1202 Algebra 2 Notes

Based on the 2016 spring lectures by Dr M L Roberts

18/01/16 1202

L1

Algebra 2

1. Number Theory
2. Groups
3. Linear Algebra (Determinants & Diagonalising)

Problem sheets on Fridays.

Chapter 1 - Number Theory

Here we are looking at the properties of $\mathbb{N}$, the natural numbers, and $\mathbb{Z}$, the integers.

Def 1.1

Let $a, b \in \mathbb{Z}$, $a \neq 0$. then a divides b (written $a|b$) if there is a $z \in \mathbb{Z}$ such that $b = az$.

We also say is a divisor or factor, or b is a multiple of a.

e.g $2|6$ since $6 = 2 \times 3$, but $2 \nmid 7$.

Proposition 1.2

Let $a, b, c, d, e \in \mathbb{Z}$, $a \neq 0$, then

(i) $a|b$ and $a|c$ then $a \,|\, bd + ce$

(ii) $a|b$ and $b|c$ then $a|c$

(iii) $a|b$ and $b|a$ then $b = \pm a$

Proof

(i) $b = ax$, $c = ay$ for some $x, y \in \mathbb{Z}$ then

$$bd + ce = axd + aye$$
$$= a(xd + ye)$$

$(xd + ye) \in \mathbb{Z}$ so $a \,|\, bd + ce$

(ii) similar

(iii) $b = ax$, $a = by$ for some $x, y \in \mathbb{Z}$

$a = by = axy$

$\Rightarrow a(1 - xy) = 0$, $a \neq 0$ so $xy = 1$

$\therefore x = \pm 1$, $y = \pm 1$, so $b = \pm a$.

Def 1.3
We say that a factorisation $a = bc$ is called trivial if $b = \pm 1$ or $c = \pm 1$.
If a has a non trivial factorisation it is called composite.
If $a > 1$ and a has no non-trivial factorisation then a is a prime.
e.g. 6 is composite, $6 = 2 \times 3$, 7 is prime, $7 = ab \Rightarrow a = \pm 1$ or $b = \pm 1$.

We can divide integers into:
- composites
- primes
- -ve primes
- units (+1 or -1)

The fundamental result about primes is that every positive integer factorises uniquely into primes.
e.g. $36 = 2 \times 2 \times 3 \times 3$.
In order to prove this, we need to develop some results about division.

The division theorem
Theorem 1.4
Let $a, b \in \mathbb{Z}$, $b > 0$.
Then there exists unique integers q and r such that $a = bq + r$ and $0 \leq r < b$.
e.g. $a = 27$, $b = 4$: $27 = 4 \times 6 + 3$
$a = -24$, $b = 5$: $-24 = 5 \times -5 + 1$

Proof
Let q be the largest integer $\leq \frac{a}{b}$
Then $\frac{a}{b} = q + \alpha$ $\quad 0 \leq \alpha < 1$
$a = bq + \alpha b$ Since $\alpha b = a - bq$, $\alpha b \in \mathbb{Z}$

L1

$0 \le \alpha b < b$

Taking $r = \alpha b$ gives the required number.

Suppose

$a = bq + r = bq' + r' \quad (0 \le r < b,\ 0 \le r' < b)$

Then $b(q - q') = r' - r$

$|r - r'| < b$

but $(r - r')$ is a multiple of b

$\therefore r - r' = 0 \quad \therefore r = r'$ and $q = q'$.

q is called the quotient and r the remainder.

Def 1.5

Let $a, b \in \mathbb{Z}$, $a, b \neq 0$.

Then the hcf (highest common factor) or gcd (greatest common divisor) of a and b is the largest positive integer d such that $d | a$ and $d | b$. Write $d = hcf(a, b)$.

eg. $hcf(6, 8) = 2$ as $2 | 6$, $2 | 8$.

We say a, b are coprime if $hcf(a, b) = 1$

Euclid's algorithm

Theorem 1.6

Let a, b be positive integers. Then there exists $n \in \mathbb{N}$, $q_1, \ldots, q_{n+1}, r_1, \ldots, r_n \in \mathbb{Z}$ with $b > r_1 > r_2 > \ldots > r_n > 0$

and $b = aq_1 + r_1$

$a = r_1 q_2 + r_2$

$r_1 = r_2 q_3 + r_3$

$\vdots$

$r_{n-2} = r_{n-1} q_n + r_n$

$r_{n-1} = r_n q_{n+1} \ (+0)$

Then $r_n = hcf(a, b)$

eg. find $hcf(1169, 560)$

$1169 = 560 \times 2 + 49$

$560 = 49 \times 11 + 21$

$49 = 21 \times 2 + 7$

$21 = 7 \times 3$

$\therefore hcf = 7.$

Ex

Find hcf (30, 18) by this method.

$30 = 18 \times 1 + 12$

$18 = 12 \times 1 + 6$

$12 = 6 \times 2$

$\therefore hcf = 6$

22/01/16 | 1201

L2

Thm 1.6

Let a, b be positive integers.
Then there is a positive integer n and integers $q_1, \dots, q_{n+1}, r_1, \dots, r_n$ with $b > r_1 > r_2 > \dots > r_n > 0$ and

$$a = bq_1 + r_1$$
$$b = r_1 q_2 + r_2$$
$$r_1 = r_2 q_3 + r_3$$
$$\vdots$$
$$r_{n-2} = r_{n-1} q_n + r_n$$
$$r_{n-1} = r_n q_{n+1}$$

Then $r_n = hcf(a, b)$.

Proof

The existence of the q_i, r_i and that $b > r_1 > r_2 > \dots$ etc follows from division theorem.
Since $r_1 > r_2 > \dots$ is a strictly decreasing sequence of non-negative integers, at some stage it is zero, say $r_{n+1} = 0$.
Now prove (i) $r_n | a$ and $r_n | b$
and (ii) $x | a$ and $x | b \Rightarrow x | r_n$.

(i) $r_{n-1} = r_n q_{n+1}$ so $r_n | r_{n-1}$.
$r_{n-2} = r_{n-1} q_n + r_n$, $r_n | r_{n-1}$ and $r_n | r_n$.
By 1.2 (i), $r_n | (r_{n-1} q_n + r_n) = r_{n-2}$.
Continuing up the list of equations, we get
$r_n | r_{n-3}$, $r_n | r_{n-4}$, ..., $r_n | r_1$, $r_n | b$, $r_n | a$.

(ii) Suppose $x | a$ and $x | b$.
Then $x | a - bq_1 = r_1$
Thus $x | b$ and $x | r_1$
$x | b - r_1 q_2 = r_2$
Continuing, $x | r_3$, ..., $x | r_n$.

Note this this proof actually shows that any common divisor of a and b actually divides $r_n = hcf(a, b)$.

Linear Combinations and the h, k - lemma

Def 1.7

A linear combination of integers a and b is an integer of the form $ax + by$ $(x, y \in \mathbb{Z})$.

e.g. 20 is a linear combination of 6 and 8,

as $20 = 6 \times 2 + 8 \times 1$

2 also: $2 = 6 \times (-1) + 8 \times 1$

1 is not a linear combination of 6 and 8.

Thm 1.8

Let a, b be non-zero integers and x an integer. Then x is a linear combination of a and b $\Leftrightarrow \text{hcf}(a, b) \mid x$

Proof

[$\Rightarrow$] Let x be a linear combina of , d b

Then $\text{hcf}(a, b) \mid a$ and $\text{hcf}(a, b) \mid b$, so $\text{hcf}(a, b) \mid x$.

[$\Leftarrow$] We need to show that $\text{hcf}(a, b)$ is a linear combination of a and b.

Use Thm 1.6 and rewrite equations:

$$r_1 = a - bq_1$$
$$r_2 = b - r_1 q_2$$
$$r_3 = r_1 - r_2 q_3$$
$$\vdots$$
$$r_n = r_{n-2} - r_{n-1} q_n$$
$$0 = r_{n-1} - r_n q_{n+1}$$

Write $L(p, q)$ for the set of linear combinations of p and q.

So $r_n \in L(r_{n-2}, r_{n-1})$

$r_{n-1} \in L(r_{n-3}, r_{n-2})$

Hence $r_n \in L(r_{n-3}, r_{n-2})$

Continue to get $r_n \in L(r_{n-4}, r_{n-3}), \ldots, r_n \in L(a, b)$. $\square$

22/01/16 1202

L2

This is easiest to see in an example.

Example

hcf(5, 7) = 1

Express 1 as a linear combination of 5 and 7.

$7 = 5 \times 1 + 2$

$5 = 2 \times 2 + 1$

$1 = 5 - 2 \times 2$

$= 5 - (7-5) \times 2$

$= 5 \times 3 - 7 \times 2$

Repeat with 42 and 19.

$42 = 19 \times 2 + 4$

$19 = 4 \times 4 + 3$

$4 = 3 \times 1 + 1$

$1 = 4 - 3 \times 1$

$= 4 - (19 - 4 \times 4)$

$= 4 - (19 - (42 - 19 \times 2) \times 4)$

$= (42 - 19 \times 2) - 19 + (42 - 19 \times 2) \times 4$

$= 42 \times 5 - 19 \times 11$

The part of this theorem we will use is:

Lemma 1.9 ("h, k - lemma")

Let a and b be coprime integers then $\exists$ integers h and k such that $ah + bk = 1$.

Unique Factorisation

Crucial result is:

Prop 1.10

Let p be prime, $a, b \in \mathbb{Z}$

Then $p | ab \Rightarrow p | a$ or $p | b$.

Proof

Assume $p | ab$.

Let $d = hcf(a, b)$.

Since p is prime, $d = 1$ or $d = p$.

Case 1: $d = p$

Then $p = d | a$

Case 2: $d = 1$

i.e. a and p are coprime.

Then $\exists h, k \in \mathbb{Z}$ s.t. $ah + pk = 1$.

Then $b = bah + bpk$

$p | ab$, $p | p$

$\therefore$ $p | bah + bpk = b$

This easily extends to:

Prop 1.11

Let p be a prime, $a_1, \ldots, a_n$ integers.

Then $p | a_1 \ldots a_n \Rightarrow p | a_i$ for some i.

Thm 1.12 (Unique Factorisation in $\mathbb{Z}$)

Let z be a positive integer. Then z can be written as a product of primes $z = p_1, \ldots, p_n$ (p_i not necessarily distinct primes) and this is unique (up to the order).

22/01/16 1202

L2

Proof

first prove existence of such a factorisation.

$z=1$ is trivial (product of no primes).

Suppose $2, 3, \dots, z-1$ can all be written as a product of primes.

z is either a prime or not.

If z is prime, it is a product of one prime, itself.

If z is not prime, $z=ab$, $1<a,b<z$.

Then by inductive hypothesis, a and b can be written as a product of primes:

hence so can $z=ab$.

Result follows by induction.

Now uniqueness.

Let $P(n)$ denote statement.

If $z=p_1 \dots p_n = q_1 \dots q_m$ where $p_1, \dots, p_n, q_1, \dots, q_m$ are primes, then $m=n$ and $q_1 \dots q_n$ is a re-ordering of $p_1 \dots p_n$.

$P(1)$ is immediately true.

Suppose $P(n-1)$ holds

Let $z = p_1 \dots p_n = q_1 \dots q_m$

Now $p_n | q_1 \dots q_m$

By Corollary 1.11, $p_n | q_i$ for some i.

But q_i is prime, so $p_n = q_i$

$z = p_1 \dots p_n = q_1 \dots q_{i-1} q_i q_{i+1} \dots q_m$

Cancel $q_i = p_n$ to get

$p_1 \dots p_{n-1} = q_1 \dots q_{i-1} q_{i+1} \dots q_m$

By $P(n-1)$, $n-1 = m-1$

and $q_1 \dots q_{i-1} q_{i+1} \dots q_m$ is a reordering of $p_1 \dots p_{n-1}$

Hence $m=n$ and $q_1 \dots q_i \dots q_m$ is a reordering of $p_1 \dots p_n$.

Ie. $P(n)$ holds.

Thus $P(n-1) \Rightarrow P(n)$.

By induction, $P(n)$ holds for all n.

e.g. $120 = 2 \times 2 \times 2 \times 3 \times 5$ (uniquely)

It is worth noting that there are other possible number systems, eg.

$\mathbb{Z}[i] = \{a + bi : a, b \in \mathbb{Z}\}$ (Gaussian integers)

or $\mathbb{Z}[\sqrt{-5}] = \{a + b\sqrt{-5} : a, b \in \mathbb{Z}\}$

in which we can define addition, multiplication, divisibility, primes, etc, <u>and</u> some of these have unique factorisation into primes, some don't.

$\mathbb{Z}[i]$ has unique factorisation and this can to prove results about the integers.

(See EX 1 Q4)

$\mathbb{Z}[\sqrt{-5}]$ does not have unique factorisation,

$6 = 2 \times 3 = (1 + \sqrt{-5})(1 - \sqrt{-5})$

and one can show $2, 3, 1+\sqrt{-5}$, -5 are all primes.

(See Excercises for similar examples)

One of the earliest results about primes is:

<u>Theorem 1.14</u> (Euclid)

There are infinitely many prime numbers.

<u>Proof</u>

Suppose not, say $p_1, \dots, p_n$ are all the primes

Let $N = p_1 \dots p_n + 1$.

N may or may not be prime, but it must have a prime factor.

But $p_1 \nmid N, \dots, p_n \nmid N$.

$\therefore$ Thus this prime factor is different from $p_1, \dots, p_n$. #

$\therefore$ There are infinitely many primes.

Another way of looking at the proof is that it gives a way of keeping on constructing new primes.

e.g. $2, 3$: $N = 2 \times 3 + 1 = 7$, $2, 3, 7$

22/01/16 1202

L2

$N = 2 \times 3 \times 7 + 1 = 43$, $\quad 2, 3, 7, 43$

$N = 2 \times 3 \times 7 \times 43 + 1 = 1807 = \underline{13} \times 139$; $2, 3, 7, 43, 13$

Chapter 2 - Groups

Definition & basic properties

Def 2.1

A group is a set G with a binary operation, $*$ on G such that

(i) $*$ is associative

(ii) G has an identity element under $*$

(iii) Each element of G has an inverse (under $*$).

Here a binary operation is a rule assigning to a pair of elements $g, h \in G$ another element of G denoted $g*h$. [This is sometimes called a closed binary operation to emphasise that $g*h \in G$.]

Formally it is a map $G \times G \to G$ $(g,h) \to g*h$.

$*$ is associative if $\forall f,g,h \in G$ $(f*g)*h = f*(g*h)$.

e is an identity element if $\forall g \in G$, $e*g = g*e = g$

$h \in G$ is an inverse of g if $g*h = e = h*g$.

If G is a group under $*$ and $g*h = h*g$ $\forall g,h \in G$ then G is called abelian or commutative.

Examples

(i) $G = \mathbb{Z}$, the integers, $* = +$, normal addition. (abelian group)

so $a+(b+c) = (a+b)+c$, $a+0 = 0+a = a$, $a+(-a) = 0$.

(ii) $G = \mathbb{R} - \{0\}$, $* =$ multiplication. (abelian group)

so $a(bc) = (ab)c$, $a \cdot 1 = a = 1 \cdot a$, $a \cdot (\frac{1}{a}) = 1$.

(iii) $G = GL_2(\mathbb{R}) =$ 2×2 invertable matrices over $\mathbb{R}$, $* =$ multiplication

$A(BC) = (AB)C$, $AI_2 = A = I_2 A$, $A \cdot A^{-1} = I_2 = A^{-1} \cdot A$ (non-abelian group).

(iv) The definition of a field F is:

ⓐ F is a group under $+$ (abelian group)

ⓑ $F - \{0\}$ is a group under $\times$ (or $\cdot$) (abelian groups)

ⓒ $a(b+c) = ab + ac$ $\forall a,b,c \in F$.

Associativity

Many familiar binary operations are associative eg. $+, \cdot$ on $\mathbb{R}$, multiplication of matrices, composite functions.

$a * b = a/b$ is not associative,

eg. $(2*2)*2 = (2/2)/2 = \frac{1}{2}$

$2*(2*2) = 2/(2/2) = 2$

Exercise

Determine whether or not the following are asso ve:

(i) $*$ on $M_2(\mathbb{R})$ (2x2 matrices) by $A*B = AB - BA$

$(A*B)*C = (AB-BA)*C = (AB-BA)C - C(AB-BA)$

$A*(B*C) = A(B*C) - (B*C)A = A(BC-CB) \quad (BC-CB)A$

$ABC - BAC - CAB + CBA \neq ABC - ACB \quad BCA + CBA$

so not associative. (needs an example - use basic matrices)

(ii) $*$ on $\mathbb{R}$ by $a*b = ab + a + b$

$(a*b)*c = (ab+a+b)*c = ab+a \; b)c \quad (ab+a+b)+c$

$a*(b*c) = a(b*c) + a + (b*c) = a(bc+b+c) + a + (b \quad)$

$abc + ac + bc + ab + a + b + c = abc + ab + ac + a + bc + b + c$

so associative.

Notice that associativity extends to more than 3 elements

eg $(a*b)*(c*d) = (a*(b*c))*d$

Identity elements

Lemma 2.3

Let $*$ be a binary operation on a set G and e and f both identity elements. Then $e = f$.

Proof

$e = e*f = f$

f identity e identity

Hence the identity element (if it exists) is unique.

Thus we can say the identity element in a group.

24/01/16 1202

L4

Identity

$e * g = g = g * e \quad \forall g \in G$

Ex

Determine which of the following have an identity element:

(i) $*$ on $\mathbb{R}$ by $a*b = ab + a + b$ — yes, $b = 0$ as $a*0 = a = 0*a$

(ii) $*$ on $\mathbb{R}$ by $a*b = a$ — no.

Suppose e identity, $e*1 = e$ by def $*$

$e*1 = 1$ by def identity

$\therefore e = 1$

Similarly $e = e*2 = 2$

so $e = 1 = 2$ ※.

Inverses

Lemma 2.4

Let G be a set and $*$ be an associative binary operation on G with identity element e. Then if g and h are both inverses of $f \in G$, $g = h$.

Proof

Consider $g*f*h$

$$(g*f)*h \underset{\text{as } g \text{ is inverse of } f}{=} e*h \underset{\text{def. of identity}}{=} h$$

$$g*(f*h) \underset{h \text{ inverse of } f}{=} g*e \underset{\text{def. of identity}}{=} g$$

Since $*$ is associative, $g = h$.

Hence in particular in a group G, any element g has a unique inverse, usually denoted g^{-1}.

Lemma 2.5

Let G be a group, $g, h \in G$. Then

(i) $(g^{-1})^{-1} = g$

(ii) $(g * h)^{-1} = h^{-1} * g^{-1}$

Proof

(i) By def^n of g^{-1}, $g * g^{-1} = e = g^{-1} * g$

Hence g is the solution to

$x * g^{-1} = e = g^{-1} * x$

$\therefore g = (g^{-1})^{-1}$

(ii) $(h^{-1} * g^{-1}) * (g * h)$

$= h^{-1} * (g^{-1} * g) * h$

$= h^{-1} * e * h$

$= h^{-1} * h = e$

$= (g * h) * (h^{-1} * g^{-1})$

By def^n $(g * h)^{-1} = h^{-1} * g^{-1}$. $\square$

[Ex of (ii): $f(x) = \sin(x^2)$ (not invertable)
$g(y) = \sqrt{\sin^{-1}(y)}$]

Ex

For each of the following find which elements have inverses & in this case what the inverse is:

(i) $G = \mathbb{R} - \{-1\}$ $a * b = ab + a + b$

$a * x = 0 = a + x + ax$

so $x = \dfrac{-a}{a+1}$

Since $a \neq -1$, $\dfrac{-a}{a+1} \in \mathbb{R}$.

Also $\dfrac{-a}{a+1} \neq -1$ $\left(\dfrac{-a}{a+1} = -1 \Rightarrow -a = -a - 1 \Rightarrow 0 = 1\right)$

$\therefore \dfrac{-a}{a+1} \in G$

$a * \dfrac{-a}{a+1} = a\left(\dfrac{-a}{a+1}\right) + a - \dfrac{a}{a+1} = \dfrac{a(a+1) - a - a^2}{a+1} = 0$

$\therefore$ every element $a \in G$ has inverse $\frac{-a}{a+1}$. In fact G forms a group under $*$, where $a * b = ab + a + b$.

29/01/16 1202

L4

Notation

We often write gh instead of $g*h$ in a general group.

Def 2.6

Define $g^2 = gg$, $g^3 = ggg$ (well-defined since we are assuming associativity), etc..., $g^0 = e$, $g^{-n} = (g^{-1})^n$.

Normal rules of indices apply.

Lemma 2.7

$\forall m, n \in \mathbb{Z}$,

(i) $g^m g^n = g^{m+n}$

(ii) $(g^m)^n = g^{mn}$.

However $(gh)^n \neq g^n h^n$ in general (true in an abelian group).

Prop 2.8

(i) Let G be a group, $f, g, h \in G$.

$fg = fh \Rightarrow g = h$

$gf = hf \Rightarrow g = h$ (cancellation)

(ii) Let G be a group, $g \in G$.

Then $gG = \{gx : x \in G\}$ contains each element of G exactly once

In particular, if G is finite, $G = \{g_1, \dots, g_n\}$, then the list $gg_1, gg_2, \dots, gg_n$ contains each element of g once, i.e. it is a re-ordering of $g_1, \dots, g_n$

Proof

(i) $fg = fh$

$\Rightarrow f^{-1}(fg) = f^{-1}(fh)$

$\Rightarrow (f^{-1}f)g = (f^{-1}f)h$

$\Rightarrow eg = eh \quad \Rightarrow g = h$ (second part similar)

(ii) Consider $\varphi : G \to G$ given by $\varphi(x) = gx$.

By (i), φ is injective.

$[\varphi(x) = \varphi(y) \Rightarrow$ g [illegible] g [illegible] $]$

φ is also surjective, since x $\varphi(g^{-1}x)$

Examples of groups

Lemma 2.9

Let X be a set and let

$S(X) = \{f : X \to X, f \text{ bijective}\}$

Then $S(X)$ forms a group under the operation of composition.

Proof

$\circ$ is a closed binary operation on $S(X)$, since f, g bijective $\Rightarrow f \circ g$ bijective (1201).

$\circ$ is associative.

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

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

Here $(f \circ g) \circ h = f \circ (g \circ h)$

id is defined by $id(x) = x \quad \forall x \in X$ is the identity element $(f \circ id)(x) = f(id(x)) = f(x)$.

So $f \circ id = f$

If $f \in S(X)$, then f has an inverse function f (since f is bijective: 1201) and $f f^{-1} = id = f$ f.

L4

One particular case is when $X = \{1, \dots, n\}$

Def 2.10

If $X = \{1, \dots, n\}$ then $S(X)$ is denoted S_n and is called the symmetric group: elements of S_n are called permutations (cf 1201).

$S(X)$ can be called the automorphism group $\mathrm{Aut}(X)$ of X. If X has some kind of structure (eg. vectorspace), then $\mathrm{Aut}(X)$ is defined to be the bijections $X \to X$ that preserve the structure.

e.g. if V is a vectorspace over $\mathbb{R}$.

$$\mathrm{Aut}(V) = \{f : V \to V, \text{ bijective s.t. } f(\underline{v}_1 + \underline{v}_2) = f(\underline{v}_1) + f(\underline{v}_2) \text{ and } f(\lambda \underline{v}_1) = \lambda f(\underline{v}_1), \ \forall \lambda \in \mathbb{R}, \underline{v}_1, \underline{v}_2 \in V\}.$$

$\mathrm{Aut}(X)$ provides information about the object X.

Def 2.11

Let n be a fixed positive integer. For $a, b \in \mathbb{Z}$ write $a \equiv b \pmod{n}$ and say a is congruent to $b \pmod{n}$ if $n | b-a$. Let $\bar{x} = \{z \in \mathbb{Z} : z \equiv x \pmod{n}\}$.

If $m \in \mathbb{Z}$, by the division theorem, m can be written as $m = nq + r$ where $0 \le r < n$. This m is congruent to exactly one of $0, 1, 2, \dots, n-1$, ie. m lies in exactly one of the sets $\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}$.

Let $\mathbb{Z}_n = \{\bar{0}, \bar{1}, \bar{2}, \dots, \overline{n-1}\}$.

eg. $2 \equiv 5 \pmod{3}$

$7 \equiv 107 \pmod{10}$

mod 3, any number is congruent to exactly one of $0, 1, 2$,

$\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$, $\bar{0} = \{\dots, -3, 0, 3, 6, \dots\}$, $\bar{1} = \{\dots, -2, 1, 4, 7, \dots\}$,

$\bar{2} = \{\dots, -1, 2, 5, 8, \dots\}$

Lemma 2.12

Let $n \in \mathbb{N}$. If $a \equiv b \pmod{n}$ and $c \equiv d \pmod{n}$ then
(i) $a + c \equiv b + d \pmod{n}$
(ii) $ac \equiv bd \pmod{n}$.
Hence the operators $+$ and $\times$ on $\mathbb{Z}_n$ given by
$\bar{a} + \bar{b} = \overline{a+b}$, $\bar{a} . \bar{b} = \overline{ab}$ are well defined.

e.g. mod 3: $\bar{2} + \bar{2} = \bar{4} = \bar{1}$
$\overline{14} + \bar{5} = \overline{19} = \bar{1}$

Proof

(i) $b - a = nx$
$d - c = ny$
$(b+d) - (a+c) = n(x+y)$
so $b + d \equiv a + c \pmod{n}$

(ii) $bd - ac = (a + nx)(c + ny) - ac$
$= nxc + nay + n^2xy$
$= n(xc + ay + nxy)$
so $ac \equiv bd \pmod{n}$

Theorem 2.18

(a) $\mathbb{Z}_n$ under $+$ forms a(n abelian) group.
(b) For any prime p, $\mathbb{Z}_p^* = \mathbb{Z}_p - \{\bar{0}\}$ forms a(n abelian) group under multiplication.

Proof

(a) This follows from the fact that $\mathbb{Z}$ is a group under $+$ e.g.
$\bar{a} + (\bar{b} + \bar{c}) = \bar{a} + \overline{b+c} = \overline{a+(b+c)}$
$= \overline{(a+b)+c} = \overline{a+b} + c + (\bar{a} + \bar{b}) + \bar{c}$
(b) First note that multiplication does given a (closed) binary operation on $\mathbb{Z}_p^*$: let $\bar{x}, \bar{y} \in \mathbb{Z}_p^*$, $\bar{x} \neq \bar{0}$, $\bar{y} \neq \bar{0}$,

so $p \nmid x$, $p \nmid y$.
If $\bar{x}, \bar{y} = \bar{0}$, $\overline{xy} = \bar{0}$, ie $p | xy$.
Since p is prime, this would imply $p | x$ or $p | y$. ※
$\therefore$ $\overline{x}\overline{y} \neq 0$, i.e. $\overline{x}\overline{y} \in \mathbb{Z}_p^*$.

Since multiplication on $\mathbb{Z}$ is associative, it is also associative on $\mathbb{Z}^*_p$.
The identity is $\bar{1}$.
Need to prove the existence of inverse. Two alternative proofs:
1). Fix $\bar{a} \in \mathbb{Z}_p^*$. Consider the set $\{\overline{1 \cdot a}, \overline{2a}, \dots, \overline{(p-1) \cdot a}\} \in \mathbb{Z}_p^*$
These are all distinct if $\overline{xa} = \overline{ya}$, then $p | xa - ya = (x-y)a$.
$p \nmid a$, so $p | x-y$.
So $1 \leq x, y \leq p-1$
$\therefore$ $|x-y| < p$ $\therefore$ $x - y = 0$, $x = y$
Since there are $p-1$ elements, $\{\overline{1 \cdot a}, \overline{2a}, \dots, \overline{(p-1)a}\} = \mathbb{Z}_p^*$.
$\therefore$ One of these elements, say $\overline{ra} = \bar{1}$
$\therefore$ $\bar{r} = \bar{a}^{-1}$

01/02/16 | 1202

L5

<u>Thm 2.13 (6)</u>

$\mathbb{Z}_p^* = \{\bar{1}, \bar{2}, \ldots, \overline{p-1}\}$ forms a group under multiplication.

$\vdots$

Second proof of existence of inverses:

Let $\bar{a} \in \mathbb{Z}_p^*$, $\bar{a} \neq 0$, so $p \nmid a$.

Since p prime, a and p are coprime, so by h,k - lemma, $\exists h, k \in \mathbb{Z}$ s.t. $ah + pk = 1$.

Then in $\mathbb{Z}_p$, $\bar{a}\bar{h} = \bar{1}$, ie. $\bar{h}$ is the inverse of $\bar{a}$.

Thus $\mathbb{Z}_p^*$ is a group. $\square$

The two proofs give two methods of finding $\bar{a}^{-1}$.

e.g. what is the inverse of $\bar{3}$ in $\mathbb{Z}_{11}^*$?

① look at $\bar{3} \times \bar{1} = \bar{3}$, $\bar{3} \times \bar{2} = \bar{6}$, $\bar{3} \times \bar{3} = \bar{9}$, $\bar{3} \times \bar{4} = \overline{12} = \bar{1}$,

so $\bar{3}^{-1} = \bar{4}$.

② $\underline{11} = \underline{3} \times 3 + \underline{2}$

$\underline{3} = \underline{2} \times 1 + \underline{1}$

$\underline{1} = \underline{3} - \underline{2}$

$= \underline{3} - (\underline{11} - \underline{3} \times 3)$

$= \underline{3} \times 4 - \underline{11}$

so $\bar{1} = \bar{3} \times \bar{4} \pmod{11}$

<u>Ex</u>

(i) Find $\bar{5}^{-1}$ in $\mathbb{Z}_{17}$ by both methods

(ii) Solve $5x \equiv 12 \pmod{17}$

(i) ① $\bar{5} \times \bar{1} = \bar{5}$, $\bar{5} \times \bar{2} = \overline{10}$, $\bar{5} \times \bar{3} = \overline{15}$, $\bar{5} \times \bar{4} = \overline{20} = \bar{3}$, $\bar{5} \times \bar{5} = \overline{25} = \bar{8}$,

$\bar{5} \times \bar{6} = \overline{30} = \overline{13}$, $\bar{5} \times \bar{7} = \overline{35} = \bar{1}$, so $\bar{5}^{-1} = \bar{7}$

② $\underline{17} = \underline{5} \times 3 + \underline{2}$

$\underline{5} = \underline{2} \times 2 + \underline{1}$

$\underline{1} = \underline{5} - \underline{2} \times 2 = \underline{5} - 2(\underline{17} - \underline{5} \times 3) = \underline{5} \times \underline{7}$

so $\bar{1} = \bar{5} \times \bar{7} \pmod{17}$

(ii) $\bar{7} \times \bar{5}x = \bar{7} \times \overline{12}$

so $x = \overline{84} = \overline{16}$

The simplest way of specifying a group is to give the <u>group table</u>

e.g. $G = \{a, b, c\}$

	a	b	c
a	a	b	c
b	b	c	a
c	c	a	b

tells us $b * c = a$

It is very easy to see the identity element in a group table, e.g. a is the identity above. The inverses are also apparent, e.g. a^{-1} , b , c

In fact each element must appear exacty once, ach row and each column.

Associativity is not evident, and just writing down a table which does have identity and inverses won't usually give a group.

e.g. $\mathbb{Z}_4, +$

	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$

$\mathbb{Z}_5^*, \times$

	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$
$\bar{2}$	$\bar{2}$	$\bar{4}$	$\bar{1}$	$\bar{3}$
$\bar{3}$	$\bar{3}$	$\bar{1}$	$\bar{4}$	$\bar{2}$
$\bar{4}$	$\bar{4}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

Symmetries

Def 2.14

(i) An <u>isometry</u> of the plane $\mathbb{R}^2$ is a bijective function $f: \mathbb{R}^2 \to \mathbb{R}^2$ which preserves the distance between points, i.e. $\forall \underline{x}, \underline{y} \in \mathbb{R}^2 \quad d(f(x), f(y)) = d(x, y)$,

eg rotations, reflections, translations are all isometries.

(ii) If T is a set of points in $\mathbb{R}^2$, then $\mathrm{Sym}(T)$ is the

01/02/16 1202

L5

set of isometries f such that $f(T) = T$.

$\left[f(T) = \{ f(\underset{\sim}{x}) : \underset{\sim}{x} \in T \} \right]$

The set of all isometries forms a (very big) group under composition, but we will look at $\mathrm{Sym}(T)$.

<u>Lemma 2.15</u>

Sym(T) forms a group under composition.

<u>Proof</u>

Let $f, g \in \mathrm{Sym}(T)$.

Then $f \circ g$ is a bijection and for all $\underset{\sim}{x}, \underset{\sim}{y} \in \mathbb{R}^2$

$$d((f\circ g)(\underset{\sim}{x}), (f\circ g)(\underset{\sim}{y})) = d(f(g(\underset{\sim}{x})), f(g(\underset{\sim}{y})))$$
$$= d(g(\underset{\sim}{x}), g(\underset{\sim}{y}))$$
$$= d(\underset{\sim}{x}, \underset{\sim}{y})$$

So $f \circ g$ is an isometry.

$(f\circ g)(T) = f(g(T)) = f(T) = T$

$\therefore f \circ g \in \mathrm{Sym}(T)$.

Composition of functions is associative.

$id \in \mathrm{Sym}(T)$, where $id(\underset{\sim}{x}) = \underset{\sim}{x} \;\; \forall \underset{\sim}{x} \in \mathbb{R}^2$.

If $f \in \mathrm{Sym}(T)$, f^{-1} exists (since f bijective) and

$d(f^{-1}(\underset{\sim}{x}), f^{-1}(\underset{\sim}{y})) = d(f(f^{-1}(\underset{\sim}{x})), f(f^{-1}(\underset{\sim}{y}))) = d(\underset{\sim}{x}, \underset{\sim}{y})$

So f^{-1} is an isometry, and $f^{-1}(T) = T$.

$\therefore f^{-1} \in \mathrm{Sym}(T)$

$\therefore \mathrm{Sym}(T)$ is a group

e.g. consider T an equilateral triangle. ${}_3\overset{1}{\Delta}_2$

Obvious elements of Sym(T) are:

id : ${}_3\overset{1}{\Delta}_2 \to {}_3\overset{1}{\Delta}_2$

x_1 = reflection in vertical line: ${}_3\overset{1}{\Delta}_2 \to {}_2\overset{1}{\Delta}_3$

x_2 = reflection in line from bottom left corner: ${}_3\overset{1}{\Delta}_2 \to {}_3\overset{2}{\Delta}_1$

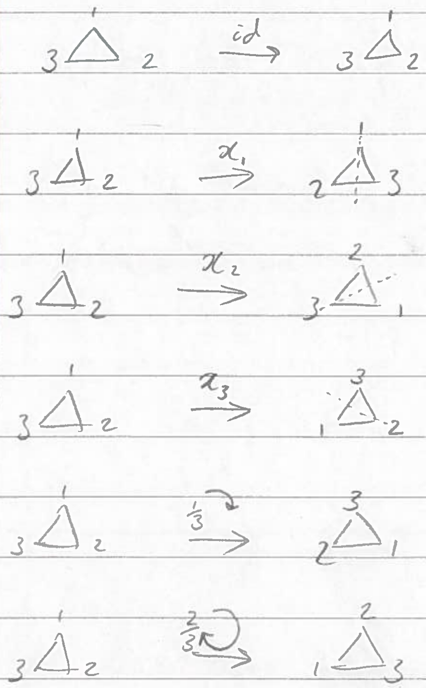
x_3 = reflection in line from bottom right corner: ${}_3\overset{1}{\Delta}_2 \to {}_1\overset{3}{\Delta}_2$

y_1 = rotation by $120°$ $\circlearrowright$: ${}_3\overset{1}{\triangle}_2 \to {}_2\overset{3}{\triangle}_1$

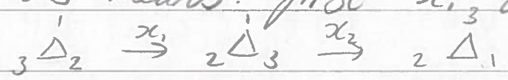
y_2 = rotation by $240°$ $\circlearrowleft$: ${}_3\overset{1}{\triangle}_2 \to {}_1\overset{2}{\triangle}_3$

05/01/16 1202

L6



Are there any more? No - there are 3 choices for where vertex 1 goes, then 2 choices for vertex 2 then no choices for vertex 3.
ie. no more than 6 symmetries.
$\therefore$ $Sym(T) = \{e, x_1, x_2, x_3, y_1, y_2\}$
The structure of $Sym(T)$ is given by how they compose.
e.g. what is $x_2 \circ x_1$?
We think of these as functions acting on the left, so this means: first x_1 then x_2.



$\therefore$ $x_2 \circ x_1 = y_1$
The direct way of showing the group structure is to write down the group table.

	e	x_1	x_2	x_3	y_1	y_2
e	e	x_1	x_2	x_3	y_1	y_2
x_1	x_1	e	y_2	y_1	x_3	x_2
x_2	x_2	y_1	e	y_2	x_1	x_3
x_3	x_3	y_2	y_1	e	x_2	x_1
y_1	y_1	x_2	x_3	x_1	y_2	e
y_2	y_2	x_3	x_1	x_2	e	y_1

A more efficient way of writing down the group structure is as follows:
Write $x = x_1$, $y = y_1$. Then every element of the group can be found by combining x and y:

$$y_2 = y_1^2 = y^2$$
$$yx = y_1 x_1 = x_2$$
$$y^2 x = y_2 x_1 = x_3$$

$$\therefore \text{Sym}(T) = \{e, y, y^2, x, yx, y^2 x\} \quad (*)$$

We say x and y <u>generate</u> Sym(T) and also (*) is a <u>normal form</u> for the elements.
We now need to know how to combine two elements from (*) to get the answer in the same form
eg. $(yx)(y^2x) = ?$
To do this we need enough relations, e.g.
$x^2 = e$, $y^3 = e$, $xy = x_1 y_1 = x_3 = y^2 x$.

Now
$$\begin{aligned}(yx)(y^2x) &= yxyyx \\ &= yy^2xyx \\ &= y^3xyx \\ &= xyx \\ &= y^2xx \\ &= y^2\end{aligned}$$

Thus the group structure is completely specified by generators x and y and relations $x^2 = e$, $y^3 = e$, $xy = y^2 x$.
Write:

$$\underbrace{\text{Sym}(T) = \langle \overbrace{x, y}^{\text{generators}} : \overbrace{x^2 = e,\ y^3 = e,\ xy = y^2 x}^{\text{relations}} \rangle}_{\text{presentation of Sym(T)}}$$

The order of an element and cyclic groups

Def^n 2.16

(i) The order of a group G, denoted $|G|$, is the number of elements in G.

(ii) Let G be a group, $g \in G$. Then the order of g, denoted $o(g)$, is the least positive integer n s.t. $g^n = e$, or ∞ if no such element exists.

eg. in last example

$o(x_1) = 2$, $x_1 \neq e$, $x_1^2 = e$

$o(y_1) = 3$, $y_1 \neq e$, $y_1^2 \neq e$, $y_1^3 = e$

$o(e) = 1$

Ex

Find the orders of:

(i) $\bar{2}$ in $\mathbb{Z}_6$, $o(\bar{2} \text{ in } \mathbb{Z}_6) = 3$ [note $a^2 = a * a$]

$\bar{3}$ in $\mathbb{Z}_6$, $o(\bar{3} \text{ in } \mathbb{Z}_6) = 2$ [here '$*$' = '$+$']

$\bar{5}$ in $\mathbb{Z}_6$, $o(\bar{5} \text{ in } \mathbb{Z}_6) = 6$

(ii) $\bar{2}$ in $\mathbb{Z}_5^*$, $o(\bar{2} \text{ in } \mathbb{Z}_5^*) = 4$

$\bar{3}$ in $\mathbb{Z}_5^*$, $o(\bar{3} \text{ in } \mathbb{Z}_5^*) = 4$

(iii) 2 in $\mathbb{R} - \{0\}$ under multiplication, $o(2) = \infty$

-1 in $\mathbb{R} - \{0\}$ under multiplication, $o(-1) = 2$

Lemma 2.17

Let G be a group, $g \in G$.

(a) Suppose $o(g) = n < \infty$. Then

(i) $g^m = e \Leftrightarrow n \mid m$

(ii) any power of g is equal to exactly one of the elements $e, g, g^2, \ldots, g^{n-1}$.

(b) Suppose $o(g) = \infty$. Then any power of g is equal to exactly one of $\ldots g^{-2}, g^{-1}, e, g, g^2, \ldots$

Proof

(a) (i) ($\Leftarrow$)

If $n | m$ then $m = nr$. for some $r \in \mathbb{Z}$

Hence $g^m = g^{nr} = (g^n)^r = e^r = e$

($\Rightarrow$)

Suppose $g^m = e$. Write $m = nq + r$, $0 \leq r < n$.

Then $g^r = g^{m-nq} = (g^m)(g^n)^{-q}$

$= e \cdot e^{-q} = e$

Now $0 \leq r < n$ and by def of n as $o(g)$, there is no positive integer $< n$ so that g to that power is e. $\therefore r = 0$, so $n | m$.

(ii) Let $m \in \mathbb{Z}$ then $m = nq + r$

$0 \leq r < n$, so

$g^m = g^{nq+r}$

$= (g^n)^q g^r$

$= e^q g^r$

$= g^r$

Also $g^r = g^s$, $0 \leq r, s < n$

Then $g^{r-s} = e$. By def of order, $r - s = 0$, so $r = s$.

(b) If $g^r = g^s$, say $r \leq s$, then $g^{s-r} = e$, $s - r \geq 0$.

Since $o(g) = \infty$, $s - r = 0$, ie. $r = s$. $\square$

Def 2.18

Let G be a group, $g \in G$.

Define $\langle g \rangle = \{g^i : i \in \mathbb{Z}\} \subseteq G$. If $\langle g \rangle = G$ then g is said to generate G. If G is generated by some element, G is cyclic.

eg. $\mathbb{Z}$ under $+$ is cyclic, generated by 1

$2 = 1 + 1$, $3 = 1 + 1 + 1$, ...

$\mathbb{Z}_5^*$ (under $\times$) is cyclic

$\bar{2}$, $\bar{2}^2 = \bar{4}$, $\bar{2}^3 = \bar{3}$, $\bar{2}^4 = \bar{1}$ [not generated by $\bar{4}$: $\bar{4}^2 = \bar{1}$]

05/01/16 1202

L6

Sym(T) (as above) is not cyclic,

$\langle x_1 \rangle = \{x_1, e\} \neq G$

$\langle x_2 \rangle = \{x_2, e\} \neq G$

$\langle x_3 \rangle = \{x_3, e\} \neq G$

$\langle y_1 \rangle = \{e, y_1, y_2\} \neq G$

$\langle y_2 \rangle = \{e, y_1, y_2\} \neq G$

$\langle e \rangle = \{e\} \neq G$

Lemma 2.19

Let $|G| = n < \infty$.

Then G is cyclic $\Leftrightarrow$ G contains an element of order n.

Proof

($\Leftarrow$) Suppose $o(g) = n$. Then by 2.17

$\langle g \rangle = \{e, g, \ldots, g^{n-1}\}$ and $|\langle g \rangle| = n = |G|$

$\therefore \langle g \rangle = G$ and G is cyclic as it is generated by g.

($\Rightarrow$) Suppose $G = \langle g \rangle$.

$|\langle g \rangle| = |G| = n$

By 2.17, $o(g) = n$ □

Def 2.20

Let G be a cyclic group generated by g.

(i) If $o(g) = n$, then $G = \{e, g, \ldots, g^{n-1}\}$ and $G = \langle g, g^n = e \rangle$ is a cyclic group of order n, denoted C_n.

(ii) If $o(g) = \infty$, then $G = \{g^i : i \in \mathbb{Z}\} = \langle g | \rangle$ is the <u>infinite</u> cyclic group, denoted C_∞

$G = \{e, a, a^2\}, \quad a^3 = e$
$H = \{e, b, b^2\}, \quad b^3 = e$

We say G is <u>isomorphic</u> to H, written $\cong H$. We can usually regard isomorphic groups as the same, and in this sense there is only one cyclic group of order 3, for example.

$C_n \cong \mathbb{Z}_n$
$\{e, g, g^2, \ldots, g^{n-1}\} \cong \{\bar{0}, \bar{1}, \ldots, \overline{n-1}\}$

eg $n=3$: $\{e, g, g^2\} \cong \{\bar{0}, \bar{1}, \bar{2}\}$

$g \times g^2 = e \qquad \bar{1} + \bar{2} = 0$

$C_\infty \cong \mathbb{Z}$
[like taking $\log_g$]

<u>Subgroups</u>
<u>Def 2.21</u>
Let $H \subseteq G$, G is a group. Then H is a <u>subgroup</u> of G, written $H \leq G$, if
(i) $e \in H$
(ii) $g, h \in H \Rightarrow gh \in H$
(iii) $g \in H \Rightarrow g^{-1} \in H$

e.g. $G = C_6 = \{e, g, g^2, g^3, g^4, g^5\}$, $g^6 = e$

$H = \{e, g^2, g^4\} \leq G$

$K = \{e, g^4\} \not\leq G \quad$ since $g^4 \cdot g^4 = g^2 \notin K$.
(ii)&(iii) can be combined into $g, h \in H \Rightarrow g^{-1}h \in H$, and (i) can be replaced, by $H \neq \emptyset$.

08/02/16 1202

L7

Fermat's "little" Theorem

Let $\bar{a} \in \mathbb{Z}_p^*$

Then $\bar{a}^{p-1} = \bar{1}$ (so $o(\bar{a}) \mid p-1$)

eg. in $\mathbb{Z}_7^*$, $o(\bar{a}) \mid 6$

eg. $\bar{2}$, $\bar{2}^2 = \bar{4}$, $\bar{2}^3 = \bar{1}$ so $o(\bar{2}) = 3$

$\bar{3}$, $\bar{3}^2 = \bar{2}$, $\bar{3}^3 = \bar{6}$, $\bar{3}^4 = \bar{4}$, $\bar{3}^5 = \bar{5}$, $\bar{3}^6 = \bar{1}$, $o(\bar{3}) = 6$.

Proof

Let $\bar{a} \in \mathbb{Z}_p^*$

Consider the set $\{\bar{a}, \overline{2a}, \dots, \overline{(p-1)a}\}$

By 2.8 (a), this is $\mathbb{Z}_p^*$ again,

ie. $\{\bar{a}, \overline{2a}, \dots, \overline{(p-1)a}\} = \{\bar{1}, \bar{2}, \dots, \overline{(p-1)}\}$

[eg in $\mathbb{Z}_5^*$

$$\begin{array}{cccc} \bar{3}, & 2 \times \bar{3}, & 3 \times \bar{3}, & 4 \times \bar{3} \\ \| & \| & \| & \| \\ \bar{3}, & \bar{1}, & \bar{4}, & \bar{2} \end{array}$$
]

Multiplying all the elements together

$$\bar{a} \times \overline{2a} \times \overline{3a} \times \dots \times \overline{(p-1)a} = \bar{1} \times \bar{2} \times \bar{3} \times \dots \times \overline{(p-1)}$$

$$\overline{(p-1)!}\,\bar{a}^{p-1} = \overline{(p-1)!}$$

$$\text{so } \bar{a}^{p-1} = \bar{1}$$

Another proof will be given after proving Lagrange's Theorem on the size of subgroups.

e.g. $\bar{2}^{64}$ in $\mathbb{Z}_{31}^*$

$\bar{2}^{30} = \bar{1}$

$$\text{so } \bar{2}^{64} = \bar{2}^{2 \times 30 + 4}$$
$$= (\bar{2}^{30})^2 \times \bar{2}^4$$
$$= \bar{2}^4 = \overline{16}$$

Let G be a group, $H \leq G$.
Then H is a <u>subgroup</u> of G, written $H \leq G$, if
(i) $e \in H$ or $H \neq \emptyset$
(ii) $g, h \in H \Rightarrow gh \in H$
(iii) $g \in H \Rightarrow g^{-1} \in H$
$\Big\}$ or $g, h \in H \Rightarrow g^{-1}h \in H$

<u>Lemma 2.22</u>
Let G be a group, $H \leq G$.
Then H is a subgroup of G iff H forms a group under the same operations as G.

<u>Proof</u>
($\Leftarrow$) clear
($\Rightarrow$) Condition (ii) means that we have a closed binary operation on H.
By (i) and (iii), H has an identity element & every element of H has an inverse.
Associativity is automatic, since it holds in G.

eg. $3\mathbb{Z} \leq \mathbb{Z}$ under $+$
(i) $0 \in 3\mathbb{Z}$
(ii) let $x, y \in 3\mathbb{Z}$, say $x = 3a$, $y = 3b$ $(a, b \in \mathbb{Z})$
Then $x + y = 3(a+b) \in 3\mathbb{Z}$
(iii) $-x = 3x - a \in 3\mathbb{Z}$

$\mathbb{Q} \leq \mathbb{R}$ under $+$
$\mathbb{Q}^* \leq \mathbb{R}^*$ under $\times$

$Sym(T) = \langle x, y : y^3 = e, x^2 = e, xy = y^2 x \rangle$
$\{e, y\}$ is not a subgroup $yy = y^2 \notin \{e, y\}$
$\{e, y, y^2\}$ is a subgroup
$\{e, yx\}$ is a subgroup $(yx)^2 = yxyx = yy^2xx = ee = e$

08/02/16 1202

L7

<u>Recall:</u>

S_n = permutation group
= set of bijections $\{1, \dots, n\} \to \{1, \dots, n\}$
under composition

Any $\sigma \in S_n$ can be written as a product of transpositions say $\sigma = \tau_1 \dots \tau_m$. If m is even, σ is called <u>even</u>, if m is odd, σ is called <u>odd</u>.

<u>Theorem 3.23</u>

Let $A_n = \{\sigma \in S_n : \sigma \text{ even}\}$.
A_n is a subgroup of S_n, called the <u>alternating group</u>.
$|A_n| = \frac{1}{2}|S_n| = \frac{1}{2}n!$

<u>Proof</u>

(i) $e \in A_n$ (e is a product of 0 transpositions)
(ii) let $g, h \in A_n$, say $g = \tau_1 \dots \tau_{2m}$, $h = \sigma_1 \dots \sigma_{2p}$
(τ_i, σ_i transpositions).
Then $gh = \tau_1 \sigma_1 \dots \tau_{2m} \sigma_p \in A_n$ and $g^{-1} = \tau_{2m} \dots \tau_1 \in A_n$.
$|S_n| = n!$
Define $\phi : A_n \to S_n - A_n$ by $\phi(\sigma) = (12)\sigma$.
ϕ is well defined: σ is even, so $(12)\sigma$ is odd.
ϕ is injective. $\phi(\sigma) = \phi(\psi) \Rightarrow (12)\sigma = (12)\psi$
$\Rightarrow \sigma = \psi$.
ϕ is surjective, $\psi = \phi((12)\psi)$
$\therefore |A_n| = |S_n - A_n| = \frac{1}{2}|S_n| = \frac{1}{2}n!$

Lagrange's Theorem

One problem in group theory is:
given a group, find all its subgroups.
In general a difficult problem, but the next theorem gives an important necessary condition for H to be a subgroup of G.

Theorem 2.24 (Lagrange)
Let G be a finite group and H a subgroup of G. Then $|H|$ divides $|G|$.

e.g. suppose $|G| = 7$, then if $H \leq G$, $|H| \mid 7$
ie. $|H| = 1$ or 7, i.e $H = \{e\}$ or $H = G$.
ie. a group of order 7 has no non-trivial subgroups.

If $|G| = 6$, then any subgroups must be of order $1, 2, 3, 6$.

12/02/16 | 1202

L8

Thm 2.24 (Lagrange's Theorem)
Let G be a finite group, H a subgroup. Then $|H|$ divides $|G|$.

Example
$G = C_6 = \{e, x, x^2, x^3, x^4, x^5\}$ $(x^6 = e)$
$H = \{e, x^3\} \leq G$

Proof of Thm 2.24

Def. of cosets!

① For any $g \in G$, the coset $Hg = \{hg : H \in H\}$.
eg. $He = \{ee, x^3e\} = \{e, x^3\} = H$
$Hx = \{ex, x^3x\} = \{x, x^4\}$
$Hx^2 = \{ex^2, x^3x^2\} = \{x^2, x^5\}$
$Hx^3 = \{ex^3, x^3x^3\} = \{x^3, x^6\} = \{e, x^3\} = H$
$Hx^4 = \{ex^4, x^3x^4\} = \{x^4, x^7\} = \{x, x^4\}$
$Hx^5 = \{ex^5, x^3x^5\} = \{x^5, x^8\}$

② G is the union of all the cosets.
Since $g = eg \in Hg$, $\bigcup_{g \in G} Hg = G$

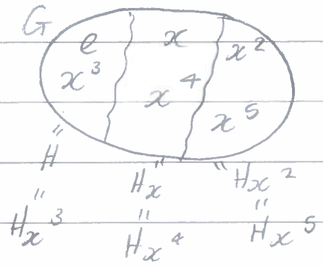
③ Two cosets are either equal or disjoint.
So suppose $Hg \cap Hg' \neq \emptyset$.
Say $x \in Hg \cap Hg'$.
Then $x = h_1 g = h_2 g'$ for some $h_1, h_2 \in H$
Hence $g = h_1^{-1} h_2 g'$.
For any $h \in H$, $hg = \underbrace{h h_1^{-1} h_2}_{\in H \text{ since } H \text{ subgroup.}} g \in Hg'$

$\therefore Hg \leq Hg'$.
Similarly $Hg' \leq Hg$ $\therefore Hg = Hg'$
ie. $Hg \cap Hg' = \emptyset$ or $Hg = Hg'$

④ G is the disjoint union of some of the cosets
ie. $\exists g_1, \ldots, g_r$ st.
$G = Hg_1 \cup \ldots \cup Hg_r$ and $Hg_i \cap Hg_j \neq \emptyset$ $(i \neq j)$
(eg. $G = He \cup Hx \cup Hx^2$)
We know $G = \bigcup_{g \in G} Hg$ and each $Hg \cap Hg' = \varphi$ or $Hg = Hg'$.

So leaving out repetitions, G is disjoint union of some of the cosets.

$G = He \cup Hx \cup Hx^2$



⑤ All cosets are the same size $|H|$.

Define $\varphi : H \to Hg$ by $\phi(h) = hg$.

ϕ is surjective by definition of Hg.

Suppose $\phi(h) = \varphi(h')$. Then $hg = h'g$, so since G is a group $h = h'$, ie. φ is injective.

$\therefore$ φ bijective and $|H| = |Hg|$

eg. $\varphi : H \to Hx^2$

$$e \to ex^2 = x^2$$
$$x^3 \to x^3x^2 = x^5$$

$|Hg| = 2$

⑥ Result: $|G| = |H| r$, so $|H|$ divides $|G|$

$$G = Hg_1 \cup \ldots \cup Hg_r \quad \text{(disjoint)}$$
$$|G| = |Hg_1| + \ldots + |Hg_r|$$
$$= |H| + \ldots + |H|$$
$$= r|H|$$

Corollary

Let G be a finite group, $g \in G$. Then $o(g) \mid |G|$.

Proof

Let $H = \langle g \rangle = \{g^i : i \in \mathbb{Z}\}$. Then $H \leq G$, and $|H| = o(g)$. (2.17)

12/02/16 | 1202

L8

By theorem, $o(g) \mid |G|$.
e.g. if G has 6 elements, the only possible orders of elements are 1, 2, 3, 6.

Ex
Find the order of each element in C_6 and in S_3.
What does this tell you about the two groups?

C_6 : e, x, x^2, x^3, x^4, x^5
order: 1 6 3 2 3 6

S_3 : e, $(1,2)$, $(1,3)$, $(2,3)$, $(1,2,3)$, $(1,3,2)$
order: 1 2 2 2 3 3

The fact that the orders are different tells you that the two groups are "genuinely" different.
ie. non isomorphic.

Corollary 2.26
Let G be a group of order p, prime. Then G is cyclic ($G \cong C_p$)

Proof
Pick any elements $g \in G$ different from e.
Then $o(g) \mid p$ and $o(g) \neq 1$ $\therefore o(g) = p$.
B 2.19, G is cyclic, generated by g.

So groups of prime order are rather simple: there is just one group C_p of order p for each prime, p. Groups of composite order are much more complicated.

Groups of small order:

2	C_2	6	C_6 or S_3 ← non-abelian
3	C_3	7	C_7
4	C_4 or $C_2 \times C_2$		
5	C_5		

We can apply these results to $\mathbb{Z}_p^*$.

Thm 2.27 (Fermat's Little Theorem)
Let $\bar{a} \in \mathbb{Z}_p^*$. Then $\bar{a}^{p-1} = \bar{1}$

Proof
$|\mathbb{Z}_p^*| = p-1$
$\therefore$ By 2.25, $o(\bar{a}) \mid p-1$
Say $p-1 = o(\bar{a})r$.
Then $\bar{a}^{p-1} = \bar{a}^{o(\bar{a})r} = (\bar{a}^{o(\bar{a})})^r = \bar{1}^r = \bar{1}$

eg. $2^{75} \pmod{37}$
$2^{36} \equiv 1$
$2^{72} \equiv 1$
$2^{75} \equiv 2^3 \equiv 8$

12/02/16 1202

L8

Chapter 3 - Determinants

Definitions and the 2x2 case

Def 3.1

Let A be an $n \times n$ matrix. The the determinant of A is

$$\det(A) = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1,\sigma(1)} \ldots\ldots a_{n,\sigma(n)}$$

Here S_n is the permutation group, ie. the group of all bijections $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$.
$\operatorname{sgn}(\sigma)$ is the sign of σ, i.e. $\operatorname{sign}(\sigma) = \begin{cases} 1, & \sigma \text{ even} \\ -1, & \sigma \text{ odd.} \end{cases}$

Prop 3.2 (2x2 case)

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

Then (i) $\det A = ad - bc$

(ii) A is invertable $\Leftrightarrow \det A \neq 0$

In this case $A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

(iii) Let L_A be the linear map $\mathbb{R}^2 \to \mathbb{R}^2$ given by $L_A(\underset{\sim}{x}) = A\underset{\sim}{x}$.
Then if S is a shape in $\mathbb{R}^2$,
then $\text{area}(L_A(S)) = \text{area}(S) \times \det A$
i.e. L_A multiplies areas by $\det A$.

(iv) If B is another 2×2 matrix, $\det(AB) = \det A \det B$.

Proof

(i) $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1,\sigma(1)} a_{2,\sigma(2)}$

$S_2 = \{e, (1,2)\}$

$\sigma = e \quad : \quad (+1)\, a_{1,1} \quad a_{2,2}$

$\sigma = (1,2) : \quad (-1)\, a_{1,2} \quad a_{2,1}$

so $\det A = a_{11} a_{22} - a_{12} a_{21} = ad - bc$

e.g. $\det \begin{pmatrix} 2 & 1 \\ 4 & 3 \end{pmatrix} = 2 \times 3 - 1 \times 4 = 2$

$\det \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix} = 2 \times 3 - 6 \times 1 = 0$

(ii) Try to find inverse of $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

Naively: want to solve

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

so $\begin{pmatrix} ax + bz & ay + bt \\ cx + dz & cy + dt \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

so $\begin{cases} ax + bz = 1 \\ ay + bt = 0 \\ cx + dz = 0 \\ cy + dt = 0 \end{cases}$

$adx + bdz = d$

$bcx + bdz = 0$

so $(ad - bc) = d$

If $ad - bc \neq 0$, $x = \dfrac{d}{ad - bc}$

Similarly, if $ad - bc \neq 0$, get $\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \dfrac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

and it is very easy to check this is A^{-1}.

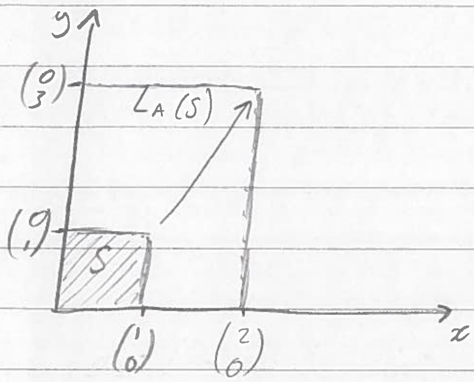
If $ad - bc = 0$ and we assume A has an inverse, we get $d = 0$, and similarly $b = c = a = 0$, i.e. $A = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$ #

(iii) e.g. $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$, $\det A = 6$

$$L_A \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix}$$

12/02/16 | 1202

L8



Area $L_A(S) = 6 = 6 \times 1$

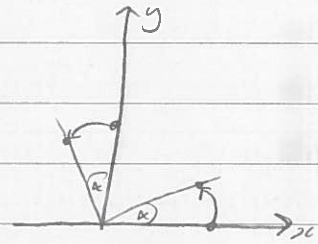
$= \det A \times \text{Area}(S)$

$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$

$L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$

$L_A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$

$L_A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$



L_A rotates by angle α anticlockwise about origin.

Area $(L_A(S))$ = area (S)

$= 1 \times$ area (S)

$\det A = \cos^2\alpha + \sin^2\alpha = 1$ [det = 1 ⇒ map maintains area size]

e.g. $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $\det A = 0$

$L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix}$

22/02/16 1202

L9

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

2x2 $\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$

Prop 3.2 (2x2 case) $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$

(i) $\det A = ad - bc$

(ii) A invertable $\Leftrightarrow \det A \neq 0$

In this case $A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

(iii) $L_A : \mathbb{R}^2 \to \mathbb{R}^2$, $L_A(\underline{x}) = A\underline{x}$.

? Then if $S \leq \mathbb{R}^2$, area $(L_A(S)) = |\det A| \times$ area (S)

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

Proof-(iv)

This can be checked by direct calculation. Alternatively it can be seen from (iii)

L_A multiplies areas by $|\det A|$

L_B " " " $|\det B|$

$\therefore L_A L_B$ " " " $|\det A| |\det B|$

$\shortparallel$ $\shortparallel$

L_{AB} " " " $|\det AB|$

3x3 case (Prop 3.3)

$$\det A = \sum_{\sigma \in S_3} (\operatorname{sgn} \sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)}$$

where $S_3 = \{e, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$

[ex: $\sigma = (1\ 2\ 3) \Rightarrow \sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 1$.]

So
$$\begin{aligned}\det A = &(\operatorname{sgn}(id)) a_{1,id(1)} a_{2,id(2)} a_{3,id(3)} \\
&+ (\operatorname{sgn}(1\,2)) a_{1,(1\,2)(1)} a_{2,(1\,2)(2)} a_{3,(1\,2)(3)} \\
&+ (\operatorname{sgn}(1\,3)) a_{1,(1\,3)(1)} a_{2,(1\,3)(2)} a_{3,(1\,3)(3)} \\
&+ (\operatorname{sgn}(2\,3)) a_{1,(2\,3)(1)} a_{2,(2\,3)(2)} a_{3,(2\,3)(3)} \\
&+ (\operatorname{sgn}(1\,2\,3)) a_{1,(1\,2\,3)(1)} a_{2,(1\,2\,3)(2)} a_{3,(1\,2\,3)(3)} \\
&+ (\operatorname{sgn}(1\,3\,2)) a_{1,(1\,3\,2)(1)} a_{2,(1\,3\,2)(2)} a_{3,(1\,3\,2)(3)}\end{aligned}$$

$$\text{So } \det A = (+1)a_{1,1}a_{2,2}a_{3,3}$$
$$+(-1)a_{1,2}a_{2,1}a_{3,3}$$
$$+(-1)a_{1,3}a_{2,2}a_{3,1}$$
$$+(-1)a_{1,1}a_{2,3}a_{3,3}$$
$$+(+1)a_{1,2}a_{2,3}a_{3,1}$$
$$+(+1)a_{1,3}a_{2,1}a_{3,2}$$

<u>Pattern:</u>

$$\begin{array}{cccccc} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} & a_{33} \end{array}$$

add the $\searrow$ diagonals and subtract the $\swarrow$ diagonals.

$$\text{eg } \det\begin{pmatrix} 2 & 1 & 7 \\ 1 & 2 & -1 \\ 3 & 4 & 5 \end{pmatrix} = (2\times 2\times 5) + (1\times(-1)\times 3) + (7\times 1\times 4) - (7\times 2\times 3) - (1\times 1\times 5) - (2\times(-1)\times 4)$$

NB: This simple pattern does not work for 4x4 determinants where there are $4! = 24$ terms!

<u>Ex</u>

$$\text{Find } \det\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix} = 8 + 1 + 1 - 2 - 2 - 2 = 4$$

<u>Properties of determinants</u>

The first result is about transposes.

Recall A^T has (i,j)-entry a_{ji}

$$\text{e.g. } \begin{pmatrix} 1 & 2 & 7 \\ 1 & 3 & 1 \\ -1 & -2 & 0 \end{pmatrix}^T = \begin{pmatrix} 1 & 1 & -1 \\ 2 & 3 & -2 \\ 7 & 1 & 0 \end{pmatrix}$$

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}, \quad \det A = ad - bc, \quad \det A^T = ad - bc$$

Prop 3-4

Let A be an $n \times n$ matrix.
Then $\det(A^T) = \det A$.

Proof

Write $B = A^T$.

$$\det(A^T) = \det(B) = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, b_{1,\sigma(1)} \dots b_{n,\sigma(n)} = \sum_{\sigma \in S_n} (\operatorname{sgn}\sigma)\, a_{\sigma(1),1} \dots a_{\sigma(n),n}$$

Let $\sigma = \psi^{-1}$. Then as σ ranges over S_n, so does ψ.

$$\det(A^T) = \sum_{\psi \in S_n} (\operatorname{sgn}\psi^{-1})\, a_{\psi^{-1}(1),1} \dots a_{\psi^{-1}(n),n} = \sum_{\psi \in S_n} (\operatorname{sgn}\psi)\, a_{\psi^{-1}(1),1} \dots a_{\psi^{-1}(n),n}$$

$a_{\psi^{-1}(1),1} \dots a_{\psi^{-1}(n),n} = a_{1,\psi(1)} \dots a_{n,\psi(n)}$

Suppose $\psi(1) = i$, then $\psi^{-1}(i) = 1$
The term $a_{\psi^{-1}(i),i} = a_{1,i} = a_{1,\psi(1)}$

or

write $a_{\psi^{-1}(1),1} \dots a_{\psi^{-1}(n),n} = \prod_{i=1}^{n} a_{\psi^{-1}(i),i}$

let $j = \psi^{-1}(i)$. As i varies from 1 to n, so does j, so
$\prod_{i=1}^{n} a_{\psi^{-1}(i),i} = \prod_{j=1}^{n} a_{j,\psi(j)}$

$$\therefore \det(A^T) = \sum_{\psi \in S_n} (\operatorname{sgn}\psi)\, a_{1,\psi(1)} \dots a_{n,\psi(n)} = \det A$$

This result means any results about rows immediately translate into results about columns.

Prop 3.5

Let A be a <u>lower triangular matrix</u>, i.e. $a_{ij} = 0 \ \forall j > i$
Then $\det A = a_{11} a_{22} \dots a_{nn}$

e.g. $\det\begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11} a_{22} a_{33}$

Proof

$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}$

One term in this sum is when $\sigma = id$, giving

$\operatorname{sgn}(id)\, a_{1,id(1)} \dots a_{n,id(n)} = a_{11} a_{22} \dots a_{nn}$

We claim all other terms are zero.

So suppose $a_{1,\sigma(1)} \dots a_{n,\sigma(n)} \neq 0$

Then $a_{1,\sigma(1)} \neq 0, \dots, a_{n,\sigma(n)} \neq 0$

Since $a_{1,\sigma(1)} \neq 0$, $\sigma(1) \leq 1$, i.e. $\sigma(1) = 1$

Since $a_{2,\sigma(2)} \neq 0$, $\sigma(2) \leq 2$, i.e. $\sigma(2) = 2$

Continuing $\sigma(3) = 3, \dots$

i.e. $\sigma = id$.

26/02/16 1202

L10

$$\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$$

$$\det A^T = \det A$$

$$\det \begin{pmatrix} a_1 & & 0 \\ & \ddots & \\ 0 & & a_n \end{pmatrix} = a_1 \cdots a_n$$

Elementary matrices

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{p(1,2)} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = P(1,2)$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{e(1,2;\lambda)} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = E(1,2;\lambda)$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \xrightarrow{d(2;\lambda)} \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} = D(2;\lambda)$$

eg. $\underset{A:}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \xrightarrow{e(1,2;\lambda)} \begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix}$

$$E(1,2;\lambda)A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix}$$

Thm 3.6

(a) Exchanging 2 rows of a matrix multiplies the determinant by -1.

(b) Multiplying a row by λ multiplies the determinant by λ. (correct this by multiplying det A by $\frac{1}{\lambda}$)

(c) Adding a multiple of one row to another doesn't change the determinant.

e.g. $\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{p(1,2)} \begin{pmatrix} c & d \\ a & b \end{pmatrix}$

$ad - bc \qquad\qquad bc - ad$

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{d(2;\lambda)} \begin{pmatrix} a & b \\ \lambda c & \lambda d \end{pmatrix}$

$ad - bc \qquad\qquad a\lambda d - b\lambda c$

$= \lambda(ad - bc)$ so $\det\begin{pmatrix} a & b \\ \lambda c & \lambda d \end{pmatrix} = \lambda \det\begin{pmatrix} a & b \\ c & d \end{pmatrix}$

$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow{e(1,2;\lambda)} \begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix}$

$ad - bc \qquad\qquad (a+\lambda c)d - c(b+\lambda d)$

$= ad - bc + \lambda\cancel{(cd - cd)}$

$= ad - bc$

Proof

(a) Consider $p(1,2)$, say $A \xrightarrow{p(1,2)} B$

$b_{1j} = a_{2j} \quad (j = 1, \dots, n)$

$b_{2j} = a_{1j} \quad (j = 1, \dots, n)$

$b_{rj} = a_{rj} \quad (j = 1, \dots, n) \qquad r \geq 3$

so $\det B = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) b_{1,\sigma(1)} b_{2,\sigma(2)} \dots b_{n,\sigma(n)}$

$= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma) a_{2,\sigma(1)} a_{1,\sigma(2)} a_{3,\sigma(3)} \dots a_{n,\sigma(n)}$

let $\tau = (1,2)$. As σ ranges over S_n, so does $\sigma\tau$

$= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma\tau) a_{1,\sigma\tau(2)} a_{2,\sigma\tau(1)} a_{3,\sigma\tau(3)} \dots a_{n,\sigma\tau(n)}$

$= \sum_{\sigma \in S_n} -(\operatorname{sgn} \sigma) a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)} \dots a_{n,\sigma(n)}$.

$= -\det A$.

(b) Similar but easy.

(c) First note that as a consequence of (a), if a matrix

26/02/16 1202

L10

has 2 rows the same it must have determinant = 0.

eg. $A = \begin{pmatrix} a_1 \\ a_1 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} \xrightarrow{p(1,2)} \begin{pmatrix} a_1 \\ a_1 \\ a_3 \\ \vdots \\ a_n \end{pmatrix} = A$

$$\det A = -\det A \quad \therefore \det A = 0$$

Suppose $A \xrightarrow{e(1,2;\lambda)} B$

$$\begin{cases} b_{1j} = a_{1j} + \lambda a_{2j} \\ b_{rj} = a_{rj} \qquad (r \geq 2) \end{cases}$$

$$\det B = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, b_{1,\sigma(1)} \dots b_{n,\sigma(n)}$$

$$= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)(a_{1,\sigma(1)} + \lambda a_{2,\sigma(2)})\, a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$$

$$= \underbrace{\sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1,\sigma(1)} \dots a_{n,\sigma(n)}}_{\det A} + \lambda \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{2,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$$

Let C be the matrix obtained from A by replacing the first row by the second row

$A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \qquad C = \begin{pmatrix} a_2 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{pmatrix}$

$c_{1j} = a_{2j}$, $c_{rj} = a_{rj}$ $(r \geq 2)$

$$\det C = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, c_{1,\sigma(1)} c_{2,\sigma(2)} \dots c_{n,\sigma(n)}$$

$$= \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{2,\sigma(1)} a_{2,\sigma(2)} \dots a_{n,\sigma(n)}$$

so $\det B = \det A + \underbrace{\lambda \det C}_{=0}$

$\therefore \det B = \det A$

$$\det \begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix} = ad - bc + \lambda(cd - cd).$$

This gives us an effective way of calculating determinants using row reductions to triangular form

e.g. $\det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 2 & 2 & 0 & 2 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{e(2,1;-2)}{=} \det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & -2 & -2 & 2 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

$\overset{d(2,-\frac{1}{2})}{=} -2\det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 3 & -1 & -1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \overset{e(3,2;-3)}{=} -2\det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & -4 & 2 \\ 0 & 0 & 1 & 1 \end{pmatrix}$

$\overset{p(3,4)}{=} 2\det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & -4 & 2 \end{pmatrix} = 2\det\begin{pmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 6 \end{pmatrix}$

$= 2\times(1\times1\times1\times6) = 12$

e.g. $\det\begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^2 & b^2 & c^2 \end{pmatrix} = \det\begin{pmatrix} 1 & 0 & 0 \\ a & b-a & c-a \\ a^2 & b^2-a^2 & c^2-a^2 \end{pmatrix}$

$=(b-a)(c-a)\det\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 1 \\ a^2 & b+a & c+a \end{pmatrix}$

$=(b-a)(c-a)\det\begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ a^2 & b+a & c-b \end{pmatrix}$

$=(b-a)(c-a)(c-b)$

Hence $\det A \neq 0 \iff a, b, c$ distinct.

Example of Vandermonde determinant.

26/02/16 1202

<10

Exercise

Find det of

(i) $\begin{pmatrix} 0 & 2 & 3 & 1 \\ 1 & 0 & 1 & -1 \\ 2 & 2 & 0 & 1 \\ 3 & 4 & 2 & -2 \end{pmatrix} = A$

$\downarrow$

$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 2 & 2 & 0 & 1 \\ 3 & 4 & 2 & -2 \\ 0 & 2 & 3 & 1 \end{pmatrix}$

$\downarrow$

$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 0 & 3 \\ 0 & 4 & -1 & 1 \\ 0 & 2 & 3 & 1 \end{pmatrix}$

$\downarrow$

$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & -1 & -5 \\ 0 & 0 & 3 & -2 \end{pmatrix}$

$\downarrow$

$\begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 0 & 3 \\ 0 & 0 & -1 & -5 \\ 0 & 0 & 0 & 13 \end{pmatrix}$

so $\det A = 1 \times 2 \times (-1) \times 13$

$= -26$

$\times$

-38 !!

(ii) $\det \begin{pmatrix} 1 & 1 & 1 \\ a & b & c \\ a^3 & b^3 & c^3 \end{pmatrix}$

$= \det \begin{pmatrix} 1 & 0 & 0 \\ a & b-a & c-a \\ a^3 & b^3-a^3 & c^3-a^3 \end{pmatrix}$

$= (b-a)(c-a) \det \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 1 \\ a^3 & b^2+ab+a^2 & c^2+ac+a^2 \end{pmatrix}$

$= (b-a)(c-a) \det \begin{pmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ a^3 & (b^2+ab+a^2) & (c^2+ac+a^2-b^2-ab-a^2) \end{pmatrix}$

$= (b-a)(c-a)(c^2-b^2+ac-ab)$

$= (b-a)(c-a)(c-b)(a+b+c)$

Two main results

We saw that for 2×2 matrices A , is le
$\iff \det A \neq 0$ and $\det(AB)$ $\det A$ B
Now we want to show the for $(n\times n)$ matrices
using elementary matrices.

Prop 3.7

Let A be a square matrix, E an elementary matrix (both $n\times n$).
Then $\det(EA) = \det E \det A$ and det 0

Proof

First let $E = P(i,j)$. Then EA (i,j) is he es
of applying $p(i,j)$ to A (Fac from he sou)
By 3.6 (a), $\det(P(i,j)A) =$ A
Applied with $A = I$ gives $\det(P(i,j)) = -\det I = -1$
$\therefore \det(P(i,j)A) = \det(P(i,j)) \det$
Similarly for $E(i,j;\lambda)$ and $D(i,$ $)$.
So $\det P(i,j) = -1$, $\det E(i,j;\lambda) =$, (i,λ) λ.

Corrollary 3.7

Let A be a square matr , , ..., elementary matrices (all $n\times n$). Then $\det(E_n \dots E_2 E_1 A) = \det(E_n) \dots \det(E_1) \det(A)$

Thm 3.8

Let A be $n\times n$.
A is invertable $\iff \det A \neq 0$.

Proof

By Facts 1 and 2, we can f elem y
$E_1, \dots, E_n$ s.t.
$E_n \dots E_1 A = T$ (in RRE form)
By corollary 3.7, $\det T = \det E_n \dots \det E_1 \det A$.

26/02/16 1202

L10

Each $\det E_i \neq 0$, so $\det T \neq 0 \Leftrightarrow \det A \neq 0$.

Suppose A is invertable, by fact 5 $T = I_n$.

$\therefore \det T \neq 0$ so $\det A \neq 0$.

Suppose A is not invertable, by fact 5 T has a zero row, so $\det T = 0$ $\therefore$ $\det A = 0$.

29/02/16 1202

L11

Thm 3.10

Let A, B be $n \times n$ matrices. Then $\det(AB) = \det A \det B$.

Proof

Suppose the elementary row operations $e_1, \ldots, e_n$ reduce A to RRE form, so $E_n \ldots E_1 A = T$ (RRE).

Each E_i has an inverse - another elementary matrix, say F_i, so

$A = F_1 \ldots F_n T$ ①

Then $AB = F_1 \ldots F_n (TB)$ ②

Case 1: A invertable.

Then $T = I$, so $A = F_1 \ldots F_n$,

$AB = F_1 \ldots F_n B$.

By 3.8, $\det(AB) = \det F_1 \det F_2 \ldots \det F_n \det B$

$= \det A \det B$.

Case 2: A not invertable.

T has a zero row, $T = \begin{pmatrix} & & \\ 0 & — & 0 \end{pmatrix}$

and hence $TB = \begin{pmatrix} & & \\ 0 & — & 0 \end{pmatrix}$,

so $\det T = 0$, $\det TB = 0$.

By ①, ② and 3.8,

$\det A = 0$, $\det(AB) = 0$

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

Expansion by minors

Def 3.11

The (i, j)-minor M_{ij} of a $n \times n$ matrix A is the determinant of the matrix obtained from A by deleting the i^{th} row and j^{th} column.

The (i,j)-cofactor $C_{ij} = (-1)^{i+j} M_{ij}$

eg $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{23} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$M_{23} = \det \begin{pmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{pmatrix} = a_{11} a_{32} - a_{12} a_{31}$$

$$C_{23} = (-1)^{2+3} M_{23} = -M_{23}$$

We can form a matrix of minors, M_{ij}, and of cofactors, C_{ij}, and the second matrix is obtained from the first by multiplying each entry by $+1$ or -1 in the following pattern: $\begin{pmatrix} + & - & + & \cdots \\ - & + & - & \\ + & - & + & \vdots \\ - & & & \\ \vdots & & & \end{pmatrix}$

Ex

(i) For $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ find the matrix of minors and the matrix of cofactors.

$$M = \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \quad C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$$

(ii) Do the same, with $A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 3 & 1 \\ -1 & 2 & -2 \end{pmatrix}$

$$M = \begin{pmatrix} -8 & 1 & 3 \\ -10 & 1 & 4 \\ -7 & 1 & 3 \end{pmatrix}, \quad C = \begin{pmatrix} -8 & -1 & 3 \\ 10 & 1 & -4 \\ -7 & -1 & 3 \end{pmatrix}$$

29/02/16 | 1202

L11

Prop 3.12

Let A be $n \times n$. For any i $\det A = \sum_{j=1}^{n} a_{ij} C_{ij}$
(expansion along the i^{th} row)
and $\det A = \sum_{j=1}^{n} a_{ji} C_{ji}$ (expanding down the i^{th} collumn).

eg. $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$

$i=1$ $\det A = a_{11} C_{11} + a_{12} C_{12}$
$= a_{11} a_{22} + a_{12}(-a_{21})$
$= a_{11} a_{22} - a_{12} a_{21}$

$i=2$ $\det A = a_{21} C_{21} + a_{22} C_{22}$
$= a_{21}(-a_{12}) + a_{22} a_{11}$
$= a_{11} a_{22} - a_{12} a_{21}$

3×3 case

$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$

$$\det A = a_{11} C_{11} + a_{12} C_{12} + a_{13} C_{13}$$
$$= a_{11} M_{11} - a_{12} M_{12} + a_{13} M_{13}$$
$$= a_{11} \det\begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det\begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det\begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$
$$= a_{11}(a_{22}a_{33} - a_{23}a_{32}) - a_{12}(a_{21}a_{33} - a_{23}a_{31}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})$$
$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{23}a_{31} - a_{13}a_{22}a_{31}$$

Proof

Online.

This gives us an improved way of finding determinants (combined with row operations).

e.g. $\det \begin{pmatrix} 1 & 0 & 3 & 4 \\ 0 & 0 & 2 & 0 \\ 2 & 1 & 4 & -5 \\ 11 & 0 & 2 & 1 \end{pmatrix}$

$$= -0 + 0 - 2 \det \begin{pmatrix} 1 & 0 & 4 \\ 2 & 1 & -5 \\ 11 & 0 & 1 \end{pmatrix} + 0$$ (expanding along 2nd row)

$$= -2\left[0 + 1 \det \begin{pmatrix} 1 & 4 \\ 11 & 1 \end{pmatrix} + 0 \right]$$ (expanding down 2nd column)

$$= -2 \det \begin{pmatrix} 1 & 4 \\ 11 & 1 \end{pmatrix} = -2 \times -43 = 86$$

Adjugate and Inverse

Def 3.13

Let A be $n \times n$. Then the adjugate of A, $\mathrm{adj}(A)$, is the transpose of the matrix of cofactors.

$$(\mathrm{adj}\, A)_{ij} = C_{ji}$$

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

$$M = \begin{pmatrix} d & c \\ b & a \end{pmatrix}, \quad C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}, \quad \mathrm{adj}\, A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Recall $A^{-1} = \dfrac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

$$= \frac{1}{\det A} \mathrm{adj}\, A.$$

04/03/16 1202

L12

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

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

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

$$\operatorname{adj} A = C^T = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

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

$$= \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix}$$

$$= (ad-bc) I$$

$$= \det A . I_2$$

so $A^{-1} = \frac{1}{\det A} \operatorname{adj} A$

__Thm 3.14__

Let A be $n \times n$. then $A \operatorname{adj} A = (\det A) I_n = \operatorname{adj} A . A$

In particular if $\det A \neq 0$,

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A.$$

__Proof__

The (i,i)-entry of $A . \operatorname{adj} A$ is

$$\sum_{k=1}^{n} a_{ik} (\operatorname{adj} A)_{ki}$$

$$= \sum_{k=1}^{n} a_{ik} C_{ik} = \det A$$ (expansion along the ith row

Now consider the $(1,2)$-entry of $A \operatorname{adj} A$.

This is $\sum_{k=1}^{n} a_{1k} (\operatorname{adj} A)_{k2}$

$$= \sum_{k=1}^{n} a_{1k} C_{2k}$$

This is the determinant of

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{31} & a_{32} & \cdot & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & & & n \end{pmatrix} = D$$

expanded along the d row.

D differs from A only in row 2, so the $(2,k)$-cofactors of A and D are the same.
Since D has 2 row the same, $\det D$ 0.
i.e. $(1,2)$-entry of Adj-A 0, similarly the (i,j)-entry is 0 $\forall\, i \neq j$.

$\therefore A \operatorname{adj} A$

$$= \begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & & \vdots \\ \vdots & & \ddots & \\ 0 & & 0 & \det A \end{pmatrix}$$

$= \det A . I_n$

Similarly $\operatorname{adj} A . A = \det A \, I_n$

eg. $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix}$, $M = \begin{pmatrix} 3 & 3 & -1 \\ 5 & 1 & -1 \\ -4 & -8 & -4 \end{pmatrix}$

$C = \begin{pmatrix} 3 & -3 & -1 \\ -5 & 1 & 1 \\ -4 & 8 & -4 \end{pmatrix}$, $\operatorname{adj} A = \begin{pmatrix} 3 & -5 & -4 \\ -3 & 1 & 8 \\ -1 & 1 & -4 \end{pmatrix}$

$\det A = 1 \times 3 + 2 \times -3 + 3 \times -3 = -12 \neq 0$ ← found by multiplying circled row & column

so A invertable, $A^{-1} = \frac{-1}{12} \begin{pmatrix} 3 & -5 & -4 \\ -3 & 1 & 8 \\ -1 & 1 & -4 \end{pmatrix}$

04/03/16 1202

L12

Ex

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

find A^{-1} by this method.

(ii) $A = \begin{pmatrix} \alpha & 1 & 2 \\ 0 & \beta & 1 \\ 1 & \gamma & 2 \end{pmatrix}$ For which α, β, γ is A invertable? Find a formula for A^{-1} in this case.

(i) $M = \begin{pmatrix} -1 & 5 & 3 \\ 1 & -1 & -1 \\ 0 & -2 & -2 \end{pmatrix}$ $\det A = 0 \times -1 + 1 \times -5 + 1 \times 3 = -2$

$C = \begin{pmatrix} -1 & -5 & 3 \\ -1 & -1 & 1 \\ 0 & 2 & -2 \end{pmatrix}$, $\operatorname{adj} A = \begin{pmatrix} -1 & -1 & 0 \\ -5 & -1 & 2 \\ 3 & 1 & -2 \end{pmatrix}$

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

(ii) $\det A = 2\alpha\beta + 1 - 2\beta - \alpha\gamma$

$= 2\alpha\beta - 2\beta - \alpha\gamma + 1 \neq 0$ if A^{-1} exists

$M = \begin{pmatrix} 2\beta-\gamma & -1 & -\beta \\ 2-2\gamma & 2\alpha-2 & \alpha\gamma-1 \\ 1-2\beta & \alpha & \alpha\beta \end{pmatrix}$, $C = \begin{pmatrix} 2\beta-\gamma & 1 & -\beta \\ 2\gamma-2 & 2\alpha-2 & 1-\alpha\gamma \\ 1-2\beta & -\alpha & \alpha\beta \end{pmatrix}$

$\operatorname{adj} A = \begin{pmatrix} 2\beta-\gamma & 2\gamma-2 & 1-2\beta \\ 1 & 2\alpha-2 & -\alpha \\ -\beta & 1-\alpha\gamma & \alpha\beta \end{pmatrix}$

Note: $|A| = \det A$

so $A^{-1} = \frac{1}{2\alpha\beta-2\beta-\alpha\gamma+1}\begin{pmatrix} 2\beta-\gamma & 2\gamma-2 & 1-2\beta \\ 1 & 2\alpha-2 & -\alpha \\ -\beta & 1-\alpha\gamma & \alpha\beta \end{pmatrix}$

Chapter 4 - Diagonalisation

Def^n and basic criterion

Recall that an $n \times n$ matrix D is diagonal if $d_{ij} = 0 \ \forall i \neq j$.

e.g. $\begin{pmatrix} d_{11} & 0 \\ 0 & d_{22} \end{pmatrix}$, $\begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{pmatrix}$ etc..

We write $D = diag(d_1, \dots, d_n)$ for

$$\begin{pmatrix} d_{11} & & 0 \\ & \ddots & \\ 0 & & d_{nn} \end{pmatrix}.$$

Diagonal matrices are in a very simple form. However, most matrices are not diagonal. Most matrices are closely related to a diagonal matrix.

Def^n 4.1

A matrix A is diagonalisable if there exists an invertable matrix P such that $P^{-1}AP$ is diagonal.

Suppose there is such a P. How can it be found?

$P^{-1}AP = D$ (diag)

so $AP = PD$

Write P in columns, $P = (\underline{v}_1, \dots, \underline{v}_n)$, $D = diag(d_1, \dots, d_n)$

so $A(\underline{v}_1, \dots, \underline{v}_n) = (\underline{v}_1 \ \dots \ \underline{v}_n)\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix}$ [note: $\begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} = \begin{pmatrix} d_1 p & d_2 q \\ d_1 r & d_2 s \end{pmatrix}$]

$\Rightarrow (A\underline{v}_1 \ \dots \ A\underline{v}_n) = (d_1\underline{v}_1 \ \dots \ d_n\underline{v}_n)$

i.e. to find columns of P, we need to solve

$A\underline{x} = d\,\underline{x}$

04/03/12 1202

L12

Prop^n 4.2

Let $\underline{v}_1, \dots, \underline{v}_n \in \mathbb{R}^n$
and let P be the $n \times n$ matrix whose columns are
? $\underline{v}_1, \dots, \underline{v}_n$. Then T.F.A.E

(i) $\{\underline{v}_1, \dots, \underline{v}_n\}$ is LI

(ii) $\{\underline{v}_1, \dots, \underline{v}_n\}$ is a basis for $\mathbb{R}^n$

(iii) P is invertable.

Def^n 4.3

Let A be an $n \times n$ matrix over $\mathbb{R}$. Then λ is called an <u>eigenvalue</u> of A if there exists a non-zero $\underline{v} \in \mathbb{R}^n$ s.t.

$$A\underline{v} = \lambda \underline{v}.$$

$\underline{v}$ is then called an eigenvector of A (associated to λ).

Prop^n 4.4 (Basic criterion for diagonalisability)

Let A be an $n \times n$ matrix over $\mathbb{R}$. Then A is diagonalisable if and only if there exists a basis for $\mathbb{R}^n$ consisting of eigenvectors (equivalently if there is a set of n LI eigenvectors).

Proof

Suppose A is diagonalisable, say $P^{-1}AP = D$
Then $AP = PD$, so the columns of P are eigenvectors.
Since P is invertable, by 4.2, the columns of P form a basis for $\mathbb{R}^n$.
Conversely, suppose $\underline{v}_1, \dots, \underline{v}_n$ is a basis for $\mathbb{R}^n$ consisting of eigenvectors $\underline{v}_1, \dots, \underline{v}_n$. Let $P = (\underline{v}_1, \dots, \underline{v}_n)$.
By 4.2, P is invertable and

$$\begin{aligned} AP &= (A\underline{v}_1 \ \dots \ A\underline{v}_n) \\ &= (\lambda_1 \underline{v}_1 \ \dots \ \lambda_n \underline{v}_n) \\ &= (\underline{v}_1 \dots \underline{v}_n)\begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} = PD. \end{aligned}$$

Finding eigenvalues and eigenvectors

Given A, want to find $\underline{v} \neq \underline{0}$, λ s.t.

$$A\underline{v} = \lambda \underline{v}.$$

Prop^n 4.5

Let A be an $n \times n$ matrix over $\mathbb{R}$, $\lambda \in \mathbb{R}$.

Then T.F.A.E.

(i) λ is an eigenvalue

(ii) $\lambda I_n - A$ is not invertable

(iii) $\det(\lambda I_n - A) = 0$

Proof

(i) $\Rightarrow$ (ii)

Suppose $A\underline{v} = \lambda \underline{v}$ $(\underline{v} \neq 0)$

$$A\underline{v} = (\lambda I)\underline{v}$$

$$(\lambda I - A)\underline{v} = 0$$

If $\lambda I - A$ were invertable this would imply $\underline{v} = \underline{0}$, contradiction.

$\therefore$ $\lambda I - A$ is not invertable.

(ii) $\Rightarrow$ (i)

$\lambda I - A$ is not invertable.

Then $(\lambda I - A)\underline{x} = \underline{0}$ has a non trivial solution, say $\underline{v}$. Then $A\underline{v} = \lambda \underline{v}$, so $\lambda \underline{v}$ is an eigenvalue.

(ii) $\Rightarrow$ (iii)

From chapter 3.

To find eigenvalues, solve $\det(\lambda I - A) = 0$.

Example

$$A = \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}$$

Solve $\det(tI - A) = 0$

$$\det\left(\begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} - \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}\right) = 0$$

04/03/16 1202

L12

$$\Rightarrow \det\begin{pmatrix} t-1 & -2 \\ -6 & t-2 \end{pmatrix} = 0$$

so $(t-1)(t-2) - (-2)(-6) = 0$

$\Rightarrow t^2 - 3t - 10 = 0$

$\Rightarrow (t-5)(t+2) = 0$

Roots are $+5$ and -2.

Now find corresponding eigenvectors

$\lambda = 5$

$A\underline{v} = 5\underline{v}$

$(A - 5I)\underline{v} = 0$

$$\Rightarrow \begin{pmatrix} -4 & 2 \\ 6 & -3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} -4 & 2 \\ 6 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -\frac{1}{2} \\ 6 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}$$

so $\begin{pmatrix} 1 & -\frac{1}{2} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

$\Rightarrow x = \frac{1}{2}y$

so general soln : $\begin{pmatrix} \frac{1}{2}y \\ y \end{pmatrix} = y\begin{pmatrix} \frac{1}{2} \\ 1 \end{pmatrix}$

take $\underline{v}_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$

$\lambda = -2$

$$\begin{pmatrix} 3 & 2 \\ 6 & 4 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

take $\underline{v}_2 = \begin{pmatrix} -2 \\ 3 \end{pmatrix}$

$$P = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}$$

$\det P = 3 + 4 = 7 \neq 0$

so P invertable.

Check:

$$AP = \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}\begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 10 & -6 \end{pmatrix}$$

$$PD = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix} = \begin{pmatrix} 5 & 4 \\ 10 & -6 \end{pmatrix}$$

so $P^{-1}AP = D$

Ex

Diagonalise $\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$

$\det(AI - A) = 0$

$\det \begin{pmatrix} \lambda - 2 & 1 \\ 1 & \lambda - 2 \end{pmatrix} = 0$

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

$\Rightarrow \lambda = 2 \pm 1$, $\lambda = 3$ or 1

07/03/16 1202

L13

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

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

$$P^{-1}AP = D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

<u>Applications</u>

1). Find A^m (4.6)

If D is diagonal, easy to find D^m

$$D = \begin{pmatrix} d_1 & & & 0 \\ & d_2 & & \\ & & \ddots & \\ 0 & & & d_n \end{pmatrix}$$

$$D^m = \begin{pmatrix} d_1^m & & & 0 \\ & d_2^m & & \\ & & \ddots & \\ 0 & & & d_n^m \end{pmatrix}$$

Now suppose $P^{-1}AP = D$

Then $A = PDP^{-1}$

$$A^m = PD\underbrace{P^{-1}\cdot P}D\underbrace{P^{-1}\cdots P}DP^{-1}$$

$$= PD^mP^{-1}$$

eg. $A = \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}$, $P = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}$, $P^{-1} = \frac{1}{7}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$

$$P^{-1}AP = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}$$

So $A = P\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}P^{-1} \Rightarrow A^m = P\begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}^m P^{-1}$

$$= \begin{pmatrix} 1 & -2 \\ 2 & 7 \end{pmatrix}\begin{pmatrix} 5^m & 0 \\ 0 & (-2)^m \end{pmatrix}\frac{1}{7}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$$

$$A^m = \frac{1}{7}\begin{pmatrix} 5^m & -2(-2)^m \\ 2(5^m) & 3(-2)^m \end{pmatrix}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$$

$$= \frac{1}{7}\begin{pmatrix} 3(5^m)+4(-2)^m & 2(5^m)-2(-2)^m \\ 6(5^m)-6(-2)^m & 4(5^m)+3(-2)^m \end{pmatrix}$$

<u>Ex</u>

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \quad P = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$P^{-1}AP = D = \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix}$$

Find A^m. What does the formula give for $m=-1$, $m=\frac{1}{2}$?

$$A^m = PD^mP^{-1}$$

$$= \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}\begin{pmatrix} 3^m & 0 \\ 0 & 1^m \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 3^m & -1 \\ 3^m & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \frac{1}{2}\begin{pmatrix} 3^m+1 & 3^m-1 \\ 3^m-1 & 3^m+1 \end{pmatrix}$$

$m = -1$: $A^{-1} = \frac{1}{2}\begin{pmatrix} \frac{4}{3} & -2/3 \\ -2/3 & 4/3 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$

$m = \frac{1}{2}$: $A^{\frac{1}{2}} = \frac{1}{2}\begin{pmatrix} \sqrt{3}+1 & \sqrt{3}-1 \\ \sqrt{3}-1 & \sqrt{3}+1 \end{pmatrix}$ 「$(A^{\frac{1}{2}})^2 = A$」

4.7 - Solving simultaneous linear difference equations.

A linear difference equation $x_{n+1} = ax_n$
has solution $x_n = a^n x_0$.
We can have difference equations involving 2 variables,
e.g. $x_{n+1} = ax_n + by_n$
$y_{n+1} = cx_n + dy_n$

$$\underline{z}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$$

$$\underline{z}_{n+1} = A\underline{z}_n \quad , \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

Solution: $\underline{z}_n = A^n z_0$

4.8 - Solving simultaneous linear differential equations

A very simple type of differential equation is

$$\frac{dx}{dt} = ax$$

$$\int \frac{dx}{x} = \int a\,dt$$

$\Rightarrow \ln x = at + c$
$\Rightarrow x = e^{at+c}$ so $x = Ke^{at}$

Consider simultaneous linear 1st order ODEs:

$$\frac{dx_1}{dt} = ax_1 + bx_2 \quad , \quad \frac{dx_2}{dt} = cx_1 + dx_2$$

$$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad , \quad \underline{x}' = \begin{pmatrix} x_1' \\ x_2' \end{pmatrix} = \begin{pmatrix} ax_1 + bx_2 \\ cx_1 + dx_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = A\underline{x}$$

Let $\underline{x} = P\underline{y}$

$$(P\underline{y})' = A(P\underline{y})$$
$$P\underline{y}' = AP\underline{y}$$
$$\underline{y}' = (P^{-1}AP)\underline{y}$$

Choose P s.t. $P^{-1}AP = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix}$

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

$$\left.\begin{array}{l} y_1' = d_1 y_1 \\ y_2' = d_2 y_2 \end{array}\right\} \text{2 separate equations.}$$

e.g. $\frac{dx_1}{dt} = x_1 + 2x_2$, $\frac{dx_2}{dt} = 6x_1 + 2x_2$

given that $x_1(0) = 2$, $x_2(0) = 1$

$\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ $\quad \underline{x}' = A\underline{x}$ $\quad A = \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}$

Let $\underline{x} = P\underline{y}$ $\quad P = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}$ ← found previously.

$$(Py)' = A(Py)$$
$$y' = (P^{-1}AP)y$$
$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
$$y_1' = 5y_1 \Rightarrow y_1 = k_1 e^{5t}$$
$$y_2' = -2y_2 \Rightarrow y_2 = k_2 e^{-2t}$$

$$\underline{x} = P\underline{y}$$
$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$
$$= \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} k_1 e^{5t} \\ k_2 e^{-2t} \end{pmatrix}$$

so $\begin{cases} x_1 = k_1 e^{5t} - 2k_2 e^{-2t} \\ x_2 = 2k_1 e^{5t} + 3k_2 e^{-2t} \end{cases}$

Initial conditions: $2 = k_1 - 2k_2$, $1 = 2k_1 + 3k_2$

so $k_1 = 8/7$, $k_2 = -3/7$

$\therefore x_1 = \frac{1}{7}(8e^{5t} + 6e^{-2t})$, $x_2 = \frac{1}{7}(16e^{5t} - 9e^{-2t})$

L14

A is diagonalisable if
$\exists$ inv P s.t. $P^{-1}AP = D$

$\underline{v}$ is an *eigenvector* of A if $\underline{v} \neq \underline{0}$, $A\underline{v} = \lambda \underline{v}$ for some λ. λ is an *eigenvalue*.
We find the eigenvalues by solving
$\det(tI - A) = 0$
this then lets us find the eigenvectors.

Basic criterion
A $(n \times n)$ is digonalisable $\Leftrightarrow$ there is a set of n linear independent eigenvectors.
In this case $P = (\underline{v}_1 \; \dots \; \underline{v}_n)$
where $\underline{v}_1, \dots, \underline{v}_n$ are eigenvectors and $P^{-1}AP$ is diagonal.

Which matrices can be diagonalised?
Def 4.9
Let A be an $n \times n$ matrix. Then the *characteristic polynomial* of A is
$$c_A(t) = \det(tI - A)$$
$c_A(t)$ is a polynomial of degree n.
Its roots are the eigenvalues.

How can a matrix fail to be diagonalisable?
The first way is "not having enough eigenvalues".

eg. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ $\quad c_A(t) = \det(tI - A)$
$$= \det\begin{pmatrix} t & -1 \\ 1 & t \end{pmatrix} = t^2 + 1$$

t^2+1 has no real roots so over $\mathbb{R}$, A cannot be digonalised.
However over $\mathbb{C}$, there are two eigenvalues $(\pm i)$ and 2 LI eigenvectors, and so A can

be diagonalised.
Over $\mathbb{C}$ this problem can't arise.

Thm 4.10 (Fundamental Theorem of Algebra)
Any polynomial with complex coefficients factorises into linear factors.

So we will assume from now on that $c_A(t)$ factorises into linear factors:
$$c_A(t) = (t-\lambda_1)^{f_1} \dots (t-\lambda_r)^{f_r}$$
where $\lambda_1, \dots, \lambda_r$ are the eigenvalues and
$f_1 + \dots + f_r = n$.
The simplest case is where $r=n$ and all $f_i = 1$.

Thm 4.11
Let A be an $n \times n$ matrix with n distinct eigenvalues.
Then A is diagonalisable.

Proof
Let $\lambda_1, \dots, \lambda_n$ be the eigenvalues with corresponding eigenvectors $\underline{v}_1, \dots, \underline{v}_n$.
By the basic criterion, it is enough to prove that $\{\underline{v}_1, \dots, \underline{v}_n\}$ is LI
We prove by contradiction.
Suppose $\underline{v}_1, \dots, \underline{v}_n$ are linearly dependent.
Pick a shortest possible relation of dependence (i.e. involving as few as possible non-zero terms).
By re-ordering, we can assume this is:
$$\alpha_1 \underline{v}_1 + \dots + \alpha_r \underline{v}_r = \underline{0} \quad (\text{all } \alpha_i \neq 0) \quad (1)$$
$$A(\alpha_1 \underline{v}_1 + \dots + \alpha_r \underline{v}_r) = A\,\underline{0}$$
$$\Rightarrow \alpha_1 A\underline{v}_1 + \dots + \alpha_r A\underline{v}_r = \underline{0}$$
$$\Rightarrow \alpha_1 \lambda_1 \underline{v}_1 + \dots + \alpha_r \lambda_r \underline{v}_r = \underline{0} \quad (2)$$
$$\alpha_1 \lambda_r \underline{v}_1 + \dots + \alpha_r \lambda_r \underline{v}_r = \underline{0} \quad (1) \times \lambda_r$$

Subtracting the last two lines gives

$$\alpha_1(\lambda_1 - \lambda_r)\underline{v}_1 + \dots + \alpha_r(\lambda_r - \lambda_r)\underline{v}_r = \underline{0} \qquad (3)$$

Since all λ_i distinct, $\alpha_i(\lambda_i - \lambda_r)\underline{v}_i \neq 0 \quad (i = 1, \dots, r-1)$

i.e. (3) is a shorter relation of dependence that (1). ✱

[note $r = 1$ is not possible since then $\alpha_1 \underline{v} = \underline{0}$, $\alpha_1 \neq 0$, $\underline{v}_1 \neq 0$] $\square$

Method for diagonalising an $n \times n$ matrix with n distinct eigenvalues.

1). Find $c_A(t) = \det(tI - A)$

2). Factorise into linear factors: by assumption
$$c_A(t) = (t - \lambda_1) \dots (t - \lambda_n)$$

3). For each λ_i, solve $A\underline{x} = \lambda_i \underline{x}$ to find a (non-zero) eigenvector $\underline{v}_i$.

4). The set $\{\underline{v}_1, \dots, \underline{v}_n\}$ is LI.

5). Let $P = (\underline{v}_1 \dots \underline{v}_n)$ (P invertable)

6). Then $P^{-1}AP = D$
$$= \operatorname{diag}(\lambda_1, \dots, \lambda_n)$$

Check: $AP = PD$

and that P is invertable.

Ex

Diagonalise $A = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix}$

$$c_A(t) = \det \begin{pmatrix} t-1 & 0 & 0 \\ 1 & t-2 & 4 \\ 0 & 0 & t \end{pmatrix} = t(t-1)(t-2)$$

so $\lambda_1 = 0$, $\lambda_2 = 1$, $\lambda_3 = 2$

$\lambda = 0$: $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$ so $\begin{cases} x_1 = 0 \\ -x_1 + 2x_2 - 4x_3 = 0 \\ 0x_3 = 0 \end{cases}$

$\Rightarrow \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}$

$\lambda = 1$: $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

so $\begin{cases} x_1 = x_1 \\ -x_1 + 2x_2 - 4x_3 = x_2 \\ 0 = x_3 \end{cases}$

$\Rightarrow \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$

$\lambda = 2$: $\begin{pmatrix} 1 & 0 & 0 \\ -1 & 2 & -4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = 2\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$

so $\begin{cases} x_1 = 2x_1 \\ -x_1 + 2x_2 - 4x_3 = 2x_2 \\ 0 = x_3 \end{cases}$

$\Rightarrow \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$

so $P = \begin{pmatrix} 0 & 1 & 0 \\ 2 & 1 & 1 \\ 1 & 0 & 0 \end{pmatrix}$

so $P^{-1}AP = D = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

What else can stop diagonalisation?
We need to look at repeated eigenvalues

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

$c_A(t) = \det \begin{pmatrix} t-3 & -1 \\ 0 & t-3 \end{pmatrix} = (t-3)^3$

One eigen-value (3) repeated.

Eigenvectors?

$A\underline{v} = 3\underline{v}$

$(A - 3I)\underline{v} = \underline{0}$

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

solution: $\begin{pmatrix} \alpha \\ 0 \end{pmatrix}$, $\alpha \in \mathbb{R}$

Hence there are <u>not</u> 2 LI eigenvectors so A is not diagonalisable.

However it is not that a repeated root necessarily stops diagonalisation.

e.g. $B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$ $c_B(t) = (t-3)^2$

$\lambda = 3$ (twice)

Eigenvectors:

$A\underline{v} = 3\underline{v}$

$(A - 3I)\underline{v} = 0$

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

So any $\begin{pmatrix} x \\ y \end{pmatrix}$ is a solution

e.g. $\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ are eigen vectors.

B is diagonalisable.

We need ideas about subspaces, sums, direct sums, etc.

Def 4.13 (Revision)

A subset W of a vector space V is a subspace if $W \neq \emptyset$ and $\underline{u}, \underline{v} \in W$, $\lambda, \mu \in \mathbb{R}$ $\Rightarrow \lambda \underline{u} + \mu \underline{w} \in W$.

A subspace of V forms a vector space itself under the same operations.

Examples

1). Subspaces of $\mathbb{R}^2$ are:

ⓐ $\{\underline{0}\}$

ⓑ a line through the origin

ⓒ $\mathbb{R}^2$

2). $\{\underline{x} : A\underline{x} = \underline{0}\}$ is a subspace of $\mathbb{R}^n$

3). Subspaces of $\mathbb{R}^3$:

ⓐ $\{\underline{0}\}$

ⓑ line through the origin

ⓒ plane through the origin

ⓓ $\mathbb{R}^3$

Def 4.14 (Revision)

Let $U, W \leq V$. subspace of

Then $U + W = \{\underline{u} + \underline{w} : u \in U, w \in W\}$

Prop 4.15

Let $U, W \leq V$. Then $U + W$, $U \cap W \leq V$.

Proof

$U \neq \emptyset$, $W \neq \emptyset$, so $U + W \neq \emptyset$.

Let $v_1, v_2 \in U + W$, $\lambda, \mu \in \mathbb{R}$

Then $v_1 = u_1 + w_1$, $v_2 = u_2 + w_2$ for some $u_i \in U$, $w_i \in W$.

$\lambda_1 v_1 + \lambda_2 v_2 = \lambda_1 (u_1 + w_1) + \lambda_2 (u_2 + w_2)$

$$= (\lambda_1 u_1 + \lambda_2 u_2) + (\lambda_1 w_1 + \lambda_2 w_2) \in U + W.$$

$U \cap W$ similar.

eg. $V = \mathbb{R}^2$

$U = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R} \right\} \leq \mathbb{R}^2$

$W = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R} \right\} \leq \mathbb{R}^2$ ← x is a dummy variable!

$$U \cap W = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$$

$$U + W = \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} x \\ x \end{pmatrix} = \begin{pmatrix} 2x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}$$ no!

$$= \left\{ \begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} y \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}$$ ✓

$$= \left\{ \begin{pmatrix} x + y \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\} = \mathbb{R}^2$$

<u>Ex</u>

Let $V = \mathbb{R}^3$

$$V = \left\{ \begin{pmatrix} x \\ x \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\} \leq \mathbb{R}^3$$

$$W = \left\{ \begin{pmatrix} x \\ y \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\} \leq \mathbb{R}^3$$

Find $U + W$, $U \cap W$ and dind the dimensions of U, W, $U + W$, $U \cap W$. What is the relation?

$$U + W = \left\{ \begin{pmatrix} x + u \\ x + v \\ y + v \end{pmatrix} : x, y, u, v \in \mathbb{R} \right\}$$

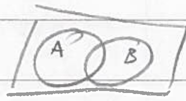
$$U \cap W = \left\{ \begin{pmatrix} x \\ x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}$$

$\dim U = 2$
$\dim W = 2$
$\dim(U + W) = 3$
$\dim(U \cap W) = 1$

14/03/16 | 1202

L15

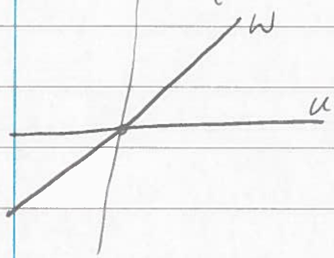
Thm 4.16 (Revision)
Let $U, W \leq V$,
then $\dim(U+W) = \dim(U) + \dim(W) - \dim(U \cap W)$

[Compare for sets $|A \cup B| = |A| + |B| - |A \cap B|$ 

]

Def 4.17
Let $U, W \leq V$.
The sum $U+W$ is direct if $U \cap W = \{\underline{0}\}$
and then we write $U+W = U \oplus W$.
e.g. $U = \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\} \leq \mathbb{R}^2$

$W = \left\{\begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R}\right\} \leq \mathbb{R}^2$



$U+W$ is direct ($U \cap W = \{\underline{0}\}$)
$U \oplus W = \mathbb{R}^2$

from 4.16, $\dim(U \oplus W) = \dim U + \dim W$

Now we want to define a direct sum for several subspaces. First the sum.

Def 4.18
Let $U_i \leq V$ $(i = 1, \dots, r)$. Then
$\sum_{i=1}^{r} U_i = U_1 + U_2 + \dots + U_r = \left\{ \sum_{i=1}^{r} \underline{u}_i : \underline{u}_i \in \mathbb{R}^2 \right\}$

It is very easy to check $\sum_{i=1}^{r} U_i \leq V$
What about directness? eg. $U+W+X$,
requiring $U \cap W = \{\underline{0}\}$, $U \cap X = \{\underline{0}\}$, $W \cap X = \{\underline{0}\}$ doesn't work well.
eg. $V = \mathbb{R}^2$, $U = \{\begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R}\}$, $W = \{\begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R}\}$, $X = \{\begin{pmatrix} 0 \\ x \end{pmatrix} : x \in \mathbb{R}\}$.

Need $(U+W) \cap X = \{\underline{0}\}$

Def 4.19

$\sum_{i=1}^{r} U_i$ is direct (write $\sum_{i=1}^{r} U_i = \bigoplus_{i=1}^{r} U_i$) if for all i, $U_i \cap \left(\sum_{j \neq i} U_j\right) = \{\underline{0}\}$

e.g. $V = \mathbb{R}^3$

$U_1 = \left\{\begin{pmatrix} x \\ 0 \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\}$

$U_2 = \left\{\begin{pmatrix} 0 \\ x \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\}$

$U_3 = \left\{\begin{pmatrix} 0 \\ 0 \\ x \end{pmatrix} : x \in \mathbb{R}\right\}$

then $U_1 + U_2 + U_3 = U_1 \oplus U_2 \oplus U_3$

Lemma 4.20

Let $U_i \leq V$ $(i = 1, \dots, r)$.

Then the following are equivalent:

(i) $\sum_{i=1}^{r} U_i$ is direct

(ii) $\sum_{i=1}^{r} \underline{u}_i = \underline{0} \Rightarrow$ all $\underline{u}_i = \underline{0}$

Proof

$(\Rightarrow)$ Suppose $\sum U_i$ is direct and $\sum \underline{u}_i = \underline{0}$ $(\underline{u}_i \in U_i)$.

Then $\underline{u}_1 = -\sum_{i=2}^{r} \underline{u}_i \in U_1 \cap \left(\sum_{i=2}^{r} U_i\right) = \{\underline{0}\}$

$\therefore$ $\underline{u}_1 = \underline{0}$. Similarly all $\underline{u}_i = \underline{0}$.

$(\Leftarrow)$ Suppose (ii) holds and

$\underline{v} \in U_1 \cap \left(\sum_{i=2}^{r} U_i\right)$

$\underline{v} = \underline{u}_1 = \sum_{i=2}^{r} \underline{u}_i$ $(\underline{u}_j \in U_j)$

$\underline{u}_1 - \sum_{i=2}^{r} \underline{u}_i = 0$

? By ( ) $\underline{u}_1 = 0$

$\therefore U_1 \cap \left(\sum_{i=2}^{r} U_i\right) = \{\underline{0}\}$

14/03/16 | 1202

L15

Similarly $U_i \cap (\sum_{j \neq i} U_j) = \{\underline{0}\}$

<u>Lemma 4.21</u>

Let $U_i \leq V$ $(i = 1, \ldots, r)$ and suppose $\sum_{i=1}^{r} U_i$ is direct. Let B_i be a basis for U_i.

Then

(i) $B = \bigcup_{i=1}^{r} B_i$ is a basis for $\bigoplus_{i=1}^{r} U_i$

(ii) $\dim(\bigoplus_{i=1}^{r} U_i) = \sum_{i=1}^{r} \dim(U_i)$

<u>Proof</u>

(i) <u>Spanning</u>

Let $\underline{v} \in \bigoplus_{i=1}^{r} U_i$

By def, $\underline{v} = \sum_{i=1}^{r} \underline{u_i}$ $(\underline{u_i} \in U_i)$

Since $B_i = \{ \underline{b}_1^{(i)}, \ldots, \underline{b}_{n_i}^{(i)} \}$.

Each $\underline{u_i} = \sum_{j=1}^{n_i} \lambda_j^{(i)} \underline{b}_j^{(i)}$

$\therefore \underline{v} = \sum_{i=1}^{r} \sum_{j=1}^{n_i} \lambda_j^{(i)} \underline{b}_j^{(i)}$

$B = \{ \underline{b}_1^{(1)}, \ldots, \underline{b}_{n_1}^{(1)}, \underline{b}_1^{(2)}, \ldots, \underline{b}_{n_2}^{(2)}, \ldots \}$

i.e. $\underline{v}$ is in the linear span of B.

(ii) <u>LI</u>

Suppose $\sum_{i,j} \lambda_j^{(i)} \underline{b}_j^{(i)} = \underline{0}$

$$= \sum_{i=1}^{r} \underbrace{\left(\sum_{j=1}^{n_i} \lambda_j^{(i)} \underline{b}_j^{(i)} \right)}_{\in U_i} = \underline{0}$$

Since $\sum U_i$ is direct, each $\sum_{j=1}^{n_i} \lambda_j^{(i)} \underline{b}_j^{(i)} = \underline{0}$.

Since $\mathcal{B}_i$ LI, each $\lambda_j^{(i)} = 0$.

Def 4.22

Let λ be an eigenvalue of the $n \times n$ matrix A. Then the eigenspace associated with λ is

$$E_\lambda = \{ \underline{v} \in \mathbb{R}^n : A\underline{v} = \lambda \underline{v} \}$$

18/03/16 1202

L18

Can A be diagonalised? ($n \times n$ matrix)

Basic Criterion:

A can be diagonalised $\Leftrightarrow \exists\, n$ LI eigenvectors

For each eigenvalue λ, there is at least one eigenvector.
There are n eigenvalues, counting multiplicity.
If all eigenvalues are distinct (ie. no repeated roots) then this gives n eigenvectors, and we proved these were LI: hence A can be diagonalised.

What happens with repeated eigenvalues?
We need to look at all eigenvectors associated to a given eigenvalue.

Def^n 4.22

Let λ be an eigenvalue of A. Then the eigenspace (associated to λ) is

$$E_\lambda = \{\underline{v} \in \mathbb{R}^n : A\underline{v} = \lambda \underline{v}\}$$

Prop^n 4.23

E_λ is a subspace of $\mathbb{R}^n$.

Proof

$\underline{0} \in E_\lambda$, since $A\underline{0} = \underline{0} = \lambda \underline{0}$.
Let $\underline{u}, \underline{w} \in E_\lambda$, $c \in \mathbb{R}$. Then

$A\underline{u} = \lambda \underline{u}$, $A\underline{w} = \lambda\underline{w}$.

Hence $A(\underline{u}+\underline{w}) = A\underline{u} + A\underline{w} = \lambda\underline{u} + \lambda\underline{w} = \lambda(\underline{u}+\underline{w})$,
so $\underline{u}+\underline{w} \in E_\lambda$,
and $A(c\underline{u}) = cA\underline{u} = c\lambda\underline{u} = \lambda(c\underline{u})$, so $c\underline{u} \in E_\lambda$.

Prop^n 4.24

Let $\lambda_1, \dots, \lambda_r$ be the eigenvalues of the $n \times n$ matrix A.

Then the sum $\sum_{i=1}^{r} E_{\lambda_i}$ is direct.

Proof

WTP: $\sum_{i=1}^{r} \underline{u}_i = \underline{0}$ $(\underline{u}_i \in E_{\lambda_i}) \Rightarrow$ all $\underline{u}_i = \underline{0}$

Suppose not.

Pick a shortest non-trivial sum of form (*) and re-number to get:

(1) $\sum_{i=1}^{p} \underline{u}_i = \underline{0}$, $\underline{u}_i \neq 0$ $(i = 1, \dots, p)$, $\underline{u}_i \in E_{\lambda_i}$

and there is no non-zero sum involving $< p$ terms giving $\underline{0}$. Note $p > 1$ ($p = 1$ says $\underline{u}_1 = 0$ but $\underline{u}_1 \neq 0$)

$$A \sum_{i=1}^{p} \underline{u}_i = A\underline{0}$$

$$\sum_{i=1}^{p} A\underline{u}_i = \underline{0}$$

$$\text{so} \quad \sum_{i=1}^{p} \lambda_i \underline{u}_i = 0 \qquad (2)$$

(2) $- \lambda_p \times$ (1) : $\sum_{i=1}^{p-1} (\lambda_i - \lambda_p)\underline{u}_i = 0$

Let $\underline{u}_i' = (\lambda_i - \lambda_p)\underline{u}_i$

Then $\underline{u}_i' \in E_{\lambda_i}$, $\underline{u}_i' \neq 0$ (since $\lambda_i \neq \lambda_p$, $\underline{u}_i \neq 0$)

and $\sum_{i=1}^{p-1} \underline{u}_i' = \underline{0}$. ※ Contradiction to (1) being the shortest such relation.

$\therefore$ There is no such relation (1), ie. the sum is direct.

18/03/16 1202

L16

Def^n 4.25

Let A be an $n \times n$ matrix over $\mathbb{R}$ with eigenvalues $\lambda_1, \dots, \lambda_r$ (distinct). Suppose $c_A(t)$ factorises into linear factors over $\mathbb{R}$, say

$$c_A(t) = (t-\lambda_1)^{f_1} \dots (t-\lambda_r)^{f_r} \qquad (f_i \geq 1).$$

Then

(i) the <u>algebraic multiplicity</u> of λ_i is f_i

(ii) the <u>geometric multiplicity</u> of λ_i is $e_i = \dim(E_{\lambda_i})$

Thm 4.26

Let A be as above.

Then A is diagonalisable

$$\iff e_i = f_i \qquad (i = 1, \dots, r)$$

Proof

($\Leftarrow$) Suppose $e_i = f_i$.

Sum $\sum_{i=1}^{r} E_{\lambda_i}$ is direct.

Pick a basis B_i for each E_{λ_i}.

By 4.21 $B = \bigcup_{i=1}^{r} B_i$ is a basis for $\bigoplus_{i=1}^{r} E_{\lambda_i}$.

Now $\dim \bigoplus_{i=1}^{r} E_{\lambda_i} = \sum_{i=1}^{r} \dim(E_{\lambda_i})$

$$= \sum_{i=1}^{r} e_i = \sum_{i=1}^{r} f_i = \deg(c_A(t)) = n.$$

$\therefore \bigoplus_{i=1}^{r} E_{\lambda_i}$ is a subspace of $\mathbb{R}^n$ of dimension n

$\therefore$ Hence B is a basis for $\mathbb{R}^n$ consisting of eigenvectors

So by Basic Criterion, A is diagonalisable.

This gives a method of diagonalising

e.g. $A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{pmatrix}$

$$c_A(t) = \det \begin{pmatrix} t-3 & 1 & 0 \\ 1 & t-3 & 0 \\ -1 & 1 & t-4 \end{pmatrix}$$

$$= (t-4)\left[(t-3)^2 - 1\right]$$
$$= (t-4)^2(t-2)$$

So $\lambda_1 = 4, \quad \lambda_2 = 2$

$f_1 = 2, \quad f_2 = 1$

$E_4 = \{\underline{v} : A\underline{v} = 4\underline{v}\}$

$$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ -1 & 1 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$$

$$= \left\{ \begin{pmatrix} y \\ y \\ z \end{pmatrix} : y, z \in \mathbb{R} \right\}$$

$$= \left\{ y \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : y, z \in \mathbb{R} \right\}$$

So E_4 has a basis : $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}$

$E_2 = \{\underline{v} : A\underline{v} = 2\underline{v}\}$

gives basis : $\left\{ \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$

so $e_1 = 2, \quad e_2 = 1$

Since $e_1 = f_1$, $e_2 = f_2$, A is diagonalisable and a basis of eigenvectors is $\left\{ \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \right\}$

So $P = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}$

Then $P^{-1}AP = D = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$

18/03/16 1202

L16

Check:

$\det P = 2 \neq 0$ so P is invertable

$$AP = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 0 & 2 \\ 4 & 0 & -2 \\ 0 & 4 & 2 \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & -1 \\ 0 & 1 & 1 \end{pmatrix}\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} = PD$$

Continuation of proof.

$(\Rightarrow)$

To prove the converse, we need a lemma:

<u>Lemma 4.27</u>

With the above notation, $1 \leq e_i \leq f_i$

<u>Proof</u>

$1 \leq e_i$ is just in the definition of eigenvalues.

We'll prove $e_1 \leq f_1$. Write $\lambda = \lambda_1$, $e = e_1$, $f = f_1$.

Let $\{\underline{v}_1, \ldots, \underline{v}_e\}$ be a basis for E_λ: extend to a basis $\{\underline{v}_1, \ldots, \underline{v}_n\}$ for $\mathbb{R}^n$.

Let $P = (\underline{v}_1 \ldots \underline{v}_n)$

P is invertable and

$$\begin{aligned} AP &= A(\underline{v}_1 \ldots \underline{v}_n) \\ &= (A\underline{v}_1 \ldots A\underline{v}_n) \\ &= (\lambda\underline{v}_1 \ldots \lambda\underline{v}_e \; A\underline{v}_{e+1} \ldots A\underline{v}_n) \\ &= (\underline{v}_1 \ldots \underline{v}_n)\begin{pmatrix} \lambda & 0 & 0 & \\ 0 & \lambda & \vdots & \\ \vdots & 0 & \lambda & * \\ \vdots & \vdots & \vdots & \\ 0 & 0 & 0 & \end{pmatrix} \end{aligned}$$

?

$$= P\begin{array}{c|cc} & e & n-e \\ e & \lambda I_e & X \\ n-e & 0 & Y \end{array} \quad \text{(say)}$$

$$P^{-1}AP = \begin{pmatrix} \lambda I_e & X \\ 0 & Y \end{pmatrix} = B \quad \text{(say)}$$

$$c_B(t) = \det(tI - B)$$
$$= \det\begin{pmatrix} (t-\lambda)I_e & X \\ 0 & tI_{n-e} - Y \end{pmatrix}$$

$$= (t-\lambda)^e g(t)$$

$$c_A(t) = c_B(t) = (t-\lambda)^e g(t)$$
$\therefore$ $(t-\lambda)^e$ divides $c_A(t) = (t-\lambda)^f (t-\lambda_2)^{f_2} \dots (t-\lambda_n)^{f_n}$
$\therefore$ $e \leq f$

Proof ($\Rightarrow$ of 4.26)
Suppose A is diagonalisable.
Each $e_i \leq f_i$ Hence if some $e_i < f_i$ then
$$\sum_{i=1}^{r} e_i < \sum_{i=1}^{r} f_i = n$$

Hence $\dim\left(\bigoplus_{i=1}^{r} E_{\lambda_i}\right) = \sum e_i < n$
But since A is diagonalisable, there are n LI eigenvectors which all lie in $\bigoplus_{i=1}^{r} E_{\lambda_i}$. ※
This is a contradiction. (can't have n LI vectors in a space of dimension $< n$)
$\therefore$ each $e_i = f_i$.

Exercise
Let $A = \begin{pmatrix} 2 & 1 & 0 \\ -1 & 4 & 0 \\ 0 & 0 & 3 \end{pmatrix}$

Find eigenvalues, eigenspaces, f_i, e_i & determine if A is diagonalisable.
$$c_A(t) = \det\begin{pmatrix} t-2 & 1 & 0 \\ -1 & t-4 & 0 \\ 0 & 0 & t-3 \end{pmatrix}$$
$$= (t-3)\big((t-2)(t-4)+1\big)$$
$$= (t-3)(t-3)(t-3)$$

18/03/16 . 1202

L16

so $\lambda_1 = 3$
$f_1 = 3$

$E_3 = \{\underline{v} : A\underline{v} = 3\underline{v}\}$

$= \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : \begin{pmatrix} -1 & 1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \right\}$

$= \left\{ \begin{pmatrix} x \\ x \\ z \end{pmatrix} : x, z \in \mathbb{R} \right\}$

$= \left\{ x\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + z\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} : x, z \in \mathbb{R} \right\}$

$\therefore e_1 = \dim E_3 = 2 < 3 = f_1$
$\therefore$ A is not diagonalisable.

In fact, there is an "almost" diagonal form, called Jordan Normal Form. (Math 2201)

<u>The minimum polynomial and the Cayley Hamilton Theorem.</u>

<u>Def. 4.29</u>
To matrices A and B are <u>similar</u> if $\exists$ invertable matrix P s.t. $P^{-1}AP = B$.

<u>Lemma 4.30</u>
Suppose A and B are similar. Then $c_B(t) = c_A(t)$.

<u>Proof</u>
$B = P^{-1}AP$
$c_B(t) = \det(tI - B)$
$= \det(tI - P^{-1}AP)$
$= \det(P^{-1}(tI - A)P)$

$= \det P^{-1} \cdot \det(tI - A) \cdot \det P$

$= (\det P)^{-1} \cdot \det P \cdot \det(tI - A)$

$= \det(tI - A) = c_A(t).$

21/03/16 1202

L17

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

$$c_A(t) = (t-2)(t-3) = t^2 - 5t + 6$$

$$A^2 - 5A + 6I = \begin{pmatrix} 4 & 0 \\ 0 & 9 \end{pmatrix} - \begin{pmatrix} 10 & 0 \\ 0 & 15 \end{pmatrix} + \begin{pmatrix} 6 & 0 \\ 0 & 6 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$f(t) = t^2 - 5t + 6$

$f(A) = 0$

Prop^n 4.31

Let A be an $n \times n$ matrix over $\mathbb{R}$.
Then there exists a non-zero polynomial $f(t) \in \mathbb{R}[t]$ s.t. $f(A) = 0$.

Proof

Consider $M_n(\mathbb{R})$ (the $n \times n$ matrices over $\mathbb{R}$) as a vector space over $\mathbb{R}$.
This has basis $\{E(i,j) : 1 \le i, j \le n\}$.

Hence $\dim_{\mathbb{R}}(M_n(\mathbb{R})) = n^2$.
Hence the set of n^2+1 matrices $I, A, A^2, \dots, A^{n^2}$ must be linearly dependent,
say $\alpha_0 I + \alpha_1 A + \dots + \alpha_{n^2} A^{n^2} = 0$ (not all $\alpha_i = 0$)
let $f(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_{n^2} t^{n^2} \in \mathbb{R}[t]$
f is non zero and $f(A)$ is non zero.

A polynomial is called monic if it has leading coefficient 1.
eg. $t^3 - 3t + 2$ is monic.

Thm 4.32

Let $A \in M_n(\mathbb{R})$

(i) there exists a unique monic polynomial of least degree s.t. $m(A) = 0$.

(ii) if $f \in \mathbb{R}[t]$, and $f(A) = 0$, then m divides f.

Proof

(i) By 4.31, $\exists$ non-zero f s.t. $f(A) = 0$.
Let m be a monic polynomial of least degree s.t. $m(A) = 0$.
Suppose m' is another such.
Let $f = m - m'$, a polynomial of degree $< \deg(m)$
and $f(A) = m(A) - m'(A) = 0 - 0 = 0$.
If $f \neq 0$ dividing by the coefficient of top term in f gives a monic polynomial, q of degree $< \deg(m)$ s.t. $q(A) = 0$. # this is a contradiction.

$\therefore f = 0$ and $m = m'$

$\therefore m$ is unique.

(ii) Suppose $f(A) = 0$.
Write $f(t) = m(t)q(t) + r(t)$ with $\deg(r) < \deg(m)$

$$\underset{0}{\underset{\shortparallel}{f(A)}} = \underset{0}{\underset{\shortparallel}{m(A)}}\, q(A) + r(A)$$

so $r(A) = 0$.
This again yields a contradiction unless $r = 0$.

$\therefore r(t) = 0$

$\therefore f = mq$.

m is called the minimal polynomial of A.

21/03/16 1202

L17

eg. (i) $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$

$m_A(t) = (t-2)(t-3) = c_A(t)$

since $m_A(A) = 0$ and if $f(t) = t + c$, $f(A) \neq 0$.

(ii) $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$

$f(t) = (t-2)^2$, $f(A) = 0$

$f(t) = t - 2$, $f(A) = A - 2I = 0$

so $m_A(t) = t - 2$. $c_A(t) = (t-2)^2$

(iii) $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$

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

$(A - 2I)^2 = 0$

$m_A(t) = (t-2)^2 = c_A(t)$

<u>Thm 4.33</u> (Cayley - Hamilton Theorem)

Let $A \in M_n(\mathbb{R})$. Then $c_A(A) = 0$,

i.e. $m_A(t) \mid c_A(t)$.

<u>Proof</u>

See notes.