Contents

I Syllabus and learning outcomes **2**

II Lecture notes **5**

1 Modular Arithmetic **6**
1.1 Basics 6
1.2 Some applications 7
1.3 Solving equations 8
1.3.1 The Euclidean Algorithm 9
1.4 Chinese Remainder Theorem 10
1.5 Cryptography 11
1.6 Public Key Cryptosystems 12

2 Matrix Algebra **14**
2.1 Matrix equations 16
2.2 Affine matrix cryptosystems 17
2.3 Linear Transformations 18
2.4 Matrices as linear transformations 19
2.5 Matrix inverses and systems of equations 20
2.6 Gauss-Jordan method for computing the inverse of a matrix . . . 20
2.7 Systems of linear equations 22
2.8 Determinants 23
2.9 Parallelograms 24

3 Eigenvalues and Eigenvectors **25**
3.1 Problems with rabbits 26
3.2 Finding the n^{th} Fibonacci number 28
3.3 Finding F_n 29

III Sample Exam questions **31**

IV Sample Exercises **33**

Part I

Syllabus and learning outcomes

Elementary Number Theory

Syllabus

- Modular arithmetic, Euclidean algorithm, applications to ISBNs and cryptography
- Euler's Phi function, Fermat's little theorem (and its proof), application to public key cryptography
- Chinese Remainder Theorem

Learning outcomes

- You will be able to use modular arithmetic and Euler's Phi function to: detect errors in ISBNs; encipher messages using 1-dimensional affine and RSA cryptosystems; attack 1-dimensional affine cryptosystems; calculate with Chinese remainders. You will also be able to present a proof of Fermat's little theorem.

Matrix arithmetic

Syllabus

- Matrix addition, multiplication, row operations, inversion (via row operations), systems of equations, applications to resource allocation problems
- Linear transformations, applications to cryptography and geometry

Learning outcomes

- You will be able to use matrices to: solve resource allocation problems; encipher messages using higher dimensional affine cryptosystems; attack higher dimensional affine cryptosystems; solve some geometric problems.

Eigenvalues and eigenvectors

Syllabus

- Calculation of eigenvalues, eigenvectors and matrix powers for 2x2 matrices, Hamilton-Cayley theorem (with proof for 2x2 matrices)
- Fibonacci sequence, Golden Ratio, applications to practical recurrence problems

Learning outcomes

- You will be able to use eigenvalues, eigenvectors and the Principle of Induction to solve practical and theoretical problems about recurrence. You will also be able to state the Hamilton-Cayley theorem and prove it in the 2x2 case.

About this document

This booklet is meant as a companion set of notes to your lecture course. It is not intended as a stand alone course. The text is divided into five main chapters, within each there are sections for particular topics. The last chapter contains some exercises. It is highly recommended that you attempt all exercises and work through all examples by hand to ensure that you fully understand each topic.
If you are viewing this in electronic form, you will notice the red internal links and cyan links to external websites. If you can not see these you may need a newer pdf viewer (such as `http://okular.kde.org/` or `http://www.foxitsoftware.com/Secure_PDF_Reader/`).
There is no need to print this document in colour, and there is ample room for notes on the margins, please try to conserve paper
The following texts and links may be of use for self study, unfortunately there is no single text that will cover exactly this course, so only certain chapters of each of the following are relevant to this course.

- Discrete Mathematics - Biggs. N. L.
- Schaum's Outline of Linear Algebra - Schaums
- `http://en.wikipedia.org/wiki/Linear_algebra`
- `http://joshua.smcvt.edu/linearalgebra/`
- `http://www.math.rutgers.edu/~erowland/modulararithmetic.html`
- `http://ocw.mit.edu/courses/mathematics/18-06sc-linear-algebra-fall-2011/`
- `https://www.coursera.org/`

Part II

Lecture notes

Chapter 1

Modular Arithmetic

1.1 Basics

Consider a standard 12 hour clock. If the time is "10 o'clock", what time will it be in 5 hours? The answer is of course "3 o'clock". We would like to be able to state this result in an equation.

$$10 + 5 = 3 \quad \text{"on a 12 hour clock"} \tag{1.1}$$

Henceforth we will write the above equation as

$$10 + 5 \equiv 3 \bmod 12 \tag{1.2}$$

Consider another example. If today is a Wednesday, what day of the week will it be in 72 days time ? Since there are 7 days in a week, we should use "*mod* 7". So that,

$$72 \equiv 2 \bmod 7 \tag{1.3}$$

So in "72" days, it will be a Friday (the day, two days away from Wednesday). We will now formalize this notion.
Modular arithmetic (sometimes called clock arithmetic) is a system of arithmetic for integers, where numbers "wrap around" upon reaching a certain value—the modulus. For a positive integer n, two integers a and b are said to be congruent modulo n, written:

$$a \equiv b \bmod n \tag{1.4}$$

if their difference $a - b$ is an integer multiple of n. For example,

$$38 \equiv 14 \bmod 12 \tag{1.5}$$

since $38-14 = 24 = (2)(12)$. If two numbers when added, subtracted, multiplied or divided have the same remainder when they are divided by n, we say that they are *congruent modulo n.* For example on the clock face example, we say that *(10+5) is congruent to 3 mod 12*, or for the days of the week *72 is congruent to 2 mod 7.* As an example $15 \equiv 3 \bmod 12$ since $\frac{15}{12}$ is 1 with remainder 3, similarly $\frac{72}{7} = 10$ remainder 2, so the remainder is the part of interest. Let us look at

some more examples:

$$10 \cdot 5 \equiv 2 \bmod 12 \tag{1.6}$$
$$35 \equiv 11 \bmod 12 \tag{1.7}$$
$$2 \cdot 4 \equiv 1 \bmod 7 \tag{1.8}$$
$$-7 \equiv 5 \bmod 12. \tag{1.9}$$

In the last example, perhaps try to think of traveling around the clock face 7 steps anti-clock wise (the minus direction), and then 5 steps clock wise (the positive direction).

1.2 Some applications

Every book can be identified by its "ISBN" number (International Standard Book Number , we will only consider 10 digit ISBN numbers here). For example "Earth" by "Stephen Marshak" has the ISBN - 0393118266. This number is chosen so that

$$(0) \times 1 + (3) \times 2 + (9) \times 3 + (3) \times 4 + (1) \times 5 + (1) \times 6 + (8) \times 7 + (2) \times 8 + (6) \times 9 + (6) \times 10 \equiv 0 \bmod 11$$

We can check this,

$$0 + 6 + 27 + 12 + 5 + 6 + 56 + 16 + 54 + 60 = 242 \equiv 0 \bmod 11 \tag{1.10}$$

The International Bank Account Number (IBAN) is an internationally agreed means of identifying bank accounts across national borders with a reduced risk of propagating transcription errors. It was originally adopted by the European Committee for Banking Standards (ECBS), and later adopted as an international standard under ISO 13616:1997. The current standard is ISO 13616:2007, which indicates SWIFT as the formal registrar. Initially developed to facilitate payments within the European Union, it has now also been implemented by most European countries and many other countries, especially in the Middle East and in the Caribbean.
Consider the sample IBAN:

$$GB \;\; 82 \;\; WEST \;\; 12345698765432 \tag{1.11}$$

There are three steps to validate an IBAN. First - Rearrange, i.e.

$$WEST \quad GB \;\; 82 \;\; 12345698765432 \tag{1.12}$$

Second - convert the letters into numbers, using the following method. Let $A = 10, B = 11, \cdots, Z = 35$. So that we now have,

$$321428291234569876543 2 \tag{1.13}$$

The last step is to check that this number is congruent to 1 mod 97. This is a very large number to deal with. We can simplify calculations, for example

suppose you wanted to compute 4321 mod 97, well

$$\begin{aligned} 4321 &= 4 \times 1000 + 3 \times 100 + 21 && (1.14)\\ &= 4 \times 30 + 3 \times 3 + 21 \bmod 97 && (1.15)\\ &= 23 + 9 + 21 \bmod 97 && (1.16)\\ &= 53 \bmod 97 && (1.17) \end{aligned}$$

So to check the IBAN, we use a similar approach, it is left as an exercise to validate the above IBAN.
Suppose we want to find the last digit of a number, we can use modular arithmetic to find it. For example, the last digit of 1357 is 7, and

$$\begin{aligned} 1357 &= 1000 + 300 + 50 + 7\\ &\equiv 0 + 0 + 0 + 7 \bmod 10\\ &\equiv 7 \bmod 10 \end{aligned}$$

So in general the last digit of any number is the number mod 10.

1.3 Solving equations

Solve the following equation for x,

$$3x \equiv 13 \bmod 26$$

So we would like to multiple across by 3^{-1} (we have no concept of division in modular arithmetic). So the answer should be $x \equiv 3^{-1} \cdot 13 \bmod 26$. How do we find $3^{-1} \bmod 26$, let $p := 3^{-1} \bmod 26$. Then

$$3 \cdot p \equiv 1 \bmod 26$$

After some trial an error it is easy to see that $p = 9$ is a solution, hence $3^{-1} = 9 \bmod 26$ (since $3 \cdot 9 = 27$ and $\frac{27}{26} = 1$ remainder 1). Now we can find x,

$$\begin{aligned} 3x &\equiv 13 \bmod 26 && (1.18)\\ x &\equiv 3^{-1} \cdot 13 \bmod 26 && (1.19)\\ x &\equiv 9 \cdot 13 \bmod 26 && (1.20)\\ x &\equiv 117 \bmod 26 && (1.21)\\ x &\equiv 13 \bmod 26 && (1.22) \end{aligned}$$

Lets try another example, solve $4x \equiv 12 \mod 26$. In this case we notice that $x = 3$ is a solution, but $x = 16$ is also a solution. As a last example think about trying to find $12^{-1} \bmod 15$. So we would like a number a such that $12a \equiv 1 \bmod 15$. Lets try picking a few values for a, say $\{0, 1, 2, \cdots\}$, then $12a \bmod 15$ takes the values $\{0, 12, 9, 6, 3, 0, 12, \cdots\}$ (the sequence just repeats!). This is a problem, since this means that $12^{-1} \bmod 15$ does not exist. This prompts us to ask some questions about the calculation of inverses in modular arithmetic.
First, is it always possible to find the inverse, and if so is it unique? Before we answer these questions we introduce the $gcd(x, y)$ *the greatest common divisor*

of two natural numbers x and y , this is the largest natural number that divides both x and y. For example, $gcd(5, 10) = 5$, $gcd(3, 7) = 1$. For small x and y you can determine the *gcd* by trail and error, but for larger values we will use a method known as *The Euclidean Algorithm*, this a is a procedure for finding the *gcd* of any two natural numbers.

1.3.1 The Euclidean Algorithm

We rely on the observation that the *gcd* of two numbers also divides their difference. To compute $gcd(48, 18)$, divide 48 by 18 to get a quotient of 2 and a remainder of 12. Then divide 18 by 12 to get a quotient of 1 and a remainder of 6. Then divide 12 by 6 to get a remainder of 0, which means that 6 is the *gcd*. Note that we ignored the quotient in each step except to notice when the remainder reached 0, signaling that we had arrived at the answer. We will illustrate this with another example, find $gcd(1071, 462)$?

$$\begin{aligned} 1071 &= 2 \cdot 462 + 147 \\ 462 &= 3 \cdot 147 + 21 \\ 147 &= 7 \cdot 21 + 0 \end{aligned}$$

So $gcd(1071, 462) = 21$. Again, find $gcd(99, 87)$, so

$$\begin{aligned} 99 &= 1 \cdot 87 + 12 \\ 87 &= 7 \cdot 12 + 3 \\ 12 &= 4 \cdot 3 + 0 \end{aligned}$$

hence $gcd(99, 87) = 3$. Lastly, find $gcd(7, 3)$, it is easy to compute that 1 is the answer. When the $gcd(x, y) = 1$ we say that x and y are relatively prime (or *co-prime*), other examples of relatively prime numbers are $(17, 4)$ and $(6565, 8768)$.
We turn our attention back to computing inverses in modular arithmetic. Suppose we want to compute $x^{-1} \bmod m$ (i.e the inverse of $x \mod m$), furthermore suppose that $gcd(m, x) = 1$ (they are co-prime). Then we know that

$$1 = gcd(m, x) = am + bx$$

for some numbers a and b. But this means that $bx \equiv 1 \bmod m$, so b is the inverse of $x \bmod m$. This is exactly what we wanted, we will need the numbers a and b, we use a modified version of the Euclidean algorithm for this. This is like the Euclidean algorithm but in reverse, lets look at an example. Find $gcd(26, 15)$ and use this to calculate $15^{-1} \bmod 26$. First we use the Euclidean algorithm to find $gcd(26, 15)$,

$$\begin{aligned} 26 &= 1 \cdot 15 + 11 \\ 15 &= 1 \cdot 11 + 4 \\ 11 &= 2 \cdot 4 + 3 \\ 4 &= 1 \cdot 3 + 1 \end{aligned}$$

hence $gcd(26, 15) = 1$. Now we run our calculations backward to find $15^{-1} \bmod 26$, that is

$$\begin{aligned} 1 &= 4 - 1 \cdot 3 \\ &= 4 - 1 \cdot (11 - 2 \cdot 4) \\ &= 4 - 1 \cdot 11 + 2 \cdot 4 \\ &= 3 \cdot 4 - 1 \cdot 11 \\ &= 3 \cdot (15 - 1 \cdot 11) - 1 \cdot 11 \\ &= 3 \cdot 15 - 4 \cdot 11 \\ &= 3 \cdot 15 - 4 \cdot (26 - 1 \cdot 15) \\ &= 7 \cdot 15 - 4 \cdot 26 \end{aligned}$$

Hence $1 = 7(15) - 4(26)$, (divide across by 26 and observe the remainders), hence $15^{-1} \equiv 7 \bmod 26$. As a check $15 \cdot 7 = 105$ and $\frac{105}{26} = 4\frac{1}{26}$. The alert reader will notice that this method will only work when $gcd(x, y) = 1$. Moreover $x^{-1} \bmod m$ only exists if $gcd(x, m) = 1$.

1.4 Chinese Remainder Theorem

Suppose we want to solve the following system of congruence relations.

$$\begin{aligned} x &\equiv 1 \bmod 5 \\ x &\equiv 2 \bmod 6 \\ x &\equiv 3 \bmod 7 \end{aligned}$$

One way would be to list the set of all possible numbers that satisfy each relation, and then find the intersection of these sets. While this will work it is not practical for most problems. Thankfully there is a solution, known as the *Chinese remainder theorem.*

Theorem 1.1. *Let $n_1, \cdots, n_k$ be pairwise relatively prime integers and $M := n_1 \cdot n_2 \cdots n_k$. If $a_1, \cdots, a_k$ are any integers, then there exists an integer $x \bmod M$ that satisfies the system of linear congruence relations:*

$$\begin{aligned} x &\equiv a_1 \bmod n_1 \\ x &\equiv a_2 \bmod n_2 \\ &\vdots \\ x &\equiv a_k \bmod n_k \end{aligned}$$

Moreover, $x \equiv a_1 M_1 y_1 + a_2 M_2 y_2 + \cdots a_k M_k y_k \bmod M$, where $M_i := \frac{M}{n_i}$ and $M_i y_i \equiv 1 \bmod n_i$.

The best way to get to grips with this theorem is to consider an example. Solve the following system of congruence relations for x.

$$\begin{aligned} x &\equiv 1 \bmod 5 \\ x &\equiv 2 \bmod 6 \\ x &\equiv 3 \bmod 7 \end{aligned}$$

We use the Chinese remainder theorem, this means we need to compute a_i,M_i and y_i for $i \in \{1,2,3\}$. Firstly $a_1 = 1, a_2 = 2$ and $a_3 = 3$. Next $M = 5 \cdot 6 \cdot 7 = 210$. Thus $M_1 = \frac{210}{5} = 6 \cdot 7 = 42$,$M_2 = 35$ and $M_3 = 30$. Finally we need to calculate the y_i's, these are just the inverses of the M_i. We could use the method outlined in the previous section, but some simple computation will suffice here. Lets start with y_1. So we want to solve $42y_1 \equiv 1 \bmod 5$.

$$\begin{aligned} 42y_1 &\equiv 1 \bmod 5 \\ 2y_1 &\equiv 1 \bmod 5 \\ 3 \cdot 2y_1 &\equiv 3 \bmod 5 \\ 6y_1 &\equiv 3 \bmod 5 \\ y_1 &\equiv 3 \bmod 5. \end{aligned}$$

So $y_1 = 3$, similarly we can find that $y_2 = 5$ and $y_3 = 4$. We can now write down x, $x \equiv a_1M_1y_1 + a_2M_2y_2 + \cdots a_kM_ky_k \bmod M$, or $x \equiv 836 \bmod 210$, this can be reduced to $x \equiv 206 \bmod 210$.
Lets check our solution (always a good idea!). If our answer is correct then each of the three congruence relations will be satisfied, so $x \equiv 1 \bmod 5$ since $\frac{206}{5} \equiv 1 \bmod 5$. Check the others as an exercise.

1.5 Cryptography

We motivate ourself with an example, suppose you intercept the following cipher text (encrypted message)

$$HKB377B3033CO55BB9 \tag{1.23}$$

We will also assume some details, a 37 letter alphabet was used (that is $\{0,1,2,\cdots,9,10,11,12,\cdots,36,37\}$ where $A = 10, B = 11, \cdots, Z = 36, \_ = 37$). We also assume that an *affine cryptosystem* system was used to encrypt the message in the first place.

Definition 1.1. *An Affine cryptosystem works by changing a character $x \to \alpha x + \beta$, where $\alpha x + \beta$ is the ciphered version of x.*

Let us also assume that our spy has intercepted the last three characters of the original message - the were 007. Using these assumptions we wish to decipher the original message.
In order to do this, we first must determine the numbers α and β used in the affine cryptosystem. We do this using the plain text part of the message. So the last three characters were "0 0 7" which were $BB9$ in the encrypted message (remember B is 11 in this alphabet). Hence

$$0 \to (0)\alpha + \beta = \beta \tag{1.24}$$
$$7 \to 7\alpha + \beta \tag{1.25}$$

So $\beta = 11(= B)$ and $9 \equiv 7\alpha + 11 \bmod 37$ (we use mod37 since our alphabet is

only 37 characters long). Now

$$\begin{aligned} 9 &\equiv 7\alpha + 11 \bmod 37 \\ -2 &\equiv 7\alpha \bmod 37 \\ 35 &\equiv 7\alpha \bmod 37 \\ \alpha &= 5 \end{aligned}$$

Hence the enciphering function is

$$x \to 5x + 11 \bmod 37$$

To decipher the message all we need is the inverse of the above function, which clearly is

$$x \to 5^{-1}(x - 11) \bmod 37$$

We can use the Euclidean algorithm to compute 5^{-1}. The $gcd(37, 5) = 1$ and $1 = 15 \cdot 5 - 2 \cdot 37$, hence $15 \equiv 5^{-1} \bmod 37$. So the deciphering function is

$$\begin{aligned} x &\to 5^{-1}x - 5^{-1} \cdot 11 \bmod 37 \\ &\to 15x - 15 \cdot 11 \bmod 37 \\ &\to 15x - 17 \bmod 37 \\ &\to 15x + 20 \bmod 37 \end{aligned}$$

Can you decipher message 1.23?

1.6 Public Key Cryptosystems

A public key cryptosystem is a cryptosystem with the property that someone who knows the enciphering key can not (without prohibitory long computations) discover the deciphering key. The first such system was found by *Resi,Shamir and Adelman* in 1977. Before we discuss the system itself, we need some mathematical preliminaries.

Definition 1.2. *Let $n \in \mathbb{N}$. Euler's quotient (or phi) function of n, denoted by $\phi(n)$, is the number of integers in the range $1, \cdots, n$ that are co-prime with n.*

For example, $\phi(6) = 2$, since 6 is co-prime with only two numbers (1 and 5) in the set $\{1, 2, 3, 4, 5, 6\}$, verify that $\phi(13) = 12$. Next we require some propositions.

Proposition 1.2. *If p is prime, then $\phi(p^n) = p^n - p^{n-1}$*

This is easy to see, since the only numbers that are not co-prime with p^n are the multiples of p, so $\phi(p) = p^n - \frac{p^n}{p}$.

Proposition 1.3. *Let n and m be co-prime, then $\phi(n \cdot m) = \phi(n)\phi(m)$*

For example, $\phi(15) = \phi(3 \cdot 5) = 2$ *cdot*$4 = 8$. Lastly we require a result known as Euler's theorem.

Theorem 1.4. *Let a and n be co-prime then,*

$$a^{\phi(n)} \equiv 1 \bmod n$$

.

We now give the procedure for *RSA public key encryption*. Pick two distinct primes p and q (about 1000 digits each to be safe). Also choose (randomly) an invertible element $e \in \mathbb{Z} \bmod \phi(pq)$ (Recall that $\phi(pq) = \phi(p)\phi(q) = (p-1)(q-1)$). Now compute $e^{-1} \bmod \phi(pq)$, call it d (i.e. $d \cdot e \equiv 1 \bmod \phi(pq)$). Lastly let $n := p \cdot q$. Then

- Public Key - (n, e) *public.*
- Private Key - (n, d) *Keep secret!*

The encryption is

$$x \to x^e \bmod n \tag{1.26}$$

Chapter 2

Matrix Algebra

To this point we have been dealing with numbers. In many situations (computers, geometry, etc...) collections or arrays of numbers are more useful. We call these arrays matrices.

Definition 2.1. *A $m \times n$ matrix is an array of m rows and n columns of numbers.*

For example a 2×2 matrix has four elements $\{a, b, c, d\}$ and is written as,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

A 1×2 matrix looks like

$$\begin{pmatrix} a & b \end{pmatrix}$$

A 2×1 matrix is

$$\begin{pmatrix} a \\ b \end{pmatrix}$$

and so on. As with numbers, matrices can be added and subtracted. Lets look at each of these operations, with examples:

Matrix addition

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a+e & b+f \\ c+g & d+h \end{pmatrix}$$

Matrix subtraction

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} - \begin{pmatrix} e & f \\ g & h \end{pmatrix} = \begin{pmatrix} a-e & b-f \\ c-g & d-h \end{pmatrix}$$

We also have the additive identity - the zero matrix, such that when added to a matrix, it does not change anything, e.g

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We can also multiply a matrix by a number (or scalar). For example,

$$\lambda \cdot \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} \lambda \cdot a & \lambda \cdot b \\ \lambda \cdot c & \lambda \cdot d \end{pmatrix}$$

So far these operations are straightforward, our next operation is *matrix multiplication*. When multiplying matrices, the elements of the rows in the first matrix are multiplied with corresponding columns in the second matrix. One may compute each entry in the third matrix one at a time. Lets take a simple example of two 2×2 matrices.

$$\begin{pmatrix} 1 & 2 \\ 3 & -1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \cdot 0 + 2 \cdot 1 & 1 \cdot 2 + 2 \cdot 1 \\ 3 \cdot 0 + (-1) \cdot 1 & 3 \cdot 2 + (-1) \cdot 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ -1 & 5 \end{pmatrix}$$

The first thing we notice is that, this operation requires more attention and mistakes are easier to make. Secondly, the above example only considered two matrices of equal dimension (size of the array). In general we can multiply and $m \times n$ matrix by a $n \times p$ matrix, with the product being a $m \times p$ matrix. Lets take an example:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{pmatrix}.$$

Finally we take a look at the special case of multiplication of 2×1 and 1×2 matrices. A 2×1 matrix is called a column vector, and a 1×2 matrix is called a row vector. As an example,

$$\begin{pmatrix} 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = 2 + 1 = 3$$

This is a special case where, the result is a 1×1 matrix, i.e. a number. The above multiplication is sometimes referred to as the *dot/scalar product*. At this point we remark that if we had carried out the above multiplication in reverse order, i.e.

$$\begin{pmatrix} 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 2 & 1 \end{pmatrix}$$

we get a very different answer! *This is an important fact the order of multiplication matters for matrices.* To make sure you understand this point, try reversing the order of all the above examples and compute the new answers.
We now introduce a very special matrix, known as the identity matrix I. The identity matrix in the two dimensional case is

$$I := \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

In any dimension the identity matrix has 1's down the main diagonal,and 0's everywhere else. This matrix acts like the number 1 in standard multiplication.

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$$

and

$$\begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 4 & 5 \end{pmatrix}$$

In fact, given any matrix A, we have that $I \cdot A = A \cdot I = A$ (can you prove this for any 2×2 matrix?). So we have learned some basic facts about matrix algebra.

- We can add or subtract two matrices of equal "size". Matrix addition/subtraction is preformed on each entry of the matrix.
- The order of matrix multiplication matters. In general we can multiply a $m \times p$ matrix by a $p \times n$ matrix to get a $m \times n$ matrix.

2.1 Matrix equations

Since we have no concept of matrix division, how will we solve a matrix equation such as,

$$A \cdot B = C \tag{2.1}$$

where B is the unknown and A, B, C are all 2×2 matrices. The answer is to find the *inverse matrix* of A, denoted A^{-1}. This matrix will have the property that $A^{-1} \cdot A = I = A \cdot A^{-1}$. At this point we need to make sure, that this matrix exists, and if so how to find it.
Let

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

then

$$A^{-1} := \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Check yourself that $A^{-1} \cdot A = I = A \cdot A^{-1}$ holds for the above definition of A^{-1}. We note that division by the number $ad - bc$ is only allowed when $ad - bc \neq 0$. This number is a very important number in matrix algebra, so henceforth we will refer to $ad - bc$ as the *determinant of* A - denoted by $det(A)$ or $|A|$.
So we can only find A^{-1} when $det(A) \neq 0$. Since we will only be using 2×2 matrices, we omit the general definition of inverses and determinants for higher dimensional square matrices (see `http://en.wikipedia.org/wiki/Invertible_matrix`).
Lets try an example, find the matrix A such that

$$\begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \cdot A = \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

So we need to isolate the unknown A on the left hand side, to do this we multiply both sides on the left by the inverse of

$$B := \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$$

It is very important we multiply both sides on the **left** since matrix multiplication does not commute. So lets compute the inverse of B,

$$B^{-1} := \frac{1}{1 - 6} \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} \tag{2.2}$$

Next we multiply both sides on the left by B^{-1}. So

$$\frac{1}{-5} \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix} \cdot A = \frac{1}{-5} \begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

As expected, the left hand side is just the identity (that was the whole point of finding B^{-1}). Finally we have

$$I \cdot A = \frac{1}{-5}\begin{pmatrix} 1 & -2 \\ -4 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}$$

$$A = \frac{1}{-5}\begin{pmatrix} 3 & -2 \\ -5 & 1 \end{pmatrix}$$

While it may look messy, the procedure itself is straightforward. Try the following exercise.
Solve for A (hint: this time multiplication will be on the right),

$$A \cdot \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 2 \end{pmatrix}$$

2.2 Affine matrix cryptosystems

Recall the notion of an affine cryptosystem 1.5, this method is far from secure. Since the letter "e" is very common in english (followed by "T", etc..) a simple frequency analysis could break the code (`http://en.wikipedia.org/wiki/Frequency_analysis`). To try an make the message more secure we could break the plaintext message into small message units each of length two say. Now there are many many contenders for the most frequent pair of letter in English - so we have made the job of trying to decipher the message much more difficult!
To do this we would encipher using,

$$\begin{pmatrix} x \\ y \end{pmatrix} \to A \cdot \begin{pmatrix} x \\ y \end{pmatrix} + B$$

where $\binom{x}{y}$ is the two letter message part, A is a 2×2 matrices and B a fixed 2×1 vector. We must choose the matrix A so that it is invertible! Just as in section 1.5, we often use mod37, so we also need the matrix A to be invertible mod37. Since the way to decipher the message is

$$\begin{pmatrix} x \\ y \end{pmatrix} \to A^{-1} \cdot \left[\begin{pmatrix} x \\ y \end{pmatrix} - B\right]$$

So we are not always just interested in matrices with real numbers, in fact matrices over other number systems (modular, etc..) are a rich and deep area of interest.

2.3 Linear Transformations

Take the $x-y$ plane ($\mathbb{R}^2$), any point in the plane can be represented by two numbers, x *and* y*,coordinates* .A *transformation* on the plane is a function $T : \mathbb{R}^2 \to \mathbb{R}^2$, which sends each point (x, y) in the plane to another point $T(x, y)$. It is often useful to think of these points as matrices (either 2×1 or 1×2). These special matrices are called *vectors*, generally people use *column vectors* (i.e. 2×1 matrices) to represent vectors in the pane. You can use *row vectors* if you prefer, but once you make a choice you must stick to it. For convenience of printing all vectors here are represented as *row vectors*. Like we did with matrices, we introduce some algebra to these vectors, we can add points, e.g.

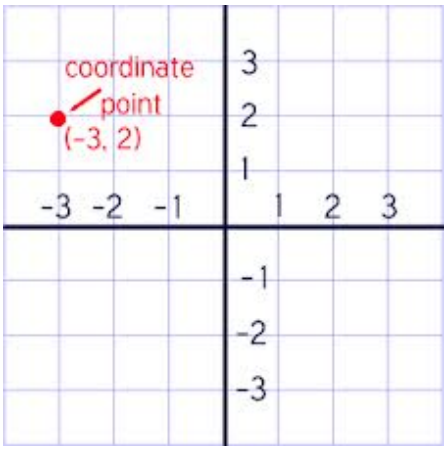


Figure 2.1: The plane

$$(2, 1) + (-3, 1) = (-1, 2)$$

In general $(x, y) + (x', y') = (x + x', y + y')$. We can also multiply a point $P = (x, y)$ by a scalar $\lambda \in \mathbb{R}$ by

$$\lambda \cdot P = (\lambda x, \lambda y)$$

For example, $3 \cdot (1, 2) = (3, 6)$. These operations allow us to introduce some algebra to the geometry (and vice versa for a historical view point see `http://en.wikipedia.org/wiki/Euclidean_geometry`).

Definition 2.2. *A transformation* $T : \mathbb{R}^2 \to \mathbb{R}^2$ *is said to be Linear if given two points* $P, Q \in \mathbb{R}^2$ *and* $\lambda \in \mathbb{R}$

(i) $T(P + Q) = T(P) + T(Q)$

(ii) $T(\lambda \cdot P) = \lambda \cdot T(P)$

Lets look at an example, consider the transformation,

$$T(x, y) \to (x + 2y, 3x + 4y)$$

So $T(-1, 2) = (-1 + 2 \cdot 2, 3 \cdot (-1) + 4 \cdot 2) = (3, 5)$. Now we can ask the question, is T linear? To show this we need to consider to generic points, say (x, y) and (a, b) and check the properties in the definition above. So

$$\begin{aligned} T((x, y) + (a, b)) &= T((x + a, y + b)) && (2.3)\\ &= (x + a + 2y + 2b, 3x + 3a + 4y + 4b) && (2.4)\\ &= (x + 2y + a + 2b, 3x + 4y + 3a + 4b) && (2.5)\\ &= ((x + 2y), (3x + 4y)) + ((a + 2b), (3a + 4b)) && (2.6)\\ &= T(x, y) + T(a, b) && (2.7) \end{aligned}$$

It is left as an exercise to verify that $T(\lambda(x,y)) = \lambda T((x,y))$. Hence T is linear. It is important to realize that not every function is linear, for example, $M(x,y) \rightarrow (x^2, y)$ is not linear. To show this, observe that

$$\begin{aligned} M(\lambda \cdot x, \lambda \cdot y) &= (\lambda^2 x^2, \lambda y) && (2.8) \\ &\neq \lambda \cdot (x^2, y) && (2.9) \end{aligned}$$

Consider the following transformation $T(x,y) = (-x, y)$, is the linear? what does the transformation do? (trying experimenting with points). Can you construct more examples of well known geometric operations that are (i) linear or (ii) non-linear.

2.4 Matrices as linear transformations

We already noticed that points in $\mathbb{R}^2$ can be viewed as row or column vectors (i.e. $2 \times 1 / 1 \times 2$ matrices). We now show that any linear transformation can also be viewed as a matrix. Take the linear transformation $T(x,y) \rightarrow (x + 2y, 3x + 4y)$. We can represent this linear transformation as

$$\begin{pmatrix} x \\ y \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + 2y \\ 3x + 4y \end{pmatrix}$$

We say that the matrix $\left(\begin{smallmatrix} 1 & 2 \\ 3 & 4 \end{smallmatrix}\right)$ represents the linear transformation T. This is a key step in our understanding of plane geometry, we now have a link between algebra and geometry, so if we find a problem hard to solve geometrically we can instead think about it algebraically and hope it is easier to solve (and vice versa).

Theorem 2.1. *Any linear transformation $T : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ can be represented by some matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.*

Proof: How do we go about proving this? We know that $T(1,0) = (a,b)$ and $T(0,1) = (c,d)$. Using the linearity of T we know that $T(x,y) = T(x(1,0) + y(0,1)$.

$$\begin{aligned} T(x,y) &= T(x(1,0) + y(0,1) && (2.10) \\ &= T(x(1,0)) + T(y(0,1)) && (2.11) \\ &= xT(1,0) + yT(0,1) && (2.12) \\ &= x(a,b) + y(c,d) && (2.13) \\ &= (ax + by, cx + dy) && (2.14) \end{aligned}$$

and finally,

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} ax + by \\ cx + dy \end{pmatrix}$$

Hence the theorem is proved.
We are now in a position to know what is, and what is not a linear transformation (if you can represent the transformation by a matrix, its linear, otherwise its not). For example

- Any reflection is linear
- Any rotation is linear
- Any composition of reflections and rotations is linear

Can you describe the matrix that would represent a reflection in the (i) x axis (ii) y axis (iii) about the line $y = x$?
Since any linear transformation is represented by a matrix, what does it mean when we multiply two matrices?
What might the following matrix represent?

$$\begin{pmatrix} cos(\theta) & -sin(\theta) \\ sin(\theta) & cos(\theta) \end{pmatrix}$$

(hint: `http://math.usask.ca/~maclean/JavaPage.html`-Linear Algebra, or pick some sample values and play around), after you have this done, try reading the following article (just skip over the complicated bits!) `http://en.wikipedia.org/wiki/Orthogonal_matrix`.

Theorem 2.2. *Let S and T be linear transformations, represented by matrices A and B respectively. Then the linear transformation $T \circ S$ (i.e. $T(S)(x, y)$) is represented by the matrix $B \cdot A$*

2.5 Matrix inverses and systems of equations

So far we have only dealt with 2×2 matrices, but what about higher dimensional ones (needed to represent linear transformations from $\mathbb{R}^n \to \mathbb{R}^n$). The same rules for matrix addition, scalar multiplication still hold, but calculating the inverse of a 3×3 matrix is different. So how can we find A^{-1} when,

$$A := \begin{pmatrix} 3 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}$$

So we need to find another matrix B such that $A \cdot B = I$, that is

$$A \cdot B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

Much like the 2×2 case there is a formula, but it is rather awkward to use, we will use a different method, one that is often used in practice since it reduces the number of operations.

2.6 Gauss-Jordan method for computing the inverse of a matrix

We first introduce the idea of row operations, let the i^{th} row of the matrix be denoted by ρ_i, also let $i, j \in \mathbb{Z}$ with $i \neq j$ and λ a non zero scalar.

(i) $\rho_i \to \rho_i + \lambda\rho_j$

(ii) $\rho_i \leftrightarrow \rho_j$

(iii) $\rho_i \to \lambda\rho_i$

To find A^{-1} we first write down the matrix A next to the identity matrix, as follows

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

Our goal is to apply row operations to the above until, we have the 3×3 identity matrix on the left, and another matrix on the right, this other matrix will be A^{-1}. We now illustrate this
Find the inverse of A above, so we first apply the row operation, $\rho_1^{'} := \rho_1 + \rho_2$,

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & 1 & 0 \\ 2 & 0 & -2 & 0 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \end{array}\right)$$

apply the row operation, $\rho_3 \leftrightarrow \rho_2$,

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 2 & 0 & -2 & 0 & 1 & 0 \end{array}\right)$$

then $\rho_3^{'} = 2\rho_1 - 5\rho_2$

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 10 & 2 & -3 & 0 \end{array}\right)$$

next, $\rho_2^{'} := -10\rho_2 + \rho_3$

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 0 & 10 & 2 & -3 & 0 \end{array}\right)$$

next we scale all the rows (i.e $\rho_1^{'} \to \frac{1}{5}\rho_1$, etc...),

$$\left(\begin{array}{ccc|ccc} 5 & 0 & 0 & 1 & 1 & 0 \\ 0 & -10 & 0 & 2 & -3 & -10 \\ 0 & 0 & 10 & 2 & -3 & 0 \end{array}\right)$$

Notice, how we have found numbers along the main diagonal on the left first, rather than forcing each number to be a 1 as soon as possible, this is preferred since will delay the use of fractions, which can be a source of errors. All that is left to do now, is to scalar multiply each row, to get the desired result. So $\frac{1}{5} \cdot \rho_1$,$-\frac{1}{10} \cdot \rho_2$ and $\frac{1}{10} \cdot \rho_3$.

$$\left(\begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{5} & \frac{1}{5} & 0 \\ 0 & 1 & 0 & -\frac{1}{5} & \frac{3}{5} & 1 \\ 0 & 0 & 1 & \frac{1}{5} & -\frac{3}{5} & 0 \end{array}\right)$$

So this is A^{-1}. As an exercise check that $A \cdot A^{-1} = I = A^{-1} \cdot A$. This method will work for inverting any $n \times n$ matrix - but why? We will explore this. Let us no explore why the *Gauss-Jordan method* works.
Consider the matrix, E_{ij}^{λ} this matrix has 1 on the main diagonal and λ in the ij (intersection of i^{th} row and j^{th}) column, and 0 elsewhere. Lets see what happens when we compute $E_{13}^{2} \cdot A$,

$$E_{13}^{2} \cdot A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 3 & 0 & 2 \\ 2 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 4 \\ 2 & 0 & -2 \\ 0 & 1 & 1 \end{pmatrix}$$

Now notice that if we had applied the row operation $\rho_1^{'} = \rho_1 + 2\rho_3$ to the matrix we would have found the same. We have just represented a row operation in terms of these E_{ij}^{λ} matrices. In fact we can represent all the row operations in terms of matrices similar to the above. For example the matrix P_{ij} that has 1 in both the ij and ji position and zero's everywhere else, will swap rows i and j. As an exercise can you write each of the row operations we used to compute A^{-1} as matrix multiplication?
So when we computed A^{-1} we were just multiplying the matrix A by a sequence of matrices, lets call them $E_1, \cdots, E_k$ (for k steps), hence

$$(E_k \cdot \cdots \cdot E_1) \cdot A = I \tag{2.15}$$
$$(E_k \cdot \cdots \cdot E_1) \cdot A \cdot A^{-1} = I \cdot A^{-1} \tag{2.16}$$
$$(E_k \cdot \cdots \cdot E_1) \cdot I = A^{-1} \tag{2.17}$$

The above just says that the *Gauss-Jordan* method computes A^{-1}.

2.7 Systems of linear equations

So far we have seen how to manipulate matrices, lets use them now to solve systems of equations. Consider a factory that requires *energy,steel and labour* to make three machines *A,B and C*. We will summarize the factory details in the table, What level of production ensures all resources are used?

Resource	A	B	C	Weekly available
Energy	$2MWh$	$3MWh$	$2MWh$	$100MWh$
Steel	1 Ton	1 Ton	4 Tons	70 Ton
Labour	$20hrs$	$10hrs$	$10hrs$	$500hrs$

Table 2.1: Maximal abelian ideals

Lets suppose we manufacture x units of A, y units of B and z units of C. If **all** resources are to be used then,

$$2x + 3y + 2z = 100 \tag{2.18}$$
$$x + y + 4z = 70 \tag{2.19}$$
$$20x + 10y + 10z = 500 \tag{2.20}$$

This collection of linear equations, is often referred to as a system of linear equations. Lets use matrices to solve this system, we first write the system in terms of matrices, that is,

$$\begin{pmatrix} 2 & 3 & 2 \\ 1 & 1 & 4 \\ 20 & 10 & 10 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 100 \\ 70 \\ 500 \end{pmatrix}$$

and we want to find the unknowns x, y and z. Now we are going to apply some row operations to the above system of equations. When we apply our row operations, we treat both sides of the equation equally!. Lets start with: $\rho_1^{'} \to \rho_2 - \frac{1}{2}\rho_1$ and $\rho_3^{'} \to \rho_3 - 10\rho_1$, giving us,

$$\begin{pmatrix} 2 & 3 & 2 \\ 0 & -\frac{1}{2} & 3 \\ 0 & -20 & -10 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 100 \\ 20 \\ -500 \end{pmatrix}$$

Next, $\rho_3^{'} \to \rho_3 - 40\rho_2$, to give

$$\begin{pmatrix} 2 & 3 & 2 \\ 0 & -\frac{1}{2} & 3 \\ 0 & 0 & -130 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 100 \\ 20 \\ -1300 \end{pmatrix}$$

Now we observe that we have an upper triangular matrix, this allows us to read the solutions very easily. So $-130z = -1300$ hence $z = 10$. Use this to find y, so that $-\frac{1}{2}y + 3(10) = 20$ hence $y = 20$ and use this to find $x = 10$. In the above matrix of coefficients, the numbers on the main diagonal are referred to as *pivots*. This procedure for solving systems of equations is known as *Gaussian elimination*, and can be applied to very large systems of equations.
Will this method always work? If not in what cases does it fail? The answer lies in the pivots, the above system of equations is solved since we found three pivots and had three unknowns.

2.8 Determinants

Here we formally introduce introduce *Determinants* (as seen in 2.1), an important concept when working with matrices. Henceforth we will mainly work with 2×2 matrices.

Definition 2.3. *Let* $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ *then the* ***adjoint matrix*** *is*

$$adj(A) := \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Observe that $A^{-1} = \frac{1}{ad-bc} \cdot adj(A)$, where $ad - bc = det(A)$ the determinant that we first in section 2.1.
Let $A := \left(\begin{smallmatrix} 3 & 1 \\ 4 & 2 \end{smallmatrix}\right)$, calculate $det(A)$ and $adj(A)$ and use these to solve the following system of equations.

$$3x + y = 2 \tag{2.21}$$
$$4x + 2y = 4 \tag{2.22}$$

This is equivalent to solving the following matrix equation,

$$\begin{pmatrix} 3 & 1 \\ 4 & 2 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix}$$

To solve for the unknown vector $\binom{x}{y}$, we multiply both sides of the above equation on the left, by A^{-1}, where $A^{-1} = \frac{1}{det(A)} Adj(A)$. To find $\binom{x}{y} = \binom{0}{2}$, so $x = 0$ and $y = 2$.

Proposition 2.3. *If $det(A) = 0$ then A has no inverse.*

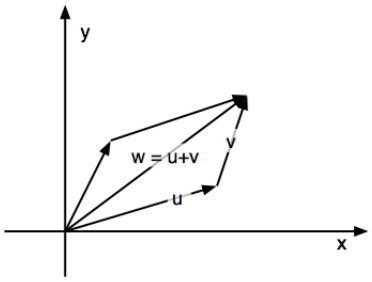
Proof The proof is left as an exercise.

Proposition 2.4. *Given two matrices A and B, then*

$$|AB| = |A| \cdot |B|$$

2.9 Parallelograms

Recall section 2.4 where we showed that linear transformation and matrices were related, it turns out that the determinant of a matrix also has a geometrical meaning. Consider two vectors in the plane. Let $\mathbf{u} = \binom{3}{0}$ and $\mathbf{v} = \binom{1}{2}$. If we draw these vectors (i.e join the origin to the points) we have the following:



It is clear that these two vectors span (or define) a parallelogram P in the plane. The area of this parallelogram is given by the formula $Area(P)$ =base $\times$ perpendicular height. In this case $Area(P) = 3 \cdot 2 = 6$. (to see this yourself, try drawing P on some graph paper, and dropping perpendiculars to form two triangles and a square).
Now observe the matrix $\begin{pmatrix} 3 & 1 \\ 0 & 2 \end{pmatrix}$ - that is the matrix whose columns are the vectors in question. Next we compute $|A|$,

$$det(A) = 3 \cdot 2 - 1 \cdot 0 = +6$$

So we get back the area again!, you will notice the + sign above, we do this since we can also find negative areas. If we had swapped the columns in A, the area would have been -6.

Theorem 2.5. *The determinant of a 2×2 matrix is equal to + or − area of the parallelogram determinant by its two columns*

In fact the above theorem holds for any $n \times n$ matrix, but instead of "area"we need an n dimensional "volume" (see `http://en.wikipedia.org/wiki/Orientation_%28vector_space%29`

Chapter 3

Eigenvalues and Eigenvectors

Once again let $A := \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$.

Definition 3.1. *A non-zero vector $v = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ is called an eigenvector for A if there exists a number λ such that*

$$Av = \lambda v$$

we call the number λ the eigen value associated to v.

Lets see this for $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, and $v = \left(\begin{smallmatrix} 4 \\ 4 \end{smallmatrix}\right)$. Then

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 12 \\ 12 \end{pmatrix}$$

Also,

$$3 \cdot \begin{pmatrix} 4 \\ 4 \end{pmatrix} = \begin{pmatrix} 12 \\ 12 \end{pmatrix}$$

Thus $v = \left(\begin{smallmatrix} 4 \\ 4 \end{smallmatrix}\right)$ is an eigenvector of A with eigenvalue of $\lambda = 3$.

Theorem 3.1. *Let A be a 2×2 matrix and let $v = \left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right)$ be an eigenvector and suppose that $Av = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$. Then $det(A) = 0$*

Proof: If A^{-1} exists, then

$$A^{-1} \cdot Av = A^{-1} \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$$

hence $v = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, a contradiction. Thus A^{-1} does not exist, hence $det(A) = 0$.
How do we go about finding all the eigenvalues and eigenvectors for a given matrix. Once again let $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$, suppose v is some eigenvector with eigenvalue λ. So

$$Av = \lambda v$$

Then $Av - \lambda v = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$, which we can write as:

$$(A - \lambda \cdot I)v = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$$

This is the same as saying, $det(A - \lambda I) = 0$. So we now have a general way to find eigenvalues and eigenvectors, lets apply this to the matrix A.
Find all eigenvectors and eigenvalues for $A = \left(\begin{smallmatrix} 2 & 1 \\ 1 & 2 \end{smallmatrix}\right)$. First we compute

$$(A - \lambda I) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} - \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix}$$

So

$$\begin{aligned} det(A - \lambda I) &= (2-\lambda)\cdot(2-\lambda) - 1 \cdot 1 && (3.1)\\ &= \lambda^2 - 4\lambda + 3 && (3.2)\\ &= (\lambda - 3)(\lambda - 1) = 0 && (3.3) \end{aligned}$$

Hence $\lambda = 3{,}1$ are the two eigenvalues for A. Next we need to find the eigenvectors associated to each of the eigenvalues. We know that if v is an eigenvector with eigenvalue λ then,

$$\begin{aligned} Av &= \lambda v && (3.4)\\ (A - \lambda I)v &= 0 && (3.5)\\ & && (3.6) \end{aligned}$$

So we need to solve:

$$\begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

First, let $\lambda = 1$ then

$$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

So $v = \left(\begin{smallmatrix} 1 \\ -1 \end{smallmatrix}\right)$ is an eigenvector for $\lambda = 1$. In similar fashion we find that $u = \left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right)$ is an eigenvector for $\lambda = 3$. Observe that any scalar multiply of the eigenvectors is also an eigenvector, so there is a certain degree of freedom when solving for eigenvectors.

3.1 Problems with rabbits

While eigenvalues and eigenvectors have vast geometric uses (see `http://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors`), we will concentrate on some algebraic problems in this final section. Consider the problem of breeding rabbits.
One newly born male rabbit and one newly born female rabbit are placed in a field. We assume that they mate at a rate of 1 a month (as rabbits do), and one month later the female produces one male/female pair. We also assume that rabbits don't die (awwwww). Our question, how fast does the rabbit population grow? Let us write the information in a simple table:

*	*	*	*	*	Month	Number of Pairs
			MF		0	1
			MF		1	1
		MF	MF		2	2
	MF	MF	MF	MF	3	3
MF	MF	MF	MF	MF	4	5

Table 3.1: Rabbit population, MF represents a male/female pair.

We may want to ask the question, how many rabbits are there are 12 months. You may have seen this problem before, as an introduction to *Fibonacci numbers*. For that reason let F_n = the number of pairs of rabbits after n months. So

$$\begin{aligned} F_0 &= 1 && (3.7)\\ F_1 &= 1 && (3.8)\\ F_2 &= F_1 + F_0 && (3.9)\\ F_3 &= F_2 + F_1 && (3.10)\\ &\vdots && (3.11)\\ F_n &= F_{n-1} + F_{n-2} && (3.12) \end{aligned}$$

So the sequence looks like $1, 1, 2, 3, 5, 8, \cdots$.
Now look at the sequence $\frac{1}{1}, \frac{2}{1}, \frac{3}{2}, \frac{5}{3}, \cdots, \frac{F_{n+1}}{F_n}$. Is there anything special about this sequence? For example, the term $\frac{55}{34} = 1.617647\cdots$ and the term $\frac{233}{141} = 1.6180555\cdots$. So it seems like the series is converging to a number,

$$\lim_{n\to\infty} \frac{F_{n+1}}{F_n} = \phi$$

where ϕ is some number "close to " 1.6180.
If this number ϕ exists (we still have not shown that !), then the population of the rabbits increases by (roughly) a factor of ϕ once a month. If we progress far enough in the sequence then $F_n = F_{n-1} + F_{n-2}$ and $F_{n-1} = F_{n-1}$ should hold, we write this as,

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix}$$

Of course there is nothing special about the n^{th} term, so we can write the above using the the previous terms, and their previous terms and so on, eventually we find

$$\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}^{n-1} \cdot \begin{pmatrix} F_1 \\ F_0 \end{pmatrix}$$

If the limit ϕ does exist, then $\frac{F_n}{F_{n-1}} \to \phi$ or $F_n \simeq \phi F_{n-1}$ or,

$$\phi \cdot \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix} \simeq \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix}$$

So ϕ is like an eigenvalue for the matrix $A := \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$. Lets try to find the eigenvalues for A, so we need to solve

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

hence, $\lambda = \frac{1 \pm \sqrt{1+4}}{2}$, logic tells us that $\phi > 0$ (in the absence of time traveling bunnies and a causality paradox) so pick $\phi := \lambda = \frac{1+\sqrt{5}}{2}$. This is the number we were after, apart from breeding rabbits this number appears vast numbers of times in nature (`http://en.wikipedia.org/wiki/Golden_ratio` and `http://en.wikipedia.org/wiki/Fibonacci_number`). This number $\phi = \frac{1+\sqrt{5}}{2}$ is known as the *Golden ratio*. See the two links for the many alternative definitions of ϕ.

Definition 3.2. *The **Golden ration** is*

$$\phi := \frac{1+\sqrt{5}}{2}$$

Definition 3.3. *Two numbers $a, b \in \mathbb{Z}_{\geq 1}$ are said to be in the Golden ration if*

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

As an exercise check the above definition by solving the equations $\frac{a+b}{a} = \lambda = \frac{a}{b}$ for λ.

Proposition 3.2. *A "beautiful body" is such that the ratio of heights, from feet to hips and hips to head is golden.*

Proposition 3.3. *A "beautiful" window has the form, that its sides are in the golden ratio.*

3.2 Finding the n^{th} Fibonacci number

We still have to find a formula for calculating the n^{th} term in the Fibonacci sequence. Let us start with a theorem.

Theorem 3.4. *If a 2×2 matrix has eigenvalues λ_1, λ_2 with corresponding eigenvectors v_1, v_2, and if the matrix T whose columns are the eigenvectors v_1 and v_2 is invertible, then*

$$T^{-1}AT = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

Proof:

$$\begin{aligned}
T^{-1}AT &= T^{-1}A\left(\begin{smallmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{smallmatrix}\right) && (3.13)\\
&= T^{-1}\left(\begin{smallmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{smallmatrix}\right) && (3.14)\\
&= T^{-1}\left(\begin{smallmatrix} \lambda_1 v_{11} & \lambda_2 v_{12} \\ \lambda_1 v_{21} & \lambda_2 v_{22} \end{smallmatrix}\right) && (3.15)\\
&= T^{-1}\left(\begin{smallmatrix} v_{11} & v_{12} \\ v_{21} & v_{22} \end{smallmatrix}\right)\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right) && (3.16)\\
&= I\left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right) && (3.17)\\
&= \left(\begin{smallmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{smallmatrix}\right) && (3.18)
\end{aligned}$$

We will use the above theorem to examine the eigenvectors associated to eigenvalue ϕ. Once again let $A = \left(\begin{smallmatrix} 1 & 1 \\ 1 & 0 \end{smallmatrix}\right)$ we already know this has eigenvalues $\phi = \frac{1+\sqrt{5}}{2}$ and $\bar{\phi} = \frac{1-\sqrt{5}}{2}$. Now lets find the associated eigenvectors.
So we need to solve

$$(A - \lambda I)\left(\begin{smallmatrix} x \\ y \end{smallmatrix}\right) = \left(\begin{smallmatrix} 0 \\ 0 \end{smallmatrix}\right)$$

for $\lambda = \phi$ and $\bar{\phi}$. First, let $\lambda = \phi$ then

$$\begin{pmatrix} \frac{1-\sqrt{5}}{2} & 1 \\ 1 & \frac{-1-\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} \phi \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Similarly for $\lambda = \bar{\phi}$ we have,

$$\begin{pmatrix} \frac{1+\sqrt{5}}{2} & 1 \\ 1 & \frac{-1+\sqrt{5}}{2} \end{pmatrix} \begin{pmatrix} 1 \\ -\phi \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

Hence we have our two eigenvectors, $\left(\begin{smallmatrix} \phi \\ 1 \end{smallmatrix}\right)$ (for $\lambda = \phi$) and $\left(\begin{smallmatrix} 1 \\ -\phi \end{smallmatrix}\right)$ (for $\lambda = \bar{\phi}$). We want to use this information in conjunction with Theorem 3.4, so let

$$T := \begin{pmatrix} \phi & 1 \\ 1 & -\phi \end{pmatrix}$$

Then,

$$T^{-1}AT = \left(\begin{smallmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{smallmatrix}\right)$$

3.3 Finding F_n

Recall our sequence $F_n = F_{n-1} + F_{n-2}$, can be written as

$$\begin{aligned}
\begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} &= \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} F_{n-1} \\ F_{n-2} \end{pmatrix} && (3.19)\\
&= A \cdot A \begin{pmatrix} F_{n-2} \\ F_{n-3} \end{pmatrix} && (3.20)\\
&= A^2 \begin{pmatrix} F_{n-2} \\ F_{n-3} \end{pmatrix} && (3.21)\\
&= \vdots && (3.22)\\
&= A^{n-1} \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} && (3.23)\\
&= A^{n-1} \begin{pmatrix} F_1 \\ F_0 \end{pmatrix} && (3.24)\\
&= A^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} && (3.25)
\end{aligned}$$

Now

$$T^{-1}AT = \begin{pmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{pmatrix} \tag{3.26}$$

$$A = T \begin{pmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{pmatrix} T^{-1} \tag{3.27}$$

Let $D := \begin{pmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{pmatrix}$, then

$$A^n = (TDT^{-1}) \cdot (TDT^{-1}) \cdot (TDT^{-1}) \cdots (TDT^{-1}) \tag{3.28}$$

$$= TD^nT^{-1} \tag{3.29}$$

$$= T \begin{pmatrix} \phi & 0 \\ 0 & \bar{\phi} \end{pmatrix}^n T^{-1} \tag{3.30}$$

$$= T \begin{pmatrix} \phi^n & 0 \\ 0 & \bar{\phi}^n \end{pmatrix} T^{-1} \tag{3.31}$$

Hence we finally have,

$$\begin{aligned} \begin{pmatrix} F_n \\ F_{n-1} \end{pmatrix} &= A^{n-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= T \begin{pmatrix} \phi^{n-1} & 0 \\ 0 & \bar{\phi}^{n-1} \end{pmatrix} T^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ &= \frac{1}{\phi^2 - 1} \begin{pmatrix} \phi & 1 \\ 1 & -\phi \end{pmatrix} \begin{pmatrix} \phi^{n-1} & 0 \\ 0 & \bar{\phi}^{n-1} \end{pmatrix} \begin{pmatrix} -\phi & -1 \\ -1 & \phi \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \end{aligned}$$

After this it is a simple exercise to show

$$F_n = \frac{1}{\sqrt{5}}\phi^n - \frac{1}{\sqrt{5}}\bar{\phi}^n \tag{3.32}$$

Part III

Sample Exam questions

1.

1. Find the inverse of 15 modulo 26.

2. The enciphered message

$$VTQ$$

was produced by applying the enciphering function $f_E\colon \mathbb{Z}_{26} \to Z_{26}, x \mapsto 15x + 3$ to single letter message units over the alphabet $A = 0, B = 1, \ldots, Z = 25$. Determine the corresponding deciphering function and use it to decipher the message.

3. Factorise 120 as a product of primes. Then calculate the number $\phi(120)$ of integers from 1 to 120 that are coprime to 120. Finally, calculate

$$7^{34} \mod 120.$$

4. Determine the smallest positive integer x that satisfies

$$x \equiv 4 \mod 5,$$
$$x \equiv 3 \mod 7.$$

2.

1. The ciphertext

$$WMHCYMRRQSDD$$

was produced by applying the function

$$f_E\colon \begin{pmatrix} x \\ y \end{pmatrix} \mapsto A \begin{pmatrix} x \\ y \end{pmatrix} \qquad \text{with } A = \begin{pmatrix} 4 & 5 \\ 1 & 5 \end{pmatrix}$$

to 2-letter message units over the alphabet $A = 0, B = 1, \ldots, Z = 25$. Use your answer to Question 1(a) to calculate $A^{-1} \mod 26$, and hence determine the first FOUR letters of plaintext.

2. Use row operations to find the inverse of

$$A = \begin{pmatrix} 1 & 2 & 2 \\ 4 & 4 & 4 \\ 5 & 3 & 4 \end{pmatrix}.$$

3.

1. Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be reflection in the line $y = -x$ and let $g\colon \mathbb{R}^2 \to \mathbb{R}^2$ be anticlockwise rotation through 90^o about the origin. Find the point $v = (x, y) \in \mathbb{R}^2$ such that $g(f(v)) = (3, 4)$.

2. Consider

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}, \quad \phi = \frac{1+\sqrt{5}}{2}, \quad v_1 = \begin{pmatrix} \phi \\ 1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ -\phi \end{pmatrix}.$$

   (i) Determine the characteristic polynomial $p_A(\lambda)$ of A.

   (ii) Verify that v_1 and v_2 are eigenvectors of A and determine the corresponding eigenvalues.

   (iii) Determine the eigenvalues of A^{-1}.

Part IV

Sample Exercises

Some exercises

stuff