88} \end{aligned}$$

and now the second way, 40 = (0,7) and 66 = (4,0) and

$$\begin{aligned} 66 \times 40 &= (0,7) \times (4,0) \\ &= (0,0) \\ &\equiv 0 \pmod{88} \end{aligned}$$

For the second example, we notice that there is no need to stop at just two distinct prime factors. We let $m = 975 = 3 \times 5^2 \times 13$, and multiply 900 and 647 (mod 975),

$$\begin{aligned} 900 \times 647 &= 582300 \\ &\equiv 225 \pmod{975} \end{aligned}$$

For the other way, we note that 900 ≡ 0 (mod 3) ≡ 0 (mod 25) ≡ 3 (mod 13), and for 647 ≡ 2 (mod 3) ≡ 22 (mod 25) ≡ 10 (mod 13),

$$\begin{aligned} 900 \times 647 &= (0,0,3) \times (2,22,10) \\ &\equiv (0,0,30) \\ &\equiv (0,0,4) \end{aligned}$$

now if we solve the following congruences

$$x \equiv 0 \pmod{3}$$

$$x \equiv 0 \pmod{25}$$

$$x \equiv 4 \pmod{13}$$

then we will get $x \equiv 225$!

Why? If anything, breaking modular arithmetic in *m* into smaller components seems to be quite a bit of work. Take the example of multiplications, firstly, we need to express each number as a *n*-tuple (*n* is the number of distinct prime factors of *m*), multiply component-wise and then solve the resultant *n* congruences. Surely, it must be more complicated than just multiplying the two numbers and then reduce the result modulo *m*. So why bother studying it at all?

By breaking a number *m* into prime factors, we have gained insight into how the arithmetic really works. More importantly, many problems in modular *m* can be difficult to solve, but when broken into components it suddenly becomes quite easy, e.g. Wilson's Theorem for a general *m* (discussed below).

Exercises

1. Show that addition can also be done component-wise.

2. Multiply component-wise 32 and 84 (mod 134).

...

Euler totient

To discuss the more general form of Wilson's Theorem and Fermat's Little Theorem in mod *m* (not prime), it's nice to know a simple result from the famous mathematician Euler. More specifically, we want to discuss a function, called the Euler totient function (or Euler Phi), denoted φ.

The φ functions does a very simple thing. For any natural number *m*, $\varphi(m)$ gives the number of $n < m$, such that $\gcd(n,m) = 1$. In other words, it calculates how many numbers in mod *m* have an inverse. We will discuss the value of φ (m) for simple cases first and then derive the formula for a general *m* from the basic results.

For example, let $m = 5$, then $\varphi(m) = 4$. As 5 is prime, all non-zero natural numbers below 5 (1,2,3 and 4) are coprimes to it. So there are 4 numbers in mod 5 that have inverses. In fact, if *m* is prime then $\varphi(m) = m - 1$.

We can generalise the above to $m = p^r$ where *p* is prime. In this case, we try to employ a counting argument to calculate $\varphi(m)$. Note that there are p^r natural numbers below *m* (0, 1, 2 ... $p^r - 1$), and so $\varphi(m) = p^r$ - (number of $n < m$ such that $\gcd(n,m) \neq 1$). We did that because it is easier to count the number of *n* 's without an inverse mod *m*.

An element, *n*, in mod *m* does not have an inverse if and only if it shares a common factor with

m. But all factors of m (not equal to 1) are a multiple of p. So how many multiples of p are there in mod m? We can list them, they are

$$0, p, 2p, \cdots, p^r - p$$

where the last element can be written as $(p^{r-1} - 1)p$, and so there are p^{r-1} multiples of p. Therefore we can conclude

$$\phi(p^r) = p^r - p^{r-1}$$

We now have all the machinery necessary to derive the formula of φ (m) for any m.

By the Fundamental Theorem of Arithmetic, any natural number m can be uniquely expressed as the product of primes, that is

$$m = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

where p_i for i = 1, 2 ... r are distinct primes and k_i are positive integers. For example 1225275 = 3Ã-5^2Ã-17Ã-31^2. From here, the reader should try to derive the following result (the CRT may help).

Euler totient function φ

```
Suppose m can be uniquely expressed as below
```

$$m = p_1^{k_1} p_2^{k_2} \cdots p_r^{k_r}$$

```
then
```

$$\phi(m) = (p_1^{k_1} - p_1^{k_1-1})(p_2^{k_2} - p_2^{k_2-1}) \cdots (p_r^{k_r} - p_r^{k_r-1})$$

Wilson's Theorem

Wilson's Theorem for a general m states that the product of all the invertible element in mod m

equals -1 if m has only one prime factor, or $m = 2p^k$ for some prime p

equals 1 for all other cases

An invertible element of mod m is a natural number $n < m$ such that gcd(n, m) = 1. A self-invertible element is an element whose inverse is itself.

In the proof of Wilson's Theorem for a prime p, the numbers 1 to p - 1 all have inverses. This is not true for a general m. In fact it is certain that $(m - 1)! \equiv 0 \pmod{m}$, for this reason we instead consider the product of all invertible elements in mod m.

For the case where $m = p$ is prime we also appealed to the fact 1 and p - 1 are the only elements when squared gives 1. In fact for $m = p^k$, 1 and m - 1 (≡-1)are the only self-invertible elements (see exercise). But for a general m, this is not true. Let's take for example m = 21. In arithmetic modulo 21 each of the following numbers has itself as an inverse

1, 20, 8, 13

so how can we say the product of all invertible elements equal to 1?

We use the CRT described above. Let us consider the case where $m = 2p^k$. By the CRT, each element in mod m can be represented as a two tuple (a,b) where a can take the value 0 or 1 while b can take the value 0, 1, ..., or $p^k - 1$. Each two tuple corresponds uniquely to a pair of congruence equations and multiplication can be performed component-wise.

Using the above information, we can easily list all the self-invertible elements, because $(a,b)^2 \equiv 1$ means $(a^2,b^2) = (1,1)$, so a is an invertible element in mod 2 and b is an invertible element in mod p^k, so $a \equiv 1$ or -1, $b \equiv 1$ or -1. But in mod 2 $1 \equiv -1$, so $a = 1$. Therefore, there are two elements that are self invertible in mod $m = 2p^k$, they are (1,1) = 1, and (1, -1) = m - 1 . So in this case, the result is the same as when m has only a single prime factor.

For the case where m has more than one prime factors and $m \neq 2p^k$. Let say m has n prime factors then m can be represented as a n-tuple. Let say m has 3 distinct prime factors, then all the self-invertible elements of m are

14. (1,1,1)
15. (1,1,-1)
16. (1,-1,1)
17. (1,-1,-1)
18. (-1,1,1)
19. (-1,1,-1)
20. (-1,-1,1)
21. (-1,-1,-1)

their product is (1,1,1) which corresponds to 1 in mod m.

Exercise

1. Let p be a prime. Show that in arithmetic modulo p^k, 1 and $p^k - 1$ are the only self-invertible elements.

...more to come

Fermat's Little Theorem

As mentioned in the previous section, not every element is invertible (i.e. has an inverse) mod *m*. A generalised version of Fermat's Last Theorem uses Euler's Totient function, it states

$$a^{\phi(m)} \equiv 1 \pmod{m}$$

for all $a \neq 0$ satisfying gcd(a,m) = 1. This is easy to see from the generalised version of Wilson's Theorem. We use a similar techique from the prove of Fermat's Little Theorem. We have

$$(ab_1)(ab_2)\cdots(ab_{\phi(m)}) \equiv b_1 b_2 \cdots b_{\phi(m)} \pmod{m}$$

where the b_i's are all the invertible elements mod *m*. By Wilson's theorem the product of all the invertible elements equals to, say, *d* (= 1 or -1). So we get

$$a^{\phi(m)} d \equiv d \pmod{m}$$

whihc is essentially the statement of Fermat's Little Theorem.

Although the FLT is very neat, it is *imprecise* in some sense. For example take *m* = 15 = 3 Ã- 5, we know that if *a* has an inverse mod 15 then $a^{\varphi(15)} = a^8 \equiv 1 \pmod{15}$. But 8 is too large, all we need is 4, by that we mean, $a^4 \equiv 1 \pmod{15}$ for all *a* with an inverse (*the reader can check*).

The Carmichael function Î»(*m*) is the smallest number such that $a^{\text{Î»(m)}} \equiv 1 \pmod{m}$. A question in the Problem Set deals with this function.

Exercises

...more to come

Two-torsion Factorisation

It it quite clear that factorising a *large* number can be extremely difficult. For example, given that 76372591715434667 is the product of two primes, can the reader factorise it? Without the help of a good computer algebra software, the task is close to being impossible. As of today, there is no known *efficient* all purpose algorithm for factorising a number into prime factors.

However, under certain special circumstances, factorising can be easy. We shall consider the two-torsion factorisation method. A 2-torsion element in modular *m* arithmetic is a number *a* such that $a^2 \equiv 1 \pmod{m}$.

Let's consider an example in arithmetic modulo 21. Note that using the CRT we can represent any number in mod 21 as a two-tuple. We note that the two-torsion elements are 1 = (1,1), 13 = (1,-1), 8 = (-1,1) and 20 = (-1,-1). Of interest are the numbers 13 and 8, because 13 + 1 = (1,-1)

+ (1,1) = (2,0). Therefore 13 + 1 = 14 is an element sharing a common factor with 21, as the second component in the two-tuple representation of 14 is zero. Therefore GCD(14,21) = 7 is a factor of 21.

The above example is very silly because anyone can factorise 21. But what about 24131? Factorising it is not so easy. But, if we are given that 12271 is a non-trivial (i.e. ≠ 1 or -1) two-torsion element, then we can conclude that both gcd(12271 + 1,24131) and gcd(12271 - 1,24131) are factors of 24131. Indeed gcd(12272,24131) = 59 and gcd(12270,24131) = 409 are both factors of 24131.

More generally, let *m* be a composite, and *t* be a non-trivial two-torsion element mod *m* i.e. $t \neq 1, -1$. Then

gcd(t + 1,m) divides *m*, and

gdc(t - 1,m) divides *m*

this can be explained using the CRT.

We shall explain the case where $m = pq$ and *p* and *q* are primes. Given *t* is a non-trivial two-torsion element, then *t* has representaion (1,-1) or (-1,1). Suppose t = (-1,1) then t + 1 = (-1,1) + (1,1) = (0,2), therefore t + 1 must be a multiple of *p* therefore gcd(t,m) = p. In the other case where t - 1 = (-1,1) - (1,1) = (-2,0) and so gcd(t - 1,m) = q.

So if we are given a non-trivial two-torsion element then we have effectively found one (and possibly more) prime factors, which goes a long way in factorising the number. In most modern public key cryptography applications, to break the system we need only to factorise a number with two prime factors. In that regard two-torsion factorisation method is frightening effectively.

Of course, finding a non-trivial two-torsion element is not an easy task either. So internet banking is still safe for the moment. By the way 76372591715434667 = 224364191 Ã-340395637.

Exercises

1. Given that 18815 is a two-torsion element mod 26176. Factorise 26176.

...more to come'

Mathematical programming

Before we begin

This chapter will not attempt to teach you how to program rigorously. Therefore a basic working knowledge of the C programming language is highly recommended. It is recommended that you learn as much about the C programming language as possible before learning the materials in this chapter.

Please read the first 7 lessons of "C Programming Tutorial" by About.com

http://cplus.about.com/library/blctut.htm

if you are unfamiliar with programming or the C programming language.

Introduction to programming

Programming has many uses. Some areas where programming and computer science in general are extremely important include artificial intelligence and statistics. Programming allows you to use computers flexibly and process data very quickly.

When a program is written, it is written into a textual form that a human can understand. However, a computer doesn't directly understand what a human writes. It needs to be transformed into a way that the computer can directly understand.

For example, a computer is like a person who reads and speaks German. You write and speak in English. The letter you write to the computer needs to be translated for the computer to speak. The program responsible for this work is referred to as the *compiler*.

You need to *compile* your English-like instructions, so that the computer can understand it. Once a program has been compiled, it is hard to "un-compile" it, or transform it back into English again. A programmer writes the program (to use our analogy, in English), called *source code*, which is a human-readable definition of the program, and then the *compiler* translates this into "machine code". We recommend using the widely available gcc compiler.

When we look at mathematical programming here, we will look at how we can write programs that will solve some difficult mathematical problems that would take us normally a lot of time to solve. For example, if we wanted to find an approximation to the root of the polynomial x^5+x+1 - this is very difficult for a human to solve. However a computer can do this no sweat -- how?

Programming language basics

We will be using the C programming language throughout the chapter, please learn about the basics of C by reading the first seven chapters of "C programming Tutorial" at About.com

http://cplus.about.com/library/blctut.htm

Discrete Programming

Discrete programming deals with integers and how they are manipulated using the computer.

Understanding integral division

In C, the command

int number;

number = 3 / 2;

will set aside some space in the computer memory, and we can refer to that space by the *variable* name **number**. In the computer's mind, number is an integer, nothing else. After

number = 3 / 2;

numbers equals 1, not 1.5, this is due to that fact that **/** when applied to two integers will give only the integral part of the result. For example:

5 / 2 equals 2

353 / 3 equals 117

99 / 7 equals 14

-19 / 2 equals -9

78 / -3 equals -29

in C.

Exercises Evaluate *x*

a) x = 7 / 2

b) x = -9 / -4

c) x = 1000 / 999

d) x = 2500 / 2501

Modelling Recursively defined functions

The factorial function n! is recursively defined:

0! = 1

n! = nÃ-(n-1)! if n ≠¥ 1

In C, if fact(n) is the functions as described above we want

fact(0) = 1;

fact(*n*) = *n* * *fact*(*n* âˆ’ 1); if $n \geq 1$

we should note that all recursively defined functions have a *terminating condition*, it is the case where the function can give a direct answer to, e.g. fact(0) = 1.

We can model the factorial functions easily with the following code and then execute it:

```
int fact (int n)

{

if (n == 0)

return 1;

if (n >= 1)

return n * fact(n - 1);

}
```

The C function above models the factorial function very naturally. To test the results, we can compile the following code:

```
#include <stdio.h> /* STanDard Input & Output Header file */

int fact (int n)

{

if (n == 0)

return 1;

if (n >= 1)

return n * fact(n - 1);

}

void main()

{

int n = 5;

printf("%d", fact(n)); /* printf is defined in stdio.h */
```

}

We can also model the Fibonacci number function. Let fib(n) return the (n + 1)th Fibonacci number, the following should be clear

fib(0) should return 1

fib(1) should return 1

fib(n) should return fib(n - 1) + fib(n - 2); for n ≥ 2

we can model the above using C:

int fib (int n)

{

if (n == 0 || n == 1) /* if n = 0 or if n = 1 */

return 1;

if (n >= 2)

return fib(n - 1) + fib(n - 2);

}

Again, you shall see that modelling a recursive function is not hard, as it only involves translating the mathematics in C code.

Modelling non-recursive functions

There are functions that involve only integers, and they can be modelled quite nicely using *functions* in C. The factorial function

f(n) = n! = n(n - 1)(n - 2) ... 3Ã-2Ã-1

can be modelled quite simply using the following code

```
int n = 10;  //get factorial for 10
int f = 1;   //start f at 1
while(n > 0) //keep looping if n bigger then 0
{
   f = n * f; //f is now product of f and n
   n = n - 1; //n is one less (repeat loop)
}
```

Feedback

What do you think? Too easy or too hard? Too much information or not enough? How can we improve? Please let us know by leaving a comment in the discussion section. Better still, edit it yourself and make it better.

Basic counting

Counting

All supplementary chapters contain materials that are part of the standard high school mathematics curriculum, therefore the material is only provided for completeness and should mostly serve as revision.

Ordered Selection

Suppose there are 20 songs in your mp3 collection. The computer is asked to randomly selected 10 songs and play them in the order they are selected, how many ways are there to select the 10 songs? This type of problems is called ordered selection counting, as the order in which the things are selected is important. E.g. if one selection is

1, 2, 3, 4, 5, 6, 7, 8, 9 and 10

then

2, 1, 3, 4, 5, 6, 7, 8, 9 and 10

is considered a different selection.

There are 20 ways to choose the first song since there are 20 songs, then there are 19 ways to choose the second song and 18 ways to choose the third song ... and so on. Therefore the total number of ways can be calculated by considering the following product:

20 Ã- 19 Ã- 18 Ã- 17 Ã- 16 Ã- 15 Ã- 14 Ã- 13 Ã- 12 Ã- 11

or denoted more compactly:

$$\frac{20!}{10!}$$

Here we use the *factorial* function, defined by 0! = 1 and $n! = (n-1)! \times n$. (In other words, $n! = 1 \times 2 \times 3 \times \ldots \times n$)

In general, the number of ordered selections of *m* items out of *n* items is:

$$\frac{n!}{(n-m)!}$$

The idea is that we cancel off all but the first *m* factors of the *n*! product.

Unordered Selection

Out of the 15 people in your mathematics class, five will be chosen to represent the class in a school wide mathematics competition. How many ways are there to choose the five students? This problem is called an unordered selection problem, i.e. the order in which you select the students is *not* important. E.g. if one selection is

Joe, Lee, Sue, Britney, Justin

another selection is

Lee, Joe, Sue, Justin, Britney

the two selections are considered equivalent.

There are

$$\frac{15!}{10!}$$

ways to choose the 5 candidates in *ordered selection*, but there are 5! *permutations* of the same five candidates. (That is, 5! different permutations are actually the same combination). Therefore there are

$$\frac{15!}{10!5!}$$

ways of choosing 5 students to represent your class.

In general, to *choose* (unordered selection) *m* candidates from *n*, there are

$$\frac{n!}{m!(n-m)!} = \binom{n}{m}$$

ways. We took the formula for ordered selections of *m* candidates from *n*, and then divided by *m*! because each unordered selection was counted as *m*! ordered selections.

Note: $\binom{n}{m}$ is read "*n* choose *m*".

Examples

Example 1 How many different ways can the letters of the word BOOK be arranged?

Solution 1 4! ways if the letters are all distinct. Since O is repeated twice, there are 2! permutations. Therefore there are 4!/2! = 12 ways.

Example 2 How many ways are there to choose 5 diamonds from a deck of cards?

Solution 2 There are 13 diamonds in the deck. So there are $\binom{13}{5}$ ways.

$$\binom{13}{5} = \frac{13!}{8!5!} = \frac{13 \times 12 \times 11 \times 10 \times 9}{120} = 1287$$

Binomial expansion

The binomial expansion deals with the expansion of following expression

$(a + b)^n$

Take n = 3 for example, we shall try to expand the expression manually we get

$(a + b)^3 = (aa + ab + ba + bb)(a + b)$

$= aaa + aab + aba + abb + baa + bab + bba + bbb$

$= aaa + aab + aba + abb + baa + bab + bba + bbb$

We deliberately did not simplify the expression at any point during the expansion, we didn't even use the well known $(a + b)^2 = a^2 + 2ab + b^2$. As you can see, the final expanded form has 8 terms. They are all the possible terms of powers of *a* and *b* with three factors!

Since there are 3 factors in each term and all the possibles terms are in the expanded expressoin. How many terms are there with only one *b*? The answer should be $\binom{3}{1}$, i.e. from 3 possible positions, choose 1 for *b*. Similarly we can work out all the coefficient of like-terms. So

$(a + b)^3$

And more generally

$(a + b)^n$

or more compactly using the summation sign (otherwise known as sigma notation)

$(a + b)^n$

Partial fractions

Method of Partial Fractions

All supplementary chapters contain materials that are part of the standard high school mathematics curriculum, therefore the material is only provided for completeness and should mostly serve as revision.

Introduction

Before we begin, consider the following: $\frac{1}{1\times 2}+\frac{1}{2\times 3}+\frac{1}{3\times 4}......+\frac{1}{99\times 100}$

How do we calculate this sum? At first glance it may seem difficult, but if you think carefully you will find: $\frac{1}{4\times 5}=\frac{5-4}{4\times 5}=\frac{5}{4\times 5}-\frac{4}{4\times 5}=\frac{1}{4}-\frac{1}{5}$

Thus the original problem can be rewritten as follows,

$$=\frac{1}{1}-\frac{1}{2}+\frac{1}{2}-\frac{1}{3}+\frac{1}{3}-\frac{1}{4}......-\frac{1}{99}+\frac{1}{99}-\frac{1}{100}$$

So all terms except the first and the last cancelled out, and therefore

$$=1-\frac{1}{100}=\frac{99}{100}$$

In fact, you've just done partial fractions! Partial fractions is a method of breaking down complex fractions that involve products into sums of simpler fractions.

Method

So, how do we do partial fractions? Look at the example below: $\frac{4z-5}{z^2-3z+2}$

Factorize the denominator.

$$\frac{4z - 5}{(z-1)(z-2)}$$

Then we suppose we **can** break it down into the fractions with denominator (z-1) and (z-2) respectively. We let their numerators be a and b.

$$\frac{4z-5}{(z-1)(z-2)} \equiv \frac{a}{z-1} + \frac{b}{z-2}$$

$$\frac{4z-5}{(z-1)(z-2)} \equiv \frac{a(z-2)}{(z-1)(z-2)} + \frac{b(z-1)}{(z-1)(z-2)}$$

$$\frac{4z-5}{(z-1)(z-2)} \equiv \frac{az-2a+bz-b}{(z-1)(z-2)}$$

$$\frac{4z-5}{(z-1)(z-2)} \equiv \frac{(a+b)z-(2a+b)}{(z-1)(z-2)}$$

$$4z-5 \equiv (a+b)z-(2a+b)$$

Therefore by matching coefficients of like power of z, we have:

$$\begin{cases} a+b=4 & ...(1) \\ 2a+b=5 & ...(2) \end{cases}$$

(2)-(1):a=1

Substitute a=1 into (1):b=3

Therefore

$$\frac{4z-5}{z^2-3z+2} = \frac{1}{z-1} + \frac{3}{z-2}$$

(Need Exercises!)

More on partial fraction

Repeated factors

On the last section we have talked about factorizing the denominator, and have each factor as the denominators of each term. But what happens when there are repeating factors? Can we apply the same method? See the example below:

$$\frac{4x-1}{(x+2)^2(x-1)}$$

$$\equiv \frac{A}{x+2}+\frac{B}{x+2}+\frac{C}{x-1}$$

$$\equiv \frac{A+B}{x+2}+\frac{C}{x-1}$$

$$\equiv \frac{(A+B)(x-1)}{(x+2)(x-1)}+\frac{C(x+2)}{(x+2)(x-1)}$$

$$\equiv \frac{(A+B)(x-1)+C(x+2)}{(x+2)(x-1)}$$

$$\equiv \frac{(A+B+C)x+(2C-A-B)}{(x+2)(x-1)}$$

Indeed, a factor is missing! Can we multiply both the denominator and the numerator by that factor? No! Because the numerator is of degree 1, multiplying with a linear factor will make it become degree 2! (You may think:can't we set A+B+C=0? Yes, but by substituting A+B=-C, you will find out that this is impossible)

From the above failed example, we see that the old method of partial fraction seems not to be working. You may ask, can we actually break it down? Yes, but before we finally attack this problem, let's look at the denominators at more detail.

Consider the following example:
$$\frac{1}{2^3 7^2}+\frac{1}{2^5 7}=\frac{2^2}{2^5 7^2}+\frac{7}{2^5 7^2}=\frac{2^2+7}{2^5 7^2}$$
We can see that the power of a prime factor in the product denominator is the maximum power of that prime factor in all term's denominator.

Similarly, let there be factor $P_1,P_2,...,P_n$, then we may have in general case:

$$\frac{A}{P_1^{\alpha_1}P_2^{\alpha_2}...P_n^{\alpha_n}}+\frac{B}{P_1^{\beta_1}P_2^{\beta_2}...P_n^{\beta_n}}+...\frac{Z}{P_1^{\zeta_1}P_2^{\zeta_2}...P_n^{\zeta_n}}$$

If we turn it into one big fraction, the denominator will be: $P_1^{max(\alpha_1,\beta_1,...,\zeta_1)}P_2^{max(\alpha_2,\beta_2,...,\zeta_2)}...P_n^{max(\alpha_n,\beta_n,...,\zeta_n)}$

Back to our example, since the factor (x+2) has a power of 2, at least one of the term has $(x+2)^2$ as the denominator's factor. You may then try as follows:

$$\frac{4x-1}{(x+2)^2(x-1)}$$

$$\equiv\frac{A}{(x+2)^2}+\frac{B}{x-1}$$

$$\equiv\frac{A(x-1)}{(x+2)^2(x-1)}+\frac{B(x+2)^2}{(x+2)^2(x-1)}$$

$$\equiv\frac{A(x-1)+B(x+2)^2}{(x+2)^2(x-1)}$$

$$\equiv\frac{Ax-A+Bx^2+4Bx+4B}{(x+2)^2(x-1)}$$

$$\equiv\frac{Bx^2+(A+4B)x+(4B-A)}{(x+2)^2(x-1)}$$

But again, we can't set B=0, since that would means the latter term is 0! What is missing? To handle it properly, let's use a table to show all possible combinations of the denominator:

Possible combinations of denominator

Power of (x+2)	Power of (x-1)	Result	Used?
0	0	1	Not useful
1	0	(x+2)	Not used
2	0	(x+2)^2	Used
0	1	(x-1)	Used
1	1	(x+2)(x-1)	Not useful

2	1	(x+2)^2(x-1)	Not useful

So, we now know that X/(x+2) is missing, we can finally happily get the answer:

$$\frac{4x-1}{(x+2)^2(x-1)}$$

$$\equiv \frac{A}{(x+2)^2}+\frac{B}{x+2}+\frac{C}{x-1}$$

$$\equiv \frac{A(x-1)}{(x+2)^2(x-1)}+\frac{B(x+2)(x-1)}{(x+2)^2(x-1)}+\frac{C(x+2)^2}{(x+2)^2(x-1)}$$

$$\equiv \frac{A(x-1)+B(x^2+x-2)+C(x^2+4x+4)}{(x+2)^2(x-1)}$$

$$\equiv \frac{(B+C)x^2+(A+B+4C)x-(A+2B-4C)}{(x+2)^2(x-1)}$$

Therefore by matching coefficient of like power of x, we have

As a conclusion, for a repeated factor of power n, we will have n terms with their denominator being X^n, X^(n-1), ...,X^2, X

Works continuing, don't distrub :)

Alternate method for repeated factors

Other than the method suggested above, we would like to use another approach to handle the problem. We first leave out some factor to make it into non-repeated form, do partial fraction on it, then multiply the factor back, then apply partial fraction on the 2 fractions.

$$\frac{4x-1}{(x+2)^2(x-1)}$$

$$\equiv \frac{1}{x+2}\times\frac{4x-1}{(x+2)(x-1)}$$

Then we do partial fraction on the latter part:

$$\frac{4x-1}{(x+2)(x-1)} \equiv \frac{A}{x+2} + \frac{B}{x-1}$$

$$\frac{4x-1}{(x+2)(x-1)} \equiv \frac{A(x-1)}{(x+2)(x-1)} + \frac{B(x+2)}{(x+2)(x-1)}$$

$$\frac{4x-1}{(x+2)(x-1)} \equiv \frac{A(x-1)+B(x+2)}{(x+2)(x-1)}$$

$$\frac{4x-1}{(x+2)(x-1)} \equiv \frac{(A+B)x+(2B-A)}{(x+2)(x-1)}$$

$$4x-1 \equiv (A+B)x+(2B-A)$$

By matching coefficients of like powers of x, we have

$$\begin{cases} A+B=4 & \ldots(1) \\ 2B-A=-1 & \ldots(2) \end{cases}$$

Substitute A=4-B into (2),

$$2B-(4-B)=-1$$

Hence B = 1 and A = 3.

We carry on:

$$\equiv \frac{1}{x+2} \times \left(\frac{3}{x+2} + \frac{1}{x-1}\right)$$

$$\equiv \frac{3}{(x+2)^2} + \frac{1}{(x+2)(x-1)}$$

Now we do partial fraction once more:

$$\frac{1}{(x+2)(x-1)} \equiv \frac{A}{x+2} + \frac{B}{x-1}$$

$$\frac{1}{(x+2)(x-1)} \equiv \frac{A(x-1)}{(x+2)(x-1)} + \frac{B(x+2)}{(x+2)(x-1)}$$

$$\frac{1}{(x+2)(x-1)} \equiv \frac{A(x-1)+B(x+2)}{(x+2)(x-1)}$$

$$\frac{1}{(x+2)(x-1)} \equiv \frac{(A+B)x+(2B-A)}{(x+2)(x-1)}$$

$$0x+1 \equiv (A+B)x+(2B-A)$$

By matching coefficients of like powers of x , we have:

$$\begin{cases} A+B=0 & ...(1) \\ 2B-A=1 & ...(2) \end{cases}$$

Substitute A=-B into (2), we have:

2B-(-B) = 1

Hence B=1/3 and A=-1/3

So finally,

$$\frac{4x-1}{(x+2)^2(x-1)} \equiv \frac{3}{(x+2)^2} - \frac{1}{3(x+2)} + \frac{1}{3(x-1)}$$

Summation sign

Summation Notation

All supplementary chapters contain materials that are part of the standard high school mathematics curriculum, therefore the material is only provided for completeness and should mostly serve as revision.

We normally use the "+" sign to represent a sum, but if the sum expression involved is complex and long, it can be confusing.

For example: $\frac{1}{1\times 2}+\frac{1}{2\times 3}+\frac{1}{3\times 4}......+\frac{1}{100\times 101}$

Writing the above would be a tedious and messy task!

To represent expression of this kind more compactly and nicely, people use the summation notation, a capital greek letter "Sigma". On the right of the sigma sign people write the expression of each term to sum, and write the upper and lower limit of the variable on top and under the sigma sign.

Example 1: $\sum_{k=3}^{10} 2k+1$

$= \ (2(3)+1) \ + \ (2(4)+1) \ + \ (2(5)+1) \ + ... + \ (2(10)+1)$
$= \ 7 \ + \ 9 \ + \ 11 \ + ... + \ 21$

Misconception: From the above there is a common misconception that the number on top of the Sigma sign is the number of terms. This is wrong. The number on top is the number to subsitute back in the last term.

Example 2: $\frac{1}{4}-\frac{1}{9}+\frac{1}{16}-\frac{1}{25}......-\frac{1}{9801}+\frac{1}{10000}$

$=\frac{1}{2^2}-\frac{1}{3^2}+\frac{1}{4^2}-\frac{1}{5^2}......-\frac{1}{99^2}+\frac{1}{100^2}$
$=\sum_{k=2}^{100}(-1)^k\frac{1}{k^2}$

Tip:If the terms alternate between plus and minus, we can use the sequence $(-1)^k = -1,\ 1,\ -1,\ 1...$

Exercise

14. Use the summation notation to represent the expression in the first example.

Change the following into sum notation:

6. $23 + 24 + 25 + 26 + \ldots\ldots + 1927$

7. $13 + 16 + 19 + 22 + \ldots\ldots + 301$

8. $*1 - 2 - 3 + 4 + 5 - 6 - 7 + 8 \ldots\ldots + 400$ (Hint:reorder the terms, or get more than one term in the expression)

9. *

$$1000 - \frac{3}{1 \times (1+3+5)} - \frac{5}{(1+3) \times (1+3+5+7)} - \frac{7}{(1+3+5) \times (1+3+5+7+9)} \ldots\ldots$$

(Hint:You need to use more than one sigma sign)

Change the following sum notation into the normal representation:

11. $$\pi = 4 \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1}$$

12. $$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}$$

(Need more exercise,especially "reading" sigma notation and change back into the old form)

Operations of sum notation

Although most rules related to sum makes sense in the ordinary system, in this new system of sum notation, things may not be as clear as before and therefore people summarize some rules related to sum notation (see if you can identify what they correspond to!)

- $$\sum_{i=p}^{q} A_i \pm c \; = \; \pm(q-p+1)c \; + \; \sum_{i=p}^{q} A_i$$
- $$\sum_{i=p}^{q} A_i \pm B_i \; = \; \sum_{i=p}^{q} A_i \pm \sum_{i=p}^{q} B_i$$
- $$\sum_{i=p}^{q} cA_i \; = \; c\sum_{i=p}^{q} A_i$$

- $$\sum_{i=p}^{q}\left[\sum_{j=r}^{s}A_{ij}\right] = \sum_{j=r}^{s}\left[\sum_{i=p}^{q}A_{ij}\right]$$

  (Note:I suggest getting a visual aid on this one:showing that you can sum a two dimensional array in either direction)

- $$\sum_{i=p}^{q}A_i = \sum_{i=p-k}^{q-k}A_{i+k}$$ (Index substutition)

- $$\sum_{i=p}^{q}A_i = \sum_{i=p}^{r}A_i + \sum_{i=r+1}^{q}A_i \text{ , } where\ p \le r < q$$ (Decomposition)

- $$\left(\sum_{i=p}^{q}a_i\right) \times \left(\sum_{j=r}^{s}b_j\right) = \sum_{i=p}^{q}\sum_{j=r}^{s}a_ib_j$$ (Factorization/Expansion)

Exercise

(put up something here please)

Beyond

"To iterate is human; to recurse, divine."

When human repeated summing, they have decided to use a more advanced concept, the concept of product. And of course everyone knows we use $\times$. And when we repeat product, we use exponential. Back to topic, we now have a notation for complex sum. What about complex product? In fact, there is a notation for product also. We use the capital greek letter "pi" to denote product, and basically everything else is exactly the same as sum notation, except that the terms are not summed, but multiplied.

Example: $\prod_{h=2}^{5} 2h - 3$ = $[(2 \times 2) - 3] \times [(2 \times 3) - 3] \times [(2 \times 4) - 3] \times [(2 \times 5) - 3]$

Exercise

1.It has been known that the factorial is defined inductively by:
$0! = 1$
$n! \times (n+1) = (n+1)!$
Now try to define it by product notation.
(more to go...)

Complex numbers

Introduction

All supplementary chapters contain materials that are part of the standard high school mathematics curriculum, therefore the material is only provided for completeness and should mostly serve as revision.

Although the real numbers can, in some sense, represent any natural quantity, they are in another sense incomplete. We can write certain types of equations with real number coefficients which we desire to solve, but which have no real number solutions. The simplest example of this is the equation:

$$\begin{aligned} x^2+1 &= 0 \\ x^2 &= -1 \\ x &= \sqrt{-1} \end{aligned}$$

Your high school math teacher may have told you that there is no solution to the above equation. He/she may have even emphasised that there is no *real* solution. But we can, in fact, extend our system of numbers to include the *complex* numbers by declaring the solution to that equation to exist, and giving it a name: the *imaginary unit*, *i*.

Let's *imagine* for this chapter that $i = \sqrt{-1}$ exists. Hence $x = i$ is a solution to the above question, and $i^2 = -1$.

A valid question that one may ask is "Why?". Why is it important that we be able to solve these quadratics with this seemingly artificial construction? It is interesting delve a little further into the reason why this imaginary number was introduced in the first place - it turns out that there was a valid reason why mathematicians realized that such a construct was useful, and could provide deeper insight.

The answer to the question lies not in the solution of quadratics, but rather in the solution of the intersection of a cubic and a line. The mathematician Cardano managed to come up with an ingenious method of solving cubics - much like the quadratic formula, there is also a formula that gives us the roots of cubic equations, although it is far more complicated. Essentially, we can express the solution of a cubic $x^3 = 3px + 2q$ in the form

$$x = \sqrt[3]{q+\sqrt{q^2-p^3}} + \sqrt[3]{q-\sqrt{q^2-p^3}}$$

An unsightly expression, indeed!

You should be able to convince yourself that the line $y = 3px + 2q$ must always hit the cubic $y = x^3$. But try solving some equation where $q^2 < p^3$, and you run into a problem - the problem is that we are forced to deal with the square root of a negative number. But, we know that in fact there is a solution for x; for example, $x^3 = 15x + 4$ has the solution x = 4.

It became apparent to the mathematician Bombelli that there was some piece of the puzzle that was missing - something that explained how this seemingly perverse operation of taking a square root of a negative number would somehow simplify to a simple answer like 4. This was in fact the motivation for considering imaginary numbers, and opened up a fascinating area of mathematics.

The topic of **Complex numbers** is very much concerned with this number *i*. Since this number doesn't exist in this real world, and only lives in our imagination, we call it the *imaginary unit*. (Note that *i* is not typically chosen as a variable name for this reason.)

The imaginary unit

As mentioned above

$$i^2 = -1.$$

Let's compute a few more powers of *i*:

$$\begin{aligned} i^1 &= i \\ i^2 &= -1 \\ i^3 &= -i \\ i^4 &= 1 \\ i^5 &= i \\ i^6 &= -1 \\ &\dots \end{aligned}$$

As you may see, there is a pattern to be found in this.

Exercises

12. Compute i^{25}

13. Compute i^{100}

14. Compute i^{1000}

Exercise Solutions

Complex numbers as solutions to quadratic equations

Consider the quadratic equation:

$$\begin{aligned} x^2 - 6x + 13 &= 0 \\ x &= \frac{6 \pm \sqrt{36 - 4 \times 13}}{2} \\ x &= \frac{6 \pm \sqrt{-16}}{2} \\ x &= \frac{6 \pm \sqrt{-1}\sqrt{16}}{2} \\ x &= \frac{6 \pm 4i}{2} \\ x &= 3 + 2i \;,\; 3 - 2i \end{aligned}$$

The *x* we get as a solution is what we call a complex number. (To be nitpicky, the solution set of this equation actually has two complex numbers in it; either is a valid value for x.) It consists of *two* parts: a *real* part of 3 and an *imaginary* part of ± 2. Let's call the real part *a* and the imaginary part *b*; then the sum $a + bi = 3 \pm 2i$ is a complex number.

Notice that by merely defining the square root of negative one, we have already given ourselves the ability to assign a value to a much more complicated, and previously unsolvable, quadratic equation. It turns out that 'any' polynomial equation of degree *n* has exactly *n* zeroes if we allow complex numbers; this is called the Fundamental Theorem of Algebra.

We denote the *real* part by *Re*. E.g.:

$\mathrm{Re}(x) = 3$

and the *imaginary* part by *Im*. E.g.:

$$\mathrm{Im}(x) = \pm 2$$

Let's check to see whether $x = 3 + 2i$ really is solution to the equation:

$$\begin{aligned} x &= 3 + 2i \\ x^2 &= (3)^2 + 2(3)(2i) + (2i)^2 \\ &= 5 + 12i \\ x^2 - 6x + 13 &= 5 + 12i - 6(3 + 2i) + 13 \\ &= 0 \end{aligned}$$

Exercises

8. Convince yourself that $x = 3 - 2i$ is also a solution to the equation.

9. Plot the points A(3, 2) and B(3, -2) on a XY plane. Draw a line for each point joining them to the origin.

10. Compute the length of AO (the distance from point A to the Origin) and BO. Denote them by r_1 and r_2 respectively. What do you observe?

11. Compute the angle between each line and the x-axis and denote them by φ_1 and φ_2. What do you observe?

12. Consider the complex numbers:

$$\begin{aligned} z &= r_1 \cdot (\cos\phi_1 + i\sin\phi_1) \\ w &= r_2 \cdot (\cos\phi_2 + i\sin\phi_2) \end{aligned}$$

Substitute *z* and *w* into the quadratic equation above using the values you have computed in Exercise 3 and 4. What do you observe? What conclusion can you draw from this?

Arithmetic with complex numbers

Addition and multiplication

Adding and multiplying two complex number together turns out to be quite straightforward. Let's illustrate with a few examples. Let $x = 3 - 2i$ and $y = 7 + 11i$, and we do addition first

$$\begin{aligned} x + y &= (3 + 7) + (-2 + 11)i \\ &= 10 + 9i \end{aligned}$$

and now multiplication

$$\begin{aligned} x \times y &= (3 - 2i)(7 + 11i) \\ &= 3 \cdot 7 + 3 \cdot 11i - 2i \cdot 7 - 2 \cdot 11i^2 \\ &= 43 + 19i \end{aligned}$$

Let's summarise the results here.

- When adding complex numbers we add the real parts with real parts, and add the imaginary parts with imaginary parts.
- When multiplying two complex numbers together, we use normal expansion. Whenever we see i^2 we put in its place -1. We then collect like terms.

But how do we calculate:

$$\frac{3 + 2i}{7 - \sqrt{5}i}$$

Note that the square root is only above the 5 and not the *i*. This is a little bit tricky, and we shall cover it in the next section.

Exercises:

$$\begin{aligned} x &= 3-2i \\ y &= 3+2i \end{aligned}$$

Compute:

6. x + y
7. x - y
8. x^2
9. y^2
10. xy
11. (x + y)(x - y)

Division

One way to calculate:

$$\frac{1}{2\sqrt{3}+\sqrt{2}}$$

is to rationalise the denominator:

$$\frac{1}{2\sqrt{3}+\sqrt{2}} = \frac{2\sqrt{3}-\sqrt{2}}{(2\sqrt{3}+\sqrt{2})(2\sqrt{3}-\sqrt{2})} = \frac{2\sqrt{3}-\sqrt{2}}{10}$$

Utilising a similar idea, to calculate

$$\frac{3+2i}{7-\sqrt{5}i}$$

we *real*ise the denominator.

$$z = \frac{3+2i}{7-\sqrt{5}i}$$

$$z = \frac{3+2i}{7-\sqrt{5}i} \times \frac{7+\sqrt{5}i}{7+\sqrt{5}i}$$

The denominator is the sum of two squares. We get:

$$z = \frac{(3+2i)\times(7+\sqrt{5}i)}{49+5}$$

$$z = \frac{21-2\sqrt{5}}{54} + \frac{14+3\sqrt{5}}{54}i$$

If somehow we can always find a complex number whose product with the denominator is a real number, then it's easy to do divisions.

If

$$z = a + ib$$

and

$$w = a - ib$$

Then *zw* is a real number. This is true for any 'a' and 'b' (provided they are real numbers).

Exercises

Convince yourself that the product of *zw* is always a real number.

Complex Conjugate

The exercise above leads to the idea of a complext conjugate. The complex congugate of *a + ib* is *a - ib*. For example, the conjugate of *2 + 3i* is *2 - 3i*. It is a simple fact that the product of a complex number and its conjugate is always a real number. If *z* is a complex number then its conjugate is denoted by $\bar{z}$. Symbolically if

z = a + ib

then,

$$\bar{z} = a - ib$$

The conjugate of *3 - 9i* is *3 + 9i*.

The conjugate of *100* is *100*.

The conjugate of *9i - 20* is *-20 - 9i*.

Conjugate laws

Here are a few simple rules regarding the complex congugate

$$\overline{z+w} = \bar{z} + \bar{w}$$

and

$$\overline{zw} = \bar{z}\bar{w}$$

The above laws simply says that the sum of congugates equals the congugate of the sum; and similarly, the congugate of the product equals the product of the congugates.

Consider this example:

(3 + 2*i*) + (89 âˆ’ 100*i*) = 92 âˆ’ 98*i*

and we can see that

$$\overline{92 - 98i} = 92 + 98i$$

which equals to

$$\overline{3+2i} + \overline{89 - 100i} = 3 - 2i + 89 + 100i = 92 + 98i$$

This confirms the addition conjugate law.

Exercise

Convince yourself that the multiplication law is also true.

The complex root

Now that you are equipped with all the basics of complex numbers, you can tackle the more advanced topic of root finding.

Consider the question:

$$\begin{aligned} z &= -3 + 4i \\ w &= \sqrt{z} \end{aligned}$$

Express *w* in the form of *a* + *ib*.

That is easy enough.

$$\begin{aligned} w &= \sqrt{z} \\ w^2 &= z \\ w &= a + ib \\ w^2 &= a^2 - b^2 + 2abi \end{aligned}$$

$$\begin{aligned} -3 &= a^2 - b^2 && (1) \\ 4 &= 2ab && (2) \end{aligned}$$

Solve (1) and (2) simultaneously to work out *a* and *b*.

Observe that if, after solving for *a* and *b*, we replace them with -*a* and -*b* respectively, then they would still satisfy the two simultaneous equations above, we can see that (as expected) if $w = a + ib$ satisfies the equation $w^2 = z$, then so will $w = -(a + ib)$. With real numbers, we always take the non-negative answer and call the solution $\sqrt{x}$. However, since there is no notion of "greater than" or "less than" with complex numbers, there is no such choice of $\sqrt{z}$. In fact, which square root to take as "the" value of $\sqrt{z}$ depends on the circumstances, and this choice is very important to some calculations.

info -- Finding the square root

Finding the root of a real number is a very difficult problem to start with. Most people have no hope of finding a close estimate of $\sqrt{2}$ without the help of a calculator. The modern method of approximating roots involves an easy to understand and ingenius piece of mathematics called the Taylor series expansion. The topic is usually covered in first year university maths as it requires an elementary understanding of an important branch of mathematics called calculus. The Newton-Raphson method of root finding is also used extensively for this purpose.

Now consider the problem

$$\begin{aligned} z &= -2 + 2i \\ w &= z^{1/3} \end{aligned}$$

Express *w* in the form of "$a + ib$".

Using the methodology developed above we proceed as follows,

$$\begin{aligned}
w &= z^{1/3} \\
w^3 &= z
\end{aligned}$$

$$\begin{aligned}
w &= (a+ib) \\
w^3 &= (a^2-b^2+2abi)\times(a+ib) \\
z &= (a^3-3ab^2)+i(3a^2b-b^3)
\end{aligned}$$

$$\begin{aligned}
-2 &= a^3-3ab^2 \qquad (1) \\
2 &= 3a^2b-b^3 \qquad (2)
\end{aligned}$$

It turns out that the simultaneous equations (1) & (2) are hard to solve. Actually, there is an easy way to calculate the roots of complex numbers called the De Moivre's theorem, it allows us to calculate the *n*th root of any complex number with ease. But to set the method, we need understand the geometric meaning of a complex number and learn a new way to *represent* a complex number.

Exercises

9. Find $(3 + 3i)^{1/2}$
10. Find $(1 + 1i)^{1/2}$
11. Find $i^{1/3}$

The complex plane

Complex numbers as ordered pairs

It is worth noting, at this point, that every complex number, *a* + *bi*, can be completely specified with exactly two real numbers: the *real part a*, and the *imaginary part b*. This is true of *every* complex number; for example, the number 5 has real part 5 and imaginary part 0, while the number 7*i* has real part 0 and imaginary part 7. We can take advantage of this to adopt an alternative scheme for writing complex numbers: we can write them as ordered pairs, in the form *(a, b)* instead of *a+bi*.

Instead of	We could write
$5+4i$	$(5,4)$
$3i$	$(0,3)$
$\frac{4+5i}{3}$	$(\frac{4}{3},\frac{5}{3})$
42	$(42,0)$
$\sqrt{2}+\sqrt{2}i$	$(\sqrt{2},\sqrt{2})$

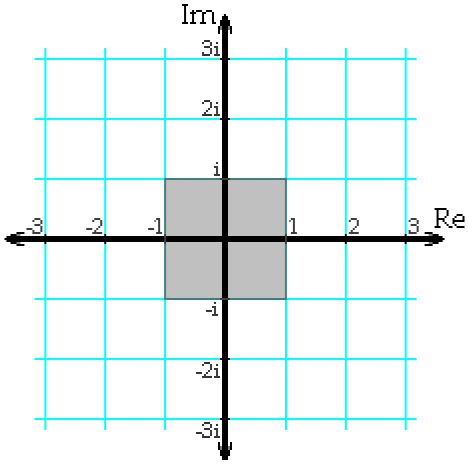
These should look familiar: they are exactly like the ordered pairs we use to represent poins in the plane. In fact, we can use them that way; the plane which results is called the *complex plane*. We refer to its x axis as the *real axis*, and to its y axis as the *imaginary axis*.

The complex plane

We can see from the above that a single complex number is a point in the complex plane. We can also represent *sets* of complex numbers; these will form *regions* on the plane. For example, the set

$$\{a + bi| - 1 \leq a \leq 1, -1 \leq b \leq 1\}$$

is a square of edge length 2 centered at the origin (See following diagram).



Complex-valued functions

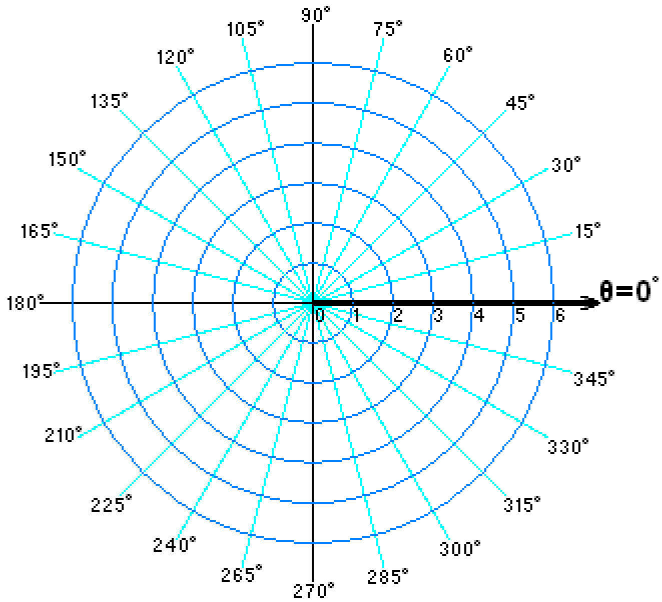
Just as we can make functions which take *real* values and output *real* values, so we can create functions from complex numbers to real numbers, or from complex numbers to complex numbers. These latter functions are often referred to as *complex-valued* functions, because they evaluate to (output) a complex number; it is implicit that their argument (input) is complex as well.

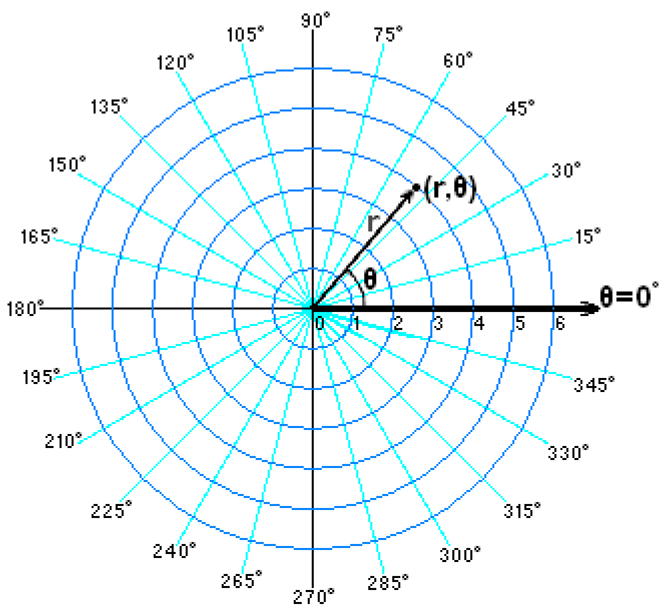
Since complex-valued functions map complex numbers to other complex numbers, and we have already seen that complex numbers correspond to points on the complex plane, we can see that a complex-valued function can turn regions on the complex plane into other regions. A simple example: the function

$f(z) = z + (0 + 1i)$

takes a point in the complex plane and shifts it up by 1. If we apply it to the set of points making up the square above, it will move the entire square up one, so that it "rests" on the x-axis.

{To make more complicated examples, I will first have to go back and introduce the polar representation of complex numbers. Makes for much more interesting functions, :-) You can use the diagrams below or modify them to make new diagrams. I will make links to these diagrams in other places in Wikibooks:math. In the 2nd diagram showing the point, r=4 and theta= 50 degrees. These types of diagrams can be used to introduce phasors, which are notations for complex numbers used in electrical engineering.}





de Moivre's Theorem

If

$z = re^{i\text{Î¸}} = r(\cos(\text{Î¸}) + i\sin(\text{Î¸}))$

then

$z^n = r^n(\cos(n\text{Î¸}) + i\sin(n\text{Î¸}))$

Complex root of unity

The complex roots of unity to the nth degree is the set of solutions to the equation x^n = 1. Clearly they all have magnitude 1. They form a cyclic group under multiplication. For any given n, there are exactly n many of them, and they form a regular n-gon in the complex plane over the unit circle.

A closed form solution can be given for them, by use of Euler's formula: u^n = {cos(2*pi*j/n)+i*sin(2*pi*j/n) | 0 <= j < n}

The sum of the nth roots of unity is equal to 0, except for n=1, where it is equal to 1.

The product of the nth roots of unity alternates between -1 and 1.

Problem set

Simplify: $(1-i)^{2i}$ Ans: $2^i e^{-\pi/2}$

The imaginary unit

22. Compute $i^{25} = i$

23. Compute $i^{100} = 1$

24. Compute $i^{1000} = 1$

$$i^{25} = i^{24} \times i^{1} = i$$

The pattern of $i^1, i^2, i^3, \ldots$ shows that $i^{4n} = 1$ where n is any integer. This case applies to questions 2 and 3. For question 1, .

Complex numbers as solutions to quadratic equations

15. Convince yourself that $x = 3 - 2i$ is also a solution to the equation.

$$\begin{aligned} x &= 3-2i \\ x^2 &= (3)^2 + 2(3)(-2i) + (-2i)^2 \\ &= 5-12i \\ x^2 - 6x + 13 &= 5 - 12i - 6(3-2i) + 13 \\ &= 0 \end{aligned}$$

10 on a XY plane. Draw a line for each point joining

11. r_1 and r_2 ...pute the length of AO (the distance from point A to the Origin) and BO. Denote them by respectively. What do you observe?

- $r_1 = \sqrt{3^2 + 2^2} = \sqrt{13}$
- $r_2 = \sqrt{3^2 + (-2)^2} = \sqrt{13}$
- r_1 and r_2 are the same

12 ϕ_1 and ϕ_2 ute the angle between each line and the x-axis and denote them by . What do you

observe?

$$\phi_1 = \tan^{-1}(3/2) \approx 56$$

$$\phi_2 = \tan^{-1}(3/-2) \approx -56$$

-

ϕ_1 and ϕ_2

-
- differ only in the sign of the number

$$\begin{aligned} z &= r_1 \cdot (\cos\phi_1 + i\sin\phi_1) \\ w &= r_2 \cdot (\cos\phi_2 + i\sin\phi_2) \end{aligned}$$

13. Consider the complex numbers:

Substitute *z* and *w* into the quadratic equation above using the values you have computed in Exercise 3 and 4. What do you observe? What conclusion can you draw from this?

$$\begin{aligned} z &= \sqrt{13} \cdot (\cos\tan^{-1}(3/2) + i\sin\tan^{-1}(3/2)) &= 2+3i \\ w &= \sqrt{13} \cdot (\cos\tan^{-1}(3/-2) + i\sin\tan^{-1}(3/-2)) &= 2-3i \end{aligned}$$

Thus the quadratic equation will equal 0 since z and w are equal to the solutions we found when solving the equation.

Addition and multiplication

$$\begin{aligned} x &= 3-2i \\ y &= 3+2i \end{aligned}$$

Compute:

13. x + y

- (3 - 2i) + (3 + 2i) = 6

14. x - y

- 3 - 2i - (3 + 2i) = -4i

15. x^2

- (3 - 2i)(3 - 2i) = 9 + (2)(3)(-2i) + $4i^2$ = 5 - 12 i

16. y^2

- (3 + 2i)(3 + 2i) = 9 + (2)(3)(2i) + $4i^2$ = 5 + 12 i

17. xy

- (3 - 2i)(3 + 2i) = 9 + 6 i - 6i + $4i^2$ = 5

18. (x + y)(x - y)

- ((3 - 2i) + (3 + 2i))((3 - 2i) - (3 + 2i)) = (6)(-4i) = -24i

Division Convince yourself that the product of zw is always a real number. $zw = (a + bi)(a - bi) = a^2 - abi + abi - b^2i^2 = a^2 - b$

Complex Conjugate Convince yourself that the multiplication law is also true.

$$\begin{aligned}
z &= a + bi \\
w &= c - di \\
\\
x &= z \times w \\
&= (a + bi)(c - di) \quad = \quad ac + bci - adi - bdi^2 \quad = \quad ac - bd + (bc - ad)i \\
\bar{x} &= ac - bd - (bc - ad)i \\
\\
\bar{z} \times \bar{w} &= (a - bi)(c + di) \\
&= ac - bic + adi - bdi^2 \\
&= ac - bd - (bc + ad)i \\
&= \bar{x}
\end{aligned}$$

The complex root

1. Find $(3 + 3i)^{1/2}$

$$\begin{aligned}
z &= 3+3i \\
w &= \sqrt{z} \\
w^2 &= z \\
w &= a+ib \\
w^2 &= a^2-b^2+2abi \\
\\
3 &= a^2-b^2 && (1) \\
3 &= 2ab && (2) \\
\\
a &= \tfrac{3}{2b} \\
\\
3 &= (\tfrac{3}{2b})^2-b^2 \\
&= \tfrac{9}{4b^2}-b^2 \\
0 &= \tfrac{9}{4b^2}-3-b^2 \\
&= \tfrac{9}{4}-3b^2-b^4 \\
\\
c &= b^2 \\
0 &= \tfrac{9}{4}-3c-c^2 \\
c &= \frac{--3\pm\sqrt{(-3)^2-4\times\frac{9}{4}\times-1}}{2\times-1} \\
&= \frac{-3\pm\sqrt{18}}{2} \\
b^2 &= \frac{-3\pm\sqrt{18}}{2} \\
b &= \sqrt{\frac{-3+\sqrt{18}}{2}} \\
&= \sqrt{\frac{-3+\sqrt{18}}{2}} \\
&= \sqrt{\frac{-3+3\sqrt{2}}{2}} \\
&= \sqrt{\frac{3(\sqrt{2}-1)}{2}} \\
&= \sqrt{\tfrac{3}{2}}\sqrt{(\sqrt{2}-1)} \\
&= \tfrac{\sqrt{6}}{2}\sqrt{(\sqrt{2}-1)} \\
&= \frac{\sqrt{6(\sqrt{2}-1)}}{2}
\end{aligned}$$

$$\begin{aligned}
3 &= 2ab \\
a &= \frac{3}{2b} \\
&= \frac{3}{2\frac{\sqrt{6(\sqrt{2}-1)}}{2}} \\
&= \frac{3}{\sqrt{6}}\times\frac{1}{\sqrt{\sqrt{2}-1}} \\
&= \frac{\sqrt{6}}{2}\times\frac{1}{\sqrt{\sqrt{2}-1}} \\
&= \frac{\sqrt{6}}{2}\times\frac{\sqrt{\sqrt{2}+1}}{\sqrt{\sqrt{2}-1\times\sqrt{\sqrt{2}+1}}} \\
&= \frac{\sqrt{6}}{2}\times\sqrt{\sqrt{2}+1} \\
&= \frac{\sqrt{6(\sqrt{2}+1)}}{2}
\end{aligned}$$

Thus the solution for $(3 + 3i)^{1/2}$ is:

$$\sqrt{3+3i} = \pm(\frac{\sqrt{6(\sqrt{2}+1)}}{2} + \frac{\sqrt{6(\sqrt{2}-1)}}{2} \times i)$$

$$\begin{aligned}
z &= 1 + 1i \\
w &= \sqrt{z} \\
w^2 &= z \\
w &= a + ib \\
w^2 &= a^2 - b^2 + 2abi \\
\\
1 &= a^2 - b^2 \qquad (1) \\
1 &= 2ab \qquad (2) \\
\\
a &= \tfrac{1}{2b} \\
\\
1 &= (\tfrac{1}{2b})^2 - b^2 \\
&= \tfrac{1}{4b^2} - b^2 \\
0 &= \tfrac{1}{4b^2} - 1 - b^2 \\
&= \tfrac{1}{4} - b^2 - b^4 \\
\\
c &= b^2 \\
0 &= \tfrac{1}{4} - 1c - c^2 \\
c &= \frac{--1\pm\sqrt{(-1)^2-4\times\frac{1}{4}\times-1}}{2\times-1} \\
&= \frac{-1\pm\sqrt{2}}{2} \\
b^2 &= \frac{-1\pm\sqrt{2}}{2} \\
b &= \sqrt{\frac{-1+\sqrt{2}}{2}} \\
&= \frac{\sqrt{\sqrt{2}-1}}{\sqrt{2}} \\
&= \frac{\sqrt{\sqrt{2}-1\times\sqrt{2}}}{\sqrt{2}\times\sqrt{2}} \\
&= \frac{\sqrt{2(\sqrt{2}-1)}}{2}
\end{aligned}$$

Thus the solution for $(1 + 1i)^{1/2}$ is:

$$\sqrt{1+1i} = \pm(\frac{\sqrt{2(\sqrt{2}+1)}}{2} + \frac{\sqrt{2(\sqrt{2}-1)}}{2} \times i)$$

$$\begin{array}{rcl} z & = & i \\ w & = & z^{1/3} w^3 \quad = \quad z \end{array}$$

$$\begin{array}{rcl} w & = & (a+ib) \\ w^3 & = & (a^2 - b^2 + 2abi) \times (a+ib) \\ z & = & (a^3 - 3ab^2) + i(3a^2b - b^3) \end{array}$$

$$\begin{array}{rcll} 0 & = & a^3 - 3ab^2 & (1) \\ 1 & = & 3a^2b - b^3 & (2) \end{array}$$

3. Find $i^{1/3}$

Thus the solution for $i^{1/3}$ is:

$$1 = \sqrt[3]{i} = \pm(\frac{\sqrt{3}}{2} + \frac{1}{2} \times i)$$

$$\begin{array}{rcl} 0 & = & a^3 - 3ab^2 \\ a^3 & = & 3ab^2 \\ a^2 & = & 3b^2 \end{array}$$

$$\begin{array}{rcl} 1 & = & 3a^2b - b^3 \\ 1 & = & 3(3b^2)b - b^3 \\ 1 & = & 9b^3 - b^3 \\ 1 & = & 8b^3 \\ \frac{1}{8} & = & b^3 \\ b & = & \sqrt[3]{\frac{1}{8}} \\ & = & \frac{1}{2} \end{array}$$

$$\begin{array}{rcl} a^2 & = & 3b^2 \\ & = & 3(\frac{1}{2})^2 \\ & = & \frac{3}{4} \\ a & = & \sqrt{\frac{3}{4}} \\ & = & \frac{\sqrt{3}}{2} \end{array}$$

Bases

15. The following numbers are written in base 2. Write them out in base 10:

1. 101011 (base 2) = 43 (base 10) = $2^5 + 2^3 + 2^1 + 2^0$= 32+8+2+1

2. 001101 = 13

3. 10 = 2

4. 011 = 3

16. Write those numbers out in base 10 as if they were originally in base 5.

1. 101011 (base 5) = 3256 (base 10) = $5^5 + 5^3 + 5^1 + 5^0$ = 3125+125+5+1

2. 001101 = 751

3. 10 = 25

4. 011 = 26

17. How many numbers could I write out in base 5 with only the first 4 columns?

Answer: 625 = 5^4 (each new column multiplies the number of possibilities by 5)

13. In computing, each 1 or 0 is called a *bit*. They are stored in groups of 8. Each group is called a *byte*. How many bytes are possible?

Answer: 256 = 2^8 (values from 00000000 through 11111111 binary, or 0 through 255 base 10)

12. Question: When editing the bytes directly, writing out 10110001 is too long and hexadecimal is used instead (in this case, B1). How many digits of hexadecimal are needed to cover all possible bytes?

Answer: 2 digits (16^2 = number of possible 2-digit hexadecimal numbers = 256 = number of possible bytes)

Differentiation

Differentiate from first principle

All supplementary chapters contain materials that are part of the standard high school mathematics curriculum, therefore the material is only provided for completeness and should mostly serve as revision. This section and the *differentiation technique* section can be skipped if you are already familiar with calculus/differentiation.

In calculus, differentiation is a very important operation applied to functions of real numbers. To differentiate a function f(x), we simply evaluate the limit

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

where the $\lim_{h \to 0}$ means that we let *h* approach 0. However, for now, we can simply think of it as putting *h* to 0, i.e., letting *h* = 0 at an appropriate time. There are various notations for the result of differentiation (called the derivative), for example

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

and

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

mean the same thing. We say, f'(x) is the derivative of f(x). Differentiation is useful for many purposes, but we shall not discuss why calculus was invented, but rather how we can apply calculus to the study of generating functions.

It should be clear that if

g(x) = f(x)

then

g'(x) = f'(x)

the above law is important. If g(x) a closed-form of f(x), then it is valid to differentiate both sides to obtain a new generating function.

Also if

h(x) = g(x) + f(x)

then

$h'(x) = g'(x) + f'(x)$

This can be verified by looking at the properties of limits.

Example 1

Differentiate from first principle f(x) where

$f(x) = x^2$

Firstly, we form the *difference quotient*

$$f'(x) = \lim_{h\to 0} \frac{(x+h)^2 - x^2}{h}$$

We can't set *h* to 0 to evaluate the limit at this point. Can you see why? We need to expand the quadratic first.

$$= \lim_{h\to 0} \frac{x^2 + 2xh + h^2 - x^2}{h}$$

$$= \lim_{h\to 0} \frac{2xh + h^2}{h}$$

We can now factor out the *h* to obtain now

$$\lim_{h\to 0} 2x + h$$

from where we can let *h* go to zero safely to obtain the derivative, $2x$. So

$f'(x) = 2x$

or equivalently:

$(x^2)' = 2x$

Example 2

Differentiate from first principles, $p(x) = x^n$.

We start from the difference quotient:

$$p'(x) = \lim_{h\to 0} \frac{(x+h)^n - x^n}{h}$$

By the binomial theorem, we have:

$$= \lim_{h \to 0} \frac{1}{h}(x^n + nx^{n-1}h + ... + h^n - x^n)$$

The first x^n cancels with the last, to get

$$= \lim_{h \to 0} \frac{1}{h}(nx^{n-1}h + ... + h^n)$$

Now, we bring the constant 1/*h* inside the brackets

$$= \lim_{h \to 0} nx^{n-1} + ... + h^{n-1}$$

and the result falls out:

$= nx^{n - 1}$

Important Result

If

$p(x) = x^n$

then

$p'(x) = nx^{n - 1}$

As you can see, differentiate from first principle involves working out the derivative of a function through algebraic manipulation, and for that reason this section is algebraically very difficult.

Example 3

Assume that if

h(x) = f(x) + g(x)

then

h'(x) = f'(x) + g'(x)

Differentiate $x^2 + x^5$

Solution Let $h(x) = x^2 + x^5$

$h'(x) = 2x + 5x^4$

Example 4

Show that if g(x) = AÃ-f(x) then

g'(x) = AÃ-f'(x)

Solution

$$\begin{aligned} g(x) &= Af(x) \\ g'(x) &= \lim_{h\to 0} \tfrac{A}{h}(f(x+h) - f(x)) \\ &= A\lim_{h\to 0} \tfrac{1}{h}(f(x+h) - f(x)) \\ &= Af'(x) \end{aligned}$$

Example 5

Differentiate from first principle

$f(x) = \frac{1}{1-x}$

Solution

$$\begin{aligned} f'(x) &= \lim_{h\to 0} \tfrac{1}{h}\left(\tfrac{1}{1-(x+h)} - \tfrac{1}{1-x}\right) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left(\tfrac{1-x-(1-(x+h))}{(1-(x+h))(1-x)}\right) \\ &= \lim_{h\to 0} \tfrac{1}{h}\left(\tfrac{h}{(1-(x+h))(1-x)}\right) \\ &= \lim_{h\to 0} \tfrac{1}{(1-(x+h))(1-x)} \\ &= \tfrac{1}{(1-x)^2} \end{aligned}$$

Exercises

1. Differentiate

$f(z) = z^3$

2. Differentiate

$f(z) = (1 - z)^2$

3. Differentiate from first principle

$$f(z) = \frac{1}{(1-z)^2}$$

4. Differentiate

$f(z) = (1 - z)^3$

5. Prove the result assumed in example 3 above, i.e. if

$f(x)=g(x)+h(x)$

then

$f'(x)=g'(x)+h'(x)$.

Hint: use limits.

Differentiating f(z) = (1 - z)^n

We aim to derive a vital result in this section, namely, to derive the derivative of

$f(z) = (1 - z)^n$

where $n \geq 1$ and n an integer. We will show a number of ways to arrive at the result.

Derivation 1

Let's proceed:

$f(z) = (1 - z)^n$

expand the right hand side using binomial expansion

$$f(z) = 1 - \binom{n}{1} z + \binom{n}{2} z^2 + ... + (-1)^n z^n$$

differentiate both sides

$$f'(z) = -\binom{n}{1} + \binom{n}{2}2z + ... + (-1)^n n z^{n-1}$$

now we use $\binom{n}{i} = \frac{n!}{i!(n-i)!}$

$$f'(z) = -\frac{n!}{1!(n-1)!} + \frac{n!}{2!(n-2)!}2z + ... + (-1)^n n z^{n-1}$$

and there are some cancelling

$$f'(z) = -\frac{n!}{1!(n-1)!} + \frac{n!}{1!(n-2)!}z + ... + (-1)^n n z^{n-1}$$

take out a common factor of -n, and recall that 1! = 0! = 1 we get

$$f'(z) = -n(1 + \frac{n-1!}{1!(n-2)!}z + ... + (-1)^{n-1} z^{n-1})$$

let j = i - 1, we get

$$f'(z) = -n(1 + \frac{n-1!}{1!(n-2)!}z + ... + (-1)^{n-1} z^{n-1})$$

but this is just the expansion of $(1 - z)^{n-1}$

$f'(z) = -n(1 - z)^{n-1}$

Derivation 2

Similar to Derivation 1, we use instead the definition of a derivative:

$$f'(z) = \lim_{h \to 0} \frac{(1-(z+h))^n - (1-z)^n}{h}$$

expand using the binomial theorem

$$f'(z) = \lim_{h \to 0} \frac{\sum_{i=0}^{n} \binom{n}{i}(-1)^i (z+h)^i - \sum_{i=0}^{n} \binom{n}{i}(-1)^i z^i}{h}$$

factorise

$$f'(z) = \lim_{h \to 0} \frac{\sum_{i=0}^{n} \binom{n}{i} (-1)^i ((z+h)^i - z^i)}{h}$$

take the limit inside (recall that [Af(x)]' = Af'(x))

$$f'(z) = \sum_{i=0}^{n} \binom{n}{i} (-1)^i \lim_{h \to 0} \frac{(z+h)^i - z^i}{h}$$

the inside is just the derivative of z^i

$$f'(z) = \sum_{i=1}^{n} \binom{n}{i} (-1)^i i z^{i-1}$$

exactly as derivation 1, we get

$f'(z) = -n(1 - z)^{n-1}$

Example Differentiate $(1 - z)^2$

Solution 1

$f(z) = (1 - z)^2 = 1 - 2z + z^2$

$f'(z) = -2 + 2z$

$f'(z) = -2(1 - z)$

Solution 2 By the result derived above we have

$f'(z) = -2(1 - z)^{2-1} = -2(1 - z)$

Exercises

Imitate the method used above or otherwise, differentiate:

1. $(1 - z)^3$
2. $(1 + z)^2$
3. $(1 + z)^3$
4. (Harder) $1/(1 - z)^3$ (Hint: Use definition of derivative)

Differentiation technique

We will teach how to differentiate functions of this form:

$$f(z) = \frac{1}{g(z)}$$

i.e. functions whose reciprocals are also functions. We proceed, by the definition of differentiation:

$$f(z) = \frac{1}{g(z)}$$

$$\begin{aligned} f'(z) &= \lim_{h \to 0} \tfrac{1}{h}\left(\tfrac{1}{g(z+h)} - \tfrac{1}{g(z)}\right) \\ &= \lim_{h \to 0} \tfrac{1}{h}\left(\tfrac{g(z)-g(z+h)}{g(z+h)g(z)}\right) \\ &= \lim_{h \to 0} \tfrac{g(z+h)-g(z)}{h} \tfrac{-1}{g(z+h)g(z)} \\ &= \lim_{h \to 0} g'(z) \tfrac{-1}{g(z+h)g(z)} \\ &= -\tfrac{g'(z)}{g(z)^2} \end{aligned}$$

Example 1

$$\begin{aligned} \tfrac{1}{1-z} &= 1 + z + z^2 + z^3 + \ldots \\ \left(\tfrac{1}{1-z}\right)' &= 1 + 2z + 3z^2 + \ldots \end{aligned}$$

by

$$\left(\frac{1}{g}\right)' = \frac{-g'}{g^2}$$

where *g* is a function of *z*, we get

$$\tfrac{1}{(1-z)^2} = 1 + 2z + 3z^2 + \ldots$$

which confirmed the result derived using a counting argument.

Exercises

Differentiate

1. $1/(1-z)^2$

2. $1/(1-z)^3$

3. $1/(1+z)^3$

4. Show that $(1/(1 - z)^n)' = n/(1-z)^{n+1}$

Differentiation applied to generating functions

Now that we are familiar with differentiation from first principle, we should consider:

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

we know

$$\frac{1}{1-x^2} = 1 + x^2 + x^4 + x^6 + ...$$

differentiate both sides

$$\left(\tfrac{1}{1-x^2}\right)' = 2x + 4x^3 + 6x^5 + ...$$

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

therefore we can conclude that

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

Note that we can obtain the above result by the substituion method as well,

$$\frac{1}{(1-z)^2} = 1 + 2z + 3z + ...$$

letting $z = x^2$ gives you the require result.

The above example demonstrated that we need not concern ourselves with difficult differentiations. Rather, to get the results the easy way, we need only to differentiate the basic forms and apply the substitution method. By basic forms we mean generating functions of the form:

$$\frac{1}{(1-z)^n}$$

for n ≥ 1.

Let's consider the number of solutions to

$a_1 + a_2 + a_3 + ... + a_n = m$

for $a_i \geq 0$ for i = 1, 2, ... n.

We know that for any *m*, the number of solutions is the coefficient to:

$$(1 + x + x^2 + ...)^n = \frac{1}{(1-z)^n}$$

as discussed before.

We start from:

$$\frac{1}{1-z} = 1 + x + x^2 + ... + x^n + ...$$

differentiate both sides (note that 1 = 1!)

$$\frac{1!}{(1-z)^2} = 1 + 2x + 3x^2 ... + nx^{n-1} + ...$$

differentiate again

$$\frac{2!}{(1-z)^3} = 2 + 2 \times 3x ... + n(n-1)x^{n-2} + ...$$

and so on for (n-1) times

$$\frac{(n-1)!}{(1-z)^n} = (n-1)! + \frac{n!}{1!}x + \frac{(n+1)!}{2!}x^2 + \frac{(n+2)!}{3!}x^3 + ...$$

divide both sides by (n-1)!

$$\frac{1}{(1-z)^n} = 1 + \frac{n!}{(n-1)!1!}x + \frac{(n+1)!}{(n-1)!2!}x^2 + \frac{(n+2)!}{(n-1)!3!}x^3 + ...$$

the above confirms the result derived using a counting argument.

Differentiate from first principle

1. . $f'(z) = 3z^2$ (We know that if $p(x) = x^n$ then $p'(x) = nx^{n-1}$)

2.

$f(z) = (1 - z)^2 = z^2 - 2z + 1$

$f'(z) = 2z - 2$

3.

$$\begin{aligned}
f'(z) &= \lim_{h\to 0} \frac{\frac{1}{(1-z-h)^2} - \frac{1}{(1-z)^2}}{h} \\
&= \lim_{h\to 0} \frac{1}{h}\left(\frac{1}{(1-z-h)^2} - \frac{1}{(1-z)^2}\right) \\
&= \lim_{h\to 0} \frac{1}{h}\left(\frac{(1-z)^2}{(1-z-h)^2(1-z)^2} - \frac{(1-z-h)^2}{(1-z-h)^2(1-z)^2}\right) \\
&= \lim_{h\to 0} \frac{1}{h}\frac{(1-z)^2-(1-z-h)^2}{(1-z-h)^2(1-z)^2} \\
&= \lim_{h\to 0} \frac{1}{h}\frac{z^2-2z+1-(z^2+2hz-2z+h^2-2h+1)}{(1-z-h)^2(1-z)^2} \\
&= \lim_{h\to 0} \frac{1}{h}\frac{z^2-2z+1-z^2-2hz+2z-h^2+2h-1)}{(1-z-h)^2(1-z)^2} \\
&= \lim_{h\to 0} \frac{1}{h}\frac{-2hz-h^2+2h}{(1-z-h)^2(1-z)^2} \\
&= \lim_{h\to 0} \frac{-2z-h+2}{(1-z-h)^2(1-z)^2} \\
&= \frac{-2z+2}{(1-z)^2(1-z)^2} \\
&= \frac{-2z+2}{(1-z)^4} \\
&= \frac{2(1-z)}{(1-z)^4} \\
&= \frac{2}{(1-z)^3}
\end{aligned}$$

4.

$f(z) = (1 - z)^3 = - z^3 + 3z^2 - 3z + 1$

$f'(z) = - 3z^2 + 6z - 3$

5. if

f(*x*)=*g*(*x*)+*h*(*x*)

then

$$\begin{aligned} f'(x) &= \lim_{k\to 0} \frac{f(x+k)-f(x)}{k} \\ &= \lim_{k\to 0} \frac{(g(x+k)+h(x+k))-(g(x)+h(x))}{k} \\ &= \lim_{k\to 0} \frac{g(x+k)-g(x)+h(x+k)-h(x)}{k} \\ &= \lim_{k\to 0}\left(\frac{g(x+k)-g(x)}{h} + \frac{h(x+k)-h(x)}{k}\right) \\ &= \lim_{k\to 0} \frac{g(x+k)-g(x)}{h} + \lim_{k\to 0} \frac{h(x+k)-h(x)}{k} \\ &= g'(x)+h'(x) \end{aligned}$$

Differentiating f(z) = (1 - z)^n

1.

$f(z) = (1 - z)^3 = - z^3 + 3z^2 - 3z + 1$

$$\begin{aligned} f'(z) &= -3z^2+6z-3 \\ &= -3(z^2-2z+1) \\ &= -3(z-1)^2 \end{aligned}$$

2.

$f(z) = (1 + z)^2 = z^2 + 2z + 1$

$$\begin{aligned} f'(z) &= 2z+2 \\ &= 2(z+1) \end{aligned}$$

3.

$f(z) = (1 + z)^3 = z^3 + 3z^2 + 3z + 1$

$$\begin{aligned} f'(z) &= 3z^2+6z+3 \\ &= 3(z^2+2z+1) \\ &= 3(z+1)^2 \end{aligned}$$

4.

$$f(z) = \frac{1}{(1-z)^3}$$

$$\begin{aligned}
f'(z) &= \lim_{k\to 0} \frac{\frac{1}{(1-z-k)^3} - \frac{1}{(1-z)^3}}{k} \\
&= \lim_{k\to 0} \frac{1}{k}\left(\frac{1}{(1-z-k)^3} - \frac{1}{(1-z)^3}\right) \\
&= \lim_{k\to 0} \frac{1}{k} \frac{(1-z)^3-(1-z-k)^3}{(1-z-k)^3(1-z)^3} \\
&= \lim_{k\to 0} \frac{1}{k} \frac{-z^3+3z^2-3z+1-(-z^3-3kz^2+3z^2-3k^2z+6kz-3z-k^3+3k^2-3k+1)}{(1-z-k)^3(1-z)^3} \\
&= \lim_{k\to 0} \frac{1}{k} \frac{-z^3+3z^2-3z+1+z^3+3kz^2-3z^2+3k^2z-6kz+3z+k^3-3k^2+3k-1}{(1-z-k)^3(1-z)^3} \\
&= \lim_{k\to 0} \frac{1}{k} \frac{3kz^2+3k^2z-6kz+k^3-3k^2+3k)}{(1-z-k)^3(1-z)^3} \\
&= \lim_{k\to 0} \frac{3z^2+3kz-6z+k^2-3k+3)}{(1-z-k)^3(1-z)^3} \\
&= \frac{3z^2-6z+3}{(1-z)^3(1-z)^3} \\
&= \frac{3z^2-6z+3}{(1-z)^6} \\
&= \frac{3(z^2-2z+1)}{(1-z)^6} \\
&= \frac{3(1-z)^2}{(1-z)^6} \\
&= \frac{3}{(1-z)^4}
\end{aligned}$$

Differentiation technique

1.

$$\begin{aligned}
f(z) &= \frac{1}{(1-z)^2} \\
f'(z) &= -\frac{((1-z)^2)'}{((1-z)^2)^2} \\
&= -\frac{-2(1-z)}{(1-z)^4} \\
&= \frac{2}{(1-z)^3}
\end{aligned}$$

We use the result of the differentation of f(z)=(1-z)^n (f'(z) = -n$(1-z)^{n-1}$)

2.

$$
\begin{aligned}
f(z) &= \frac{1}{(1-z)^3} \\
f'(z) &= -\frac{((1-z)^3)\prime}{((1-z)^3)^2} \\
&= -\frac{-3(1-z)^2}{(1-z)^6} \\
&= \frac{3}{(1-z)^4}
\end{aligned}
$$

3.

$$
\begin{aligned}
f(z) &= \frac{1}{(1+z)^3} \\
f'(z) &= -\frac{((1+z)^3)'}{((1+z)^3)^2} \\
&= -\frac{3(1+z)^2}{(1+z)^6} \\
&= \frac{-3}{(1+z)^4}
\end{aligned}
$$

We use the result of exercise 3 of the previous section f(z)= $(1+z)^3$ -> f'(z)=3(1+z)^2

4.

$$
\begin{aligned}
f(z) &= \frac{1}{(1-z)^n} \\
f'(z) &= -\frac{((1-z)^n)'}{((1-z)^n)^2} \\
&= -\frac{-n(1-z)^{n-1}}{(1-z)^{2n}} \\
&= -\frac{-n}{(1-z)^{2n-(n-1)}} \\
&= -\frac{-n}{(1-z)^{2n-n+1}} \\
&= \frac{n}{(1-z)^{n+1}}
\end{aligned}
$$

We use the result of the differentation of f(z)=(1-z)^n (f'(z) = -n$(1-z)^{n-1}$)

License

GNU Free Documentation License

Version 1.2, November 2002

0. PREAMBLE

The purpose of this License is to make a manual, textbook, or other functional and useful document "free" in the sense of freedom: to assure everyone the effective freedom to copy and redistribute it, with or without modifying it, either commercially or noncommercially. Secondarily, this License preserves for the author and publisher a way to get credit for their work, while not being considered responsible for modifications made by others.

This License is a kind of "copyleft", which means that derivative works of the document must themselves be free in the same sense. It complements the GNU General Public License, which is a copyleft license designed for free software.

We have designed this License in order to use it for manuals for free software, because free software needs free documentation: a free program should come with manuals providing the same freedoms that the software does. But this License is not limited to software manuals; it can be used for any textual work, regardless of subject matter or whether it is published as a printed book. We recommend this License principally for works whose purpose is instruction or reference.

1. APPLICABILITY AND DEFINITIONS

This License applies to any manual or other work, in any medium, that contains a notice placed by the copyright holder saying it can be distributed under the terms of this License. Such a notice grants a world-wide, royalty-free license, unlimited in duration, to use that work under the conditions stated herein. The "Document", below, refers to any such manual or work. Any member of the public is a licensee, and is addressed as "you". You accept the license if you copy, modify or distribute the work in a way requiring permission under copyright law.

A "Modified Version" of the Document means any work containing the Document or a portion of it, either copied verbatim, or with modifications and/or translated into another language.

A "Secondary Section" is a named appendix or a front-matter section of the Document that deals exclusively with the relationship of the publishers or authors of the Document to the Document's overall subject (or to related matters) and contains nothing that could fall directly within that overall subject. (Thus, if the Document is in part a textbook of mathematics, a Secondary Section may not explain any mathematics.) The relationship could be a matter of historical connection with the subject or with related matters, or of legal, commercial, philosophical, ethical or political position regarding them.

The "Invariant Sections" are certain Secondary Sections whose titles are designated, as being those of Invariant Sections, in the notice that says that the Document is released under this License. If a section does not fit the above definition of Secondary then it is not allowed to be designated as Invariant. The Document may contain zero Invariant Sections. If the Document does not identify any Invariant Sections then there are none.

The "Cover Texts" are certain short passages of text that are listed, as Front-Cover Texts or Back-Cover Texts, in the notice that says that the Document is released under this License. A Front-Cover Text may be at most 5 words, and a Back-Cover Text may be at most 25 words.

A "Transparent" copy of the Document means a machine-readable copy, represented in a format whose specification is available to the general public, that is suitable for revising the document straightforwardly with generic text editors or (for images composed of pixels) generic paint programs or (for drawings) some widely available drawing editor, and that is suitable for input to text formatters or for automatic translation to a variety of formats suitable for input to text formatters. A copy made in an otherwise Transparent file format whose markup, or absence of markup, has been arranged to thwart or discourage subsequent modification by readers is not Transparent. An image format is not Transparent if used for any substantial amount of text. A copy that is not "Transparent" is called "Opaque".

Examples of suitable formats for Transparent copies include plain ASCII without markup, Texinfo input format, LaTeX input format, SGML or XML using a publicly available DTD, and standard-conforming simple HTML, PostScript or PDF designed for human modification. Examples of transparent image formats include PNG, XCF and JPG. Opaque formats include proprietary formats that can be read and edited only by proprietary word processors, SGML or XML for which the DTD and/or processing tools are not generally available, and the machine-generated HTML, PostScript or PDF produced by some word processors for output purposes only.

The "Title Page" means, for a printed book, the title page itself, plus such following pages as are needed to hold, legibly, the material this License requires to appear in the title page. For works in formats which do not have any title page as such, "Title Page" means the text near the most prominent appearance of the work's title, preceding the beginning of the body of the text.

A section "Entitled XYZ" means a named subunit of the Document whose title either is precisely XYZ or contains XYZ in parentheses following text that translates XYZ in another language. (Here XYZ stands for a specific section name mentioned below, such as "Acknowledgements", "Dedications", "Endorsements", or "History".) To "Preserve the Title" of such a section when you modify the Document means that it remains a section "Entitled XYZ" according to this definition.

The Document may include Warranty Disclaimers next to the notice which states that this License applies to the Document. These Warranty Disclaimers are considered to be included by reference in this License, but only as regards disclaiming warranties: any other implication that these Warranty Disclaimers may have is void and has no effect on the meaning of this License.

2. VERBATIM COPYING

You may copy and distribute the Document in any medium, either commercially or noncommercially, provided that this License, the copyright notices, and the license notice saying this License applies to the Document are reproduced in all copies, and that you add no other conditions whatsoever to those of this License. You may not use technical measures to obstruct or control the reading or further copying of the copies you make or distribute. However, you may accept compensation in exchange for copies. If you distribute a large enough number of copies you must also follow the conditions in section 3.

You may also lend copies, under the same conditions stated above, and you may publicly display copies.

3. COPYING IN QUANTITY

If you publish printed copies (or copies in media that commonly have printed covers) of the Document, numbering more than 100, and the Document's license notice requires Cover Texts, you must enclose the copies in covers that carry, clearly and legibly, all these Cover Texts: Front-Cover Texts on the front cover, and Back-Cover Texts on the back cover. Both covers must also clearly and legibly identify you as the publisher of these copies. The front cover must present the full title with all words of the title equally prominent and visible. You may add other material on the covers in addition. Copying with changes limited to the covers, as long as they preserve the title of the Document and satisfy these conditions, can be treated as verbatim copying in other respects.

If the required texts for either cover are too voluminous to fit legibly, you should put the first ones listed (as many as fit reasonably) on the actual cover, and continue the rest onto adjacent pages.

If you publish or distribute Opaque copies of the Document numbering more than 100, you must either include a machine-readable Transparent copy along with each Opaque copy, or state in or with each Opaque copy a computer-network location from which the general network-using public has access to download using public-standard network protocols a complete Transparent copy of the Document, free of added material. If you use the latter option, you must take reasonably prudent steps, when you begin distribution of Opaque copies in quantity, to ensure that this Transparent copy will remain thus accessible at the stated location until at least one year after the last time you distribute an Opaque copy (directly or through your agents or retailers) of that edition to the public.

It is requested, but not required, that you contact the authors of the Document well before redistributing any large number of copies, to give them a chance to provide you with an updated version of the Document.

4. MODIFICATIONS

You may copy and distribute a Modified Version of the Document under the conditions of sections 2 and 3 above, provided that you release the Modified Version under precisely this License, with the Modified Version filling the role of the Document, thus licensing distribution and modification of the Modified Version to whoever possesses a copy of it. In addition, you must do these things in the Modified Version:

A. Use in the Title Page (and on the covers, if any) a title distinct from that of the Document, and from those of previous versions (which should, if there were any, be listed in the History section of the Document). You may use the same title as a previous version if the original publisher of that version gives permission.

B. List on the Title Page, as authors, one or more persons or entities responsible for authorship of the modifications in the Modified Version, together with at least five of the principal authors of the Document (all of its principal authors, if it has fewer than five), unless they release you from this requirement.

C. State on the Title page the name of the publisher of the Modified Version, as the publisher.

D. Preserve all the copyright notices of the Document.

E. Add an appropriate copyright notice for your modifications adjacent to the other copyright notices.

F. Include, immediately after the copyright notices, a license notice giving the public permission to use the Modified Version under the terms of this License, in the form shown in the Addendum below.

G. Preserve in that license notice the full lists of Invariant Sections and required Cover Texts given in the Document's license notice.

H. Include an unaltered copy of this License.

I. Preserve the section Entitled "History", Preserve its Title, and add to it an item stating at least the title, year, new authors, and publisher of the Modified Version as given on the Title Page. If there is no section Entitled "History" in the Document, create one stating the title, year, authors, and publisher of the Document as given on its Title Page, then add an item describing the Modified Version as stated in the previous sentence.

J. Preserve the network location, if any, given in the Document for public access to a Transparent copy of the Document, and likewise the network locations given in the Document for previous versions it was based on. These may be placed in the "History" section. You may omit a network location for a work that was published at least four years before the Document itself, or if the original publisher of the version it refers to gives permission.

K. For any section Entitled "Acknowledgements" or "Dedications", Preserve the Title of the section, and preserve in the section all the substance and tone of each of the contributor acknowledgements and/or dedications given therein.

L. Preserve all the Invariant Sections of the Document, unaltered in their text and in their titles. Section numbers or the equivalent are not considered part of the section titles.

M. Delete any section Entitled "Endorsements". Such a section may not be included in the Modified Version.

N. Do not retitle any existing section to be Entitled "Endorsements" or to conflict in title with any Invariant Section.

O. Preserve any Warranty Disclaimers.

If the Modified Version includes new front-matter sections or appendices that qualify as Secondary Sections and contain no material copied from the Document, you may at your option designate some or all of these sections as invariant. To do this, add their titles to the list of Invariant Sections in the Modified Version's license notice. These titles must be distinct from any other section titles.

You may add a section Entitled "Endorsements", provided it contains nothing but endorsements of your Modified Version by various parties--for example, statements of peer review or that the text has been approved by an organization as the authoritative definition of a standard.

You may add a passage of up to five words as a Front-Cover Text, and a passage of up to 25 words as a Back-Cover Text, to the end of the list of Cover Texts in the Modified Version. Only one passage of Front-Cover Text and one of Back-Cover Text may be added by (or through arrangements made by) any one entity. If the Document already includes a cover text for the same cover, previously added by you or by arrangement made by the same entity you are acting on behalf of, you may not add another; but you may replace the old one, on explicit permission from the previous publisher that added the old one.

The author(s) and publisher(s) of the Document do not by this License give permission to use their names for publicity for or to assert or imply endorsement of any Modified Version.

5. COMBINING DOCUMENTS

You may combine the Document with other documents released under this License, under the terms defined in section 4 above for modified versions, provided that you include in the combination all of the Invariant Sections of all of the original documents, unmodified, and list them all as Invariant Sections of your combined work in its license notice, and that you preserve all their Warranty Disclaimers.

The combined work need only contain one copy of this License, and multiple identical Invariant Sections may be replaced with a single copy. If there are multiple Invariant Sections with the same name but different contents, make the title of each such section unique by adding at the end of it, in parentheses, the name of the original author or publisher of that section if known, or else a unique number. Make the same adjustment to the section titles in the list of Invariant Sections in the license notice of the combined work.

In the combination, you must combine any sections Entitled "History" in the various original documents, forming one section Entitled "History"; likewise combine any sections Entitled "Acknowledgements", and any sections Entitled "Dedications". You must delete all sections Entitled "Endorsements."

6. COLLECTIONS OF DOCUMENTS

You may make a collection consisting of the Document and other documents released under this License, and replace the individual copies of this License in the various documents with a single copy that is included in the collection, provided that you follow the rules of this License for verbatim copying of each of the documents in all other respects.

You may extract a single document from such a collection, and distribute it individually under this License, provided you insert a copy of this License into the extracted document, and follow this License in all other respects regarding verbatim copying of that document.

7. AGGREGATION WITH INDEPENDENT WORKS

A compilation of the Document or its derivatives with other separate and independent documents or works, in or on a volume of a storage or distribution medium, is called an "aggregate" if the copyright resulting from the compilation is not used to limit the legal rights of the compilation's users beyond what the individual works permit. When the Document is included in an aggregate, this License does not apply to the other works in the aggregate which are not themselves derivative works of the Document.

If the Cover Text requirement of section 3 is applicable to these copies of the Document, then if the Document is less than one half of the entire aggregate, the Document's Cover Texts may be placed on covers that bracket the Document within the aggregate, or the electronic equivalent of covers if the Document is in electronic form. Otherwise they must appear on printed covers that bracket the whole aggregate.

8. TRANSLATION

Translation is considered a kind of modification, so you may distribute translations of the Document under the terms of section 4. Replacing Invariant Sections with translations requires special permission from their copyright holders, but you may include translations of some or all Invariant Sections in addition to the original versions of these Invariant Sections. You may include a translation of this License, and all the license notices in the Document, and any Warranty Disclaimers, provided that you also include the original English version of this License and the original versions of those notices and disclaimers. In case of a disagreement between the translation and the original version of this License or a notice or disclaimer, the original version will prevail.

If a section in the Document is Entitled "Acknowledgements", "Dedications", or "History", the requirement (section 4) to Preserve its Title (section 1) will typically require changing the actual title.

9. TERMINATION

You may not copy, modify, sublicense, or distribute the Document except as expressly provided for under this License. Any other attempt to copy, modify, sublicense or distribute the Document is void, and will automatically terminate your rights under this License. However, parties who have received copies, or rights, from you under this License will not have their licenses terminated so long as such parties remain in full compliance.

10. FUTURE REVISIONS OF THIS LICENSE

The Free Software Foundation may publish new, revised versions of the GNU Free Documentation License from time to time. Such new versions will be similar in spirit to the present version, but may differ in detail to address new problems or concerns. See http://www.gnu.org/copyleft/.

Each version of the License is given a distinguishing version number. If the Document specifies that a particular numbered version of this License "or any later version" applies to it, you have the option of following the terms and conditions either of that specified version or of any later version that has been published (not as a draft) by the Free Software Foundation. If the Document does not specify a version number of this License, you may choose any version ever published (not as a draft) by the Free Software Foundation

External links

- GNU Free Documentation License (Wikipedia article on the license)
- Official GNU FDL webpage