Vectors & Matrices

J.M. Evans @ damtp . cam...

1. Complex numbers

1.1 Definitions

Construct $\mathbb{C}$ by adding an element i to $\mathbb{R}$ with $i^2 = -1$.

Any complex number $z \in \mathbb{C}$ has form

$z = x + iy$ with $x, y \in \mathbb{R}$

$x = \text{Re}(z)$ <u>real</u> part

$y = \text{Im}(z)$ <u>imaginary</u> part

(i) <u>Addition</u> (and substraction)

Define $z_1 \pm z_2 = (x_1 + iy_1) \pm (x_2 + iy_2)$

$$= (x_1 \pm x_2) + i(y_1 \pm y_2)$$

(ii) <u>Multiplication</u>

Define $z_1 z_2 = (x_1 + iy_1) \cdot (x_2 + iy_2)$

$$= (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$$

Note: if $z \neq 0$ $(x \neq 0 \vee y \neq 0)$ then

$z^{-1} = \dfrac{x - iy}{x^2 + y^2}$ is a multiplicative inverse

For $x + iy = z$ we also define the following:

(iii) <u>Complex conjugate</u>

$\bar{z} = z^* = x - iy$

$\Rightarrow \text{Re}(z) = \frac{1}{2}(z + \bar{z})$ & $\text{Im}(z) = \frac{1}{2i}(z - \bar{z})$

(iv) <u>Modulus</u> $r = |z|$, real ≥ 0 given by

$$r^2 = |z|^2 = z\bar{z} = x^2 + y^2$$

(v) <u>Argument</u> $\theta = \arg z$, real, given by

$z = r(\cos\theta + i\sin\theta)$ $\qquad r \neq 0$

$$\Leftrightarrow \cos\theta = \frac{x}{r} \quad \wedge \quad \sin\theta = \frac{y}{r}$$

$$\Rightarrow \tan\theta = y/x$$

$\arg(z)$ is determined mod 2π

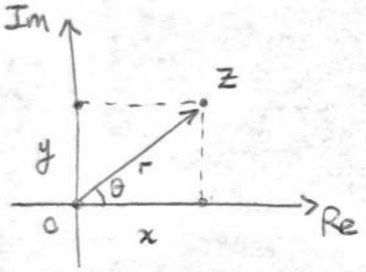
i.e. you can change $\theta \to \theta + 2n\pi$ $(n \in \mathbb{Z})$

To make it unique we can restrict to the principal value defined by $-\pi < \theta \leq \pi$ (often inconvenient)

For $z = 0$ $(r = 0)$ $\arg z$ is not defined

(vi) Argand diagram or complex plane:

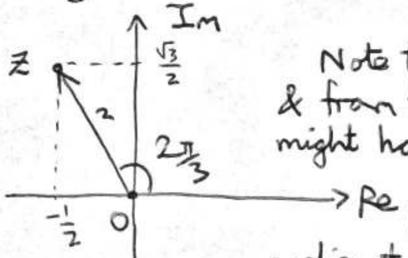
plot $\mathrm{Re}(z)$ and $I(z)$ on orthogonal axes and often draw an arrow to the point z



Then $r = |z|$ and $\theta = \arg z$ are length and angle shown.

Example $\quad z = -1 + i\sqrt{3} = 2\left(-\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)$

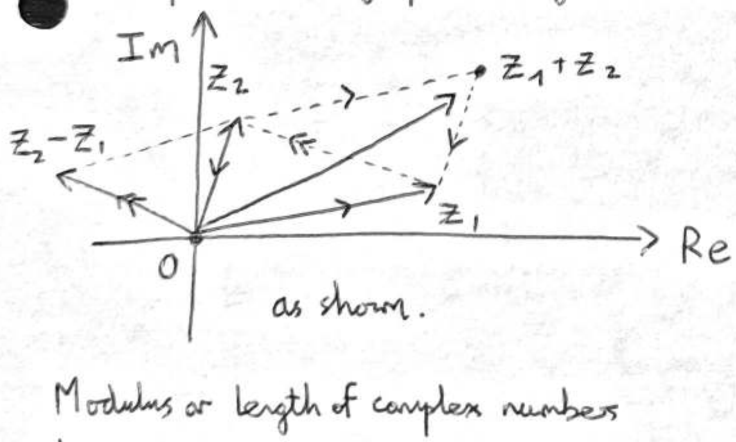
$$\Rightarrow |z| = 2 \ \& \ \arg z = 2\pi/3 + (2n\pi)$$



Note $\tan\theta = -\sqrt{3}$ & from this we might have deduced $\theta = -\pi/3$ but this applies to $-z \neq z$ $= 1 - i\sqrt{3}$

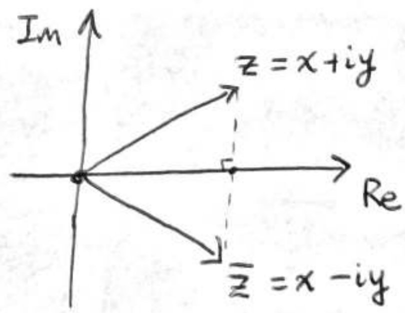
1.2 Basic properties & Consequences

Addition & subtraction of complex numbers is represented by parallelogram constructions



as shown.

Complex conjugation is equivalent to reflection in the real axis



Modulus or length of complex numbers obeys $|z_1 z_2| = |z_1||z_2|$ COMPOSITION

& the triangle inequality

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

This is clear geometrically.

To check it algebraically, square both sides:

Also $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$

$\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$

L1.3 want to show

$$(z_1+z_2)(\overline{z_1+z_2}) \leq |z_1|^2 + 2|z_1||z_2| + |z_2|^2$$

$$\Leftrightarrow \quad z_1\bar{z}_2 + z_2\bar{z}_1 \leq 2|z_1||z_2|$$

$$\Leftrightarrow \quad \tfrac{1}{2}\left(z_1\bar{z}_2 + (\overline{z_1\bar{z}_2})\right) \leq |z_1||\bar{z}_2|$$

$$\Leftrightarrow \quad \operatorname{Re}(z_1\bar{z}_2) \leq |z_1\bar{z}_2| \quad \checkmark \text{ true}$$

Alternative versions of Δ inequality:

replace z_2 by $z_2 - z_1$:

$$|z_2| \leq |z_1| + |z_2 - z_1|$$

$$\Rightarrow \quad |z_2 - z_1| \geq |z_2| - |z_1|$$

or $|z_1| - |z_2|$ by exchanging $1 \leftrightarrow 2$

so $|z_2 - z_1| \geq \big||z_2| - |z_1|\big|$.

Now consider

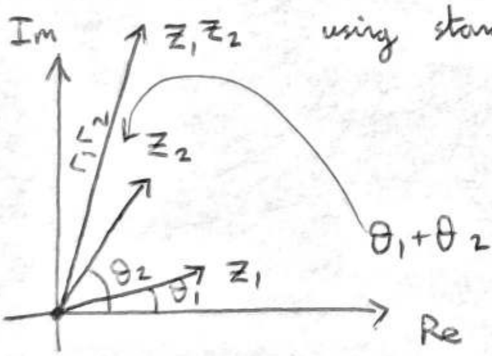
$$z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$$

& $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$

Then their product $z_1 z_2 = r_1 r_2(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2)$

$$= r_1 r_2\left(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\right)$$

using standard trig addition



$\ast$ Under multiplication of complex numbers moduli multiply & arguments add

1.3 Exponential and trigonometric functions

- Define $\exp(z) = e^z = \sum_{n=0}^{\infty} \frac{1}{n!} z^n \quad \forall z \in \mathbb{C}$

(converges $\forall z$). Multiplying series and rearranging (justification in Analysis I)

gives $e^z e^w = e^{z+w}$.

Note also $e^0 = 1$ & hence $(e^z)^n = e^{nz}$ $\quad n \in \mathbb{Z}$.

for +ve n clear

$(e^z)^{-n} = \frac{1}{(e^z)^n} = \frac{1}{e^{nz}}$

while

$e^{nz}(e^{-nz}) = e^0 = 1$

$\Rightarrow (e^z)^{-n} = e^{-zn}$

(similar to DeMoivre)

Define $\cos(z) = \frac{1}{2}(e^{iz} + e^{-iz})$

$= 1 - \frac{1}{2!} z^2 + \frac{1}{4!} z^4 - \dots$

- & $\sin(z) = \frac{1}{2i}(e^{iz} - e^{-iz})$

$= z - \frac{1}{3!} z^3 + \frac{1}{5!} z^5 - \dots$

Differentiate series for real x gives

$\frac{d}{dx}(e^x) = e^x, \quad \frac{d}{dx}(\cos x) = -\sin x, \quad \frac{d}{dx}(\sin x) = \cos x.$

But $e^0 = 1$ & $\cos 0 = 1$ & $\sin 0 = 0$.

Hence series definitions agree with geometrical definitions of cos/sin for real x.

Note $e^{iz} = \cos z + i \sin z \quad \forall z \in \mathbb{C}$

- but $\mathrm{Re}(e^{ix}) = \cos x$ for $x \in \mathbb{R}$ FUNFACT! $\mathrm{Re}(e^{iz}) = \cos z \Leftrightarrow z \in \mathbb{R}$

$\mathrm{Im}(e^{ix}) = \sin x$

Lemma: $e^z = 1 \Leftrightarrow z = 2\pi n i$ for some $n \in \mathbb{Z}$

Proof: $e^{x+iy} = e^x e^{iy} = e^x(\cos y + i \sin y) = 1$

iff $x = 0$ $(e^x = 1)$

& $y = 2n\pi$ ($\cos y = 1$ & $\sin y = 0$)

From results/comments above

modulus/argument or polar form of a general complex number (§1.1(v1)) is

- $z = r(\cos\theta + i \sin\theta) = re^{i\theta}$

$r = |z|$ & $\theta = \arg(z)$

De Moivre's Theorem

$(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for $n \in \mathbb{Z}$

follows from $(e^{i\theta})^n = e^{in\theta}$.

[or simple to prove directly using trig. properties]

Roots of unity

$z = re^{i\theta}$ satisfies $z^N = r^N e^{iN\theta} = 1$

$\iff r^N = 1 \wedge N\theta = 2n\pi$ for some $n \in \mathbb{Z}$

$\iff r = 1 \wedge \theta = \frac{2\pi}{N} n$

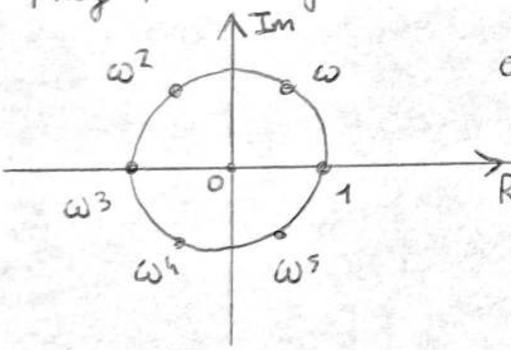
This gives N distinct solutions $z = e^{2\pi i n/N}$ $\quad n \in \{0, 1, \ldots, N-1\}$

$= \cos\frac{2\pi n}{N} + i \sin\frac{2\pi n}{N}$

$= \omega^n$ for $\omega = e^{2\pi i/N}$.

These are N^{th} roots of unity.

They form a regular N-gon



e.g. N=6

1.4 Logarithm & Complex Powers

Define $\log z$ by the condition $e^{\log z} = z \quad \forall z \in \mathbb{C}, z \neq 0$.

Consider $z = re^{i\theta} = e^{\log r} e^{i\theta} = e^{\log r + i\theta}$

$\Rightarrow \log z = \log r + i\theta$

$= \log|z| + i \arg(z)$

But arg & hence log, are multi-valued:

$\theta \to \theta + 2n\pi$

$\log z \longrightarrow \log z + 2n\pi i$ $\quad n \in \mathbb{Z}$

(Become single-valued if we restrict range e.g. $0 \leq \theta < 2\pi$)

Example $\quad z = -1 = e^{i(\pi + 2n\pi)}$

$$\Rightarrow \log(-1) = i\arg(-1) = (2n+1)\pi i$$

If we restrict $0 \leq \theta < 2\pi$ (for example)

then $\arg(-1) = \pi$

$\quad \log(-1) = i\pi$

Define complex powers by $z^\alpha = e^{\alpha \log z} \quad \forall\, (\alpha, z) \in \mathbb{C} \times \mathbb{C}\setminus\{0\}$

This ensures $z^\alpha z^\beta = z^{\alpha+\beta}$ & $(z^\alpha)^\beta = z^{\alpha\beta}$.

But these expressions are multi-valued in general, since

if $\log z \longrightarrow \log z + 2n\pi i$ $\qquad z^\alpha \to e^{\alpha(\log z + 2n\pi i)} = z^\alpha e^{2\pi i n\alpha}$

then $z^\alpha \longrightarrow z^\alpha e^{2\pi i n \alpha}$

(i) $\alpha = p \in \mathbb{Z} \Rightarrow z^p$ has a unique value

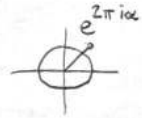
(ii) $\alpha = p/q \in \mathbb{Q} \Rightarrow z^{p/q}$ finitelly many answers

But in general we get infinitely many values

in fact $\forall\, z \in \mathbb{C}\setminus\mathbb{Q}$:

z not real give blowy-upies

z irrational loops 'n loops



Example $\quad i^i \qquad i = e^{i\pi(\frac{1}{2} + 2n)}$

$$\arg i = \pi\left(\tfrac{1}{2} + 2n\right)$$

$$\log i = i\pi\left(\tfrac{1}{2} + 2n\right)$$

$$i^i = e^{i \log i} = e^{-\pi(\frac{1}{2} + 2n)}$$

Could choose e.g. $-\pi < \theta \leq \pi$ and then $n = 0$ in expressions above.

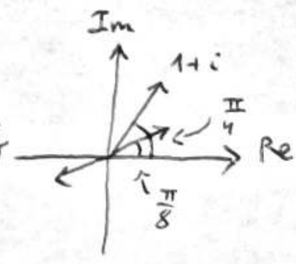
Example $\quad (1+i)^{\frac{1}{2}}: \qquad (1+i) = \sqrt{2}\, e^{i\pi(\frac{1}{4} + 2n)}$

$$\log(1+i) = \log\sqrt{2} + i\pi\left(\tfrac{1}{4} + 2n\right)$$

$$\therefore (1+i)^{\frac{1}{2}} = e^{\frac{1}{2}[\log\sqrt{2} + i\pi(\frac{1}{4} + 2n)]}$$

$$= 2^{\frac{1}{4}} e^{i\frac{\pi}{8}} e^{in\pi}$$

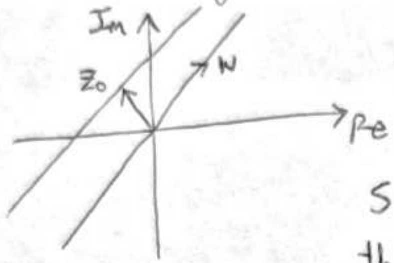
$$= \pm 2^{\frac{1}{4}} e^{i\frac{\pi}{8}}$$ depending on n's parity



L2.4!!

1.5 Lines and Circles

For a fixed $w \in \mathbb{C}$ $(\neq 0)$ the set of points $z = \lambda w$ $(\lambda \in \mathbb{R})$ is a line through the origin in direction of w as shown:



Similarly $z = z_0 + \lambda w$ is a parallel line through z_0.

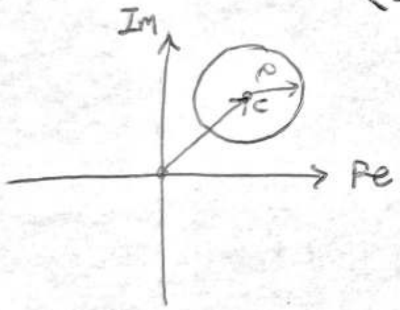
To write this in a form without the real parameter, take the conjugate

$$\bar{z} = \bar{z}_0 + \lambda \bar{w}$$ & combine to eliminate λ:

$$\bar{w} z - w \bar{z} = \bar{w} z_0 - w \bar{z}_0 .$$

The equation of a circle centre c & radius ρ is given by

$$|z-c| = \rho \iff (z-c)(\bar{z}-\bar{c}) = \rho^2$$
$$\iff z\bar{z} - c\bar{z} - \bar{c}z = \rho^2 - c\bar{c}$$



Note general point is $z = c + \rho e^{i\theta}$

2. Vectors in 3 Dimensions

A vector is a quantity with magnitude (length, size) & direction. ← obscure, ignores abstract vector spaces

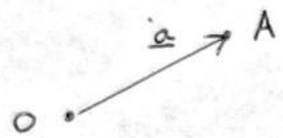
Examples: my ass

But all modelled on position.

Start with geometrical approach to position vectors in 3d space or 2d plane with standard (Euclidean) length and angle.

Pick point O as origin, then any point A has position vector

$\underline{a} = \overrightarrow{OA}$



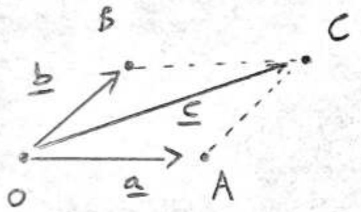
Length $|\underline{a}| = |\overrightarrow{OA}|$

write $\underline{O}$ for position vector of O

Note $|\underline{a}| = 0 \iff \underline{a} = \underline{O}$

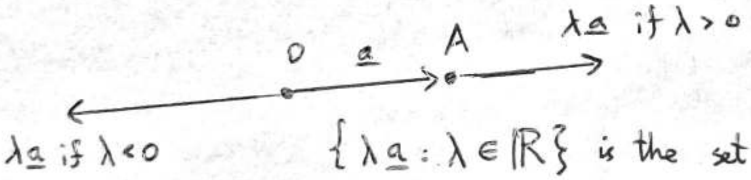
§2.1 Vector addition and Scalar multiplication

(i) Addition Given $\underline{a}, \underline{b}$ position vectors of points A, B, construct parallelogram $OACB$ as shown.



then $\underline{a} + \underline{b} = \underline{c}$ by def^n

(ii) Scalar multiplication Given $\underline{a}$ and $\lambda \in \mathbb{R}$, $\lambda\underline{a}$ is a vector with length $|\lambda\underline{a}| = |\lambda||\underline{a}|$ in direction of $\underline{a}$ as shown



$\{\lambda\underline{a} : \lambda \in \mathbb{R}\}$ is the set of all points on line through O and A

(iii) Properties

$$\underline{a} + \underline{O} = \underline{a}$$

$$\underline{a} + (-\underline{a}) = \underline{O}$$

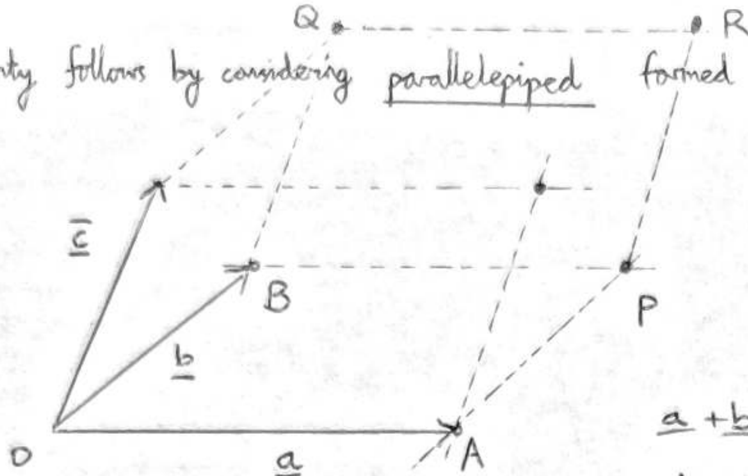
(such $-\underline{a}$ exists $\forall\, \underline{a}$)

$$\underline{a} + \underline{b} = \underline{b} + \underline{a} \quad \text{commutativity}$$

$$\underline{a} + (\underline{b} + \underline{c}) = (\underline{a} + \underline{b}) + \underline{c} \quad \text{associativity}$$

[vectors under addition are an abelian group]

Associativity follows by considering <u>parallelepiped</u> formed by $\underline{a}, \underline{b}, \underline{c}$ as shown



$\underline{a} + \underline{b} = \underline{p}$

and $\underline{p} + \underline{c} = \underline{r}$

since OPRC is parallelogram

$\underline{b} + \underline{c} = \underline{q}$

$\underline{a} + \underline{q} = \underline{r}$

since OARQ is parallelogram

(iv) <u>Linear Combinations and Span</u>

Vectors $\underline{a}$ and $\underline{b}$ parallel, written $\underline{a} \parallel \underline{b} \Leftrightarrow \underline{a} = \lambda \underline{b}$ for some $\lambda \in \mathbb{R}$.
or $\underline{b} = \lambda \underline{a}$

We allow
- $\lambda = 0$, so $\underline{0} \parallel$ any vector
- $\lambda < 0$, so we do <u>not</u> distinguish anti-parallel

An expression
$\alpha \underline{a} + \beta \underline{b}$ is called a <u>linear combination</u>
of the vectors $\underline{a}$ and $\underline{b}$.

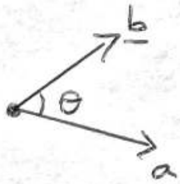
The set of all vectors $\{\alpha \underline{a} + \beta \underline{b} : \alpha, \beta \in \mathbb{R}\}$ is called the <u>span</u> of $\underline{a}$ and $\underline{b}$.

If $\underline{a} \nparallel \underline{b}$ (neither is 0) then this span is a plane containing O, A, B.

<u>§2.2 Scalar or dot product</u>

Given $\underline{a}$ & $\underline{b}$, let θ be the angle between them (in the plane the span)

as shown



& define $\underline{a} \cdot \underline{b} = |\underline{a}||\underline{b}|\cos\theta$

(also known as <u>inner</u> product)

<u>Properties</u>

$$\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a}$$

$$\underline{a} \cdot \underline{a} = |\underline{a}|^2 \geq 0 \quad \text{with equality iff } \underline{a} = \underline{0}$$

$$(\lambda \underline{a}) \cdot \underline{b} = \lambda(\underline{a} \cdot \underline{b}) = \underline{a} \cdot (\lambda \underline{b})$$

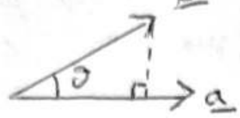
Also $\underline{a} \cdot (\underline{b} + \underline{c}) = \underline{a} \cdot \underline{b} + \underline{a} \cdot \underline{c}$

L3.3

Resolving & Components

For $\underline{a} \neq \underline{0}$

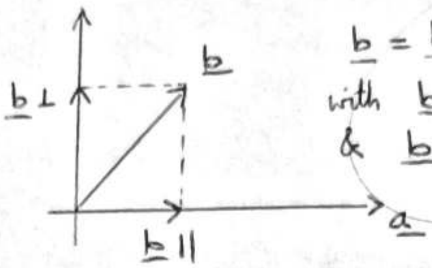
$$\frac{\underline{a}\cdot\underline{b}}{|\underline{a}|} = |\underline{b}|\cos\theta$$



is the component of $\underline{b}$ along $\underline{a}$

(we are resolving along $\underline{a}$)

This allows us to write



$\underline{b} = \underline{b}_\parallel + \underline{b}_\perp$

with $\underline{b}_\parallel \parallel \underline{a}$

& $\underline{b}_\perp \perp \underline{a}$

construct using

$$\underline{b}_\parallel = \frac{\underline{a}\cdot\underline{b}}{|\underline{a}|}\cdot\frac{\underline{a}}{|\underline{a}|}$$

then

$$(\underline{b} - \underline{b}_\parallel)\cdot\underline{a} \overset{?}{=} 0$$

actually not clear :/ I mean in $\mathbb{R}^3$ it's pretty obvious

Then $\underline{a}\cdot\underline{b} = \underline{a}\cdot\underline{b}_\parallel$

Now $(\underline{b}+\underline{c})_\parallel = \underline{b}_\parallel + \underline{c}_\parallel$ & $(\underline{b}+\underline{c})_\perp = \underline{b}_\perp + \underline{c}_\perp$.

Hence $\underline{a}\cdot(\underline{b}_\parallel + \underline{c}_\parallel) = a\cdot(\underline{b}+\underline{c})_\parallel$

$\Rightarrow \underline{a}\cdot(\underline{b}+\underline{c}) = \underline{a}\cdot\underline{b} + \underline{a}\cdot\underline{c}$

since $\underline{a}\cdot\underline{b}_\parallel + \underline{a}\cdot\underline{c}_\parallel = \underline{a}\cdot(\underline{b}_\parallel + \underline{c}_\parallel)$.

Orthogonality

Say $\underline{a}$ & $\underline{b}$ are perpendicular/orthogonal iff

$\underline{a}\cdot\underline{b} = 0$ & ~~[illegible]~~ write $\underline{a}\perp\underline{b}$

This includes $\underline{a}$ or $\underline{b} = \underline{0}$, so $\underline{0}$ is $\perp$ to any vector with this definition.

2.3 Orthonormal Bases

Choose vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ which are perpendicular & have length 1

or $\underline{e}_i\cdot\underline{e}_j = \delta(i,j)$ Kronecker delta

equivalent to choosing perpendicular cartesian axes along these directions.

Any vector can be written $\underline{a} = \sum_i a_i \underline{e}_i$

& each component is uniquely determined by

$a_i = \underline{e}_i\cdot\underline{a}$ since $\cdot$ with the rest is 0

Often convenient to identify the vector with this list of components

$$\underline{a} = (a_1, a_2, a_3) \text{ or } \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$$

row vector column vector

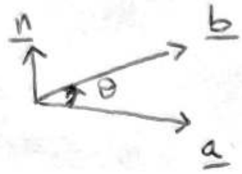
$|\underline{a}|^2 = a_1^2 + a_2^2 + a_3^2$ (Pythagoras)

Now $\underline{a}\cdot\underline{b} = \sum_{i,j} a_i b_j \underline{e}_i\underline{e}_j = \sum_i a_i b_i$ ☺

L4.1

2.4 Vector or Cross Product

Given $\underline{a}$ and $\underline{b}$, let θ be the angle in the plane measured in sense shown, relative to a unit vector $\underline{n}$ normal to plane



"right-handed sense"

Define $\underline{a} \wedge \underline{b}$ (or $\underline{a} \times \underline{b}$)

as $|\underline{a}||\underline{b}| \sin\theta \, \underline{n}$.

If $\underline{a} \parallel \underline{b}$, then $\underline{n}$ is not uniquely defined but $\sin\theta = 0$ and $\underline{a} \wedge \underline{b} = \underline{0}$ in any case.

Properties

$$\underline{a} \wedge \underline{b} = -\underline{b} \wedge \underline{a}$$

$$(\lambda \underline{a}) \wedge \underline{b} = \underline{a} \wedge (\lambda \underline{b}) = \lambda (\underline{a} \wedge \underline{b})$$

$$\underline{a} \wedge (\underline{b} + \underline{c}) = \underline{a} \wedge \underline{b} + \underline{a} \wedge \underline{c}$$

(return to this below)

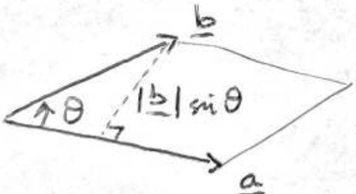
$\underline{a} \wedge \underline{b} = \underline{0}$ iff $\underline{a} \parallel \underline{b}$

$\underline{a} \wedge \underline{b} \perp \underline{a}$ & $\underline{b}$ by def^n

so $\underline{a} \cdot (\underline{a} \wedge \underline{b}) = \underline{b} \cdot (\underline{a} \wedge \underline{b}) = 0$

Interpretations

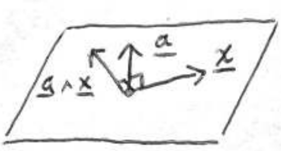
(i) $\underline{a} \wedge \underline{b}$ is vector area of parallelogram shown:



$$|\underline{a} \wedge \underline{b}| = \underbrace{|\underline{a}|}_{\text{base}} \underbrace{|\underline{b}| \sin\theta}_{\text{height}} \qquad (\sin\theta \geq 0)$$

area, & direction of $\underline{n}$ specifies orientation of ▱ in 3d space

(ii) Fix $\underline{a}$ & consider any $\underline{x} \perp \underline{a}$.



Then $\underline{x} \to \underline{a} \wedge \underline{x}$ scales $\underline{x}$ by $|\underline{a}|$ since $|\underline{a} \wedge \underline{x}| = |\underline{a}||\underline{x}|$ & rotates $\underline{a}$ by $90°$ as shown.

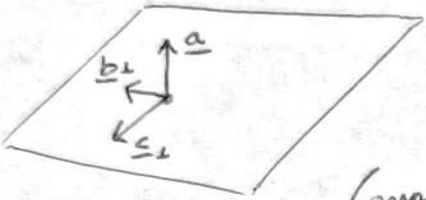
Note: use this to show $\underline{a} \wedge (\underline{b} + \underline{c}) = \underline{a} \wedge \underline{b} + \underline{a} \wedge \underline{c}$

by first writing $\underline{b} = \underline{b}_{\parallel} + \underline{b}_{\perp}$ and $\underline{c} = \underline{c}_{\parallel} + \underline{c}_{\perp}$

& observing $\underline{a} \wedge \underline{b} = \underline{a} \wedge \underline{b}_{\perp}$.

Then we need to show $\underline{a} \wedge (\underline{b} + \underline{c})_{\perp} = \underline{a} \wedge \underline{b}_{\perp} + \underline{a} \wedge \underline{c}_{\perp}$

but this can be understanded as so:



geometrically by scaling & rotating the parallelogram involved in addition

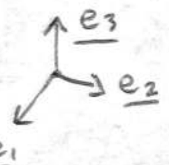
Component Expressions

Consider orthonormal basis vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ (as before)

but now satisfying, in addition

$$\underline{e}_1 \wedge \underline{e}_2 = \underline{e}_3 = -\underline{e}_2 \wedge \underline{e}_1$$

This is a right-handed orthonormal set.



For $\underline{a} = \sum_i a_i \underline{e}_i$, $\underline{b} = \sum_i b_i \underline{e}_i$

we can compute

$$\underline{a} \wedge \underline{b} = \sum_{ij} a_i b_j \, \underline{e}_i \wedge \underline{e}_j = (a_2 b_3 - a_3 b_2)\underline{e}_1 + (a_3 b_1 - a_1 b_3) e_2 + (a_1 b_2 - a_2 b_1)\underline{e}_3$$

Also use $\underline{i}, \underline{j}, \underline{k}$ for such basis vectors, or write vectors as (a_1, a_2, a_3).

Example $\underline{a} = (2, 0, -1)$ $\underline{b} = (7, -3, 5)$

$$\underline{a} \wedge \underline{b} = \begin{vmatrix} \underline{e}_1 & \underline{e}_2 & \underline{e}_3 \\ 2 & 0 & -1 \\ 7 & -3 & 5 \end{vmatrix} = (-3\underline{e}_1 - 17\underline{e}_2 - 6\underline{e}_3) = (-3, -17, -6)$$

We can check that $\underline{a} \cdot (\underline{a} \wedge \underline{b}) = \underline{b} \cdot (\underline{a} \wedge \underline{b}) = 0$.

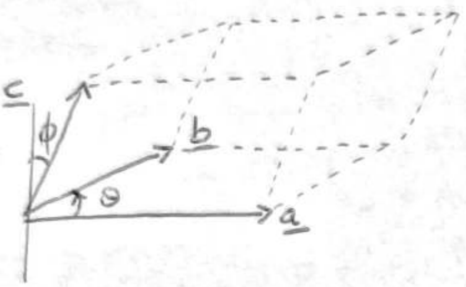
L4.3

2.5 Triple products

(a) Scalar triple product

$$\underline{a}\cdot(\underline{b}\wedge\underline{c}) = \underline{b}\cdot(\underline{c}\wedge\underline{a}) = \underline{c}\cdot(\underline{a}\wedge\underline{b})$$
$$= -\underline{a}\cdot(\underline{c}\wedge\underline{b}) = -\underline{b}\cdot(\underline{a}\wedge\underline{c}) = -\underline{c}\cdot(\underline{b}\wedge\underline{a})$$ noice

<u>Interpretation</u>



$$|\underline{c}\cdot(\underline{a}\wedge\underline{b})| = |\underline{c}||\underline{a}\wedge\underline{b}|\cos\phi$$ ← should be

$$|\underline{a}\cdot(\underline{b}\wedge\underline{c})| = |\underline{a}||\underline{b}\wedge\underline{c}|\cos\phi$$ ← mismp :(

= volume of parallelepiped

= (area of base)($\perp$ height)

$\underline{a}\cdot(\underline{b}\wedge\underline{c}) = 0$ iff $\underline{a}, \underline{b}, \underline{c}$ are <u>coplanar</u> meaning <u>one</u> of them is a linear combination of the others e.g. $\underline{c} = \alpha\underline{a} + \beta\underline{b}$.

$\underline{a}\cdot\underline{b}\wedge\underline{c}$ is a "signed" volume

$\underline{a}\cdot\underline{b}\wedge\underline{c} > 0$ say $\underline{a}, \underline{b}, \underline{c}$ are a right-handed set of vectors

(b) Vector triple products

Repeated vector products can be simplified:

$$\underline{a}\wedge(\underline{b}\wedge\underline{c}) = (\underline{a}\cdot\underline{c})\underline{b} - (\underline{a}\cdot\underline{b})\underline{c}$$

$$(\underline{a}\wedge\underline{b})\wedge\underline{c} = (\underline{a}\cdot\underline{c})\underline{b} - (\underline{b}\cdot\underline{c})\underline{a}$$

(second eq^{n} follows from first)

Note cross product is not associative

This can be proved by brute force using components or more efficiently by combining this with summation convention.

Note, however, $\underline{a}\wedge(\underline{b}\wedge\underline{c}) \perp \underline{b}\wedge\underline{c}$

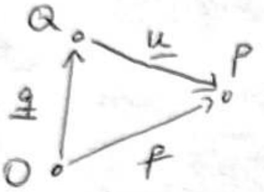
$\Rightarrow \underline{a}\wedge(\underline{b}\wedge\underline{c}) = \beta\underline{b} + \gamma\underline{c}$

& $\underline{a}\cdot(\underline{a}\wedge(\underline{b}\wedge\underline{c})) = 0 \Rightarrow \beta(\underline{a}\cdot\underline{b}) + \gamma(\underline{a}\cdot\underline{c}) = 0$

L4.4:

2.6 Lines, Planes & Vector Equations

Vectors were introduced as position vectors from origin O, but definitions of addition etc. mean we can also use them to describe displacements between

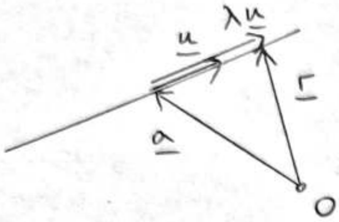


$\underline{u} = \overrightarrow{QP} = \underline{p} - \underline{q}$

L5.1

(a) Lines

General point on a line through $\underline{a}$ with direction given by $\underline{u}\ (\neq \underline{0})$ has position vector $\underbrace{\underline{r} = \underline{a} + \lambda \underline{u}}_{\text{Parametric form}}$ for $\lambda \in \mathbb{R}$.



Alternative form, without parameter, obtained by crossing with $\underline{u}$

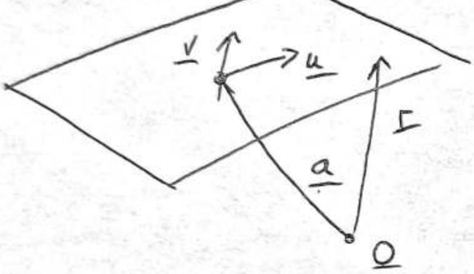
$$\underline{u} \wedge (\underline{r} - \underline{a}) = \underline{0}.$$

Conversely, this holds iff $\underline{u} \parallel (\underline{r} - \underline{a}) \iff (\underline{r} - \underline{a}) = \lambda \underline{u}$ for some λ in $\mathbb{R}$, since $\underline{u} \neq \underline{0}$. Hence two forms are equivalent.

(b) Planes

General point on a plane through $\underline{a}$ with directions $\underline{u}, \underline{v}$ in plane ($\underline{u} \nparallel \underline{v}$) has position vector $\underline{r} = \underline{a} + \lambda \underline{u} + \mu \underline{v}$ for some $\lambda, \mu \in \mathbb{R}$.

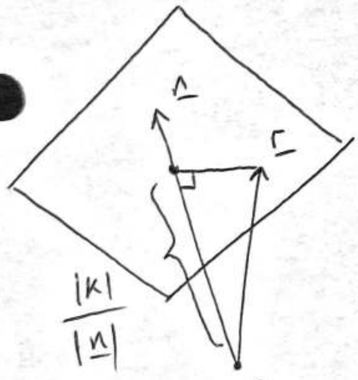
parametric form



Alternative form without parameters obtained by dotting with normal $\underline{n} = \underline{u} \wedge \underline{v}$, giving

$$\underline{n} \cdot (\underline{r} - \underline{a}) = 0$$

or $\underline{n} \cdot \underline{r} = \kappa$ where $\kappa = \underline{n} \cdot \underline{a} = \text{const}$.



Component of $\underline{r}$ along $\underline{n}$ is $\dfrac{\underline{n} \cdot \underline{r}}{|\underline{n}|} = \dfrac{\kappa}{|\underline{n}|}$.

clearly a plane.

Moreover, $\dfrac{|\kappa|}{|\underline{n}|}$ is perp. distance of plane from $\underline{O}$.

(c) Vector Equations

Start with an example:

$$\underline{u} \wedge \underline{r} = \underline{c} \quad \text{where } \underline{u}, \underline{c} \text{ are const.} \ \& \ \underline{u} \neq \underline{0}.$$

Note $\underline{u} \cdot \underline{u} \wedge \underline{r} = \underline{u} \cdot \underline{c}$ and $\underline{u} \cdot \underline{u} \wedge \underline{r} = 0$

so eq^n is inconsistent unless $\underline{u} \cdot \underline{c} = 0$.

In this case, try a particular solution by considering

$$\underline{u} \wedge (\underline{u} \wedge \underline{c}) = \cancel{(\underline{u}\cdot\underline{c})}^{\,0}\,\underline{u} - (\underline{u}\cdot\underline{u})\underline{c} = -|\underline{u}|^2 \underline{c}$$

Hence if $\underline{a} = -\frac{1}{|u|^2} \cancel{x}\,(\underline{u}\wedge\underline{c})$ is a particular solution : $\underline{u} \wedge \underline{a} = \underline{c}$

General solution: add $\lambda \underline{u}$ since this is most general vector with

$\underline{u} \wedge (\lambda \underline{u}) = \underline{0}$; this gives $\underline{r} = \underline{a} + \lambda \underline{u}$

In general, we can attempt to solve vector equations by dotting or crossing with appropriate constant vectors.

Example

$$\left.\begin{aligned} \underline{r} + \underline{a} \wedge (\underline{b} \wedge \underline{r}) &= \underline{c} \\ \text{or} \quad \underline{r} + (\underline{a}\cdot\underline{r})\underline{b} - (\underline{a}\cdot\underline{b})\underline{r} &= \underline{c} \end{aligned}\right\} (1)$$

Dot with $\underline{a}$: $\underline{a}\cdot\underline{r} = \underline{a}\cdot\underline{c}$ (2)

(equation of a plane)

substitute into (1) to get

$$(1 - \underline{a}\cdot\underline{b})\underline{r} = \underline{c} - (\underline{a}\cdot\underline{c})\underline{b}$$

If $\underline{a}\cdot\underline{b} \neq 1$ we get $\underline{r} = \dfrac{1}{1-\underline{a}\cdot\underline{b}}\left(\underline{c} - (\underline{a}\cdot\underline{c})\underline{b}\right)$. (point)

If $\underline{a}\cdot\underline{b} = 1$ then (3) is inconsistent (no sol^n) unless $\underline{c} = (\underline{a}\cdot\underline{c})\underline{b}$,

in which case (1) becomes $(\underline{a}\cdot\underline{r} - \underline{a}\cdot\underline{c})\,\underline{b} = \underline{0}$

which follows from (2), so solution is

$$\underline{a}\cdot\underline{r} = \underline{a}\cdot\underline{c} \quad \text{(plane)}.$$

2.7 Index (suffix) notation and Summation Convention

(a) Components; δ & ε

Write vectors $\underline{a}, \underline{b}, \ldots$ in terms of components $a_i, b_i, \ldots$ w.r.t. basis

$\underline{\hat{e}}_1, \underline{e}_2, \underline{e}_3$ (right handed orthonormal)

Indices or suffices $i, j, k, l, p, q, \ldots$ take all values $1, 2, 3$.

Then e.g. $\underline{c} = \alpha \underline{a} + \beta \underline{b}$

$\Leftrightarrow c_i = \alpha a_i + \beta b_i$ true for $i = 1, 2, 3$

a free index

$$\underline{a} \cdot \underline{b} = \sum_i a_i b_i$$

$$= \sum_j a_j b_j$$

$$\underline{x} = \underline{a} + (\underline{b} \cdot \underline{c}) \underline{d}$$

$\Leftrightarrow x_j = a_j + \sum_k b_k c_k d_j$ for $j = 1, 2, 3$ free index

Define Kronecker delta

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \qquad \delta_{ij} = \delta_{ji} \quad \text{symmetric}$$

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

Note $\underline{e}_i \cdot \underline{e}_j = \delta_{ij}$

Also $\underline{a} \cdot \underline{b} = \left(\sum_i a_i \underline{e}_i \right) \left(\sum_j b_j \underline{e}_j \right)$

$$= \sum_{j,i} a_i b_j \underline{e}_i \cdot \underline{e}_j$$

$$= \sum_{i,j} a_i b_j \delta_{ij}$$

$$= \sum_i a_i b_i$$

Define Levi-Civita epsilon

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{if } (i,j,k) \text{ even permutation of } (1,2,3) \\ -1 & \text{if } \quad \text{"} \quad \text{odd} \quad \text{"} \quad \text{"} \quad \text{"} \\ 0 & \text{otherwise} \end{cases}$$

i.e. $\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = +1$

$\varepsilon_{132} = \varepsilon_{321} = \varepsilon_{213} = -1$

& $\varepsilon_{ijk} = 0$ otherwise

LS.4

ε_{ijk} is totally antisymmetric:

$$\varepsilon_{ijk} = \varepsilon_{jki} = \varepsilon_{kij} = -\varepsilon_{ikj} = -\varepsilon_{kji} = -\varepsilon_{jik}$$

Note $\underline{e}_i \wedge \underline{e}_j = \sum_k \varepsilon_{ijk} \underline{e}_k$

e.g. $\underline{e}_2 \wedge \underline{e}_{\cancel{x}1} = \sum_k \varepsilon_{21k} \underline{e}_k = \varepsilon_{213} \underline{e}_3 = -\underline{e}_3$

Now consider

$$\underline{a} \wedge \underline{b} = \left(\sum_i a_i \underline{e}_i \right) \wedge \left(\sum_j b_j \underline{e}_j \right)$$

$$= \sum_{j,i} a_i b_j \, \underline{e}_i \wedge \underline{e}_j$$

$$= \sum_{i,j,k} a_i b_j \, \varepsilon_{ijk} \underline{e}_k$$

Or $(\underline{a} \wedge \underline{b})_k = \sum_{i,j} \varepsilon_{ijk} a_i b_j$

$$(a \times b) \cdot (\overset{a}{\cancel{b}} \times c) = (a \times b)_i (\overset{a}{\cancel{b}} \times c)_i$$

$$= \varepsilon_{ijk} a_j b_k \varepsilon_{ipq} \overset{a}{\cancel{b}}_p c_q$$

$$= (\delta_{jp} \delta_{kq} - \delta_{jq} \delta_{kp})(a_j b_k \overset{a}{\cancel{b}}_p c_q)$$

$$= a_p a_p b_q c_q - a_q b_p a_p c_q$$

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

Notice

(b) Summation convention

In expressions above we notice that whenever an index is repeated it is summed over. In the Einstein summation convention we omit the $\sum$ signs for repeated indices: the sum is then understood.

Examples (i) $a_i \delta_{ij} = a_1 \delta_{1j} + a_2 \delta_{2j} + a_3 \delta_{3j}$ ($\sum_i$ understood)

$$= \begin{cases} a_1 \text{ if } j=1 \\ a_2 \text{ if } j=2 \\ a_3 \text{ if } j=3 \end{cases} \quad \text{i.e. } a_i \delta_{ij} = a_j \quad \text{for } j=1,2,3$$

(ii) $\underline{a} \cdot \underline{b} = \delta_{ij} a_i b_j = a_i b_i$

(iii) $(\underline{a} \wedge \underline{b})_i = \varepsilon_{ijk} a_j b_k$ ($\sum_{jk}$ understood)

(iv) $\underline{a} \cdot (\underline{b} \wedge \underline{c}) = \delta_{pq} a_p (\underline{b} \wedge \underline{c})_q$

$$= a_q \, \varepsilon_{qjk} b_j c_k$$

$$= \varepsilon_{ijk} a_i b_j c_k$$

($\sum_{ijk}$ understood)

(v) $\delta_{ii} = \delta_{11} + \delta_{22} + \delta_{33} = 3 \neq 1$!

More precisely, we follow these rules:

(i) An index occurring exactly once must appear in every term in an equation and can take any value - free index

(ii) An index occurring exactly twice in a given term is summed over - repeated, contracted or dummy index

(iii) No index can occur more than twice in a term

Example $[\underline{a} \wedge (\underline{b} \wedge \underline{c})]_i = \varepsilon_{ijk} a_j (\underline{b} \wedge \underline{c})_k$

$$= \varepsilon_{ijk} a_j \varepsilon_{kpq} b_p c_q$$

$$= \varepsilon_{kij} \varepsilon_{kpq} a_j b_p c_q \quad (\textstyle\sum_{jkpq} \text{ understood})$$

Claim: $\varepsilon_{kij} \varepsilon_{kpq} = \delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp}$ $(*)$

$(*)$ implies $[\underline{a} \wedge (\underline{b} \wedge \underline{c})]_i = (\delta_{ip} \delta_{jq} - \delta_{iq} \delta_{jp}) a_j b_p c_q$

$$= \delta_{jq} a_j c_q \delta_{ip} b_p - \delta_{jp} a_j b_p \delta_{iq} c_q = (a_j c_j) b_i - (a_j b_j) c_i = [(\underline{a} \cdot \underline{c}) \underline{b} - (\underline{a} \cdot \underline{b}) \underline{c}]_i$$

L6.2 True for $i=1,2,3$ hence $\underline{a}\wedge(\underline{b}\wedge\underline{c}) = (\underline{a}\cdot\underline{c})\underline{b} - (\underline{a}\cdot\underline{b})\underline{c}$.

(c) $\varepsilon\varepsilon$ identities

The following can be checked by considering all possible index values, but simplified using symmetry

- $$\varepsilon_{ijk}\varepsilon_{pqr} = \delta_{ip}\delta_{jq}\delta_{kr} - \delta_{jp}\delta_{kr}\delta_{iq} + \delta_{jp}\delta_{kq}\delta_{ir} - \delta_{kp}\delta_{jq}\delta_{ir} + \delta_{kp}\delta_{iq}\delta_{jr} - \delta_{ip}\delta_{kq}\delta_{jr}$$

  LHS gives relative parity

  only one term on the right is non-zero: sign gives rel. parity

  Total antisymmetry $\Rightarrow$ assume $\{i,j,k\} = \{p,q,r\} = \{1,2,3\}$

  we will not see this often

- $$\varepsilon_{ijk}\varepsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp}$$

<u>check:</u> RHS & LHS are both anti-symmetric

(change sign under $i \leftrightarrow j$, $p \leftrightarrow q$)

so both vanish if $i=j$ or $p=q$.

From $\sum_k$ implicit on LHS it suffices to check $i=p=1$, $j=q=2$

$$\text{LHS} = \varepsilon_{123}\varepsilon_{123} = +1$$
$$\text{RHS} = \delta_{11}\delta_{22} - \delta_{12}\delta_{21} = +1 \checkmark$$

or $i=q=1$, $j=p=2$

$$\text{LHS} = \varepsilon_{123}\varepsilon_{213} = -1$$
$$\text{RHS} = \delta_{12}\delta_{21} - \delta_{11}\delta_{22} = -1 \checkmark$$

- $$\varepsilon_{ijk}\varepsilon_{pjk} = 2\delta_{ip} \qquad \delta_{jj} = 3!$$

All follow from contractions of the first, most general identity.

$$\varepsilon_{ijk}\varepsilon_{pqk} = \varepsilon_{ijk}\varepsilon_{pqr}\delta_{kr} = \delta_{ip}\delta_{jq}\delta_{kk} - \delta_{jp}\delta_{iq}\delta_{kk} + \delta_{jp}\delta_{iq}\cancel{\delta_{kr}} - \delta_{ip}\delta_{jq}\cancel{\delta_{kr}} + \cancel{\delta_{kp}}\delta_{iq}\delta_{jp} - \delta_{ip}\delta_{jq}\cancel{\delta_{jr}}$$

L6.3

3. Vectors in General; $\mathbb{R}^n$ & $\mathbb{C}^n$

3.1 $\mathbb{R}^n$ & Real Vector Spaces

If we regard a vector as a set of components, we can easily generalise from 3 to n dimensions. Let $\mathbb{R}^n = \{ \underline{x} = (x_1, x_2, \dots, x_n) : x_i \in \mathbb{R} \}$ & define

(i) <u>Addition</u> $\underline{x} + \underline{y} = (x_1 + y_1, \dots, x_n + y_n)$

& (ii) <u>Scalar multiplication</u> $\lambda \underline{x}\ (\lambda x_1, \dots, \lambda x_n)$

for any $\underline{x}, \underline{y} \in \mathbb{R}^n$ & $\lambda \in \mathbb{R}$.

These operations share certain algebraic properties with the operations we defined geometrically in 3d.

<u>Axiomatic approach</u> Let V be a set of objects, <u>vectors</u>, $\underline{v}, \underline{w} \in V$ with operations defined (i) $\underline{v} + \underline{w} \in V$ addition

(ii) $\lambda \underline{v} \in V$ scalar multiplication

Then V is called a (real) <u>vector</u> <u>space</u> if $V, +$ is an abelian group, and

$$\lambda(\underline{v} + \underline{w}) = \lambda \underline{v} + \lambda \underline{w}$$
$$(\lambda + \mu)\underline{v} = \lambda \underline{v} + \mu \underline{v}$$
$$\lambda(\mu \underline{v}) = (\lambda\mu)\underline{v}$$
$$1\underline{v} = \underline{v}$$

These axioms apply to our geometrical vectors in 3d & to $\mathbb{R}^n$ as above but lots of other cases too.

<u>Example</u> Let $V = \{ f : [0,1] \to \mathbb{R} : f \text{ is smooth } \& \ f(0) = f(1) = 0 \}$.

Define $(f+g)(x) = f(x) + g(x)$ & $(\lambda f)(x) = \lambda f(x)$, and then can check axioms.

<u>(a) Subspaces & Span</u> A subset $U \subseteq V$ that is also a vector space (with operations inherited from V) is called a <u>subspace</u>. A necessary & sufficient condition is that $\underline{v}, \underline{w} \in U \Rightarrow \lambda \underline{v} + \mu \underline{w} \in U$ for any $\lambda, \mu \in \mathbb{R}$.

L6.4

Given vectors $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ in a vector space V then span is defined by

$$\text{span}\{\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n\} = \left\{ \sum_i \lambda_i \underline{v}_i : \lambda_i \in \mathbb{R} \right\}$$ this is a subspace.

L7.1 It consists of all linear combinations of $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_n}$
[subscripts are labels, not indices]

In 3d or $\mathbb{R}^3$ any line or plane through O is a subspace. However, a line or plane that does not contain O is not a subspace [need $\underline{0}$ in subspace].

e.g. $\underline{v_1} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ & $\underline{v_2} = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ then $\underline{r} = \lambda_1 \underline{v_1} + \lambda_2 \underline{v_2}$,

a general point on plane given by span$\{v_1, v_2\}$

$= \{\underline{r} : \underline{n} \cdot \underline{r} = 0\}$ for $\underline{n} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ normal to plane.

But $\underline{r} = \underline{a} + \lambda_1 \underline{v_1} + \lambda_2 \underline{v_2}$ with $\underline{a} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$ then this is not a subspace. It is a plane parallel to plane above given by $\{\underline{r} : \underline{n} \cdot \underline{r} = 3\}$.

Similarly in $\mathbb{R}^n$: given $\alpha_1, \ldots, \alpha_n$

$$U = \{ \underline{x} \in \mathbb{R}^n : \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n = 0 \}$$

is a subspace. Check $\underline{x}, \underline{y} \in U \Rightarrow \lambda \underline{x} + \mu \underline{y} \in U$ since

$$\lambda \underline{x} + \mu \underline{y} = \begin{pmatrix} (\lambda \underline{x} + \mu \underline{y})_1 \\ \vdots \\ (\lambda \underline{x} + \mu \underline{y})_n \end{pmatrix} = \begin{pmatrix} \lambda x_1 + \mu y_1 \\ \vdots \\ \lambda x_n + \mu y_n \end{pmatrix}$$

$$\sum_i \alpha_i (\lambda x_i + \mu y_i) = \lambda \sum_i \alpha_i x_i + \mu \sum_i \alpha_i y_i = 0.$$

However

$$\{ \underline{x} \in \mathbb{R}^n : \alpha_1 x_1 + \cdots + \alpha_n x_n = 1 \}$$

not a subspace [repeat calculation $\sum_i \alpha_i (\lambda x_i + \mu y_i) = \lambda \sum_i \alpha_i x_i + \mu \sum_i \alpha_i y_i$

$= \lambda + \mu \neq 1$ in general]

(b) Linear Independence

~~Vec~~ Consider vectors $\underline{v_1}, \underline{v_2}, \ldots, \underline{v_r}$ in V (a vector space) and a linear relation

$$\lambda_1 \underline{v_1} + \lambda_2 \underline{v_2} + \cdots + \lambda_r \underline{v_r} = \underline{0} \qquad (*)$$

If $(*) \Rightarrow \lambda_i = 0 \ \forall i$ then the vectors are called linearly independent.

(they obey only the trivial linear relation, with $\lambda_i = 0$)

- If (*) is satisfied with at least one $\lambda_i \neq 0$ then the vectors are <u>linearly dependent</u> (they obey a non-trivial linear relation)

The set $\{\underline{v}_1, \underline{v}_2, \dots, \underline{v}_r\}$ is also called linearly independent as appropriate.

<u>Examples</u>

In $\mathbb{R}^2$ $\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}$ are linearly independent since $\lambda_1\begin{pmatrix}1\\0\end{pmatrix} + \lambda_2\begin{pmatrix}0\\1\end{pmatrix} = \begin{pmatrix}\lambda_1\\\lambda_2\end{pmatrix}$ is $\underline{0}$ iff $\lambda_1, \lambda_2 = 0$.

$\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}2\\-2\end{pmatrix}$ are linearly independent since $\lambda_1\begin{pmatrix}1\\1\end{pmatrix} + \lambda_2\begin{pmatrix}2\\-2\end{pmatrix} = \begin{pmatrix}\lambda_1 + 2\lambda_2\\\lambda_1 - 2\lambda_2\end{pmatrix}$

is $\underline{0}$ iff $\lambda_1 = \lambda_2 = 0$.

$\left\{\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\2\end{pmatrix}\right\}$ linearly dependent since

$$(0)\begin{pmatrix}1\\0\end{pmatrix} + (2)\begin{pmatrix}0\\1\end{pmatrix} + (-1)\begin{pmatrix}0\\2\end{pmatrix} = \underline{0}$$

a non-trivial relation. Note not every vector in this set can be expressed in terms of the others (linear): $\begin{pmatrix}1\\0\end{pmatrix}$ not a linear comb. of $\begin{pmatrix}0\\1\end{pmatrix}$ & $\begin{pmatrix}0\\2\end{pmatrix}$.

- In $\mathbb{R}^3$ vectors $\underline{a}, \underline{b}, \underline{c}$ are linearly independent iff they are not coplanar, i.e. $\underline{a} \cdot \underline{b} \wedge \underline{c} \neq 0$.

Consider $\alpha\underline{a} + \beta\underline{b} + \gamma\underline{c} = \underline{0}$

& take dot with $\underline{b} \wedge \underline{c}$ to get $\alpha\,\underline{a} \cdot \underline{b} \wedge \underline{c} = 0 \Rightarrow \alpha = 0$ & $\beta = \gamma = 0$ similarly

3.2 Bases and Dimension

For a vector space V, a basis is a set of vectors $\{\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n\}$ that

- (i) $\text{span}\{\underline{e}_1, \dots, \underline{e}_n\} = V$, i.e. any $\underline{v} \in V$ can be written

$$\underline{v} = \sum_{i=1}^{n} v_i \underline{e}_i$$

& (ii) is linearly independent.

Given (ii), the coefficients of (i) are unique, since

$$\sum_i v_i \underline{e_i} = \sum_i v_i' \underline{e_i} \Rightarrow \sum_i (v_i - v_i') \underline{e_i} = 0$$

$\Rightarrow v_i - v_i' = 0$ (linear indep.)

v_i are the components of $\underline{v}$ w.r.t. the basis.

Examples

For $\mathbb{R}^n$: $\underline{e_1} = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, $\underline{e_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \underline{e_n} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$ is a basis since

(i) $\underline{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = x_1 \underline{e_1} + x_2 \underline{e_2} + \cdots + x_n \underline{e_n}$ spans ✓

(ii) $\underline{x} = \underline{0}$ iff $x_1 = x_2 = \cdots = x_n = 0$ independent ✓

This is called the standard basis for $\mathbb{R}^n$.

For $\mathbb{R}^2$ $\underline{f_1} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\underline{f_2} = \begin{pmatrix} 2 \\ -2 \end{pmatrix}$ also a basis since lin. independent (above)

& they span since

$$\underline{e_1} = \tfrac{1}{2} \underline{f_1} + \tfrac{1}{4} \underline{f_2} \quad \text{and} \quad \underline{e_2} = \tfrac{1}{2} \underline{f_1} - \tfrac{1}{4} \underline{f_2}$$

ergo $\text{span}\{\underline{f_1}, \underline{f_2}\} = \text{span}\{\underline{e_1}, e_2\}$.

Theorem: If $\{e_1, \ldots, e_n\}$ & $\{f_1, \ldots, f_m\}$ are bases for V, then $n = m$.

i.e. all bases contain same number of vectors.

Proof: $\underline{f_a} = \sum_i A_{ai} \underline{e_i}$ for some constants A_{ai} $\boxed{\begin{matrix} 1 \le i, j \le n \\ 1 \le a, b \le m \end{matrix}}$

Similarly $\underline{e_i} = \sum_a B_{ia} \underline{f_a}$ for constants B_{ia}.

$$\Rightarrow \underline{f_a} = \sum_i A_{ai} \Big(\sum_b B_{ib} \underline{f_b} \Big) = \sum_b \Big(\sum_i A_{ai} B_{ib} \Big) \underline{f_b}$$

&

$$\underline{e_i} = \sum_a B_{ia} \Big(\sum_j A_{aj} \underline{e_j} \Big) = \sum_j \Big(\sum_a B_{ia} A_{aj} \Big) \underline{e_j}$$

Now $\{e_i\}$ and $\{f_a\}$ bases $\Rightarrow$ coeff. are unique, so

$$\sum_i A_{ai} B_{ib} = \delta_{ab} = \begin{cases} 1 & \text{if } a=b \\ 0 & \text{otherwise} \end{cases}$$

& $$\sum_a B_{ia} A_{aj} = \delta_{ij}.$$

But then

$$\sum_{i\,a} A_{ai} B_{ia} = \sum_a \delta_{aa} = m$$

$$= \sum_{a\,i} B_{ia} A_{ai} = \sum_i \delta_{ii} = n.$$

Hence $m=n$, as claimed. $\square$

<u>Def^n</u>: The number of vectors in any basis of V is the <u>dimension</u> of V.

Note: $\mathbb{R}^n$ has dimension n.

L8.1 Shown $\exists$ a well-defined notion of dimension for a vector space V:

$\dim V = n$, the # elements in any basis

- Consider any spanning set for a vector space V, $Y = \{\underline{w_1}, \underline{w_2}, \dots, \underline{w_m}\}$ with m elements.

If this is lin. independent then $m = n$ (it is a basis).

If it is not lin. indep. then

$$\sum_{i=1}^{m} \lambda_i \underline{w_i} = \underline{0} \quad \text{with at least one } \lambda_i \neq 0$$

Wlog $\lambda_m \neq 0$, then $\underline{w_m} = -\frac{1}{\lambda_m} \sum_{i=1}^{m-1} \lambda_i \underline{w_i}$ & hence span $Y' =$ span Y

where $Y' = Y / \{\underline{w_m}\}$. We can repeat this until we produce a basis of size n. Hence $m \geq n$.

unless $V \stackrel{?}{=} \emptyset$ (cannot happen)

if $V = \{\underline{0}\}$ it's shit

- Consider any linearly independent set $X = \{\underline{u_1}, \underline{u_2}, \dots, \underline{u_k}\}$ with k elements.

If span $X = V$, it is a basis and $k = n$. If not, $\exists\ \underline{u_{k+1}} \in V$ that is not expressible as $\sum_{i=1}^{k} \lambda_i \underline{u_i}$. Then $X' = X \cup \{\underline{u_{k+1}}\}$ is still linearly independent.

This follows since $\sum_{i=1}^{k+1} \mu_i \underline{u_i} = \underline{0} \Rightarrow$ $\mu_{k+1} = 0$ & then $\mu_i = 0$ for $1 \leq i \leq k$ (since X was linearly independent)

or $\mu_{k+1} \neq 0$, but then we can write $\underline{u_{k+1}}$ as a linear combi of elements of X, giving a contradiction. We can repeat until we produce a basis (of size n) and hence $k \leq n$. (unless ∞ dimensions?)

3.3 Inner Product

(a) Definition & Properties

Generalising from component expression in 3D, we define an <u>inner product</u> or <u>scalar product</u> on $\mathbb{R}^n$ by $\underline{x} \cdot \underline{y} = \sum_{i=1}^{n} x_i y_i = x_1 y_1 + \dots + x_n y_n$,

for any $\underline{x} = (x_1, \dots, x_n)$ $\underline{y} = (y_1, \dots, y_n) \in \mathbb{R}^n$.

L8.2

Properties

(i) Clearly $\underline{x} \cdot \underline{y} = \underline{y} \cdot \underline{x}$ *symmetric*

(ii) $(\lambda \underline{x} + \lambda' \underline{x}') \cdot \underline{y} = \lambda \underline{x} \cdot \underline{y} + \lambda' \underline{x}' \cdot \underline{y}$ *bi-linearity*

& $\underline{x} \cdot (\mu \underline{y} + \mu' \underline{y}') = \mu \underline{x} \cdot \underline{y} + \mu' \underline{x} \cdot \underline{y}'$

(iii) $\underline{x} \cdot \underline{x} = \sum_i x_i^2 \geq 0$ *positive definite*

w/ equality iff $x_i = 0$, i.e. $\underline{x} = \underline{0}$

The length of a vector (or norm) is $|\underline{x}| = \sqrt{\underline{x} \cdot \underline{x}}$

More generally still, an inner product on an abstract (real) vector space can be defined axiomatically based on these properties. Then inner product often written $(\underline{u}, \underline{v})$.

From definition above, the standard basis for $\mathbb{R}^n$

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

is orthonormal. $\underline{e_i} \cdot \underline{e_j} = \delta_{ij}$ [δ_{ij} works in $\mathbb{R}^n$, but $\delta_{ii} = n$]

(b) Cauchy-Schwarz & Δ Inequalities

Cauchy-Schwarz Inequality: $|\underline{x} \cdot \underline{y}| \leq |\underline{x}||\underline{y}|$ for any $\underline{x}, \underline{y} \in \mathbb{R}^n$

& equality iff $\underline{x} = \lambda \underline{y}$ or $\underline{y} = \lambda \underline{x}$ for some $\lambda \in \mathbb{R}$.

This allows us to recover geometrical properties of the inner product

(i) We can set $\underline{x} \cdot \underline{y} = |\underline{x}||\underline{y}| \cos\theta$

(C-S ensures $-1 \leq \cos\theta \leq 1$) to define angle θ between $\underline{x}$ and $\underline{y}$ in $\mathbb{R}^n$.

(ii) $|\underline{x} + \underline{y}| \leq |\underline{x}| + |\underline{y}|$ Δ inequality, since

$LHS^2 = (\underline{x} + \underline{y}) \cdot (\underline{x} + \underline{y}) = |\underline{x}|^2 + 2 \underline{x} \cdot \underline{y} + |\underline{y}|^2$

$RHS^2 = |\underline{x}|^2 + 2|\underline{x}||\underline{y}| + |\underline{y}|^2$ & compare using C-S.

L8.3 <u>Proof of C-S:</u>

If $\underline{y} = \underline{0}$, result is immediate.

If $\underline{y} \neq \underline{0}$, consider

$$|\underline{x} - \lambda \underline{y}|^2 = (\underline{x} - \lambda \underline{y})\cdot(\underline{x} - \lambda \underline{y})$$

$$= |\underline{x}|^2 - 2\lambda\, \underline{x}\cdot\underline{y} + \lambda^2 |\underline{y}|^2 \geq 0$$

for any real λ.

Inequality $\Rightarrow$ quadratic expression in λ has at most one real root

$\Rightarrow$ discriminant ≤ 0

$$(-2\lambda\, \underline{x}\cdot\underline{y})^2 - 4|\underline{y}|^2|\underline{x}|^2 \leq 0$$

& result follows.

Furthermore, equality holds iff (dis.) $= 0$ $\Rightarrow$ single real root $\Rightarrow$ $\underline{x} = \lambda \underline{y}$ for some λ.

$\square$

<u>Example</u>

Vector space of functions $V = \{ f : [0,1] \to \mathbb{R} : f \text{ smooth},\ f(0) = f(1) = 0 \}$

Define an inner product by

$$(f, g) = \int_0^1 f(x) g(x)\, dx$$

(symmetric, bilinear and pos. def.)

Cauchy-Schwarz: $|(f,g)| \leq \|f\|\,\|g\|$ where $\|f\| = \sqrt{(f,f)}$, i.e.

$$\left| \int_0^1 f(x) g(x) \right| \leq \left(\int_0^1 f(x)^2\, dx \right)^{\frac{1}{2}} \left(\int_0^1 g(x)^2\, dx \right)^{\frac{1}{2}}.$$

L8.4

§3.4 $\mathbb{C}^n$ & Complex Vector Spaces

Let $\mathbb{C}^n = \{ \underline{z} = (z_1, z_2, \dots, z_n) : z_i \in \mathbb{C} \}$ & define

(i) Addition: $\underline{z} + \underline{w} = (z_1 + w_1, \dots, z_n + w_n)$ for $\underline{z}, \underline{w} \in \mathbb{C}^n$

(ii) Scalar multiplication: $\lambda \underline{z} = (\lambda z_1, \dots, \lambda z_n)$ for $\underline{z} \in \mathbb{C}^n$, $\lambda \in \mathbb{C}$

These operations obey the axioms of a vector space with scalars now complex rather than real.

Define an inner product on $\mathbb{C}^n$ by

$$(\underline{z}, \underline{w}) = \sum_i \overline{z}_i w_i = \overline{z}_1 w_1 + \dots + \overline{z}_n w_n \quad \text{complex valued}$$

Properties

(i) $(\underline{w}, \underline{z}) = \overline{(\underline{z}, \underline{w})}$ hermitian

(ii) $(\lambda \underline{z} + \lambda' \underline{z}', \underline{w}) = \overline{\lambda} (\underline{z}, \underline{w}) + \overline{\lambda}' (\underline{z}', \underline{w})$ anti-linear in 1st arg

& $(\underline{z}, \lambda \underline{w} + \lambda' \underline{w}') = \lambda (\underline{z}, \underline{w}) + \lambda' (\underline{z}', \underline{w}')$ linear in 2nd arg

(iii) $(\underline{z}, \underline{z}) \in \mathbb{R}^+_0$ with equality to 0 iff $\underline{z} = 0$ positive definite

L9.1

4. Matrices & Linear Maps

4.1 Introduction & Definitions

A <u>linear map</u> or <u>linear transformation</u> is a function $T : V \to W$ between vector spaces V and W satisfying $T(\lambda \underline{x} + \mu \underline{y}) = \lambda T(\underline{x}) + \mu T(\underline{y})$ for all $\lambda, \mu \in \mathbb{R}$ if V, W are real vector spaces e.g. $V = \mathbb{R}^n$, $W = \mathbb{R}^m$ or for all $\lambda, \mu \in \mathbb{C}$ if V, W are complex vector spaces e.g. $V = \mathbb{C}^n$, $W = \mathbb{C}^m$.

[V is the domain, W is the (range or) codomain]

- $\underline{x}' = T(\underline{x})$ is the image of $\underline{x}$ under T

  $\operatorname{Im} T = \{ \underline{x}' \in W : T(\underline{x}) = x' \text{ for some } \underline{x} \in V \}$

  is the <u>image</u> of T.

  $\operatorname{Im} T$ is a subspace of W.

  $\operatorname{Ker} T = \{ \underline{x} \in V : T(\underline{x}) = \underline{0} \}$

  is the <u>kernel</u> of T.

  $\operatorname{Ker} T$ is a subspace of V.

$\dim \operatorname{Im}(T)$ is the <u>rank</u> of T

$\dim \operatorname{Ker}(T)$ is the <u>nullity</u> of T

<u>Example</u> $V = W = \mathbb{R}^3$

$$T : (x_1, x_2, x_3) \to (3x_1 + x_2 + 5x_3,\ -x_1 - 2x_3,\ 2x_1 + x_2 + 3x_3)$$

$\operatorname{Im} T = \{ \lambda(1,0,1) + \mu(3,-1,2) : \lambda, \mu \in \mathbb{R} \}$

& $\operatorname{rank} T = 2$

$\operatorname{Ker} T = \{ \lambda(2,-1,-1) : \lambda \in \mathbb{R} \}$

& $\operatorname{nullity} T = 1$

- Note we can take linear combinations of linear maps:

  if T, S both linear $V \to W$

  then $\alpha T + \beta S : V \to W$

  is also linear where

  $(\alpha T + \beta S)(\underline{x}) = \alpha T(\underline{x}) + \beta S(\underline{x})$.

- We can also compose them: given linear maps $U \xrightarrow{S} V \xrightarrow{T} W$ as shown, then $(T \circ S)(\underline{x}) = T(S(\underline{x}))$.

4.2 Matrices for linear maps $\mathbb{R}^n \to \mathbb{R}^n$

- A linear map $T: \mathbb{R}^n \to \mathbb{R}^n$ can be defined by a <u>matrix</u>: an $n \times n$ array M with <u>entries</u> $M_{ij} \in \mathbb{R}$ where i labels <u>rows</u> and j labels <u>columns</u>

e.g. $n=3$

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

$T(\underline{x}) = \underline{x}' = M\underline{x}$ where second equality means

$$x_i' = M_{ij} x_j \quad \left(\begin{matrix}\text{summation}\\ \text{convention}\end{matrix}\right)$$

- $$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_n' \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ M_{n1} & M_{n2} & \cdots & M_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Helpful to regard rows & cols of M as vectors e.g. $n=3$

$$M = \begin{pmatrix} \leftarrow \underline{R}_1 \rightarrow \\ \leftarrow \underline{R}_2 \rightarrow \\ \leftarrow \underline{R}_3 \rightarrow \end{pmatrix} = \begin{pmatrix} \uparrow & \uparrow & \uparrow \\ \underline{C}_1 & \underline{C}_2 & \underline{C}_3 \\ \downarrow & \downarrow & \downarrow \end{pmatrix}$$

$$x_i' = M_{ij} x_j = \underline{R}_i \cdot \underline{x}$$

$\underline{e}_i$ standard basis vectors

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ \underline{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

- behave under T as follows:

$$\underline{e}_i \to T(\underline{e}_i) = \underline{e}_i' = M\underline{e}_i = \underline{C}_i$$

i.e. columns of M are images of standard basis vectors & $\mathrm{Im}(M)$

$$= \mathrm{Im}(T) = \text{span cols} = \text{span}(\underline{C}_1, \underline{C}_2, \ldots, \underline{C}_n)$$

<u>Example</u>

Consider linear map written previously

matrix $M = \begin{pmatrix} 3 & 1 & 5 \\ -1 & 0 & -2 \\ 2 & +1 & 3 \end{pmatrix}$ corresponds to map $T(\underline{x}) = \underline{x}' = M\underline{x}$ with

$$\begin{pmatrix} x_1' \\ x_2' \\ x_3' \end{pmatrix} = \begin{pmatrix} 3x_1 + x_2 + 5x_3 \\ -x_1 - 2x_3 \\ 2x_1 + x_2 + 3x_3 \end{pmatrix}.$$

$$\mathrm{Im}\, T = \text{Span}\left\{ \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 5 \\ -2 \\ 3 \end{pmatrix} \right\}$$

- $$= \text{Span}\left\{ \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$ (3^{rd} col linear combi of other two)

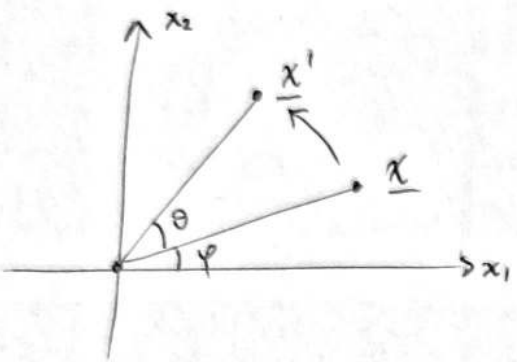
so $\mathrm{rank}(T) = 2$.

L9.3

4.3 Geometrical Examples $\mathbb{R}^3 \to \mathbb{R}^3$

(a) Rotations

(i) Rotation through angle θ about vector $\underline{e}_3$ (axis)



$$\underline{x} = (\rho\cos\varphi, \rho\sin\varphi, x_3)$$

$$\underline{x}' = \left(\rho\cos(\theta+\varphi), \rho\sin(\theta+\varphi), x_3\right)$$

$$= \left(x_1\cos\theta - x_2\sin\theta,\ x_1\sin\theta + x_2\cos\theta,\ x_3\right)$$

using trig. formulae

i.e. $x_i' = R_{ij}x_j$ where $R = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$

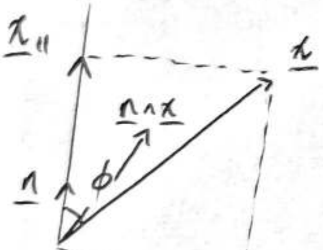
Alternative derivation: cols of R are images of standard basis vecs

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \to \begin{pmatrix} \cos\theta \\ \sin\theta \\ 0 \end{pmatrix} = \underline{e}_1' = \underline{C}_1$$

$$\underline{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \to \begin{pmatrix} -\sin\theta \\ \cos\theta \\ 0 \end{pmatrix} = \underline{e}_2' = \underline{C}_2$$

$$\underline{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \underline{e}_3' = \underline{C}_3$$

(ii) Rotation through angle θ about an axis $\underline{n}$, unit vector



Write $\underline{x} = \underline{x}_{\parallel} + \underline{x}_{\perp}$ where $\underline{x}_{\parallel} = (\underline{n}\cdot\underline{x})\underline{n}$

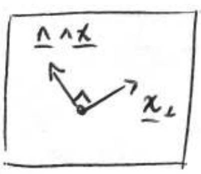
& $\underline{x}_{\perp} = \underline{x} - (\underline{n}\cdot\underline{x})\underline{n} \perp \underline{n}$.

Note $|\underline{x}_{\parallel}| = |\underline{x}|\cos\phi$ but $\underline{n}\wedge\underline{x} \perp \underline{n}, \underline{x}, \underline{x}_{\perp}$

$|\underline{x}_{\perp}| = |\underline{x}|\sin\phi$ & $|\underline{n}\wedge\underline{x}| = |\underline{x}|\sin\phi$.

Under rotation $\underline{x}_{\parallel} \to \underline{x}_{\parallel}$, $\underline{x}_{\perp} \to \cos\theta\,\underline{x}_{\perp} + \sin\theta\,\underline{n}\wedge\underline{x}$

Looking down on plane $\perp \underline{n}$



Hence $\underline{x} \to \underline{x}_{\parallel} + \cos\theta\,\underline{x}_{\perp} + \sin\theta\,\underline{n}\wedge\underline{x}$

i.e. $\underline{x}' = R\underline{x} = \underline{x}\cos\theta + \left(1 - \underset{?}{\overset{\cos}{\cancel{\sin}}}\theta\right)(\underline{n}\cdot\underline{x})\underline{n} + \sin\theta\,\underline{n}\wedge\underline{x}$

L9.4

In components

$$x_i' = R_{ij} x_j = x_i \cos\theta + (1 - \underset{\cos}{\cancel{\sin}}\,\theta) n_j x_j n_i + \sin\theta \, \varepsilon_{ijk} n_j x_k$$

RHS can be arranged into

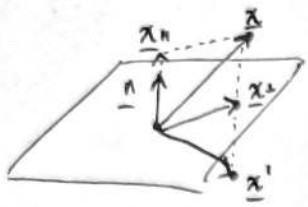
$$R_{ij} = \delta_{ij} \cos\theta + (1 - \underset{\cos}{\cancel{\sin}}\,\theta) n_i n_j - \sin\theta \, \varepsilon_{ijk} n_k$$

L10.1

check $\theta = 0$:

$R_{ij} = \delta_{ij}$ and $R = I$ is the unit matrix

(b) <u>Reflection</u> in plane with unit normal $\underline{n}$



$$\underline{x} = \underline{x}_{\parallel} + \underline{x}_{\perp} \quad \longrightarrow \quad \underline{x}' = -\underline{x}_{\parallel} + \underline{x}_{\perp} = \underline{x} - 2(\underline{x}\cdot\underline{n})\underline{n}$$

In components $x_i' = H_{ij} x_j$

$$= (\delta_{ij} - 2n_i n_j) x_j$$

so $H_{ij} = \delta_{ij} - 2n_i n_j$.

For example, reflection in plane with $\underline{n} = \frac{1}{\sqrt{3}}(1,1,1)$

$n_i n_j = \frac{1}{3} \ \forall i,j$

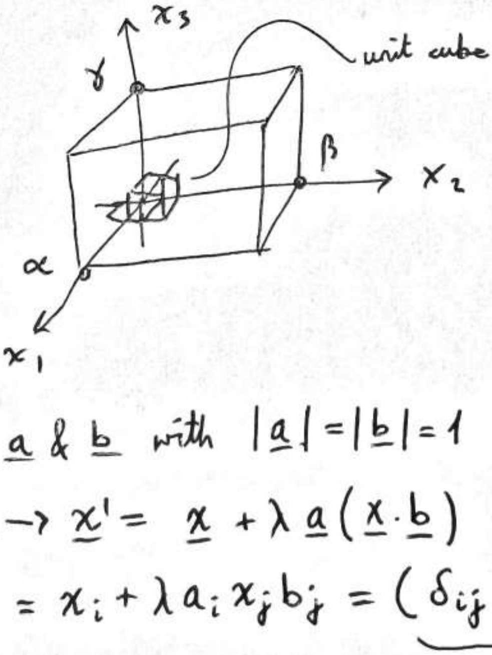
$$\Rightarrow H = \begin{pmatrix} 1/3 & -2/3 & -2/3 \\ -2/3 & 1/3 & -2/3 \\ -2/3 & -2/3 & 1/3 \end{pmatrix}.$$

[related map: projection onto plane given by $\underline{x} \to \underline{x} - (\underline{x}\cdot\underline{n})\underline{n}$]

(c) <u>Dilation</u> with scale factors α, β, γ along directions of $\underline{e_1}, \underline{e_2}, \underline{e_3}$ is given by

$$\underline{x} = x_i \underline{e}_i \quad \to \quad \underline{x}' = \alpha x_1 \underline{e_1} + \beta x_2 \underline{e_2} + \gamma x_3 \underline{e_3} = M\underline{x} \quad \text{where}$$

$$M = \begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$$



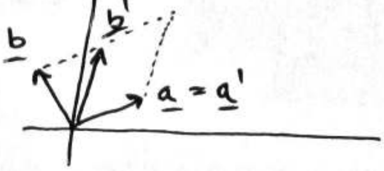
(d) <u>Shear</u> given vectors $\underline{a}$ & $\underline{b}$ with $|\underline{a}| = |\underline{b}| = 1$ & $\underline{a}\cdot\underline{b} = 0$, a shear with parameter λ is defined by $\underline{x} \to \underline{x}' = \underline{x} + \lambda\underline{a}(\underline{x}\cdot\underline{b})$

$$\text{so} \quad x_i' = x_i + \lambda a_i x_j b_j = \underbrace{(\delta_{ij} + \lambda a_i b_j)}_{S_{ij}} x_j$$

Note $\underline{a} \to \underline{a}$

$\underline{b} \to \underline{b} + \lambda\underline{a}$

$\underline{u} \to \underline{u}$ whenever $\underline{u} \perp \underline{b}$



4.4 Matrices in General

Consider a linear map $T: V \to W$ with V & W real or complex vectors spaces of dimension n & m resp., & consider $\{\underline{e}_1, \dots, \underline{e}_n\}$ basis for V & $\{\underline{f}_1, \dots, \underline{f}_m\}$ basis for W. The matrix for T w.r.t. these bases is an $m \times n$ array with entries M_{ai} where $a = 1, \dots, m$, $i = 1, \dots, n$, in $\mathbb{R}$ or $\mathbb{C}$, defined by $T(\underline{e}_i) = \sum_a \underline{f}_a M_{ai}$

row cols

(note index positions). If $\underline{x}' = T(\underline{x})$ with $\underline{x} = \sum_i x_i \underline{e}_i$ & $\underline{x}' = \sum_a x_a' \underline{f}_a$ then

$$\sum_a x_a' \underline{f}_a = T\Big(\sum_i x_i \underline{e}_i\Big) = \sum_i x_i T(\underline{e}_i) = \sum_i x_i \sum_a \underline{f}_a M_{ai}$$

$$= \sum_a \Big(\sum_i M_{ai} x_i\Big) \underline{f}_a$$

Equate coeffs ($\underline{f}_a$ basis vecs) so $x_a' = \sum_i M_{ai} x_i = M_{ai} x_i$

$$\begin{pmatrix} x_1' \\ \vdots \\ x_m' \end{pmatrix} = \begin{pmatrix} M_{11} & \cdots & M_{1n} \\ \vdots & \ddots & \vdots \\ M_{m1} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

Moral: given choices of bases V is identified with $\mathbb{R}^n$, W is identified with $\mathbb{R}^m$, T is identified with a matrix M. (or same with $\mathbb{C}$)

If $S: V \to W$ is a linear map with matrix N (w.r.t. same bases) then

$\alpha T + \beta S$ has matrix $\alpha M + \beta M$ with entries

$$(\alpha M + \beta N)_{ai} = \alpha M_{ai} + \beta N_{ai}.$$

(2) Matrix multiplication as composition of linear maps

Given linear maps $U \xrightarrow{S} V \xrightarrow{T} W$ with matrices A for T, B for S,

dim: p $\quad n \quad$ m

bases: $\{\underline{u}_r\}$ $\quad \{\underline{e}_i\}$ $\quad \{\underline{f}_a\}$

then the composite map $T \circ S$ has matrix $C = AB$ where $\underset{m \times p}{C_{ar}} = \underset{m \times n}{A_{ai}} \underset{n \times p}{B_{ir}}$

10.3 This is the defn of matrix multiplication – essential that have common range of index i above. In terms of columns and rows

$$AB = \begin{pmatrix} \vdots \\ \longleftarrow \underline{R}_a(A) \longrightarrow \\ \vdots \end{pmatrix} \begin{pmatrix} & \uparrow & \\ \cdots & \underline{C}_r(B) & \cdots \\ & \downarrow & \end{pmatrix}$$

$$(AB)_{ar} = \underbrace{\underline{R}_a(A) \cdot \underline{C}_r(B)}_{\text{both vectors in } \mathbb{R}^n}$$

<u>Example</u> $\mathbb{R}^3 \longrightarrow \mathbb{R}^2 \longrightarrow \mathbb{R}^3$

$$B = \begin{pmatrix} 1 & 0 & -1 \\ 2 & -1 & 3 \end{pmatrix} \quad \& \quad A = \begin{pmatrix} 1 & 3 \\ -5 & 0 \\ 2 & 1 \end{pmatrix} \quad \text{then} \quad AB = \begin{pmatrix} 7 & -3 & 8 \\ -5 & 0 & 5 \\ 4 & -1 & 1 \end{pmatrix}$$

(b) <u>Transpose & Hermitian Conjugate</u>

If M is an $m \times n$ matrix, the <u>transpose</u>

M^T is an $n \times m$ matrix defined by $(M^T)_{ia} = M_{ai}$

(exchanging rows & cols)

$i = 1, \dots, n; \ a = 1, \dots, m$

If M is $n \times n$ (square) then M is <u>symmetric</u> iff $M^T = M$ or $M_{ij} = M_{ji}$

& <u>antisymmetric</u> iff $M^T = -M$ or $M_{ij} = -M_{ji}$.

Clearly, in general for $m \times n$ matrices A & B we have

$$(\alpha A + \beta B)^T = \alpha A^T + \beta B^T.$$

Also $(AB)^T = B^T A^T$ for A $p \times m$, B $m \times n$.

Check using indices $[(AB)^T]_{ra} = (AB)_{ar} = A_{ai} B_{ir} = B_{ir} A_{ai}$

$$= (B^T)_{\cancel{ir}\, ri} (A^T)_{ia} = (B^T A^T)_{ra} \text{ as claimed.}$$

The <u>hermitian conjugate</u> of a matrix M is written $M^\dagger$ & defined by

$(M^\dagger)_{ia} = \overline{M_{ai}}$ or $M^\dagger = \overline{M^T} = \overline{M}^T$ where complex conjugation is defined component-wise. If M is square, it is <u>hermitian</u> iff $M^\dagger = M$ & <u>anti-hermitian</u> iff $M^\dagger = -M$

(c) Trace

- For a complex $n \times n$ matrix M, the trace is defined by

$$\mathrm{tr}(M) = M_{ii} = M_{11} + \cdots + M_{nn}$$ sum of diag. entries

Note: (i) $\mathrm{tr}(\alpha M + \beta N) = \alpha\, \mathrm{tr} M + \beta\, \mathrm{tr} N$

$$[\, (\alpha M + \beta N)_{ii} = \alpha M_{ii} + \beta N_{ii} \,]$$

(ii) $\mathrm{tr}(MN) = \mathrm{tr}(NM)$

$$[\, (MN)_{ii} = M_{ij} N_{ji} = N_{ji} M_{ij} = (NM)_{jj} \,]$$

- But $MN \neq NM$ in general.

(iii) $\mathrm{tr}(M) = \mathrm{tr}(M^T)$

(iv) $n \times n$ identity or unit matrix $I = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$ has $I_{ij} = \delta_{ij}$

& $\mathrm{tr}\, I = \delta_{ii} = n$.

Example $R_{ij} = \delta_{ij} - 2 n_i n_j$ is reflection matrix with $\underline{n}$ unit vector introduced in §4.3 (b) for reflection in plane in $\mathbb{R}^3$ with normal $\underline{n}$; same formula applies in $\mathbb{R}^2$ for reflection in line with normal $\underline{n}$

- $$\mathrm{tr}\, R = R_{ii} = \delta_{ii} - 2 n_i n_i = \begin{cases} 3 - 2 = 1 & \text{in } \mathbb{R}^3 \\ 2 - 2 = 0 & \text{in } \mathbb{R}^2 \end{cases}$$

This captures eigenvalues of map (see later)

(d) Decomposition of an $n \times n$ matrix

An $n \times n$ matrix M can be written as a sum of symmetric and anti-symmetric parts. Define $\underbrace{S = \tfrac{1}{2}(M + M^T)}_{\text{symmetric}}$, or $S_{ij} = \tfrac{1}{2}(M_{ij} + M_{ji})$

- & $\underbrace{A = \tfrac{1}{2}(M - M^T)}_{\text{antisymmetric}}$, or $A_{ij} = \tfrac{1}{2}(M_{ij} - M_{ji})$

& then $S + A = M$, or $S_{ij} + A_{ij} = M_{ij}$.

The symmetric part can be further decomposed into a multiple of I & a symmetric traceless part. Let

$$T = S - \frac{1}{n}\operatorname{tr}(S) I \quad \text{or} \quad T_{ij} = S_{ij} - \frac{1}{n}\operatorname{tr}(S)\delta_{ij}$$

then $\operatorname{tr}(T) = T_{ii} = S_{ii} - \frac{1}{n}\operatorname{tr}(S)\,\delta_{ii} = 0$.

But $\operatorname{tr}(S) = \operatorname{tr}(M)$ since $\operatorname{tr}(A) = 0$.

Hence $$M_{ij} = T_{ij} + A_{ij} + \frac{1}{n}(\operatorname{tr} M)\,\delta_{ij}.$$

Example For $n=3$, suppose $T_{ij} = 0$. Set $A_{ij} = \varepsilon_{ijk} a_k$ with

$A = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}$ & $\operatorname{Tr}(M) = 3\lambda$. Then $M\underline{x} = \underline{x} \times \underline{a} + \lambda \underline{x}$.

(e) Matrix inverses

Consider $m \times n$ matrix A & $n \times m$ matrices B, C. B is a left inverse for A if $BA = I$ $(n \times n)$. C is a right inverse for A if $AC = I$ $(m \times m)$. If A is square, $m = n$, then $B = B(AC) = (BA)C = C$ & we write $B = C = A^{-1}$ the inverse: $AA^{-1} = A^{-1}A = I$. Not every matrix A has an inverse; if it does A is called invertible or non-singular. If P & Q are $n \times n$ invertible matrices then PQ is invertible with $(PQ)^{-1} = Q^{-1}P^{-1}$.

[check: $PQQ^{-1}P^{-1} = PIP^{-1} = PP^{-1} = I$]

Examples (i) In §4.3 (a) found matrix for rotation $R(\theta)$ with given axis $\underline{n}$:
$$R(\theta)_{ij} = \delta_{ij}\cos\theta + (1-\cos\theta)n_i n_j - \varepsilon_{ijk}\sin\theta\, n_k$$

$$R(\theta)_{ij} R(-\theta)_{jk} = \left(\delta_{ij}\cos\theta + (1-\cos\theta)n_i n_j - \varepsilon_{ijp}\sin\theta\, n_p\right)$$
$$\times \left(\delta_{jk}\cos\theta + (1-\cos\theta)n_j n_k + \varepsilon_{jkq}\sin\theta\, n_q\right)$$

$$= \delta_{ik}\cos^2\theta + \cos\theta(1-\cos\theta)\,2n_in_k + (1-\cos\theta)^2 n_in_k \qquad (n_jn_j=1)$$

$$- \varepsilon_{ijp}\varepsilon_{jkq}\, n_pn_q \sin^2\theta \qquad \text{[other terms cancel]}$$

$(1-\cos\theta)\sin\theta\, n_in_jn_q\varepsilon_{jkq}$ is zero by symmetry

$$= \delta_{ik}\cos^2\theta + \cancel{(1-\cos^2\theta)n_in_k}$$

$$+ \delta_{ik}n_pn_p\sin^2\theta - \cancel{n_in_k\sin^2\theta}$$

$\varepsilon_{ikq}\sin\theta\cos\theta\, n_q = \varepsilon_{ikp}\sin\theta\cos\theta\, n_p$ by relabelling

$$= \delta_{ik}. \quad \text{Hence } R(-\theta) = R(\theta)^{-1}.$$

(ii) Shear in §4.3 (b) was defined for fixed orthonormal vectors $\underline{a}, \underline{b} \in \mathbb{R}^3$.

Specialise to $\underline{a} = \underline{e_1}$, $\underline{b} = \underline{e_2}$, then shear with parameter λ is given by

$$S(\lambda)_{ij} = \delta_{ij} + \lambda a_ib_j = \delta_{ij} + \lambda\delta_{i1}\delta_{j2}$$

or $S(\lambda) = \begin{pmatrix} 1 & \lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ & $S(-\lambda) = \begin{pmatrix} 1 & -\lambda & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ & $S(-\lambda)S(\lambda) = I$.

4.5 Orthogonal & Unitary Matrices

A real $n\times n$ matrix U is orthogonal iff $U^TU = UU^T = I$,

i.e. $U^T = U^{-1}$. Example: $U = R(\theta)$ as in example above

then $U^T = R(\theta)^T = R(-\theta) = U^{-1}$

(by inspection)

Equivalent statement: U is orthogonal iff it preserves inner products:

$$(U\underline{x})\cdot(U\underline{y}) = \underline{x}\cdot\underline{y} \quad \forall\, \underline{x},\underline{y} \in \mathbb{R}^n$$

i.e. it preserves lengths & angles (defined via Cauchy-Schwarz)

To see equivalence, note that the inner product can be written

$\underline{x}\cdot\underline{y} = \underline{x}^T\underline{y}$ with $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ col vec. or $n\times 1$ matrix

& $\underline{x}^T = (x_1, \dots, x_n)$ row vec or $1\times n$ matrix

Then $(U\underline{x})^T(U\underline{y}) = \underline{x}^TU^TU\underline{y} = \underline{x}^T\underline{y}$ $\forall\, x, y$ iff $U^TU = I$.

Another point of view: cols of U are $U\underline{e_1}, \dots, U\underline{e_n}$. But then U preserving inner products is equivalent to $(U\underline{e_i})\cdot(U\underline{e_j}) = \underline{e_i}\cdot\underline{e_j} = \delta_{ij}$ $\Leftrightarrow$ cols. of U being orthonormal basis. Compare with $U_{ik}U^T_{kj} = \delta_{ij}$

Example General 2×2 orthogonal matrix U:

$$U\begin{pmatrix}1\\0\end{pmatrix}=\begin{pmatrix}\cos\theta\\ \sin\theta\end{pmatrix}$$ general unit vector in $\mathbb{R}^2$ for some θ

$$U\begin{pmatrix}0\\1\end{pmatrix}=\pm\begin{pmatrix}-\sin\theta\\ \cos\theta\end{pmatrix}$$ general unit vector such that $U\begin{pmatrix}0\\1\end{pmatrix}\perp U\begin{pmatrix}1\\0\end{pmatrix}$.

Two cases:

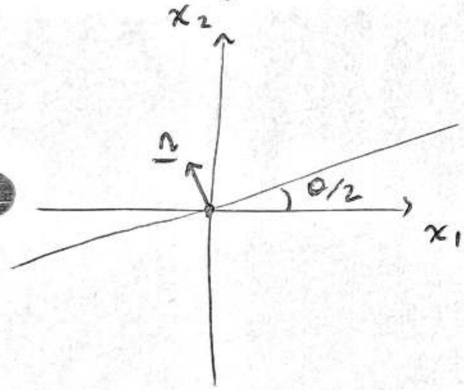
$$U=R=\begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$$ rotation in $\mathbb{R}^2$ by comparison with §4.3(a)

or $$U=H=\begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$$ reflection in $\mathbb{R}^2$ by comparison with §4.3(b)

More precisely: H is reflection in line through origin with normal $\begin{pmatrix}-\sin\theta/2\\ \cos\theta/2\end{pmatrix}$

since $H_{ij}=\delta_{ij}-2n_in_j \Rightarrow$ ~~δ_{ij}~~ $H=\begin{pmatrix}1-2\sin^2\theta/2 & 2\sin\theta/2\cos\theta/2\\ 2\sin\theta/2\cos\theta/2 & 1-2\cos^2\theta/2\end{pmatrix}$,

which matches using double angle formulae.



reflection in line making angle $\theta/2$ with x-axis

For rotation $R(\theta)^T=R(\theta)^{-1}=R(-\theta)$

but for reflection $H^T=H$ & $H^2=I$

so $H^T=H^{-1}=H$.

A complex $n\times n$ matrix U is unitary iff $U^\dagger U=UU^\dagger=I$, i.e. $U^\dagger=U^{-1}$.

Equivalent statement: U is unitary iff it preserves complex inner product:

$$(U\underline{z},U\underline{w})=(\underline{z},\underline{w})\quad\forall\,\underline{z},\underline{w}\in\mathbb{C}^n$$

Inner product can be written $(\underline{z},\underline{w})=\underline{z}^\dagger\underline{w}=\bar{z}_iw_i$

with $\underline{z}=\begin{pmatrix}z_1\\ \vdots\\ z_n\end{pmatrix}$ col vec $(n\times1)$ matrix & $\underline{z}^\dagger=(\bar{z}_1,\dots,\bar{z}_n)$ row vec $(1\times n)$ matrix

$$(U\underline{z})^\dagger(U\underline{w}) = \underline{z}^\dagger U^\dagger U\underline{w} = \underline{z}^\dagger \underline{w} \quad \forall\, \underline{z},\underline{w} \iff U^\dagger U = I$$

- Example 1×1 unitary matrix U:

$u \in \mathbb{C}$ with $u^\dagger = \bar{u}$ & $u^\dagger u = \bar{u}u = 1 \Rightarrow u = e^{i\theta}$ for some real θ

Relate $\mathbb{C}$ to $\mathbb{R}^2$ by $z = x_1 + ix_2 \longleftrightarrow \underline{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$,

$$w = y_1 + iy_2 \longleftrightarrow \underline{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

Complex inner product on $\mathbb{C}$ is $\bar{z}w = (x_1 - ix_2)(y_1 + iy_2)$

$$= \underline{x}\cdot\underline{y} + i[\underline{x},\underline{y}]$$

Real part $\underline{x}\cdot\underline{y}$, standard inner product on $\mathbb{R}^2$; Im part, real scalar

$[\underline{x},\underline{y}] = x_1y_2 - x_2y_1$ defined for all $\underline{x},\underline{y} \in \mathbb{R}^2$.

[note $\begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} y_1 \\ y_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ [\underline{x},\underline{y}] \end{pmatrix}$]

Now linear map on $\mathbb{C}$ $\quad z \to uz = e^{i\theta}z$

$(x_1 + ix_2) \to (\cos\theta + i\sin\theta)(x_1 + ix_2)$

$\iff \underline{x} \to R\underline{x}$ with $R = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

Know $\underline{x}\cdot\underline{y} \to (R\underline{x})\cdot(R\underline{y}) = \underline{x}\cdot\underline{y}$ & can check that

$$[\underline{x},\underline{y}] \to [R\underline{x}, R\underline{y}] = [\underline{x},\underline{y}]$$

but both these follow from $\bar{z}w \to (e^{i\theta}z)^\dagger e^{i\theta}w = \bar{z}w$

5. Determinants, Rank & Inverses

5.1 Determinants & Inverses in $\mathbb{R}^3$ & $\mathbb{R}^2$

Consider a linear map $\mathbb{R}^n \to \mathbb{R}^n$ given by $\underline{x} \to M\underline{x}$ with M $n\times n$ matrix.

For $n=2$ or 3 we will define a related matrix $\widetilde{M}$ and a scalar, the <u>determinant</u>

$\det(M)$ such that $\widetilde{M}M = (\det M) I$ (*)

Moreover $\det M$ is the factor by which areas ($n=2$) or volumes ($n=3$) get scaled under the action of M.

- If $\det M \neq 0$ then M is invertible with $M^{-1} = \frac{1}{\det M}\widetilde{M}$.

In later sections we generalise this to $\mathbb{C}^n \to \mathbb{C}^n$ for all n.

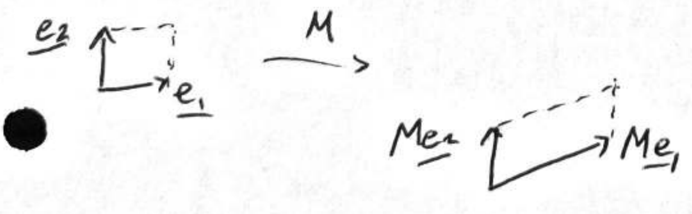
(a) $\mathbb{R}^2 \to \mathbb{R}^2$

For vectors $\underline{a}, \underline{b}$ in $\mathbb{R}^2$ the scalar $[\underline{a}, \underline{b}] = a_1b_2 - a_2b_1 = -[\underline{b}, \underline{a}]$ gives area of parallelogram (with orientation) defined by $\underline{a}$ and $\underline{b}$. For standard orthonormal basis vectors $\underline{e}_1 = \begin{pmatrix}1\\0\end{pmatrix}$ & $\underline{e}_2 = \begin{pmatrix}0\\1\end{pmatrix}$ we have

- $[\underline{e}_1, \underline{e}_2] = 1$. For a map $\mathbb{R}^2 \to \mathbb{R}^2$ with $\underline{x} \to M\underline{x}$ recall cols of M are images of basis vectors

$$\underline{c}_1 = M\underline{e}_1 = M_{11}\underline{e}_1 + M_{21}\underline{e}_2$$
$$\underline{c}_2 = M\underline{e}_2 = M_{12}\underline{e}_1 + M_{22}\underline{e}_2.$$

Under M, areas scale by a factor $[M\underline{e}_1, M\underline{e}_2] = [\underline{c}_1, \underline{c}_2] = \det M$, where $\det M = M_{11}M_{22} - M_{12}M_{21}$ the <u>determinant</u> of $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22}\end{pmatrix}$.



L12.4

Now define $\tilde{M} = \begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix}$ & (*) holds, as claimed.

- $\det M \neq 0 \iff M\underline{e_1}, M\underline{e_2}$ lin. independent $\rightsquigarrow$ say $M_{11}M_{22} = M_{12}M_{21}$

  $\iff \mathrm{Im}(M) = \mathbb{R}^2$

  $\iff \mathrm{rank}\, M = 2$

then
$(-M_{12}M_{22})\begin{pmatrix} M_{11} \\ M_{21} \end{pmatrix} + (M_{11}M_{22})\begin{pmatrix} M_{12} \\ M_{22} \end{pmatrix}$
is linear combi which is trivial only if 0s in same row/col which are obviously ~~independent~~ dependent

also kinda obvious haha dependence constrains to line

Example

$H = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}$ $\quad \det H = -1$

- $\tilde{H} = \begin{pmatrix} -\cos\theta & \sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ & $H^{-1} = -\tilde{H}$.

(b) $\mathbb{R}^3 \to \mathbb{R}^3$

For vectors $\underline{a}, \underline{b}, \underline{c} \in \mathbb{R}^3$ let $[\underline{a}, \underline{b}, \underline{c}] = \underline{a} \cdot \underline{b} \wedge \underline{c}$ (new notation)

$= \varepsilon_{ijk} a_i b_j c_k$

volume (with orientation) for parallelepiped defined by $\underline{a}, \underline{b}, \underline{c}$.

For $\underline{e_1}, \underline{e_2}, \underline{e_3}$ basis vectors,

- $[\underline{e_i}, \underline{e_j}, \underline{e_k}] = \varepsilon_{ijk}$

For a map $\mathbb{R}^3 \to \mathbb{R}^3$ with $\underline{x} \to M\underline{x}$ we have cols

$\underline{C_1}(M) = M\underline{e_1} = M_{i1}\underline{e_i}$

$\underline{C_2}(M) = M\underline{e_2} = M_{j2}\underline{e_j}$

$\underline{C_3}(M) = M\underline{e_3} = M_{k3}\underline{e_k}$

Under M, volumes scale by a factor

$[M\underline{e_1}, M\underline{e_2}, M\underline{e_3}] = [\underline{C_1}(M), \underline{C_2}(M), \underline{C_3}(M)]$

$= M_{i1}M_{j2}M_{k3}[\underline{e_i}, \underline{e_j}, \underline{e_k}]$

$= M_{i1}M_{j2}M_{k3}\,\varepsilon_{ijk}$

$= \det M$ $\quad$ def^n for 3×3

Now define $\tilde{M}$ in terms of its rows:

- $\underline{R_1}(\tilde{M}) = \underline{C_2}(M) \times \underline{C_3}(M)$

and same cyclically ...

& note

$\det M \neq 0 \iff \mathrm{Im}\, M = \mathbb{R}^3 \iff \mathrm{rank}\, M = 3$

$$(\tilde{M}M)_{ij} = R_i(\tilde{M}) \cdot C_j(M) = (\overbrace{\underline{C_1}(M) \cdot \underline{C_2}(M) \times \underline{C_3}(M)}^{\det M})\,\delta_{ij}$$

come on ...

Example

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

$$\underline{c_2} \times \underline{c_3} = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \wedge \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 6 \end{pmatrix}$$

$$\underline{c_3} \times \underline{c_1} = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \wedge \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ -1 \\ -2 \end{pmatrix}$$

$$\underline{c_1} \times \underline{c_2} = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ 11 \\ -1 \end{pmatrix}$$

$$\det M = \underline{c_1} \cdot \underline{c_2} \wedge \underline{c_3} = 23$$

$$\tilde{M} = \begin{pmatrix} -1 & 3 & 6 \\ 8 & -1 & -2 \\ 4 & 11 & -1 \end{pmatrix}$$

& $\tilde{M}M = 23\, I$.

Notation Determinants are written $\det(M) = |M|$ for a matrix M

e.g. $\begin{vmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{vmatrix} = M_{11}M_{22} - M_{12}M_{21}$

For M 3×3 we can calculate $\det M$ in terms of 2×2 determinants

e.g. $\begin{vmatrix} a & l & p \\ b & m & q \\ c & n & r \end{vmatrix} = a\begin{vmatrix} m & q \\ n & r \end{vmatrix} - b\begin{vmatrix} l & p \\ n & r \end{vmatrix} + c\begin{vmatrix} l & p \\ m & q \end{vmatrix}$

5.2 Rank

(a) Rank and nullity

Recall given $T: V \to W$ linear map with $\dim V = n$ & $\dim W = m$,

$\operatorname{Im} T = \{ \underline{x}' \in W : \underline{x}' = T\underline{x} \text{ for some } \underline{x} \in V \}$ is a subspace of W

[check: $\underline{x}', \underline{y}' \in \operatorname{Im} T \Rightarrow \underline{x}' = T\underline{x}$ & $\underline{y}' = T\underline{y}$ for some $\underline{x}, \underline{y}$

$\Rightarrow \lambda \underline{x}' + \mu \underline{y}' = T(\lambda \underline{x} + \mu \underline{y}) \in \operatorname{Im} T$

& $\underline{0} = T(\underline{0}) \in \operatorname{Im} T$]

Note $\operatorname{rank}(T) = \dim \operatorname{Im}(T) \leq m$. Also $\operatorname{Ker} T = \{ \underline{x} \in V : T\underline{x} = \underline{0} \}$ is a subspace of V [check: $T(\underline{x}) = T(\underline{y}) = \underline{0} \Rightarrow T(\lambda \underline{x} + \mu \underline{y}) = \underline{0}$]

nullity, $\operatorname{null}(T) = \dim \operatorname{Ker} T \leq n$.

Theorem (Rank Nullity): For a linear map T as above,

$$\operatorname{rank}(T) + \operatorname{null}(T) = \dim \operatorname{Im}(T) + \dim \operatorname{Ker}(T) = n.$$

~~Proof~~ n: Extreme cases: (i) $T(\underline{x}) = \underline{0} \;\; \forall \underline{x} \in V$

$\operatorname{Ker} T = V$, dim n

$\operatorname{Im} T = \{\underline{0}\}$, dim 0

(ii) $T(\underline{x}) = \underline{x} \;\forall \underline{x} \in V$ $\quad \operatorname{Ker} T = \{\underline{0}\}$, dim 0 $\quad \operatorname{Im} T = V$ "identity map"

Proof: * non-examinable *

Let $\underline{e}_1, \dots, \underline{e}_k$ be a basis for $\ker T$, so $k = \dim \ker T = \text{null } T$.

[V, W general spaces, $\{\underline{e}_i\}$ general basis]

Extend to a basis of V by adding $\underline{e}_{k+1}, \dots, \underline{e}_n$ & claim

$\{T(\underline{e}_{k+1}), \dots, T(\underline{e}_n)\} = \mathcal{B}$ is a basis for $\text{Im } T$.

To check this: $\mathcal{B}$ spans, since general

$$\underline{x} = \sum_{i=1}^{n} x_i \underline{e}_i \quad \in V$$

$$T(\underline{x}) = \sum_{i=k+1}^{n} x_i T(\underline{e}_i).$$

$\mathcal{B}$ is also linearly independent, since

$$\sum_{i=k+1}^{n} \lambda_i T(\underline{e}_i) = \underline{0} \Rightarrow \sum_{i=k+1}^{n} \lambda_i \underline{e}_i \in \ker T$$

$$\Rightarrow \sum_{i=k+1}^{n} \lambda_i \underline{e}_i = \sum_{i=1}^{k} \mu_i \underline{e}_i \quad \text{for some } \mu_i$$

so,
$$\sum_{i=1}^{n} \nu_i \underline{e}_i = \underline{0} \quad \text{where } \nu_i = \begin{cases} \lambda_i & \text{for } 1 \le i \le k \\ -\mu_i & \text{for } k+1 \le i \le n \end{cases}$$

but this implies $\lambda_i = 0$ since $\{\underline{e}_i\}$ is a basis for V.

Hence $\dim \text{Im}(T) = n - k$

so $\text{rank } T + \text{null } T = n$. $\square$

Examples $V = \mathbb{R}^3$, $W = \mathbb{R}^4$

(i) Linear map given by matrix $M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & 1 & 0 \end{pmatrix}$

$\text{Im } M = \text{span cols of } M = \text{Span}\left\{\begin{pmatrix} 0 \\ 1 \\ -1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 1 \\ 0 \end{pmatrix}\right\}$ dim 2, so rank $M = 2$

L13.3

$$\underline{x} \in \ker M \iff \begin{aligned} x_1 + x_3 &= 0 \\ x_2 - x_3 &= 0 \\ -x_2 + x_3 &= 0 \\ x_1 + x_2 &= 0 \end{aligned} \iff \underline{x} = \lambda \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix} \text{ for any real } \lambda$$

dim 1

null $M = 1$, and $1+2=3$ ☺

(ii)

$$N = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & -1 & 0 \end{pmatrix} \quad \begin{aligned} \text{rank } N &= 3 \\ \text{null } N &= 0 \end{aligned}$$

(b) Rank & Inverses

Consider linear map $T : V \to W$ with properties as in part (a).

Suppose $\operatorname{Ker}(T) = \{\underline{0}\}$ or $\operatorname{rank}(T) = n$. (This requires $m \geq n$)

Then T has a left inverse, i.e. a linear map $S : W \to V$ with $ST : V \to V$ the identity.

As in proof of rank-nullity, consider $T(\underline{e_1}), T(\underline{e_2}), \ldots, T(\underline{e_n})$ basis for $\operatorname{Im}(T)$ where $\underline{e_1}, \underline{e_2}, \ldots, \underline{e_n}$ is basis for V (now $\operatorname{Ker}(T) = \{\underline{0}\}$)

Extend this to a basis for W by vectors $\underline{u_1}, \ldots, \underline{u}_{m-n}$.

A linear map is specified by its action on basis vectors, so define

$$S(T(\underline{e_i})) = \underline{e_i} \quad i = 1, \ldots, n$$
$$S(\underline{u}_a) = \underline{0} \quad a = 1, \ldots, m-n$$

ST is the identity map by construction (true for basis vecs $\underline{e_i}$ of V)

Example $V = \mathbb{R}^3$ $W = \mathbb{R}^4$

& $N = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \\ 1 & -1 & 0 \end{pmatrix}$ as in (ii) of part (a)

Let $\underline{e_i}$ & $\underline{f_a}$ be the standard basis vectors for $\mathbb{R}^3$ and $\mathbb{R}^4$ respectively.

L13.4

Im N contains $2\underline{f}_1 = 2\begin{pmatrix}1\\0\\0\\0\end{pmatrix} = \underline{C}_1 + \underline{C}_2 + \underline{C}_3$

$= N\underline{e}_1 + N\underline{e}_2 + N\underline{e}_3$,

$$2\underline{f}_4 = 2\begin{pmatrix}0\\0\\0\\1\end{pmatrix} = N\underline{e}_1 \overset{-}{\cancel{+}} N\underline{e}_2 - N\underline{e}_3, \text{ and}$$

$$2(\underline{f}_2 - \underline{f}_3) = 2\begin{pmatrix}0\\1\\-1\\0\end{pmatrix} = N\underline{e}_1 + N\underline{e}_2 - N\underline{e}_3$$

Extend to a basis of $W = \mathbb{R}^4$ by taking $\underline{u} = \underline{f}_3$.

Inverse P given by $P(N\underline{e}_i) = \underline{e}_i$ for $i = 1, 2, 3$

$P(\underline{f}_3) = \underline{0}$.

But by linearity:

$$P\underline{f}_1 = \tfrac{1}{2}(\underline{e}_1 + \underline{e}_2 + \underline{e}_3)$$

$$P\underline{f}_2 = \tfrac{1}{2}(\underline{e}_1 + \underline{e}_2 - \underline{e}_3)$$

$$P\underline{f}_3 = \underline{0}$$

$$P\underline{f}_4 = \tfrac{1}{2}(\underline{e}_1 - \underline{e}_2 - \underline{e}_3).$$

Hence $P = \frac{1}{2}\begin{pmatrix}1 & 1 & 0 & 1\\ 1 & 1 & 0 & -1\\ 1 & -1 & 0 & -1\end{pmatrix}$.

Check $PN = \frac{1}{2}\begin{pmatrix}1 & 1 & 0 & 1\\ 1 & 1 & 0 & -1\\ 1 & -1 & 0 & -1\end{pmatrix}\begin{pmatrix}1 & 0 & 1\\ 0 & 1 & -1\\ 0 & -1 & 1\\ 1 & -1 & 0\end{pmatrix}$

$= \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix}$ ✓

(c) Row & Column Rank

For an $m \times n$ matrix M_{pi} with $p = 1, \ldots, m$, $i = 1, \ldots, n$,

the columns of M are $\underline{C}_i \in \mathbb{R}^m$

& the rows of M are $\underline{R}_p \in \mathbb{R}^n$

& $(\underline{C}_i)_p = M_{pi} = (\underline{R}_p)_i$

L13.5 Jeez!

The col rank is the maximal number of linearly indep cols, i.e. dim of column space, the subspace the cols span, & this is the Im of M.

Similarly the row rank is the max. # of lin. indep. rows, i.e. dim of row space.

$$\text{col rank} = \text{row rank}$$

To see this, suppose $\underline{C}_i = \sum_\alpha U_{i\alpha} \underline{V_\alpha}$ $\quad \alpha = 1, \ldots, r$

for some vectors $\underline{V_\alpha} \in \mathbb{R}^m$

then $(\underline{c}_i)_p = \sum_\alpha U_{i\alpha} V_{p\alpha}$ where $V_{p\alpha} = (\underline{v}_\alpha)_p$

$\Rightarrow (\underline{R}_p)_i = \sum_\alpha V_{p\alpha} (\underline{u}_\alpha)_i$ where $\underline{u}_\alpha \in \mathbb{R}^n$

with $(\underline{u}_\alpha)_i = U_{i\alpha}$.

Now if $\{\underline{V_\alpha}\}$ is a basis for ~~row~~ col space, the same number of vectors $\{\underline{u}_\alpha\}$ span the row space. $\therefore$ row rank $\leq$ col rank

Now reverse argument

$$\text{col rank} \leq \text{row rank}.$$

* Proof above non-examinable *

~~5.8~~ L14.1

5.3 Determinants in $\mathbb{R}^n$ & $\mathbb{C}^n$

(2) Permutations & the ε symbol

A permutation σ on a set of size n is a bijection from $\{1, 2, \dots, n\}$ to itself (i.e. a one-to-one & onto map), specified by an ordered list $\sigma(1), \sigma(2), \dots, \sigma(n)$.

Permutations can be multiplied or composed (as maps or functions); the identity map is a permutation; if σ is a permutation then so is σ^{-1}. They form a group of order $n!$, the symmetric group.

A transposition τ is a permutation that swaps two elements and fixes all others: $\tau(p) = q$, $\tau(q) = p$, $p \neq q$

$$\& \ \tau(i) = i \quad \forall\, i \notin \{p, q\}.$$

We denote this $(p \ q)$.

Any permutation is a product of transpositions

Examples (i) $\sigma(1) = 3$, $\sigma(2) = 2$, $\sigma(3) = 4$, $\sigma(4) = 1$

then $\sigma = (14)(13)$ (read from right to left)

(ii) $\sigma(1) = 2$, $\sigma(2) = 3$, $\sigma(3) = 4$, $\sigma(4) = 1$

then $\sigma = (14)(13)(12)$.

(iii) $id = (13)(23)(13)(12)$ so representing a perm. as product of transpositions can't be done uniquely.

Proposition & Definition: If σ is the product of k transpositions $\sigma = \tau_k \cdots \tau_2 \tau_1$ then k is always even, or always odd, depending on σ.

$\varepsilon(\sigma) = (-1)^k = \pm 1$ is the signature or sign of σ

and we say σ is even or odd.

Examples as above:

(i) σ is even $\varepsilon(\sigma) = +1$

(ii) σ is odd $\varepsilon(\sigma) = -1$

(iii) $\varepsilon(\text{id}) = +1$.

The alternating or ε symbol in $\mathbb{R}^n$ or $\mathbb{C}^n$ is an n-index object (w.r.t. standard orthonormal bases) or tensor, defined by

$$\underbrace{\varepsilon_{ij\ldots l}}_{n \text{ indices}} = \begin{cases} +1 \text{ if } ij\ldots l \text{ is an } \underline{\text{even}} \text{ perm. of } 1, 2, \ldots, n \\ -1 \quad " \quad " \quad " \quad " \ \underline{\text{odd}} \quad " \quad " \quad " \\ 0 \text{ if any indices } i, j \text{ coincide} \end{cases}$$

Thus, for a permutation σ, $\varepsilon_{\sigma(1)\sigma(2)\cdots\sigma(n)} = \varepsilon(\sigma)$.

Note $\varepsilon_{ij\ldots l}$ is totally antisymmetric; exchanging any pair of indices changes its sign.

Examples In $\mathbb{R}^4$ $\varepsilon_{1234} = 1 = \varepsilon_{3241}$

but $\varepsilon_{2341} = -1$.

(b) Alternating forms & linear (in)dependence

Given vectors $\underline{v}_1, \ldots, \underline{v}_n$ in $\mathbb{R}^n$ or $\mathbb{C}^n$ we define the alternating form (of rank n) by its action on these vectors

$$\begin{aligned} \text{alt}(\underline{v}_1, \ldots, \underline{v}_n) &= [\underline{v}_1, \ldots, \underline{v}_n] \\ &= \varepsilon_{ij\ldots l}(\underline{v}_1)_i(\underline{v}_2)_j \cdots (\underline{v}_n)_l \quad \leftarrow \Sigma \text{ convention} \\ &= \sum_\sigma \varepsilon(\sigma) V_{\sigma(1)1} V_{\sigma(2)2} \cdots V_{\sigma(n)n} \end{aligned}$$

where $\underline{v}_a = V_{ia}\underline{e}_i$ or $(\underline{v}_a)_i = V_{ia}$ (Σ convention applies)

Properties (i) Multi-linear i.e. $[\underline{v}_1, \ldots, \underline{v}_{p-1}, \alpha\underline{u} + \beta\underline{w}, \underline{v}_{p+1}, \ldots, \underline{v}_n]$

$$= \alpha[\underline{v}_1, \ldots, \underline{u}, \ldots, \underline{v}_n] + \beta[\underline{v}_1, \ldots, \underline{w}, \ldots, \underline{v}_n].$$

(ii) Totally antisymmetric

$[\underline{v}_{\sigma(1)}, \dots, \underline{v}_{\sigma(n)}] = \varepsilon(\sigma)[\underline{v}_1, \dots, \underline{v}_n]$ Check this below.

(iii) $[\underline{e}_1, \dots, \underline{e}_n] = +1$ compare with $\varepsilon_{12\dots n} = +1$.

Properties (i), (ii) and (iii) fix alt completely.

Further properties following from these:

(iv) If $\underline{v}_p = \underline{v}_q$ for some $p \neq q$, then from (ii) $[\underline{v}_1, \dots, \underline{v}_p, \dots, \underline{v}_q, \dots, \underline{v}_n] = 0$.

(v) Using (iv) and (i) if $\underline{v}_p = \sum_{i \neq p} \lambda_i \underline{v}_i$, then $[\underline{v}_1, \dots, \underline{v}_n] = 0$.

(sub in and use (iv))

Properties (i) and (iii) are immediate. To check (ii) it is sufficient to check transpositions. For $p < q$ consider

$$[\underline{v}_1, \dots, \underline{v}_{p-1}, \underline{v}_q, \underline{v}_{p+1}, \dots, \underline{v}_{q-1}, \underline{v}_p, \underline{v}_{q+1}, \dots, \underline{v}_n]$$

$$= \sum_\sigma \varepsilon(\sigma) V_{\sigma(1)1} \cdots V_{\sigma(p)q} \cdots V_{\sigma(q)p} \cdots V_{\sigma(n)n}$$

$$= \sum_\sigma \varepsilon(\sigma) V_{\sigma'(1)1} \cdots V_{\sigma'(q)q} \cdots V_{\sigma'(p)p} \cdots V_{\sigma'(n)n}$$

where $\sigma' = \sigma\tau$ & $\tau = (p\ q)$ is a transposition.

But $\varepsilon(\sigma') = (-1)\,\varepsilon(\sigma)$ & $\sum_{\sigma'}$ is equivalent to $\sum_\sigma$ & result follows.

Proposition: $[\underline{v}_1, \dots, \underline{v}_n] \neq 0 \iff \underline{v}_1, \dots, \underline{v}_n$ linearly independent

To show this "$\Rightarrow$" consider property (v) above: if $\underline{v}_1, \dots, \underline{v}_n$ linearly dependent, then we can express some $\underline{v}_p$ as a linear combi of the others & $[\underline{v}_1, \dots, \underline{v}_n] = 0$.

To show "$\Leftarrow$" note that $\underline{v}_1, \dots, \underline{v}_n$ linearly indep means they span ($\mathbb{R}^n$ or $\mathbb{C}^n$) so we can write $\underline{e}_i = U_{ai}\underline{v}_a$. Then

$$[\underline{e_1}, \ldots, \underline{e_n}] = U_{a1} U_{b2} \cdots U_{cn} [\underline{v_a}, \underline{v_b}, \ldots, \underline{v_c}]$$

$$= U_{a1} \cdots U_{cn} \varepsilon_{ab\ldots c} [\underline{v_1}, \ldots, \underline{v_n}].$$

But $LHS = +1$, so $[\underline{v_1}, \ldots, \underline{v_n}] \neq 0$.

(c) Definition of determinant in $\mathbb{R}^n$ & $\mathbb{C}^n$

For an $n \times n$ matrix M with columns $\underline{C}_a = M \underline{e_a}$ define the determinant $\det M \in \mathbb{R}, \mathbb{C}$ by

$$\det M = [\underline{C_1}, \underline{C_2}, \ldots, \underline{C_n}]$$

$$= [M\underline{e_1}, \ldots, M\underline{e_n}]$$

$$= \varepsilon_{ij\ldots l} M_{i1} M_{j2} \cdots M_{ln}$$

$$= \sum_{\sigma} \varepsilon(\sigma) M_{\sigma(1)1} M_{\sigma(2)2} \cdots M_{\sigma(n)n}$$

Σ convention applies unless we say otherwise &

$\sum_\sigma$ means sum over all permutations ; we can write $\sigma \in S_n$

set (or group) of size $n!$ if there is confusion.

Determinant is also written

$$\det(M) = \begin{vmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{vmatrix}.$$

<u>Examples</u>

(i) General 2×2 case $\quad M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$.

only two perms

id & $\tau = (1,2)$ $\qquad \varepsilon_{12} = -\varepsilon_{21} = +1$

$$\det M = \varepsilon_{ij} M_{i1} M_{j2} = M_{11}M_{22} - M_{12}M_{21}$$

(ii) M is $n\times n$ & A is $(n-1)\times(n-1)$ with

$$M = \left(\begin{array}{c|c} A & \begin{matrix} 0 \\ \vdots \\ 0 \end{matrix} \\ \hline 0 \cdots 0 & 1 \end{array}\right) \ddagger \quad \begin{array}{ll} \varepsilon_{i_1 \cdots i_{n-1} n} & (\mathbb{R}^n \text{ or } \mathbb{C}^n) \\ = \varepsilon_{i_1 \cdots i_{n-1}} & (\mathbb{R}^{n-1} \text{ or } \mathbb{C}^{n-1}) \end{array}$$

$$\det(M) = \det(A)$$

By repeatedly applying similar idea $M = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ diagonal, then

$$\det(M) = \lambda_1 \lambda_2 \cdots \lambda_n$$

(iii) $M = \begin{pmatrix} i & 0 & 3 & 0 \\ 0 & 0 & 2i & 0 \\ 0 & 5i & 0 & -i \\ 2 & 0 & 0 & 1 \end{pmatrix}$ $\quad \det M = [i\underline{e_1} + 2\underline{e_4}, 5i\underline{e_3}, 3\underline{e_1} + 2i\underline{e_2}, -i\underline{e_3} + \underline{e_4}]$

$$= i[\underline{e_1}, 5i\underline{e_3}, \cancel{3\underline{e_1}} + 2i\underline{e_2}, \cancel{-i\underline{e_3}} + \underline{e_4}]$$

$$+ 2[\underline{e_4}, 5i\underline{e_3}, 3\underline{e_1} + 2i\underline{e_2}, \cancel{-i\underline{e_3}} + \cancel{\underline{e_4}}]$$

$$= i(5i)(2i)[\underline{e_1}, \underline{e_3}, \underline{e_2}, \underline{e_4}] = -10i(-1)[\underline{e_1}, \underline{e_2}, \underline{e_3}, \underline{e_4}]$$

$$= 10i.$$

(d) Transpose Property

- For any $n \times n$ matrix M (entries in $\mathbb{R}$ or $\mathbb{C}$)

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

$$\det M = [\underline{R_1}, \underline{R_2}, \dots, \underline{R_n}] = \underbrace{\varepsilon_{ij\dots l}}_{n \text{ indices}} M_{1i} M_{2j} \cdots M_{nl}$$

$$= \sum_{\sigma} \varepsilon(\sigma) M_{1\sigma(1)} M_{2\sigma(2)} \cdots M_{n\sigma(n)}$$

[recall $(M^T)_{ij} = M_{ji}$ & rows $\underline{R_a}$ of M are cols of M^T]

- <u>Proof</u>: Note that for any permutation ρ

$$M_{\sigma(1)1} M_{\sigma(2)2} \cdots M_{\sigma(n)n} = M_{\sigma'(1)\rho(1)} \cdots M_{\sigma'(n)\rho(n)}$$

where $\sigma' = \sigma\rho$ (reorder factors)

For each term in formula, choose ρ to be σ^{-1} so $\sigma' = id$, &

note $\varepsilon(\rho) = \varepsilon(\sigma)$ & $\sum_{\sigma}$ is equiv. to $\sum_{\sigma^{-1} = \cancel{\sigma} \rho}$, giving formula for $\det M^T$. □

- Now know that $\det M$ is a totally antisymmetric function of columns and rows of M. Use this via row / col operations to simplify $\det M$.

$\underline{C_a} \to \underline{C_a} + \lambda \underline{C_b} \quad b \neq a$

or $\underline{R_a} \to \underline{R_a} + \lambda \underline{R_b} \quad b \neq a$

leave $\det M$ unchanged. In addition $\det M = 0$ if there is some linear relation with not all coeffs. zero obeyed by rows/cols.

-

Multiplicative property

- For any $n \times n$ matrices M & N, $\det(MN) = \det M \cdot \det N$.

Lemma: $\varepsilon_{i_1 \cdots i_n} M_{i_1 a_1} \cdots M_{i_n a_n} = (\det M)\, \varepsilon_{a_1 a_2 \cdots a_n}$

Proof: LHS $= [M\underline{e}_{a_1}, \ldots, M\underline{e}_{a_n}]$. This is zero if any $a_i = a_j$ for $i \neq j$, but then RHS also zero. Otherwise

$$LHS = [M\underline{e}_{\sigma(1)}, \ldots, M\underline{e}_{\sigma(n)}]$$ for some permutation σ with $\sigma(r) = a_r$.

So $LHS = \varepsilon(\sigma)[M\underline{e}_1, \ldots, M\underline{e}_n] = \varepsilon(\sigma)(\det M) = RHS$.

Proof of mult. property:

$$\begin{aligned}
LHS &= \varepsilon_{i_1 \cdots i_n} (MN)_{i_1 1} (MN)_{i_2 2} \cdots (MN)_{i_n n} \\
&= \varepsilon_{i_1 \cdots i_n} M_{i_1 k_1} \cdots M_{i_n k_n} \\
&\quad \times N_{k_1 1} \cdots N_{k_n n} \qquad (\Sigma \text{ convention}) \\
&= \det(M)\, \varepsilon_{k_1 \cdots k_n} \qquad (\text{by Lemma}) \\
&\quad \times N_{k_1 1} \cdots N_{k_n n} \\
&= \det(M) \det(N), \text{ as required}
\end{aligned}$$

Example: If R is an orthogonal matrix (real & $n \times n$) then

$$R^T R = I \Rightarrow \det(R^T R) = \det(R^T)\det(R) = \det(R)^2$$
$$= \det(I) = 1$$

$\therefore \det(R) = \pm 1$

LIS.4

5.4 Minors, Cofactors & Inverses

(a) Expanding determinants

For a matrix M, select a particular column $\underline{C}_a$ (fixed a) & write $\underline{C}_a = \sum_i M_{ia} \underline{e}_i$ in def^n:

$$\det M = [\underline{C}_1, \dots, \underline{C}_n] = [\underline{C}_1, \dots, \underline{C}_{a-1}, \sum_i M_{ia}\underline{e}_i, \underline{C}_{a+1}, \dots, \underline{C}_n]$$

$$= \sum_i M_{ia}\Delta_{ia} \qquad \text{(no sum on } a)$$

where $\Delta_{ia} = [\underline{C}_1, \dots, \underline{C}_{a-1}, \underline{e}_i, \underline{C}_{a+1}, \dots, \underline{C}_n]$

$$= \det \left(\begin{array}{c|c|c} A & \begin{matrix}0\\ \vdots \\ 0\end{matrix} & B \\ \hline 0 \cdots & 0\ 1\ 0 & \cdots 0 \\ \hline C & \begin{matrix}0\\ \vdots \\ 0\end{matrix} & D \end{array} \right) \leftarrow \text{row } i$$

$\uparrow$ col a

(using only block form & 0's in row i appear using antisymmetry)

Δ_{ia} is called the <u>cofactor</u>. It is the determinant of the matrix obtained from M by replacing ia entry by 1 & all other entries in row i, col a, by 0.

Now by re-ordering rows & columns

$$\Delta_{ia} = (-1)^{n-a-1}(-1)^{n-i-1} \times \det \left(\begin{array}{cc|c} A & B & 0 \\ & & \vdots \\ C & D & 0 \\ \hline 0 & \cdots\ 0 & 1 \end{array} \right)$$

$= (-1)^{a+i} M^{ia}$ where M^{ia} is called the <u>minor</u> & is the det of the $(n-1)\times(n-1)$ matrix obtained by deleting row i & col a from M.

Hence $\det M = \sum_i M_{ia}\Delta_{ia} = \sum_i (-1)^{i+a} M_{ia} M^{ia}$ <u>a fixed</u>

expansion about column a

or $\det M = \sum_a M_{ia} \Delta_{ia} = \sum_a (-1)^{i+a} M_{ia} M^{ia}$ <u>i fixed</u>

expansion about row i

<u>Example</u>

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

Expand around 2^{nd} row

$M_{23} = 2i$

$$M^{23} = \begin{vmatrix} i & 0 & 0 \\ 0 & 5i & -i \\ 2 & 0 & 1 \end{vmatrix}$$

$$\det M = (-1)^{2+3} M_{23} M^{23} = -2i \begin{vmatrix} i & 0 & 0 \\ 0 & 5i & -i \\ 2 & 0 & 1 \end{vmatrix}$$

But then $\begin{vmatrix} i & 0 & 0 \\ 0 & 5i & -i \\ 2 & 0 & 1 \end{vmatrix} = i \begin{vmatrix} 5i & -i \\ 0 & 1 \end{vmatrix} = i(5i) = -5$ (expand around 1^{st} row)

So $\det M = (-2i)(-5i) = 10i$.

Formulae above in general known as *Laplace expansion formulae*.

(b) Adjugates & Inverses

We showed above $\sum_i M_{ia} \Delta_{ia} = [\underline{C_1}, \ldots, \underline{C_a}, \ldots, \underline{C_n}] = \det M$

But now by exactly similar argument

$$\sum_i M_{ib} \Delta_{ia} = [\underline{C_1}, \ldots, \underline{C_{a-1}}, \underline{C_b}, \underline{C_{a+1}}, \ldots, \underline{C_n}] = 0 \text{ for } a \neq b$$

Hence $M_{ib} \Delta_{ia} = \det M \, \delta_{ab}$ for any a, b.

The <u>adjugate</u> matrix

$$\widetilde{M} = \mathrm{adj}(M) = \Delta^T$$

then satisfies $\widetilde{M} M = (\det M) I$.

(16.1)

5.5 Systems of Linear Equations

(a) Introduction and Nature of Solutions

- Consider system of n linear equations in n simultaneous unknowns written in vector/matrix form $A\underline{x} = \underline{b}$.

There are three possibilities:

(i) $\det A \neq 0 \Rightarrow A^{-1}$ exists & $\underline{x} = A^{-1}\underline{b}$ is the unique solution

(ii) $\det A = 0$ & $b \notin \operatorname{Im} A \Rightarrow$ no solution

(iii) $\det A = 0$ & $b \in \operatorname{Im} A \Rightarrow \infty$ solutions $\underline{x} = \underline{x}_0 + \underline{u}$

with $\underline{x}_0$ a particular solution and $\underline{u} \in \operatorname{Ker} A$

Elaboration: A solution exists iff $A\underline{x}_0 = \underline{b} \iff \underline{b} \in \operatorname{Im} A$ for some $\underline{x}_0$

- Then $\underline{x}$ also a solution iff $A(\underline{x} - \underline{x}_0) = \underline{0} \iff \underline{x} - \underline{x}_0 \in \operatorname{Ker} A$

i.e. $\underline{x} = \underline{x}_0 + \underline{u}$ where $\underline{u}$ is a solution of the homogeneous problem

$A\underline{u} = \underline{0} \iff \underline{u} \in \operatorname{Ker} A$

Now $\det A \neq 0 \iff \operatorname{Im} A = \mathbb{R}^n$

$\iff \operatorname{Ker} A = \{\underline{0}\}$

so in case (i) we have a unique solution which can be computed using A^{-1}.

- But $\det A = 0 \iff \operatorname{rank} A < n \iff \operatorname{null} A > 0$

Now could have either $\underline{b} \notin \operatorname{Im}(A)$ as in case (ii) or $\underline{b} \in \operatorname{Im} A$ as in case (iii).

If $\underline{u}_1, \dots \underline{u}_k$ basis for $\operatorname{Ker} A$ ($k = \operatorname{null} A$) then general sol^n in case (iii) has form $\underline{x} = \underline{x}_0 + \sum_{i=1}^{k} \lambda_i \underline{u}_i$

(b) Extended example

$A\underline{x} = \underline{b}$ $\quad (n=3)$

- $$A = \begin{pmatrix} 1 & 1 & a \\ a & 1 & 1 \\ 1 & a & 1 \end{pmatrix} \quad \& \quad \underline{b} = \begin{pmatrix} 1 \\ \cancel{x}c \\ 1 \end{pmatrix}$$

L16.2

$$\det A = \begin{vmatrix} 1-a & 1-a & a \\ a-1 & 0 & 1 \\ 0 & a-1 & 1 \end{vmatrix} \quad \text{by col. ops } \underline{c_1 \to c_1 - c_3},\ \underline{c_2 \to c_2 - c_3}$$

$$= (a-1)^2 \begin{vmatrix} -1 & -1 & a \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} = (a-1)^2 \begin{vmatrix} 0 & 0 & a+2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{vmatrix} \quad \text{by } \underline{R_1} \to \underline{R_1} + \underline{R_2} + \underline{R_3}$$

$$= \underline{(a-1)^2(a+2)}.$$

For $a \neq 1$ and $a \neq -2$, $\det A \neq 0$ and A^{-1} exists.

$A^{11} = \begin{vmatrix} 1 & 1 \\ a & 1 \end{vmatrix} = 1-a \quad \begin{pmatrix} \cancel{1} & \cancel{1} & \cancel{a} \\ \cancel{a} & 1 & 1 \\ \cancel{1} & a & 1 \end{pmatrix} \qquad A^{12} = \begin{vmatrix} a & 1 \\ 1 & 1 \end{vmatrix} = a-1 \quad \begin{pmatrix} \cancel{1} & \cancel{1} & \cancel{a} \\ a & \cancel{1} & 1 \\ 1 & \cancel{a} & 1 \end{pmatrix}$

$A^{13} = \begin{vmatrix} a & 1 \\ 1 & a \end{vmatrix} = a^2-1$

IW/

Matrix of cofactors $\Delta_{ij} = (-1)^{i+j} M^{ij}$ is given by

$$\Delta = \begin{pmatrix} 1-a & 1-a & a^2-1 \\ a^2-1 & 1-a & 1-a \\ 1-a & a^2-1 & 1-a \end{pmatrix}$$

Then $\tilde{A} = \Delta^T$ & $A^{-1} = \frac{1}{|A|}\tilde{A} = \frac{1}{(1-a)(a+2)} \begin{pmatrix} 1 & -(a+1) & 1 \\ 1 & 1 & -(1+a) \\ -(1+a) & 1 & 1 \end{pmatrix}$

and solution of $A\underline{x} = \underline{b}$ for any $\underline{c}$ is $\underline{x} = A^{-1}\underline{b} = \frac{1}{(1-a)(a+2)} \begin{pmatrix} 2-c-ca \\ c-a \\ c-a \end{pmatrix}$

- $\underline{a=1}$ $\quad A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \operatorname{Im} A = \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$\operatorname{Ker} A = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$b \in \operatorname{Im} A$ iff $c = 1$ & then p.i. is $\underline{x_0} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

General solution is $\underline{x} = \underline{x}_0 + \underline{u} = \begin{pmatrix} 1-\lambda-\mu \\ \lambda \\ \mu \end{pmatrix}$

This solution (a plane) holds for $a=1$ & $c=1$.

No solution for $a=1$, $c\neq 1$

- $\underline{a=-2}$ $\quad A=\begin{pmatrix} 1 & 1 & -2 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{pmatrix}$ $\quad \operatorname{Im} A = \operatorname{Span}\left\{\begin{pmatrix}1\\-2\\1\end{pmatrix}, \begin{pmatrix}1\\1\\-2\end{pmatrix}\right\}$

& $\underline{b} \in \operatorname{Im}(A)$ iff $c=-2$ (in general test wether this vector lies in span using scalar triple product.

Particular solution $\underline{x}_0 = \begin{pmatrix}1\\0\\0\end{pmatrix}$.

$\operatorname{Ker} A = \left\{\lambda\begin{pmatrix}1\\1\\1\end{pmatrix}\right\}$ & general solution $\underline{x} = \begin{pmatrix}1+\lambda\\ \lambda \\ \lambda\end{pmatrix}$ a line

for $a=-2$, $c=-2$ with no solution if $a=-2$, $c\neq -2$

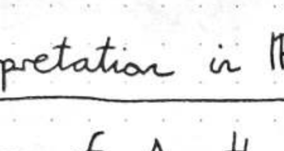
(c) Geometrical interpretation in $\mathbb{R}^3$

If $\underline{R_1}$, $\underline{R_2}$, $\underline{R_3}$ rows of A then

$A\underline{u} = \underline{0}$ homogeneous problem $\iff \begin{cases} \underline{R_1}\cdot\underline{u} = 0 \\ \underline{R_2}\cdot\underline{u} = 0 \\ \underline{R_3}\cdot\underline{u} = 0 \end{cases}$ equations for planes through $\underline{0}$ with normals R_i

rank$(A) = 3 \Rightarrow$ normals are lin. indep & planes intersect at $\underline{0}$
(det $A \neq 0$)

rank$(A) = 2 \Rightarrow$ normals lie in a plane; normal to plane through $\underline{0}$ lies on all planes
(det $A = 0$)



top down view onto normals' plane

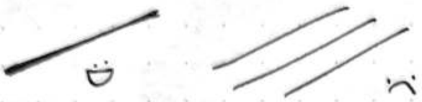
rank$(A) = 1 \Rightarrow$ normals all parallel & planes coincide

But now consider inhomogeneous problem

$A\underline{x} = \underline{b} \iff \begin{cases} \underline{R_1}\cdot\underline{x} = b_1 \\ \underline{R_2}\cdot\underline{x} = b_2 \\ \underline{R_3}\cdot\underline{x} = b_3 \end{cases}$ equations for planes not through $\underline{0}$ in general rank$(A)=3 \Rightarrow$ planes still intersect in a unique point

rank$(A)=2 \Rightarrow$ planes may intersect in a line as before,
but may not e.g. 

Similarly, rank$(A)=1$ planes may coincide as before or they may not

 depending on $\underline{b}$.

(d) Gaussian elimination

Consider a system of m equations in n unknowns $x_1, \dots, x_n$

$$\begin{cases} A_{11}x_1 + A_{12}x_2 + \dots + A_{1n}x_n = b_1 \\ A_{21}x_1 + A_{22}x_2 + \dots + A_{2n}x_n = b_2 \\ \vdots \qquad \ddots \qquad \vdots \\ A_{m1}x_1 + A_{m2}x_2 + \dots + A_{mn}x_n = b_n \end{cases}$$

Assume $A_{11} \neq 0$ (otherwise rearrange order)

& subtract off multiple of first equation to make coeffs. of x_1 vanish in remaining equations.

Then repeat with second equation to eliminate x_2, and so on...

We end up with
$$\begin{aligned} A_{11}x_1 + \dots + A_{1n}x_n &= b_1 \\ A^{(2)}_{22}x_2 + \dots + A^{(2)}_{2n}x_n &= b^{(2)}_2 \\ \ddots \quad A^{(r)}_{rr}x_r + \dots + A^{(r)}_{rn}x_n &= b^{(r)}_r \\ 0 &= b^{(r)}_{r+1} \\ &\vdots \\ 0 &= b^{(r)}_m \end{aligned}$$

where superscript (r) denotes values obtained at stage r.

L17.1

Possibilities: (i) $r = n \leq m$ & $b_{r+i}^{(r)} = 0$

Then unique solution : start by finding x_n & sub back into others.

(ii) $r \overset{?}{=} n < m$ & $b_{r+i}^{(r)} \neq 0$ for some i

Then there are no solutions (inconsistent)

(iii) $r = m$, $x_{r+1}, \dots, x_n$ undetermined & for any given values we can solve for $x_1, \dots, x_r$, so ∞ many solutions.

Connection with matrix approach – consider case (i)

New set of equations has form $M\underline{x} = \underline{b}$ with

$$M = \begin{pmatrix} A_{11} & \cdots & * & * \\ 0 & A_{22}^{(2)} & \ddots & \vdots \\ \vdots & & & A_{nn}^{(n)} \\ 0 & 0 & \cdots & \end{pmatrix} \quad \& \quad \underline{b} = \begin{pmatrix} b_1 \\ b_2^{(2)} \\ \vdots \\ b_n^{(n)} \end{pmatrix}$$

M is upper triangular $\det M = \sum_{\sigma} \varepsilon(\sigma) M_{\sigma(1)1} \cdots M_{\sigma(n)n}$

$$= A_{11} A_{22}^{(2)} \cdots A_{nn}^{(n)} \neq 0$$

[for any $\sigma \neq id$ have $M_{\sigma(j)j} = 0$ for some j]

So M is invertible and there is a unique sol^n.

In fact $\det M = \pm \det A$, since operations used change det by at most its sign.

Gaussian elimination is very effective computationally ☺

6. Eigenvalues & Eigenvectors

6.1 Introduction

Consider a linear map $T: V \to V$

(a) Definitions

with V a real or complex vector space.

A vector $\underline{v} \in V$ with $\underline{v} \neq \underline{0}$ is an eigenvector of T with eigenvalue λ if $T\underline{v} = \lambda \underline{v}$.

We consider $V = \mathbb{R}^n$ or $\mathbb{C}^n$ & T given by $n \times n$ matrix A, then

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

& for a given λ this is satisfied for some $\underline{v} \neq \underline{0}$ iff

$$\det(A - \lambda I) = 0 \qquad \underline{\text{characteristic equation}}$$

Thus, λ is an eigenvalue iff it is a root of the characteristic polynomial defined by $\chi_A(t) = \det(A - tI)$.

We find eigenvalues by finding the roots of characteristic equation & then determining the corresponding eigenvectors.

<u>(b) Examples</u>

(i) $V = \mathbb{C}^2$ & $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & i \\ -i & 2-\lambda \end{vmatrix} = (2-\lambda)^2 - 1$$

$= 0$ iff $\lambda = 2 \pm 1 = 1$ or 3.

To find eigenvectors $\underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$:

$\underline{\lambda = 1}$ $\quad (A - I)\underline{v} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0}$

$$\Rightarrow \underline{v} = \alpha \begin{pmatrix} 1 \\ i \end{pmatrix} \text{ for any complex } \alpha \neq 0.$$

$\underline{\lambda = 3}$ $\quad (A - 3I)\underline{v} = \begin{pmatrix} -1 & i \\ -i & -1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0}$

$$\Rightarrow \underline{v} = \beta \begin{pmatrix} 1 \\ -i \end{pmatrix} \text{ for any complex } \beta \neq 0.$$

(ii) $V = \mathbb{R}^2$ $\quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad$ [shear]

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

$= 0$ iff $\lambda = 1$ (repeated root)

Eigenvectors: $(A - I)\underline{v} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0}$

$\Rightarrow \underline{v} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\quad$ Only one linearly indep. eigenvector $\therefore$

L17.3

(iii) $V = \mathbb{R}^2$ or $\mathbb{C}^2$ & $U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

- rotation through θ in $\mathbb{R}^2$

$$\det(U - \lambda I) = \begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix} = (\cos\theta - \lambda)^2 + \sin^2\theta$$

$$(= \lambda^2 - 2\lambda\cos\theta + 1)$$

$$= 0 \text{ iff } \lambda = \cos\theta \pm i\sin\theta = e^{\pm i\theta}$$

Hence no real eigenvalues unless $\theta = n\pi$.

- Eigenvectors for eigenvalues $e^{\pm i\theta}$ are $\begin{pmatrix} 1 \\ \pm i \end{pmatrix}$.

(iv) $V = \mathbb{C}^n$ $\qquad A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ diagonal

$\det(A - \lambda I) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)$, so
eigenvalues are λ_i and eigenvectors are $\underline{e_i}$ (standard basis)

(c) Deductions involving $\chi_A(t)$

- Recall, Fundamental Theorem of Algebra states that any polynomial

$$p(z) = \sum_{j=0}^{n} c_j z^j \qquad c_j \in \mathbb{C}, \; c_n \neq 0$$

can be factorised over $\mathbb{C}$:

$$p(z) = c_n (z - w_1) \cdots (z - w_n) \qquad w_i \in \mathbb{C}$$

& $p(z) = 0$ iff $z = w_i$ for some i, i.e. there are n complex roots, counted with multiplicity [w a root with multiplicity r if $(z-w)^r$ is a factor of p but $(z-w)^{r+1}$ is not].

- Consider $\chi_A(t)$ for $n \times n$ matrix A as map $\mathbb{C}^n \to \mathbb{C}^n$,

then $\chi_A(t) = \det(A - tI) = \varepsilon_{j_1 \cdots j_n}(A_{j_1 1} - \delta_{j_1 1} t) \cdots (A_{j_n n} - \delta_{j_n n} t)$

$$= c_n (t - \lambda_1) \cdots (t - \lambda_n).$$

where $c_n = (-1)^n$ & $\lambda_1, \ldots, \lambda_n$ are eigenvalues.

Now observe

(i) An $n \times n$ matrix A has n eigenvalues in $\mathbb{C}$ (counted with multiplicity)

(ii) Terms of order t^{n-1} arise only from factors with $j_r = r$ giving

$$-c_n(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = (-1)^{n-1}(A_{11} + A_{22} + \cdots + A_{nn})$$

i.e. $\sum_i \lambda_i = \sum_i A_{ii}$

sum of eigenvalues $= \mathrm{tr}(A)$.

(iii) $c_0 = \chi_A(0) = \det A = c_n(-1)^n \lambda_1 \cdots \lambda_n$

so product of eigenvalues $= \det A$

(iv) If A is real, $\chi_A(t)$ has real coefficients, so

$\chi_A(t) = 0 \iff \chi_A(\bar{t}) = 0$, i.e. non-real roots appear in conjugate pairs.

6.2 Eigenspaces & Multiplicities

(a) Definitions

For an eigenvalue λ, of a matrix A, we define the eigenspace

$$E_\lambda = \{ \underline{v} : A\underline{v} = \lambda \underline{v} \} = \mathrm{Ker}(A - \lambda I)$$

This is a subspace of $\mathbb{C}^n$ consisting of all eigenvectors & $\underline{0}$

The geometric multiplicity of λ is defined as

$$m_\lambda = \dim E_\lambda$$

i.e. the # linearly independent eigenvectors with that eigenvalue

The algebraic multiplicity is the multiplicity of λ as a root of $\chi_A(t)$, denoted by M_λ.

For any eigenvalue λ, $M_\lambda \geq m_\lambda \geq 1$.

4/9.1 Deep in - diagonalisation

Examples

$$A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ 1 & -2 & 0 \end{pmatrix} \quad \begin{matrix} \lambda_1 = 5 \\ \lambda_2 = \lambda_3 = -3 \end{matrix} \quad \begin{matrix} M_5 = m_5 = 1 \\ M_{-3} = m_{-3} = 2 \end{matrix}$$

$$P = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, \quad P^{-1} = \frac{1}{8}\begin{pmatrix} 1 & 3 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{pmatrix}$$

$$\& \quad P^{-1}AP = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}.$$

Consider instead

$$A = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix}$$ found $\lambda = -2$ only root of charac. equation

Now $M_{-2} = 3$ but $m_{-2} = 2 < 3 = M_{-2}$

$\therefore$ matrix not diagonalisable.

Note in this case, if $\exists P$ with $P^{-1}AP = \begin{pmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix} = -2I$

$\Rightarrow A = -2PIP^{-1} = -2I$, contradiction ☺

(c) Similarity In general, matrices A & B are similar if they are related by $B = P^{-1}AP$ for some matrix P. We will see later that similar matrices represent same linear map w.r.t. different bases.

Proposition: If A & B are similar, then

(i) $\operatorname{tr} B = \operatorname{tr} A$ (ii) $\det B = \det A$

(iii) $\chi_B = \chi_A$

Proof: $\operatorname{tr} B = B_{ii} = P^{-1}_{ij} A_{jk} P_{ki} = A_{jk} \delta_{jk} = \operatorname{tr} A$

$\det B = \det P^{-1} \det A \det P = \det A$, duh

$\det(B - tI) = \det(P^{-1}AP - tI) = \det(P^{-1}AP - tP^{-1}IP)$

L19.2 $\quad = \det\{P^{-1}[A - tI]P\}$

$\quad = \det(A - tI)$ ☺

For the special case where $B = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ these properties imply $\operatorname{tr} A = \sum \lambda_i \quad \det A = \prod \lambda_i$

& $\chi_A(t) = (\lambda_1 - t) \cdots (\lambda_n - t)$

(d) Random proof that $M_\lambda \geq m_\lambda$

* non-examinable *

Let λ be an eigenvalue for A, and choose a basis $\underline{v}_1, \ldots, \underline{v}_r$ for E_λ where $r = m_\lambda$.

Extend this by any vectors $\underline{w}_{r+1}, \ldots, \underline{w}_n$ that give a basis for $V = \mathbb{R}^n$ or $\mathbb{C}^n$.

Define $n \times n$ matrix P with

$$\underline{C}_i(P) = \underline{v}_i \quad \text{for } 1 \leq i \leq r$$

$$\underline{C}_a(P) = \underline{w}_a \quad \text{for } r+1 \leq a \leq n$$

Then $A\underline{C}_i(P) = A\underline{v}_i = \lambda \underline{v}_i$ by definition

$$A\underline{C}_a(P) = A\underline{w}_a = \sum_{i=1}^{r} B_{ia}\underline{v}_i + \sum_{b=r+1}^{n} B_{ba}\underline{w}_b$$

In matrix form

$$AP = PB$$

where $B_{ij} = \lambda\delta_{ij} \qquad \left.\begin{matrix} B_{ia} \\ B_{ab} \end{matrix}\right\}$ unknown.

$B_{ai} = 0$

Hence $P^{-1}AP = B = \left(\begin{array}{c|c} \lambda I & ////\!/ \\ \hline 0 & \hat{B} \end{array}\right)$ block form

where $\hat{B}$ has entries B_{ab}. [P invertible since cols are basis]

Now $\chi_A(t) = \chi_B(t) = \det\left(\begin{array}{c|c} \lambda - tI & /// \\ \hline 0 & \hat{B} - tI \end{array}\right)$

$$= (\lambda - t)^r \det(\hat{B} - tI)$$

[expanding about columns] so we are done. □

L19.3

6.4 Hermitian & Symmetric Matrices

Recall, A $(n\times n)$ is hermitian iff

$$A^\dagger = \overline{A}^T = A \quad \text{or} \quad A_{ij} = \overline{A_{ji}}.$$

A real & symmetric matrix is a special case

$$\overline{A} = A \;\&\; A^T = A.$$

Recall definition of complex inner product of vectors $\underline{v}, \underline{w} \in \mathbb{C}^n$:

$$\underline{v}^\dagger \underline{w} = \overline{v_i}\, w_i$$

If $\underline{v}, \underline{w} \in \mathbb{R}^n$ this reduces to the real inner product.

Observe that if A is hermitian, then

$$(A\underline{v})^\dagger \underline{w} = \underline{v}^\dagger (A\underline{w})$$

since $\underline{v}^\dagger A^\dagger \underline{w} = \underline{v}^\dagger A \underline{w}$

or $\overline{v_j}\,\overline{A_{ij}}\, w_i = \overline{v_j}\, A_{ji}\, w_i$.

<u>Theorem</u>: For an $n\times n$ hermitian matrix A,

(i) Every eigenvalue is <u>real</u>

(ii) Eigenvectors $\underline{v}, \underline{w}$ with distinct eigenvalues (λ, μ resp, with $\lambda \neq \mu$) are orthogonal $\underline{v}^\dagger \underline{w} = 0$.

(iii) If A is real and symmetric then for each eigenvalue λ we can choose a real eigenvector $\underline{v}$ & (ii) becomes $\underline{v}^T \underline{w} = \underline{v}\cdot\underline{w} = 0$ (with $\underline{w}$ a real evector for μ)

<u>Proof</u>: (i) From observation: if $\underline{v}$ is an evec with eval λ then $\underline{v}^\dagger (A\underline{v}) = (A\underline{v})^\dagger \underline{v} \Rightarrow \underline{v}^\dagger(\lambda\underline{v}) = (\lambda\underline{v})^\dagger \underline{v}$

$\Rightarrow \lambda\, \underline{v}^\dagger v = \overline{\lambda}\, \underline{v}^\dagger \underline{v}$ but $\underline{v} \neq 0$ so $\lambda = \overline{\lambda}$.

(ii) $\underline{v}^\dagger (A\underline{w}) = (A\underline{v})^\dagger \underline{w} \Rightarrow \mu \underline{v}^\dagger \underline{w} = (\lambda\underline{v})^\dagger \underline{w}$

$$= \lambda\, \underline{v}^\dagger \underline{w} \qquad \text{since } \lambda \in \mathbb{R}$$

Hence if $\lambda \neq \mu$, $\underline{v}^\dagger \underline{w} = 0$.

(iii) Given $A\underline{v} = \lambda \underline{v}$ with $\underline{v} \in \mathbb{C}^n$

write $\underline{v} = \underline{u} + i\underline{u}'$ with $\underline{u}, \underline{u}' \in \mathbb{R}^n$

& then, for A & λ real, have

$A\underline{u} = \lambda \underline{u}$ & $A\underline{u}' = \lambda \underline{u}'$, so $\underline{u}$ & $\underline{u}'$ are eigenvectors with eval λ unless they are zero. But $\underline{v} \neq \underline{0}$ so $\underline{u} \neq \underline{0}$ or $\underline{u}' \neq \underline{0}$.

<u>Example</u> Refer to §6.1 (b)(i)

recall $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$, $A^\dagger = A$

evals $\lambda = 1$ evec $\begin{pmatrix} 1 \\ i \end{pmatrix} = \underline{v}_1$

$\lambda = 3$ evec $\begin{pmatrix} 1 \\ -i \end{pmatrix} = \underline{v}_2$

$\underline{v}_1^\dagger \underline{v}_2 = \begin{pmatrix} 1 & -i \end{pmatrix} \begin{pmatrix} 1 \\ -i \end{pmatrix} = 0$ ✓

Now consider $P = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ orthonormal columns

Then $P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$ but $P^\dagger = P^{-1}$ is unitary.

L20.1

Recall $\mathcal{B}_\lambda$ is basis for eigenspace with eigenvalue λ.

- It is straightforward to construct an orthonormal basis for $\mathcal{B}_\lambda$ by following the Gram–Schmidt construction:

Let $\mathcal{B}_\lambda = \{ \underline{w}_1, \dots, \underline{w}_k \}$ so $k = \dim E_\lambda = m_\lambda$.

Rescale $\underline{w}_1$ by defining $\underline{u}_1 = \frac{1}{|\underline{w}_1|} \underline{w}_1$ where $|\underline{w}_1|^2 = \underline{w}_1^\dagger \underline{w}_1$.

Define $\underline{w}_j' = \underline{w}_j - (\underline{u}_1^\dagger \underline{w}_j)\underline{u}_1 \quad j = 2, \dots, k$

so by construction $\underline{u}_1^\dagger \underline{w}_j' = 0$.

But $\underline{u}_1, \underline{w}_2', \dots, \underline{w}_k'$ is a basis for E_λ since has same span.

- Now carry on:

define $\underline{u}_2 = \frac{\underline{w}_2'}{|\underline{w}_2'|}$ = unit vector & then

$$\underline{w}_j'' = \underline{w}_j' - (\underline{u}_2^\dagger \underline{w}_j')\underline{u}_2 \quad j = 3, \dots, k$$

so by construction $\underline{u}_1^\dagger \underline{w}_j'' = \underline{u}_2^\dagger \underline{w}_j'' = 0$ &

$\underline{u}_1, \underline{u}_2, \underline{w}_3'', \dots, \underline{w}_k''$ is still a basis.

Repeat process to arrive at an orthonormal basis

$\underline{u}_1, \dots, \underline{u}_k$.

- Gram–Schmidt orthogonalisation or orthonormalisation.

[small variations – sometimes rescale at end of process].

Example $\underline{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\underline{w}_2 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$, $\underline{w}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$, $\underline{w}_2' = \underline{w}_2 - (\underline{u}_1 \cdot \underline{w}_2)\underline{u}_1$

$$= \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} - \frac{1}{\sqrt{2}}(-1)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

-

$$\underline{w}_3' = \underline{w}_3 - (\underline{u}_1 \cdot \underline{w}_3)\underline{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{\sqrt{2}}(2)\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Next step:

$$|\underline{w_2}'|^2 = \frac{1}{4}(1+4+1) = \frac{3}{2}$$

so $\underline{u_2} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$.

$$\underline{w_3}'' = \underline{w_3}' - (\underline{u_2}\cdot\underline{w_3}')\underline{u_2} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{\sqrt{6}}(2)\frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$$

& $\underline{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}$.

From results above, for a hermitian matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_r$ we can choose an orthonormal basis for each eigenspace $B_{\lambda_1} \cup \cdots \cup B_{\lambda_r}$ to be an orthonormal set of vectors.

Theorem: any hermitian matrix A is diagonalisable. Moreover for A $n\times n$ acting on $V = \mathbb{C}^n$ then

(i) there is an orthonormal basis of eigenvectors $\underline{u}_1, \ldots, \underline{u}_n$
$A\underline{u}_i = \lambda_i \underline{u}_i$ for some λ_i or something
with $\underline{u}_i^\dagger \underline{u}_j = \delta_{ij}$ or something. Yes, true. Equivalently

(ii) there exists ~~of~~ an $n\times n$ unitary matrix P with
$P^{-1}AP = P^\dagger AP = D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ where columns of P are the eigenvectors $\underline{u}_i$.

For the special case when A is real & symmetric, we have a basis of eigenvectors for $\mathbb{R}^n$ such that

$$A\underline{u}_i = \lambda\underline{u}_i$$

$$\& \quad \underline{u}_i \cdot \underline{u}_j = \delta_{ij}$$

Equivalently, there exists an $n\times n$ real orthogonal matrix with $P^{-1}AP = P^TAP = D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$ & columns of P are eigenvectors $\underline{u}_i$.

Proof: * outline - non examinable *

We prove the result for any matrix A acting on a space V of dimension n, a subspace of some $\mathbb{C}^m$ where $m > n$, by induction on n.

Assume true for spaces of dimension $< n$. Given A, we know $\chi_A(t)$ has at least one root $t = \lambda$ & eigenvector $\underline{v} \in V$ with $A\underline{v} = \lambda \underline{v}$.

Define $W = \{ \underline{w} \in V : \underline{v}^\dagger \underline{w} = 0 \}$ a subspace of dimension $< n$.

But $\underline{v}^\dagger (A\underline{w}) = (A\underline{v})^\dagger \underline{w} = (\lambda \underline{v})^\dagger \underline{w} = \lambda \underline{v}^\dagger \underline{w} = 0$

So $A\underline{w} \in W$ and A maps $W \to W$.

By inductive hypothesis W has a basis of orthonormal evecs $\underline{u}_1, \dots \underline{u}_{n-1}$

Then define $\underline{u}_n = \frac{1}{|\underline{v}|} \underline{v}$ and $\underline{u}_1, \dots, \underline{u}_n$ is an orthonormal basis for V.

But for $n = 1$ result is immediate & hence Theorem is true. $\square$

§6.5 Quadratic Forms

On $\mathbb{R}^n$, define a <u>quadratic form</u> to be a function

$$\mathcal{F}(\underline{x}) = \underline{x}^T A \underline{x} = x_i A_{ij} x_j$$

for some real, symmetric $n \times n$ matrix.

[Note if $B_{ij} = -B_{ji}$ then $x_i B_{ij} x_j = 0$ so no loss of generality in taking A to be symmetric]

From diagonalisability of real symmetric matrices, we know that $P^{-1}AP = P^T AP = D$ diagonal matrix of eigenvalues if P has cols = orthormal evecs.

Then $f(\underline{x}) = \underline{x}^T P D P^T \underline{x} = (\underline{x}')^T D \underline{x}'$

where $\underline{x}' = P^T \underline{x}$.

If $\underline{u}_1, \dots, \underline{u}_n$ are evecs, columns of P, then

$x_i' = \underline{u}_i \cdot \underline{x}$

i.e. $\underline{x} = x_1 \underline{e}_1 + \dots + x_n \underline{e}_n$ ($\underline{e}_i$ standard basis)

$= x_1' \underline{u}_1 + \dots + x_n' \underline{u}_n$

so x_i' components w.r.t. new orthonormal basis

$$f(\underline{x}) = \sum_i \lambda_i x_i'^2 \quad \& \quad |\underline{x}|^2 = x_i x_i = x_i' x_i'$$

L21.1

P orthogonal ⇒ lengths, inner products preserved

- Note: can regard x_i' as new coordinates w.r.t. basis vectors $\underline{u}_i$ (orthonormal), equivalent to choosing new axis.

Directions defined by $\underline{u}_i$ are called the <u>principal axes</u> of f (related to axes corresponding to standard basis vectors $\underline{e}_i$ by a rotation or reflection).

<u>Examples in $\mathbb{R}^2$</u>

Consider $f(\underline{x}) = \underline{x}^T A \underline{x}$ with $A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$

- <u>Evals</u> λ given by $(\alpha - \lambda)^2 - \beta^2 = 0$

  solutions $\lambda_1 = \alpha + \beta$, $\lambda_2 = \alpha - \beta$

<u>Evecs</u> $\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $\underline{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ (orthormal)

$$f(\underline{x}) = \alpha x_1^2 + 2\beta x_1 x_2 + \alpha x_2^2$$
$$= (\alpha+\beta) x_1'^2 + (\alpha - \beta) x_2'^2$$

where $x_1' = \underline{u}_1 \cdot \underline{x} = \frac{1}{\sqrt{2}}(x_1 + x_2)$

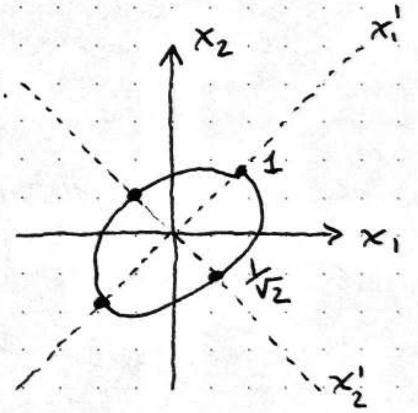
$x_2' = \underline{u}_2 \cdot \underline{x} = \frac{1}{\sqrt{2}}(-x_1 + x_2)$

- Often interested in solutions of $f(\underline{x}) = \text{const}$.

e.g. $\alpha = \frac{3}{2}$, $\beta = -\frac{1}{2}$ ⇒ $\lambda_1 = 1$, $\lambda_2 = 2$

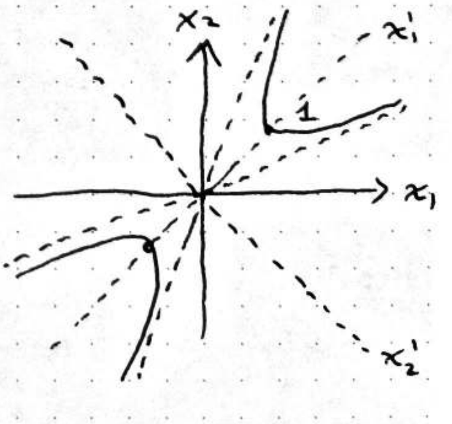
$f(\underline{x}) = x_1'^2 + 2x_2'^2 = 1$ is an ellipse

axes: $1, \frac{1}{\sqrt{2}}$



e.g. $\alpha = -\frac{1}{2}$, $\beta = \frac{3}{2}$ ⇒ $\lambda_1 = 1$, $\lambda_2 = -2$

$f(\underline{x}) = x_1'^2 - 2x_2'^2 = 1$ is a hyperbola



-

Examples in $\mathbb{R}^3$

(i) If A has eigenvalues $\lambda_1, \lambda_2, \lambda_3 > 0$ then

$$f(\underline{x}) = \lambda_1 x_1^2 + \lambda_2 x_2^2 + \lambda_3 x_3^2$$

& $f(\underline{x})$ defines an ellipsoid

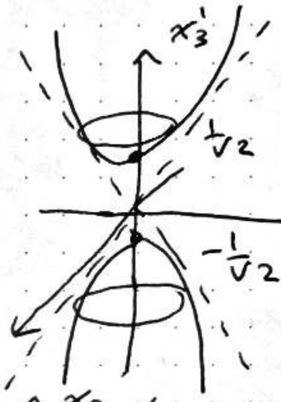
(ii) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ has evecs $\underline{u_1} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $\underline{u_2} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$

with $\lambda_1 = \lambda_2 = -1$

& $\underline{u_3} = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ with $\lambda_3 = 2$

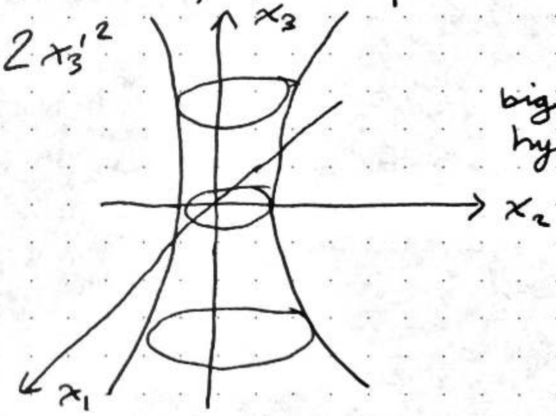
$$f(\underline{x}) = 2x_1x_2 + 2x_2x_3 + 2x_3x_1$$
$$= -x_1'^2 - x_2'^2 + 2x_3'^2$$

$f(\underline{x}) = 1 \iff 2\underline{x_3}'^2 = 1 + x_1'^2 + x_2'^2$



two sheeted hyperboloid

$f(\underline{x}) = -1 \iff x_1'^2 + x_2'^2 = 1 + 2x_3'^2$



big ol hyperboloid

6.6 Cayley-Hamilton Theorem

Theorem: Let A be an $n \times n$ matrix with $\chi_A(t) = \det(A - tI) = \sum_{r=0}^{n} c_r t^r$.

Then $\chi_A(A) = \sum_{r=0}^{n} c_r A^r = 0$

i.e. A always satisfies its own characteristic equation.

Proof: (i) For any 2×2 matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ we can check directly: $\det(A - tI) = \det\begin{pmatrix} a-t & b \\ c & d-t \end{pmatrix} = (a-t)(d-t) - bc$

$$= t^2 - (a+d)t + (ad - bc)$$

Check $\chi_A(A) = 0$ by substitution.

(ii) For an $n \times n$ diagonalisable matrix, write $P^{-1}AP = D$

$$= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Note that $D^r = \begin{pmatrix} \lambda_1^r & & 0 \\ & \ddots & \\ 0 & & \lambda_n^r \end{pmatrix}$ is also diagonal, & similarly

if $p(t)$ is any polynomial, $p(D) = \begin{pmatrix} p(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p(\lambda_n) \end{pmatrix}$.

But $\chi_A(\lambda_i) = 0$ & hence $\chi_A(D) = 0$ ← the 0 matrix

Now note also $A = PDP^{-1} \Rightarrow A^r = (PDP^{-1})^r = PD^rP^{-1}$.

Hence $\chi_A(A) = \chi_A(PDP^{-1}) = P\chi_A(D)P^{-1} = 0$, as required.

Relationship with A^{-1}

Cayley-Hamilton says $c_0 I + c_1 A + \cdots + c_n A^n = 0$

$$\Rightarrow (c_1 I + c_2 A + \cdots + c_n A^{n-1})A = -c_0 I$$

& $c_0 = \det A$. Hence if $\det A \neq 0$ then

A is invertible & $A^{-1} = \frac{-1}{\det A}(c_1 I + c_2 A + \cdots + c_n A^{n-1})$.

7. Changing bases, Canonical forms & Symmetries

7.1 Changing bases in general

Recall, as in §4.4, given a linear map $T: V \to W$ real or complex vector spaces of dimension n, m respectively, & a choice of bases $\{\underline{e}_1, \ldots, \underline{e}_n\}$ for V & $\{\underline{f}_1, \ldots, \underline{f}_m\}$ for W, the $m \times n$ matrix

A_{ai} w.r.t. these basis is defined by $T(\underline{e}_i) = \sum_a A_{ai} \underline{f}_a$.

This is defined so that $\underline{y} = T(\underline{x}) \Leftrightarrow y_a = A_{ai} x_i$,

where $\underline{y} = y_a \underline{f}_a$ & $\underline{x} = x_i \underline{e}_i$, This applies to any chosen bases. So given another choice $\{\underline{e}'_1, \ldots, \underline{e}'_n\}$ & $\{\underline{f}'_1, \ldots, \underline{f}'_m\}$ we have $T(\underline{e}'_i) = \sum_a \underline{f}'_a A'_{ai}$ defines A'.

How are A and A' related? We need to specify how the bases are related; let

$$\underline{e}_i' = \sum_j \underline{e}_j P_{ji} \quad \& \quad \underline{f}'_a = \sum_b \underline{f}_b Q_{ba}.$$

The matrices P $(n \times n)$ & Q $(m \times m)$ must be invertible.

<u>Proposition</u>: With matrices A & A' for a linear map T & bases related as above $A' = Q^{-1}AP$.

TUTURU

Comments (i) Def^n of A wrt bases $\{\underline{e}_i\}$ & $\{\underline{f}_a\}$ can be interpreted: column i of A consists of components of $T(\underline{e}_i)$ wrt $\{\underline{f}_a\}$.

Similarly, def^n of change of basis matrices P & Q can be interpreted as: column i consists of components of new i basis vectors wrt old basis vectors.

(ii) Invertibility of P & Q follow from reversing roles of $\underline{e}_i$, $\underline{e}_i'$ to get matrix P' instead of P & compute $PP' = I$.

(a) Example $n = 2$, $m = 3$

$$T(\underline{e}_1) = \underline{f}_1 + 2\underline{f}_2 - \underline{f}_3 = \sum_a \underline{f}_a A_{a1}$$

$$T(\underline{e}_2) = -\underline{f}_1 + 2\underline{f}_2 + \underline{f}_3 = \sum_a \underline{f}_a A_{a2}$$

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

New basis for V:

$$\underline{e}_1' = \underline{e}_1 - \underline{e}_2 = \sum_i \underline{e}_i P_{i1}$$

$$\underline{e}_2' = \underline{e}_1 + \underline{e}_2 = \sum_i \underline{e}_i P_{i2}$$

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

$$Q = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

$\Uparrow$

& new basis for W:

$$\underline{f}_1' = \underline{f}_1 - \underline{f}_3 = \sum_a \underline{f}_a Q_{a1}$$

$$\underline{f}_2' = \underline{f}_2 = \sum_a \underline{f}_a Q_{a2}$$

$$\underline{f}_3' = \underline{f}_1 + \cancel{\underline{f}_2} + \underline{f}_3 = \sum_a \underline{f}_a Q_{a3}$$

Apply change of basis formula

$$A' = Q^{-1}AP$$

$$= \begin{pmatrix} \tfrac12 & 0 & -\tfrac12 \\ 0 & 1 & 0 \\ \tfrac12 & 0 & \tfrac12 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

$$= \begin{pmatrix} \tfrac12 & 0 & -\tfrac12 \\ 0 & 1 & 0 \\ \tfrac12 & 0 & \tfrac12 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ -2 & 0 \end{pmatrix}$$

$$= \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 0 \end{pmatrix}$$

Direct check:

$$T(\underline{e}_1') = 2\underline{f}_1 - 2\underline{f}_3 = 2\underline{f}_1'$$

$$T(\underline{e}_2') = 4\underline{f}_2 = 4\underline{f}_2'$$

(b) Important special cases

(i) $V = W$ and $\underline{f}_i = \underline{e}_i$ (same basis)
and $\underline{f}_i' = \underline{e}_i'$ (same new basis) then $Q = P$
and $A' = P^{-1}AP$, so matrices representing the same linear map are always similar. Conversely, if A and A' are similar we can regard them as representing the same linear map wrt different bases. (Any invertible matrix gives a change of basis)

Saw earlier that $\operatorname{tr}(A') = \operatorname{tr}(A)$ and $\det(A') = \det(A)$, and more generally $\chi_A(t) = \chi_{A'}(t)$ for similar matrices – from above we can regard these as properties of a linear map.

(ii) $V = W = \mathbb{R}^n$ or $\mathbb{C}^n$ $\quad \underline{e}_i$ standard basis
$\underline{e}_i' = \underline{v}_i$ eigenvectors so $T(\underline{v}_i) = \lambda_i \underline{v}_i$
then $A' = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ & $\underline{v}_i = \sum_j \underline{e}_j P_{ji}$
so columns of P are eigenvectors

Such a basis of evecs does not exist in general, but does if A is hermitian, for example.

(c) Proof of Proposition

$$T(\underline{e}_i') = T\left(\sum_j \underline{e}_j P_{ji}\right) = \sum_j T(\underline{e}_j) P_{ji}$$

$$= \sum_{j,a} \underline{f}_a A_{aj} P_{ji} \qquad \left(\begin{matrix}\text{def}^n \text{ of} \\ A\end{matrix}\right)$$

But also $T(\underline{e}_i') = \sum_a \underline{f}_a' A'_{ai} \qquad \left(\begin{matrix}\text{def}^n \text{ of} \\ A'\end{matrix}\right)$

$$= \sum_{a,b} \underline{f}_b Q_{ba} A'_{ai}$$

$\therefore \sum_j A_{bj} P_{ji} = \sum_a Q_{ba} A'_{ai} \Rightarrow AP = QA'$ & $A' = Q^{-1}AP$
as claimed. $\square$

(d) Changes in Vector Components under changes of basis

Consider a general vector in V

$\underline{x} = x_i \underline{e}_i = x_i' \underline{e}_i'$ where

x_i, x_i' are components wrt bases $\{\underline{e}_i\}$, $\{\underline{e}_i'\}$

related by matrix P as above. $\left\{ \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} = P \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix} \right\}$

Then $x_i \underline{e}_i = x_j' \underline{e}_i P_{ij}$

$$= (P_{ij} x_j') e_i$$

so $x_i = P_{ij} x_j'$.

Convenient to write $X = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ & $X' = \begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$

as column vectors, and then $X = PX'$.

Similarly for general vector in W

$\underline{y} = y_a \underline{f}_a = y_b' \underline{f}_b'$

& have $y_a = Q_{ab} y_b'$

So with $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$, $Y' = \begin{pmatrix} y_1' \\ \vdots \\ y_n' \end{pmatrix}$ then $Y = QY'$.

But def^n of matrices is designed to ensure $\underline{y} = T(\underline{x}) \Leftrightarrow y_a = A_{ai} x_i$

& $y_a' = A_{ai}' x_i$ or $Y = AX$ and $Y' = A'X'$

Hence $Y' = Q^{-1} Y = Q^{-1} AX = (Q^{-1}AP) X'$

and so $A' = Q^{-1}AP$ alternative derivation.

Previously saw for eigenvectors of real symmetric matrix (& similarly for hermitian) that

$\underline{x} = x_i \underline{e}_i = x_i' \underline{u}_i$

where $\underline{e}_i$ standard basis vectors for $V = \mathbb{R}^n$ & $\underline{u}_i$ new set of orthonormal basis vectors.

L22.*

7.2 Jordan Canonical (or Normal) Form

Consider a linear map $T: \mathbb{C}^n \to \mathbb{C}^n$ with matrix A wrt a chosen basis.

<u>Proposition</u>: For $n=2$, any complex matrix A is similar to one of the following

(i) $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ with $\lambda_1 \neq \lambda_2$

(ii) $A = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ for some λ

(iii) $A = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ for some λ

where $\lambda, \lambda_1, \lambda_2 \in \mathbb{C}$

<u>Proof</u>: Consider $\chi_A(t)$.

(i) If there are distinct eigenvalues λ_1, λ_2 then $M_1 = m_1 = M_2 = m_2 = 1$ & evecs $\underline{v_1}, \underline{v_2}$ are a basis. Wrt to this basis A' is diagonal.

(ii) If there is a repeated eigenvalue $M_\lambda = 2$, and $m_\lambda = 2$ then we get a basis of eigenvectors. In fact $A = A' = \lambda I$.

(iii) If ~~there~~ we have $M_\lambda = 2$, and $m_\lambda = 1$ only one evec $\underline{v}$, with $A\underline{v} = \lambda \underline{v}$. Let $\underline{w}$ be any other linearly indep. vector, and consider $A\underline{w} = \alpha \underline{v} + \beta \underline{w}$. With respect to $\{\underline{v}, \underline{w}\}$ we have matrix $\begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix}$. But then $\beta = \lambda$ since only one eval, and $\alpha \neq 0$ since $\underline{w}$ not an evec. Set $\underline{u} = \alpha \underline{v}$ & then wrt basis $\{\underline{u}, \underline{w}\}$ have matrix $\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ since $A(\alpha\underline{v}) = \lambda(\alpha\underline{v})$ & $A\underline{w} = \alpha\underline{v} + \beta\underline{w} = \underline{u} + \lambda\underline{w}$. $\square$

<u>Alternative approach / construction for case (iii)</u>

If A has characteristic polynomial $\chi_A(t) = (t-\lambda)^2$ but $A \neq \lambda I$, then $\exists$ some vector $\underline{w}$ such that
$$\underline{u} = (A - \lambda I)\underline{w} \neq \underline{0},$$
since not every vector is an evec of A.
But $(A-\lambda I)\underline{u} = (A - \lambda I)^2 \underline{w} = \underline{0}$ by Cayley-Hamilton. Hence $A\underline{u} = \lambda \underline{u}$, $A\underline{w} = \underline{u} + \lambda \underline{w}$.
So, changing basis to $\{\underline{u}, \underline{w}\}$ get new matrix
$$A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = P^{-1}AP$$
where columns of P are $\underline{u}, \underline{w}$.

<u>Example</u> $A = \begin{pmatrix} 1 & 4 \\ -1 & 5 \end{pmatrix}$

$\chi_A(t) = (t-3)^2$ & $A - 3I = \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix}$

Choose $\underline{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ not an evec ✓

Then $\underline{u} = (A - 3I)\underline{w} = \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -2 \\ -1 \end{pmatrix}$

This is an evec (could check) so $A\underline{u} = 3\underline{u}$, $A\underline{w} = \underline{u} + 3\underline{w}$. Check $P = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix}$ gives $P^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}$ and $P^{-1}AP = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}\begin{pmatrix} -6 & 1 \\ -3 & -1 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}$.

<u>Theorem</u>: any $n \times n$ complex matrix A is similar to a matrix A' with block form
$$A' = \begin{pmatrix} \boxed{J_{n_1}(\lambda_1)} & & 0 \\ & \ddots & \\ 0 & & \boxed{J_{n_r}(\lambda_r)} \end{pmatrix}$$

where each block on diagonal is square &
$$J_p(\lambda) = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$$
$p \times p$ Jordan block

L23.2

Here $n_1 + n_2 + \cdots + n_r = n$.

The same eigenvalue can appear in more than one block.

A diagonalisable $\Longleftrightarrow$ A' consists of 1×1 blocks

Proof: Discussed in IB Linear Algebra, done in IB Groups, Rings & Modules

Example
$$A = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix}$$ considered for §6.2 (b)(ii)

$$\chi_A(t) = -(2+t)^3$$

Take $\underline{w} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$, not an evec

Let $\underline{u} = (A + 2I)\underline{w} = \begin{pmatrix} -1 \\ -1 \\ -2 \end{pmatrix}$.

In this case $\underline{u}$ is an evec $A\underline{u} = -2\underline{u}$.

Find one additional linearly indep evec $\underline{\tilde{u}} = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$. x

For new basis $\{\underline{u}, \underline{w}, \underline{\tilde{u}}\}$ construct $P = \begin{pmatrix} -1 & 1 & 1 \\ -1 & 0 & 0 \\ -2 & 0 & 1 \end{pmatrix}$ to give

$$P^{-1} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 1 & -1 \\ 0 & -2 & 1 \end{pmatrix} \text{ and } P^{-1}AP = \left(\begin{array}{cc|c} -2 & 1 & 0 \\ 0 & -2 & 0 \\ \hline 0 & 0 & -2 \end{array}\right).$$

L23.3

7.3 Quadrics & Conics

(a) Quadrics in general

A quadric in $\mathbb{R}^n$ is (a hypersurface) defined by an eq^n of the form

$$Q(\underline{x}) = \underline{x}^T A \underline{x} + \underline{b}^T \underline{x} + c = 0$$

where $A^T = A$ real $n \times n$ matrix

$\underline{b} \in \mathbb{R}^n$ & c a constant.

i.e. $Q = A_{ij} x_i x_j + b_i x_i + c = 0$.

Consider classifying solutions of this equation, up to <u>geometrical equivalence</u>: we will not distinguish solutions which differ by isometries in $\mathbb{R}^n$, i.e. which differ by

(i) translations – change in origin

(ii) orthogonal transformations w.r.t. some chosen origin

If A is invertible $[A^T = A \Rightarrow (A^{-1})^T = A^{-1}]$, let

$\underline{y} = \underline{x} + \frac{1}{2} A^{-1} \underline{b}$ then

$$\begin{aligned}
\underline{y}^T A \underline{y} &= (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b})^T A (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b}) \\
&= (\underline{x}^T + \tfrac{1}{2} \underline{b}^T A^{-1}) A (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b}) \\
&= \underline{x}^T A \underline{x} + \tfrac{1}{2} \underline{b}^T \underline{x} + \tfrac{1}{2} \underline{x}^T \underline{b} + \tfrac{1}{4} \underline{b}^T A^{-1} \underline{b}
\end{aligned}$$

Hence $Q(\underline{x}) = 0 \iff \mathcal{F}(\underline{y}) = k$ where

$$\mathcal{F}(\underline{y}) = \underline{y}^T A \underline{y} \quad \& \quad k = \tfrac{1}{4} \underline{b}^T A^{-1} \underline{b} - c$$

quadratic form. Now diagonalise $\mathcal{F}$, i.e. find orthonormal eigenvectors, defining <u>principal axes</u> (new coordinate axes)

Geometrical nature of quadric determined by evals of A and value of k.

Refer to examples in $\mathbb{R}^3$ in §6.5

(i) $\lambda_i > 0$, $k > 0$ get ellipsoid

(ii) λ_i different signs (but non-zero), $k \neq 0$

get hyperboloids

If A not invertible - at least one zero eigenvalue, then may have one linear term.

(b) <u>Conics</u>

Quadrics in $\mathbb{R}^2$ are curves, <u>conics</u>.

Consider possible cases for $A \neq 0$.

Suppose $\det A \neq 0$. By construction above (completing square) we can assume $\underline{b} = \underline{0}$ & have $f(\underline{x}) = \underline{x}^T A \underline{x} = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 = k$

Suppose evals have same sign, then $\lambda_1, \lambda_2, k > 0$ wlog (else no sol^n) & this represents an ellipse. (if $k=0$ point)

Suppose evals have different signs, then $\lambda_1 > 0$, $\lambda_2 < 0$ & $k \geq 0$ wlog. If $k=0$ get pair of straight lines Lol.
If $k > 0$ get hyperbola.

Suppose $\det A = 0$ with $\lambda_1 > 0$, $\lambda_2 = 0$ wlog. Refer to original equation
$$\lambda_1 x_1'^2 + b_1' x_1' + b_2' x_2' + c = 0$$
after diagonalisation of A, so x_i' new orthogonal basis.

Let $x_1' + \frac{1}{2\lambda_1} b_1' = x_1''$ & we obtain
$$\lambda_1 (x_1'')^2 + b_2' x_2' + c - \frac{b_1'^2}{4\lambda_1^2} = 0$$
or $\lambda_1 (x_1'')^2 + b_2' x_2''$ where $x_2'' = x_2' + \frac{1}{b_2'}\left(c - \frac{b_1'^2}{4\lambda_1^2}\right)$ if $b_2' \neq 0$
if $b_2' = 0$ lines.

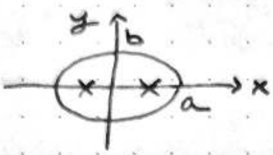
L 23.5 now L 24.1

This is a parabola.

Standard forms for Conics in Cartesians

For coordinates x, y in $\mathbb{R}^2$

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$



ellipse $(a > b)$

eccentricity $e < 1$

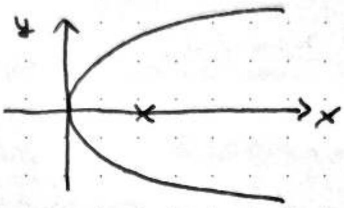
$b^2 = a^2(1-e^2)$

foci

$x = \pm ae$
$y = 0$

$y^2 = 4ax$

parabola



eccentricity $e = 1$

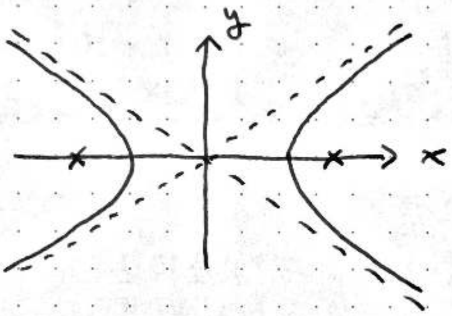
focus $x = a$
$y = 0$

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

hyperbola

eccentricity $e > 1$

$b^2 = a^2(e^2-1)$



foci

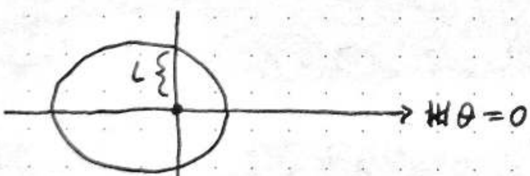
$x = \pm ae$
$y = 0$

Standard forms for Conics in Polars

Take origin to be focus.

$$r = \frac{l}{1 + e\cos\theta}$$

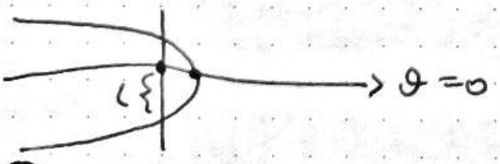
$\begin{cases} \frac{l}{1+e} = a - ae \\ \frac{l}{1+e} = ae - a \end{cases}$



$l = a(1-e^2)$

$e < 1$

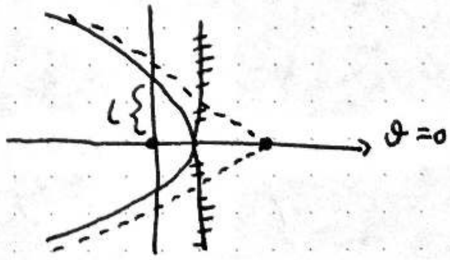
ellipse



$l = 2a$

$e = 1$

parabola



$l = a(e^2-1)$

$e > 1$

hyperbola

7.4 Symmetries and transformation groups

(2) Orthogonal transformations & Rotations in $\mathbb{R}^n$

From § 4.5 we know R orthogonal $\Leftrightarrow R^T R = R R^T = I$

$\Leftrightarrow (R\underline{x}) \cdot (R\underline{y}) = \underline{x} \cdot \underline{y} \quad \forall\, \underline{x}, \underline{y} \in \mathbb{R}^n$

$\Leftrightarrow$ cols (or rows) are orthonormal

The set of all such matrices forms a group, $O(n)$, the <u>orthogonal group</u>

[Easy to check : $R_1, R_2 \in O(n) \Rightarrow (R_1 R_2)^T = R_2^T R_1^T = R_2^{-1} R_1^{-1} = (R_1 R_2)^{-1}$]

associativity inherited;

$I \in O(n)$; $R \in O(n) \Rightarrow R^{-1} \in O(n)$

For $R \in O(n)$, $\det(R^T R) = \det(R^T)\det(R) = \det(R)^2 = \det(I) = 1$

$\Rightarrow \det(R) = \pm 1$

The matrices R in $O(n)$ with $\det R = 1$ form a subgroup $SO(n)$, the <u>special orthogonal group</u>.

~~Reflection~~ This subgroup contains precisely the <u>rotations</u> in $\mathbb{R}^n$ (by defn). These preserve length and orientation.

Reflections belong to $O(n)$ but not to $SO(n)$. Choosing some specific H in $O(n) \setminus SO(n)$ any element of $O(n)$ is of the form R or RH, with $R \in SO(n)$.

In §4.5 we found this explicitly for $n=2$: the most general orthogonal matrix in this case is

$$R = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix} \in SO(2) \quad \text{or} \quad RH \text{ with } H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

L24.3

(b) 2D Minkowski Space & Lorentz Transformations

- Consider a new "inner-product" on $\mathbb{R}^2$ defined by

$$(\underline{x}, \underline{y}) = \underline{x}^T J \underline{y} \qquad \text{with } J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$
$$= x_1 y_1 - x_2 y_2$$

This is not positive definite:

$$(\underline{x}, \underline{x}) = x_1^2 - x_2^2$$

& basis vectors are "orthonormal" in generalised sense

$$(\underline{e}_1, \underline{e}_1) = 1, \quad (\underline{e}_1, \underline{e}_2) = 0, \quad (\underline{e}_2, \underline{e}_2) = -1$$

This is called the <u>Minkowski metric</u>.

$$\left\{ \cancel{\begin{pmatrix} \sinh & -\cosh \\ -\cosh & -\sinh \end{pmatrix}} \quad \begin{pmatrix} -\cosh & -\sinh \\ -\sinh & -\cosh \end{pmatrix} \right.$$

- Now consider $M : \mathbb{R}^2 \to \mathbb{R}^2$ which preserve it

$$(M\underline{x}, M\underline{y}) = (\underline{x}, \underline{y})$$

or equivalently

$$M^T J M = J.$$

$$\left[(M\underline{x})^T J (M\underline{y}) = \underline{x}^T M^T J M \underline{y} = \underline{x}^T J \underline{y} \quad \forall \underline{x}, \underline{y} \right]$$

As with rotations, we can check such matrices form a group & $(\det M)^2 = 1$.

- The subgroup with $\det M = 1$ & $M_{11} > 0$ is the <u>Lorentz group</u> in 2D.

General form of M can be determined by requiring columns of M to be orthonormal in same sense as $\underline{e}_1, \underline{e}_2$ above.

$$M(\theta) = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$

for some real θ given $M_{11} > 0$

overall sign fixed by $\det M = +1$

Note that $M(\theta_1) M(\theta_2) = M(\theta_1 + \theta_2)$

<u>Physical interpretation</u>

- Write $M(\theta) = \gamma(v) \begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix}$

where $\gamma(v) = \dfrac{1}{\sqrt{1 - v^2}}$ so

$$v = \tanh\theta \quad \& \quad |v| < 1$$

L24.4

In units with $c=1$, taking $x_1 =$ time and $x_2 =$ space.
M gives the Lorentz transformation relating observers in special relativity with relative velocity v.

$$x_1' = \gamma(x_1 + v x_2)$$
$$x_2' = \gamma(x_2 + v x_1)$$

γ factor responsible for time dilation & length contraction effects.

Group property $\tanh(\theta_3) = \tanh(\theta_1 + \theta_2)$ $\qquad \theta_3 = \theta_1 + \theta_2$

$$= \frac{\tanh\theta_1 + \tanh\theta_2}{1 + \tanh\theta_1 \tanh\theta_2}$$

$$\Rightarrow \quad v_3 = \frac{v_1 + v_2}{1 + v_1 v_2}$$