1202 Algebra 2 Notes

Based on the 2012 spring lectures by Dr M L Roberts

1202 — Algebra 2.

18 January 2012
Dr Mark ROBERTS
CLT.

Course outline.

Key Topics

① Linear algebra → determinants → diagonalisation

② Groups

③ Number theory.

CHAPTER 1
NUMBER THEORY.

Here we consider questions about $\mathbb{Z} = \{\dots, -4, -3, -2, -1, 0, 1, 2, \dots\}$.

$\boxed{\text{Definition}}$ 1.1 Let $a, b \in \mathbb{Z}$. Then a divides b (written $a|b$) if $b = az$ for some $z \in \mathbb{Z}$.

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

For instance, $2|6$ because $6 = 3(2)$; $2 \nmid 11$ $\because$ $11 \neq 2z$ for any $z \in \mathbb{Z}$.

Some simple properties of divisibility.

$\boxed{\text{Proposition}}$ 1.2 Let $a, b, c, d, e \in \mathbb{Z}$, $a \neq 0$, then

(i) if $a|b$ and $a|c$, then $a | bd + ce$. (any linear combination)

(ii) if $a|b$ and $b|c$, then $a|b$.

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

Proof — (i) since $a|b$, $\exists z \in \mathbb{Z}$ s.t. $b = az$; since $a|c$, $\exists y \in \mathbb{Z}$ s.t. $c = ay$.

then $bd + ce = azd + aye = a(zd + ye)$. and since $zd + ye \in \mathbb{Z}$, we have $a | bd+ce$, q.e.d.

parts (ii) and (iii) are analogous, proven in Exercise 1.

$\boxed{\text{Definition}}$ 1.3 We say that a factorisation $a = bc$, if b or c is ± 1, is trivial.

If $a > 1$ and a has no non-trivial factorisation, then a is prime.

If it has a non-trivial factorisation, then a is composite.

If $a = \pm 1$, a is a unit.

For instance, 25 is composite because $25 = 5 \times 5$; 7 is prime because $7 = ab \Rightarrow a$ or $b = \pm 1$.

All of the subject of Number theory starts from the fact that any number can be expressed uniquely as a product of primes.

For instance, $24 = 2^3 \times 3$, and apart from re-ordering of factors, there is no other way to express 24 as a product of primes.

To prove this fact, we need to develop some results about divisibility.

DIVISION THEOREM.

$\boxed{\text{Theorem}}$ 1.4 (Division theorem)

Let $a, b \in \mathbb{Z}$, $b > 0$. Then $\exists_{\text{unique}}\, q, r \in \mathbb{Z}$ s.t. $a = bq + r$ with $0 \leq r < b$.

We sometimes call q the quotient and r the remainder.

For instance, $13, 3$: $13 = 3(4) + (1)$ and $-17, 4$: $-17 = 4(-5) + (3)$.

Proof — let q be the largest integer $\leq \frac{a}{b} \in \mathbb{Q}$; i.e. $q = \lfloor \frac{a}{b} \rfloor$.

then $\frac{a}{b} = q + \alpha$, where $0 \leq \alpha < 1$ $\Rightarrow$ $a = bq + \alpha b$, where $0 \leq \alpha b = r < b$.

by closure of $\mathbb{Z}$ under $+$, $a, bq \in \mathbb{Z} \Rightarrow r \in \mathbb{Z}$ $\therefore$ $q, r \in \mathbb{Z}$ exist, q.e.d.

suppose q, r are not unique, then $a = bq + r = bq' + r'$, with $0 \leq r, r' < b$.

$\Rightarrow b(q - q') = r' - r$

since $|r' - r| < b$, $b|q - q'| < b \Rightarrow |q - q'| < 1 \Rightarrow q - q' = 0$ ($\because q, q' \in \mathbb{Z}$) $\Rightarrow q = q'$ and $r = r'$ $\Rightarrow$ q, r are unique, q.e.d.

$\boxed{\text{Definition}}$ 1.5. Let a, b be non-zero integers. Then the <u>highest common factor / greatest common divisor</u> of a and b is the largest positive integer which divides both a and b, denoted $hcf(a,b)$ / $gcd(a,b)$.

e.g. $gcd(6,9)=3$, $gcd(8,2,5)=1$.

If $gcd(a,b)=1$, a and b are called <u>coprime</u>.

EUCLIDEAN ALGORITHM.

This is a method for finding the gcd of 2 numbers by repeated division. It is much more efficient than factorising, for large numbers.

It is much easier to find gcd of 2 numbers than determining if a number is prime.

$\boxed{\text{Theorem}}$ 1.6 (Euclidean algorithm).

Let a, b be positive integers. Then there exist $q_1, \dots, q_{n+1}, r_1, \dots, r_n \in \mathbb{Z}$ [note: one more q than r!]

with $a > b > r_1 > r_2 > \cdots > r_n > 0$ s.t.

$$\left.\begin{aligned} a &= bq_1 + r_1 \\ b &= r_1 q_2 + r_2 \\ &\vdots \\ r_1 &= r_2 q_3 + r_3 \\ &\vdots \\ r_{n-2} &= r_{n-1} q_n + r_n \\ r_{n-1} &= r_n q_{n+1} \end{aligned}\right\}$$

continued iterations of the division theorem on (a,b), then (b, r_1), then (r_1, r_2) ...... until we obtain a pair that divides evenly to produce no remainder.

then $gcd(a,b) = r_n$.

example before proving...

$\boxed{\text{Ex}}$ Find $gcd(1169, 560)$

$1169 = 2 \cdot 560 + 49$

$560 = 11 \cdot 49 + 21$

$49 = 2 \cdot 21 + 7$

$21 = 3 \cdot 7 \quad \Rightarrow gcd(1169, 560) = 7$.

$\boxed{\text{Ex}}$ Find $gcd(30,18)$ by this method. Check answer by definition of gcd.

$30 = 1 \cdot 18 + 12$

$18 = 1 \cdot 12 + 6$

$12 = 2 \cdot 6 \quad \Rightarrow gcd(30,18) = 6$.

check: $\left.\begin{aligned} 30 &= 2 \cdot 3 \cdot 5 \\ 18 &= 2 \cdot 3^2 \end{aligned}\right\} gcd(30,18) = 2 \cdot 3 = 6$ (verified).

Proof of theorem 1.6 — the existence of the q_i and the r_i and the fact that $b > r_1 > r_2 > \cdots$ follow immediately from thm 1.4.

The r_i terms form a strictly decreasing sequence of non-negative integers. Hence, at some stage it would become 0, say $r_{n+1} = 0$.

We now prove, to show that $r_n = gcd(a,b)$, the following.

(i) $r_n | a$ and $r_n | b$, and (ii) if $x|a$ and $x|b$, then $x | r_n$ (and hence $x \le r_n$).

for part (i) — $r_{n-1} = r_n q_{n+1}$, so $r_n | r_{n-1}$

$r_{n-2} = r_{n-1} q_n + r_n$, so since $r_n | r_{n-1}$ and $r_n | r_n$, by Proposition 1.2, $r_n | r_{n-2}$.

likewise, $r_n | r_{n-2}$ and $r_n | r_{n-1} \Rightarrow r_n | r_{n-2} q_{n-1} + r_{n-1} = r_{n-3}$ etc.

by induction then, $r_n | b$ and $r_n | a$.

for part (ii) — suppose $x|a$ and $x|b$, then we have

$a = bq_1 + r_1 \Rightarrow r_1 = a - bq_1$, so $x | r_1$; and likewise $r_2 = b - r_1 q_2 \Rightarrow x | r_2$, etc.

by induction then, $x | r_n$. q.e.d.

LINEAR COMBINATIONS AND THE h,k-LEMMA.

$\boxed{\text{Definition}}$ 1.7 A <u>linear combination</u> of $a, b \in \mathbb{Z}$ is an integer of the form $ar + bs$ $(r, s \in \mathbb{Z})$.

e.g. 20 is a linear combination of 6 and 8, since $20 = 2(6) + 1(8)$.

$\boxed{\text{Ex}}$ Find 1 as a linear combination of 5 and 7. $1 = -4(5) + 3(7) = 3(5) - 2(7)$... etc.

Can you express 1 as a linear combination of 9 and 21? No, because $gcd(9,21) = 3 \nmid 1$.

$\boxed{\text{Theorem}}$ 1.8 Let a, b be non-zero integers and $x \in \mathbb{Z}$. Then

x is a linear combination of a and b $\Leftrightarrow$ $\gcd(a,b) \mid x$.

Proof — forward relation:

$\gcd(a,b) \mid a$ and $\gcd(a,b) \mid b$ $\Rightarrow$ $\gcd(a,b) \mid$ linear combination of a and b $\Rightarrow$ $\gcd(a,b) \mid x$.

backward relation:

recall from the Euclidean algorithm...

$$\left.\begin{aligned} r_1 &= a - bq_1 \\ r_2 &= b - r_1 q_2 \\ r_3 &= r_1 - r_2 q_3 \\ &\vdots \\ r_n &= r_{n-2} - r_{n-1} q_n \end{aligned}\right\}$$

so $r_n = r_{n-2} - r_{n-1} q_n \Rightarrow r_n$ is a linear combination of r_{n-2} and r_{n-1}

$= r_{n-2} - (r_{n-3} - r_{n-2} q_{n-1}) q_n = r_{n-2}(1 + q_{n-1} q_n) - r_{n-3} q_n \Rightarrow r_n$ is a linear comb. of r_{n-3} and r_{n-2}

$\vdots$

proceeding inductively, we see that r_n is a linear combination of a and b.

$r_n \mid x \Rightarrow \exists k \in \mathbb{Z}$ s.t. $x = k r_n \Rightarrow x$ is a linear combination of a and b (since it is a multiple of r_n)// q.e.d.

$\boxed{\text{Ex}}$ Express 1 as a linear combination of 5 and 7.

$7 = 1(5) + 2$ | $1 = 1(5) - 2[1(7) - 1(5)] = 3(5) - 2(7)$//

$5 = 2(2) + 1$ | $1 = 1(5) - 2(2)$

$2 = 2 \cdot 1$

Express 1 as a linear combination of 42 and 19.

$42 = 2(19) + 4$ | $1 = 5[1(42) - 2(19)] - 1(19) = 5(42) - 11(19)$//

$19 = 4(4) + 3$ | $1 = 1(4) - 1[1(19) - 4(4)] = 5(4) - 1(19)$

$4 = 1(3) + 1$ | $1 = 1(4) - 1(3)$

$3 = 3(1)$

20 January 2011
Dr Mark ROBERTS.
Darwin LT.

A particular case of Theorem 1.8 is when $\gcd(a,b)$ is a linear combination of a and b. The most often used is:

$\boxed{\text{Lemma}}$ 1.9 If a and b are coprime, then $\exists h, k \in \mathbb{Z}$ s.t. $ah + bk = 1$.

(also called the h,k-lemma)

UNIQUE FACTORISATION.

$\boxed{\text{Proposition}}$ 1.10 Let p be a prime number. Then if $p \mid ab$, then $p \mid a$ or $p \mid b$.

Proof — consider $\gcd(a,p)$. We know that $\gcd(a,p) \mid p$ by definition.

then $\gcd(a,p)$ is either 1 or p.

- if $\gcd(a,p) = p$, then $p \mid a$
- if $\gcd(a,p) = 1$, then by h,k-lemma, $\exists h,k$ s.t. $ah + pk = \gcd(a,p) = 1$.

multiplying throughout by b, we get $abh + pbk = b$. by hypothesis, $p \mid ab \Rightarrow ab = px$ for some $x \in \mathbb{Z}$.

then equation becomes $px + pbk = b \Rightarrow p(x + bk) = b \Rightarrow p \mid b$// q.e.d.

Note: this theorem does not hold for composite p, e.g. $6 \mid 3 \times 4$ but $6 \nmid 3$ and $6 \nmid 4$.

$\boxed{\text{Corollary}}$ 1.11 Let p be a prime number and $a_1, a_2, \ldots\ldots, a_n \in \mathbb{Z}$. Then $p \mid a_1 \cdots a_n \Rightarrow p \mid$ some a_i.

Proof — use formal inductive proof on n, applying Prop 1.10.

e.g. for $p \mid a_1 a_2 a_3 \Rightarrow p \mid a_1 (a_2 a_3) \Rightarrow p \mid a_1$ or $p \mid (a_2)(a_3) \Rightarrow p \mid a_1$ or $p \mid a_2$ or $p \mid a_3$// q.e.d.

This is the key result to prove the unique factorisation theorem.

$\boxed{\text{Theorem}}$ 1.12 Let $z \in \mathbb{N}$. Then z can be written as a product of prime numbers $z = p_1 \cdots p_n$ (p_i primes),

and this is a unique expression (up to order)

i.e. if $z = q_1 \cdots\cdots q_m$ (q_i primes), then $n = m$ and $\{q_1, q_2, \ldots, q_m\}$ is a rearrangement of $\{p_1, \ldots, p_n\}$.

[formally, $\exists \underbrace{\sigma}_{\text{permutation}} \in S_n$ s.t. $q_i = p_{\sigma(i)}$ $\forall i = 1, 2, \ldots, n$]

for instance, $30 = 2\times3\times5$, $20 = 2\times2\times5$. (unique!)

Proof — Prove first statement: existence of prime factorisation.

Proof by induction on z. $z=1$ is true ($n=0$). (or start at $z=2$ is true).

Suppose that all numbers $< z$ can be written as a product of primes.

Then z is either prime or composite. If z is prime, then statement automatically holds.

If z is not prime, then it has a non-trivial factorisation where $1 < a,b < z$ and $z = ab$.

by hypothesis, $a = p_1 \cdots p_r$, $b = q_1 \cdots q_s$ for primes p_i, q_j; $\therefore z = ab = p_1 \cdots p_r q_1 \cdots q_s$

all numbers $< z$ can be written as pdt of primes $\Rightarrow$ z can be written as pdt of primes.

hence by induction, statement is true for all z.

Prove second statement: uniqueness of factorisation.

Proof by induction on $P(n)$, where $P(n)$ denotes the statement:

"$p_1 \cdots p_n = q_1 \cdots q_m \Rightarrow n = m$ and $\{q_1, \cdots, q_m\}$ is a rearrangement of $\{p_1, \cdots, p_n\}$".

$P(1)$ is true: Suppose $p_1 = \prod^m q_j$; since p_1 is prime, $m=1$ and $p_1 = q_1$.

Suppose $P(n-1)$ holds. consider $p_1 \cdots p_n = q_1 \cdots q_m$. then we have

$p_n | p_1 \cdots p_n \Rightarrow p_n | q_1 \cdots q_m$. By corollary 9.11, $p_n |$ some q_j.

since q_j is prime, $p_n = q_j$, so by cancelling terms, we have

$z \cdot \frac{1}{p_n} = z \cdot \frac{1}{q_j} = p_1 \cdots p_{n-1} = q_1 \cdots q_{j-1} q_{j+1} \cdots q_n$.

by $P(n-1)$, $n-1 = m-1$ and $\{q_1, \cdots, q_{j-1}, q_{j+1}, \cdots, q_n\}$ is a rearrangement of $\{p_1, \cdots, p_{n-1}\}$.

multiplying p_n and q_j respectively back in, we see that $P(n-1)$ holds $\Rightarrow P(n)$ holds.

By induction, factorisation is unique// q.e.d.

This is an important result about $\mathbb{Z}$. It is also true about some other systems.

For instance, the Gaussian integers $\mathbb{Z}[i] = \{a+bi : a,b \in \mathbb{Z}\} \subseteq \mathbb{C}$; whereas Exercise 1 Q4 is an example of a number system with non-unique factorisations.

$\boxed{\text{Theorem}}$ 1.13 (Euclid's proof for infinite prime numbers).

There are infinitely many prime numbers.

Proof — by contradiction. Suppose the opposite, there are a finite number (n) of primes only.

say $p_1, \dots, p_n$ are all the primes. Then we consider $N = p_1 \cdots p_n + 1$.

N must have a prime factor, say q, by unique factorisation theorem.

since $p_1, \dots, p_n$ represent all the primes, $q = p_i$; but $p_i \nmid N = p_i(p_1 \cdots p_{i-1} p_{i+1} \cdots p_n) + 1$

this contradicts the assumption, so we conclude that there are infinitely many primes// q.e.d.

One can also think of this as an outline of a method of constructing more and more primes.

CHAPTER 2
GROUPS.

25 January 2012
Dr Mark L ROBERTS.
ALT.

$\boxed{\text{Definition}}$ 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 $*$.

where the terms have the following meanings:

- a binary operation $*$ on G is a rule assigning to any two elements $a, b \in G$ an element denoted $(a*b)$ in G. Formally, this is a function $G \times G \to G$, $(a,b) \mapsto a*b \in G$ (also called a closed binary operation to emphasise that $a*b \in G$).
- a binary operation is associative: if $a*(b*c) = (a*b)*c$ $\forall a,b,c \in G$.
- e is an identity element for G (under $*$) if $e*a = a = a*e$.
- $b \in G$ is an inverse of $a \in G$ if $a*b = e = b*a$.

If we also have $a*b = b*a$, then G is called abelian or commutative.

$\boxed{\text{Ex}}$ some examples of groups include:

(i) $G=\mathbb{Z}$ and $*$ is addition $(+)$. then this is an abelian group –

clearly associative binary operation. 0 is identity element, and inverse of z is $-z$. commutative.

(ii) $G=\mathbb{R}\setminus\{0\}=\{x\in\mathbb{R}: x\neq 0\}$ and $*$ is multiplication. this is an abelian group –

clearly associative binary operation, 1 is identity element, and inverse of r is r^{-1}. commutative.

(iii) $G=GL_n(\mathbb{R})=\{\text{invertible } n\times n \text{ matrices with real entries}\}$, nomenclature: general linear for some n.

$*$ is ordinary matrix multiplication. This is a non-abelian group (for $n>1$).

Here the product of two invertible matrices is invertible, matrix multiplication is associative, identity $= I_n$, and inverse is normal matrix inverse.

Not abelian e.g. $\begin{pmatrix}1&1\\0&1\end{pmatrix}\begin{pmatrix}1&0\\1&1\end{pmatrix}\neq\begin{pmatrix}1&0\\1&1\end{pmatrix}\begin{pmatrix}1&1\\0&1\end{pmatrix}$.

Examine each of the three properties: associativity, identity, inverse

<u>Associativity</u>: Many familiar binomial operations are associative e.g. addition, multiplication on $\mathbb{R}$ or $\mathbb{Z}$ or M_n. It is quite easy to find non-associative operations (e.g. division on $\mathbb{R}\setminus\{0\}$: $(1/2)/2 = 1/4$ but $1/(2/2)=1$.

$\boxed{\text{Ex}}$ Determine if the following operations are associative or not.

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

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

solutions:

(i) $(A*B)*C \overset{?}{=} A*(B*C) \Rightarrow (AB-BA)*C \overset{?}{=} A*(BC-CA)$

the two expressions are not formally the same, but this by itself does not show that it is not associative – we need a specific example.

counter-example: e.g. $A=\epsilon(1,1)$, $B\in(2,2)$, $C\in(1,2)$. $\left[\text{recall. } \epsilon(a,b)\epsilon(c,d)=\begin{cases}0 & \text{if } b\neq c\\ \epsilon(a,d) & \text{if } b=c\end{cases}\right]$.

$(A*B)*C \overset{?}{=} A*(B*C) \Rightarrow (AB-BA)*C \overset{?}{=} A*(BC-CA) \Rightarrow 0*C \overset{?}{=} A*-\epsilon(1,2) \Rightarrow 0 \overset{?}{=} -\epsilon(1,2)$

equality does not hold. $\Rightarrow$ not associative$_{//}$.

(ii) $a*b = (a+1)(b+1)-1$

$(a*b)*c \overset{?}{=} a*(b*c) \Rightarrow ((a+1)(b+1)-1)*c \overset{?}{=} a*((b+1)(c+1)-1) \Rightarrow (a+1)(b+1)(c+1)-1 \overset{?}{=} (a+1)(b+1)(c+1)-1$

equality holds $\Rightarrow$ associative$_{//}$.

$\boxed{\text{Lemma}}$ 2.2 If $*$ is an associative binary operation on S, then for any $x_1,\cdots,x_n\in S$, any bracketing of $x_1*x_2*\cdots*x_n$ yields the same answer.

Proof – obtained by induction on n, under associativity.

<u>Identity</u>: Consider the presence of an identity element.

$\boxed{\text{Lemma}}$ 2.3 Let $*$ be a binary operation on S, and suppose e and f are identity elements. Then $e=f$.

Proof – $e \overset{f \text{ identity}}{=} e*f \overset{e \text{ identity}}{=} f_{//}$ q.e.d.

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

$\boxed{\text{Ex}}$ Find, if it exists, the identities of

(i) $*$ on $\mathbb{R}$ by $a*b = ab+a+b$. ; 0 is identity$_{//}$ $\because$ $0*a = a = a*0$.

(ii) $*$ on $\mathbb{R}$ by $a*b = a-b$; Suppose e is an identity, $a*e=a \Rightarrow a-e=a \Rightarrow e=0$; but $0*1 = 0-1 = -1 \neq 1$.

no identity element exists$_{//}$.

<u>Inverse</u>: Consider the existence of an inverse under $*$.

$\boxed{\text{Lemma}}$ 2.4 Let G be a set and $*$ an associative binary operation on G with identity element e.

Then if both g and h are inverses of $f\in G$, then $g=h$.

Proof – we know that $f*g = e = g*f$ and $f*h = e = h*f$.

by associativity, $(g*f)*h = g*(f*h) \Rightarrow e*h = g*e \Rightarrow h=g_{//}$ q.e.d.

This means that the inverse, if it exists, is unique. In particular, within a group, each element g has a unique inverse, which is usually denoted g^{-1}.

$\boxed{\text{Lemma}}$ 2.5 For any $g\in G$, a group.

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

Proof — (i) by definition of g^{-1}, $g * g^{-1} = e = g^{-1} * g \Rightarrow g$ is the inverse of g^{-1}

$\therefore (g^{-1})^{-1} = g$ // q.e.d.

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

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

hence $(h^{-1} * g^{-1})$ is the inverse of $(g*h)$.

Analogously, $(h^{-1} * g^{-1}) * (g*h) = e$ // q.e.d.

Note the reversal of order: in general, $(g*h)^{-1} \neq g^{-1} * h^{-1}$. Also note that in $\mathbb{Z}$ under $+$, "a^{-1}" $= -a$. Do not mismatch notation

$\boxed{\text{Ex}}$ $G = \mathbb{R}$, $a*b = ab + a + b$. Which elements have inverses?

To find inverse a^{-1}, let $x = a^{-1}$ and solve $a * x = e$

thus $a + x + ax = 0 \Rightarrow x(a+1) = -a \Rightarrow x = -\frac{a}{a+1}$

for $a \neq -1$, $a * -\frac{a}{a+1} = a - \frac{a}{a+1} - \frac{a^2}{a+1} = \frac{a^2 + a - a - a^2}{a+1} = 0$.

Hence for $a \neq -1$, $a^{-1} = -\frac{a}{a+1}$; and hence, we conclude that $\mathbb{R} \setminus \{-1\}$ under $*$ forms a group.

Notation: In a general group G, we normally write gh instead of $g*h$.

$\boxed{\text{Definition}}$ 2.6 Define $g^3 = g.g.g$ (well-defined by associativity), $g^4 = gggg$ (well-defined by lemma 2.2) etc.

Define $g^{-n} = (g^{-1})^n$ and $g^0 = e$.

$\boxed{\text{Lemma}}$ 2.7 For $m, n \in \mathbb{Z}$, $g \in G$, a group,

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

i.e. usual rules for indices hold.

Proof — Formally, by induction.

$\boxed{\text{Proposition}}$ 2.8 (i) Let G be a group, $f, g, h \in G$; and $fg = fh$ then $g = h$. (left- or right-cancellation only).

(ii) Suppose G is a group with a finite number of elements $g_1, \dots, g_n$ and $g \in G$;

then the list $gg_1, gg_2, \dots, gg_n$ contains each element of G exactly once.

Proof — (i) $fg = fh \Rightarrow f^{-1}(fg) = f^{-1}(fh) \Rightarrow (ff^{-1})g = (ff^{-1})h \Rightarrow eg = eh \Rightarrow g = h$ // q.e.d.

(ii) Define function $\varphi: G \to G$ by $\varphi(g_i) = gg_i$.

by part (i), φ is injective, i.e. all gg_i are distinct. But $\{gg_1, \dots, gg_n\}$ is a set of n distinct elements contained in G

$\Rightarrow$ set of size n in $G \Rightarrow$ set is G.

27 January 2012
Dr. Mark L Roberts
Darwin LT.

$\boxed{\text{Lemma}}$ 2.9 Let X be a set and let $S(X) = \{f: X \to X : f \text{ is a bijection}\}$.

Let $\circ$ denote composition of functions [i.e. $(f \circ g)(x) = f(g(x))$]. Then $S(X)$ is a group under $\circ$.

Proof — if f and g are bijections, then so is $f \circ g$. Hence $f \circ g \in S(X)$

$\circ$ is associative $((f \circ g) \circ h)(x) = (f \circ g)[h(x)] = f(g(h(x))) = f((g \circ h)(x)) = (f \circ (g \circ h))(x)$.

$\therefore (f \circ g) \circ h = f \circ (g \circ h)$.

The function $\text{Id}: X \to X$ defined by $\text{Id}(x) = x$ $(\forall x \in X)$ is in $S(X)$. and $\text{Id} \circ f = f = f \circ \text{Id}$, Id is an identity element.

Since f is a bijection, $f \in S(X)$ and it has an inverse function $g: X \to X$ such that $f \circ g = g \circ f$, and $g \in S(X)$.

$\therefore$ every element of $S(X)$ has an inverse

$\Rightarrow S(X)$ is a group under $\circ$.

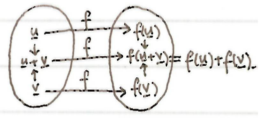
An important case is when X is a finite set, say $X = \{1, 2, \dots, n\}$. Then $S(X)$ is the permutation group, S_n.

$S(X)$ can be called the <u>automorphism group</u> of X: particularly if X is not just a set, but has some structure. e.g. X is a vector space, group, metric space.

In this case, we look at the set of bijections preserving that structure.

For example, if V is a vector space over $\mathbb{R}$, then

$\text{Aut}(V) = \{f: V \to V : f \text{ is a bijection, and } f(\underline{u}+\underline{v}) = f(\underline{u}) + f(\underline{v})\ \forall \underline{u}, \underline{v} \in V,\ f(\lambda \underline{u}) = \lambda f(\underline{u})\ \forall \lambda \in \mathbb{R}, \underline{u} \in V\}$



1202-006.

$\boxed{\text{Definition}}$ 2.11 Let n be a fixed positive integer. For $a, b \in \mathbb{Z}$ we write $a \equiv b \pmod{n}$ and say a is <u>congruent</u> to $b \pmod{n}$ if $b-a$ is a multiple of n.

e.g. $1 \equiv 7 \pmod 6$, $1015 \equiv 775 \pmod{10}$

Let $\bar{i}$ denote the set of integers congruent to $i \pmod n$

e.g. in mod 5, $\bar{2} = \{x \in \mathbb{Z}, x \equiv 2 \pmod 5\}$
$= \{x : x-2 = 5p \ \forall p \in \mathbb{Z}\}$.

If $m \in \mathbb{Z}$, by the division theorem, m can be written uniquely as

$m = nq + r$, $0 \leq r < n$. So $m \equiv r \pmod n$ and $m \in \bar{r}$.

Thus, m lies in exactly one of the n equivalence classes $\bar{0}, \bar{1}, \dots, \overline{n-1}$.

e.g. in mod 5, $12 \equiv 2 \pmod 5$, $12 \in \bar{2}$.
$12 \notin \bar{0}, \bar{1}, \bar{3}, \bar{4}$.
likewise, $-51 \in \bar{4}$.

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

Consider an illustration of the number line below, evaluated in $\mathbb{Z}_3$.

$\bar{0} = \{\dots, -6, -3, 0, 3, 6, \dots\}$, $\bar{1} = \{\dots, -5, -2, 1, 4, 7, \dots\}$, $\bar{2} = \{\dots, -4, -1, 2, 5, 8, \dots\}$.

There are 3 disjoint sets in $\mathbb{Z}_3$, which contains 3 elements $\bar{0}, \bar{1}$ and $\bar{2}$.

Let $\mathbb{Z}_n = \{\bar{0}, \bar{1}, \dots, \overline{(n-1)}\}$. We want to introduce an algebraic structure on $\mathbb{Z}_n$, i.e. addition, multiplication, etc.

We need to verify that these operations are well-defined.

$\boxed{\text{Lemma}}$ 2.12 Let $n \in \mathbb{N}$. If $a \equiv b \pmod n$ and $c \equiv d \pmod n$, then

(i) $a + c \equiv b + d \pmod n$ and (ii) $ac \equiv bd \pmod n$.

Proof -- (i)

(ii) $b - a = nr$, $d - c = ns$. Then $bd - ac = bd - bc + bc - ac = b(d-c) + c(b-a)$

$= bns - cnr = n(bs + cr) \Rightarrow n \mid bd - ac \Rightarrow bd \equiv ac \pmod n$ // q.e.d.

$\boxed{\text{Theorem}}$ 2.13 (a) $\mathbb{Z}_n$ forms a group under the operation "+" defined by $\bar{a} + \bar{b} = \overline{a+b}$.

(b) If p is prime, then $\mathbb{Z}_p^* = \mathbb{Z}_p \setminus \{\bar{0}\}$ forms a group under multiplication defined by $\bar{a} \cdot \bar{b} = \overline{ab}$.

Proof -- (a) By lemma 2.12, + is well-defined.

Then various group properties follow from properties in $\mathbb{Z}$. e.g. associativity: $(\bar{a} + \bar{b}) + \bar{c} = \overline{a+b} + \bar{c} = \overline{(a+b)+c} = \overline{a+(b+c)} = \bar{a} + \overline{b+c}$
$= \bar{a} + (\bar{b} + \bar{c})$.

identity element is $\bar{0}$,

inverse of $\bar{a}$ is $\overline{-a}$.

Hence $\mathbb{Z}_n$ forms a group under +.

(b). Again, multiplication is well-defined by 2.12.

Also, we need to check that since $\bar{a}, \bar{b} \in \mathbb{Z}_p^*$, then $\bar{a} \neq \bar{0}$, $\bar{b} \neq \bar{0}$.

i.e. $p \nmid a$ and $p \nmid b$. Since p is prime, $p \nmid ab$, i.e. $\overline{ab} \neq \bar{0}$ and $\overline{ab} \in \mathbb{Z}_p^*$

Associativity follows as for +, identity is $\bar{1}$.

To show that every element has an inverse, we fix $\bar{a} \in \mathbb{Z}_p^*$ and consider $S = \{\bar{a}, \overline{2a}, \dots, \overline{(p-1)a}\}$.

We want to show that $\bar{1} \in S$.

Each element of S is in $\mathbb{Z}_p^*$ (since $\bar{a} \in \mathbb{Z}_p^*$, $\bar{1}, \bar{2}, \dots, \overline{p-1}$ is in $\mathbb{Z}_p^*$ then $\bar{1}\bar{a}, \dots, \overline{p-1}\,\bar{a} \in \mathbb{Z}_p^*$).

No two elements of S are equal (suppose $\overline{ra} = \overline{sa}$ for some $1 \leq r, s < p$, then $\overline{(r-s)}\,\bar{a} = \bar{0}$ i.e. $p \mid (r-s)a$

$p \nmid a$ so $p \mid r-s$. but $|r-s| < p$, i.e. $r - s = 0$ and $r = s$).

So S is a set of size $(p-1)$ contained in $\mathbb{Z}_p^*$ also of size $(p-1)$. Hence $S = \mathbb{Z}_p^*$, and $\bar{1} \in S$.

As such, $\bar{a}$ has an inverse, and $\mathbb{Z}_p^*$ is a group under "×" // q.e.d.

$\boxed{\text{Ex}}$ Find $\bar{2}^{-1}$ in $\mathbb{Z}_5^*$.

$\{\bar{2}, \bar{2} \times \bar{2}, \bar{2} \times \bar{3}, \bar{2} \times \bar{4}\} = \{\bar{2}, \bar{4}, \bar{1}, \bar{3}\}$.

$\bar{2} \times \bar{3} = \bar{1} = e$, so $\bar{2}^{-1} = \bar{3}$ //.

inverses (mod p) can also be found using the Euclidean algorithm.

Recall that if $a \not\equiv 0 \pmod p$, $\exists h, k$ s.t. $ah + pk = 1 \Rightarrow ah \equiv 1 \pmod p$

thus $\bar{a}\bar{h} = \bar{1}$ in $\mathbb{Z}_p^*$, and $\bar{a}^{-1} = \bar{h}$ //.

1 February 2012
Dr Mark L ROBERTS.
9LT.

Symmetries:

The idea of symmetry is important in Maths and Mathematical Physics.

A symmetry of something is a bijective map that preserves something.

For example, we might consider isometries of the plane $\mathbb{R}^2$.

$\boxed{\text{Definition}}$ 2.14 (i) An isometry of $\mathbb{R}^2$ is a bijective map $\mathbb{R}^2 \to \mathbb{R}^2$ which preserves the distance between points.

i.e. $\forall x, y \in \mathbb{R}^2$, $|f(x) - f(y)| = |x-y|$ x —— y $\rightarrow$ $f(x)$ —— $f(y)$ distances are the same.

(ii) If T is any set of points in $\mathbb{R}^2$, $\mathrm{Sym}(T)$ is the set of isometries which send T to itself. (not necessarily each element in T to itself!).

e.g. T — T under reflection on axes, T under rotation.

$\boxed{\text{Lemma}}$ 2.15 Sym(T) forms a group under composition.

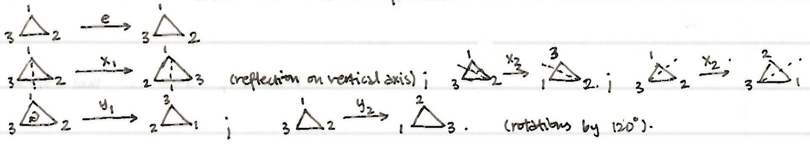
Proof — If $f, g \in \mathrm{Sym}(T)$, then $f \circ g$ is bijective an is an isometry $\Rightarrow$ $f \circ g$ sends T to T, i.e. $f \circ g \in \mathrm{Sym}(T)$.

① Composition of functions is associative. ② Identity element is $\mathrm{Id}: \mathbb{R}^2 \to \mathbb{R}^2$ where $\mathrm{Id}(x) = x$, which is an isometry sending T to $T \Rightarrow \mathrm{Id} \in \mathrm{Sym}(T)$, and ③ $f \in \mathrm{Sym}(T) \Rightarrow f$ is bijective and thus has inverse f^{-1}, which is again in $\mathrm{Sym}(T)$.

Thus, Sym(T) forms a group under $\circ$. // q.e.d.

$\boxed{\text{Ex}}$ Take $T = \Delta$, an equilateral triangle in $\mathbb{R}^2$. calculate Sym(T).

Label vertices: there are various obvious symmetries:



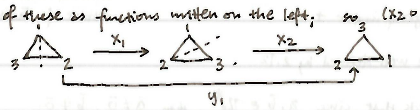
These are all clearly elements of Sym(T). Could there be any more? No.

Any $f \in \mathrm{Sym}(T)$ is determined by where it sends the vertices, and has to send a vertex to the position of another vertex.

Total possible elements in group $\mathrm{Sym}(T) = 3! = 6$. $\therefore \mathrm{Sym}(T) = \{e, x_1, x_2, x_3, y_1, y_2\}$ // $\rightarrow$ forms a group under $\circ$.

To specify the group, we have to state how these elements compose with each other. For instance, what does $x_2 \circ x_1$ mean?

We think of these as functions written on the left; so $(x_2 \circ x_1)(r) = x_2(x_1(r))$.



by comparing end-result to the list above, we see that

$x_2 \circ x_1 = y_1$.

We can conveniently write down all the group information in a group table. This does specify the group entirely, but not very efficiently.

$\circ$	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

i.e. $x_1 = x_2 \, y_1$

A more efficient way is by means of generators and relations.

Write $x = x_1$, $y = y_1$. Then note that

$y_2 = y^2$, $x_2 = x^2 y$. $x_3 = xy$

So another way of writing the group is $\mathrm{Sym}(T) = \{e, \underset{y_1}{y}, \underset{y_2}{y^2}, \underset{x_1}{x}, \underset{x_3}{xy}, \underset{x_2}{xy^2}\}$.

To specify the group table, we need to give enough rules (relations) to combine any two elements of this form and obtain the answer in the same form.

$x^2 = e$, $y^3 = e$. However, these are insufficient:

for instance, what is $(xy)(xy)$? what happens when y preceeds x?

$yx = y_1 x_1 = x_2 = xy^2$. Now these are sufficient. e.g. $(xy)(xy^2) = x(yx)y^2 = x(xy^2)y^2$

$= x^2 y^4 = ey = y$.

we write $\mathrm{Sym}(T) = \langle \underbrace{x, y}_{\text{generators}} : \underbrace{x^2 = e, y^3 = e, yx = xy^2}_{\text{relations}} \rangle$.

The order of an element and cyclic groups:

$\boxed{\text{Definition}}$ 2.16 (i) The order of a group G, denoted $|G|$, is the number of elements in G.

A group is called finite if $|G| < \infty$ and infinite if $|G| = \infty$.

(ii) The order of an element $g \in G$ is the least positive integer n s.t. $g^n = e$ (or ∞ if no such n exists).

This is denoted $\sigma(g)$.

$\boxed{\text{Ex}}$ For $\mathbb{R}$ under $+$, $\sigma(2) = \infty$.

For $\mathbb{R}\setminus\{0\}$ under $\times$, $\sigma(-1) = 2$ $\because$ $(-1)^2 = 1$.

For $\mathbb{C}\setminus\{0\}$ under $\times$, $\sigma(i) = 4$ $\because$ $(-1)^n = 1 = e$ for $n_{min} = 4$.

For $\mathbb{Z}_6$ under $+$, $\sigma(\bar{1}) = 6$, $\sigma(\bar{2}) = 3$, $\sigma(\bar{3}) = 2$, $\sigma(\bar{4}) = 3$, $\sigma(\bar{5}) = 6$.

For $\mathbb{Z}_5^*$ under $\times$, $\sigma(\bar{1}) = 1$, $\sigma(\bar{2}) = 4$, $\sigma(\bar{3}) = 4$, $\sigma(\bar{4}) = 2$.

$\boxed{\text{Lemma}}$ 2.17 Let G be a group and $g \in G$ with $\sigma(g) = n$. Then

(i) $g^m = e \Rightarrow n | m$.

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

Proof – (i) we have $g^n = e$ and $g^x \neq e$ for any $1 \leq x < n$.

suppose $g^m = e$; by the division theorem, $\exists q, r \in \mathbb{Z}$ s.t. $m = nq + r$ with $0 \leq r < n$

then $g^r = g^{m-nq} = g^m (g^n)^{-q} = e \cdot e^{-q} = e$

but $0 \leq r < n$, so $r = 0$ i.e. $m = nq$ and $n | m$ // q.e.d.

(ii) as before, $\forall m \in \mathbb{Z}$, $g^m = g^{nq+r}$ for some $0 \leq r < n$.

then $g^m = (g^n)^q g^r = e^q g^r = g^r$

$\therefore$ any power of g is of form g^r $(0 \leq r < n)$.

to prove uniqueness – if $g^r = g^s$ $(0 \leq r \leq s < n)$, then $g^{s-r} = e$ and $0 \leq s-r < n$.

hence, as $n = \sigma(g)$, $s - r = 0 \Rightarrow s = r$. // q.e.d.

e.g. if $\sigma(g) = 3$, then $\dots, \underset{e}{g^{-3}}, \underset{g}{g^{-2}}, \underset{g^2}{g^{-1}}, e, g, g^2, \underset{e}{g^3}, \underset{g}{g^4}, \underset{g^2}{g^5}, \underset{e}{g^6}, \dots$ (repeats periodically).

We now deal with the problem of classifying groups.

This is a very large area – we look at a small part of the theory. First the simplest class of groups – cyclic groups.

$\boxed{\text{Definition}}$ 2.18 Let G be a group and $g \in G$. Then define $\langle g \rangle = \{g^n : n \in \mathbb{Z}\}$.

If $G = \langle g \rangle$, then G is said to be generated by g; if G is generated by some element $g \in G$, G is cyclic.

For instance, $\mathbb{Z}$ under $+$ is cyclic because $\mathbb{Z} = \langle 1 \rangle$. However, $\mathbb{Z} \neq \langle 2 \rangle$. (incl. neg. numbers).

$\langle 1 \rangle = \{\dots, -3, -2, -1, 0, 1, 2, 3, 4, \dots\}$

$\mathbb{Z}_5^*$ is cyclic, e.g. $\mathbb{Z}_5^* = \langle \bar{2} \rangle$. $\bar{2}^0 = \bar{1}$, $\bar{2}^1 = \bar{2}$, $\bar{2}^2 = \bar{4}$, $\bar{2}^3 = \bar{3}$ (however $\mathbb{Z}_5^* \neq \langle \bar{4} \rangle = \{\bar{1}, \bar{4}\}$).

However, we see that S_3 is not cyclic because none of its elements in it are generators.

$\langle e \rangle = \{e\}$, $\langle (1\ 2) \rangle = \{e, (1\ 2)\}$, $\langle (1\ 3) \rangle = \{e, (1\ 3)\}$, $\langle (2\ 3) \rangle = \{e, (2\ 3)\}$,

$\langle (1\ 2\ 3) \rangle = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$, $\langle (1\ 3\ 2) \rangle = \{e, (1\ 3\ 2), (1\ 2\ 3)\}$. $\Rightarrow$ none have order 6.

$\boxed{\text{Lemma}}$ 2.19 Let G be a finite group of order n. Then G is cyclic $\Leftrightarrow$ G contains an element of order n. 3 February 2012. Dr Mark L ROBERTS. Darwin LT.

Proof – (forward relation) Suppose G is cyclic, say $G = \langle g \rangle$. Then $|\langle g \rangle| = |G| = n$. But $\langle g \rangle = \{e, g, \dots, g^{m-1}\}$ where $\sigma(g) = m$.

Hence $|\langle g \rangle| = \sigma(g)$. $\therefore \sigma(g) = n$

(backward relation). Suppose g has order n. Then $|\langle g \rangle| = n$. $\langle g \rangle \subseteq G$ and $|G| = n$ $\therefore$ $\langle g \rangle = G$, and G is cyclic, q.e.d.

$\boxed{\text{Definition}}$ 2.20 Let G be a cyclic group, say $G = \langle g \rangle$. Then we note that

(i) if $o(g) = n < \infty$, then $G = \{e, g, \dots, g^{n-1}\}$ where $g^n = e$;

and G is called the cyclic group of order n, denoted C_n.

(ii) if $o(g) = \infty$, then $G = \{\dots, g^{-2}, g^{-1}, e, g, g^2, \dots\}$,

and G is called the infinite cyclic group, denoted C_∞.

observe that C_∞ is isomorphic to $\mathbb{Z}$ under $+$.

$$\begin{aligned} g &\longleftrightarrow 1 & & \\ g * g = g^2 &\longleftrightarrow 2 = 1+1 & e &\longleftrightarrow 0 \\ g^3 &\longleftrightarrow 3 & g^{-1} &\longleftrightarrow -1. \\ g^n &\longleftrightarrow n & & \end{aligned}$$

we write $C_\infty \cong \mathbb{Z}$.

Another example of isomorphism is where $G = \{e, a, a^2\}$ where $a^3 = e$ and $H = \{e, b, b^2\}$ where $b^3 = e$; then $G \cong H$.

is $\mathbb{Z}_5^*$ under $\times$ $\cong C_4$ $\quad \mathbb{Z}_5^* = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}\}$ and $C_4 = \{e, g, g^2, g^3\}$.

$\bar{2} \longleftrightarrow g$, $\quad \bar{4} = \bar{2}^2 \longleftrightarrow g^2$, $\quad \bar{3} = \bar{8} = \bar{2}^3 \longleftrightarrow g^3$ $\quad \bar{1} \longleftrightarrow e$.

Isomorphism will be further explained in later courses.

SUBGROUPS.

Definition 2.21 Let G be a group and H be a subset of G. Then H is a subgroup of G, denoted $H \leq G$,

if H itself forms a group under the same operations as G.

(i.e. same binary operation, which implies same identity and inverses.)

A more convenient, and equivalent, form of the definition is the following.

Lemma Let $H \subseteq G$. then H is a subgroup of G iff

(i) $e \in H$

(ii) $g, h \in H \Rightarrow gh \in H$.

(iii) $g \in H \Rightarrow g^{-1} \in H$.

(forward relation)

Proof — Suppose H is a subgroup of G, then H has an identity element f. so $\forall h \in H$, $f * h = h$.

Hence $f = e$, the identity of G. so $e \in H$. (ii) and (iii) are automatic.

(backward relation).

The operation on H, as in G, is associative. H is well-defined by (ii). By (i) and (iii), H has an identity element and inverses. $\therefore$ H forms a group. // q.e.d.

Examples: • $2\mathbb{Z}$ = {even integers} is a subgroup of $\mathbb{Z}$ under $+$.

(i) $0 \in 2\mathbb{Z}$ (ii) suppose $g, h \in 2\mathbb{Z}$, $g = 2a$, $h = 2b$ for some $a, b \in \mathbb{Z}$; $g+h = 2a + 2b = 2(a+b)$ and $a+b \in \mathbb{Z}$ so $g+h \in \mathbb{Z}$.

(iii) if $g \in 2\mathbb{Z}$, say $g = 2a$, then $g^{-1} = -g = -2a = 2(-a) \in 2\mathbb{Z}$.

The three conditions for a subgroup are equivalent to: $H \neq \emptyset$ and $h, k \in H \Rightarrow h^{-1}k \in H$.

EX (i) let $G = \mathbb{Z}$ under $+$.

$H = \{x \in \mathbb{Z},\ x \equiv 0 \pmod 3\}$, $\quad K = \{x \in \mathbb{Z},\ x \equiv 1 \pmod 3\}$, $\quad J = \{x \in \mathbb{Z},\ x \geq 0\}$.

Which of H, K, J are subgroups?

H is a subgroup: (i) $0 \in H$, (ii) $f, g \in H \Rightarrow f = 3a, g = 3b \Rightarrow fg = 3a + 3b = 3(a+b) \equiv 0 \pmod 3$, (iii) $g \in H \Rightarrow g = 3a$, $g^{-1} = -g = -3a \equiv 0 \pmod 3$

K is not a subgroup: (i) $0 \notin K$; J is not a subgroup: (iii) $g \in H \Rightarrow g^{-1} = -g \leq 0 \notin H$.

(ii) Find all subgroups of C_6.

$\{e\}$, $\{e, g^2, g^4\}$, $\{e, g^3\}$, $\{e, g, g^2, g^3, g^4, g^5\}$. $\quad$ e has to lie in H!

$e \in H$. If $g \in H$, then $g^2 \in H \dots$ $H = G$.

suppose $g \notin H$. if $g^2 \in H$, then $g^4 \in H$ so $\{e, g^2, g^4\} \leq H$. if then $g^3 \in H$, then $g = g^3 (g^2)^{-1} \in H$ ※. thus $g^3 \notin H$; $g^5 \notin H$.

so $H = \{e, g^2, g^4\}$. Next suppose $g^2 \notin H$, if $g^3 \in H$, we similarly get $H = \{e, g^3\}$.

if $g^3 \notin H$, then $g^4 \notin H$ $\because$ $(g^4)^{-1} = g^2 \notin H$. likewise $g^5 \notin H$, so $H = \{e\}$.

A more systematic way of doing part (ii) is to look at $\min\{m > 0 \text{ s.t. } g^m \in H\}$.

this helps us find subgroups of C_n.

8 February 2011
Dr Mark L ROBERTS.
GLT.

$\boxed{\text{Theorem}}$ 2.23 let A_n denote the set of even permutations in S_n.
Then A_n forms a subgroup of S_n, called the alternating group,
and $|A_n| = \frac{1}{2}|S_n| = \frac{1}{2}n!$

Recall that if $\sigma \in S_n$, we can write $\sigma = \tau_1 \cdots \tau_m$ where τ_i are transpositions.
If $\sigma = \rho_1 \cdots \rho_n$, where $\rho_1, \cdots, \rho_n$ are transpositions, then m even $\Leftrightarrow$ n even; m odd $\Leftrightarrow$ n odd.
So if $\sigma = \tau_1 \cdots \tau_m$ (m even), we call σ an <u>even permutation</u>.

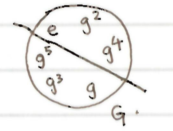
Proof (of Thm 2.23) — e is even, so $e \in A_n$.
Suppose $g, h \in A_n$. Then $g = \tau_1 \cdots \tau_m$, $h = \rho_1 \cdots \rho_n$ for transpositions τ_i, ρ_j where m, n are even.
Then $gh = \tau_1 \cdots \tau_m \rho_1 \cdots \rho_n$ and $m+n$ is even i.e. $gh \in A_n$.
Also, $g^{-1} = (\tau_1 \cdots \tau_m)^{-1} = \tau_m^{-1} \cdots \tau_1^{-1} = \tau_m \cdots \tau_1$ (since inverse of a transposition is itself).
Hence $A_n \leq S_n$. Define $\varphi: A_n \to S - A_n$ by $\varphi(\sigma) = (1\ 2)\sigma$
Note: $(1\ 2)\sigma$ is odd. φ is injective. $[\varphi(\sigma_1) = \varphi(\sigma_2) \Rightarrow (1\ 2)\sigma_1 = (1\ 2)\sigma_2 \Rightarrow \sigma_1 = \sigma_2]$.
φ is surjective as well; if $\sigma \in S - A_n$, where $(1\ 2)\sigma \in A_n$ and $\varphi((1\ 2)\sigma) = (1\ 2)(1\ 2)\sigma = \sigma$.
$\therefore$ φ is bijective: $|A_n| = |S_n| - |A_n| \Rightarrow |A_n| = \frac{1}{2}|S_n| = \frac{1}{2}n!$ // q.e.d.

For example, consider $S_3 = \{e, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$.
Then $A_3 = \{e, (1\ 2\ 3), (1\ 3\ 2)\}$ (even permutations) $\xrightarrow{\varphi,\ \text{bijection}}$ $S_3 - A_3 = \{(1\ 2), (1\ 3), (1\ 3\ 2)\}$.
Observe here that if $\varphi: A_n \to S_n A_n$, $\varphi(\sigma) = (1\ 2)\sigma$, then $\varphi(e) = (1\ 2)$, $\varphi((1\ 2\ 3)) = (1\ 2)(1\ 2\ 3) = (2\ 3)$, $\varphi((1\ 3\ 2)) = (1\ 3)$.

LAGRANGE'S THEOREM.

To analyse a group, we need to know about its subgroups. (e.g. a simple group is one that contains no subgroups but the trivial one).
This is a hard question in general, but the theorem gives some straightforward information about subgroups of finite groups.

$\boxed{\text{Theorem}}$ 2.24 (Lagrange's Theorem)
Let G be a finite group, and H a subgroup. Then $|H|$ divides $|G|$. [For instance, $G = C_6$, $H = \{e, g^2, g^4\} \leq G \Rightarrow 3|6$]
Proof — stage 1: Definition of cosets.
For any $g \in G$, define the <u>coset</u> $Hg = \{hg : h \in H\}$. [e.g. in C_6, $He = \{he : h \in H\} = \{ee, g^2e, g^4e\} = \{e, g^2, g^4\}$;
$Hg = \{hg : h \in H\} = \{eg, g^2g, g^4g\} = \{g, g^3, g^5\}$;
$Hg^2 = \{hg^2 : h \in H\} = \{eg^2, g^2g^2, g^4g^2\} = \{g^2, g^4, e\}$ ...
stage 2: G is a union of cosets.
Since $e \in H$, $g = eg \in Hg$ ← coset.
Hence, $G = \bigcup_{g \in G} Hg$.
stage 3: cosets are either the same or disjoint:
claim — either $Hg = Hg'$ or $Hg \cap Hg' = \phi$.
So suppose $Hg \cap Hg' \neq \phi$, say $x \in Hg \cap Hg'$, then for some $h_1 \in H$, $h_2 \in H$,
$x = h_1 g = h_2 g'$. Then $g' = h_2^{-1} h_1 g$. Hence for any $h \in H$, $hg' = h h_2^{-1} h_1 g$.
since H is a subgroup, $h h^{-1} h_1 \in H$ due to closure; hence $hg' \in Hg$ $\forall h \in H$, thus $Hg' \in Hg$.
by that same symmetric argument, $Hg \in Hg' \Rightarrow Hg = Hg'$. (proven).



stage 4: G is the <u>disjoint union</u> of some of the cosets.
i.e. $\exists\, g_1, \cdots, g_r$ s.t. $G = Hg_1 \cup Hg_2 \cup \cdots \cup Hg_r$ where $Hg_i \cap Hg_j = \phi$.
Since $G = \bigcup_{g \in G} Hg$, we can leave out repetitions; and since we know cosets are disjoint, G is a disjoint union.
stage 5: All cosets are the same size.
injective: $hg = h'g \Rightarrow hgg^{-1} = h'gg^{-1} \Rightarrow h = h'$
surjective: definition of Hg.
Claim that for any $g \in G$, $|Hg| = |H|$ (define $\varphi: H \to Hg$ by $\varphi(h) = hg$ — φ is bijective.)
stage 6: Result — $|G| = |Hg_1 \cup \cdots \cup Hg_r| \overset{\text{disjoint}}{=} |Hg_1| + \cdots + |Hg_r| = |H| + \cdots + |H| = r|H| \Rightarrow |H| \,\big|\, |G|$ // q.e.d.

For instance, if $|G| = 7$, and $H \leq G$, then $|H|$ divides 7, i.e. $|H| = 1$ or 7

thus G has no non-trivial subgroups, i.e. G is cyclic.

$\boxed{\text{Corollary}}$ 2.25 Let G be a finite group, $g \in G$. Then $o(g) \mid |G|$.

Proof — let $H = \langle g \rangle = \{e, g, g^2, \dots\} \leq G$. and $|H| = o(g)$. (by 2.17).

By Lagrange's Theorem, $o(g) = |H|$ divides $|G|$, q.e.d.

For example, if $|G| = 6$, the only possible orders of elements are 1, 2, 3 or 6.

$\boxed{\text{Corollary}}$ 2.26 Let G be a group of order p, where p is prime. Then $G \cong C_p$.

Proof — Let $e \neq g \in G$. Then $o(g) \neq 1$, $o(g) \mid p$ (by 2.25).

thus $o(g) = p$ and $|\langle g \rangle| = p$. and hence, $\langle g \rangle = G$, q.e.d.

Thus, groups of prime order are easy to classify, as there is exactly one group for each p, namely C_p.

Whereas on the other hand, groups of composite order are more complicated.

Yet, we see that, for instance, it is now quite easy to work out the subgroups of S_3.

$|S_3| = 6$, $\therefore$ if $H \leq S_3$, $|H| = 1, 2, 3$ or 6.

- $|H| = 1 \Rightarrow H = \{e\}$; • $|H| = 6 \Rightarrow H = S_3$.
- if $|H| = 2$, H is a group of order 2: so $H \cong C_2$, i.e. $H = \langle g \rangle$ where $o(g) = 2$. $\Rightarrow g = (1\ 2)$ or $(1\ 3)$ or $(2\ 3)$
- Similarly if $|H| = 3$, H is a group of order 3: so $H \cong C_3$, i.e. $H = \langle g \rangle$ where $o(g) = 3$. $\Rightarrow g = (1\ 2\ 3)$ or $(1\ 3\ 2)$.

so, $H = \{e\}$, $\{e, (1\ 2)\}$, $\{e, (1\ 3)\}$, $\{e, (2\ 3)\}$, $\{e, (1\ 2\ 3), (1\ 3\ 2)\}$, $\{e, (1\ 2), (1\ 3), (2\ 3), (1\ 2\ 3), (1\ 3\ 2)\}$.

and S_3 has exactly 6 subgroups.

Note: S_3 has no element of order 6 $\Rightarrow$ corollary 2.25 does not work in converse case.

Recall that $\mathbb{Z}_p^\times$ (p is prime) is the set of non-zero integers (mod p) under multiplication.

$\mathbb{Z}_p^\times$ is a group (e.g. $\mathbb{Z}_5^\times = \{\bar{1}, \bar{2}, \bar{3}, \bar{4}\}$.

$\boxed{\text{Theorem}}$ 2.27 (Fermat's Little theorem) — by Pierre de Fermat

Let p be a prime, and $a \not\equiv 0 \pmod p$, then $a^{p-1} \equiv 1 \pmod p$.

Proof — $\bar{a} \in \mathbb{Z}_p^\times$ $\because a \not\equiv 0 \pmod p$. By 2.25, $o(\bar{a}) \mid |\mathbb{Z}_p^\times| = p-1$.

so $\exists k$ s.t. $p-1 = k\, o(\bar{a})$ $\Rightarrow$ $\bar{a}^{p-1} = [\bar{a}^{o(\bar{a})}]^k = \bar{1}^k = \bar{1}$. (by 2.17).

i.e. $a^{p-1} \equiv 1 \pmod p$, q.e.d.

For instance, find $3^{2203} \pmod{23}$:

by Fermat's little theorem, $3^{22} \equiv 1$ $\Rightarrow$ $3^{2200} \equiv 1$ $\Rightarrow$ $3^{2203} \equiv 3^3 \equiv 27 \equiv 4 \pmod{23}$.

10 February 2012
Dr Mark L ROBERTS
Darwin LT.

CHINESE REMAINDER THEOREM.

This theorem tells us about solving simultaneous congruences.

e.g. Find x such that $x \equiv 7 \pmod{11}$ and $x \equiv 10 \pmod{13}$.

then $x = 11m + 7 = 13n + 10$; $m, n \in \mathbb{Z}$. $\Rightarrow$ $11m - 13n = 3$.

Since 11, 13 are coprime, by the h,k-lemma, $\exists h, k$ s.t. $11h + 13k = 1$.

By Euclidean algorithm, we find that $h = 6$, $k = -5$ i.e. $11(6) - 13(5) = 1$ $\Rightarrow$ $11(18) - 13(15) = 3$ $\Rightarrow$ $m = 18$.

$x = 11m + 7 = 11(18) + 7 = 205$. $\Rightarrow$ 205 is a solution.

Suppose x_1, x_2 are both solutions, then $x_1 \equiv x_2 \equiv 7 \pmod{11}$, $x_1 \equiv x_2 \equiv 10 \pmod{13}$ $\Rightarrow$ $x_1 - x_2 \equiv 0 \pmod{11} \equiv 0 \pmod{13}$.

11 and 13 $\mid x_1 - x_2$ $\xRightarrow{\gcd(11,13)=1}$ $(11)(13) \mid x_1 - x_2$ $\Rightarrow$ $x_1 \equiv x_2 \pmod{11 \cdot 13}$. Solution is $x \equiv 205 \pmod{143} \equiv 62 \pmod{143}$, solution is unique equivalence class.

1102-012

$\boxed{\text{Theorem}}$ (Chinese Remainder Theorem).

Let m, n be coprime integers. Then there exists a solution c to $x \equiv a \pmod m$, $x \equiv b \pmod n$.

The complete set of solutions is $\{x : x \equiv c \pmod{mn}\}$, i.e. the solution is unique $\pmod{mn}$.

Proof — By h,k-lemma, $\exists\, h,k$ s.t. $mh + nk = 1$

Then nk is a solution to $x \equiv 1 \pmod m$ and $x \equiv 0 \pmod n$

and mh is a solution to $x \equiv 0 \pmod m$ and $x \equiv 1 \pmod n$

Hence if we let $c = ank + bmh$, $a, b \in \mathbb{Z}$.

Then $c = ank + bmh \equiv ank \equiv a(1) \equiv a \pmod m$ and $c = ank + bmh \equiv bmh \equiv b(1) \equiv b \pmod n$.

$\therefore$ $c = ank + bmh$ is a solution to $x \equiv a \pmod m$, $x \equiv b \pmod n$ // q.e.d.

Now if $x \equiv c \pmod{mn}$, then $x \equiv c \pmod m$ and $x \equiv c \pmod n$; so x is also a solution.

If x is any solution, $x \equiv c \pmod m$ and $x \equiv c \pmod n$; so $x \equiv c \pmod{mn}$.

$\boxed{\text{Ex}}$ What is $2^{66} \pmod{77}$?

Consider congruences mod 7 and mod 11. By Fermat's little theorem, $2^6 \equiv 1 \pmod 7$ $\Rightarrow$ $2^{66} = (2^6)^{11} \equiv 1 \pmod 7$.

Also by Fermat's little theorem, $2^{10} \equiv 1 \pmod{11}$ $\Rightarrow$ $2^{66} = (2^{10})^6 \cdot 2^6 \equiv 1 \cdot 2^6 \equiv 64 \equiv 9 \pmod{11}$.

We need to solve $x \equiv 1 \pmod 7$, $x \equiv 9 \pmod{11}$. We find h, k s.t. $11h + 7k = 1$. By inspection, $h = 2$ and $k = -3$.

Applying the Chinese Remainder theorem, solution is $(11)(2)\cdot(1) + (7)(-3)(9) = 22 - 189 = -167 \equiv 64 \pmod{77}$.

Hence, $2^{66} \equiv 64 \pmod{77}$.

This theorem generalises for more than 2 simultaneous congruences i.e. $x \equiv a_i \pmod{n_i}$, $i = 1, 2, \cdots, m$

with each n_i, n_j coprime. This uses the fact that n_1 and $n_2 \cdots n_m$ are coprime.

Note: The theorem does not hold for non-coprime congruences.

e.g. $x \equiv 0 \pmod 2$, $x \equiv 1 \pmod 4$ has no solution.

22-February 2011
Dr Mark L ROBERTS.
CLT.

CHAPTER 3
DETERMINANTS.

$\boxed{\text{Definition}}$ 3.1 Let A be an $n \times n$ matrix (with entries a_{ij}). Then $\det A = \sum_{\sigma \in S_n} (\operatorname{sgn} \sigma)\, a_{1,\sigma(1)} \cdot a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$;

when S_n is the group of permutations of $\{1, \dots, n\}$, $\operatorname{sgn}(\sigma) = \begin{cases} +1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}$.

Note: Formula means take each possible $\sigma \in S_n$; take the product $a_{1,\sigma(1)} \cdots a_{n,\sigma(n)}$,

multiply by ± 1, and sum up terms.

For instance, in the 2x2 case, $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$, $S_2 = \{ id, (1\ 2) \}$.

$\operatorname{sgn}(id) = +1$, $\operatorname{sgn}(1\ 2) = -1$. So, $\det A = \sum_{\sigma \in S_2} (\operatorname{sgn} \sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} = \operatorname{sgn}(id)\, a_{1,id(1)} a_{2,id(2)} + \operatorname{sgn}(1\ 2)\, a_{1,(1\ 2)(1)}\, a_{2,(1\ 2)(2)}$

$= (+1) a_{11} a_{22} + (-1) a_{12} a_{21} = a_{11} a_{22} - a_{12} a_{21}$.

$\boxed{\text{Proposition}}$ 3.2 Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Then

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

(ii) A is invertible $\iff$ $\det A \neq 0$. In this case, $A^{-1} = \frac{1}{\det A} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$.

(iii) Let $L_A : \mathbb{R}^2 \to \mathbb{R}^2$ be the linear map defined by $L_A(\underline{v}) = A\underline{v}$. Then for any shape S in the plane,

Area $[L_A(S)]$ = Area$(S) \cdot |\det A|$

(iv) If B is another 2x2 matrix, then $\det(AB) = \det A \cdot \det B$

Proof — (i) by definition.

(ii) suppose A has inverse $\begin{pmatrix} x & y \\ z & t \end{pmatrix}$, then $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x & y \\ z & t \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I_2$.

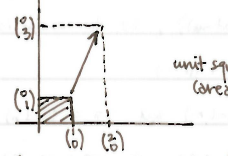
$\therefore$ $\left.\begin{array}{l} ax + bz = 1 \\ ay + bt = 0 \\ cx + dz = 0 \\ cy + dt = 1 \end{array}\right\}$ $\begin{array}{l} cax + cbz = c \\ acx + adz = 0 \\ \Rightarrow (ad - bc) z = -c. \end{array}$ Similarly, $\begin{array}{l} (ad-bc) y = -b \\ (ad-bc) x = d \\ (ad-bc) t = a \end{array}$ Hence if $ad - bc \neq 0$, we get $\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = A^{-1}$.

If $ad - bc = 0$, then $a = b = c = d = 0$; not invertible.

for example, $A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}$, $\det A = 0$, so A is not invertible

$A = \begin{pmatrix} 1 & 2 \\ 3 & 1 \end{pmatrix}$, $\det A = -5 \neq 0$, so A is invertible, $A^{-1} = -\frac{1}{5}\begin{pmatrix} 1 & -2 \\ -3 & 1 \end{pmatrix} = \begin{pmatrix} -1/5 & 2/5 \\ 3/5 & -1/5 \end{pmatrix}$.

(iii) for instance, let $A = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}$. Then $L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2x \\ 3y \end{pmatrix}$. $\Rightarrow$ scaling by factor of 2 in x-direction, 3 in y-direction.

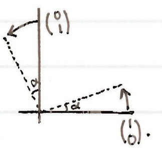


unit square (area 1) $\longrightarrow$ 2x3 rectangle (area 6).

L_A multiplies areas by $6 = \det A$.

this applies to other shapes in $\mathbb{R}^2$ as well, as each point is changed out.

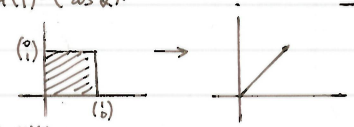
similarly, for instance, let $A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$. Then $L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x\cos\alpha - y\sin\alpha \\ x\sin\alpha + y\cos\alpha \end{pmatrix}$.



$\det(A) = \cos^2\alpha - (-\sin^2\alpha) = 1$. for instance, $L_A\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} \cos\alpha \\ \sin\alpha \end{pmatrix}$, $L_A\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -\sin\alpha \\ \cos\alpha \end{pmatrix}$.

L_A preserves area upon counter-clockwise rotation by α.

again by example, if $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, $L_A\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x+y \\ x+y \end{pmatrix}$.



L_A multiplies areas by 0 (everything is squashed onto a line), $\det A = (1)(1) - (1)(1) = 0$.

finally for instance, $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ gives a reflection about y-axis, $\det A = -1$; area is unchanged.

for proof of general case, for unit square, see Ex5 Q4.

(iv) for proof, see Ex5 Q3(a) (by direct calculation).

This suggests two reasons why det is important: it captures in a single number a lot of information about a matrix.

- first, whether or not matrix is invertible
- second, $\det A$ is a "scale factor" related to the linear transformation L_A (significant in multivariable calculus, e.g. Jacobians, Wronskians).

Return to point (iv): if we use interpretation of $\det A$ as scale factor, this is clear:

$L_A L_B = L_{AB}$ $\longleftarrow$ multiplies area by $\det A \det B$. Thus, $\det A \det B = \det AB$.

multiplies area by $\det A$, $\det B$.

We next consider the case for a 3x3 matrix:

$\boxed{\text{Proposition}}$ 3.3 Let $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Then, $\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31}$.

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

$S_3 = \{ id, (1\ 2\ 3), (1\ 3\ 2), (1\ 2), (1\ 3), (2\ 3) \}$.
sgn $+1$ $+1$ $+1$ -1 -1 -1.

Entering into the formula,

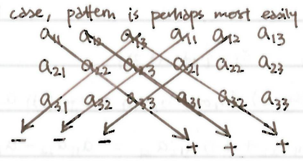
Thus, $\det A = \sum_{\sigma \in S_3} (\operatorname{sgn}\sigma)\, a_{1,\sigma(1)} a_{2,\sigma(2)} a_{3,\sigma(3)}$.

$$= (\operatorname{sgn} id)\, a_{1,id(1)} a_{2,id(2)} a_{3,id(3)} + (\operatorname{sgn}(123))\, a_{1,(123)(1)} a_{2,(123)(2)} a_{3,(123)(3)} + (\operatorname{sgn}(132))\, a_{1,(132)(1)} a_{2,(132)(2)} a_{3,(132)(3)}$$

$$+ (\operatorname{sgn}(12))\, a_{1,(12)(1)} a_{2,(12)(2)} a_{3,(12)(3)} + (\operatorname{sgn}(13))\, a_{1,(13)(1)} a_{2,(13)(2)} a_{3,(13)(3)} + (\operatorname{sgn}(23))\, a_{1,(23)(1)} a_{2,(23)(2)} a_{3,(23)(3)}$$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32}.$$

For 3x3 case, pattern is perhaps most easily remembered by.



$\boxed{\text{Ex}}$. find $\det\begin{pmatrix} 1 & 2 & -1 \\ -2 & 1 & 1 \\ 3 & -2 & 1 \end{pmatrix}$.

$\det\begin{pmatrix} 1 & 2 & -1 \\ -2 & 1 & 1 \\ 3 & -2 & 1 \end{pmatrix} = 1 + 6 + (-4) - (-3) - (-2) - (-4) = 3 - (-9) = 12$ //

Properties of nxn determinants:

Evaluating an nxn determinant from the definition involves adding up $n!$ terms, each the product of n terms -- calculationally insane.

For this reason, we develop alternative methods of finding determinants, and to prove properties of them.

The first result is that transposing a matrix does not change the determinant: i.e. $\det A = \det A^T$.

e.g. in 2x2 case $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $A^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$; then $\det A = ad - bc$, $\det(A^T) = ad - bc$.

For general case, recall that $(A^T)_{ij} = A_{ji}$; then.

$\boxed{\text{Proposition}}$ 3.4 Let A be $n\times n$. Then $\det A^T = \det A$.

Proof – let $B = A^T$. $A = (a_{ij})$, $B = (b_{ij})$; $b_{ij} = a_{ji}$.

$\det(A^T) = \det B = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{\sigma(1),1}\, a_{\sigma(2),2} \cdots\cdots a_{\sigma(n),n}$

Write $\mu = \sigma^{-1}$, and as σ ranges over S_n, then so does μ. i.e. μ is a rearrangement of S_n.

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

$= \sum_{\mu\in S_n} (\operatorname{sgn}\mu)\, a_{\mu^{-1}(1),1} \cdots a_{\mu^{-1}(n),n}$ $(\because \operatorname{sgn}\mu = \operatorname{sgn}\mu^{-1})$

Observe that $a_{\mu^{-1}(1),1} \cdots\cdots a_{\mu^{-1}(n),n} = a_{1,\mu(1)} \cdots\cdots a_{n,\mu(n)}$, because if we suppose $\mu(1) = r$,

first term on RHS is $a_{1,r}$. And one term on LHS $a_{\mu^{-1}(r),r} = a_{1,r}$; and etc.

finally, this gives us $\det(A^T) = \sum_{\mu\in S_n} (\operatorname{sgn}\mu)\, a_{1,\mu(1)} \cdots a_{n,\mu(n)} = \det A$. // q.e.d.

Interestingly, this means that any result about rows implies the corresponding result about columns (e.g. Theorem 3.6, later).

$\boxed{\text{Proposition}}$ 3.5 Let A be a lower triangular matrix (i.e. $a_{ij} = 0$ for $j > i$). $A = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$.

Then $\det A = a_{11} a_{22} \cdots\cdots a_{nn}$

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

one term is $\sigma = \mathrm{Id}$, which gives $a_{11}a_{22}\cdots a_{nn}$ contributing factor.

For the other terms, $\sigma \neq \mathrm{Id}$, and suppose $a_{1,\sigma(1)} \cdots a_{n,\sigma(n)} \neq 0$, then $a_{1,\sigma(1)}, a_{2,\sigma(2)}, \cdots, a_{n,\sigma(n)} \neq 0$.
By architecture of lower triangular matrix,
then if $\sigma(1) > 1$, $a_{1,\sigma(1)} = 0$ $\therefore \sigma(1) = 1$.

if $\sigma(2) > 2$, $a_{2,\sigma(2)} = 0$, so $\sigma(2) \leq 2$, $\sigma(2) \neq 1$ $\Rightarrow \sigma(2) = 2$.

if $\sigma(3) > 3$, $a_{3,\sigma(3)} = 0$, so $\sigma(3) \leq 3$, $\sigma(3) \neq 1, 2$ $\Rightarrow \sigma(3) = 3$.

etc. (formally by induction), $\sigma(i) = i$ $\forall i$, so $\sigma = \mathrm{Id}$.

24 February 2012.
Dr Mark L ROBERTS.
Tomlin LT.

since $\det A^T = \det A$, this property above also applies to upper triangular matrices.

e.g. $\det \begin{pmatrix} 3 & -1 & -7 \\ 0 & 2 & 15 \\ 0 & 0 & 1 \end{pmatrix} = (3)(2)(1) = 6$.

Determinants of such matrices (triangular) are easily computable, so we attempt to convert other matrices to this form.

For more details, refer to $\boxed{\text{Handout 1}}$. [Defns E1–E3 ; facts F1–F5).

$\boxed{\text{Theorem}}$ 3.6 (a) Exchanging two rows of a matrix, $P(i,j)$ [or $p(i,j)$], multiplies the determinant by -1.

(b) Multiplying a row by λ, $D(i,\lambda)$ [or $d(i,\lambda)$], multiplies determinant by λ.

(c) Adding a multiple of one row to another, $E(i,j;\lambda)$ [or $e(i,j;\lambda)$], does not change the determinant.

for instance, in a 2x2 matrix: $\underset{\det:\, ad-bc}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \xrightarrow{D(2,\lambda)} \underset{\lambda ad - \lambda bc}{\begin{pmatrix} a & b \\ \lambda c & \lambda d \end{pmatrix}}$; $\underset{\det:\, ad-bc}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \xrightarrow{P(1,2)} \underset{bc-ad}{\begin{pmatrix} c & d \\ a & b \end{pmatrix}}$; $\underset{\det:\, ad-bc}{\begin{pmatrix} a & b \\ c & d \end{pmatrix}} \xrightarrow{E(1,2;\lambda)} \underset{d(a+\lambda c) - c(b+\lambda d) = ad-bc}{\begin{pmatrix} a+\lambda c & b+\lambda d \\ c & d \end{pmatrix}}$.

Proof – (a) WLOG, consider $P(1,2)$ in an $n\times n$ matrix.

$A = (a_{ij}) \xrightarrow{P(1,2)} B = (b_{ij})$, then $b_{1j} = a_{2j}$, $b_{2j} = a_{1j}$, $b_{mj} = a_{mj}$ $(m \geq 3)$

$\det B = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{2,\sigma(1)}\, a_{1,\sigma(2)}\, a_{3,\sigma(3)} \cdots a_{n,\sigma(n)}$

let $\tau = (1\ 2)$, then as σ varies over S_n, so does $\sigma\tau$.

so $\det B = \sum_{\sigma\in S_n} \operatorname{sgn}(\sigma\tau) \cdot a_{2,\sigma\tau(1)}\, a_{1,\sigma\tau(2)}\, a_{3,\sigma\tau(3)} \cdots a_{n,\sigma\tau(n)} = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)(\underset{\text{transposition}}{\operatorname{sgn}\tau})\, a_{2,\sigma(2)}\, a_{1,\sigma(1)}\, a_{3,\sigma(3)} \cdots a_{n,\sigma(n)}$

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

(b)

(c) First note that if A has two rows which are the same, then $\det A = 0$.

e.g. suppose rows i and j are the same, $A \xrightarrow{P(i,j)} A$; from the first part (a), $\det A = -\det A$, i.e. $\det A = 0$.

Now suppose $A \xrightarrow{e(1,2;\lambda)} B$, then $b_{1j} = a_{1j} + \lambda a_{2j}$; $b_{mj} = a_{mj}$ $(m \geq 2)$.

$\det B = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, b_{1,\sigma(1)} \cdots b_{n,\sigma(n)} = \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)(a_{1,\sigma(1)} + \lambda a_{2,\sigma(1)})\, a_{2,\sigma(2)}\, a_{3,\sigma(3)} \cdots a_{n,\sigma(n)}$

$= \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{1,\sigma(1)}\, a_{2,\sigma(2)} \cdots a_{n,\sigma(n)} + \lambda \sum_{\sigma\in S_n} (\operatorname{sgn}\sigma)\, a_{2,\sigma(1)}\, a_{2,\sigma(2)}\, a_{3,\sigma(3)} \cdots a_{n,\sigma(n)} = \det A + 0 = \det A$.

Note: 0 $\because$ first two rows are the same, so there is a 0 row.

29 February 2012
Dr. Mark L ROBERTS.
BLT.

The results we have established thus far provide a good calculational method for finding the determinant of large matrices.

Note also that since $\det A = \det A^T$, we can also perform column operations to the same effect. (i.e. operations $\mathcal{E}', \mathcal{D}', \mathcal{P}'$)

$\boxed{\text{Ex}}$ (i) $\det\begin{pmatrix}1&2&1&0\\2&2&0&2\\0&3&-1&-1\\0&0&1&1\end{pmatrix} \overset{\mathcal{E}(2,1;-2)}{=} \det\begin{pmatrix}1&2&1&0\\0&-2&-2&2\\0&3&-1&-1\\0&0&1&1\end{pmatrix} \overset{\mathcal{D}(2;-\frac{1}{2})}{=} -2\det\begin{pmatrix}1&2&1&0\\0&1&1&-1\\0&3&-1&-1\\0&0&1&1\end{pmatrix} \overset{\mathcal{E}(3,2;-3)}{=} -2\det\begin{pmatrix}1&2&1&0\\0&1&1&-1\\0&0&-4&2\\0&0&1&1\end{pmatrix} \overset{\mathcal{P}(3,4)}{=} 2\det\begin{pmatrix}1&2&1&0\\0&1&1&-1\\0&0&1&1\\0&0&-4&2\end{pmatrix}$

not $-\frac{1}{2}$!

$\overset{\mathcal{E}(4,3;4)}{=} 2\det\begin{pmatrix}1&2&1&0\\0&1&1&-1\\0&0&1&1\\0&0&0&6\end{pmatrix}$, which is a triangular matrix $= 2\,(1)(1)(1)(6) = 12$.

(ii) $\det\begin{pmatrix}1&1&1\\a&b&c\\a^2&b^2&c^2\end{pmatrix} \underset{\mathcal{E}'(3,1;-1)}{\overset{\mathcal{E}'(2,1;-1)}{=}} \det\begin{pmatrix}1&0&0\\a&b-a&c-a\\a^2&b^2-a^2&c^2-a^2\end{pmatrix} \underset{\mathcal{D}'(3,\frac{1}{c-a})}{\overset{\mathcal{D}'(2;\frac{1}{b-a})}{=}} (b-a)(c-a)\det\begin{pmatrix}1&0&0\\a&1&1\\a^2&b+a&c+a\end{pmatrix} \overset{\mathcal{E}'(3,2;-1)}{=} (b-a)(c-a)\det\begin{pmatrix}1&0&0\\a&1&0\\a^2&b+a&c-b\end{pmatrix}$

$= (b-a)(c-a)(c-b)$ Note: we call this the 3×3 Vandermonde matrix.

Thus, $\det A = (b-a)(c-a)(c-b)$; a,b,c all distinct $\iff$ A is invertible.

(iii) $\det\begin{pmatrix}0&2&3&1\\1&0&1&-1\\2&2&0&1\\3&4&2&-2\end{pmatrix} \underset{\mathcal{E}(4,2;-3)}{\overset{\mathcal{E}(3,2;-2)}{=}} \det\begin{pmatrix}0&2&3&1\\1&0&1&-1\\0&2&-2&1\\0&4&-1&1\end{pmatrix} \underset{\mathcal{E}(4,3;-2)}{\overset{\mathcal{E}(1,3;-1)}{=}} \det\begin{pmatrix}0&0&5&0\\1&0&1&-1\\0&2&-2&1\\0&0&3&-1\end{pmatrix} \overset{\mathcal{D}(1,\frac{1}{5})}{=} 5\det\begin{pmatrix}0&0&1&0\\1&0&1&-1\\0&2&-2&1\\0&0&3&1\end{pmatrix} \overset{\mathcal{E}(4,1;-3)}{=} 5\det\begin{pmatrix}0&0&1&0\\1&0&1&-1\\0&2&-2&1\\0&0&0&1\end{pmatrix}$

$\overset{\mathcal{P}(1,2)}{=} -5\det\begin{pmatrix}1&0&1&-1\\0&0&1&0\\0&2&-2&1\\0&0&0&1\end{pmatrix} \overset{\mathcal{P}(2,3)}{=} 5\det\begin{pmatrix}1&0&1&-1\\0&2&-2&1\\0&0&1&0\\0&0&0&1\end{pmatrix} = 5(1)(2)(1)(1) = 10$. $\times$ -38.

(iv) $\det\begin{pmatrix}1&1&1\\a&b&c\\a^3&b^3&c^3\end{pmatrix} \underset{\mathcal{E}'(3,1;-1)}{\overset{\mathcal{E}'(2,1;-1)}{=}} \det\begin{pmatrix}1&0&0\\a&b-a&c-a\\a^3&b^3-a^3&c^3-a^3\end{pmatrix} = (b-a)(c-a)\det\begin{pmatrix}1&0&0\\a&1&1\\a^3&b^2+ab+a^2&c^2+ac+a^2\end{pmatrix} = \left[\det\begin{pmatrix}1&0&0\\a&1&0\\a^3&b^2+ab+a^2&c^2-b^2+ac-ab\end{pmatrix}\right](b-a)(c-a)$

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

We now jump past points 3.7 to 3.10, temporarily.

Expansion along rows or down columns.

$\boxed{\text{Definition}}$ 3.11 The <u>(i,j)-minor</u> M_{ij} of an $n\times n$ matrix A is the determinant of what one gets by removing the i^{th} row and j^{th} column of A.

The <u>(i,j)-cofactor</u> C_{ij} of A is $(-1)^{i+j} M_{ij}$.

For instance, if $A = \begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix}$, then $M_{12} = \det\begin{pmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{pmatrix} = a_{21}a_{33} - a_{23}a_{31}$; $C_{12} = (-1)^{1+2} M_{12} = a_{23}a_{31} - a_{21}a_{33}$.

The signs are allocated in the pattern $\begin{pmatrix}+&-&\cdots\\-&+&\\\vdots&&\ddots\end{pmatrix}$ e.g. for $A = \begin{pmatrix}a&b\\c&d\end{pmatrix}$ $M_{11} = d$, $M_{12} = c$, $M_{21} = b$, $M_{22} = a$; $C_{11} = d$, $C_{12} = -c$, $C_{21} = -b$, $C_{22} = a$.

The matrix of minors is (M_{ij}). Here, $\begin{pmatrix}M_{11}&M_{12}\\M_{21}&M_{22}\end{pmatrix} = \begin{pmatrix}d&c\\b&a\end{pmatrix}$

Similarly, the matrix of cofactors is (C_{ij}). Here, $(C_{ij}) = \begin{pmatrix}d&-c\\-b&a\end{pmatrix}$

$\boxed{\text{Proposition}}$ 3.12 Let A be an $n\times n$ matrix. Then $\det A = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n} = a_{21}C_{21} + a_{22}C_{22} + \cdots + a_{2n}C_{2n}$

$= \sum_{j=1}^{n} a_{ij}C_{ij}$ for any fixed $1 \le i \le n$. (i.e. expand along i^{th} row).

$= \sum_{i=1}^{n} a_{ij}C_{ij}$ for any fixed $1 \le j \le n$ (i.e. expand along j^{th} column).

e.g. $A = \begin{pmatrix}a_{11}&a_{12}&a_{13}\\a_{21}&a_{22}&a_{23}\\a_{31}&a_{32}&a_{33}\end{pmatrix} = a_{11}C_{11} + a_{12}C_{12} + a_{13}C_{13} = a_{11}\begin{vmatrix}a_{22}&a_{23}\\a_{32}&a_{33}\end{vmatrix} - a_{12}\begin{vmatrix}a_{21}&a_{23}\\a_{31}&a_{33}\end{vmatrix} + a_{13}\begin{vmatrix}a_{21}&a_{22}\\a_{31}&a_{32}\end{vmatrix}$.

$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$.

Proof – omitted; just a matter of matching up terms.

The best way of calculating determinants is often a mixture of expanding and using row/column operations.

$\boxed{\text{Ex}}$ $\det\begin{pmatrix}1&0&3&4\\0&0&2&0\\2&1&4&-5\\11&0&2&1\end{pmatrix} = 0 + 0 - 2\det\begin{pmatrix}1&0&4\\2&1&-5\\11&0&1\end{pmatrix} + 0 = -2\det\begin{pmatrix}1&0&4\\2&1&-5\\11&0&1\end{pmatrix} \overset{\mathcal{E}'(3,1;-4)}{=} -2\det\begin{pmatrix}1&0&0\\2&1&-13\\11&0&-43\end{pmatrix} = -2\det\begin{pmatrix}1&-13\\0&-43\end{pmatrix} = -2(-43) = 86$.

or $\hookrightarrow = -2\begin{vmatrix}1&4\\11&1\end{vmatrix} = -2(1-44) = 86$

(Back to the theory behind determinants, 3.7–3.10).

We saw earlier for 2×2 case that A is invertible $\iff \det A \neq 0$; and that $\det(AB) = \det A \det B$.

We now prove these results for $n\times n$ case:

$\boxed{\text{Proposition}}$ 3.7 Let A be an $n\times n$ matrix, and E an $n\times n$ elementary matrix. Then $\det E \neq 0$ and $\det(EA) = \det E \det A$.

Proof – (i) Let $E = P(i,j)$. Then $P(i,j)A$ is the matrix obtained by applying the row operation $\mathcal{P}(i,j)$ to A.

By Theorem 3.6(a), $\det(P(i,j)A) = -\det A$

$P(i,j)$ is what we get by applying $\mathcal{P}(i,j)$ to I. By 3.6(a), $\det(P(i,j)) = -\det I = -1$; so $\det(P(i,j)A) = \det(P(i,j))\det A$.

(ii),(iii) these apply similarly to show that $\det(E(i,j;\lambda)) = 1$ and $\det(D(i,\lambda)) = \lambda$.

By induction, this yields:

$\boxed{\text{Corollary}}$ 3.7 Let A be a square matrix, and $E_1, \dots, E_n$ be elementary matrices of the same size. Then,

$\det(E_n \cdots E_1 A) = \det(E_n)\det(E_{n-1})\cdots\det(E_1)\det(A)$.

$\boxed{\text{Theorem}}$ 3.8 Let A be an $n\times n$ matrix. Then A is invertible $\iff$ $\det A \neq 0$.

Proof – We know that there are elementary matrices $E_1, E_2, \dots, E_n$ s.t. $E_n \cdots E_1 A = T$ (RRE form).

By Corollary 3.7, $\det(E_n)\det(E_{n-1})\cdots\det(E_1)\det A = \det T$.

Also, $\det E \neq 0$, so $\det A = 0 \iff \det T = 0$.

Suppose A is invertible, then $T = I_n$ and $\det T = \det I_n = 1 \neq 0 \Rightarrow \det A \neq 0$.

Suppose A is not invertible, then T has a zero row so $\det T = 0$. Hence, $\det A = 0$.

3 March 2012.
Dr Mark L ROBERTS.
Damin LT.

$\boxed{\text{Theorem}}$ 3.10 For two $n\times n$ matrices A and B, $\det(AB) = \det A \det B$.

Proof – By F2, $\exists$ elementary matrices $E_1, E_2, \dots, E_n$ s.t. $E_n \cdots E_1 A = T$, in RRE form.

$A = E_1^{-1}E_2^{-1}\cdots E_n^{-1}T$. By F3, each E_i^{-1} is again elementary, say $E_i^{-1} = F_i$

$A = F_1 \cdots F_n T$. By Theorem 3.8, $\det A = \det F_1 \cdots \det F_n \det T$.

$AB = F_1 \cdots F_n TB$. Again by Theorem 3.8, $\det(AB) = \det F_1 \cdots \det F_n \det(TB)$.

By F4, either $T = I$ or T has a zero row.

Case 1:
If $T = I$, $\det A = \det F_1 \cdots \det F_n$; $\det AB = \det F_1 \cdots \det F_n \det(IB) = \det F_1 \cdots \det F_n \det B = \det A \det B$.

Case 2:
If T has a zero row, then TB has a zero row as well; $\therefore \det T = 0$, $\det TB = 0 \Rightarrow \det A = \det AB = 0$,

$\therefore \det(AB) = \det A \det B \quad \forall A, B \in M_n$.

Adjugate and inverse.

We aim to get a formula for A^{-1}.

$\boxed{\text{Definition}}$ 3.13 The adjugate, $\operatorname{adj} A$, of an $n\times n$ matrix A is $\operatorname{adj} A = C^T$.

i.e. $(\operatorname{adj} A)_{ij} = C_{ji}$

for instance, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, $M = \begin{pmatrix} d & c \\ b & a \end{pmatrix}$, $C = \begin{pmatrix} d & -c \\ -b & a \end{pmatrix}$, $\operatorname{adj} A = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$

Note then that $A \operatorname{adj} A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} = \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} = (ad-bc)I_2 \Rightarrow A \operatorname{adj} A = \det A \cdot I_2$

$\Rightarrow A \frac{1}{\det A}\operatorname{adj} A = I_2 \quad (= \frac{1}{\det A}\operatorname{adj} A \cdot A)$.

Hence we establish that in 2×2 case, $A^{-1} = \frac{1}{\det A}\operatorname{adj} A$.

$\boxed{\text{Theorem}}$ 3.14 Let A be an $n\times n$ matrix. Then $A \operatorname{adj} A = (\det A) I_n = (\operatorname{adj} A) A$.

Hence if $\det A \neq 0$ (i.e. A is invertible), $A^{-1} = \frac{1}{\det A}\operatorname{adj} A$.

Proof – The $(i,i)^{th}$-entry of $A(\operatorname{adj} A) = \sum_{j=1}^n a_{ij}(\operatorname{adj} A)_{ji}$ (matrix product) $= \sum_{j=1}^n a_{ij}C_{ij} = \det A$ (cofactor expansion).

The $(i,j)^{th}$ entry for $i \neq j$: consider $i=1, j=2$.

then $(1,2)^{th}$ entry of $A(\operatorname{adj} A) = \sum_{j=1}^n a_{ij}(\operatorname{adj} A)_{j2} = \sum_{j=1}^n a_{1j}C_{2j}$

Consider matrix B with first row of A duplicated, i.e. we define $B = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$; if C' represents matrix of cofactors of B,

$\det B = \sum_{j=1}^n b_{2j}C'_{2j} = \sum_{j=1}^n a_{1j}C_{2j}$ ($\because$ by removing 1^{st} col, 2^{nd} row, $C_{2j} = C'_{2j}$).

but $\det B = A \Rightarrow (1,2)^{th}$ entry of $A(\operatorname{adj} A) = 0$

$\therefore A(\operatorname{adj} A)$ has diagonal entries $\det A$, off-diagonal entries 0. Thus, $A(\operatorname{adj} A) = (\det A) I_n$.

Similarly, $(\operatorname{adj} A) A = (\det A) I_n$ // q.e.d.

$\boxed{\text{Ex}}$

(i) Let $A = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \\ 0 & -1 & 1 \end{pmatrix}$. Find A^{-1}.

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

$\det A = \sum a_{ij}C_{ij} = 1\times 3 + 2\times -3 + 3\times -3 = -12 \Rightarrow A$ is invertible, $A^{-1} = \frac{1}{-12}\begin{pmatrix} 3 & -5 & -4 \\ -3 & 1 & 8 \\ -3 & 1 & -4 \end{pmatrix} = \frac{1}{12}\begin{pmatrix} -3 & 5 & 4 \\ 3 & -1 & -8 \\ 3 & -1 & 4 \end{pmatrix}$ //.

(ii) Let $A = \begin{pmatrix} 0 & 1 & 1 \\ 2 & -1 & -1 \\ 1 & 1 & 2 \end{pmatrix}$. Find A^{-1}.

We have $M = \begin{pmatrix} -1 & 5 & 3 \\ 1 & -1 & -1 \\ 0 & -2 & -2 \end{pmatrix}$.

(iii) Let $A = \begin{pmatrix} \alpha & 1 & 2 \\ 0 & \beta & 1 \\ 1 & \gamma & 2 \end{pmatrix}$. For which values of α, β, γ is A invertible? For these values, find A^{-1}.

CHAPTER 4.
DIAGONALISATION.

7 March 2012
Dr Mark L ROBERTS
GLT.

Diagonalisation is an important result in linear algebra and its applications.

Recall that an $n\times n$ matrix A is diagonal if $a_{ij}=0$ for all $i\neq j$ i.e. all entries off the main diagonal are 0.

e.g. 2×2: $\begin{pmatrix} d_1 & 0 \\ 0 & d_2\end{pmatrix}$, 3×3: $\begin{pmatrix} d_1 & 0 & 0 \\ 0 & d_2 & 0 \\ 0 & 0 & d_3\end{pmatrix}$; we write $\operatorname{diag}(d_1, d_2, \dots, d_n)$ for $\begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \ddots & \\ & & & d_n\end{pmatrix}$, so for instance, $\operatorname{diag}(2,0,3)=\begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 3\end{pmatrix}$.

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

$\boxed{\text{Definition}}$ 4.1 An $n\times n$ matrix is diagonalisable if there exists an invertible matrix P such that $P^{-1}AP=D$ (diagonal).

How could we find such a matrix P? $P(P^{-1}AP)=PD \Rightarrow AP=PD$.

In the 2×2 case, this means that for $A=\begin{pmatrix} a & b \\ c & d\end{pmatrix}$, we seek $P=\begin{pmatrix} p & q \\ r & s\end{pmatrix}$ s.t. $\begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix} p & q \\ r & s\end{pmatrix}=\begin{pmatrix} p & q \\ r & s\end{pmatrix}\begin{pmatrix} d_1 & 0 \\ 0 & d_2\end{pmatrix}$.

Multiplying the first column, we get $\begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix} p \\ r\end{pmatrix}=\begin{pmatrix} p \\ r\end{pmatrix}d_1$; and the second column, we get $\begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix} q \\ s\end{pmatrix}=\begin{pmatrix} q \\ s\end{pmatrix}d_2$.

If we name the columns of P as $\underline{v}_1=\begin{pmatrix} p \\ r\end{pmatrix}$, $\underline{v}_2=\begin{pmatrix} q \\ s\end{pmatrix}$; then $A\underline{v}_1=d_1\underline{v}_1$, $A\underline{v}_2=d_2\underline{v}_2$. So the columns of P are solutions to $A\underline{v}=*\underline{v}$.

$\boxed{\text{Proposition}}$ 4.2 Let $\underline{v}_1,\dots,\underline{v}_n\in\mathbb{F}^n$, and let $P=(\underline{v}_1\ \underline{v}_2\cdots\underline{v}_n)$, where $\underline{v}_1,\dots,\underline{v}_n$ represent the columns of P, then the following are equivalent:

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

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

(iii) P is invertible.

For instance, $P=\begin{pmatrix} 1 & 2 \\ 2 & 4\end{pmatrix}$ is not invertible $\because \det P=0 \Rightarrow \begin{pmatrix} 1 \\ 2\end{pmatrix},\begin{pmatrix} 2 \\ 4\end{pmatrix}$ are not LI $\Rightarrow$ not a basis for $\mathbb{F}^2$.

$\boxed{\text{Definition}}$ 4.3 Let A be an $n\times n$ matrix over $\mathbb{F}$. Then λ is an eigenvalue of A if $\exists$ a non-zero vector $\underline{v}\in\mathbb{F}^n$ s.t. $A\underline{v}=\lambda\underline{v}$.

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

Nomenclature: the eigenvalues/eigenvectors are sometimes also called characteristic values/vectors.

$\boxed{\text{Proposition}}$ 4.4 (Basic criterion for diagonalisability)

The following are equivalent for an $n\times n$ matrix A over $\mathbb{F}$:

(i) A is diagonalisable;

(ii) there exists a basis for $\mathbb{F}^n$ consisting of eigenvectors;

(iii) there exist n LI eigenvectors.

Proof — (i) $\Rightarrow$ (ii): suppose P is invertible, $P^{-1}AP=D$. Then $AP=PD$. Let columns of P be $\underline{v}_1,\underline{v}_2,\dots,\underline{v}_n$; so $P=(\underline{v}_1\cdots\underline{v}_n)$.

Let $D=\operatorname{diag}(d_1,d_2,\dots,d_n)$. This gives us $A(\underline{v}_1\,\vdots\,\underline{v}_2\,\vdots\cdots\vdots\,\underline{v}_n)=(\underline{v}_1\,\vdots\,\underline{v}_2\,\vdots\cdots\vdots\,\underline{v}_n)\begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n\end{pmatrix}$.

$\Rightarrow (A\underline{v}_1\,\vdots\,A\underline{v}_2\,\vdots\cdots\vdots\,A\underline{v}_n)=(d_1\underline{v}_1\,\vdots\,d_2\underline{v}_2\,\vdots\cdots\vdots\,d_n\underline{v}_n)$. Hence $\forall i$, $1\le i\le n$, $A\underline{v}_i=d_i\underline{v}_i$,

i.e. each $\underline{v}_i$ is an eigenvector of A (associated to d_i). Since P is invertible, by 4.3, $\{\underline{v}_1,\dots,\underline{v}_n\}$ is a basis for $\mathbb{F}^n$, i.e. $\mathbb{F}^n$ has a basis consisting of eigenvectors.

(ii) $\Rightarrow$ (i): this is the same argument as (i) $\Rightarrow$ (ii), read in reverse.

(ii) $\Rightarrow$ (iii): Proof through definition 4.3.

Finding eigenvalues and eigenvectors:

Objective: we want to find non-zero solutions to $A\underline{v}=\lambda\underline{v}$, where $\underline{v}$ and λ are initially unknown.

$\boxed{\text{Proposition}}$ 4.5. Let $A\in M_n(\mathbb{F})$, $\lambda\in\mathbb{F}$. Then the following are equivalent:

(i) λ is an eigenvalue of A

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

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

Proof: (i) $\Rightarrow$ (ii): Suppose $A\underline{v}=\lambda\underline{v}$, $\underline{v}\neq\underline{0}$. Then $A\underline{v}=(\lambda I_n)\underline{v}$ ($\because \underline{v}\in\mathbb{F}^n \Rightarrow I_n\underline{v}=\underline{v}$)

$\Rightarrow (\lambda I_n - A)\underline{v}=\underline{0}$. Since $\underline{v}\neq\underline{0}$, $(\lambda I_n - A)$ is not invertible. (otherwise $\underline{v}=(\lambda I_n - A)^{-1}\cdot\underline{0}=\underline{0}$).

(ii) $\Rightarrow$ (i): Suppose $\lambda I - A$ is singular. Then the system $(\lambda I - A)\underline{x} = \underline{0}$ has a non-zero solution for $\underline{x}$.

If this solution is $\underline{v}$, then $A\underline{v} = \lambda \underline{v} \Rightarrow \lambda$ is an eigenvalue // q.e.d.

(ii) $\Leftrightarrow$ (iii): see Theorem 3.8. // q.e.d.

So, we now have a method for finding eigenvalues and eigenvectors; and hence diagonalising.

$\boxed{\text{Ex}}$ (i) Let $A = \begin{pmatrix}1&2\\6&2\end{pmatrix}$. Then $\lambda I - A = \begin{pmatrix}\lambda&0\\0&\lambda\end{pmatrix} - \begin{pmatrix}1&2\\6&2\end{pmatrix} = \begin{pmatrix}\lambda-1&-2\\-6&\lambda-2\end{pmatrix}$

λ is an eigenvalue of A if $\det(\lambda I - A) = 0$ i.e. $\det\begin{pmatrix}\lambda-1&-2\\-6&\lambda-2\end{pmatrix} = 0 \Rightarrow (\lambda-1)(\lambda-2) - 12 = 0 \Rightarrow \lambda^2 - 3\lambda - 10 = 0. \Rightarrow (\lambda-5)(\lambda+2) = 0. \Rightarrow \lambda = 5, -2.$

there are two eigenvalues, 5 and -2.

We then find the corresponding eigenvectors.

where $\lambda = 5$, $A\underline{v} = 5\underline{v} \Rightarrow \begin{pmatrix}1&2\\6&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = 5\begin{pmatrix}x\\y\end{pmatrix} \Rightarrow$ system is $\begin{cases}x + 2y = 5x\\6x + 2y = 5y\end{cases} \Rightarrow$ general solution is $\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}\alpha\\2\alpha\end{pmatrix}$, $\alpha \in \mathbb{R}$.

Pick any value of α, e.g. $\underline{v}_1 = \begin{pmatrix}1\\2\end{pmatrix}$

where $\lambda = -2$, $A\underline{v} = -2\underline{v} \Rightarrow \begin{pmatrix}1&2\\6&2\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = -2\begin{pmatrix}x\\y\end{pmatrix}$ [or equivalently, $(A+2I)\underline{v} = \underline{0}$] $\Rightarrow \begin{pmatrix}3&2\\6&4\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix}$

$\begin{pmatrix}3&2&0\\6&4&0\end{pmatrix} \longrightarrow \begin{pmatrix}1&2/3&0\\0&0&0\end{pmatrix}$. General solution, fixing y, is $\begin{pmatrix}-\frac{2}{3}y\\y\end{pmatrix}$. We pick any value, e.g. $\underline{v}_2 = \begin{pmatrix}-2\\3\end{pmatrix}$.

then $P = \begin{pmatrix}1&-2\\2&3\end{pmatrix}$. P is invertible, and $P^{-1}AP = \begin{pmatrix}5&0\\0&2\end{pmatrix}$.

check: $\det P = 3 + 4 = 7 \neq 0 \Rightarrow P$ is invertible; and we see if $AP = PD$: $AP = \begin{pmatrix}1&2\\6&2\end{pmatrix}\begin{pmatrix}1&-2\\2&3\end{pmatrix} = \begin{pmatrix}5&4\\10&-6\end{pmatrix}$, $PD = \begin{pmatrix}1&-2\\2&3\end{pmatrix}\begin{pmatrix}5&0\\0&-2\end{pmatrix} = \begin{pmatrix}5&4\\10&-6\end{pmatrix}$ // q.e.d.

Note: the order of entries in D and P must correspond! i.e. λ_1 must correspond to eigenvector $\underline{v}_1$.

(ii) Let $A = \begin{pmatrix}2&1\\1&2\end{pmatrix}$. Then $\lambda I - A = \begin{pmatrix}\lambda&0\\0&\lambda\end{pmatrix} - \begin{pmatrix}2&1\\1&2\end{pmatrix} = \begin{pmatrix}\lambda-2&-1\\-1&\lambda-2\end{pmatrix}$. $\det(\lambda I - A) = 0 \Rightarrow (\lambda-2)^2 - 1 = 0 \Rightarrow \lambda^2 - 4\lambda + 3 = 0 \Rightarrow \lambda = 1$ or 3.

$\lambda_1 = 1 \Rightarrow (\lambda_1 I - A)\underline{v}_1 = \underline{0} \Rightarrow \begin{pmatrix}-1&-1\\-1&-1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow \begin{pmatrix}-1&-1&0\\-1&-1&0\end{pmatrix} \Rightarrow \underline{v}_1 = \begin{pmatrix}1\\-1\end{pmatrix}$.

$\lambda_2 = 3 \Rightarrow (\lambda_2 - A)\underline{v}_2 = \underline{0} \Rightarrow \begin{pmatrix}1&-1\\-1&1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}0\\0\end{pmatrix} \Rightarrow \underline{v}_2 = \begin{pmatrix}1\\1\end{pmatrix}$.

$P = \begin{pmatrix}1&1\\-1&1\end{pmatrix}$, $D = \begin{pmatrix}1&0\\0&3\end{pmatrix}$. Then $AP = PD = \begin{pmatrix}1&3\\-1&3\end{pmatrix}$.

Applications of diagonalisation:

(i) Given A, find A^m — application 4.6

(ii) solve simultaneous linear difference equations.

(iii) solve simultaneous linear differential equations.

$\boxed{\text{Application}}$ 4.6 Given a diagonalisable matrix A, find A^m.

This is easy if A is diagonal: $\mathrm{diag}(d_1, d_2, \ldots, d_n)^m = \mathrm{diag}(d_1^m, d_2^m, \ldots, d_n^m)$.

e.g. $\begin{pmatrix}2&0\\0&3\end{pmatrix}^m = \begin{pmatrix}2^m&0\\0&3^m\end{pmatrix}$.

Suppose $P^{-1}AP = D$, then $(P^{-1}AP)^m = D^m$; note that $(P^{-1}AP)^m = P^{-1}AP \cdot P^{-1}AP \cdot P^{-1} \cdots\cdots P \cdot P^{-1}AP = P^{-1}AIAI\cdots IAP = P^{-1}A^mP. = D^m$.

therefore, $\boxed{A^m = PD^mP^{-1}}$

General approach: Problem about $A \xrightarrow{\text{diagonalise}}$ Problem about $D \xrightarrow{\text{solve}}$ Solution for $D \xrightarrow{\text{undo diagonalisation}}$ solution for A.

$\boxed{\text{Ex}}$ Find $\begin{pmatrix}1&2\\6&2\end{pmatrix}^m$.

From earlier work, $A = \begin{pmatrix}1&2\\6&2\end{pmatrix}$, $P = \begin{pmatrix}1&-2\\2&3\end{pmatrix}$, $D = \begin{pmatrix}5&0\\0&-2\end{pmatrix} = P^{-1}AP. \Rightarrow P^{-1}A^mP = D^m = \begin{pmatrix}5^m&0\\0&(-2)^m\end{pmatrix}$

$$A^m = P\begin{pmatrix}5^m&0\\0&(-2)^m\end{pmatrix}P^{-1} = \begin{pmatrix}1&-2\\2&3\end{pmatrix}\begin{pmatrix}5^m&0\\0&(-2)^m\end{pmatrix}\begin{pmatrix}1&-2\\2&3\end{pmatrix}^{-1} = \begin{pmatrix}1&-2\\2&3\end{pmatrix}\begin{pmatrix}5^m&0\\0&(-2)^m\end{pmatrix}\frac{1}{7}\cdot\begin{pmatrix}3&2\\-2&1\end{pmatrix} = \frac{1}{7}\begin{pmatrix}5^m&(-2)^{m+1}\\2\cdot5^m&3(-2)^m\end{pmatrix}\begin{pmatrix}3&2\\-2&1\end{pmatrix}$$
$$= \frac{1}{7}\begin{pmatrix}3\cdot5^m + (-2)^{m+2} & 2\cdot5^m + (-2)^{m+1}\\6\cdot5^m + 3\cdot(-2)^{m+1} & 4\cdot5^m + 3(-2)^m\end{pmatrix}.$$

check: where $m = 0$, $A^0 = \frac{1}{7}\begin{pmatrix}3+4&2-2\\6-6&4+3\end{pmatrix} = \frac{1}{7}\begin{pmatrix}7&0\\0&7\end{pmatrix} = I_2$; where $m = 1$, $A^1 = \frac{1}{7}\begin{pmatrix}7&14\\42&14\end{pmatrix} = \begin{pmatrix}1&2\\6&2\end{pmatrix}$.

$\boxed{\text{Application}}$ 4.7 Solving simultaneous difference equations e.g. $\begin{cases}x_{n+1} = ax_n + by_n\\y_{n+1} = cx_n + dy_n\end{cases}$.

Let $\underline{v}_n = \begin{pmatrix}x_n\\y_n\end{pmatrix}$, $A = \begin{pmatrix}a&b\\c&d\end{pmatrix}$. Then $\underline{v}_{n+1} = A\underline{v}_n$. Solution is $\boxed{\underline{v}_n = A^n\underline{v}_0}$ — find A^n by previous method.

9 March 2017
Dr Mark L Roberts.
Domain 4.

$\boxed{\text{Application}}$ 4.8 Solving simultaneous differential equations e.g. $\begin{cases}\frac{dx_1}{dt} = ax_1 + bx_2\\\frac{dx_2}{dt} = cx_1 + dx_2\end{cases}$.

Easy to solve if $b = c = 0$, then $\begin{cases}dx_1/dt = ax_1\\dx_2/dt = dx_2\end{cases} \Rightarrow \begin{matrix}x_1 = Ae^{at}\\x_2 = Be^{dt}\end{matrix}$.

Let $\underline{x} = \begin{pmatrix}x_1\\x_2\end{pmatrix}$. Use $'$ for differentiation w.r.t. t, $A = \begin{pmatrix}a&b\\c&d\end{pmatrix}$. Then $\begin{pmatrix}dx_1/dt\\dx_2/dt\end{pmatrix} = \begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}x_1\\x_2\end{pmatrix} \Rightarrow \underline{x}' = A\underline{x}$.

From $\underline{x}' = A\underline{x}$, we make a change of variables. $\underline{x} = P\underline{y}$ i.e. $\left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right]$

then differentiating w.r.t. t, $\underline{x}' = P\underline{y}'$; and $\underline{x}' = A\underline{x}$ reduces to $P\underline{y}' = AP\underline{y}$.

this gives us $\underline{y}' = P^{-1}AP\underline{y} \Rightarrow \underline{y}' = D\underline{y}$.

$\boxed{\text{Ex}}$ solve $\frac{dx_1}{dt} = x_1 + 2x_2$; $\frac{dx_2}{dt} = 6x_1 + 2x_2$; with initial conditions $x_1(0) = 2$, $x_2(0) = 1$.

Let $A = \begin{pmatrix} 1 & 2 \\ 6 & 2 \end{pmatrix}$, $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$. then $\underline{x}' = A\underline{x}$ and $\underline{x}(0) = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

From earlier example, if $P = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}$, then $P^{-1}AP = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix} = D$.

Let $\underline{x} = P\underline{y}$, then $P\underline{y}' = AP\underline{y}$, $\underline{y}' = P^{-1}AP\underline{y} = D\underline{y} \Rightarrow \begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$

$\Rightarrow y_1' = 5y_1$ and $y_2' = 2y_2 \Rightarrow \underline{y} = \begin{pmatrix} c_1 e^{5t} \\ c_2 e^{2t} \end{pmatrix}$.

We now find constants c_1, c_2 and changing variables $\underline{x} = P\underline{y} \Rightarrow \underline{y} = P^{-1}\underline{x}$ and $P^{-1} = \frac{1}{7}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}$.

so $\underline{y}(0) = \frac{1}{7}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}\underline{x}(0) = \frac{1}{7}\begin{pmatrix} 3 & 2 \\ -2 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 8/7 \\ -3/7 \end{pmatrix}$; and $\underline{y}(0) = \begin{pmatrix} c_1 e^0 \\ c_2 e^0 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$ $\therefore$ $\underline{y} = \frac{1}{7}\begin{pmatrix} 8e^{5t} \\ -3e^{-2t} \end{pmatrix}$.

$\underline{x} = P\underline{y} = \begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}\underline{y} = \frac{1}{7}\begin{pmatrix} 1 & -2 \\ 2 & 3 \end{pmatrix}\begin{pmatrix} 8e^{5t} \\ -3e^{-2t} \end{pmatrix} = \frac{1}{7}\begin{pmatrix} 8e^{5t} + 6e^{-2t} \\ 16e^{5t} - 9e^{-2t} \end{pmatrix} \Rightarrow$ solution is $x_1 = \frac{1}{7}(8e^{5t} + 6e^{-2t})$; $x_2 = \frac{1}{7}(16e^{5t} - 9e^{-2t})$ //.

Not all square matrices can be diagonalised.

$\boxed{\text{Definition}}$ 4.9 Let $A \in M_n(\mathbb{F})$. Then the <u>characteristic polynomial</u> of A is given by $C_A(t) = \det(tI - A)$

Recall that the eigenvalues of A are roots of $C_A(t) = 0$. (Proposition 4.5).

the factorisation of $C_A(t)$ into irreducible linear factors is important in determining whether A is diagonalisable.

One way that A can fail to be diagonalisable is the case of "missing eigenvalues".

For instance, let $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in M_2(\mathbb{R})$. $C_A(t) = \det\begin{pmatrix} t & 1 \\ -1 & t \end{pmatrix} = t^2 + 1 \Rightarrow$ eigenvalues are roots of $t^2 + 1 = 0$ — no real roots.

since $A \in M_2(\mathbb{R})$, there are no real eigenvalues $\Rightarrow$ no eigenvectors so A is not diagonalisable (in the reals).

If we regard $A \in M_2(\mathbb{C})$, then the two eigenvalues are i and $-i$, and A can be diagonalised. The problem with diagonalisation cannot occur with $\mathbb{C}$ in general.

$\boxed{\text{Theorem}}$ 4.10 (Fundamental theorem of Algebra).

Any polynomial in $\mathbb{C}[t]$ factorises into linear factors.

Proof — omitted. Mainly an analysis proof. closely related to Intermediate Value Theorem.

In fact, if $C_A(t)$ does not factorise into linear factors, then it cannot be diagonalised.

From here on, we consider case where it does factorise into linear factors.

$C_A(t) = (t-\lambda_1)^{f_1} \cdots (t-\lambda_r)^{f_r}$ where $\lambda_1, \lambda_2, \ldots, \lambda_r$ are the eigenvalues, and $\sum_{i=1}^{r} f_i = n$.

$\boxed{\text{Theorem}}$ 4.11 Let $A \in M_n(\mathbb{F})$, and suppose that A has n distinct eigenvalues. Then A is diagonalisable.

Proof — Let $\lambda_1, \ldots, \lambda_n$ be the eigenvalues with associated eigenvectors $\underline{v}_1, \ldots, \underline{v}_n$.

$C_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$. We claim that $\{\underline{v}_1, \ldots, \underline{v}_n\}$ is LI.

Proof by contradiction — suppose $\{\underline{v}_1, \ldots, \underline{v}_n\}$ is not LI. We pick a shortest possible relation of dependence.

By renumbering, we get $\alpha_1\underline{v}_1 + \cdots + \alpha_r\underline{v}_r = \underline{0}$, all $\alpha_i \neq 0$; with no relation involving $<r$ terms. — (1).

Manipulating (1), we get $A(\alpha_1\underline{v}_1 + \cdots + \alpha_r\underline{v}_r) = A\underline{0} = \underline{0}$.

$\alpha_1 A\underline{v}_1 + \cdots + \alpha_r A\underline{v}_r = \underline{0} \Rightarrow \alpha_1\lambda_1\underline{v}_1 + \cdots + \alpha_r\lambda_r\underline{v}_r = \underline{0}$ — (2) $\because \lambda_1, \ldots, \lambda_r$ are eigenvalues.

Taking λ_r multiples of (1), we have $\alpha_1\lambda_r\underline{v}_1 + \cdots + \alpha_r\lambda_r\underline{v}_r = \underline{0}$.

(2) − (3): $\alpha_1(\lambda_1 - \lambda_r)\underline{v}_1 + \cdots + \alpha_{r-1}(\lambda_{r-1} - \lambda_r)\underline{v}_{r-1} = \underline{0}$. Then this is a shorter relation since it involves $\leq r-1$ terms

and is non-trivial since $\alpha_1(\lambda_1 - \lambda_r) \neq 0$ as $\alpha_1 \neq 0$ and $\lambda_1 \neq \lambda_r$. $\Rightarrow$ contradiction to hypothesis that it was the shortest relation.

Hence, we conclude that no dependence relation exists; thus we have n LI eigenvectors, and

by basic criterion (4.4), A is diagonalisable.

14 March 2012
Dr Mark Roberts
ALT.

In fact, we can develop a method to diagonalise $n \times n$ matrices with n distinct eigenvalues.

$\boxed{\text{Method}}$ 4.12 How to diagonalise an $n\times n$ matrix with n different eigenvalues.

(i) Find the characteristic polynomial $c_A(t) = \det(tI_n - A)$

(ii) Factorise it into linear factors $c_A(t) = (t-\lambda_1)(t-\lambda_2)\cdots(t-\lambda_n)$.

(iii) For each eigenvalue λ_i, find a corresponding eigenvector $\underline{v}_i$.

(iv) The set $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n\}$ is LI and hence forms a basis for $\mathbb{R}^n$, so the matrix $P = (\underline{v}_1\, \underline{v}_2 \cdots \underline{v}_n)$ is invertible.

(v) then $P^{-1}AP = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$.

What else can hinder diagonalisation? It must be something to do with repeated roots.

e.g. $A = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$, $c_A(t) = \det\begin{pmatrix} t-3 & 1 \\ 0 & t-3 \end{pmatrix} = (t-3)^2$ $\Rightarrow$ A has eigenvalue 3 (twice). Then suppose $\begin{pmatrix} x \\ y \end{pmatrix}$ is an eigenvector, $A\begin{pmatrix} x \\ y \end{pmatrix} = 3\begin{pmatrix} x \\ y \end{pmatrix}$.

then $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$ $\Rightarrow$ $y = 0$, eigenvalue is of form $\begin{pmatrix} x \\ 0 \end{pmatrix}$ $\Rightarrow$ there are not two LI eigenvectors.

Note that $B = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$, then $c_B(t) = (t-3)^2$, but B is diagonalisable $\Rightarrow$ so having a repeated eigenvalue is not adequate to claim A is not diagonalisable.

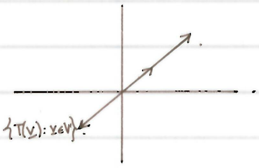
The problem is that in A, there are "not enough" eigenvectors associated with the eigenvalue.

We revise some material on subspaces:

$\boxed{\text{Definition}}$ 4.13 Let V be a vector space over $\mathbb{F}$. Then a subspace W of V is a non-empty $W \subseteq V$ s.t.

$\lambda, \mu \in \mathbb{F}$, $\underline{u}, \underline{v} \in W \Rightarrow \lambda\underline{u} + \mu\underline{v} \in W$. We write $W \leq V$.

A subspace forms a vector space itself. e.g. subspaces of $\mathbb{R}^2$ include $\{\underline{0}\}$, any line through 0, or $\mathbb{R}^2$ itself.



For instance also, if $T: V \to W$ is linear, $\ker T \leq V$, $\operatorname{Im} T \leq W$. $\because$ $\ker T = \{\underline{v} \in V : T(\underline{v}) = 0\}$, $\operatorname{Im} T = \{T(\underline{v}) : \underline{v} \in V\}$.

similarly, $\{\underline{x} : A\underline{x} = \underline{0}\}$ is a subspace of $\mathbb{R}^n$. This is called ~~an~~ affine set.

$\boxed{\text{Definition}}$ 4.14 If $U, W \leq V$, then the sum $U + W$ is $U + W = \{\underline{u} + \underline{w} : \underline{u} \in U, \underline{w} \in W\}$.

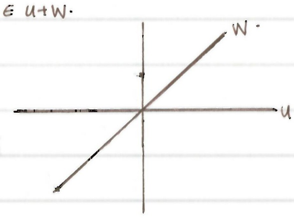
$\boxed{\text{Proposition}}$ 4.15 If $U, W \leq V$, then $U \cap W \leq V$, $U + W \leq V$

Proof — (for $U+W \leq V$). Let $\underline{x}_1, \underline{x}_2 \in U+W$; $\underline{x}_1 = \underline{u}_1 + \underline{w}_1$, $\underline{x}_2 = \underline{u}_2 + \underline{w}_2$ for some $\underline{u}_i \in U$, $\underline{w}_i \in W$.

If $\lambda, \mu \in \mathbb{F}$, then $\lambda\underline{x}_1 + \mu\underline{x}_2 = \lambda(\underline{u}_1 + \underline{w}_1) + \mu(\underline{u}_2 + \underline{w}_2) = \underbrace{(\lambda\underline{u}_1 + \mu\underline{u}_2)}_{\in U} + \underbrace{(\lambda\underline{w}_1 + \mu\underline{w}_2)}_{\in W} \in U + W$.

Also, $\underline{0} \in U$, $\underline{0} \in W \Rightarrow \underline{0} \in U+W$.

e.g. $U = \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\}$, $W = \left\{\begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R}\right\}$; then $U, W \leq \mathbb{R}^2$.



$U + W = \{\underline{u} + \underline{w} : \underline{u} \in U, \underline{w} \in W\} = \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} + \begin{pmatrix} y \\ y \end{pmatrix} : x, y \in \mathbb{R}\right\} = \left\{\begin{pmatrix} x+y \\ y \end{pmatrix} : x, y \in \mathbb{R}\right\} = \mathbb{R}^2$.

$U \cap W = \left\{\begin{pmatrix} 0 \\ 0 \end{pmatrix}\right\}$.

$\boxed{\text{Ex}}$ Let $V = \mathbb{R}^3$. $U = \left\{\begin{pmatrix} x \\ x \\ y \end{pmatrix} : x, y \in \mathbb{R}\right\}$, $W = \left\{\begin{pmatrix} x \\ y \\ y \end{pmatrix} : x, y \in \mathbb{R}\right\}$. Find $U+W$, $U \cap W$ and dimension of each of $U, W, U+W, U\cap W$. Find a relation between them.

$U + W = \{\underline{u} + \underline{w} : \underline{u} \in U, \underline{w} \in W\} = \left\{\begin{pmatrix} x \\ x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \\ y' \end{pmatrix} : x, y, x', y' \in \mathbb{R}\right\} = \left\{\begin{pmatrix} x+x' \\ x+y' \\ y+y' \end{pmatrix} : x, y \in \mathbb{R}\right\} = \mathbb{R}^3$. $\Rightarrow \dim(U+W) = 3$.

$U \cap W = \left\{\begin{pmatrix} x \\ x \\ x \end{pmatrix} : x \in \mathbb{R}\right\}$. $\dim(U \cap W) = 1$

$\dim U = 2$, $\dim W = 2$; then $\dim(U+W) = \dim U + \dim W - \dim(U \cap W)$ //

$\boxed{\text{Theorem}}$ 4.16 Let $U, W \leq V$. Let $U, W \leq V$. Then $\dim(U+W) + \dim(U \cap W) = \dim U + \dim W$.

$\boxed{\text{Definition}}$ 4.17 Let $U, W \leq V$. Then $\underline{U+W \text{ is direct}}$ (we write $U \oplus W$) if $U \cap W = \{\underline{0}\}$.

If $U+W$ is direct, then from theorem 4.16, $\dim(U \oplus W) = \dim U + \dim W$.

The idea is that if $V = U \oplus W$, then V is decomposed (broken up) into two independent bits.

This is an important technique in linear algebra, which enables us to break up a problem into simpler ones.

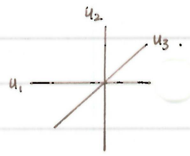
We need the analogous definition of a direct sum for more than two components.

$\boxed{\text{Definition}}$ 4.18 Let $U_1, U_2, \dots, U_n \leq V$, then the sum $\sum_{i=1}^{n} U_i = U_1 + U_2 + \cdots + U_n$ is defined as $\{\underline{u}_1 + \underline{u}_2 + \cdots + \underline{u}_n : \underline{u}_i \in U_i\}$.

then $\sum_{i=1}^{n} U_i \leq V$.

What about the directness of the sum of multiple components, such as of $U_1+U_2+U_3$?

It is not enough to take $U_i \cap U_j = \{0\}$ $\forall i \neq j$. For instance, if $V = \mathbb{R}^2$, $U_1 = \left\{\begin{pmatrix} x \\ 0 \end{pmatrix} : x \in \mathbb{R}\right\}$, $U_2 = \left\{\begin{pmatrix} 0 \\ x \end{pmatrix} : x \in \mathbb{R}\right\}$, $U_3 = \left\{\begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R}\right\}$.
but $\dim(V) = 2$



What we really want is $(U_1+U_2)+U_3$ to be direct, i.e. $(U_1+U_2) \cap U_3 = \{\underline{0}\}$.

$\boxed{\text{Definition}}$ 4.19 $\quad U_1 + \cdots + U_r$ is direct, and we write $U_1 \oplus \cdots \oplus U_r$ or $\bigoplus_{i=1}^{r} U_i$, if
$U_i \cap \left(\sum_{j \neq i} U_j\right) = \{\underline{0}\}$.

For instance, $U_1+U_2+U_3$ is direct if $U_1 \cap (U_2+U_3) = (U_1+U_2) \cap U_3 = (U_1+U_3) \cap U_2 = \{\underline{0}\}$.

$\boxed{\text{Ex}}$ $\quad V = \mathbb{R}^3$, $U_i = \{x\underline{e}_i : x \in \mathbb{R}\}$ for $i = 1,2,3$. then $U_1+U_2+U_3$ is direct.

The definition above is awkward to work with. A better condition is given by.

$\boxed{\text{Lemma}}$ 4.20 $\quad$ Consider $U_1, \dots, U_r \leq V$. Then $U_1 + \cdots + U_r$ is direct if and only if $\sum_{i=1}^{r} \underline{u}_i = \underline{0}$ $(\underline{u}_i \in U_i) \Rightarrow$ all $\underline{u}_i = \underline{0}$

Proof — ($\Rightarrow$) suppose $\sum_{i=1}^{r} U_i$ is direct. If $\sum_{i=1}^{r} \underline{u}_i = \underline{0}$, then $\underline{u}_1 = -\sum_{i=2} \underline{u}_i \in U_1 \cap \left(\sum_{i \neq 1} U_i\right) = \{\underline{0}\}$.

so $\underline{u}_1 = \underline{0}$, and in a similar way, WLOG, $\underline{u}_i = \sum_{j \neq i} \underline{u}_j \in U_i \cap \left(\sum_{j \neq i} U_j\right) = \{\underline{0}\}$, and $\underline{u}_i = \underline{0}$ $\forall i$.

($\Leftarrow$) suppose $\sum_{i=1}^{r} \underline{u}_i = \underline{0} \Rightarrow$ all $\underline{u}_i = \underline{0}$. Let $\underline{v} \in U_i \cap \left(\sum_{j \neq i} U_j\right)$. then $\underline{v} = \underline{u}_i = \sum_{j \neq i} \underline{u}_j$

$\therefore$ $\underline{u}_1 + \underline{u}_2 + \cdots + \underline{u}_{i-1} - \underline{u}_i + \underline{u}_{i+1} + \cdots + \underline{u}_r = \underline{0}$ $\quad \therefore -\underline{u}_i = \underline{0}$ and $\underline{u}_i = \underline{0}$. // q.e.d.

To prove things about directness, we almost always use $\sum \underline{u}_i = \underline{0} \Rightarrow \underline{u}_i = \underline{0}$, which is analogous to linear independence.

$\boxed{\text{Lemma}}$ 4.21 $\quad$ Let $U_i \leq V$ $(i = 1, \dots, r)$ and $\sum_{i=1}^{r} U_i$ is direct. Let $\mathcal{B}_i$ be a basis for U_i. Then

(i) $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ is a basis for $\sum_{i=1}^{r} U_i = \bigoplus_{i=1}^{r} U_i$, and

(ii) $\dim\left(\bigoplus_{i=1}^{r} U_i\right) = \sum_{i=1}^{r} \dim U_i$

Proof — (i) write $\mathcal{B}_i = \{b_1^{(i)}, \dots, b_{r_i}^{(i)}\}$. write $\bigoplus_{i=1}^{r} U_i = W$. We must prove that $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ is a basis for W.

Spanning: Let $\underline{w} \in W$. Since $W = \sum_{i=1}^{r} U_i$, $\underline{w} = \underline{u}_1 + \cdots + \underline{u}_r$ for some $\underline{u}_i \in U_i$

$= \sum_{i=1}^{r} \underline{u}_i = \sum_{i=1}^{r}\left(\sum_{j=1}^{r_i} \alpha_{ij} b_j^{(i)}\right) \in L(\mathcal{B})$ ← linear span of $\mathcal{B}$.

LI: Suppose $\sum_{i,j} \alpha_{ij} b_j^{(i)} = \underline{0}$, then $\sum_i \Big(\underbrace{\sum_j \alpha_{ij} b_j^{(i)}}_{\in U_i}\Big) = \underline{0}$. By directness, each $\sum_j \alpha_{ij} b_j^{(i)} = \underline{0}$.

Since $\mathcal{B}^i$ is a basis, all $\alpha_{ij} = 0$.

16 March 2012
Dr Mark L ROBERTS.
Darwin LT.

$\boxed{\text{Definition}}$ 4.22 $\quad$ The <u>eigenspace</u> associated to an eigenvalue λ is $E_\lambda = \{\underline{v} \in \mathbb{F}^n : A\underline{v} = \lambda \underline{v}\}$.

$\boxed{\text{Proposition}}$ 4.23 $\quad E_\lambda \leq \mathbb{F}^n$.

Proof — since $A\underline{0} = \lambda \underline{0}$, we know that $\underline{0} \in E_\lambda$. Let $\underline{v}_1, \underline{v}_2 \in E_\lambda$, $\alpha_1, \alpha_2 \in \mathbb{F}$.

$A(\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2) = \alpha_1 A\underline{v}_1 + \alpha_2 A\underline{v}_2 = \alpha_1 \lambda \underline{v}_1 + \alpha_2 \lambda \underline{v}_2 = \lambda(\alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2). \Rightarrow \alpha_1 \underline{v}_1 + \alpha_2 \underline{v}_2 \in E_\lambda$. // q.e.d.

Crucial result about eigenspaces comes as follows:

$\boxed{\text{Proposition}}$ 4.24 $\quad$ Let $\lambda_1, \lambda_2, \dots, \lambda_r$ be distinct eigenvalues of A. Then the sum $\sum_{i=1}^{r} E_{\lambda_i}$ is direct.

Proof — Suppose there is a non-trivial relation, $\sum_{i=1}^{r} \underline{u}_i = \underline{0}$ $(\underline{u}_i \in U_i)$, by contradiction. (i.e. not all $\underline{u}_i = \underline{0}$).

Choose a shortest such relation and renumber the vectors. We will get $\sum_{i=1}^{s} \underline{u}_i = \underline{0}$ —① with all $\underline{u}_i$ terms non-zero.

Left-multiplying both sides by A, $A\left(\sum_{i=1}^{s} \underline{u}_i\right) = A\underline{0} = \underline{0} \Rightarrow \sum_{i=1}^{s} A\underline{u}_i = \underline{0} \Rightarrow \sum_{i=1}^{s} \lambda_i \underline{u}_i = \underline{0}$ —② ($\because \lambda_i$ is associated eigenvalue).

② − ①λ_s: $\sum_{i=1}^{s} \lambda_i \underline{u}_i - \lambda_s \sum_{i=1}^{s} \underline{u}_i = \sum_{i=1}^{s-1} (\lambda_i - \lambda_s)\underline{u}_i = \underline{0}$. But $\lambda_i - \lambda_s \neq 0$, $\underline{u}_i \neq \underline{0}$ so $(\lambda_i - \lambda_s)\underline{u}_i \neq \underline{0}$, $(\lambda_i - \lambda_s)\underline{u}_i \in E_\lambda$.

So we get a shorter relation involving only $s-1$ terms $\Rightarrow$ contradiction $\Rightarrow$ only the trivial relation exists $\Rightarrow \sum_{i=1}^{r} E_{\lambda_i}$ is direct // q.e.d.

$\boxed{\text{Definition}}$ 4.25 $\quad$ Let A be an $n \times n$ matrix over $\mathbb{F}$, with characteristic polynomial $c_A(t) = (t-\lambda_1)^{f_1}(t-\lambda_2)^{f_2} \dots (t-\lambda_r)^{f_r}$ $(f_i \geq 1)$. Then

(i) the <u>algebraic multiplicity</u> of λ_i is f_i

(ii) the <u>geometric multiplicity</u> of λ_i is $e_i = \dim(E_{\lambda_i})$.

Note: $n = \deg(c_A) = \sum_{i=1}^{r} f_i$

this follows from property of polynomials.

1202-022.

$\boxed{\text{Theorem}}$ 4.26 A is diagonalisable $\iff e_i = f_i$

Proof – ($\Leftarrow$): From 4.24, the sum $\sum_{i=1}^{r} E_{\lambda_i}$ is direct.

Let $\mathcal{B}_i$ be a basis for E_{λ_i}, then $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ is a basis for $\bigoplus_{i=1}^{r} E_{\lambda_i}$.

Now $\dim\left(\bigoplus_{i=1}^{r} E_{\lambda_i}\right) = \sum_{i=1}^{r} \dim(E_{\lambda_i}) = \sum_{i=1}^{r} e_i = \sum_{i=1}^{r} f_i = n$.

Since $\bigoplus_{i=1}^{r} E_{\lambda_i} \leq \mathbb{F}^n$, and $\dim\left(\bigoplus_{i=1}^{r} E_{\lambda_i}\right) = \dim(\mathbb{F}^n) = n$; then $\bigoplus_{i=1}^{r} E_{\lambda_i} = \mathbb{F}^n$.

i.e. $\mathcal{B}$ is a basis for $\mathbb{F}^n$ consisting of eigenvectors. By Basis Criterion, A is diagonalisable// q.e.d.

To be continued with proof of ($\Rightarrow$), after example below.

$\boxed{\text{Ex}}$ Diagonalise $A = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{pmatrix}$.

$c_A(t) = \det(tI_3 - A) = \det\begin{pmatrix} t-3 & -1 & 0 \\ -1 & t-3 & 0 \\ 1 & -1 & t-4 \end{pmatrix} = (t-4)\det\begin{pmatrix} t-3 & -1 \\ -1 & t-3 \end{pmatrix} = (t-4)(t^2-6t+8) = (t-4)^2(t-2)$.

So $\lambda_1 = 4$, $f_1 = 2$; $\lambda_2 = 2$, $f_2 = 1$. By theorem 4.26, A is diagonalisable $\iff e_1 = 2$, $e_2 = 1$.

$\lambda_1 = 4$. $E_4 = \{\underline{v} : A\underline{v} = 4\underline{v}\} = \left\{\underline{v} : \begin{pmatrix} -1 & 1 & 0 \\ 1 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}\underline{v} = \underline{0}\right\} = \left\{\begin{pmatrix} \alpha \\ \alpha \\ \beta \end{pmatrix} : \alpha, \beta \in \mathbb{R}\right\}$

Thus E_4 has a basis $\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}\right\}$; so $e_1 = 2 = f_1$

$\lambda_2 = 2$. $E_2 = \{\underline{v} : A\underline{v} = 2\underline{v}\} = \left\{\underline{v} : \begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{pmatrix}\underline{v} = \underline{0}\right\} = \left\{\begin{pmatrix} -\alpha \\ \alpha \\ 0 \end{pmatrix} : \alpha \in \mathbb{R}\right\}$. E_2 has basis $\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\} \Rightarrow e_2 = 1 = f_2$.

$\therefore$ A is diagonalisable $\because e_i = f_i \ \forall i$. A basis for $\mathbb{R}^3$ consisting of eigenvectors is $\left\{\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}\right\}$.

Let $P = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$, then P is invertible, and $P^{-1}AP = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix}$// [order of entries match eigenvector columns in P].

Check: $\det P = -1 \det\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} = -2$, so P is invertible. $AP = \begin{pmatrix} 3 & 1 & 0 \\ 1 & 3 & 0 \\ -1 & 1 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 4 & 0 & -2 \\ 4 & 0 & 2 \\ 0 & 4 & 0 \end{pmatrix}$, $PD = \begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 2 \end{pmatrix} = \begin{pmatrix} 4 & 0 & -2 \\ 4 & 0 & 2 \\ 0 & 4 & 0 \end{pmatrix}$.

21 March 2012
Dr Mark L ROBERTS.
QLT

To continue proving other direction of Theorem 4.26, we need to introduce a lemma.

$\boxed{\text{Lemma}}$ 4.27. With notations as above, $e_i \leq f_i$.

Proof – enough to prove $e_1 \leq f_1$. Write $e = e_1$, $f = f_1$, $\lambda = \lambda_1$. ← this proof is non-examinable.

Let $\{\underline{v}_1, \dots, \underline{v}_e\}$ be a basis for E_λ. Extend to a basis $\{\underline{v}_1, \dots, \underline{v}_e, \underline{v}_{e+1}, \dots, \underline{v}_n\}$ for $\mathbb{F}^n$.

Then let $P = \begin{pmatrix} \downarrow & \downarrow & & \downarrow \\ \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{pmatrix}$. Then P is invertible as its columns form a basis (by Proposition 4.2).

$AP = A\begin{pmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_e & \underline{v}_{e+1} & \cdots & \underline{v}_n \end{pmatrix} = \begin{pmatrix} A\underline{v}_1 & A\underline{v}_2 & \cdots & A\underline{v}_e & A\underline{v}_{e+1} & \cdots & A\underline{v}_n \end{pmatrix} = \begin{pmatrix} \lambda\underline{v}_1 & \lambda\underline{v}_2 & \cdots & \lambda\underline{v}_e & * & \cdots & * \end{pmatrix}$.

$= \begin{pmatrix} \underline{v}_1 & \underline{v}_2 & \cdots & \underline{v}_n \end{pmatrix}\left(\begin{array}{cccc|c} \lambda & 0 & \cdots & 0 & \\ 0 & \ddots & & \vdots & * \\ \vdots & & \ddots & 0 & \\ 0 & \cdots & 0 & \lambda & \\ \hline 0 & 0 & \cdots & 0 & * \end{array}\right)$. Thus, $P^{-1}AP = \left(\begin{array}{c|c} \lambda I_e & X \\ \hline 0 & Y \end{array}\right) = B$ (defining matrices X, Y and B).

Then $c_B(t) = \det(tI - B) = \det\left(\begin{array}{c|c} (t-\lambda)I_e & -X \\ \hline 0 & tI_{n-e} - Y \end{array}\right) = (t-\lambda)^e g(t)$ (expanding down 1st to eth columns)

But we know also that $c_B(t) = \det(tI - B) = \det(tI - P^{-1}AP) = \det(P^{-1}(tI - A)P) = \det(P^{-1})\det(tI-A)\det(P) = \det(tI - A) = c_A(t)$. scalars

Hence, $c_A(t) = (t-\lambda)^e g(t) = (t-\lambda_1)^{f_1} \cdots (t-\lambda_r)^{f_r}$, where $\lambda \neq \lambda_2 \neq \cdots \neq \lambda_r$ $\therefore e \leq f$// q.e.d.

Returning to Theorem 4.26, (cont'd)

Proof – ($\Rightarrow$): NTP if $e_i \neq f_i \Rightarrow A$ is not diagonalisable (contrapositive).

By lemma 4.27, $e_i < f_i$, so $\sum_{j=1}^{r} e_j < \sum_{j=1}^{r} f_j = n$. $\therefore \dim\left(\bigoplus_{j=1}^{r} E_{\lambda_j}\right) = \sum e_j < n$. Thus,

since all eigenvectors are in $\bigoplus_{j=1}^{r} E_{\lambda_j}$, there cannot exist n LI eigenvectors $\Rightarrow A$ not diagonalisable by Basis Criterion q.e.d.

$\boxed{\text{Method}}$ 4.28 How to diagonalise an $n \times n$ matrix, where possible.

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

(ii) If $c_A(t)$ does not factorise into linear factors, A is not diagonalisable.

Otherwise, $c_A(t) = (t-\lambda_1)^{f_1} \cdots (t-\lambda_r)^{f_r}$.

(iii) Find a basis $\mathcal{B}$ for each eigenspace E_{λ_i}. Let $\dim(E_{\lambda_i}) = e_i$ $(1 \leq e_i \leq f_i)$

(iv) If some $e_i < f_i$, A is not diagonalisable, otherwise A is diagonalisable.

(v) Let $\mathcal{B} = \bigcup_{i=1}^{r} \mathcal{B}_i$ be a basis for $\mathbb{F}^n$

(vi) Let $P = \begin{pmatrix} \downarrow & & \downarrow \\ \underline{v}_1 & \cdots & \underline{v}_n \end{pmatrix}$. Then P is invertible, and $P^{-1}AP$ is diagonal, and

$P^{-1}AP = D = \operatorname{diag}(\underbrace{\lambda_1, \dots, \lambda_1}_{f_1 \text{ times}}, \underbrace{\lambda_2, \dots, \lambda_2}_{f_2 \text{ times}}, \dots\dots, \underbrace{\lambda_r, \dots, \lambda_r}_{f_r \text{ times}})$.

The ~~diagonal~~ minimum polynomial and the Cayley-Hamilton Theorem.

$\boxed{\text{Definition}}$ 4.29 Two matrices A and B are similar if there exists an invertible P s.t. $P^{-1}AP = B$.

$\boxed{\text{Lemma}}$ 4.30 If A and B are similar, then $c_A(t) = c_B(t)$.

Proof – done above.

So, any matrix is diagonalisable $\Leftrightarrow$ it is similar to a diagonal matrix.

In terms of linear mappings, if $T: V \to V$ is a linear mapping and $\mathcal{B}, \mathcal{B}'$ are two bases for V, with $A = M(T)^{\mathcal{B}}_{\mathcal{B}}$ and $B = M(T)^{\mathcal{B}'}_{\mathcal{B}'}$, then $B = P^{-1}AP$, i.e. A and B are similar.

$\boxed{\text{Proposition}}$ 4.31 Let $A \in M_n(\mathbb{F})$. Then there exists a non-zero polynomial $f(t) \in \mathbb{F}[t]$ s.t. $f(A) = 0$.

e.g. $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, $A^2 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$, $A^2 + I = 0 \Rightarrow f(t) = t^2 + 1$, then $f(A) = 0$.

Proof – We can think of $M_n(\mathbb{F})$ as a vector space over $\mathbb{F}$, with basis $E(i,j)$. Hence dimension of $M_n(\mathbb{F})$ is n^2.

Hence, the set $\{I, A, A^2, \dots, A^{n^2}\}$ (containing n^2+1 elements) is not linearly independent.

Say $\sum_{i=0}^{n^2} \alpha_i A^i = 0$, not all $\alpha_i = 0$. Let $f(t) = \sum_{i=0}^{n^2} \alpha_i t^i$, then $f(A) = 0$.

e.g. Take $A = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$. Then $\alpha \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \beta \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} + \gamma \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

This gives us $\alpha + 2\beta + 6\gamma = 0$, $\alpha + 4\beta + 16\gamma = 0$. One solution is $(\alpha, \beta, \gamma) = (8, -6, 1)$.

So $8 \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - 6 \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix} + \begin{pmatrix} 4 & 0 \\ 0 & 16 \end{pmatrix} = 0. \Rightarrow 8I - 6A + A^2 = 0 \Rightarrow f(t) = t^2 - 6t + 8$, then $f(A) = 0$.

Notice also that $f(t) = (t-4)(t-2) = c_A(t)$.

There will be various polynomials f s.t. $f(A) = 0$. Can we find some structure in the set $\{f(t) \in \mathbb{F}[t] : f(A) = 0\}$?

A polynomial $f(t) = a_n t^n + a_{n-1} t^{n-1} + \dots + a_0$ is called <u>monic</u> if $a_n = 1$.

Clearly, any polynomial is the product of a constant and a monic polynomial.

$\boxed{\text{Theorem}}$ 4.32 Let $A \in M_n(\mathbb{F})$. Then

(i) there exists a unique monic polynomial of least degree, $m \in \mathbb{F}[t]$, s.t. $m(A) = 0$; and

(ii) if f is such that $f(A) = 0$, then $f = m \cdot g$ for some $g(t) \in \mathbb{F}[t]$.

Then $m = m_A$ is called the <u>minimal polynomial</u> of A.

Proof – By Proposition 4.31, $\exists$ polynomials f s.t. $f(A) = 0$. Let m be a monic polynomial of least degree s.t. $m(A) = 0$.

(i) Suppose m_1, m_2 are two such monic polynomials. Let $f = m_1 - m_2$. Then $f(A) = m_1(A) - m_2(A) = 0 - 0 = 0$.

and $\deg(f) < \deg m_1$ (since m_1, m_2 are monic). Then f multiplied by a suitable constant is a monic polynomial, f'.

Yet $f'(A) = 0 \Rightarrow$ contradiction as m_1 is of least degree $\Rightarrow$ hence m is unique.

(ii) Let $f(A) = 0$. We can write $f(t) = m(t) q(t) + r(t)$, degree $r <$ degree m. Then we NTP: ~~$r(A) = 0$~~ $r = 0$;

let $t = A$, then $f(A) = m(A) q(A) + r(A) \Rightarrow 0 = 0 q(A) + r(A) \Rightarrow r(A) = 0$

Also, since $\deg r < \deg m$, it follows that $r = 0 \Rightarrow f(t) = m(t) q(t)$.

$\boxed{\text{Theorem}}$ 4.33 (Cayley-Hamilton Theorem)

Let $A \in M_n(\mathbb{F})$. Then $c_A(A) = 0$, so $m_A(t)$ divides $c_A(t)$.

e.g. if $A = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}$, $m_A(t) = (t-2)(t-4) = c_A(t)$. If $A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$, $m_A(t) = t-2$, $c_A(t) = (t-2)^2$.

if $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $c_A(t) = (t-2)^2$. Then $m_A(t)$ must be a factor of $c_A(t)$, and testing, $m_A(t) = (t-2)^2$

Proof – Omitted. (not examinable). It is quite straightforward if one assumes A is diagonalisable.

End of syllabus.

1202-024.