Vectors and Matrices Lecture 1 — 9/10/20

Course covers topics in algebra and geometry with abstract / concrete approaches / conceptual / computational

1) Complex Numbers
2) Vectors in 3D — index notation / summation convention
3) Vectors in general; $\mathbb{R}^n$ and $\mathbb{C}^n$
4) Matrices and Linear Maps
5) Determinants and Inverses
6) Eigenvalues and Eigenvectors
7) Changing Bases, Canonical Forms and Symmetries

1) Complex Numbers

1.1 Definitions

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

Any complex number $z \in \mathbb{C}$ has the form $\underline{z = x + iy}$ with $x, y \in \mathbb{R}$. $\quad x = \mathrm{Re}(z) \quad y = \mathrm{Im}(z)$

So $\mathbb{R} \subset \mathbb{C}$: $\mathbb{R}$ is the subset consisting of $x = x + i0$.

For $z_1 = x_1 + iy_1$ and $z_2 = x_2 + iy_2$:

(i) Addition: $z_1 + z_2 = (x_1 + x_2) + i(y_1 + y_2)$

(ii) Multiplication:
$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)$$

For $z = x + iy$, we define

(iii) Complex conjugate $\bar{z} = z^* = x - iy \Rightarrow \mathrm{Re}(z) = \frac{1}{2}(z + \bar{z})$, $\mathrm{Im}(z) = \frac{1}{2}(z - \bar{z})$

(iv) Modulus $r = |z|$, real and ≥ 0

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

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

(v) <u>Argument</u> $\theta = \arg(z)$, real, $z \neq 0$ by

$$z = r(\cos\theta + i\sin\theta) \quad \text{(polar form)}$$

$$\Leftrightarrow \quad \cos\theta = \frac{x}{(x^2+y^2)^{1/2}}, \quad \sin\theta = \frac{y}{(x^2+y^2)^{1/2}}$$

$$\Rightarrow \tan\theta = \frac{y}{x}$$

$\arg(z)$ is only determined "mod 2π" – can change

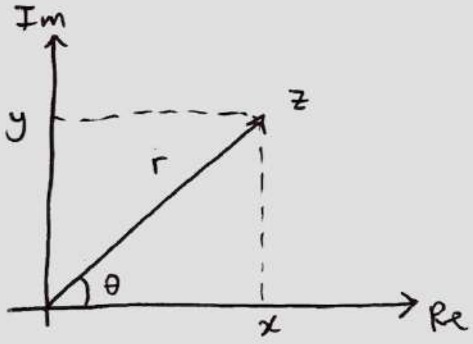
$\theta \to \theta + 2n\pi$ $(n \in \mathbb{Z})$ and $z \to z$: doesn't change.

To make θ unique, we can restrict the range

e.g. $-\pi < \theta \leq \pi$ is the <u>principal</u> value.

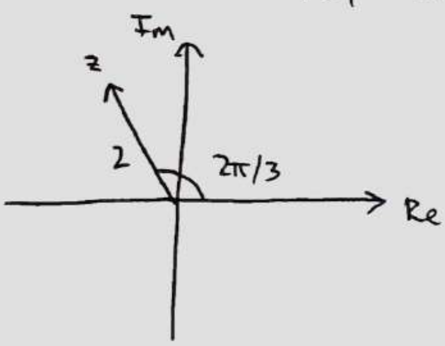
(vi) <u>Argand diagram or Complex Plane</u>

Plot $\mathrm{Re}(z)$ and $\mathrm{Im}(z)$ on orthogonal axes, then $r = |z|$ and $\arg(z) = \theta$ are the length and angle shown.



<u>Example</u> For $z = -1 + i\sqrt{3} = 2\left(-\frac{1}{2} + \frac{i\sqrt{3}}{2}\right)$

we have $|z| = 2$, $\arg(z) = \frac{2\pi}{3} + 2n\pi$



$$\tan\theta = -\sqrt{3}$$

$$\Rightarrow \theta = \frac{2\pi}{3} + 2n\pi$$

or $\theta = -\frac{\pi}{3} + 2n\pi = \arg(-z)$

Basic Properties / Consequences

(i) On $\mathbb{C}$ the operations $+$ and $\times$ are commutative and associative. Furthermore, $(\mathbb{C}, +)$ is an abelian group with identity $0 = 0 + i0$ with $z = x + iy$ having inverse $-x + i(-y)$

$\mathbb{C} \setminus \{0\}$ under multiplication is also an abelian group: $e = 1 = 1 + i0$

$z = x + iy \neq 0$ has inverse $z^{-1} = \frac{1}{z} = \frac{x}{x^2+y^2} + \frac{iy}{x^2+y^2} = \frac{\bar{z}}{|z|^2}$

Distributive laws hold:

$z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$

$\mathbb{C}$ is a <u>field</u>

(ii) Fundamental Theorem of Algebra

<u>Aside:</u> motivate definitions of number systems leading to $\mathbb{C}$ by considering solns of eqns:

$\mathbb{N}$ $\quad x + 3 = 0$

$\downarrow$ to $\mathbb{Z}$

$\mathbb{Z}$ $\quad 5x + 1 = 0$

$\downarrow$ to $\mathbb{Q}$

$\mathbb{Q}$ $\quad x^2 - 2 = (x - \sqrt{2})(x + \sqrt{2}) = 0$

$\downarrow$ to $\mathbb{R}$

$\mathbb{R}$ $\quad x^2 + 4 = (x - 2i)(x + 2i) = 0$

$\downarrow$ to $\mathbb{C}$

$\mathbb{C}$

<u>FTA:</u> A polynomial of degree n with coefficients in $\mathbb{C}$ can be written as a product of n linear factors

$$p(z) = c_n z^n + \dots + c_1 z + c_0 \quad \text{with } c_i \in \mathbb{C},\ c_n \neq 0$$
$$= c_n (z - \alpha_1)(z - \alpha_2) \dots (z - \alpha_n)$$

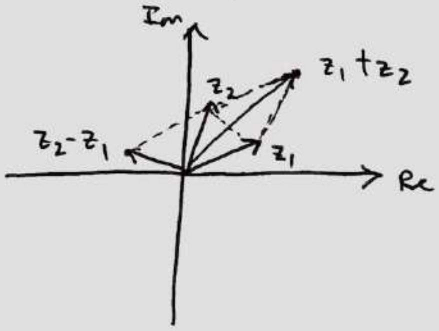
for some $\alpha_i \in \mathbb{C}$.

So the "chain" in the aside stops at $\mathbb{C}$.

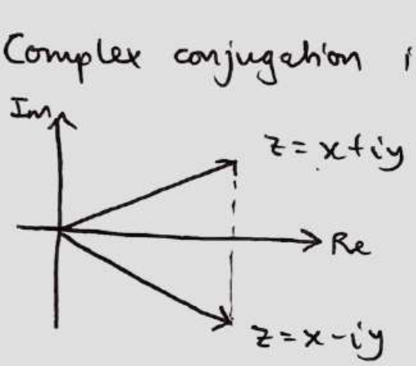
Hence $p(z) = 0$ for $n \geq 1$ has at least one solution and exactly n solutions α_i, counted with multiplicity.

(Not proved in course - see Part IB Complex Methods / Analysis)

(iii) Addition and subtraction correspond to parallelogram constructions shown



Complex conjugation is reflection in the real axis



Aside: quaternions (look up)

(iv) <u>Proposition</u> Modulus / length obey

$$|z_1 z_2| = |z_1||z_2| \quad \text{(composition property)}$$

$$|z_1 + z_2| \leq |z_1| + |z_2| \quad \text{(triangle inequality)}$$

<u>Proof</u> of composition property: compute square of each side and compare.

<u>Proof</u> of triangle inequality: compare squares of each side

$$\text{LHS}^2 = (z_1 + z_2)(\bar{z}_1 + \bar{z}_2) = |z_1|^2 + z_1\bar{z}_2 + \bar{z}_1 z_2 + |z_2|^2$$

$$\text{RHS}^2 = |z_1|^2 + |z_2|^2 + 2|z_1||z_2|$$

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

$\iff \frac{1}{2}\left(z_1\bar{z}_2 + \overline{(z_1\bar{z}_2)}\right) \leq |z_1||\bar{z}_2|$

$\iff \mathrm{Re}(z_1\bar{z}_2) \leq |z_1\bar{z}_2|$ which is true. $\square$

Alternative form of Δ inequality:

$|z_2 - z_1| \geq |z_2| - |z_1|$ by replacing z_2 with $z_2 - z_1$ in original form and rearranging

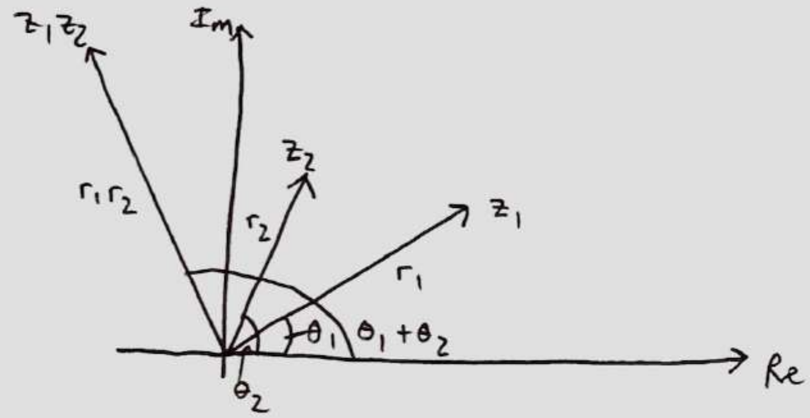
or $|z_2 - z_1| \geq |z_1| - |z_2|$

So $\underline{|z_2 - z_1| \geq \big||z_2| - |z_1|\big|}$

(v) <u>Proposition</u> $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ (P1.1)

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

$\Rightarrow \quad z_1 z_2 = r_1 r_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2))$



<u>Proof</u> just check, use trig addition formulas

<u>De Moivre's Theorem</u> $(\cos\theta + i\sin\theta)^n = \cos n\theta + i\sin n\theta$ for $n \in \mathbb{Z}$.

~~(For $z \neq 0$,~~ (For $z \neq 0$, $z^0 = 1$, $z^{-n} = (z^{-1})^n$)

(prove by induction with P1.1) <u>for $n \geq 0$</u>.

Base case $n = 0$ ✓ $= 1$

Inductive: if $(\cos\theta + i\sin\theta)^k = \cos k\theta + i\sin k\theta$

then $(\cos\theta + i\sin\theta)^{k+1} = (\cos k\theta + i\sin k\theta)(\cos\theta + i\sin\theta)$

$= \cos(k+1)\theta + i\sin(k+1)\theta$

by trig addition formulae.

For $n = -m < 0$ we use $(z^{-m}) = (z^m)^{-1}$ with $z = \cos\theta + i\sin\theta$ then use the above result. □

$(\cos m\theta + i\sin m\theta)^{-1} = \cos m\theta - i\sin m\theta$.

12/10/20

Vectors and Matrices - Lecture 2

1.3 - Exponential and Trigonometric Functions

Define exp, cos, sin on $\mathbb{C}$ by

$$\exp(z) = e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

$$\cos(z) = \frac{1}{2}\left(e^{iz} + e^{-iz}\right) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \dots$$

$$\sin(z) = \frac{1}{2i}\left(e^{iz} - e^{-iz}\right) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \dots$$

The series converge for all $z \in \mathbb{C}$ - they can be multiplied, rearranged etc. (Proved in Analysis I)

Proposition $e^z e^w = e^{z+w} \quad \forall z, w \in \mathbb{C}$

Corollary $e^z e^{-z} = e^0 = 1$ and $(e^z)^n = e^{nz} \quad n \in \mathbb{Z}$

Proof Multiply series on LHS and find terms of degree n

$$\sum_{r=0}^{n} \frac{1}{r!} z^r \frac{1}{(n-r)!} w^{n-r} = \frac{1}{n!}(z+w)^n \quad \text{by Binomial Thm}$$

Definitions reduce to familiar ones for $z = x \in \mathbb{R}$. From series (differentiation allowed)

$$\frac{d}{dx}(\exp x) = \exp x$$

$$\frac{d}{dx}(\cos x) = -\sin x$$

$$\frac{d}{dx}(\sin x) = \cos x$$

differentiating the series term by term

$\exp(0) = 1$, $\cos(0) = 1$, $\sin(0) = 0$ give "boundary conditions"

These properties characterise these functions of the real variable x.

Lemma For $z = x+iy$

(i) $e^z = e^x(\cos y + i\sin y)$

(ii) exp on $\mathbb{C}$ takes all values except 0

(iii) $e^z = 1 \iff z = 2n\pi i, \quad n \in \mathbb{Z}$.

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

(ii) $|e^z| = e^x$ (as $|e^{iy}| = 1$) and e^x takes all positive real values.

$\arg(e^z) = y$ and this takes all real values.

(iii) $e^z = 1 \iff e^x = 1$ and $\cos y = 1$, $\sin y = 0$

$\iff x = 0, \ y = 2n\pi$ as required.

Now return to polar form and note it can be written

$$z = r(\cos\theta + i\sin\theta) = re^{i\theta} \qquad (r = |z|, \ \theta = \arg z)$$

Then De Moivre's Theorem follows immediately from the $e^z e^w = e^{z+w}$ result: $(e^{i\theta})^n = e^{in\theta}$.

Roots of Unity

z is an N^{th} root of unity if $z^N = 1$. To find all solutions

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

$$\iff r^N e^{iN\theta} = 1 \iff r^N = 1 \text{ and } N\theta = 2n\pi \quad (n \in \mathbb{Z})$$

This gives N distinct solutions $z = e^{2\pi i n/N}$

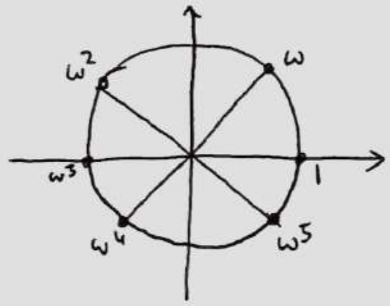
with $n = 0, 1, \dots, N-1$.

$$\downarrow$$

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

$$(\omega = e^{2\pi i/N})$$

These solutions lie on the unit circle at the vertices of a regular N-gon e.g. $N = 6$



1.4 Logarithms and Complex Powers

Define $w = \log z \qquad z \in \mathbb{C},\ z \neq 0$

by $e^w = e^{\log z} = z$

i.e. log is inverse of exp, but exp is many-to-one so log is multi-valued.

$z = re^{i\theta} = e^{\log r} e^{i\theta} = e^{\log r + i\theta}$ ← r is real, >0

so $\log z = \log r + i\theta = \boxed{\log|z| + i \arg z}$.

Multivalued behaviour of arg and log are related:

$\operatorname{Im}(\log z) = \arg z$ so $\theta \to \theta + 2n\pi \Rightarrow \log z \to \log z + 2n\pi i$ $(n \in \mathbb{Z})$.

To make them single valued we can restrict e.g.

$0 \leq \theta < 2\pi$ or $-\pi < \theta \leq \pi$.

e.g. $z = -1 = e^{i\pi}$ so $\log 1 + \underbrace{i\pi}_{\text{or } i(2n+1)\pi} = \cancel{\log z}\ \log(-1)$

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

Complex Powers $\boxed{z^\alpha = e^{\alpha \log z}}$ $\quad z \in \mathbb{C},\ z \neq 0,\ \alpha \in \mathbb{C}$

(contains log - also multivalued)

$\arg z \to \arg z + 2n\pi \Rightarrow z^\alpha \to z^\alpha e^{2\pi i n\alpha}$ (exercise: check)

Can restrict $-\pi < \theta \leq \pi$ for principal value for z^α.

If $\alpha = p \in \mathbb{Z}$ then $z^\alpha = z^p$ is <u>unique</u> because

$e^{2\pi i n p} = 1$ so $z^p = z^p e^{2n i n p} = z^p$.

If $\alpha = \frac{p}{q} \in \mathbb{Q}$ then $z^\alpha = z^{p/q}$ takes <u>finitely many</u> values (as seen with roots of unity).

But generally there are infinitely many values.

<u>Examples</u> i^i

$$i = e^{i\pi/2}, \quad \arg i = \frac{\pi}{2} + 2n\pi$$

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

so $i^i = e^{i \log i} = \underline{e^{-(\frac{\pi}{2} + 2n\pi)}}$ $\quad n \in \mathbb{Z}$.

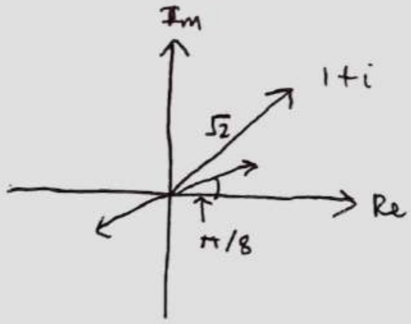
$(1+i)^{1/2}$

$$1+i = \sqrt{2} e^{i\pi/4} = e^{\frac{1}{2}\log 2 + i\pi/4}$$

so $\log(1+i) = \frac{1}{2}\log 2 + i\pi/4 + i2n\pi \quad n \in \mathbb{Z}$

so $(1+i)^{1/2} = e^{\frac{1}{2}\log(1+i)} = e^{\frac{1}{4}\log 2 + i\left(\frac{\pi}{8} + n\pi\right)}$

$= \underline{\pm 2^{1/4} e^{i\pi/8}}$ for n even or odd.



- We can immediately extend standard quadratic formula to quadratics with <u>complex</u> coefficients.
- For $x \in \mathbb{R}$ and $x > 0$, we have <u>unique</u> real value of $\log x$ and hence a unique value of $x^\alpha = e^{\alpha \log x}$ $(\alpha \in \mathbb{R})$

  Then $x^\alpha x^\beta = x^{\alpha + \beta}$ and $(x^\alpha)^\beta = x^{\alpha\beta}$.

  But for complex variables we must beware:

  $z^\alpha z^\beta$ and $z^{\alpha+\beta}$ each have sets of values which may not match.

1.5 Transformations, Lines and Circles

Consider the following transformations on $\mathbb{C}$.

$z \mapsto z + a$ (translation by $a \in \mathbb{C}$)

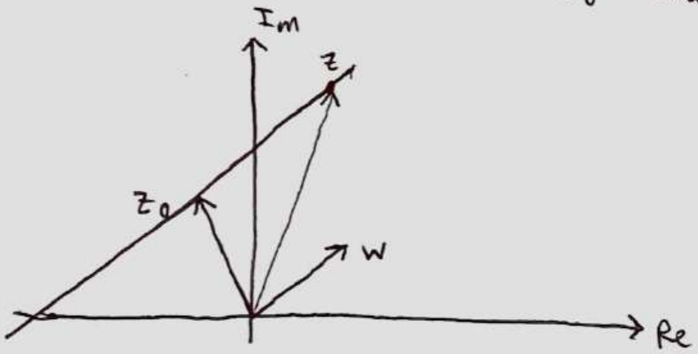
$z \mapsto \lambda z$ (scaling by $\lambda \in \mathbb{R}$)

$z \mapsto e^{i\alpha} z$ (rotation by $\alpha \in \mathbb{R}$)

$z \mapsto \bar{z}$ (reflection in real axis)

$z \mapsto \frac{1}{z}$ (inversion

Use these ideas to find the general point on a line in $\mathbb{C}$ through some fixed z_0 and parallel to $w \in \mathbb{C}$, $w \neq 0$.



General point is

$$\underline{z = z_0 + \lambda w} \qquad (\lambda \in \mathbb{R})$$

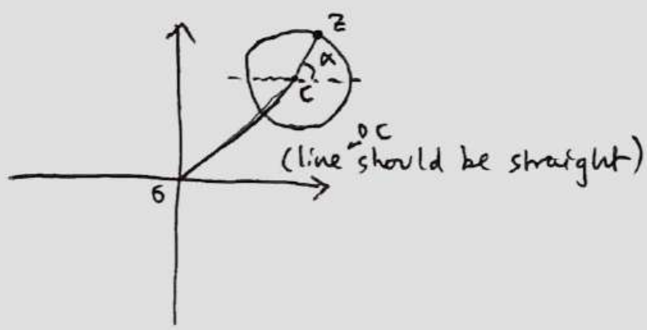
To eliminate λ, take conjugate

$$\bar{z} = \bar{z}_0 + \lambda \bar{w}$$ and combine.

Then eliminate λ to give

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

Then consider circles: centre ~~at~~ $c \in \mathbb{C}$, radius ρ (real, >0)



$z = c + \rho e^{i\alpha}$, with $\alpha \in \mathbb{R}$.

Equivalently

$|z - c| = \rho$ as you'd expect:

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

by squaring.

Vectors and Matrices – Lecture 3

14/10/20

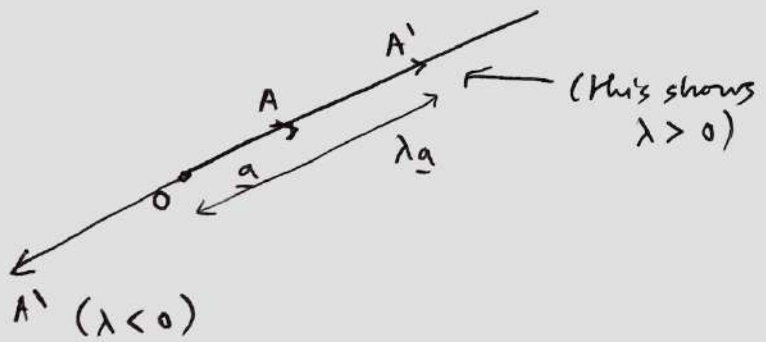
2. Vectors in 3D

A vector is a quantity with magnitude and direction.
Take a geometrical approach and consider position vectors in 3D using standard (Euclidean) notions of points/lines/planes/lengths/angles.

Choose point O as <u>origin</u>, then points A, B, ... are represented as displacements from O – position vectors $\underline{a} = \vec{OA}$ etc.

2.1 Vector Addition & Scalar Multiplication

(i) <u>Scalar Multiplication</u> Given $\underline{a}$ (p.v. for point A) and scalar $\lambda \in \mathbb{R}$, we define λa to be p.v. for A' on line OA with length $|\lambda \underline{a}| = |\vec{OA'}| = |\lambda||\underline{a}|$, and direction as shown depending on sign of λ.



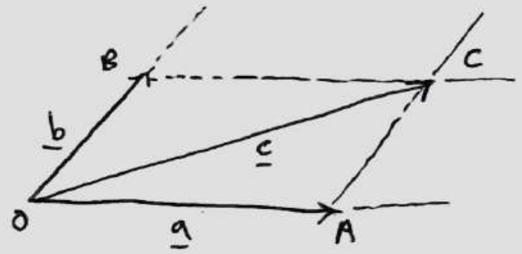
<u>Definition</u> $\text{span}\{a\} = \{\lambda \underline{a} : \lambda \in \mathbb{R}\}$; if $a \neq 0$ then this is line through O, A.

Define $\underline{a}, \underline{b}$ to be parallel ($\underline{a} \parallel \underline{b}$) iff $\underline{a} = \lambda \underline{b}$ or $\underline{b} = \lambda \underline{a}$ for some $\lambda \in \mathbb{R}$ (λ can be 0 so $\underline{a} = \lambda \underline{b} \not\Leftrightarrow \underline{b} = \lambda \underline{a}$)

This is a convenient usage for this course.

(ii) <u>Vector Addition</u> Given $\underline{a}, \underline{b}$ position vectors of A, B, if $\underline{a} \not\parallel \underline{b}$ then construct a parallelogram OACB :-

PTO



Then we define
$\underline{a} + \underline{b} = \underline{c}$.

If $\underline{a} \parallel \underline{b}$ then $\underline{a} = \alpha \underline{u}$, $b = \beta \underline{u}$ where $\alpha, \beta \in \mathbb{R}$, u is a unit vector of length 1. Then $\underline{a} + \underline{b} = \underline{(\alpha + \beta)\underline{u}}$.

Given $\underline{a}, \underline{b}, \dots, \underline{c}$ we can form a <u>linear combination</u>
$\alpha \underline{a} + \beta \underline{b} + \dots + \gamma \underline{c}$ for any $\alpha, \beta, \dots, \gamma \in \mathbb{R}$.

<u>Definition</u> $\text{span}\{\underline{a}, \underline{b}\} = \{\alpha \underline{a} + \beta \underline{b} : \alpha, \beta \in \mathbb{R}\}$
If $\underline{a} \nparallel \underline{b}$ ($\Rightarrow$ neither 0) then this is the <u>plane</u> through $0, A, B$.

(iii) <u>Properties</u> For any $\underline{a}, \underline{b}, \underline{c}$ we have

$\underline{a} + \underline{0} = \underline{0} + \underline{a} = \underline{a}$ ($\underline{0}$ is identity for +)

$\exists -\underline{a} : \underline{a} + (-\underline{a}) = (-\underline{a}) + \underline{a} = \underline{0}$ (inverse)

$\underline{a} + \underline{b} = \underline{b} + \underline{a}$ (commutative)

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

Vectors under + form an <u>abelian group</u>.

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

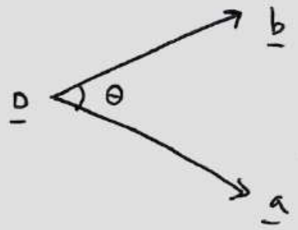
$(\lambda + \mu)\underline{a} = \lambda \underline{a} + \mu(\underline{a}$

$\lambda(\mu \underline{a}) = (\lambda\mu)\underline{a}$

Note we can check associativity by constructing a parallelepiped. (see Aside screenshot)

2.2 Scalar or Dot Product

(i) Definition Given vectors $\underline{a}$, $\underline{b}$, let θ be angle between them as shown



Then define

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

(scalar / dot product)

(θ defined unless $\underline{a}$ or $\underline{b}$ is 0 : in those cases $\underline{a} \cdot \underline{b} = 0$)

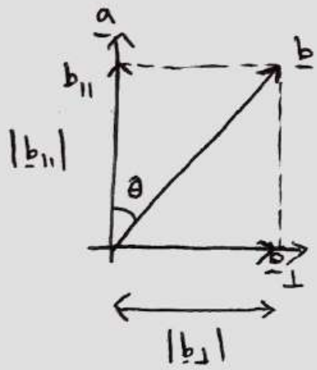
$\underline{a}$ and $\underline{b}$ are <u>orthogonal</u> or <u>perpendicular</u> iff $\underline{a} \cdot \underline{b} = 0$, iff $\theta = \frac{\pi}{2} \bmod \pi$ when π defined.

Write $\underline{a} \perp \underline{b}$, allow $\underline{a}$ or $\underline{b}$ to be $\underline{0}$ so $\underline{0} \perp \underline{a}$ $\forall \underline{a}$.

(ii) Interpretation For $\underline{a} \neq \underline{0}$, $|\underline{b}| \cos\theta$ is component of $\underline{b}$ along $\underline{a}$

$$= \frac{(\underline{a} \cdot \underline{b})}{|\underline{a}|} = \hat{\underline{u}} \cdot \underline{b} \quad \text{where } \hat{\underline{u}} = \frac{\underline{a}}{|\underline{a}|}$$

$= |\underline{b}_{\parallel}|$ where we resolve $\underline{b} = \underline{b}_{\parallel} + \underline{b}_{\perp}$.



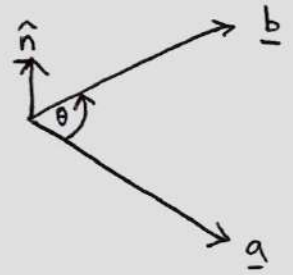
Note $\underline{a} \cdot \underline{b} = \underline{a} \cdot \underline{b}_{\parallel}$

(iii) Properties

$$\underline{a} \cdot \underline{b} = \underline{b} \cdot \underline{a}$$
$$\underline{a} \cdot \underline{a} = |\underline{a}|^2 \geq 0$$
$$(\lambda \underline{a}) \cdot \underline{b} = \lambda(\underline{a} \cdot \underline{b}) = \underline{a} \cdot (\lambda \underline{b})$$
$$\underline{a} \cdot (\underline{b} + \underline{c}) = \underline{a} \cdot \underline{b} + \underline{a} \cdot \underline{c}$$

2.3 Vector or Cross Product

(i) **Definition** Given $\underline{a}$, $\underline{b}$ let θ be angle between them measured in sense shown, wrt a unit vector $\underline{n}$ normal to plane they span.

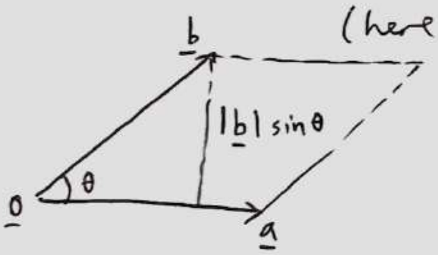


Define $\underline{a} \wedge \underline{b} = \underline{a} \times \underline{b} = |\underline{a}||\underline{b}| \sin\theta \, \hat{\underline{n}}$.

If $\underline{a} \not\parallel \underline{b}$ then n is defined ~~to~~ up to a sign, but changing sign of $\hat{\underline{n}}$ changes θ to $2\pi - \theta$ and $\underline{a} \times \underline{b}$ is unchanged.

If $\underline{a} \parallel \underline{b}$ then $\hat{\underline{n}}$ is not defined
If $|\underline{a}| = 0$ or $|\underline{b}| = 0$ then θ is not defined $\Bigg]$ $\underline{a} \times \underline{b} = \underline{0}$

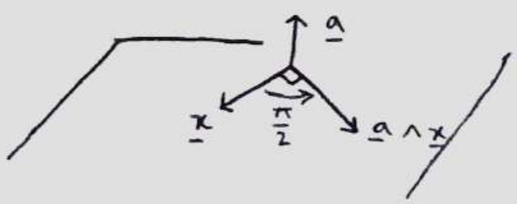
(ii) **Interpretations** $\underline{a} \times \underline{b}$ is vector area of parallelogram (here $\sin\theta > 0$)



so area $= |\underline{a}||\underline{b}| \sin\theta$ (scalar area).

Direction $\underline{n}$ gives orientation in space (making sense of vector area).

Fix $\underline{a}$ and consider $\underline{x} \perp \underline{a}$, then $\underline{x} \mapsto \underline{a} \wedge \underline{x}$ scales by $|\underline{a}|$, rotates by $\frac{\pi}{2}$ in plane $\perp$ to $\underline{a}$



Note that if we resolve as in 2.2 (ii) and write $\underline{b} = \underline{b}_{\parallel} + \underline{b}_{\perp}$

then $\underline{a} \times \underline{b} = \underline{a} \times \underline{b}_{\perp}$

Properties (iii)

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

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

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

$$\underline{a} \times \underline{b} = 0 \iff \underline{a} \parallel \underline{b} \quad (\text{allows} = \underline{0})$$

$$\underline{a} \times \underline{b} \perp \underline{a} \text{ and } \underline{b} \iff \underline{a} \cdot (\underline{a} \times \underline{b}) = \underline{b} \cdot (\underline{a} \times \underline{b}) = 0.$$

2.4 Orthonormal Bases and Components

Choose vectors $\underline{e}_1, \underline{e}_2, \underline{e}_3$ that are orthonormal i.e. unit vectors all orthogonal to each other.

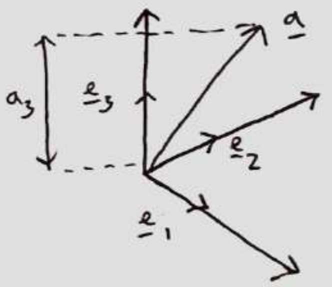
$$\underline{e}_i \cdot \underline{e}_j = \begin{cases} 1 & i=j \\ 0 & i\neq j \end{cases} \qquad i,j = 1,2,3 \qquad (*)$$

Equivalent to choosing Cartesian axes along these directions.

$\{\underline{e}_i\}$ is a basis. Any vector can be written

$$\underline{a} = \sum_i a_i \underline{e}_i = a_1\underline{e}_1 + a_2\underline{e}_2 + a_3\underline{e}_3$$

Each component is uniquely determined by $a_i = \underline{e}_i \cdot \underline{a}$.



Often identify vector $\underline{a}$ with set of components (a_1, a_2, a_3) or $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$

Scalar product in components:

$$\underline{a}\cdot\underline{b} = (a_1\underline{e}_1 + a_2\underline{e}_2 + a_3\underline{e}_3)\cdot(b_1\underline{e}_1 + b_2\underline{e}_2 + b_3\underline{e}_3)$$

By $(*)$ we have $\underline{a}\cdot\underline{b} = a_1b_1 + a_2b_2 + a_3b_3$

When $\underline{a} = \underline{b}$, $\underline{a}\cdot\underline{a} = |\underline{a}|^2$

Vector product in components

WLOG choose basis that is also right-handed: $\underline{e}_1 \times \underline{e}_2 = \underline{e}_3$,

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

(One line implies the other two).

Then $\underline{a}\times\underline{b} = (a_1\underline{e}_1 + a_2\underline{e}_2 + a_3\underline{e}_3)\times(b_1\underline{e}_1 + b_2\underline{e}_2 + b_3\underline{e}_3)$

$$= (a_2b_3 - a_3b_2)\underline{e}_1 + (a_3b_1 - a_1b_3)\underline{e}_2 + (a_1b_2 - a_2b_1)\underline{e}_3$$

Examples $\quad \underline{a} = \begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix} \quad \underline{b} = \begin{pmatrix} 7 \\ -3 \\ 5 \end{pmatrix}$

$$\underline{a} \times \underline{b} = \begin{pmatrix} 0 \times 5 - (-1)(-3) \\ 7(-1) - 2 \times 5 \\ 2(-3) - 0 \times 7 \end{pmatrix} = \begin{pmatrix} -3 \\ -17 \\ -6 \end{pmatrix}$$

Check with

$\underline{a} \cdot \underline{a} \times \underline{b} = 0$

$\underline{b} \cdot \underline{a} \times \underline{b} = 0$

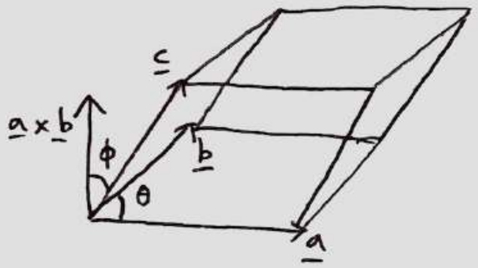
2.5 Triple Products

(a) Scalar Triple Product : $\underline{a} \cdot (\underline{b} \times \underline{c}) = \underline{b} \cdot (\underline{c} \times \underline{a}) = \underline{c} \cdot (\underline{a} \times \underline{b})$

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

$= [\underline{a}, \underline{b}, \underline{c}]$

Interpretation: $|\underline{c} \cdot (\underline{a} \times \underline{b})|$ = volume of parallelepiped as shown



Assume $\theta, \phi < \frac{\pi}{2}$

base $|\underline{a} \times \underline{b}|$

$$|\underline{c}| = |\underline{a} \times \underline{b}| \cos\phi = \underbrace{|\underline{a}||\underline{b}| \sin\theta}_{\text{scalar area of base}} \underbrace{|\underline{c}| \cos\phi}_{\text{height}}$$

$\underline{c} \cdot (\underline{a} \times \underline{b})$ is a signed volume :

if $\underline{c} \cdot (\underline{a} \times \underline{b}) > 0$ then $\underline{a}, \underline{b}, \underline{c}$ is a right-handed set

$\underline{c} \cdot (\underline{a} \times \underline{b}) = 0$ iff $\underline{a}, \underline{b}, \underline{c}$ are coplanar

Example $\quad$ Recall $\underset{\underline{a}}{\begin{pmatrix} 2 \\ 0 \\ -1 \end{pmatrix}} \times \underset{\underline{b}}{\begin{pmatrix} 7 \\ -3 \\ 5 \end{pmatrix}} = \begin{pmatrix} -3 \\ -17 \\ 6 \end{pmatrix}$

If $\underline{c} = \begin{pmatrix} 3 \\ -3 \\ 7 \end{pmatrix}$ then $\underline{c} \cdot \underline{a} \times \underline{b} = 0$ so $\underline{c}$ is spanned by $\underline{a}$ and $\underline{b}$ - lie in their plane

$$\underline{c} = \underline{b} - 2\underline{a}$$

Scalar triple product in components

$$\underline{a} \cdot \underline{b} \times \underline{c} = a_1 b_2 c_3 - a_1 b_3 c_2 + a_2 b_3 c_1 - a_2 b_1 c_3 + a_3 b_1 c_2 - a_3 b_2 c_1$$

$$= \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}$$

$\underline{b} \times \underline{c} \perp \underline{b}$ and $\underline{c}$
and $\underline{a} \times (\underline{b} \times \underline{c})$ is perp. to both $\underline{a}$ and $\underline{b} \times \underline{c}$
so $\underline{a} \times (\underline{b} \times \underline{c})$ lies in the plane spanned by $\underline{b}$ and $\underline{c}$: note $(\underline{a} \cdot \underline{c})\underline{b} - (\underline{a} \cdot \underline{b})c \in \mathrm{span}\{b, c\}$

b) Vector Triple Product

$$\underline{a} \times (\underline{b} \times \underline{c})$$

This is not associative

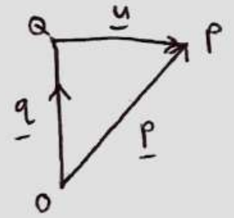
$$= (\underline{a} \cdot \underline{c})\underline{b} - (\underline{a} \cdot \underline{b})\underline{c}$$

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

(prove these by index notation later)

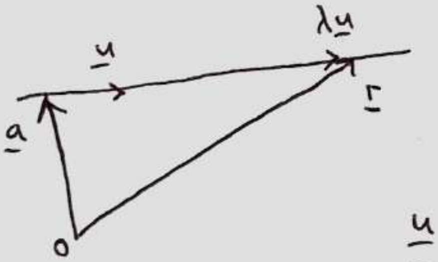
2.6 Lines, Planes and Vector Equations

Vectors defined as position vectors from origin O but definition of addition means we can use them to describe displacements between points.



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

(a) General point on a line through $\underline{a}$ with direction $\underline{u}$ has position vector $\underline{r} = \underline{a} + \lambda \underline{u}$ $\quad \lambda \in \mathbb{R}$



This is parametric form.

Alternative form : cross with $\underline{u}$

$$\underline{u} \times \underline{r} = \underline{u} \times \underline{a} \qquad (\text{as } \underline{u} \parallel \lambda \underline{u})$$

Conversely $\underline{u} \times (\underline{r} - \underline{a}) = 0 \Rightarrow \underline{r} - \underline{a} = \lambda \underline{u}$: the same thing.

" $\underline{u}$ parallel to $\underline{r} - \underline{a}$ "

Consider related eqn

$$\boxed{\underline{u} \times \underline{r} = \underline{c}}$$ where $\underline{u}, \underline{c}$ are given vectors

$$\underline{u} \cdot (\underline{u} \times \underline{r}) = \underline{u} \cdot \underline{c}$$

$$\Rightarrow \quad 0 = \underline{u} \cdot \underline{c}$$

$\Rightarrow$ If $\underline{u} \cdot \underline{c} \neq 0$ then equation is inconsistent

If $\underline{u} \cdot \underline{c} = 0$ then note $\underline{u} \times (\underline{u} \times \underline{c}) = -|\underline{u}|^2 \underline{c}$ (by VTP)

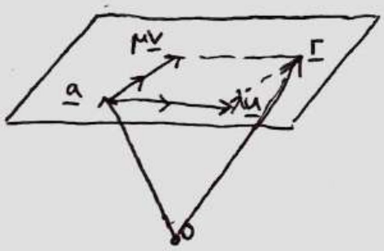
so $\underline{a} = -\dfrac{1}{|\underline{u}|^2}(\underline{u} \times \underline{c})$ is one solution.

and $\underline{r} = \underline{a} + \lambda \underline{u}$ is a general solution - a line as before. (as $\underline{u} \times \underline{a} = \underline{u} \times (\underline{a} + \lambda \underline{u})$)

b) <u>Planes</u>

General point on a plane through $\underline{a}$ with directions in plane $\underline{u}, \underline{v}$ $(\underline{u} \not\parallel \underline{v})$ is

$$\underline{r} = \underline{a} + \lambda \underline{u} + \mu \underline{v} \qquad \lambda, \mu \in \mathbb{R}$$

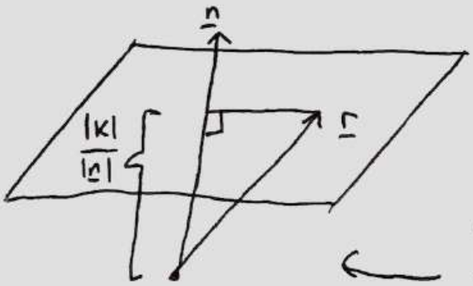


Alternatively ~~wit~~ without parameters

Dot with $\underline{n} = \underline{u} \times \underline{v}$ (normal vector)

($\neq 0$ since $\underline{u} \not\parallel \underline{v}$ but not necessarily unit)

$$\boxed{\underline{n} \cdot \underline{r} = K = \underline{n} \cdot \underline{a}}$$



$\dfrac{\underline{n} \cdot \underline{r}}{|\underline{n}|} = \dfrac{|K|}{|\underline{n}|}$ is the component of $\underline{r}$ along $\underline{n}$

(see L3 definition for $\frac{\underline{a} \cdot \underline{b}}{|\underline{a}|}$)

$\leftarrow$ $\dfrac{|K|}{|\underline{n}|}$ is $\perp$ distance from $\underline{O}$ to the plane.

(c) Other Vector Equations

* expand using dot product

Other equations with geometrical interpretation quadratic in $\underline{r}$

e.g. $|\underline{r}|^2 + \underline{r} \cdot \underline{a} = k$ for $\underline{a}$ constant :

Complete the square : $|\underline{r} \overset{*}{+} \frac{1}{2}\underline{a}|^2 = k + \frac{1}{4}|\underline{a}|^2$

Equation of sphere with centre $-\frac{1}{2}a$, radius $\left(k + \frac{1}{4}|\underline{a}|^2\right)^{1/2}$

for $k > \frac{1}{4}|\underline{a}|^2$.

Consider $\underline{r} + \underline{a} \times (\underline{b} \times \underline{r}) = \underline{c}$ $\leftarrow$ ($\underline{a}, \underline{b}, \underline{c}$ given)

(1) $\iff$ $\underline{r} + (\underline{a} \cdot \underline{r})\underline{b} - (\underline{a} \cdot \underline{b})\underline{r} = \underline{c}$ $\rightarrow$ (by VTP)

Dot with $\underline{a}$: $\underline{a} \cdot \underline{r} + \cancel{(\underline{a} \cdot \underline{r})(\underline{a} \cdot \underline{b})} - \cancel{(\underline{a} \cdot \underline{r})(\underline{a} \cdot \underline{b})} = \underline{a} \cdot \underline{c}$

$\Rightarrow$ $\underline{a} \cdot \underline{r} = \underline{a} \cdot \underline{c}$: eqn of plane

Then we can sub this back in (1) :

$$(1 - \underline{a} \cdot \underline{b})\,\underline{r} = \underline{c} - (\underline{a} \cdot \underline{c})\,\underline{b} \qquad (2)$$

then if $\underline{a} \cdot \underline{b} \neq 1$ then there is a unique solution.

$$\underline{r} = \frac{1}{1 - \underline{a} \cdot \underline{b}}\left(\underline{c} - (\underline{a} \cdot \underline{c})\,\underline{b}\right), \quad \text{a point.}$$

If $\underline{a} \cdot \underline{b} = 1$ and $\underline{c} - (\underline{a} \cdot \underline{c})\,\underline{b} \neq 0$ then (2) is inconsistent - no sol. to (1).

If $\underline{a} \cdot \underline{b} = 1$ and $\underline{c} - (\underline{a} \cdot \underline{c})\,\underline{b} = 0$

then we can put into (1) : $(\underline{a} \cdot \underline{r} - \underline{a} \cdot \underline{c})\,\underline{b} = 0$

giving the plane $\underline{a} \cdot \underline{r} = \underline{a} \cdot \underline{c}$ as the solution.

2.7 Index (suffix) notation and summation convention

(a) Components: δ and ε

Write vectors $\underline{a}, \underline{b}, \dots,$ in terms of components $a_i, b_i, \dots$ wrt to an orthonormal right-handed basis e_i.

Indices $i, j, k, l, p, q, \dots$ take values $1, 2, 3$

If $\underline{c} = \alpha \underline{a} + \beta \underline{b}$

$$\Longleftrightarrow \quad c_i = [\alpha \underline{a} + \beta \underline{b}]_i = \alpha a_i + \beta b_i \quad \text{for } i = 1, 2, 3$$

A free index can take any value.

So $\underline{a} \cdot \underline{b} = \sum_i a_i b_i = \sum_j a_j b_j$

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

$$\Longleftrightarrow \quad x_j = a_j + \left(\sum_k b_k c_k\right) d_j \qquad \text{where } j \text{ is a free index.}$$

Definition (Kronecker delta) δ_{ij}

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

$$\begin{pmatrix} \delta_{11} & \delta_{12} & \delta_{13} \\ \delta_{21} & \delta_{22} & \delta_{23} \\ \delta_{31} & \delta_{32} & \delta_{33} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

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

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

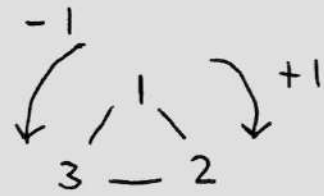
$$= \sum_{ij} a_i b_j \underline{e}_i \cdot \underline{e}_j = \sum_{ij} a_i b_j \delta_{ij}$$

<u>Definition</u> <u>Levi-Civita epsilon</u>

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

$$\varepsilon_{123} = \varepsilon_{231} = \varepsilon_{312} = +1$$

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



ε_{ijk} if i, j, k not all distinct.

ε_{ijk} is totally antisymmetric : exchanging any pair of indices changes the sign.

Then $$\boxed{\underline{e}_i \times \underline{e}_j = \sum_k \varepsilon_{ijk}\, \underline{e}_k}$$

e.g. $$\underline{e}_2 \times \underline{e}_1 = \sum_k \varepsilon_{21k}\, \underline{e}_k$$ but only $k = 3$ gives non zero one:

$$= \varepsilon_{213}\, \underline{e}_3 = \underline{-\underline{e}_3}.$$

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

$$= \sum_{ij} a_i b_j\, \underline{e}_i \times \underline{e}_j$$

$$= \underline{\sum_{ijk} a_i b_j\, \varepsilon_{ijk}\, \underline{e}_k}$$ which tells us a general component of $\underline{a} \times \underline{b}$.

$$\boxed{(\underline{a} \times \underline{b})_k = \sum_{ij} \varepsilon_{ijk}\, a_i b_j}$$

e.g. $$(\underline{a} \times \underline{b})_3 = \sum_{ij} \varepsilon_{ij3}\, a_i b_j = \varepsilon_{123} a_1 b_2 + \varepsilon_{213} a_2 b_1$$

$$= a_1 b_2 - a_2 b_1$$

(b) Summation convention

With component and index notation, indices that appear twice in any given term are usually summed over.

In summation convention, omit $\sum$ for repeated indices.

<u>Examples</u> (i) $a_i \delta_{ij}$ is understood to be a sum over i

$$= a_1 \delta_{1j} + a_2 \delta_{2j} + a_3 \delta_{3j}$$

$$= \begin{cases} a_1 & \text{if } j=1 \\ a_2 & \text{if } j=2 \\ a_3 & \text{if } j=3 \end{cases} = a_j$$

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

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

$$= \varepsilon_{jki} a_j b_k$$

(iv) ~~(a × b ×~~ $(\underline{a} \cdot \underline{b} \times \underline{c}) = \varepsilon_{ijk} a_i b_j c_k$

$\sum_{ijk}$ understood

(v) δ_{ii} means sum over

so is $\delta_{11} + \delta_{22} + \delta_{33} = 3$ (not just 1)

(vi) $[(\underline{a} \cdot \underline{c}) \underline{b} - (\underline{a} \cdot \underline{b}) \underline{c}]_i$

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

$$= a_j c_j b_i - a_j b_j c_i$$ $\qquad \sum_j$ understood

Summation Convention Rules

(i) An index occurring exactly <u>once</u> in any given term must appear once in every term in an equation. It can take any value 1, 2, 3 – free index.

(ii) An index occurring exactly <u>twice</u> in a given term is summed over – called a repeated / contracted / dummy index

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

<u>Application</u> Proof of vector triple product identity. Consider

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

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

$$= (\varepsilon_{ijk}\, \varepsilon_{pqk})\, a_j\, b_p\, c_q$$

Now $\varepsilon_{ijk}\, \varepsilon_{pqk} = \delta_{ip}\,\delta_{jq} - \delta_{iq}\,\delta_{jp}$ $\quad (*)$
(see part c)

Using $(*)$ $\quad [\underline{a} \times (\underline{b} \times \underline{c})]_i = \underbrace{\delta_{ip}\,\overbrace{\delta_{jq}\, a_j\, b_p\, c_q}^{\underline{a}\cdot\underline{c}}}_{b_i} - \delta_{iq}\,\delta_{jp}\, a_j\, b_p\, c_q$

$$= (\underline{a} \cdot \underline{c})\, b_i - (\underline{a} \cdot \underline{b})\, c_i$$

$$= [(\underline{a} \cdot \underline{c})\, \underline{b} - (\underline{a} \cdot \underline{b})\, \underline{c}]_i$$ as required. $\square$

c) $\varepsilon\varepsilon$ identities

$$\varepsilon_{ijk}\,\varepsilon_{pqk} = \delta_{ip}\delta_{jq} - \delta_{iq}\delta_{jp} = \varepsilon_{kij}\,\varepsilon_{kpq}. \qquad (*)$$

<u>Check</u>: RHS and LHS are antisymmetric under $i \leftrightarrow j$ and $p \leftrightarrow q$ so when $i=j$ or $p=q$ we get 0.

Both vanish if i and j or p and q take same value.

Sufficient to check that e.g. $i=p=1 \quad j=q=2$

$$\left.\begin{aligned} \text{LHS} &= \varepsilon_{123}\,\varepsilon_{123} = +1 \\ \text{RHS} &= \delta_{11}\delta_{22} - \delta_{12}\delta_{21} = +1 \end{aligned}\right] \checkmark$$

Or $i=q=1, \quad j=p=2$

$$\text{LHS} = \varepsilon_{123}\,\varepsilon_{213} = (+1)(-1) = -1$$

$$\text{RHS} = \delta_{12}\delta_{21} - \delta_{11}\delta_{22} = -1$$

All other index combinations which give nonzero results work similarly.

– $\boxed{\varepsilon_{ijk}\,\varepsilon_{pjk} = \cancel{2\delta_{jp}}\ 2\delta_{ip}}$ also important
Can be derived from above

LHS of $(*)$: set $q=j$ so

$$\varepsilon_{ijk}\,\varepsilon_{pjk} = \delta_{ip}\delta_{jj} - \delta_{ij}\delta_{jp}$$

$$= 3\delta_{ip} - \delta_{ip} = \underline{2\delta_{ip}}.$$

– $\boxed{\varepsilon_{ijk}\,\varepsilon_{ijk} = 6}$ Use previous expression $i=p$:

$$\varepsilon_{ijk}\,\varepsilon_{ijk} = 2\delta_{ii} = 6$$

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

ε_{ijk} is completely antisymmetric in i, j, k

ε_{pqr} is completely antisymmetric in p, q, r

$\Rightarrow$ LHS and RHS agree up to an overall factor (as RHS is constructed to have that property).

To check factor is 1, consider $i = p = 1$, $j = q = 2$, $k = r = 3$ and find that LHS = RHS = $\underline{1}$.

Can derive $(*)$ from $(\dagger)$. (Exercise).

Vectors and Matrices - Lecture 6 21/10/20

3. Vectors in General : $\mathbb{R}^n$, $\mathbb{C}^n$

3.1 Vectors in $\mathbb{R}^n$

By regarding vectors as sets of components, it is easy to generalise from 3 to n dimensions. Let

$$\mathbb{R}^n = \{ \underline{x} = (x_1, x_2, \dots, x_n) : x_i \in \mathbb{R} \}$$

a) Define

addition: $\underline{x} + \underline{y} = (x_1 + y_1, \dots, x_n + y_n)$

scalar multiplication: $\lambda \underline{x} = (\lambda x_1, \dots, \lambda x_n)$ for any $\underline{x}, \underline{y} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$.

As before, can form linear combinations $\lambda \underline{x} + \mu \underline{y}$ and we have a notion of parallel vectors: $\underline{x} \parallel \underline{y} \Leftrightarrow \underline{x} = \lambda \underline{y}$ or $\underline{y} = \lambda \underline{x}$.

Inner product (scalar product) on $\mathbb{R}^n$ is defined by

$$\underline{x} \cdot \underline{y} = \sum_i x_i y_i = x_1 y_1 + \dots + x_n y_n$$

Properties (i) Symmetric $\underline{x} \cdot \underline{y} = \underline{y} \cdot \underline{x}$

(ii) Bilinear $(\lambda \underline{x} + \lambda' \underline{x}') \cdot \underline{y} = \lambda (\underline{x} \cdot \underline{y}) + \lambda' (\underline{x}' \cdot \underline{y})$

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

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

and $= 0$ iff $\underline{x} = \underline{0}$.

Length or norm of a vector $\underline{x}$ is $|\underline{x}|$ (≥ 0) with

$$|\underline{x}|^2 = \underline{x} \cdot \underline{x}$$

We have notion of orthogonal vectors: $\underline{x} \perp \underline{y}$ iff $\underline{x} \cdot \underline{y} = 0$.

The standard basis for $\mathbb{R}^n$ is

$$\underline{e}_1 = (1, 0, \ldots, 0) \qquad \underline{e}_2 = (0, 1, \ldots, 0), \ldots,$$
$$e_n = (0, 0, \ldots, 1)$$

$$\underline{x} = \sum_i x_i \underline{e}_i \quad \text{and} \quad \underline{e}_i \cdot \underline{e}_j = \delta_{ij}$$

b) Cauchy-Schwarz and Triangle Inequalities

Proposition (Cauchy-Schwarz)

$$|\underline{x} \cdot \underline{y}| \leq |\underline{x}||\underline{y}|$$

with equality iff $\underline{x}$ and $\underline{y}$ are parallel

Geometrical deductions:

(i) Setting $\underline{x} \cdot \underline{y} = |\underline{x}||\underline{y}| \cos\theta$ allows us to define the angle θ between x and y in $\mathbb{R}^n$.

(ii) Δ inequality holds: $|\underline{x} + \underline{y}| \leq |\underline{x}| + |\underline{y}|$

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

If $\underline{y} \neq \underline{0}$ then consider $|\underline{x} - \lambda\underline{y}|^2 = (\underline{x} - \lambda\underline{y}) \cdot (x - \lambda\underline{y})$

by bilinearity it's $|\underline{x}|^2 - 2\lambda\underline{x} \cdot \underline{y} + \lambda^2|\underline{y}|^2$

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

a real quadratic in λ with at most one real root: discriminant ≤ 0

$$(-2\underline{x} \cdot \underline{y})^2 - 4|\underline{x}|^2|\underline{y}|^2 \leq 0 \Rightarrow \text{required inequality.}$$

$\Rightarrow$ equality iff discriminant is 0

$$\Leftrightarrow \quad \underline{x} = \lambda\underline{y}.$$

<u>Proof</u> (Δ inequality):

$$|\underline{x}+\underline{y}|^2 = |\underline{x}|^2 + 2\underline{x}\cdot\underline{y} + |\underline{y}|^2$$
$$\leq |\underline{x}|^2 + 2|\underline{x}||\underline{y}| + |\underline{y}|^2$$
$$= (|\underline{x}|+|\underline{y}|)^2, \quad \text{as required.}$$

(Also check theoretical physics notebook 1, chapter 5 of QM for another C/S and Δ proof).

(c) <u>Comments</u>

Inner product on $\mathbb{R}^n$ is $\underline{a}\cdot\underline{b} = \underline{\delta_{ij}\, a_i b_j}$ using SC (sum over i and j implied)

For $n=3$ this component def. matches geometrical def.

In $\mathbb{R}^3$ also have component definition of cross product

$$(\underline{a}\times\underline{b})_i = \varepsilon_{ijk}\, a_j b_k$$

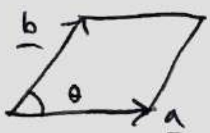
In $\mathbb{R}^n$ have $\underbrace{\varepsilon_{ij\ldots l}}_{n \text{ indices}}$ is again totally antisymmetric (see ch5)

so can't use this to define vector-valued product except for $n=3$: >3 dimensions $\Rightarrow$ no unique normal vector.

But in $\mathbb{R}^2$ have ε_{ij} with $\varepsilon_{12} = -\varepsilon_{21} = 1$

so product of 2 vectors can be defined using this as an "additional" scalar product

$$[\underline{a},\underline{b}] = \varepsilon_{ij}\, a_i b_j = \underline{a_1 b_2 - a_2 b_1}$$



Geometrically this is signed area of parallelogram

$$|[\underline{a},\underline{b}]| = |\underline{a}||\underline{b}|\sin\theta$$

Compare with $[\underline{a}, \underline{b}, \underline{c}] = \varepsilon_{ijk} a_i b_j c_k$

(signed) volume of parallelepiped in $\mathbb{R}^3$.

See invariant tensors

3.2 Vector spaces (not in this course)

(a) Axioms, span, subspaces

Let V be a set of objects called vectors with operations

$\underline{v} + \underline{w} \in V \qquad \forall \underline{v}, w \in V$

$\lambda \underline{v} \in V \qquad \forall \underline{v} \in V, \ \lambda \in \mathbb{R}$

V is called a real vector space if

- V with $+$ is an abelian group

and

- Scalar multiplication / vector addition satisfy

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

Smooth: infinitely differentiable

These axioms apply to geometrical vectors.

Example $V = \{ f : [0,1] \to \mathbb{R}^{*}: f \text{ smooth}, f(0) = f(1) = 0 \}$

This is a real vector space with

$$(f+g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda f(x)$$

(see Ch 8 of TTM QM notes on functions as vectors)

(can check conditions above).

A subspace of a real vector space V (e.g. $\mathbb{R}^n$) is a vector subset $U \subseteq V$ that is itself a vector space.

Note that a non-empty subset is a subspace iff

$$\underline{v}, \underline{w} \in U \Rightarrow \lambda \underline{v} + \mu \underline{w} \in U \quad \forall \lambda, \mu \in \mathbb{R}.$$

For any vectors $\underline{v}_1, \ldots, \underline{v}_r$ in V, their span

$$\text{span}\{\underline{v}_1, \ldots, \underline{v}_r\} = \{\lambda_1 \underline{v}_1 + \ldots + \lambda_r \underline{v}_r : \lambda_i \in \mathbb{R}\}$$

is a subspace of V.

V and $\{\underline{0}\}$ are also subspaces.

Examples in $\mathbb{R}^3$

A line or plane through 0 is a subspace.
but a line or plane not passing through 0 is not a subspace.

e.g. $\underline{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$ $\underline{v}_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ $\underline{n} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$\Rightarrow$ $\text{span}\{\underline{v}_1, \underline{v}_2\} = \{\underline{r} : \underline{n} \cdot \underline{r} = 0\}$, a plane and subspace.

But $\{r : \underline{n} \cdot \underline{r} = 1\}$ is a plane but not a subspace (not through 0).

(since $\underline{r}, \underline{r}'$ on plane $\Rightarrow \underline{n} \cdot \underline{r} = \underline{n} \cdot \underline{r}' = 1$ but $\underline{n} \cdot (\underline{r} + \underline{r}') = 2$: not on plane)

b) Linear Dependence and Independence

For $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_r \in V$ (a real vector space) consider a linear relation

$$\lambda_1 \underline{v}_1 + \lambda_2 \underline{v}_2 + \ldots + \lambda_r \underline{v}_r = \underline{0} \qquad (*)$$

If $(*) \Rightarrow \lambda_i = 0 \ \forall i$, then the vectors form a <u>linearly independent</u> set. (Obey only trivial linear relation).

If $(*)$ holds with at least one $\lambda_i \neq 0$ then they form a <u>linearly dependent</u> set.

<u>Examples</u> In $\mathbb{R}^2$ $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \end{pmatrix} \right\}$ is linearly dependent

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

Note we cannot express $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ in terms of other vectors so LD is <u>not</u> the same as "can express any in terms of others".

Any set containing $\underline{0}$ is linearly dependent

e.g. $\mathbb{R}^2$ $\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \end{pmatrix} \right\}$ Then $0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + \underbrace{(-3)}_{\text{nontrivial}} \begin{pmatrix} 0 \\ 0 \end{pmatrix} = \underline{0}$

In $\mathbb{R}^3$ $\{\underline{a}\}$ is linearly independent iff $\underline{a} \neq \underline{0}$.

$\{\underline{a}, \underline{b}, \underline{c}\}$ linearly independent if <u>$\underline{a} \cdot (\underline{b} \times \underline{c}) \neq 0$.</u>

This is since $\alpha \underline{a} + \beta \underline{b} + \gamma \underline{c} = \underline{0}$

$\Rightarrow \alpha\, \underline{a} \cdot (\underline{b} \times \underline{c}) = 0$ (by $\cdot$ with $(\underline{b} \times \underline{c})$)

$\Rightarrow \alpha = 0$ and $\beta = \gamma = 0$ similarly.

as we assumed $\underline{a} \cdot (\underline{b} \times \underline{c}) \neq 0$ so $\alpha = 0$.

c) Inner Products

This is an additional structure on a real vector space V, that can also be characterised by axioms.

For $\underline{v}, \underline{w} \in V$, inner product is $\underline{v} \cdot \underline{w}$ or $(\underline{v}, \underline{w}) \in \mathbb{R}$.

We require that this satisfies properties in 3.1(a):
(i) symmetric, (ii) bilinear, (iii) positive definite $\quad (\dagger)$

Definition of length / norm and deductions e.g. C-S inequality depend <u>just</u> on these properties.

<u>Example</u> Consider space of functions $V = \{ f : [0,1] \to \mathbb{R} : f \text{ smooth}, f(0) = f(1) = 0 \}$

We can define inner product by

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

: can check this has $(\dagger)$ properties.

Cauchy-Schwarz $\Rightarrow$ $|(f,g)| \leq \|f\| \|g\|$

with $\|f\|^2 = (f, f)$

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

<u>Lemma</u> In any real vector space V with an inner product, if $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_r$ are all nonzero vectors and orthogonal, then they are linearly independent.

<u>Proof</u> If $\sum_i \alpha_i \underline{v}_i = \underline{0}$ then

$$\underbrace{\left(\underline{v}_j, \sum_i \alpha_i \underline{v}_i\right)}_{\text{collapses to } \alpha_j (\underline{v}_j, \underline{v}_j)} = 0 \quad \text{for any specific } j$$

collapses to $\alpha_j (\underline{v}_j, \underline{v}_j) \Rightarrow \underline{\alpha_j = 0}$ (since $(v_j, v_j) \neq 0$ for $v_j \neq 0$). $\square$

3.3 Bases and Dimension

For a vector space V, a <u>basis</u> is a set $B = \{\underline{e}_1, \dots, \underline{e}_n\}$ such that (i) B spans V : $\underline{v} \in V \Rightarrow \underline{v} = \sum_{i=1}^{n} v_i \underline{e}_i$

(ii) B is linearly independent

Given (ii) the coefficients v_i in (i) are unique since

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

$\Rightarrow v_i = v_i'$ as B is linearly independent.

<u>Examples</u> Standard basis for $\mathbb{R}^n$ consists of

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

Can easily check coefficients

Many other bases can be chosen.

e.g. in $\mathbb{R}^2$ we have bases

$$\left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}, \quad \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right\}, \quad \{\underline{a}, \underline{b}\} \text{ where } \underline{a} \not\parallel \underline{b}$$

<u>Theorem</u> If $\{\underline{e}_1, \dots, \underline{e}_n\}$ and $\{\underline{f}_1, \dots, \underline{f}_m\}$ are bases for a real vector space V, then $n = m$.

(<u>Definition</u> : the no. of vectors in any basis is the <u>dimension</u> of the space). $\mathbb{R}^n$ has dimension n.

<u>Proof</u> $\underline{f}_a = \sum_i A_{ai} \underline{e}_i$ as $\underline{e}$ is a basis

and $\underline{e}_i = \sum_a B_{ia} \underline{f}_a$ for constants $A_{ai}, B_{ia} \in \mathbb{R}$

$i, j = 1, \dots, n$

$a, b = 1, \dots, m$

Combine: $$\underline{f}_a = \sum_i A_{ai} \left(\sum_b B_{ib} \underline{f}_b \right)$$

$$= \sum_b \left(\sum_i A_{ai} B_{ib} \right) \underline{f}_b$$

but $\underline{f}_a$, $\underline{f}_b$ are linearly independent and coefficients wrt a basis are unique, hence

$$\sum_i A_{ai} B_{ib} = \delta_{ab}$$

otherwise relation is nontrivial.

Similarly $$\underline{e}_i = \sum_j \left(\sum_a B_{ia} A_{aj} \right) \underline{e}_j$$

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

Now $$\sum_{ia} A_{ai} B_{ia} = \sum_a \delta_{aa} = m$$

$$= \sum_i \delta_{ii} = n$$ so $\underline{m = n}$ $\square$

Note $\{\underline{0}\}$ is sometimes called the trivial vector space, dimension 0.

Steps in proof of basis theorem are within scope of course, but the proof without prompts is non-examinable.

The same applies to the following.

Proposition Let V be a vector space with finite subsets

$Y = \{\underline{w}_1, \dots, \underline{w}_m\}$ that spans V

$X = \{\underline{u}_1, \dots, \underline{u}_k\}$ that is linearly independent.

Then $k \leq n \leq m$ where n = dimension of V.

and (i) a basis can be found as a subset of Y by discarding vectors in Y if necessary

(ii) X can be extended to a basis by adding in additional vectors from Y as necessary.

Proof (i) If Y is lin. indep, then Y is a basis as it spans V and $m = n = \dim V$.

If Y is not linearly independent then there is

$$\sum_{i=1}^{m} \lambda_i \underline{w}_i = \underline{0}, \quad \lambda_i \neq 0 \text{ for some } i.$$

WLOG take $\lambda_m \neq 0$ then

$$\underline{w}_m = -\frac{1}{\lambda_m} \sum_{i=1}^{m-1} \lambda_i \underline{w}_i$$

so span Y = span Y' with $Y' = \{\underline{w}_1, \dots, \underline{w}_{m-1}\}$.

Repeat until a basis is obtained.

(ii) If X spans V then it's already a basis and $k = n = \dim V$.

If not, then $\exists\, \underline{u}_{k+1} \in V$ (not in span X)

PTO

But then since $\underline{u}_{k+1}$ is not in span X, if

$$\sum_{i=1}^{k+1} \mu_i \underline{u}_i = \underline{0} \overset{\text{then}}{\Rightarrow} \mu_{k+1} = 0 \quad (\overset{\text{as}}{\underline{u}_{k+1}} \notin \text{span } X)$$

Then $\mu_i = 0$ for $i = 1, \ldots, k$ (X lin. indep)

Hence $X' = \{\underline{u}_1, \ldots, \underline{u}_k, \underline{u}_{k+1}\}$ is lin. indep.

Furthermore we can choose $\underline{u}_{k+1}$ from Y ($Y \overset{\text{if}}{\subseteq} \text{span } X \overset{\text{then}}{\Rightarrow} \text{span } Y \subseteq \text{span } X$)

$\Rightarrow \text{span } X = V$

Repeat $X \to X'$ until a basis is obtained.

Process stops because Y is finite. $\square$

We usually deal with finite dimensional spaces.

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

Note $s_n(x) = \sqrt{2} \sin n\pi x$, $n = 1, 2, \ldots,$
are in V

$$(s_n, s_m) = 2 \int_0^1 \sin n\pi x \sin m\pi x \, dx = \begin{cases} 1 & \text{if } m = n \\ 0 & \text{if } m \neq n \end{cases} = \underline{\delta_{mn}}$$

Orthonormal $\Rightarrow$ linearly independent

But there are infinitely many such functions so V has infinite dimension.

<u>3.4</u> Vectors in $\mathbb{C}^n$

(a) Introduction - definitions

Let $\mathbb{C}^n = \{ \underline{z} = (z_1, \ldots, z_n) : z_j \in \mathbb{C} \}$

and define <u>addition</u> : $\underline{z} + \underline{w} = (z_1 + w_1, \ldots, z_n + w_n)$

<u>scalar multiplication</u> : $\lambda \underline{z} = (\lambda z_1, \lambda z_2, \ldots, \lambda z_n)$

If scalars $\lambda, \mu \in \mathbb{R}$ then $\mathbb{C}^n$ is a real vector space - axioms or key properties satisfy those in 3.2.

If scalars $\lambda, \mu \in \mathbb{C}$ then $\mathbb{C}^n$ is a <u>complex vector space</u> with the same axioms, and definitions of linear combinations, lin. indep/dep, bases, dimension all generalise to complex scalars.

The distinction between real and complex scalars is important.

e.g. $\underline{z} = (z_1, \dots, z_n) \in \mathbb{C}^n$ with $z_j = x_j + iy_j$
$(x_j, y_j \in \mathbb{R})$

Then $z = \sum_j x_j \underline{e}_j + \sum_j y_j \underline{f}_j$ a real linear combination

with $\underline{e}_j = (0, \dots, 1, \dots, 0)$, $\underline{f}_j = (0, \dots, i, \dots, 0)$

$\{\underline{e}_1, \dots, \underline{e}_n, \underline{f}_1, \dots, \underline{f}_n\}$ is a basis for $\mathbb{C}^n$ as a <u>real</u> vector space, so real dimension is $2n$.

But $\underline{\underline{z} = \sum_j z_j \underline{e}_j}$ is a complex linear combination with <u>complex</u> coefficients

so basis runs from $\underline{e}_1$ to $\underline{e}_n$.

So as a complex vector space, dimension is n (over $\mathbb{C}$).

From now on, view $\mathbb{C}^n$ as a complex vector space unless we say otherwise.

(b) <u>Inner Product</u>

Inner product or scalar product on $\mathbb{C}^n$ is defined by

$$(\underline{z}, \underline{w}) = \sum_j \bar{z}_j w_j = \bar{z}_1 w_1 + \dots + \bar{z}_n w_n$$

PTO

Properties (i) Hermitian: $(\underline{w}, \underline{z}) = \overline{(\underline{z}, \underline{w})}$

(ii) Linear / anti-linear

$$(\underline{z}, \lambda \underline{w} + \lambda' \underline{w}') = \lambda(\underline{z}, \underline{w}) + \lambda'(\underline{z}, \underline{w}')$$

$$(\lambda \underline{z} + \lambda' \underline{z}', \underline{w}) = \bar{\lambda}(\underline{z}, \underline{w}) + \lambda'(\underline{z}', \underline{w}')$$

(iii) Positive definite

$$(\underline{z}, \underline{z}) = \sum_j |z_j|^2 \text{, real and } \geq 0$$

$$= 0 \text{ iff } z = \underline{0}$$

Define length or norm of z to be $|z| \geq 0$, $|\underline{z}|^2 = (\underline{z}, \underline{z})$

Also define $\underline{z}, \underline{w} \in \mathbb{C}^n$ to be orthogonal if their inner product is 0: $(\underline{z}, \underline{w}) = 0$.

Note that standard basis for $\mathbb{C}^n$ is then orthonormal:

$(\underline{e}_j, \underline{e}_k) = \delta_{jk}$

Example Complex inner product on $\mathbb{C}$ ($\mathbb{C}^n$, $n = 1$) is $(z, w) = \bar{z} w$

Let $z = a_1 + i a_2$, $w = b_1 + i b_2$ $a_1, a_2, b_1, b_2 \in \mathbb{R}$

Each correspond to $\underline{a}, \underline{b}$ in $\mathbb{R}^2$

then $\bar{z} w = (a_1 b_1 + a_2 b_2) + i(a_1 b_2 - a_2 b_1)$

$= (\underline{a} \cdot \underline{b}) + i [\underline{a}, \underline{b}]$

recover two scalar products in $\mathbb{R}^2$.

Vectors and Matrices – Lecture 9 28/10/20

4. Matrices and Linear Maps

4.1 Introduction

Definitions A <u>linear map</u> or linear transformation is a function $T : V \to W$ between vector spaces V (dim n) and W (dim m) such that

(1) $$T(\lambda \underline{x} + \mu \underline{y}) = T(\lambda \underline{x}) + T(\mu \underline{y}) = \underline{\lambda T(\underline{x}) + \mu T(\underline{y})}$$

for $x, y \in V$, $\lambda, \mu \in \mathbb{R}$ or $\mathbb{C}$

Note: a linear map is completely determined by its action on a basis - deduce this from (1) (exercise)

$\underline{x}' = T(\underline{x}) \in W$ is the <u>image</u> of $x \in V$ under T

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

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

Lemma $\mathrm{Ker}(T)$ is a subspace of V and $\mathrm{Im}(T)$ is a subspace of W

Check: $\underline{0} \in \mathrm{Ker}(T)$ and $\underline{x}, \underline{y} \in \mathrm{Ker}(T) \Rightarrow T(\lambda \underline{x} + \mu \underline{y})$
$= \lambda T(\underline{x}) + \mu T(\underline{y}) = \lambda \underline{0} + \mu \underline{0} = \underline{0} \in \ker(T)$.

$0 \in \mathrm{Im}(T)$ and if $\underline{x}' = T(\underline{x})$, $\underline{y}' = T(\underline{y})$
then $T(\lambda \underline{x} + \mu \underline{y}) = \lambda T(\underline{x}) + \mu T(\underline{y}) = \lambda x' + \mu y'$
so $\lambda x' + \mu y'$ is in the image. □

Examples
(i) Zero linear map $T : V \to W$ so $T(\underline{x}) = \underline{0}$ $\forall \underline{x} \in V$.
(ii) Identity map $T : V \to V$ with $T(\underline{x}) = \underline{x}$
(iii) $V, W = \mathbb{R}^3$ $\underline{x}' = T(\underline{x})$ where

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

(think of like a matrix)

Here $\mathrm{Im}(T) = \left\{ \lambda \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} + \mu \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$, a plane

$\mathrm{Ker}(T) = \left\{ \lambda \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix} \right\}$ a line

b) <u>Rank-Nullity</u>

Define $\dim \mathrm{Im}(T)$ or $\mathrm{rank}(T)$ as the <u>rank</u> of T $(\leq m)$

$\dim \mathrm{Ker}(T) = \mathrm{null}(T)$ as the <u>nullity</u> of T $(\leq n)$

<u>Theorem</u> For $T: V \to W$ a linear map as above,

$$\mathrm{rank}(T) + \mathrm{null}(T) = n = \dim V.$$

<u>Examples</u> Refer to part (a).

<u>Proof</u> (non-examinable)

Let $\underline{e}_1, \dots, \underline{e}_k$ be a basis for $\mathrm{Ker}(T)$ so $T(\underline{e}_i) = 0$ for $i = 1, \dots, k$

Extend this by adding in $\underline{e}_{k+1}, \dots, \underline{e}_n$ to get a basis for V, then claim $B = \{T(\underline{e}_{k+1}), \dots, T(\underline{e}_n)\}$ is a basis for $\mathrm{Im}(T)$. The result then <u>follows</u> since $\mathrm{null}\, T = k$ and $\mathrm{rank}\, T = n-k$.

To prove claim: B spans $\mathrm{Im}(T)$ since for any $\underline{x} = \sum_{i=1}^{n} x_i \underline{e}_i \in V$, $T(x) = \sum_{i=k+1}^{n} x_i T(\underline{e}_i)$

B is lin. independent since

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

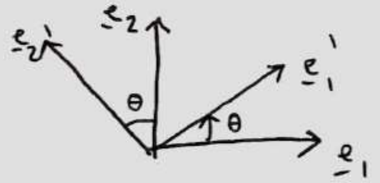
$$\Rightarrow \sum_{i=k+1}^{n} \lambda_i \underline{e}_i \in \mathrm{Ker}(T) \Rightarrow \sum_{i=k+1}^{n} \lambda_i \underline{e}_i = \sum_{i=1}^{k} \mu_i \underline{e}_i$$

$\Rightarrow \lambda_i = 0$, $\mu_i = 0$ as $\{\underline{e}_1, \dots, \underline{e}_n\}$ is linearly independent. So claim is proved. $\square$

4.2 Geometrical Examples

(a) Rotations

In $\mathbb{R}^2$, rotation about $\underline{0}$ through angle θ is defined by



$$\underline{e}_1 \mapsto \underline{e}_1' = \cos\theta\ \underline{e}_1 + \sin\theta\ \underline{e}_2$$
$$\underline{e}_2 \mapsto \underline{e}_2' = -\sin\theta\ \underline{e}_1 + \cos\theta\ \underline{e}_2$$

In $\mathbb{R}^3$, rotation about axis given by $\underline{e}_3$ defined as above but with $\underline{e}_3 \mapsto \underline{e}_3' = \underline{e}_3$ in addition.

Now extend to rotation in $\mathbb{R}^3$ about axis given by $\hat{\underline{n}}$.

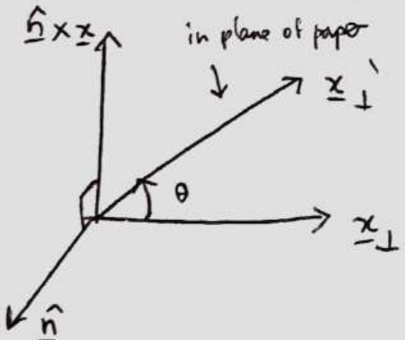
For any $\underline{x}$, resolve $\parallel$ and $\perp$ to $\hat{\underline{n}}$:

$$\underline{x} = \underline{x}_{\parallel} + \underline{x}_{\perp} \quad \text{with} \quad \underline{x}_{\parallel} = (\underline{x}\cdot\underline{n})\,\underline{n}$$

Then $\underline{x}_{\parallel} \mapsto \underline{x}_{\parallel}' = \underline{x}_{\parallel}$

$$\underline{x}_{\perp} \mapsto \underline{x}_{\perp}' = (\cos\theta)\,\underline{x}_{\perp} + (\sin\theta)\,\underline{n}\times\underline{x}$$

This follows by considering plane perpendicular to $\underline{n}$



and noting $|\underline{x}_{\perp}| = |\underline{n}\times\underline{x}|$

So result follows by comparison with $\mathbb{R}^2$ version.

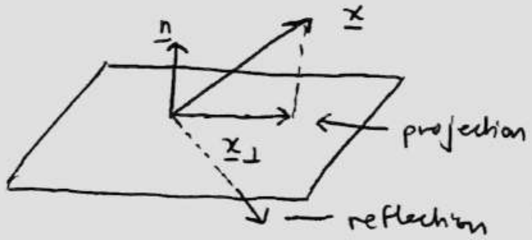
Reassemble: $\underline{x} \mapsto \underline{x}' = \underline{x}_{\parallel}' + \underline{x}_{\perp}' = (\underline{n}\cdot\underline{x})\,\underline{n} + \cos\theta(\underline{x} - (\underline{x}\cdot\underline{n})\underline{n})$
$+ \sin\theta\ \underline{n}\times\underline{x}$

or $\underline{\underline{x}' = (\cos\theta)\,\underline{x} + (1-\cos\theta)(\underline{n}\cdot\underline{x})\,\underline{n} + (\sin\theta)\,\underline{n}\times\underline{x}}$

b) Reflections and Projections

For a plane with unit normal n, define projection by

$$\underline{x} \mapsto \underline{x}' = x_\perp = \underline{x} - (\underline{x}\cdot\underline{n})\,\underline{n}$$



Reflection: $\underline{x}_{\parallel} \mapsto -\underline{x}_{\parallel}$

$\underline{x}_\perp \mapsto \underline{x}_\perp$

i.e. $\underline{x} \mapsto \underline{x}' = \underline{x} - 2(\underline{n}\cdot\underline{x})\,\underline{n}$

Same expressions apply in $\mathbb{R}^2$ with plane replaced by line

(c) Dilations Given scale factors $\alpha, \beta, \gamma > 0$ define a dilation along axes

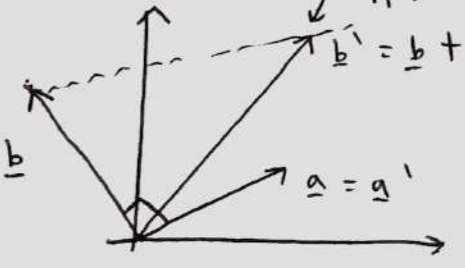
$$\underline{e}_1 \mapsto \underline{e}_1' = \alpha\underline{e}_1$$
$$\underline{e}_2 \mapsto \underline{e}_2' = \beta\underline{e}_2$$
$$\underline{e}_3 \mapsto \underline{e}_3' = \gamma\underline{e}_3$$

d) Shears Let $\underline{a}, \underline{b}$ be orthogonal unit vectors in $\mathbb{R}^3$ i.e. $|\underline{a}| = |\underline{b}| = 1$ and $\underline{a}\cdot\underline{b} = 0$, λ a real parameter

Then define shear $\underline{x} \mapsto \underline{x}' = \underline{x} + \lambda\underline{a}(\underline{x}\cdot\underline{b})$

$$\boxed{\underline{a}' = \underline{a} \qquad \underline{b}' = \underline{b} + \lambda\underline{a}}$$

dashed supposed to be parallel to $\underline{a}$



Similar in $\mathbb{R}^3$

Vectors and Matrices – Lecture 10 30/10/20

4.3 Matrices as Linear Maps: $\mathbb{R}^n \to \mathbb{R}^m$

a) Definitions

Consider a linear map $T: \mathbb{R}^n \to \mathbb{R}^m$

$$\{\underline{e}_i\} \quad \{\underline{f}_a\} \quad \text{bases}$$

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

$$T(x) = x' \quad \text{with} \quad x = \sum_i x_i \underline{e}_i = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

$$x' = \sum_a x_a' \underline{f}_a = \begin{pmatrix} x_1' \\ \vdots \\ x_m' \end{pmatrix}$$

Linearity implies that T is fixed by specifying

$$T(\underline{e}_i) = \underline{e}_i' = \underline{C}_i \in \mathbb{R}^m$$

We take these as columns of an $m \times n$ <u>array</u> or <u>matrix</u> M with rows $\underline{R}_a \in \mathbb{R}^n$:

$$\begin{pmatrix} \uparrow & & \uparrow \\ \underline{C}_1 & \cdots & \underline{C}_n \\ \downarrow & & \downarrow \end{pmatrix} = \begin{pmatrix} \underline{R}_1 & \cdots \\ \vdots & \\ \underline{R}_m & \cdots \end{pmatrix}$$

M has <u>entries</u> $M_{ai} \in \mathbb{R}$ where a labels rows and i labels columns, so

$$(\underline{C}_i)_a = M_{ai} = (\underline{R}_a)_i$$

Action of T is then given by matrix M multiplying $\underline{x}$:

$$\underline{x}' = M\underline{x}$$

defined by $\underline{x_a' = M_{ai} x_i}$ (using summation convention: $\sum_i$)

or

$$\begin{pmatrix} x_1' \\ x_2' \\ \vdots \\ x_m' \end{pmatrix} = \begin{pmatrix} M_{11} & M_{12} & \cdots & M_{1n} \\ M_{21} & M_{22} & \cdots & M_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ M_{m1} & M_{m2} & \cdots & M_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} M_{11}x_1 + M_{12}x_2 + \dots + M_{1n}x_n \\ M_{21}x_1 + M_{22}x_2 + \dots + M_{2n}x_n \\ \vdots \\ M_{m1}x_1 + M_{m2}x_2 + \dots + M_{mn}x_n \end{pmatrix}$$

~~To check~~ To check that matrix multiplication is action T:

$$x' = T\left(\sum_i x_i \underline{e}_i\right) = \sum_i x_i\, T(\underline{e}_i) = \sum_i x_i \underline{C}_i$$

$$\Rightarrow \quad x_a' = \sum_i x_i (\underline{C}_i)_a = \sum_i M_{ai} x_i$$

Now regard properties of T as properties of M:

$$\operatorname{Im}(T) = \operatorname{Im}(M) = \operatorname{span}\{\underline{C}_1, \dots, \underline{C}_n\}$$

Image of M is span of columns

Note $x_a' = M_{ai} x_i = (\underline{R}_a)_i x_i = \underline{R} \cdot \underline{x}$

$$\operatorname{Ker}(T) = \operatorname{Ker}(M) = \{\underline{x} : \underline{R}_a \cdot \underline{x} = 0 \ \ \forall a\}$$

Kernel of M is subspace $\perp$ to all rows.

b) <u>Examples</u>

(i) Zero map $\mathbb{R}^n \to \mathbb{R}^m$ is <u>zero</u> matrix

$M = 0$ with $M_{ai} = 0$

(ii) Identity map $\mathbb{R}^n \to \mathbb{R}^n$ is <u>identity</u> or <u>unit</u> matrix

$$\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \quad \text{with } M_{ai} = \delta_{ai}.$$

(iii) $\mathbb{R}^3 \to \mathbb{R}^3$, $x' = T(\underline{x}) = M\underline{x}$ with

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

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

$$\underline{C}_1 = \begin{pmatrix} 3 \\ -1 \\ 2 \end{pmatrix} \quad \underline{C}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \quad \underline{C}_3 = \begin{pmatrix} 5 \\ -2 \\ 3 \end{pmatrix}$$

and $\operatorname{Im}(M)$ is span $\{\underline{C}_1, \underline{C}_2, \underline{C}_3)$

$= \operatorname{span}(\underline{C}_1, \underline{C}_2)$ as $\underline{C}_3 \in \operatorname{span}(\underline{C}_1, \underline{C}_2)$.

$\underline{R}_1 = (3\ 1\ 5) \qquad \underline{R}_2 = (-1\ 0\ 2) \qquad \underline{R}_3 = (2\ 1\ 3)$

$R_2 \times R_3 = (2\ -1\ -1) = \underline{u}$ which is perpendicular to all rows

$\ker(T) = \ker(M) = \{ \lambda \underline{u} \}$

(iv) Rotation through θ in $\mathbb{R}^2$

$$\underline{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \mapsto \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} \qquad \underline{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} -\sin\theta \\ \cos\theta \end{pmatrix}$$

so matrix is $\begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

(v) Dilation $\underline{x}' = M\underline{x}$ with scale factors α, β, γ along axes in $\mathbb{R}^3$

$\underline{e}_1 \mapsto \alpha \underline{e}_1 \qquad \underline{e}_2 \mapsto \beta \underline{e}_2 \qquad \underline{e}_3 = \gamma \underline{e}_3$

so is $\begin{pmatrix} \alpha & 0 & 0 \\ 0 & \beta & 0 \\ 0 & 0 & \gamma \end{pmatrix}$

(v) <u>Reflection</u> in plane perpendicular to $\hat{\underline{n}}$, matrix H must have

$$\underline{x}' = H\underline{x} = \underline{x} - 2(\underline{x}\cdot\underline{n})\,\underline{n}$$

$$x_i' = x_i - 2x_j n_j \underset{\uparrow}{\cancel{m}}{}_i = H_{ij} x_j$$
$$\qquad\qquad\qquad\qquad n$$

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

e.g. $n = \frac{1}{\sqrt{3}}(1, 1, 1) \qquad n_i n_j = \frac{1}{3}$ for all i, j

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

(vii) Shear, matrix S

$$\underline{x}' = S\underline{x} = \underline{x} + \lambda(\underline{b}\cdot\underline{x})\underline{a}$$ ($\underline{a}, \underline{b}$ unit vectors with $\underline{a} \perp \underline{b}$, λ scale factor)

$$x_i' = x_i + \lambda b_j x_j a_i$$

$$= S_{ij} x_j \qquad \underline{S_{ij} = \delta_{ij} + \lambda a_i b_j}$$

e.g. in $\mathbb{R}^2$ $\underline{e}_1 = \begin{pmatrix}1\\0\end{pmatrix}$ $\underline{b} = \underline{e}_2 = \begin{pmatrix}0\\1\end{pmatrix}$:

$$S = \begin{pmatrix}1 & \lambda\\ 0 & 1\end{pmatrix}$$ by definition.

Rotation in $\mathbb{R}^3$ with axis $\underline{n}$ and angle θ; matrix R satisfies

$$\underline{x}' = R\underline{x} = (\cos\theta)\underline{x} + (1-\cos\theta)(\underline{n}\cdot\underline{x})\underline{n} + (\sin\theta)(\underline{n}\times\underline{x})$$

$$\Rightarrow x_i = \cos\theta\, x_i + (1-\cos\theta)\, n_j x_j n_i - \sin\theta\, \varepsilon_{ijk} x_j n_k$$

$$= R_{ij} x_j$$

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

(c) Matrix of a general linear map $V \to W$

Consider a linear map $T: V \to W$ between general real or complex vector spaces of dimension n, m respectively.

Choose bases $\{\underline{e}_i\}$ with $i = 1, \dots, n$ for V

$\{\underline{f}_a\}$ with $a = 1, \dots, m$ for W

The matrix representing T wrt these bases is an $m \times n$ array with entries $M_{ai} \in \mathbb{R}$ or $\mathbb{C}$ defined by

$$T(\underline{e}_i) = \sum_a \underline{f}_a M_{ai}$$

Then $\underline{x}' = T(x)$

$$\Leftrightarrow \quad x_a' = \sum_i M_{ai} x_i = M_{ai} x_i$$

(summation convention)

where $\underline{x} = \sum_i x_i \underline{e}_i$, $\underline{x}' = \sum_a x_a' \underline{f}_a$

subscripts

Given choices of bases $\{\underline{e}_i\}$ and $\{\underline{f}_a\}$,

V is identified with $\mathbb{R}^n$ and similarly

W is identified with $\mathbb{R}^m$

T is identified with an $m \times n$ matrix M

<u>Note</u> Entries in column $\underline{i}$ of M are components of $T(\underline{e}_i)$ with respect to basis $\underline{f}_a$

4.4 Matrix Algebra

(a) Linear Combinations

If $T : V \to W$, $S : V \to W$ are linear maps between real or complex vector spaces V, W of dimension n, m respectively (e.g. $\mathbb{R}^n \to \mathbb{R}^m$), then $\alpha T + \beta S : V \to W$ is also a linear map, where $(\alpha T + \beta S)(\underline{x}) = \alpha T(x) + \beta S(x) \quad \forall x \in V$

If M, N are $m \times n$ matrices for T, S then $\alpha M + \beta N$ is $m \times n$ matrix for $\alpha T + \beta S$ where

$$(\alpha M + \beta N)_{ai} = \alpha M_{ai} + \beta N_{ai} \qquad \begin{array}{l} a = 1, \dots, m \\ i = 1, \dots, n \end{array}$$

(all wrt same bases - standard bases for $\mathbb{R}^n / \mathbb{C}^n$)

(b) Matrix Multiplication

If A is an $m \times n$ matrix with entries A_{ai} and B is an $n \times p$ matrix with entries B_{ir}

then AB is an $m \times p$ matrix defined by

$$(AB)_{ar} = A_{ai} B_{ir} \qquad \text{where } \begin{array}{l} a = 1, \dots, m \\ i = 1, \dots, n \\ r = 1, \dots, p \end{array}$$

summation convention

Product is only defined if # columns of A = # rows of B

Matrix multiplication corresponds to composition of linear maps.

e.g. For real matrices corresponding to maps

$S : \mathbb{R}^p \to \mathbb{R}^n$ with $S(\underline{x}) = B\underline{x}$ $\qquad \underline{x} \in \mathbb{R}^p$

$T : \mathbb{R}^n \to \mathbb{R}^m$ with $T(\underline{y}) = A\underline{y}$ $\qquad \underline{y} \in \mathbb{R}^n$

$\Rightarrow TS : \mathbb{R}^p \to \mathbb{R}^m$ with $TS(\underline{x}) = (AB)\underline{x}$.

Check this:

$$\left[(AB)\,\underline{x} \right]_a = (AB)_{ar}\, x_r$$

summed

but $\underline{A(B\underline{x}) = A_{ai}(B\underline{x})_i = A_{ai}\, B_{ir}\, x_r}$

Matrix multiplication is defined to make this work

<u>Example</u> $\quad A = \begin{pmatrix} 1 & 3 \\ -5 & 0 \\ 2 & 1 \end{pmatrix} \qquad B = \begin{pmatrix} 1 & 0 & -1 \\ 2 & -1 & 3 \end{pmatrix}$

$3\times 2 \qquad\qquad 2\times 3$

$$AB = \begin{pmatrix} 7 & -3 & 8 \\ -5 & 0 & 5 \\ 4 & -1 & 1 \end{pmatrix} \qquad BA = \begin{pmatrix} -1 & 2 \\ 13 & 9 \end{pmatrix}$$

<u>Properties of matrix products</u>

$$(\lambda M + \mu N)P = \lambda MP + \mu NP$$
$$P(\lambda M + \mu N) = \lambda PM + \mu PN$$
$$(MN)P = M(NP)$$

λ, μ scalars
(where products defined)

Identity matrix I with $I_{ij} = \delta_{ij}$

$$IM = M, \qquad MI = M$$

<u>Helpful points of view</u>

(i) Regarding vector $\underline{x} \in \mathbb{R}^n$ as a column vector or $n\times 1$ matrix, then definitions of matrix - vector and matrix - matrix multiplication agree.

(ii) For product AB $\quad$ (A $m\times n$, B $n\times p$), columns $\underline{C}_r(B) \in \mathbb{R}^n$ and $\underline{C}_r(AB) \in \mathbb{R}^m$ $\quad (r = 1, \dots, p)$ are related by $\underline{C}_r(AB) = A\,\underline{C}_r(B)$

(iii) In terms of rows and columns

$$AB = \begin{pmatrix} & \vdots & \\ \leftarrow & R_a(A) & \rightarrow \\ & \vdots & \end{pmatrix} \begin{pmatrix} & \uparrow & \\ \cdots & \underline{C}_r(B) & \cdots \\ & \downarrow & \end{pmatrix}$$

$$(AB)_{ar} = [\underline{R}_a(A)]_i \, [\underline{C}_r(B)]_i \qquad (\Sigma i)$$

$$= \underline{R}_a(A) \cdot \underline{C}_r(B) \quad \text{dot product in } \mathbb{R}^n$$

$$(i = 1, \ldots, n).$$

(c) Matrix Inverses

If A is $m \times n$ then B $(n \times m)$ is a left inverse if $\underset{(n\times n)}{BA = I}$

If C ($\cancel{m \times n}$ $n \times m$) is such that $\underset{(m \times m)}{AC = I}$ then C is a right inverse

If $m = n$: $B = C = A^{-1}$, the inverse:

$$AA^{-1} = A^{-1}A = I$$

Not every matrix has an inverse; if it does, then A is invertible or non-singular.

Consider $\underline{x}, \underline{x}' \in \mathbb{R}^n$ or $\mathbb{C}^n$ and M an $n \times n$ matrix.

If M^{-1} exists then we can solve

$\underline{x}' = M\underline{x}$ for $\underline{x}$ given $\underline{x}'$, by writing

$$M^{-1}\underline{x}' = (M^{-1}M)\underline{x} = \underline{x}$$

Example $n = 2$ $M = \begin{pmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{pmatrix}$

$$x_1' = M_{11}x_1 + M_{12}x_2$$

$$x_2' = M_{21}x_1 + M_{22}x_2$$

$$\Rightarrow \quad M_{22}x_1' - M_{12}x_2' = (\det M)x_1$$

$$-M_{21}x_1' + M_{11}x_2' = (\det M)x_2$$

where $\det M = M_{11}M_{22} - M_{12}M_{21}$

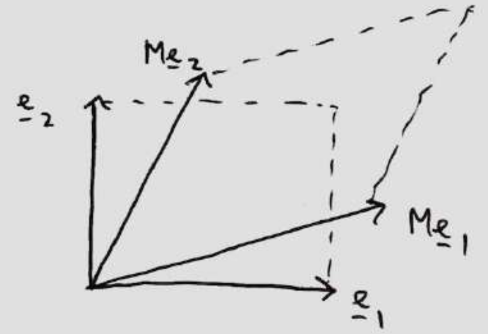
For $\det M \neq 0$, $\underline{M^{-1} = \dfrac{1}{\det M}\begin{pmatrix} M_{22} & -M_{12} \\ -M_{21} & M_{11} \end{pmatrix}}$

Note $\underline{c}_1 = M\underline{e}_1 = \begin{pmatrix} M_{11} \\ M_{21} \end{pmatrix}$ & $\underline{c}_2 = M\underline{e}_2 = \begin{pmatrix} M_{12} \\ M_{22} \end{pmatrix}$

so $\det M = [\underline{c}_1, \underline{c}_2] = [M\underline{e}_1, M\underline{e}_2]$ in $\mathbb{R}^2$

(second scalar product)

gives factor (with sign) by which <u>areas</u> are scaled under action of M.



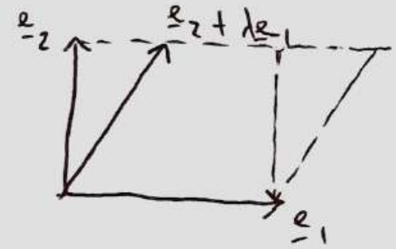
$\det M \neq 0 \iff M\underline{e}_1, M\underline{e}_2$ linearly independent

$\iff \text{Im } M$ has dim 2.

e.g. shear $S(\lambda) = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix}$ ($\underline{a} = \underline{e}_1$, $\underline{b} = \underline{e}_2$, λ parameter)

$\det S(\lambda) = 1$ so areas preserved.

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



<u>Example</u> Recall from 4.3: expression for components of matrix $R(\theta)$ for rotation about axis $\underline{n}$ (fixed) through angle θ.

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

$$= \delta_{ik}\cos^2\theta + 2\cos\theta\,(1-\cos\theta)\, n_i n_k$$
$$+ (1-\cos\theta)^2 n_i n_k - \varepsilon_{ijp}\,\varepsilon_{jkq}\, n_p n_q \sin^2\theta$$

(other cross terms cancel ; $n_j n_j = 1$)

$$= \delta_{ik}\cos^2\theta + \cancel{(1-\cos^2\theta)\, n_i n_k} + \delta_{ik}\, n_p n_p \sin^2\theta - \cancel{\sin^2\theta\, n_i n_k}$$
(by $\varepsilon\varepsilon$ identity)

$$= \delta_{ik}\cos^2\theta + \delta_{ik}\sin^2\theta = \underline{\delta_{ik}}$$ so $\underline{R(\theta)R(-\theta) = I}$.

(d) Transpose and Hermitian Conjugate

If M is an $m \times n$ matrix then the <u>transpose</u> M^T is an $n \times m$ matrix defined by

$$(M^T)_{ia} = M_{ai} \qquad \text{exchanges rows and columns}$$

<u>Properties</u> $(\alpha A + \beta B)^T = \alpha A^T + \beta B^T$

$(AB)^T = B^T A^T \qquad A \ m \times n, \ B \ n \times p$

Check: $[(AB)^T]_{ra} = (AB)_{ar} = A_{ai} B_{ir} = (A^T)_{ia} (B^T)_{ri}$

$= (A^T)_{ia} (B^T)_{ri}$

$= (B^T)_{ri} (A^T)_{ia} = (B^T A^T)_{ra}$

Note $\underline{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ col vec $\in \mathbb{R}^n$ or $m \times 1$ matrix

$\underline{x}^T = (x_1, x_2, \ldots, x_n)$ row vec $\in \mathbb{R}^n$ or $1 \times m$ matrix

Inner product on $\mathbb{R}^n$ is $\underline{x} \cdot \underline{y} = \underline{x}^T \underline{y}$ (scalar 1×1)

(but $\underline{y} x^T$ $n \times n$ matrix, $M_{ij} = x_i y_j$)

If M is $n \times n$ then M is

<u>symmetric</u> iff $M^T = M$ $(M_{ij} = M_{ji})$

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

Any M can be written $\underline{M = S + A}$

$S = \frac{1}{2}(M + M^T) \qquad A = \frac{1}{2}(M - M^T)$

Note if A is 3×3, antisymmetric

then $A_{ij} = \varepsilon_{ijk} a_k$ $\qquad A = \begin{pmatrix} 0 & a_3 & -a_2 \\ -a_3 & 0 & a_1 \\ a_2 & -a_1 & 0 \end{pmatrix}$

$$(A\underline{x})_i = \varepsilon_{ijk} a_k x_j = (\underline{x} \times \underline{a})_i$$

If M is an $m\times n$ matrix, the Hermitian conjugate $M^\dagger$ is an $n\times m$ matrix defined by

$$(M^\dagger)_{ia} = (M_{ai})^* = \overline{M}_{ai}$$

or $M^\dagger = \overline{M}^T = (\overline{M^T})$

If M is square, then M is Hermitian if $\underline{M = M^\dagger}$

M is anti-Hermitian if $\underline{M = -M^\dagger}$.

Similarly if $\underline{z} = \begin{pmatrix} z_1 \\ \cdots \\ z_n \end{pmatrix} \in \mathbb{C}^n$ then

$$\underline{z}^\dagger = (z_1^*, \ldots, z_n^*)$$

and complex inner product on $\mathbb{C}^n$ is

$$(\underline{z}, \underline{w}) = \bar{z}_1 w_1 + \ldots + \bar{z}_n w_n = \underline{z}^\dagger \underline{w}$$

e) <u>Trace</u>

For complex $n\times n$ matrix M, the <u>trace</u> is

$$\mathrm{tr}(M) = M_{ii}$$

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

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

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

$$\mathrm{tr}(I) = \delta_{ii} = n \quad \text{if } I \ n\times n$$

Check $\mathrm{tr}(MN) = \mathrm{tr}(NM)$:

$$\mathrm{tr}(MN) = (MN)_{aa} = M_{ai}N_{ia} = N_{ia}M_{ai} = (NM)_{ii}$$
$$= \mathrm{tr}(NM).$$

If S is $n \times n$ symmetric then let

$$T = S - \frac{1}{n}\mathrm{tr}(S)\, I$$

or $$T_{ij} = S_{ij} - \frac{1}{n}\mathrm{tr}(S)\,\delta_{ij}$$

Then $\mathrm{tr}(T) = 0$ by construction:

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

$$S = \underbrace{T}_{\text{traceless}} + \underbrace{\frac{1}{n}\mathrm{tr}(S)\, I}_{\text{pure trace}}$$

Note if A is $n \times n$ antisymmetric then $\mathrm{tr}(A) = A_{ii} = 0$.

4.5 Orthogonal and Unitary Matrices

A real $n \times n$ matrix U is orthogonal iff

$$U^T U = U U^T = I \qquad (U^T = U^{-1})$$

These conditions can be written $\quad U_{ki}U_{kj} = U_{ik}U_{jk} = \delta_{ij}$

cols of U are orthonormal, rows of U are orthonormal

e.g. $$\begin{pmatrix} & \vdots & \\ \leftarrow & \underline{c}_i & \rightarrow \\ & \vdots & \end{pmatrix}\begin{pmatrix} & \uparrow & \\ \cdots & \underline{c}_j & \cdots \\ & \uparrow & \end{pmatrix} = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

e.g. If $U = R(\theta)$ rotation through θ about axis $\underline{n}$

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

Equivalently we have definition: U is orthogonal iff it preserves the inner product on $\mathbb{R}^n$

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

To check equivalence

$$(U\underline{x})^T(\underline{U}\,\underline{y}) = (\underline{x}^T U^T)(U\underline{y}) = x^T \underbrace{(U^T U)}_{=I} y = x^T y$$

ignore underline

$\Leftrightarrow \quad U^T U = I$.

Moreover in $\mathbb{R}^n$ columns of U are $U\underline{e}_1, \dots, U\underline{e}_n$ so inner product is preserved when U acts on standard basis vectors iff

$$(U\underline{e}_i)\cdot(U\underline{e}_j) = \underline{e}_i\,\underline{e}_j = \delta_{ij}$$

$\Leftrightarrow$ cols of U are orthonormal.

<u>Example</u> To find general 2×2 orthogonal matrix

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

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

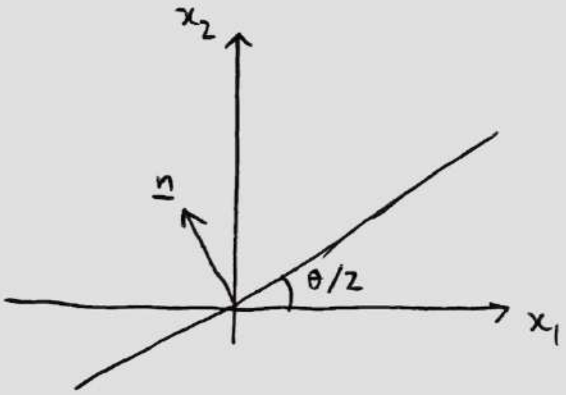
2 cases: $U = R = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}$ rotation in $\mathbb{R}^2$ by angle θ

$U = H = \begin{pmatrix}\cos\theta & \sin\theta\\ \sin\theta & -\cos\theta\end{pmatrix}$ reflection in $\mathbb{R}^2$

To match previous description consider $\underline{n} = \begin{pmatrix}n_1\\n_2\end{pmatrix} = \begin{pmatrix}-\sin\frac{\theta}{2}\\ \cos\frac{\theta}{2}\end{pmatrix}$

Previously found H for reflection in line with normal n had entries $H_{ij} = \delta_{ij} - 2n_i n_j$

$$H = \begin{pmatrix}1-2\sin^2\frac{\theta}{2} & 2\sin\frac{\theta}{2}\cos\frac{\theta}{2}\\ 2\sin\frac{\theta}{2}\cos\frac{\theta}{2} & 1-2\cos^2\frac{\theta}{2}\end{pmatrix}$$ agrees with H above.



Note that cases $U = R$ and $U = H$ are distinguished by $\det R = +1$, $\det H = -1$

Unitary Matrices

A complex $n \times n$ matrix U is unitary iff

$$UU^\dagger = U^\dagger U = I \qquad (U^{-1} = U^\dagger)$$

Equivalent: U is unitary iff it preserves complex inner product on $\mathbb{C}^n$:

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

To check equivalence

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

Orthogonal and unitary matrices form groups

See SU_3 group (special unitary group)

should be $SU(3)$

5. Determinants and Inverses

5.1 Introduction

Consider a linear map $T: \mathbb{R}^n \to \mathbb{R}^n$

If T is invertible then $\ker T = \{0\}$ (T one to one)

and $\operatorname{Im} T = \mathbb{R}^n$ (T is onto)

These conditions are equivalent (rank-nullity theorem)

Conversely if the conditions hold, then

$T(\underline{e}_1), T(\underline{e}_2), \ldots, T(\underline{e}_n)$ must be a basis

(here $\{e_i\}$ = standard basis) and we can define

T^{-1} as a linear map by $T^{-1}(T(\underline{e}_i)) = \underline{e}_i$

How can we test whether the above conditions hold from the matrix M representing T : $T(\underline{x}) = M\underline{x}$

and how can we find M^{-1} from M explicitly?

For <u>any</u> M ($n \times n$) we define a related matrix $\tilde{M}$ ($n \times n$)

and a scalar $\det M$ such that $\tilde{M}M = (\det M)I$. $(*)$

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

For <u>$n=2$</u> recall that $(*)$ holds with $M = \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix}$

and $\tilde{M} = \begin{pmatrix} M_{22} & -M_{21} \\ -M_{12} & M_{11} \end{pmatrix}$ and $\det M = [M\underline{e}_1, M\underline{e}_2]$

$= \varepsilon_{ij} M_{i1} M_{j2}$

(scalar cross product)

This is the factor by which areas scale under M.

$\det M \neq 0 \Leftrightarrow M\underline{e}_1, M\underline{e}_2$ linearly independent.

For $\underline{n = 3}$ consider similarly

$$[M\underline{e}_1, M\underline{e}_2, M\underline{e}_3] \qquad \text{scalar triple product}$$

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

$$= \det M, \quad \text{define in } 3\times 3 \text{ case}$$

This is factor by which volumes scale under M, and

$\det M \neq 0 \iff M\underline{e}_1, M\underline{e}_2, M\underline{e}_3$ linearly independent

or $\mathrm{Im}\, M = \mathbb{R}^3$

Now $\underline{\text{define}}$* $\tilde{M}$ from M using row / column notation

$$\underline{R}_1(\tilde{M}) = \underline{C}_2(M) \times \underline{C}_3(M)$$

$$\underline{R}_2(\tilde{M}) = \underline{C}_3(M) \times \underline{C}_1(M)$$

$$\underline{R}_3(\tilde{M}) = \underline{C}_1(M) \times \underline{C}_2(M)$$

* and check it works.

Note that $(\tilde{M}M)_{ij} = R_i(\tilde{M}) \cdot \underline{C}_j(M)$

$$= \underbrace{(\underline{C}_1(M) \times \underline{C}_2(M) \cdot \underline{C}_3(M))}_{\det M} \delta_{ij} = \det M\, \delta_{ij}$$

as claimed.

<u>Example</u> $\quad M = \begin{pmatrix} 1 & 3 & 0 \\ 0 & -1 & 2 \\ 4 & 1 & -1 \end{pmatrix}, \quad \tilde{M} = \cancel{\begin{pmatrix} -1 & 8 & 4 \\ 3 & -1 & \\ 6 & -2 & \end{pmatrix}} \begin{pmatrix} -1 & 3 & 6 \\ 8 & -1 & -2 \\ 4 & -11 & -1 \end{pmatrix}$

$$\underline{C}_2 \times \underline{C}_3 = \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} \times \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \\ 6 \end{pmatrix}$$

$$\underline{C}_3 \times \underline{C}_1 = \begin{pmatrix} 0 \\ 2 \\ -1 \end{pmatrix} \times \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} = \begin{pmatrix} 8 \\ -1 \\ -2 \end{pmatrix}$$

$$\underline{C}_1 \times \underline{C}_2 = \begin{pmatrix} 1 \\ 0 \\ 4 \end{pmatrix} \times \begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4 \\ -11 \\ -1 \end{pmatrix}$$

$$\det M = \underline{C}_1 \cdot \underline{C}_2 \times \underline{C}_3 = 23$$

$$\tilde{M}M = 23 I$$

5.2 ε and Alternating Forms

Recall from IA Groups: a permutation σ on the set $\{1, 2, \ldots, n\}$ is a bijection from this set to itself specified by an ordered list, $\sigma(1), \sigma(2), \ldots, \sigma(n)$. Permutations form a group S_n ($|S_n| = n!$)

A transposition $\tau = (p, q)$ $(p \neq q)$ swaps p and q

Any permutation is a product of transpositions (k transpositions) $\sigma = \tau_k \ldots \tau_2 \tau_1$ with k always even or always odd for a given σ. Take $\varepsilon(\sigma) = (-1)^k$

The <u>alternating symbol</u> or ε symbol in $\mathbb{R}^n$ or $\mathbb{C}^n$ is an n-index object (tensor) defined by

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

Thus if σ is any permutation,

$$\varepsilon_{\sigma(1)\sigma(2)\ldots\sigma(n)} = \varepsilon(\sigma)$$

so ~~ε_{ijk}~~ $\varepsilon_{ij\ldots l}$ is <u>totally antisymmetric</u> and changes sign whenever we exchange a pair of indices.

Given $\underline{v}_1, \ldots, \underline{v}_2 \in \mathbb{R}^n$ or $\mathbb{C}^n$, the <u>alternating form</u> combines them to give scalar

$$[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n] = \varepsilon_{ij\ldots l} (\underline{v}_1)_i (\underline{v}_2)_j \ldots (\underline{v}_n)_l$$

$$= \sum_\sigma \varepsilon(\sigma) (\underline{v}_1)_{\sigma(1)} (\underline{v}_2)_{\sigma(2)} \ldots (\underline{v}_n)_{\sigma(n)}$$

($\sum_\sigma$ is over $\sigma \in S_n$)

Properties

(i) Multilinear $\quad [\underline{v}_1, \dots, \underline{v}_{p-1}, \alpha \underline{u} + \beta \underline{w}, \underline{v}_{p+1}, \dots, \underline{v}_n]$

$$= \alpha [\underline{v}_1, \dots, \underline{v}_{p-1}, \underline{u}, \underline{v}_{p+1}, \dots, \underline{v}_n]$$
$$+ \beta [\underline{v}_1, \dots, \underline{v}_{p-1}, \underline{w}, \underline{v}_{p+1}, \dots, \underline{v}_n]$$

(ii) Totally antisymmetric

$$[\underline{v}_{\sigma(1)}, \underline{v}_{\sigma(2)}, \dots, \underline{v}_{\sigma(n)}] = \varepsilon(\sigma)[\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n]$$

(iii) $[\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n] = +1$

These three properties fix alternating form completely and also imply

(iv) If $\underline{v}_p = \underline{v}_q$ then $[\underline{v}_1, \dots, \underline{v}_p, \dots, \underline{v}_q, \dots, \underline{v}_n] = 0$.

from (ii) by exchanging $\underline{v}_p, \underline{v}_q$

(v) If $\underline{v}_p = \sum_{i \neq p} \lambda_i \underline{v}_i$ then

$$[\underline{v}_1, \dots, \underline{v}_p, \dots, \underline{v}_n] = 0.$$

(since subbing $\underline{v}_p$ in and using multilinearity - use (iv))

Vectors and Matrices – Lecture 14 9/11/20

Consider (ii) from Lecture 13:

$$[\underline{v}_{\sigma(1)}, \underline{v}_{\sigma(2)}, \ldots, \underline{v}_{\sigma(n)}] = \varepsilon(\sigma)\,[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n]$$

To justify, check for a transposition $\tau = (p, q)$ $\quad p < q$

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

$$= \sum_\sigma \varepsilon(\sigma)(\underline{v}_1)_{\sigma(1)} \cdots (\underline{v}_{p-1})_{\sigma(p-1)}(\underline{v}_q)_{\sigma(p)} \cdots (\underline{v}_{q-1})_{\sigma(q-1)} (\underline{v}_p)_{\sigma(q)} \cdots (\underline{v}_n)_{\sigma(n)}$$

$$= \sum_\sigma \varepsilon(\sigma)(\underline{v}_1)_{\sigma'(1)} \cdots (\underline{v}_{p-1})_{\sigma'(p-1)}(\underline{v}_q)_{\sigma'(q)} \cdots (\underline{v}_{q-1})_{\sigma'(q-1)} (\underline{v}_p)_{\sigma'(p)} \cdots (\underline{v}_n)_{\sigma'(n)}$$

with $\sigma' = \sigma(p, q) = \sigma\tau$

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

so is totally antisymmetric as required.

(because $\varepsilon(\sigma') = -\varepsilon(\sigma)$)

and $\sum_\sigma$ equiv. to $\sum_{\sigma'}$.

__Proposition__ $[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n] \neq 0 \iff \underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ are lin. independent.

__Proof__ To show "$\Rightarrow$" use property (v):

If $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ lin. dep, then $\underline{v}_p$ can be expressed in terms of the others and then $[\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n] = 0$.

To show $\Leftarrow$: $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ lin. indep $\Rightarrow$ they span, then $\underline{e}_i = U_{ai}\,\underline{v}_a$ for some U_{ai} in $\mathbb{R}$ or $\mathbb{C}$.

But then

$$[\underline{e}_1, \dots, \underline{e}_n] = [U_{a1}\underline{v}_a, U_{b2}\underline{v}_b, \dots, U_{cn}\underline{v}_c]$$

$$= U_{a1}U_{b2}\dots U_{cn}\,[\underline{v}_a, \underline{v}_b, \dots, \underline{v}_c] \quad \text{as multilinear}$$

$$= U_{a1}U_{b2}\dots U_{cn}\,\varepsilon_{ab\dots c}\,[\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n].$$

But LHS $= 1$ so $[\underline{v}_1, \underline{v}_2, \dots, \underline{v}_n] \neq 0$.

<u>Example</u> $\underline{v}_1 = \begin{pmatrix} i \\ 0 \\ 0 \\ 2 \end{pmatrix}$, $\underline{v}_2 = \begin{pmatrix} 0 \\ 0 \\ 5i \\ 0 \end{pmatrix}$, $\underline{v}_3 = \begin{pmatrix} 3 \\ 2i \\ 0 \\ 0 \end{pmatrix}$,

$\underline{v}_4 = \begin{pmatrix} 0 \\ 0 \\ -i \\ 1 \end{pmatrix}$ in $\mathbb{C}^4$.

$$[\underline{v}_1, \underline{v}_2, \underline{v}_3, \underline{v}_4] = 5i\,[\underline{v}_1, \underline{e}_3, \underline{v}_3, \underline{v}_4]$$

$$= 5i\,[i\underline{e}_1 + \cancel{2\underline{e}_4}, \underline{e}_3, \cancel{3\underline{e}_1} + 2i\underline{e}_2, \cancel{-i\underline{e}_3} + \underline{e}_4]$$

$$= 5i\,[i\underline{e}_1, \underline{e}_3, 2i\underline{e}_2, \underline{e}_4)$$

$$= 5i(i)(2i)(-1)\,[\underline{e}_1, \underline{e}_2, \underline{e}_3, \underline{e}_4]$$

$$= \underline{10i}$$

<u>5.3</u> Determinants in $\mathbb{R}^n$ and $\mathbb{C}^n$

(a) <u>Definitions</u>

For an $n \times n$ matrix M with columns $\underline{C}_a = M\underline{e}_a$, the determinant <u>$\det M$</u> or $|M| \in \mathbb{R}$ or $\mathbb{C}$, is defined by

$$\det M = [\underline{c}_1, \underline{c}_2, \dots, \underline{c}_n]$$

$$= [M\underline{e}_1, M\underline{e}_2, \dots, M\underline{e}_n]$$

$$= \varepsilon_{ij\dots l}\, M_{i1} M_{j2} \dots M_{ln}$$

$$= \sum_{\sigma} \varepsilon(\sigma)\, M_{\sigma(1)1} M_{\sigma(2)2} \dots M_{\sigma(n)n}.$$

Each of these expressions can be taken as the definition.

Examples

(i) $n = 2$ $\quad \det M = \sum_{\sigma} \varepsilon(\sigma) M_{\sigma(1)1} M_{\sigma(2)2}$

$$= M_{11} M_{22} - M_{21} M_{12}$$

(ii) M diagonal: $M_{ij} = 0$ for $i \neq j$ i.e. $M = \begin{pmatrix} M_{11} & & 0 \\ & M_{22} & \\ 0 & \cdots & M_{nn} \end{pmatrix}$

$$\Rightarrow \det M = M_{11} M_{22} \ldots M_{nn}$$

(iii) $\underset{(n\times n)}{M} = \left(\begin{array}{c|c} A & 0 \\ \hline 0 & 1 \end{array}\right)$ where A is $n-1 \times n-1$

so $M_{ni} = 0$ if $i \neq n$
$= M_{in}$.

can restrict perm σ in definition to have $\sigma(n) = n$

then $\det M = \det A$

Proposition If R_a are rows of M, then

$$\det M = [\underline{R}_1, \underline{R}_2, \ldots, \underline{R}_n]$$

$$= \varepsilon_{ij\ldots l} M_{1i} M_{2j} \ldots M_{nl}$$

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

i.e. $\underline{\det M = \det M^T}$.

(Recall $(\underline{C}_a)_i = M_{ia} = (\underline{R}_i)_a$)

Proof Show directly that $\sum_{\sigma}$ definitions agree by considering

$M_{\sigma(1)1} \ldots M_{\sigma(n)n}$

$= M_{1\rho(1)} \ldots M_{n\rho(n)}$ for $\rho = \sigma^{-1}$.

But $\varepsilon(\sigma) = \varepsilon(\rho) = \varepsilon(\sigma^{-1})$ so $\sum\limits_{\sigma}$ is $\sum\limits_{\rho}$, as required. $\square$

(b) <u>Evaluating Determinants</u>: Expanding Rows or Columns

For $M_{n\times n}$ with entries M_{ia}, define <u>minor</u> M^{ia} to be the $(n-1)\times(n-1)$ determinant of matrix obtained by deleting row i and column a from M.

(*) <u>Proposition</u>

$$\det M = \sum_i (-1)^{i+a} M_{ia} M^{ia} \quad (a \text{ fixed})$$

$$= \sum_a (-1)^{i+a} M_{ia} M^{ia} \quad (i \text{ fixed})$$

<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 by row 3:

$$M_{32} = 5i, \quad M^{32} = \begin{vmatrix} i & 3 & 0 \\ 0 & 2i & 0 \\ 2 & 0 & 1 \end{vmatrix}$$

Expand M^{32} by row 1: $= i\begin{vmatrix} 2i & 0 \\ 0 & 1 \end{vmatrix} - 3\begin{vmatrix} 0 & 0 \\ 2 & 1 \end{vmatrix} = -2$

and $M_{34} = -i$, $M^{34} = \begin{vmatrix} i & 0 & 3 \\ 0 & 0 & 2i \\ 2 & 0 & 0 \end{vmatrix} = 0$

$$\text{so } \det M = (-1)^5 M_{32} M^{32} + (-1)^7 M_{34} M^{34}$$

$$= -5i(-2) = \underline{10i}.$$

(*) proposition will be proved in next lecture.

(c) Simplifying Determinants : Row and Column Operations

Summary det M is a function of the rows $\underline{R}_i$ or the columns C_a of M.

(i) multilinear (ii) totally antisymmetric

(iii) $\det I = 1$

Consider the following consequences :

Row and Column Scalings

If $\underline{R}_i \to \lambda \underline{R}_i$ (for i fixed)

xor $\underline{C}_a \to \lambda \underline{C}_a$ (for a fixed)

$\det M \to \lambda \det M$ by multilinearity

If we scale all rows / cols then $M \to \lambda M$

so $\det M \to \lambda^n \det M$.

Row and Column Operations

If $\underline{R}_i \to \underline{R}_i + \lambda \underline{R}_j$ for $i \neq j$

or $\underline{C}_a \to \underline{C}_a + \lambda \underline{C}_b$ for $a \neq b$

then $\det M \to \det M$ (doesn't change)

Row and Column Exchanges

If $\underline{R}_i \leftrightarrow \underline{R}_j$ then $\det M \to -\det M$

or $\underline{C}_a \leftrightarrow \underline{C}_b$ by alternating property (as there's a transposition)

Example $A = \begin{pmatrix} 1 & 1 & a \\ a & 1 & 1 \\ 1 & a & 1 \end{pmatrix}$ $a \in \mathbb{C}$

$\underline{c}_1 \to \underline{c}_1 - \underline{c}_3$: $\det A = \det \begin{pmatrix} 1-a & 1 & a \\ a-1 & 1 & 1 \\ 0 & a & 1 \end{pmatrix}$

$= (1-a) \det \begin{pmatrix} 1 & 1 & a \\ -1 & 1 & 1 \\ 0 & a & 1 \end{pmatrix}$ by linearity property (see row/column scalings)

$\underline{c}_2 \to \underline{c}_2 - \underline{c}_3$: $(1-a) \det \begin{pmatrix} 1 & 1-a & a \\ -1 & 0 & 1 \\ 0 & a-1 & 1 \end{pmatrix}$

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

$\underline{R}_1 \to \underline{R}_1 + \underline{R}_2 + \underline{R}_3$: $(1-a)^2 \det \begin{pmatrix} 0 & 0 & a+2 \\ -1 & 0 & 1 \\ 0 & -1 & 1 \end{pmatrix}$

$= \underline{(1-a)^2 \, (a+2)}$

(d) **Multiplicative Property**

Theorem For $n \times n$ matrices M, N:

$$\det(MN) = \det M \det N.$$

Can prove by the following elaboration on def. of determinant

Lemma $\varepsilon_{i_1 i_2 \dots i_n} M_{i_1 a_1} M_{i_2 a_2} \dots M_{i_n a_n} = (\det M)\, \varepsilon_{a_1 a_2 \dots a_n}$

Proof of Theorem $\det(MN) = \varepsilon_{i_1 \dots i_n} (MN)_{i_1 1} \dots (MN)_{i_n n}$

$= \varepsilon_{i_1 \dots i_n} M_{i_1 k_1} \dots M_{i_n k_n}$ (matrix $\times$ def^n)
$\times N_{k_1 1} \dots N_{k_n n}$

$= (\det M)\, \varepsilon_{k_1 \dots k_n} N_{k_1 1} \dots N_{k_n n}$ by lemma

$= \underline{(\det M)(\det N)}$ $\square$

Proof of Lemma

LHS and RHS are each totally antisymmetric in $a_1, a_2, \ldots, a_n$ and so must be related by a constant factor. To fix constant, consider $a_1 = 1,\ a_2 = 2, \ldots, a_n = n$ and check it works, then result follows.

Consequences of multiplicative property

(i) $M^{-1} M = I \Rightarrow \det M^{-1} \det M = 1$

(ii) For R real and orthogonal, $R^T R = I \Rightarrow \det R^T \det R = 1$
but $\det R^T = \det R$ so $(\det R)^2 = 1 \Rightarrow \det R = \pm 1$

(iii) For U complex and unitary
$$U^\dagger U = I \Rightarrow \det(U^\dagger) \det U = 1$$
$$\Rightarrow (\overline{\det U}) \det U = 1 \qquad (\text{as } U^\dagger = (\overline{U})^T)$$
$$\Rightarrow |\det U|^2 = 1 \Rightarrow |\det U| = 1$$

5.4 Minors, Cofactors and Inverses

(a) Cofactors and Determinants

Consider a column of matrix M ($n \times n$) and write it
$$\underline{c}_a = \sum_i M_{ia} \underline{e}_i$$

$$\Rightarrow \det M = [\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_a, \ldots, \underline{c}_n]$$
$$= [\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{a-1}, \sum_i M_{ia} \underline{e}_i, \underline{c}_{a+1}, \ldots, \underline{c}_n]$$
$$= \sum_i M_{ia} \Delta_{ia} \quad \text{where } \Delta_{ia} = [\underline{c}_1, \underline{c}_2, \ldots, \underline{c}_{a-1}, \underline{e}_i, \underline{c}_{a+1}, \ldots, \underline{c}_n]$$

PTO

$$= \det \begin{array}{cc} \begin{pmatrix} A & \begin{matrix}0\\ \vdots\\ 0\end{matrix} & B \\ 0 \cdots\cdots 0 & 1 & 0 \cdots 0 \\ C & \begin{matrix}0\\ \vdots\\ 0\end{matrix} & D \end{pmatrix} & \leftarrow \text{row } i \\ \uparrow \\ \text{col } a & \end{array}$$

where additional zero entries arise from antisymmetry

$$= (-1)^{n-a}(-1)^{n-i} \det \begin{pmatrix} A & B & 0 \\ & & \vdots \\ C & D & 0 \\ 0 \cdots & 0 & 1 \end{pmatrix} = (-1)^{i+a} M^{ia}$$

where $M^{ia} = \det \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ – minor as introduced previously.

We deduce $\det M = \sum_i M_{ia} \Delta_{ia} = \sum_i (-1)^{i+a} M_{ia} M^{ia}$

Here Δ_{ia} is called the cofactor.

Similarly by considering rows:

$$\det M = \sum_a M_{ia} \Delta_{ia} = \sum_a (-1)^{i+a} M_{ia} M^{ia}$$

(b) <u>Adjugates and Inverses</u>

Consider $\underline{c}_b = \sum_i M_{ib} \underline{e}_i$

Then $[\underline{c}_1, \ldots, \underline{c}_{a-1}, \underline{c}_b, \underline{c}_{a+1}, \ldots, \underline{c}_n]$

$$= \sum_i M_{ib} \Delta_{ia} = \begin{cases} \det M & \text{if } a = b \\ 0 & \text{if } a \neq b \end{cases} = (\det M)\, \delta_{ab}$$

Similarly, expanding a row

$$\sum_a M_{ja} \Delta_{ia} = (\det M) \delta_{ij}.$$

Let Δ be the matrix of cofactors (entries Δ_{ia}) and define the <u>adjugate</u> $\tilde{M} = \operatorname{adj} M = \Delta^T$. Then

$$\Delta_{ia} M_{ib} = (\Delta^T)_{ai} M_{ib} = (\Delta^T M)_{ab} = (\det M)\delta_{ab}$$

$$\text{i.e. } \tilde{M} M = (\det M) I.$$

Similarly $M_{ja} \Delta_{ia} = M_{ja} (\Delta^T)_{ai} = (M \Delta^T)_{ji} = (\det M) \delta_{ij}$

$$\text{so } \underline{M \tilde{M} = (\det M) I}$$

Hence if $\det M \neq 0$ then $M^{-1} = \dfrac{1}{\det M} \tilde{M}$

<u>Example</u> $A = \begin{pmatrix} 1 & 1 & a \\ a & 1 & 1 \\ 1 & a & 1 \end{pmatrix}$: previously found $\det A = (1-a)^2 (a+2)$

Find cofactors: consider e.g. $A^{12} = \begin{vmatrix} a & 1 \\ 1 & 1 \end{vmatrix} = a - 1$

(det. of A without row 1 and col 2)

$$\Delta^{12} = (-1)^{1+2} A^{12} = 1 - a$$

Then we end up with matrix of cofactors

$$\Delta = \begin{pmatrix} 1-a & 1-a & a^2-1 \\ a^2-1 & 1-a & 1-a \\ 1-a & a^2-1 & 1-a \end{pmatrix}.$$

If $a \neq 1$ or -2 then $\det A \neq 0$ so can find inverse.

We end up with $A^{-1} = \begin{pmatrix} 1 & -(1+a) & 1 \\ 1 & 1 & -(1+a) \\ -(1+a) & 1 & 1 \end{pmatrix} \cdot \dfrac{1}{(1-a)(a+2)}.$

Vectors and Matrices - Lecture 16 13/11/20

5.5 Systems of Linear Equations

a) Introduction and Nature of Solutions

Consider a system of n linear eqns in n unknowns x_i written in matrix vector form:

$A\underline{x} = \underline{b}$ $\underline{x}, \underline{b} \in \mathbb{R}^n$, A $n \times n$ matrix

i.e.
$$A_{11}x_1 + \dots + A_{1n}x_n = b_1$$
$$A_{21}x_1 + \dots + A_{2n}x_n = b_2$$
$$\dots$$
$$A_{n1}x_1 + \dots + A_{nn}x_n = b_n$$

There are three possibilities.

(i) If $\det A \neq 0$ then A^{-1} exists: unique solution $\underline{x} = A^{-1}b$.

(ii) $\det A = 0$ and $\underline{b} \notin \text{Im } A \Rightarrow$ no soln by definition

(iii) $\det A = 0$ and $b \in \text{Im } A \Rightarrow \infty$ solutions.
Have form $\underline{x} = \underline{x}_0 + \underline{u}$ with $\underline{x}_0$ a particular solution and $\underline{u} \in \ker A$.

Elaboration: A solution exists iff $A\underline{x}_0 = \underline{b}$ for some x_0.
Then $\underline{x}$ is also a solution iff $\underline{u} = \underline{x} - \underline{x}_0$ satisfies $A\underline{u} = \underline{0}$ $(\underline{u} \in \ker A)$

$A\underline{u} = \underline{0}$: corresponding homogeneous problem.

Now $\det A \neq 0 \iff \text{Im } A = \mathbb{R}^n \iff \ker A = \{\underline{0}\}$ by Rank-Nullity.
So in case (i) there is always a unique solution.
$\det A = 0 \iff \text{rank}(A) < n \iff \text{null}(A) > 0$

Then either $\underline{b} \notin \operatorname{Im} A$ (ii) or $\underline{b} \in \operatorname{Im} A$ (iii)

If $\underline{u}_1, \dots, \underline{u}_k$ is a basis for $\ker A$, then the general solution to the homogeneous problem is

$$\underline{u} = \sum_{i=1}^{k} \lambda_i \underline{u}_i \qquad (k = \operatorname{null} A)$$

[Compare DEs, on functions $y(x)$, consider linear differential operator $L = p_2(x) \frac{d^2}{dx^2} + p_1(x) \frac{d}{dx} + p_0(x)$]

$Ly = f$ has particular solution / integral $y_0(x)$ with $Ly_0 = f$

then general sol^n is $y = y_0 + \underbrace{(\lambda_1 y_1 + \lambda_2 y_2)}$

general soln to homogeneous problem $Ly = 0$

<u>Example</u> $A\underline{x} = \underline{b}$, $A = \begin{pmatrix} 1 & 1 & a \\ a & 1 & 1 \\ 1 & a & 1 \end{pmatrix}$, $\underline{b} = \begin{pmatrix} 1 \\ c \\ 1 \end{pmatrix}$

Previously found $\det A = (a-1)^2(a+2)$

- $a \neq 1, -2$

  $\det A \neq 0$ and A^{-1} exists; previously found

$$A^{-1} = \frac{1}{(a-1)^2(a+2)} \begin{pmatrix} 1 & -(1+a) & 1 \\ 1 & 1 & -(1+a) \\ -(1+a) & 1 & 1 \end{pmatrix}$$

For $a \neq 1$, $a \neq -2$ there is a unique solution for any c

$$\underline{x} = A^{-1}\underline{b} = \frac{1}{(1-a)(a+2)} \begin{pmatrix} 2-c-ca \\ c-a \\ c-a \end{pmatrix}$$

Geometrically, solution is a point.

~~An~~ If $a = 1$ then $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$ $\Rightarrow \operatorname{Im} A = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$

$\operatorname{Ker} A = \operatorname{span} \left\{ \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}$

$b \in \operatorname{Im} A \iff c = 1$, so $\underline{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$ is a particular solution.

$$\rightarrow \quad \underline{x} = \underline{x}_0 + \underline{u} = \begin{pmatrix} 1-\lambda-\mu \\ \lambda \\ \mu \end{pmatrix}$$

i.e. for $a=1$, $c=1$ have case (iii), geometrically a plane

For $a=1$, $c \neq 1$ have no solutions. (case (ii))

<u>$a = -2$</u> $\quad A = \begin{pmatrix} 1 & 1 & -2 \\ -2 & 1 & 1 \\ 1 & -2 & 1 \end{pmatrix} \Rightarrow \operatorname{Im} A = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \right\}$

$\ker A = \left\{ \lambda \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ by inspection

<u>Claim</u> $\quad b \in \operatorname{Im} A \iff c = -2$; particular solution is $\underline{x}_0 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$

General solution $\underline{x} = \underline{x}_0 + \underline{u} = \begin{pmatrix} 1+\lambda \\ \lambda \\ \lambda \end{pmatrix}$

For $a = -2$, $c = -2$ we have (iii) (a line)

$a = -2$, $c \neq -2$: no solutions (ii)

b) Geometrical interpretation in $\mathbb{R}^3$

Let $\underline{R}_1, \underline{R}_2, \underline{R}_3$ be the rows of A (3×3)

$A\underline{u} = \underline{0} \iff \underline{R}_1 \cdot \underline{u} = 0, \ \underline{R}_2 \cdot \underline{u} = 0, \ \underline{R}_3 \cdot \underline{u} = 0$

Each are <u>planes</u> through $\underline{0}$ with normals $\underline{R}_i$.

So solution of homogeneous problem (finding $\ker A$) is given by intersection of these planes.

- $\dim(\operatorname{im} A) = \operatorname{rank} A = 3 \Rightarrow$ normals lin. indep, planes intersect only at $\underline{0}$.
- $\operatorname{rank} A = 2 \Rightarrow$ normals span plane, planes above intersect in a line

rank $A = 1 \Rightarrow$ normals all parallel and hence planes coincide  $\dim \ker A = 2$.

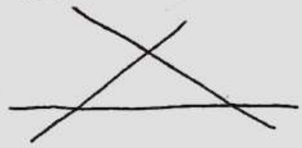
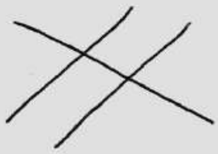
Now consider instead $A\underline{x} = \underline{b} \iff \underline{R}_1 \cdot \underline{x} = b_1,\ \underline{R}_2 \cdot \underline{x} = b_2,\ \underline{R}_3 \cdot \underline{x} = b_3$

planes with normals $\underline{R}_i$ but not through $\underline{0}$ unless $\underline{b}_i = \underline{0}$.

rank $A = 3 \iff \det A \neq 0 \iff$ normals lin. indep, planes intersect at a point: unique solution

rank $A < 3 \iff \det A = 0$; existence of soln depends on $\underline{b}$

rank $A = 2 \Rightarrow$ planes may intersect in a line (as before) but may not

e.g.  or 

(no solution)

rank $A = 1 \Rightarrow$ planes may coincide (as in homogeneous case) but may not, depending on $\underline{b}$

e.g.  or  (no solution).

Vectors and Matrices - Lecture 17 16/11/20

6. Eigenvalues and Eigenvectors

6.1 Introduction

(a) Definitions

For a linear map $T: V \to V$, a vector $\underline{v}$ with $\underline{v} \neq 0$ is an <u>eigenvector</u> of T with <u>eigenvalue</u> λ if $T(\underline{v}) = \lambda \underline{v}$.

If $V = \mathbb{R}^n$ or $\mathbb{C}^n$ and T is given by an $n \times n$ matrix A then

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

So given λ, this holds for some $\underline{v} \neq 0$ iff $\det(A - \lambda I) = 0$, called the characteristic equation for A.

i.e. λ is an eigenvalue iff it is a root of

$$\chi_A(t) = \det(A - tI) \qquad \text{characteristic polynomial}$$

$$= \det \begin{pmatrix} A_{11}-t & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22}-t & \cdots & A_{2n} \\ \cdots & & & \\ A_{n1} & A_{n2} & \cdots & A_{nn}-t \end{pmatrix}$$

We can find eigenvalues as roots of characteristic eqn, then determine corresponding eigenvectors.

b) <u>Examples</u>

(i) $V = \mathbb{C}^2$ $\quad A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$, $\quad \det(A - \lambda I) = \begin{vmatrix} 2-\lambda & i \\ -i & 2-\lambda \end{vmatrix}$

$= (2-\lambda)^2 - 1 = 0$ iff $\lambda = 2 \pm 1 \iff \lambda = 1 \text{ or } 3$

$\underline{\lambda = 1}$: find eigenvectors $\underline{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$

$$(A - I)\underline{v} = \begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0} \Rightarrow \underline{v} = \alpha \begin{pmatrix} 1 \\ i \end{pmatrix}$$

for $\alpha \neq 0$.

$\underline{\lambda = 3}$ $\quad (A - 3I)\underline{v} = \begin{pmatrix} -1 & i \\ -i & -1 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0}$

$$\Rightarrow \quad \underline{v} = \beta \begin{pmatrix} 1 \\ -i \end{pmatrix} \quad \text{for any } \beta \neq 0.$$

(ii) $V = \mathbb{R}^2$ $\quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\quad \det(A - \lambda I) = \begin{vmatrix} 1-\lambda & 1 \\ 0 & 1-\lambda \end{vmatrix}$

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

so $\lambda = 1$ repeated root

$\lambda = 1:$ $\quad (A - I)\underline{v} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \underline{0} \iff \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

for $\alpha \neq 0$.

Only one (lin. indep) eigenvector.

(iii) $V = \mathbb{R}^2$ or $\mathbb{C}^2$ $\quad U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$

$$\chi_U(t) = \det(U - tI) = t^2 - 2t\cos\theta + 1$$

Eigenvalues $\quad \lambda = e^{\pm i\theta}$

Eigenvectors $\quad \underline{v} = \alpha \begin{pmatrix} 1 \\ \mp i \end{pmatrix} \quad \alpha \neq 0 \quad$ (check)

No real EVs unless $\theta = n\pi$.

(iv) $V = \mathbb{C}^n$ $\quad A = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ diagonal

$$\chi_A(t) = (\lambda_1 - t)(\lambda_2 - t) \cdots (\lambda_n - t)$$

e. vals $\lambda = \lambda_i$ $\quad$ e. vecs $\underline{v} = \alpha \underline{e}_i$ $\quad (\alpha \neq 0)$

for each i.

(c) Deductions involving $\chi_A(t)$

For A $n \times n$, char. polynomial has degree n

$$\chi_A(t) = \sum_{j=0}^{n} c_j t^j = (-1)^n (t-\lambda_1) \dots (t-\lambda_n)$$

(i) $\exists$ at least one root of χ_A (FTA) in fact $\exists\, n$ roots.

(ii) $tr(A) = A_{ii} = \sum_{i=1}^{n} \lambda_i$ sum of eigenvalues

Compare terms of degree $n-1$ in t

and from det we get $(-t)^{n-1} A_{11} + (-t)^{n-1} A_{22} + \dots + (-t)^{n-1} A_{nn}$

Overall sign matches with expansion of

$(-1)^n (t-\lambda_1)(t-\lambda_2) \dots (t-\lambda_n)$

(iii) $\det A = \chi_A(0) = \prod_{i=1}^{n} \lambda_i$ product of eigenvalues

(iv) If A is real, coeffs c_i are real and

$$\chi_A(\lambda) = 0 \iff \chi_A(\bar{\lambda}) = 0$$

non-real roots are in conjugate pairs.

<u>6-2</u> Eigenspaces and Multiplicities

(a) <u>Definitions</u>

For an eigenvalue λ of a matrix A, define the <u>eigenspace</u>

$E_\lambda = \{ \underline{v} : A\underline{v} = \lambda \underline{v} \} = Ker(A - \lambda I)$

Not exactly set of all eigenvectors, but all <u>nonzero</u> $\underline{v} \in E_\lambda$ are the eigenvectors.

<u>Geometric multiplicity</u> $m_\lambda = \dim E_\lambda = \#$ linearly indep. e.vecs with e.val λ

$= null(A - \lambda I)$

The <u>algebraic multiplicity</u> is M_λ, multiplicity* of λ as a root of $\chi_A(t)$ i.e. $\chi_A(t) = (t-\lambda)^{M_\lambda} f(t)$ with $f(\lambda) \neq 0$

<u>Proposition</u> $M_\lambda \geqslant m_\lambda$ (and $m_\lambda \geqslant 1$ since λ a root of $\chi_A(t)$)

<u>Proof</u> see section 6.3.

* how many times it's repeated as a root

(b) <u>Examples</u>

(i) $A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix} \Rightarrow \chi_A(t) = \det(A - tI)$

$= (5-t)(t+3)^2$ (check)

roots $\lambda = 5, -3$ (twice) $\qquad M_5 = 1, \quad M_{-3} = 2$

Find eigenspaces

<u>$\lambda = 5$</u> : $E_5 = \left\{ \alpha \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} \right\}$

<u>$\lambda = -3$</u> : $E_3 = \left\{ \alpha \begin{pmatrix} -2 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 3 \\ 0 \\ 1 \end{pmatrix} \right\}$

e.g. $(A + 3I)\underline{x} = \begin{pmatrix} 1 & 2 & -3 \\ 2 & 4 & -6 \\ -1 & -2 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underline{0}$

$\Rightarrow \quad \underline{x} = \begin{pmatrix} -2x_2 + 3x_3 \\ x_2 \\ x_3 \end{pmatrix}$

dim E_5 is $m_5 = 1 = M_5$

dim E_{-3} is $m_{-3} = 2 = M_{-3}$

(ii) $$A = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix} \Rightarrow \chi_A(t) = \det(A - tI)$$
$$= -(t+2)^3$$

root $\lambda = -2$ (3 times) $\quad M_{-2} = 3$

Find eigenspace: $$(A + 2I)\underline{x} = \begin{pmatrix} -1 & -1 & 1 \\ -1 & -1 & 1 \\ -2 & -2 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underline{0}$$

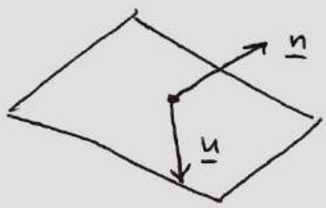
so is $$\underline{x} = \begin{pmatrix} -x_2 + x_3 \\ x_2 \\ x_3 \end{pmatrix}$$

$$E_{-2} = \left\{ \alpha \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \beta \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

$\dim E_{-2}$ is $m_{-2} = 2$
$< M_{-2} = 3$

(iii) Reflection in plane through $\underline{0}$ with unit normal $\underline{n}$

$H\underline{n} = -\underline{n}$



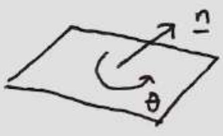
$H\underline{u} = \underline{u}$

for any $\underline{u} \perp \underline{n}$

Eigenvalues are ± 1 $\qquad E_{-1} = \{\alpha \underline{n}\}$ and $E_1 = \{\underline{x} : \underline{x}.\underline{n} = 0\}$

$M_{-1} = m_{-1} = 1 \qquad M_1 = m_1 = 2$

(iv) Rotation about axis $\underline{n}$ through θ (in $\mathbb{R}^3$)



$R\underline{n} = \underline{n}$ $\quad$ axis of rotation $\underline{n}$ is e.vec with e.val 1

No other real e.vals unless $\theta = n\pi$, rotation restricted to plane $\perp \underline{n}$ has e.vals $e^{\pm i\theta}$ as in 6.1 b) iii).

(c) Linear Independence of Eigenvectors

Proposition Let $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_r$ be eigenvectors of a matrix A $(n \times n)$ with eigenvalues $\lambda_1, \lambda_2, \dots, \lambda_r$. If eigenvalues are distinct, then $\underline{v}_i$ are linearly independent.

Proof Note that $\underline{w} = \sum_{j=1}^{r} \alpha_j \underline{v}_j \Rightarrow$

$$(A - \lambda I)\underline{w} = \sum_{j=1}^{r} \alpha_j (\lambda_j - \lambda) \underline{v}_j$$

Method 1 Suppose eigenvectors are linearly dependent. So $\exists$ a linear combination $\underline{w} = \underline{0}$ with number of nonzero coefficients $p \geqslant 2$. (eigenvectors nonzero: $p \neq 1$).
Pick such a $\underline{w}$ for which p is least, and WLOG take $\alpha_1 \neq 0$.

Then $(A - \lambda_1 I)\underline{w} = \sum_{j>1} \alpha_j (\lambda_j - \lambda_1)\underline{v}_j = \underline{0}$,
a linear relation with $p-1$ nonzero coeffs ※ (p was least).

~~Me~~ Method 2 Given a linear relation

$$\underline{w} = \underline{0} \Rightarrow \prod_{j \neq k} (A\ \lambda_j I)\underline{w} = \alpha_k \left(\prod_{j \neq k} (\lambda_k - \lambda_j) \right) \underline{v}_k$$

$$= \underline{0}$$

for k fixed

Eigenvalues are distinct so $\alpha_k = 0$.

i.e. eigenvectors are linearly indep. $\square$

Corollary With conditions as in proposition, let B_{λ_i} be a basis for the eigenspace E_{λ_i} $i = 1, 2, \ldots, r$

Then $B_{\lambda_1} \cup B_{\lambda_2} \cup \ldots \cup B_{\lambda_r}$ is linearly independent.

Proof Consider a general linear combination of all these vecs: has form $\underline{w} = \underline{w}_1 + \underline{w}_2 + \ldots + \underline{w}_r$ where $\underline{w}_i \in E_{\lambda_i}$ (each $\underline{w}_i$ is a linear combination of elements of B_{λ_i}).

Applying same arguments, we find $\underline{w} = \underline{0} \Rightarrow \underline{w}_i = 0$ for each i.

Then each $\underline{w}_i$ is the trivial linear combination of elements of B_{λ_i} and the result follows. $\square$

6.3 Diagonalisability and Similarity

(a) Introduction

Proposition: For an $n \times n$ matrix A, acting on $V = \mathbb{R}^n$ or $\mathbb{C}^n$, the following conditions are equivalent:

(i) There exists a basis of eigenvectors for V, $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ with $A\underline{v}_i = \lambda_i \underline{v}_i$ (not summation)

(ii) There exists an $n \times n$ invertible matrix P with
$$P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

(If either hold, then A is diagonalisable).

Proof Note that for any matrix P, AP has columns $A\underline{C}_i(P)$ and PD has columns $\lambda_i \underline{C}_i(P)$ (check)

Then (i) and (ii) are related by choosing $\underline{v}_i = \underline{C}_i(P)$:

$\underbrace{P^{-1}AP}_{=D} \Leftrightarrow AP = PD \Leftrightarrow A\underline{v}_i = \lambda_i \underline{v}_i$. $\square$

i.e. given an eigenvector basis as in (i), this relation defines P; conversely given a matrix P as in (ii), its columns are a basis of eigenvectors.

Examples (i) Refer to 6.1 b) ii) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$. Found $E_1 = \left\{\alpha \begin{pmatrix} 1 \\ 0 \end{pmatrix}\right\}$

A single eigenvalue $\lambda = 1$ with one lin. indep e.vec

No basis of e.vecs for $\mathbb{R}^2$ or $\mathbb{C}^2$: A not diagonalisable.

(ii) 6.1 b) iii) $U = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$ found e. vals $e^{\pm i\theta}$ distinct for $\theta \neq n\pi$

evecs $\begin{pmatrix} 1 \\ \mp i \end{pmatrix}$ lin. indep in $\mathbb{C}$

$$P = \begin{pmatrix} 1 & 1 \\ -i & i \end{pmatrix} \Rightarrow P^{-1} = \frac{1}{2}\begin{pmatrix} 1 & i \\ 1 & -i \end{pmatrix} \text{ and } P^{-1}UP = \begin{pmatrix} e^{i\theta} & 0 \\ 0 & e^{-i\theta} \end{pmatrix}.$$

So is diagonalisable over $\mathbb{C}$ (but not $\mathbb{R}$).

(b) Criteria for Diagonalisability

Proposition Consider $n \times n$ matrix A.

(i) A is diagonalisable if it has n distinct eigenvalues (sufficient)

(ii) A is diagonalisable iff for every eigenvalue λ the multiplicity coincides: $M_\lambda = m_\lambda$. (necessary and sufficient)

Proof Use proposition and corollary from (a)

(i) If n distinct eigenvalues, then n lin. indep evecs so they form a basis.

(ii) If λ_i with $i = 1, 2, \dots, r$ are all the distinct e.vals then $B_{\lambda_1} \cup \dots \cup B_{\lambda_r}$ is lin. indep but # elements is $\sum_i m_{\lambda_i}$ (dim. of each E_{λ_i}) $= \sum_i M_{\lambda_i} = n$ (degree of char. poly) where B_{λ_i} is a basis for E_{λ_i}. So we have a basis. $\square$

Note (i) is special case $M_\lambda = m_\lambda = 1 \ \forall \lambda$.

Examples Refer to 6.2(b)

(i) $A = \begin{pmatrix} -2 & 2 & -3 \\ 2 & 1 & -6 \\ -1 & -2 & 0 \end{pmatrix}$ $\lambda = 5, -3$

$M_5 = m_5 = 1, \quad M_{-3} = m_{-3} = 2$

so A is diagonalisable : $P = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$, $P^{-1} = \begin{pmatrix} 1 & 2 & -3 \\ -2 & 4 & 6 \\ 1 & 2 & 5 \end{pmatrix}$

$\Rightarrow P^{-1}AP = \begin{pmatrix} 5 & 0 & 0 \\ 0 & -3 & 0 \\ 0 & 0 & -3 \end{pmatrix}$

(ii) $A = \begin{pmatrix} -3 & -1 & 1 \\ -1 & -3 & 1 \\ -2 & -2 & 0 \end{pmatrix}$ $\lambda = -2, \quad M_{-2} = 3 > m_{-2} = 2$

$\Rightarrow$ A not diagonalisable

If it were: $P^{-1}AP = -2I \Rightarrow A = -2I$ ⁎

(c) Similarity

Matrices A and B ($n \times n$) are similar if $B = P^{-1}AP$ for some invertible P ($n \times n$), an equivalence relation.

Proposition If A and B are similar, then

(i) $\operatorname{tr} B = \operatorname{tr} A$ (ii) $\det B = \det A$

(iii) $\chi_B(t) = \chi_A(t)$

Proof (i) $\operatorname{tr} B = \operatorname{tr}(P^{-1}AP) = \operatorname{tr}(APP^{-1}) = \operatorname{tr} A$

(ii) $\det B = \det(P^{-1}AP) = \det P^{-1} \det P \det A = \det(PP^{-1}) \det A = \det A$

(iii) $\det(B - tI) = \det(P^{-1}AP - tI)$

$= \det(P^{-1}AP - tP^{-1}P)$

$= \det(P^{-1}(A - tI)P) = \det(A - tI)$

similarly to above.

6.4 Hermitian and Symmetric Matrices

(a) Real Eigenvalues and Orthogonal Eigenvectors

Recall: A $(n \times n)$ is Hermitian iff $A^\dagger = A$.

Special case: A is real and symmetric : $\bar{A} = A$, $A^T = A$.

Recall: complex inner product for $\underline{v}, \underline{w} \in \mathbb{C}^n$ is

$$\underline{v}^\dagger \underline{w} = \sum_i \bar{v}_i w_i$$

and for $\underline{v}, \underline{w} \in \mathbb{R}^n$ this reduces to $\underline{v}^T \underline{w} = \underline{v} \cdot \underline{w} = \sum_i v_i w_i$

<u>Observation</u> If A is hermitian then $(A\underline{v})^\dagger \underline{w} = \underline{v}^\dagger (A\underline{w})$ for all $\underline{v}, \underline{w} \in \mathbb{C}^n$ since $(\underline{v}^\dagger A^\dagger)\underline{w} = \underline{v}^\dagger A^\dagger \underline{w}$

$= \underline{v}^\dagger A \underline{w} = \underline{v}^\dagger (A\underline{w})$)

<u>Theorem</u> For a matrix A that is hermitian $(n \times n)$

(i) Every eigenvalue λ is real

(ii) Eigenvectors $\underline{v}, \underline{w}$ with distinct eigenvalues $\lambda \neq \mu$ are orthogonal. $(\underline{v}^\dagger \underline{w} = 0)$

(iii) If A is real and symmetric, then for each λ in (i) we can choose a real eigenvector $\underline{v}$ and (ii) becomes $\underline{v}^T \underline{w} = \underline{v} \cdot \underline{w} = 0$.

<u>Proof</u> (i) $\underline{v}^\dagger (A\underline{v}) = (A\underline{v})^\dagger \underline{v}$

$\Rightarrow \underline{v}^\dagger (\lambda \underline{v}) = (\lambda \underline{v})^\dagger \underline{v}$

$\Rightarrow \lambda \underline{v}^\dagger \underline{v} = \bar{\lambda} \underline{v}^\dagger \underline{v}$

$\underline{v} \neq \underline{0}$ so $\underline{v}^\dagger \underline{v} \neq 0$, so $\lambda = \bar{\lambda}$ so real.

(ii) $\underline{v}^\dagger (A\underline{w}) = (A\underline{v})^\dagger \underline{w}$

$\Rightarrow \underline{v}^\dagger (\mu \underline{w}) = (\lambda \underline{v})^\dagger \underline{w}$

$\Rightarrow \mu \underline{v}^\dagger \underline{w} = \lambda \underline{v}^\dagger \underline{w}$

$\lambda \neq \mu$, so $\underline{v}^\dagger \underline{w} = 0$.

(iii) Given $A\underline{v} = \lambda \underline{v}$ with $v \in \mathbb{C}^n$ but A and λ real,

let $\underline{v} = \underline{u} + i\underline{u}'$ with $\underline{u}, \underline{u}' \in \mathbb{R}^n$

Then $A\underline{u} = \lambda \underline{u}$, $A\underline{u}' = \lambda \underline{u}'$ (Re and Im parts)

but $\underline{v} \neq \underline{0} \Rightarrow$ one of $\underline{u}, \underline{u}' \neq 0$, so $\exists \geq 1$ real eigenvector. $\square$

<u>Note</u>: in 6.3 we showed that sets of evecs with distinct evals are linearly independent, but for Hermitian matrices we have that they are <u>orthogonal</u> $\Rightarrow$ linear independence.

Furthermore, previously considered bases B_λ for each eigenspace E_λ, now natural to choose bases B_λ to be orthonormal.

<u>Examples</u>

(i) see 6.1 (b) (i) $A = \begin{pmatrix} 2 & i \\ -i & 2 \end{pmatrix}$ note $A^\dagger = A$

Eigenvalues $\lambda = \begin{cases} 1 \\ 3 \end{cases}$ choose e.vecs $\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ i \end{pmatrix}$

$\underline{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -i \end{pmatrix}$

Note $\underline{u}_1^\dagger \underline{u}_2 = \frac{1}{2}(1 - i)\begin{pmatrix} 1 \\ -i \end{pmatrix} = 0$ as expected

Distinct eigenvalues $\Rightarrow$ diagonalisable.

Set $P = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix}$ then $P^{-1}AP = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$

Note choice of evecs $\Rightarrow P^{-1} = P^\dagger$; P unitary

(ii) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ A real, symmetric

Evals $\lambda = \begin{cases} -1 \\ 2 \end{cases}$ $M_{-1} = 2$

$M_2 = 1$

$E_{-1} = \text{span}\{\underline{w}_1, \underline{w}_2\}$ where $\underline{w}_1 = \begin{pmatrix}1\\-1\\0\end{pmatrix}$, $\underline{w}_2 = \begin{pmatrix}1\\0\\-1\end{pmatrix}$

$m_{-1} = 2$

To find orthonormal basis for E_{-1}, take

$$\underline{u}_1 = \frac{\underline{w}_1}{|\underline{w}_1|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}$$ then consider

$$\underline{w}_2' = \underline{w}_2 - (\underline{u}_1 \cdot \underline{w}_2)\,\underline{u}_1 = \begin{pmatrix}1/2\\1/2\\-1\end{pmatrix} \perp \underline{u}_1$$ by construction

Then $\underline{u}_2 = \frac{1}{|\underline{w}_2'|}\underline{w}_2' = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}$

so $B_{-1} = \{\underline{u}_1, \underline{u}_2\}$ is an orthonormal basis.

For E_2, choose $B_2 = \{\underline{u}_3\}$ where $\underline{u}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$ then

$B_{-1} \cup B_2$ is an orthonormal basis for $\mathbb{R}^3$.

Let P be the matrix with columns $\underline{u}_1, \underline{u}_2, \underline{u}_3$, then

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

(b) Unitary and Orthogonal Diagonalisation

__Theorem__ Any $n \times n$ hermitian matrix is diagonalisable

(i) $\exists$ a basis ~~for~~ of eigenvectors $\underline{u}_1, \ldots, \underline{u}_n \in \mathbb{C}^n$ with $A\underline{u}_i = \lambda \underline{u}_i$, equivalently

(ii) $\exists$ $n \times n$ invertible matrix P with $P^{-1}AP = D$

$$= \begin{pmatrix}\lambda_1 & 0 & 0 & 0\\ & \ddots & & \\ 0 & & & \lambda_n\end{pmatrix}$$

In addition: ~~eves~~ $\underline{u}_i$ can be chosen to be orthonormal

$\underline{u}_i^\dagger \underline{u}_j = \delta_{ij}$ equivalently

the matrix P can be chosen to be unitary.

Special case: For $n \times n$ real symmetric A, the evecs can be chosen to be $\underline{u}_1, \dots, \underline{u}_n \in \mathbb{R}^n$

so $\underline{u}_i^{\dagger} \underline{u}_j = \underline{u}_i \cdot \underline{u}_j = \delta_{ij}$.

(Equivalently P can be chosen to be orthogonal, $P^T = P^{-1}$ so $P^T A P = D$).

(Note proof of first part is non-examinable).

Vectors and Matrices - Lecture 20

24/11/20

See lecture 20 intro / separate notes for non-examinable proof of diagonalisability of Hermitian matrices.

Section 6.5 - Quadratic Forms

(a) Definition and Diagonalisability

Consider $F : \mathbb{R}^2 \to \mathbb{R}$ defined by $F(\underline{x}) = 2x_1^2 - 4x_1x_2 + 5x_2^2$

This can be simplified by writing

$$F(\underline{x}) = x_1'^2 + 6x_2'^2 \quad \text{where} \quad x_1' = \frac{1}{\sqrt{5}}(2x_1 + x_2)$$

$$x_2' = \frac{1}{\sqrt{5}}(-x_1 + 2x_2)$$

This can be found by writing $F(\underline{x}) = \underline{x}^T A \underline{x}$ where $A = \begin{pmatrix} 2 & -2 \\ -2 & 5 \end{pmatrix}$ and then diagonalising A. We find e.vals $\lambda = 1, 6$ and e.vecs $\frac{1}{\sqrt{5}}\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ and $\frac{1}{\sqrt{5}}\begin{pmatrix} -1 \\ 2 \end{pmatrix}$ leading to expressions above.

In general, a quadratic form is a function $\mathbb{R}^n \to \mathbb{R}$ given by $F(\underline{x}) = \underline{x}^T A \underline{x} = x_i A_{ij} x_j$ where A is a real symmetric $n \times n$ matrix. (Any antisymmetric part of A wouldn't contribute).

From 6.4, we can write

$$P^T A P = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

where λ_i are the evals of A and P is a real orthogonal $n \times n$ matrix with columns $\underline{u}_i$ orthonormal e.vecs.

Set $\underline{x}' = P^T \underline{x}$ or $\underline{x} = P\underline{x}'$ which gives

$$F(\underline{x}) = \underline{x}^T A \underline{x} = (P\underline{x}')^T A (P\underline{x}')$$

$$= \underline{x}'^T P^T A P \underline{x}' = \underline{(\underline{x}')^T D \underline{x}'}$$

supposed to be $\underline{x}'^T$

$$= \sum_i \lambda_i x_i'^2$$

We say F has been diagonalised.

Now note $\underline{x}' = x_1' \underline{e}_1 + \ldots + x_n' \underline{e}_n$ $\quad$ $\underline{e}_i$ standard basis

$$\underline{x} = x_1 \underline{e}_1 + \ldots + x_n \underline{e}_n$$

and also $= x_1' \underline{u}_1 \ldots + x_n' \underline{u}_n$

since $x_i' = \underline{u}_i \cdot \underline{x} \iff \underline{x}' = P^T \underline{x}$.

Hence x_i' can be regarded as coordinates wrt a new set of axes given by basis vectors $\underline{u}_i$, the <u>principal axes</u> of F. They are related to standard axes ($\underline{e}_i$) by an <u>orthogonal</u> transformation (P^T).

(b) <u>Examples in $\mathbb{R}^2$ and $\mathbb{R}^3$</u>

In $\mathbb{R}^2$ $\quad F(\underline{x}) = \underline{x}^T A \underline{x}$ with $A = \begin{pmatrix} \alpha & \beta \\ \beta & \alpha \end{pmatrix}$

evals: $\lambda_1 = \alpha + \beta, \quad \lambda_2 = \alpha - \beta$

evecs: $\underline{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \underline{u}_2 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \end{pmatrix}$

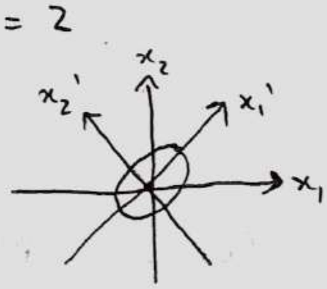
$$F(\underline{x}) = \alpha x_1^2 + 2\beta x_1 x_2 + \alpha x_2^2$$
$$= (\alpha+\beta) x_1'^2 + (\alpha - \beta) x_2'^2$$

with $x_1' = \underline{u}_1 \cdot \underline{x} = \frac{1}{\sqrt{2}}(x_1 + x_2)$

$x_2' = \underline{u}_2 \cdot \underline{x} = \frac{1}{\sqrt{2}}(-x_1 + x_2)$

e.g. $\alpha = \frac{3}{2}, \quad \beta = -\frac{1}{2} \Rightarrow \lambda_1 = 1, \quad \lambda_2 = 2$

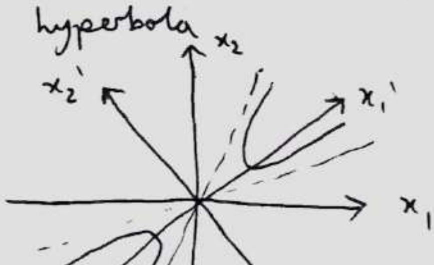
$F = 1$ represents an ellipse $x_1'^2 + 2x_2'^2 = 1$



Instead consider $\alpha = -\frac{1}{2}, \quad \beta = \frac{3}{2} \Rightarrow \lambda_1 = 1,$
$\lambda_2 = -2$

Then $F = 1$ is a hyperbola

$x_1'^2 - 2x_2'^2 = 1$



In $\mathbb{R}^3$ $\quad F(\underline{x}) = \underline{x}^T A \underline{x} = \lambda_1 x_1'^2 + \lambda_2 x_2'^2 + \lambda_3 x_3'^2$

(i) If A has evals $\lambda_1, \lambda_2, \lambda_3 > 0$ then $F = 1$ will define an ellipsoid

(ii) $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$ Found in 6.4 that evals $\lambda_1 = \lambda_2 = -1$

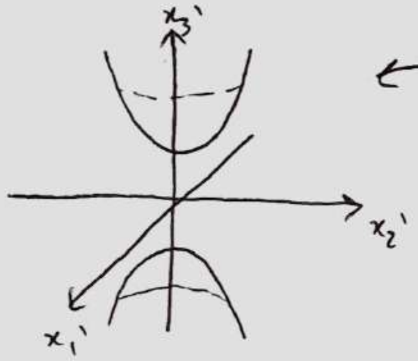
$\underline{u}_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, $\underline{u}_2 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}$ evecs

and $\lambda_3 = 2$ $\quad \underline{u}_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$.

Then $F = 2x_1 x_2 + 2x_2 x_3 + 2x_3 x_1$

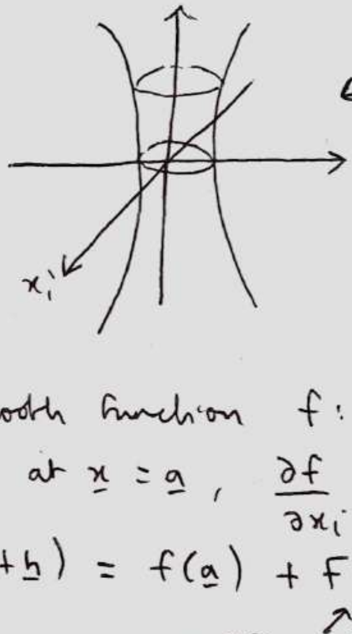
$= -x_1'^2 - x_2'^2 + 2x_3'^2$

$F = 1 \iff 2x_3'^2 = 1 + x_1'^2 + x_2'^2$



← (2 sheeted) hyperboloid

or $F = -1 \iff x_1'^2 + x_2'^2 = 1 + 2x_3'^2$



← (1 sheeted) hyperboloid

<u>Example</u> Consider a smooth function $f : \mathbb{R}^n \to \mathbb{R}$ with a stationary point at $\underline{x} = \underline{a}$, $\frac{\partial f}{\partial x_i} = 0$ at $\underline{x} = \underline{a}$

Taylor's Theorem: $f(\underline{a} + \underline{h}) = f(\underline{a}) + F(\underline{h}) + O(|\underline{h}|^3)$

a QF with $A_{ij} = \frac{1}{2} \frac{\partial^2 f}{\partial x_i \partial x_j}$ (Hessian at $\underline{x} = \underline{a}$)

$F(\underline{h}) = \lambda_1 h_1'^2 + \lambda_2 h_2'^2 + \dots + \lambda_n h_n'^2$

wrt principal axes.

$\lambda_i > 0 \ \forall i$: minimum at $\underline{x} = \underline{a}$

$\lambda_i < 0 \ \forall i$: maximum

both signs: saddle

<u>Note</u> For $n = 2$ consider $\operatorname{tr} A = \lambda_1 + \lambda_2$, $\det A = \lambda_1 \lambda_2$.

Vectors and Matrices - Lecture 21 25/11/20

6.6 Cayley-Hamilton Theorem

If A is an $n \times n$ complex matrix and

$p(t) = c_0 + c_1 t + \dots + c_k t^k$ is a polynomial, then

$p(A) = c_0 I + c_1 A + \dots + c_k A^k$

We can define power series (subject to convergence)

e.g. $\exp(A) = I + A + \frac{1}{2} A^2 + \dots + \frac{1}{r!} A^r + \dots$

(always converges)

For $D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ then $D^r = \begin{pmatrix} \lambda_1^r & & 0 \\ & \ddots & \\ 0 & & \lambda_n^r \end{pmatrix}$ and

$p(D) = \begin{pmatrix} p(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & p(\lambda_n) \end{pmatrix}$ and $\exp(D) = \begin{pmatrix} e^{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & e^{\lambda_n} \end{pmatrix}$

If $B = P^{-1} A P$ (similar to A) with P $n \times n$ invertible, then

$B^r = (P^{-1} A P)^r = P^{-1} A^r P$ for any r

and $p(B) = p(P^{-1} A P) = P^{-1} p(A) P$

Of special interest is the characteristic polynomial

$\chi_A(t) = \det(A - tI) = c_0 + c_1 t + \dots + c_n t^n$

where $c_0 = \det A$, $c_n = (-1)^n$

<u>Theorem</u> (Cayley-Hamilton) $\chi_A(A) = c_0 I + c_1 A + \dots + c_n A^n = 0$.

"a matrix satisfies its own characteristic equation"

<u>Note</u> $-c_0 I = A(c_1 A + \dots + c_n A^{n-1})$

If $c_0 = \det A \neq 0$ then A is invertible and

$A^{-1} = -\frac{1}{c_0}(c_1 A + \dots + c_n A^{n-1})$

General case is <u>non-examinable</u>.

<u>Proof</u> (i) General 2×2 matrix (special case)

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

$$\chi_A(t) = t^2 - (a+d)t + (ad-bc)$$

and check by substitution that $\chi_A(A) = 0$.

(ii) Diagonalisable $n\times n$ matrix

Consider A with evals λ_i and invertible P such that $P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$

Note $\chi_A(D) = \begin{pmatrix} \chi_A(\lambda_1) & & 0 \\ & \ddots & \\ 0 & & \chi_A(\lambda_n) \end{pmatrix} = 0$ since

λ_i are e.vals (roots of $\chi_A(t)$).

Then $\chi_A(A) = \chi_A(PDP^{-1}) = P\,\overset{0}{\chi_A(D)}\,P^{-1} = \underline{0}$.

<u>General case</u> Let $M = A - tI$

$$\det M = \det(A-tI) = \chi_A(t) = \sum_{r=0}^{n} c_r t^r$$

and $\underset{\text{adjugate}}{\tilde{M}} = \sum_{r=0}^{n-1} B_r t^r$ (highest power in t is $n-1$: $\tilde{M}$ generated, removing row/column so up to $n-1$)

This obeys $\tilde{M}M = (\det M)I = \left(\sum_{r=0}^{n-1} B_r t^r\right)(A - tI)$

$= B_0 A + (B_1 A - B_0)t + (B_2 A - B_1)t^2 + \dots + (B_{n-1}A - B_{n-2})t^{n-1}$
$\quad - B_{n-1}t^n$

Comparing coefficients: $c_0 I = B_0 A$, $c_1 I = B_1 A - B_0$, ...,
$c_{n-1} I = B_{n-1}A - B_{n-2}$, $c_n I = -B_{n-1}$

PTO

$$\left.\begin{aligned}
c_0 I &= B_0 A \\
c_1 I &= B_1 A - B_0 \\
&\vdots \\
c_{n-1} I &= B_{n-1} A - B_{n-2} \\
c_n I &= -B_{n-1}
\end{aligned}\right\} \quad
\begin{aligned}
&c_0 I \\
&+ c_1 A \\
&+ c_2 A^2 \\
&+ \dots \\
&+ c_{n-1} A^{n-1} \\
&+ c_n A^n
\end{aligned}
\quad = \quad
\begin{aligned}
&\cancel{B_0 A} \\
&+ \cancel{B_1 A^2} - \cancel{B_0 A} \\
&+ \cancel{B_2 A^3} - \cancel{B_1 A^2} \\
&+ \dots \\
&+ \cancel{B_{n-1} A^n} - \cancel{B_{n-2} A^{n-1}} \\
&- \cancel{B_{n-1} A^n}
\end{aligned}$$

$$= \underline{0}. \qquad \square$$

Chapter 7 Changing Bases, Canonical Forms and Symmetries

7.1 Changing Bases in General

Recall from 4.3c: given a linear map

$T : V \to W$ real or complex vector spaces

and choices of bases $\{\underline{e}_i\}$ $i = 1, \dots, n$ for V and

$\{\underline{f}_a\}$ $a = 1, \dots, m$ for W

The matrix A $(m \times n)$ wrt these bases is defined by

$T(\underline{e}_i) = \sum_a \underline{f}_a \underbrace{A_{ai}}_{\text{coeffs}}$: entries in column i of A are components of $T(\underline{e}_i)$ wrt basis $\{\underline{f}_a\}$.

This is chosen to ensure $\underline{y} = T(\underline{x}) \iff y_a = A_{ai} x_i$ (Σ convention)

where $\underline{y} = \sum_a y_a \underline{f}_a$, $\underline{x} = \sum_i x_i \underline{e}_i$

The equivalence holds since $T\left(\sum_i x_i \underline{e}_i\right) = \sum_i x_i T(\underline{e}_i)$

$= \sum_i x_i \left(\sum_a \underline{f}_a A_{ai}\right) = \sum_a \underbrace{\left(\sum_i A_{ai} x_i\right)}_{y_a} \underline{f}_a$ as required.

Same linear map T has different matrix A' wrt different bases $\underset{(V)}{\{\underline{e}_i'\}}$ and $\underset{(W)}{\{\underline{f}_a'\}}$.

To relate A and A' must specify how bases are related.

Change of base matrices P $(n \times n)$ and Q $(m \times m)$ are defined by

$$\underline{e}_i' = \sum_j \underline{e}_j P_{ji} \ , \quad \underline{f}_a' = \sum_b \underline{f}_b Q_{ba}$$

$\therefore$ entries in column i of P are components of a new basis vector $\underline{e}_i'$ wrt old basis vectors $\underline{e}_j$, similar for Q.

Note P and Q are invertible; in relation above we can exchange $\{\underline{e}_i\}$ and $\{\underline{e}_i'\}$ with $P \to P^{-1}$, similarly for Q.

<u>Proposition</u> With definitions above:

$$\underline{A' = Q^{-1} A P} \ , \quad \text{change of base formula for linear map.}$$

<u>Example</u> $n = 2$, $m = 3$

$$T(\underline{e}_1) = \underline{f}_1 + 2\underline{f}_2 - \underline{f}_3 = \sum_a \underline{f}_a A_{a1}$$

$$T(\underline{e}_2) = -\underline{f}_1 + 2\underline{f}_2 + \underline{f}_3 = \sum_a \underline{f}_a A_{a2}$$

defined this way

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

New basis for V:

$$\left.\begin{aligned} \underline{e}_1' &= \underline{e}_1 - \underline{e}_2 = \sum_i \underline{e}_i P_{i1} \\ \underline{e}_2' &= \underline{e}_1 + \underline{e}_2 = \sum_i \underline{e}_i P_{i2} \end{aligned}\right\} \quad P = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

get by reading across

for W:

$$\begin{aligned} \underline{f}_1' &= \underline{f}_1 - \underline{f}_3 = \sum_a \underline{f}_a Q_{a1} \\ \underline{f}_2' &= \underline{f}_2 = \sum_a \underline{f}_a Q_{a2} \\ \underline{f}_3' &= \underline{f}_1 + \underline{f}_3 = \sum_a \underline{f}_a Q_{a3} \end{aligned} \quad \Rightarrow Q = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}$$

From change of base formula

$$A' = Q^{-1}AP = \begin{pmatrix} 1/2 & 0 & -1/2 \\ 0 & 1 & 0 \\ 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 2 & 2 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 0 \end{pmatrix}$$

Then check directly:

$$\left.\begin{aligned} T(\underline{e}_1') &= 2\underline{f}_1 - 2\underline{f}_3 = 2\underline{f}_1' \\ T(\underline{e}_2') &= 4\underline{f}_2 = 4\underline{f}_2' \end{aligned}\right\} A' = \begin{pmatrix} 2 & 0 \\ 0 & 4 \\ 0 & 0 \end{pmatrix} \text{ as claimed.}$$

<u>Proof of Proposition</u> $T(\underline{e}_i') = T\left(\sum_j \underline{e}_j P_{ji}\right)$ by defn of P

$$= \sum_j T(\underline{e}_j) P_{ji} \quad (T \text{ linear})$$

$$= \sum_{ja} \underline{f}_a A_{aj} P_{ji} \quad (\text{def of } A)$$

but also $T(\underline{e}_i') = \sum_b \underline{f}_b' A'_{bi}$ (def. of A')

$$= \sum_{ab} \underline{f}_a Q_{ab} A'_{bi} \quad (\text{def. of } Q)$$

Vectors are a basis: equate coefficients of $\underline{f}_a$

$$\sum_j A_{aj} P_{ji} = \sum_b Q_{ab} A'_{bi}$$

Hence $AP = QA'$ or $A' = Q^{-1}AP$ <u>as required</u>. $\square$

Vectors and Matrices - Lecture 22

4/12/20

Consider changes in vector components

$$\underline{x} = \sum_i x_i \underline{e}_i = \sum_j x_j' \underline{e}_j' = \sum_i \left(\sum_j P_{ij} x_j' \right) \underline{e}_i$$

$$\Rightarrow \quad x_i = P_{ij} x_j' \quad (\text{using } \Sigma \text{ conv})$$

Write X for $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, X' for $\begin{pmatrix} x_1' \\ \vdots \\ x_n' \end{pmatrix}$

Then $X = PX'$ or $X' = P^{-1} X$ P: change of basis matrix

Similarly $\underline{y} = \sum_a y_a \underline{f}_a = \sum_b y_b' \underline{f}_b$ $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix}$

$$\Rightarrow \quad y_a = Q_{ab} y_b'$$

Then $Y = QY'$ or $Y' = Q^{-1} Y$ $Y' = \begin{pmatrix} y_1' \\ \vdots \\ y_m' \end{pmatrix}$

Matrices are defined to ensure

$$Y = AX \quad \text{and} \quad Y' = A'X'$$

$$\Rightarrow QY' = APX' \Rightarrow Y' = (Q^{-1}AP)X'$$

so $\underline{A' = Q^{-1}AP.}$

Special Cases

(i) $V = W$ with $\underline{e}_i = \underline{f}_i$ and $\underline{e}_i' = \underline{f}_i'$, $P = Q$

and $A' = P^{-1}AP$

Matrices representing the same linear map wrt different basis are similar.
Conversely: if A, A' similar then we can regard them as representing the same linear map with P (invertible) defining a change of basis.
In 6.3: noticed $\mathrm{tr}(A') = \mathrm{tr}(A)$, $\det(A') = \det A$, $\chi_A(t) = \chi_{A'}(t)$
so these are properties of a linear map.

(ii) $V = W = \mathbb{R}^n$ or $\mathbb{C}^n$

$\underline{e}_i$ standard basis, wrt which T has matrix A

If $\exists$ basis of evecs $\underline{e}_i' = \underline{v}_i$ with $A\underline{v}_i = \lambda_i \underline{v}_i$ then

$$A' = P^{-1}AP = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

and $\underline{v}_i = \sum_j \underline{e}_j P_{ji}$ e.vecs cols of P.

(iii) Hermitian: $A^\dagger = A$ then always have a basis of orthonormal e.vecs $\underline{e}_i' = \underline{u}_i$ and then (ii) applies with $P^\dagger P = I$. (P unitary).

7.2 Jordan Canonical / Normal Form

This result classifies $n \times n$ complex matrices up to similarity.

Proposition Any 2×2 complex matrix A is similar to one of:

(i) $A' = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$ (ii) $A' = \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix}$ (iii) $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$

$\lambda_1 \neq \lambda_2$

$\chi_A(t) = (t-\lambda_1)(t-\lambda_2)$ $\underbrace{\hspace{6em}}_{\chi_A(t) = (t-\lambda)^2}$

Proof $\chi_A(t)$ has 2 roots over $\mathbb{C}$

(i) For distinct roots / e.vals, (λ_1, λ_2) we have $M_{\lambda_1} = m_{\lambda_1} = M_{\lambda_2} = m_{\lambda_2} = 1$

So e.vecs $\underline{v}_1, \underline{v}_2$ provide a basis. Hence $A' = P^{-1}AP$ with e.vecs cols of P.

(ii) Repeated root / e.val λ: if $M_\lambda = m_\lambda = 2$ then same argument applies.

(iii) For repeated root / e.val λ with $M_\lambda = 2$, $m_\lambda = 1$, let $\underline{v}$ be e.vec and $\underline{w}$ be any other vector with $\{\underline{v}, \underline{w}\}$ linearly independent. Then $A\underline{v} = \lambda \underline{v}$, $A\underline{w} = \alpha\underline{v} + \beta\underline{w}$

Then the matrix of map wrt basis $\{\underline{v}, \underline{w}\}$ is $\begin{pmatrix} \lambda & \alpha \\ 0 & \beta \end{pmatrix}$

But $\beta = \lambda$ else have case (i) and $\alpha \neq 0$, else have case 2.

Now set $\underline{u} = \alpha \underline{v}$ and note $A(\alpha \underline{v}) = \lambda(\alpha \underline{v})$

and $A\underline{w} = \alpha \underline{v} + \lambda \underline{w}$.

So wrt basis $\{\underline{u}, \underline{w}\}$ have matrix $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$, as claimed.

Alternative approach for case 3:

If $\chi_A(t) = (t-\lambda)^2$ but $A \neq \lambda I$ then $\exists$ some $\underline{w}$ with $\underline{u} = (A - \lambda I)\underline{w} \neq 0$

Now $(A - \lambda I)\underline{u} = (A - \lambda I)^2 \underline{w} = \underline{0}$ (Cayley-Hamilton)

so $A\underline{u} = \lambda \underline{u}$, $A\underline{w} = \underline{u} + \lambda \underline{w}$

and with basis $\{\underline{u}, \underline{w}\}$ have matrix $A' = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}$ as above.

<u>Example</u> $A = \begin{pmatrix} 1 & 4 \\ -1 & 5 \end{pmatrix} \Rightarrow \chi_A(t) = (t-3)^2$

$A - 3I = \begin{pmatrix} -2 & 4 \\ -1 & 2 \end{pmatrix}$ Try $\underline{w} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ and find $\underline{u} = (A-3I)\underline{w} = \begin{pmatrix} -2 \\ -1 \end{pmatrix} \neq \underline{0}$

$\underline{w}$ not e.vec ✓

Then $A\underline{u} = 3\underline{u} \Rightarrow A\underline{w} = \underline{u} + 3\underline{w}$

Check $P = \begin{pmatrix} -2 & 1 \\ -1 & 0 \end{pmatrix} \Rightarrow P^{-1} = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix}$, $P^{-1}AP = \underline{\begin{pmatrix} 3 & 1 \\ 0 & 3 \end{pmatrix}} = A'$.

To extend arguments above to larger matrices, consider

$$N = \begin{pmatrix} 0 & 1 & & 0 \\ & 0 & 1 & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}$$

When applied to standard basis vecs in $\mathbb{C}^n$ we get $\underline{e}_n \mapsto \underline{e}_{n-1} \mapsto \dots \mapsto \underline{e}_1 \mapsto \underline{0}$.

kernel is 1-dimensional. $N^n = 0$ (nilpotent)

Then $J = \lambda I + N = \begin{pmatrix} \lambda & 1 & & 0 \\ & \lambda & 1 & \\ & & \ddots & 1 \\ 0 & & & \lambda \end{pmatrix}$ $(n \times n)$ has $\chi_J(t) = (\lambda - t)^n$

$M_\lambda = n$, $m_\lambda = 1$.

Theorem Any $n \times n$ complex matrix A is similar to a matrix of the following form

$$A' = \begin{pmatrix} J_{n_1}(\lambda_1) & & & 0 \\ & J_{n_2}(\lambda_2) & & \\ & & \ddots & \\ 0 & & & J_{n_r}(\lambda_r) \end{pmatrix}$$

where each diagonal block is a Jordan block:

$J_{n_r}(\lambda_r)$ matrix J $n_r \times n_r$ with $\lambda = \lambda_r$.

$\lambda_1, \dots, \lambda_r$ are e.vals of A and A' and the same e.val may appear in different blocks.

(Non-examinable proof)

$n_1 + n_2 + \dots + n_r = n$

A is diagonalisable iff A' consists entirely of 1×1 blocks.

The expression above is the Jordan Normal Form (proof in IB Linear Algebra / GRM).

7.3 Conics and Quadrics

(a) Quadrics in General

A quadric in $\mathbb{R}^n$ is a hypersurface defined by

$$Q(\underline{x}) = \underline{x}^T A \underline{x} + \underline{b}^T \underline{x} + c = 0$$

for some non-zero symmetric real matrix A, $\underline{b} \in \mathbb{R}^n$, $\underline{c} \in \mathbb{R}$

$$Q(\underline{x}) = A_{ij} x_i x_j + b_i x_i + c = 0$$

Consider classifying solutions for $\underline{x}$ up to geometrical equivalence: no distinction between solutions related by isometries in $\mathbb{R}^n$, i.e. related by translation or rotation / orthogonal transformation about origin.

Note A is invertible iff it has no zero eigenvalues. In this case we can complete the square in eqn $Q(\underline{x}) = 0$ by setting $\underline{y} = \underline{x} + \frac{1}{2} A^{-1} \underline{b}$

$$\Rightarrow \underline{y}^T A \underline{y} = (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b})^T A (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b})$$

$$= (\underline{x}^T + \tfrac{1}{2} \underline{b}^T (A^{-1})^T) A (\underline{x} + \tfrac{1}{2} A^{-1} \underline{b})$$

$$= \underline{x}^T A \underline{x} + \underline{b}^T \underline{x} + \tfrac{1}{4} \underline{b}^T A^{-1} \underline{b} \quad (\text{using } (A^T)^{-1} = (A^{-1})^T).$$

Then $Q(\underline{x}) = 0 \iff \mathcal{F}(\underline{y}) = k$ with $\mathcal{F}(\underline{y}) = \underline{y}^T A \underline{y}$ a quadratic form wrt a new origin: $\underline{y} = \underline{0}$

and $k = \frac{1}{4} \underline{b}^T A^{-1} \underline{b} - c$

Now diagonalise $\mathcal{F}$ as in 6.5: orthonormal evecs give principal axes, e.vals of A and value of k then determine geometrical nature of solution (of quadric).

Examples in $\mathbb{R}^3$ given in 6.5

(i) e.vals > 0 $k > 0$: ellipsoid

(ii) e.vals different signs, $k \neq 0$: hyperboloid

If A has one or more zero eigenvalues, then analysis above changes and simplest standard form may have both linear and quadratic terms.

(b) Conics as Quadrics

Quadrics in $\mathbb{R}^2$ are curves called conics.

det $A \neq 0$ By completing the square and diagonalising A, we get std. form $\lambda_1 x_1'^2 + \lambda_2 x_2'^2 = k$

New variables x_i' correspond to principal axes and new origin

$\lambda_1, \lambda_2 > 0 \Rightarrow$ ellipse for $k > 0$, point for $k = 0$, no soln for $k < 0$.

$\lambda_1 > 0$, $\lambda_2 < 0 \Rightarrow$ hyperbola for $k > 0$ or $k < 0$
pair of lines for $k = 0$ e.g. $x_1'^2 - x_2'^2 = (x_1' - x_2')(x_1' + x_2') = 0$.

det $A = 0$ Suppose (wlog) $\lambda_1 > 0$, $\lambda_2 = 0$ $(A \neq 0)$

Diagonalise A in original formula to get $\lambda_1 x_1'^2 + b_1' x_1' + b_2 x_2' + c = 0$

$\Leftrightarrow \lambda_1 x_1''^2 + b_2' x_2' + c' = 0$

where $x_1'' = x_1' + \frac{1}{2\lambda_1} b_1'$ and $c' = c - \frac{b_1'^2}{4\lambda_1^2}$

If $b_2' = 0$: get pair of lines for $c' < 0$, single line for $c' = 0$, no soln. for $c' > 0$.

If $b_2' \neq 0$: eqn becomes $\lambda_1 x_1''^2 + b_2' x_2'' = 0$ parabola

where $x_2'' = x_2' + \frac{1}{b_2'} c'$

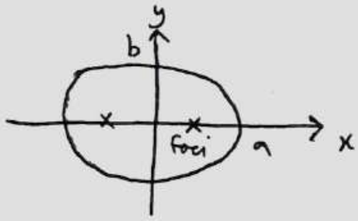
All changes of coordinates correspond to translation, or shift of origin, or orthogonal transformation — preserves lengths and angles.

c) Summary: Standard Forms for Conics

General forms can be written in terms of lengths a, b, semi-major / minor axes, length scale l, eccentricity

(i) Cartesian Coordinates

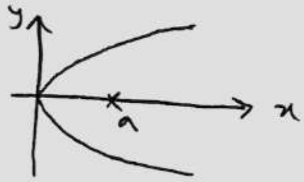
Ellipse $\quad \dfrac{x^2}{a^2} + \dfrac{y^2}{b^2} = 1$



$e < 1 \quad b^2 = a^2(1-e^2) \qquad$ foci $x = \pm ae$

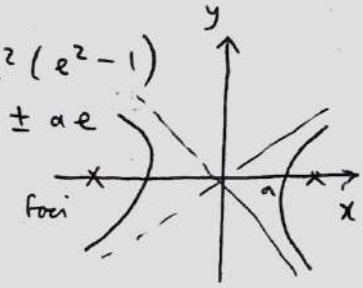
Parabola $\quad e = 1: \quad y^2 = 4ax$

focus $x = +a$



Hyperbola $\quad e > 1 \quad \dfrac{x^2}{a^2} - \dfrac{y^2}{b^2} = 1 \qquad b^2 = a^2(e^2-1)$

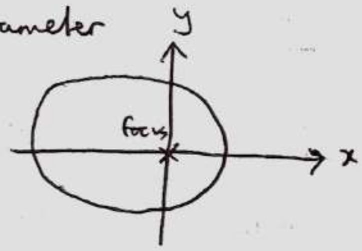
foci $= \pm ae$



(ii) Polar Coordinates (choose origin as focus)

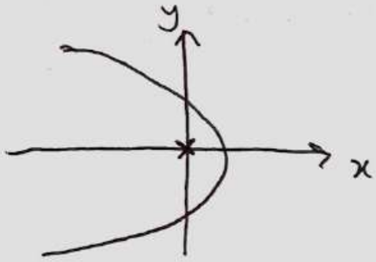
$r, \theta \qquad r = \dfrac{l}{1 + e\cos\theta} \qquad l$ a parameter

Ellipse $\quad e < 1 \quad l = a(1-e^2)$

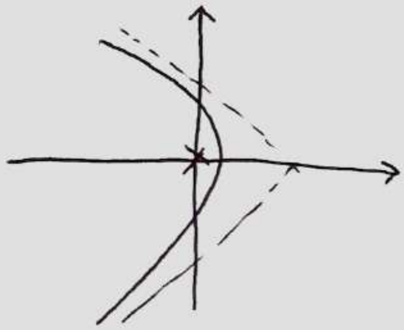


Parabola $\quad e = 1$

$l = 2a$



Hyperbola $\quad e > 1 \quad l = a(e^2-1)$

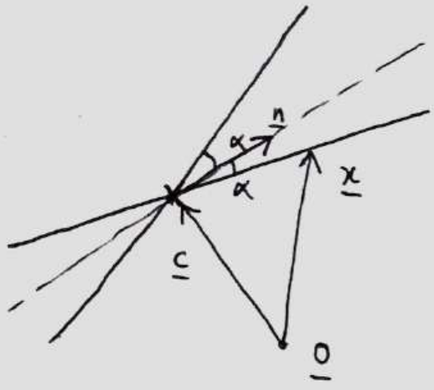


d) <u>Conics as Sections of a Cone</u>

Eqn for cone in $\mathbb{R}^3$: let $\underline{c}$ be apex, $\underline{n}$ axis, (unit vec)

α ($< \frac{\pi}{2}$) angle

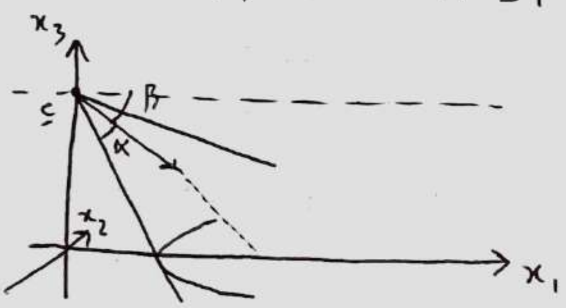
$$(\underline{x} - \underline{c}) \cdot \underline{n} = |\underline{x} - \underline{c}| \cos\alpha$$



Squaring gives double cone:

$$((\underline{x} - \underline{c}) \cdot \underline{n})^2 = |\underline{x} - \underline{c}|^2 \cos^2\alpha$$

Let $\underline{c} = c\underline{e}_3$, $\underline{n} = \cos\beta\, \underline{e}_1 - \sin\beta\, \underline{e}_3$



Consider intersection with plane $x_3 = 0$

$$(x_1 \cos\beta - c\sin\beta)^2 = (x_1^2 + x_2^2 + c^2)\cos^2\alpha$$

$$\Leftrightarrow (\cos^2\alpha - \cos^2\beta)\, x_1^2 + (\cos^2\alpha)\, x_2^2 + 2x_1 c \sin\beta\cos\beta = \text{constant}$$

<u>ellipse</u> if $\cos^2\alpha > \cos^2\beta$ $(\alpha < \beta)$

<u>parabola</u> if $\cos^2\alpha = \cos^2\beta$ $(\alpha = \beta)$

<u>hyperbola</u> if $\cos^2\alpha < \cos^2\beta$ $(\alpha > \beta)$

7.4 Symmetries and Transformation Groups

(a) Orthogonal Transformations and Rotations in $\mathbb{R}^n$

R is orthogonal $\Leftrightarrow$ $R^T R = I \Leftrightarrow (R\underline{x}) \cdot (R\underline{y}) = \underline{x} \cdot \underline{y} \quad \forall \underline{x}, \underline{y}$

$\Leftrightarrow$ rows or cols of R orthonormal

The set of matrices R $(n \times n)$ forms the orthogonal group $O(n)$

$R \in O(n) \Rightarrow \det R = \pm 1$ (check)

$SO(n) = \{ R \in O(n) : \det R = 1 \}$ is a subgroup, the special orthogonal group.

$R \in O(n) \Rightarrow R$ preserves lengths and |n-dim volume|

$R \in SO(n) \Rightarrow R$ also preserves orientation or sign of n-dim volume.

$SO(n)$ consists of all rotations in $\mathbb{R}^n$

Reflections belong to $O(n)$ but not $SO(n)$. For a specific $H \in O_n \setminus SO_n$ have any element of O_n is of the form R or RH with $R \in SO(n)$.

e.g. if n odd, then can choose $H = -I$

Now we can regard transformation $x_i' = R_{ij} x_j$ in two ways:

Active : rotation transforms vectors x_i' components of new vector $\underline{x}' = R\underline{x}$ wrt standard basis

Passive : x_i' components of same vector $\underline{x}$ but wrt new orthonormal basis vecs $\underline{u}_i$:

$$\underline{x} = \sum_i x_i \underline{e}_i = \sum_i x_i' \underline{u}_i \quad \text{with } \underline{u}_i = \sum_j R_{ij} \underline{e}_j$$

(keep $\underline{x}$ the same, change choice of axes)

$$= \sum_j \underline{e}_j P_{ji}$$

so $P = R^{-1} = R^T$

(see 6.5)

(b) 2d Minkowski Space and Lorentz Transformations

Consider a new "inner-product" on $\mathbb{R}^2$ given by

$$(\underline{x}, \underline{y}) = \underline{x}^T J \underline{y} \quad \text{where } J = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

$$= \underline{x_0 y_0 - x_1 y_1}$$

and label components in $\mathbb{R}^2$ by $\underline{x} = \begin{pmatrix} x_0 \\ x_1 \end{pmatrix}$, $\underline{y} = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$

This is not positive definite: $(\underline{x}, \underline{x}) = \underline{x}^T J \underline{x} = x_0^2 - x_1^2$

(a quadratic form with e.vals ± 1)

But still bilinear and symmetric. Standard basis vecs

$\underline{e}_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $\underline{e}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ obey

$(\underline{e}_0, \underline{e}_0) = -(\underline{e}_1, \underline{e}_1) = 1$ and $(\underline{e}_0, \underline{e}_1) = 0$

Still "orthonormal" in a sense.

New inner product is called the <u>Minkowski metric</u> on $\mathbb{R}^2$

$\mathbb{R}^2$ with this metric is called <u>Minkowski space</u>.

Orthogonal transformations preserve standard inner product.

What transformations preserve Minkowski metric?

Now consider matrix $M = \begin{pmatrix} M_{00} & M_{01} \\ M_{10} & M_{11} \end{pmatrix}$ giving map $\mathbb{R}^2 \to \mathbb{R}^2$

Preserves Minkowski metric iff

$$(M\underline{x}, M\underline{y}) = (\underline{x}, \underline{y}) \quad \forall \underline{x}, \underline{y} \in \mathbb{R}^2$$

$$\Leftrightarrow (M\underline{x})^T J (M\underline{y}) = \underline{x}^T M^T J M \underline{y} = \underline{x}^T J \underline{y} \quad \forall x, y \in \mathbb{R}^2$$

$$\Leftrightarrow \underline{M^T J M = J}$$

The set of such matrices form a group. (Note again $\det M = \pm 1$)

Also $\det M = \pm 1$ (wrote it again: check from $M^T J M = J$)

Furthermore, $|M_{00}|^2 \geq 1$ so $M_{00} \geq 1$ or $M_{00} \leq -1$

The subgroup with $\det M = +1$ and $M_{00} \geq 1$ is the <u>Lorentz group</u>.

General form for M: find this by using cols $M\underline{e}_0$, $M\underline{e}_1$ are orthonormal in the same sense as $\underline{e}_0$ and $\underline{e}_1$

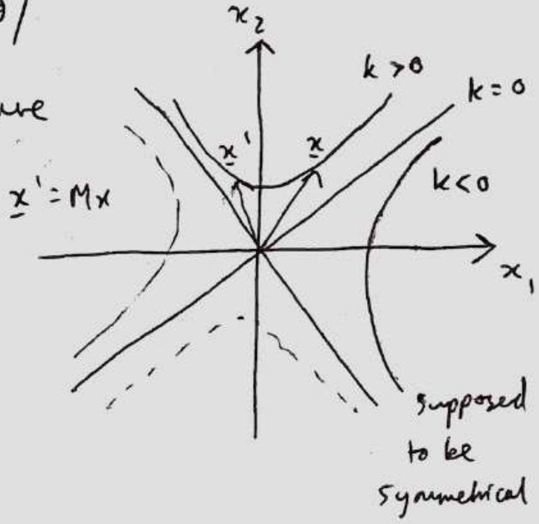
$$(M\underline{e}_0, M\underline{e}_0) = M_{00}^2 - M_{10}^2 = (\underline{e}_0, \underline{e}_0) = 1 \quad (\text{hence } |M_{00}|^2 \geq 1)$$

Taking $M_{00} \geq 1$, can write $M\underline{e}_0 = \begin{pmatrix} \cosh\theta \\ \sinh\theta \end{pmatrix}$ for some real θ.

$$\left.\begin{array}{l} (M\underline{e}_0, M\underline{e}_1) = 0 \\ (M\underline{e}_1, M\underline{e}_1) = -1 \end{array}\right\} \Rightarrow M\underline{e}_1 = \pm\begin{pmatrix} \sinh\theta \\ \cosh\theta \end{pmatrix}$$

Finally, imposing $\det M = +1$, we have

$$M = \begin{pmatrix} \cosh\theta & \sinh\theta \\ \sinh\theta & \cosh\theta \end{pmatrix}$$



curves with $(\underline{x}, \underline{x}) = k$, constant as shown.

Note that matrices found obey

$$M(\theta_1)M(\theta_2) = M(\theta_1 + \theta_2)$$

using hyperbolic addition formulas.

<u>Physical interpretation / application</u>

Set $M(\theta) = \gamma(v)\begin{pmatrix} 1 & v \\ v & 1 \end{pmatrix}$ $\quad v = \tanh\theta$, $\gamma(v) = (1-v^2)^{-1/2}$

new parameter $-1 < v < 1$

Rename $x_0 \to t$ time coordinate

$x_1 \to x$ space coordinate

Then $\underline{x}' = M\underline{x} \iff \begin{cases} t' = \gamma(t + vx) \\ x' = \gamma(x + vt) \end{cases}$

Lorentz transformation / Lorentz boost relating time/space coordinates for observers moving with relative speed v.

γ factor in Lorentz transformation gives rise to time dilation and length contraction effects.

Group property:

$M(\theta_3) = M(\theta_1) M(\theta_2)$ with $\theta_3 = \theta_1 + \theta_2$

(recall $v = \tanh \theta$)

$\Rightarrow$ related composition of velocities with $v_i = \tanh \theta_i$

$(i = 1, 2, 3)$

gives $v_3 = \dfrac{v_1 + v_2}{1 + v_1 v_2}$ (by addition formula)

Velocities don't add in the same way as nonrelativistic mechanics.

$\boxed{\text{End of examinable content}}$