MA 106 - Linear algebra

- Associativity:

  if AB, BC are defined:

  (AB)C , A(BC) also defined & equal

- $(AB)^T = B^T A^T$

  AB → compatible for multiplication ⇒ $B^t A^t$ also comp.

  but not $A^t B^t$

- Dot product: (Inner product/Scalar product)

  - defined for two column vectors of same size n

    ↳ represent pts. in $\mathbb{R}^n$

$$v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} \qquad w = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$$v \cdot w = \sum_j v_j w_j$$

note: $v \cdot w = v^T w$ as a 1x1 matrix

$w \cdot v = w^T v$

* to map from $\mathbb{R}^n \longrightarrow \mathbb{R}^m$, use a $m \times n$ matrix A and the transformation is:

$$x \longmapsto Ax$$

x : $n \times 1$ $\rightarrow R^n$ (domain)

A : $m \times n$

Ax = $m \times 1$ $\rightarrow R^m$ (range)

* Anticlockwise rotation matrix:

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

* Scaling matrix:

$$\begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}$$

* linear systems & matrices:

m linear eqns in n variables:

$$a_{11}x_1 + a_{12}x_2 \dots a_{1n}x_n = b_1 \quad -(1)$$

$$a_{21}x_1 + a_{22}x_2 \dots a_{2n}x_n = b_2 \quad -(2)$$

$$\vdots$$

$$a_{m1}x_1 + a_{m2}x_2 \dots a_{mn}x_n = b_m \quad \dots (m)$$

Can be written as:

$$\underset{\substack{\downarrow\\ \text{coefficient}\\ \text{matrix}\\ (m\times n)}}{A}\quad \underset{\substack{\downarrow\\ (n\times 1)}}{x} \quad = \quad \underset{\substack{\downarrow\\ (m\times 1)}}{b}$$

* $A^{+} = [A \mid b]$ = augmented matrix $(m \times n+1)$

$$\left[\begin{array}{c|c} \quad A \quad & \begin{matrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{matrix} \end{array}\right]$$

* Gauss elimination via ERM's:

(i) P_{jk} : interchange j^{th} & k^{th} row

$$\underset{\substack{\downarrow\\ \text{modified}\\ \text{matrix}}}{A'} = \underset{\substack{\downarrow\\ \text{ERM}}}{P_{jk}}\ \underset{\substack{\downarrow\\ \text{original}\\ \text{matrix}}}{A}$$

= the matrix obtained by interchanging jth & k^{th} rows of I

$$P_{13} = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

(ii) $E_{jk}(c)$: Row j = Row j + (c × Row k)

$E_{jk}(C) = I + c\,\mathcal{E}_{jk}$; the matrix obtained by:

→ taking I
→ adding (c×Row k) to Row j

(iii) $M_j(\lambda)$: Row j = Row j * λ

$M_j(\lambda)$ = Take I, multiply J^{th} row by λ.

Row Echelon Form:

- Each row (except maybe the first) should start w/ a string of 0's.
- Each row has strictly more zeroes than the prev. row.

★ The first non zero entry in the j^{th} row is the j^{th} pivot.

below & strictly to the right of $(j-1)^{th}$ pivot.

- Below a pivot, all entries must be 0.
- To get the REF, do clever ERO's.
- REF is not unique, reduced REF is.
- Value of pivot isn't unique but position is unalterable.
- reduced REF : all pivots = 1 .

★ The columns in which we find the pivots indicate the "non-free variables". The remaining colums are the free variables which can take any value in $\mathbb{R}^x$ (x = no. of f.v's) and the non f.v's are expressed in terms of the f.v's.

★ Theorem:

Let A be a sq. matrix of order n. There exists ERM's such that

$E_N \cdots E_2 E_1 A$ = identity matrix OR its last row = 0's.

(Basically, every matix is either invertible or not invertible)

★ if there exists a matrix B such that:

$$AB = BA = I$$

$B \rightarrow$ inverse of A

- inverse of a matrix is unique
- $(AB)^{-1} = B^{-1}A^{-1}$
- Every ERM is invertible : $E_{jk}(c)^{-1} = E_{jk}(-c)$

$$P_{jk}^{-1} = P_{jk}$$

$$M_j(\lambda)^{-1} = M_j(1/\lambda)$$

★ Theorem: A square matrix is invertible if & only if it is a product of ERM's.

Proof:

the reduced REF

$E_N \cdots E_3 E_2 E_1 A = I \longrightarrow$ invertible

OR

last row is zero $\rightarrow$ not invertible

$A = E_1^{-1} E_2^{-1} E_3^{-1} \cdots E_N^{-1}$ (product of ERM's)

★ A subset V of $\mathbb{R}^n$ is called a vector subspace if :

(i) V is non empty

(ii) $X \in V$ and $a \in R \Rightarrow aX \in V$

(closed under scalar multiplication)

(iii) $x, y \in V \Rightarrow x + y \in V$

(closed under addition)

★ vector subspaces of $\mathbb{R}^3$:

(i) zero dimensional : singleton $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$

(ii) one dimensional : line through origin (basis : one vector)

(iii) two dimensional : plane through origin (basis : two vectors)

~~III~~ (iv) three dimensional : $\mathbb{R}^3$ itself (basis : three vectors)

Suppose : $v_1, v_2 \dots v_k \in \mathbb{R}^n$

& $a_1, a_2 \dots a_k \in \mathbb{R}$

$a_1v_1 + a_2v_2 \dots a_kv_k$ is called a <u>linear combination</u>

of the vectors w/ their coeffs.

$V = \{ a_1v_1 + a_2v_2 \dots a_kv_k \;/\; a_1, a_2 \dots \in \mathbb{R} \}$

└→ <u>linear span</u> of those vectors

" V is generated by $v_1, v_2 \dots v_k$ "

★ To generate a 3 dimensional span (or 'n' dimensional)

→ choosing <u>more than n</u> basis vectors will cause redundancy

★ optimal is n.

★ if the null vector is part of a set $\{v_1, v_2 \dots v_k\}$ we can directly say, its linearly dependent bc coeff of null vector can be made non zero.

★ <u>Linearly dependent</u>: A set of vectors is lin. dep. if

$a_1v_1 + a_2v_2 \dots a_kv_k = 0$

such that at least one among $\{a_1, a_2 \dots a_k\}$ is non zero.

- if the set isn't lin dep., its obv. linearly independent.

i.e $a_1v_1 + a_2v_2 \quad \dots \quad a_kv_k = 0$

$$\Rightarrow \quad \underline{a_1 = a_2 \dots = a_k = 0}$$

(no other possible way of getting 0 using them)

★ convention : Empty set is linearly indep.

subset of L.I set is also L.I.

★ Linear span of ϕ is $\{0\}$ (ϕ - empty set)

★ First Fundamental Lemma :

suppose V is a vector subspace of R^n, generated by k vectors then, any set of $(k+1)$ or more vectors in V are linearly dependent.

Basis and dimension :

- V is a vector subspace of R^n, then a subset S of V is called a basis of V if :

(i) S is linearly independent

(ii) linear span of $S = V$

Theorem:

If V is a vector subspace of $\mathbb{R}^n$ then

(i) V has a basis

(ii) Any two bases have same no. of elements

★ A vector subspace in $\mathbb{R}^n$ will have a basis of at-most 'n' elements

(n+1 = lin. independent condition fails due to 1st Lemma)

Dimension of V is the no. of elements in its basis. (cardinality)

Rank of a matrix:

- let A be a $m \times n$ matrix
- columns of A are like elements of $\mathbb{R}^m$ & linear span is called the Column Space of A

  rows of A " $\mathbb{R}^n$ " Row Space of A.

- Row Rank of A = dimension of row space of A

  Column " " column "

★ If the linear span of a set of vectors = V, the set can be reduced down to the basis of V by deleting those vectors which are linear combs. of the rest (causing linear dependence)

Lemma: Performing ERO's, does not alter Row rank or Column Rank

Theorem: Row rank A = Column rank A = Rank of A

- Suppose dim V = k
  - if S is a linearly ind. subset of V and it contains k vectors $\longrightarrow$ S is a basis
  - if S is a set of k vectors and linear span of S = V $\Longrightarrow$ S is lin. indep. & S is a basis

Theorem:

Let A be a square matrix :
the following are equivalent :

- Rows of A are linearly indep.
- Columns of A are linearly indep.
- A is invertible
- det A $\neq$ 0

Kroencker - Capelli Theorem :

A system of linear eqⁿs $Ax = b$ has a soln. iff
Rank A = Rank $[A|b]$

Null space and Nullity of a matrix :

Let A be an $m \times n$ matrix.

Null space (A) $= \{ v \in R^n \mid Av = 0 \}$

- It is a vector subspace of R^n.
- Nullity (A) = dim (Null space (A))

★ Rank - Nullity theorem :

$$\text{Rank } A + \text{Nullity } A = n = \text{no. of columns. of } A$$

- structure of space of solns. of $Ax = b$:

assume the system has a soln. x_0 ($\Rightarrow$ Rank A = Rank A^+)

the set of all solns. is given by:

$$\{ x_0 + n \mid n \in \text{Null space}(A) \}$$

$$A(x_0+n) = Ax_0 + An = Ax_0 = b \checkmark$$

$\therefore$ we see, if nullity > 0 → infinite solns. of form $x_0 + n$

if nullity $= 0$ → unique soln.

OR

→ no soln.

Determinantal Rank

Lemma: if A is an $m \times n$ matrix and a $k \times k$ submatrix of it has non zero det., then rank $\geq k$

- converse also true.

- reason: nonzero det. $\Rightarrow$ linear independence of those columns

$$\Downarrow$$

rank has to be that, or more

Theorem: Let A be a $m \times n$ matrix

we say A has determinantal rank k if:

(i) $\exists$ some $(k \times k)$ submatrix of A w/ non zero det.

(ii) all $(k+1) \times (k+1)$ matrices have zero determinant.

- Row rank = column rank = determinantal rank

- Performing ERO's on A does not change its null space / nullity.

It does not change row / column rank either.

Rank A = Rank $\tilde{A}$

Let $\tilde{A}$ = REF of A

* Rank $\tilde{A}$ = no. of pivots = k $\rightarrow$ no. of non free variables

l = no. of free variables

$k + l = n$
$\quad \hookrightarrow$ no. of columns

- set diff values to the fv's $\begin{pmatrix} 0, 1, \ldots & 0 \\ 1, 0 \ldots & 0 \\ 0, 0 \ldots & 1 \end{pmatrix}$

  'i' such vectors

- now solve $Ax = 0$

  ~~ans~~

- the solns. obtained are obv. lying in the null space

  ↳ which form a linearly independent set

- the 'i' different solns. obtained now are also linearly independent & forms the basis for the null space of A.

★ the columns w/ the pivots are linearly independent that's why we cannot write other variables using those. this makes them 'pivoted'.

★ Rank ≤ no. of columns

≤ no. of rows

Determinants :

- det A is a fn. that assigns to each matrix, a specific real number

$$f : \underbrace{\mathbb{R}^n \times \mathbb{R}^n \ldots \mathbb{R}^n}_{n \text{ times}} \longrightarrow \mathbb{R}$$

that satisfies the following properties :

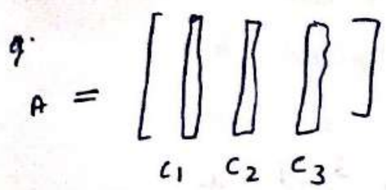
(i) $f(\hat{e}_1, \hat{e}_2 \ldots \hat{e}_n) = 1$ $\quad \hat{e}_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \quad \hat{e}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$

. . .

$\hat{e}_n$ = m membered column vector w/ n^{th} value = 1

* The inputs to this fn. are the columns of matrix.

g.



$$f(c_1, c_2, c_3) = \det A$$

(ii) $f(c_1 \cdots c_i, \cdots c_j \cdots c_n) = \underset{\uparrow}{-} f(c_1 \cdots c_j \cdots c_i \cdots c_n)$

skew-symmetry
interchange of two colums changes sign of det

(iii) multilinearity :

~~$f(\alpha c_1' + \beta c_2')$~~

$$f(\alpha c_1' + \beta c_1'', c_2 \ldots c_n)$$

$$= \alpha f(c_1', c_2 \ldots c_n) + \beta f(c_1'', c_2 \ldots c_n)$$

* This is a func. which maps a square matrix to a unique num.!
domain is : <u>n</u> tuples of column vectors in <u>R^n</u> only!

* notation thing :

$$f(\hat{e_i}, \hat{e_j}, \hat{e_k}) = \epsilon_{ijk}\, f(\hat{e_1}, \hat{e_2}, \hat{e_3})$$

$= \epsilon_{ijk}$ = parity
= no. of exchanges
required to order i, j, k

parity = 1 ; even swaps

= -1 ; odd swaps

eg. $\overline{2,3,1} \longrightarrow 1,\underbrace{3,2} \longrightarrow \underline{1,2,3}$!

2 swaps ⇒ parity = 1

★ $\epsilon_{ijk} = 0$ if two of i, j, k are same.

∴ the det fn. is:

$$f\left(\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}\right)$$

$$= \sum a_{i1} a_{j2} a_{k3} \epsilon_{ijk}$$ (summed over distinct i,j,k)

↓

fancy way of writing our usual det. expansion

★ general det. formula:

$$\sum \epsilon_{i_1 . i_2 .. i_n} \times a_{i_1 1} \times a_{i_2 2} \ldots a_{i_n n}$$

★ on writing out the expression, there are n! terms

. Computing det : use smart col. ops.

- ★ adding multiple of R_j to R_i keeps det same
- ★ multiplying R_i by k multiplies det by k
- ★ two rows same ⇒ det = 0
- ★ row interchange ⇒ det = −det ;

second fundamental lemma in Linear Algebra:

suppose $f : \mathbb{R}^n \times \mathbb{R}^n \dots \mathbb{R}^n \longrightarrow \mathbb{R}$ is a multilinear function
then the foll. are equivalent :

(i) f is skew symm

(ii) $f(c_1 \dots c_i \dots c_j \,..\, c_n) = 0$ whenever $c_i = c_j$

(iii) $f(c_1 \dots c_i, c_{i+1} \dots c_n) = 0$ whenever $c_i = c_{i+1}$

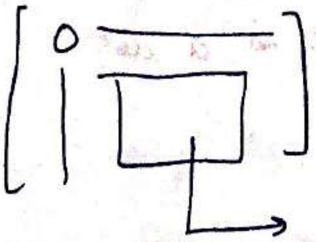
Minors & cofactors:

In an n×n matrix,

for a_{ij} : delete row & column containing it, and the value of det of $(n-1) \times (n-1)$ submatrix is the minor.

cofactor of a_{ij} is minor w/ appropriate sign (A_{ij})

$$\boxed{A_{ij} = (-1)^{i+j} M_{ij}}$$



→ cofactor of a_{11}

★ det A = ~~$a_{11}A_{11} + a_{22}A_{22} + a_{33}$~~

$= a_{11}A_{11} + a_{21}A_{21} + a_{31}A_{31}$

= many other ways of writing along any row/column, it gives the same answer. (surprisingly)

★ $\det A = \det A^T$

★ if A is an $n \times n$ square matrix :

then if Rank $A = n$

$\Rightarrow \det A \neq 0$

nullity $= 0 \Rightarrow$ unique soln.

Rank $< n \Rightarrow$ nullity $> 0 \Rightarrow$ infinite solns.

★ $Ax = 0$ has the trivial soln - null vector iff $\det A \neq 0$

if $\det A = 0$, there are ∞ non-trivial solns.

Gram determinant:

- → used to check linear dependency of n vectors
- → $n \times n$ square matrix where every element is a dot product of two vectors.

Scanned with CamScanner

eg. v_1 , v_2 , v_3 in R^3

$$\text{grammian} = \begin{bmatrix} \bar{v}_1 \cdot \bar{v}_1 & \bar{v}_1 \cdot \bar{v}_2 & \bar{v}_1 \cdot \bar{v}_3 \\ \bar{v}_2 \cdot \bar{v}_1 & \bar{v}_2 \cdot \bar{v}_2 & \bar{v}_2 \cdot \bar{v}_3 \\ \bar{v}_3 \cdot \bar{v}_1 & \bar{v}_3 \cdot \bar{v}_2 & \bar{v}_3 \cdot \bar{v}_3 \end{bmatrix}$$

* The vectors are linearly indep. iff grammian det $[\bar{v}_j \cdot \bar{v}_i]$

or $[v_j^T v] \neq 0$

(converse also true!)

* Gram det ≥ 0 always

Adjugate matrix & Cramer's Rule:

. To form adjugate matrix:

(i) Take A

(ii) Replace every element w/ its cofactor (minor with sign)

(iii) Take transpose

. We see also that:

$$\sum_{j=1}^{n} a_{ij} A_{i'j} = \begin{cases} 0 & ; \ i \neq i' \\ \det A & ; \ i = i' \end{cases}$$

* if you multiply ~~minor~~ cofactor of that guy w/ that guy itself,

only then you get det.

* if you multiply it with some other guy, you get 0.

that means we can write this all packaged up in one matrix equation :

$$\boxed{A \cdot \text{adj } A = \det A \, (I)}$$

★ if $\det A \neq 0$:

$$A^{-1} = \frac{\text{Adj } A}{\det A}$$

and :

$$Ax = B \Rightarrow x = A^{-1} B$$

$$= \frac{\text{Adj } A \times B}{(\det A)}$$

★ Product theorem:

$$\det AB = \det A \times \det B$$

★ $\det A = \det A^{T}$

Block multiplication of matrices:

A $2n \times 2n$ matrix can be written as 2×2 matrix where every element is $n \times n$. This is called a partitioned matrix.

if $$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{matrix} \}n \\ \}n \end{matrix}$$ with columns $\underbrace{\quad}_{n}$ $\underbrace{\quad}_{n}$ $$N = \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

MN is $$\begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} P & Q \\ R & S \end{bmatrix}$$

$$= \begin{bmatrix} AP + BR & AQ + BS \\ CP + DR & CQ + DS \end{bmatrix}$$

★ Let there be 4 $n \times n$ matrices I, P, Q, O

$$\det \begin{bmatrix} I & P \\ 0 & Q \end{bmatrix} = \det Q$$

* suppose $$J = \begin{bmatrix} 0 & -I_n \\ I_n & 0 \end{bmatrix}$$

A is a $2n \times 2n$ matrix such that

$$A^T J A = J$$

$\det A = \pm 1$ using product theorem

but

$\det A = 1$ is the scene, not -1. don't ask why

☆ Let G be the set of all matrices such that $A^T J A = J$

→ $I_{2n} \in G$

→ $J \in G$

→ $A, B \in G \implies AB \in G$

☆ → $A \in G \implies A^{-1} \in G$ (A is invertible bc $\det A = \pm 1$)

☆ if S is a set of m×m matrices such that:

$$I_m \in S$$

$$A, B \in S \implies AB \in S$$

$$A \in S \implies A^{-1} \in S$$

we say S is a group of matrices

The G above is called the group of symplectic matrices.

Inner products :

let V be a vector space (subspace of $\mathbb{R}^n$)
An inner product on V is a function from $V \times V \rightarrow \mathbb{R}$
which assigns to each pair of vectors $(v_1, v_2) \in V \times V$
a real no. $= \langle v_1, v_2 \rangle$ satisfying the foll. properties:

Bilinearity: $\langle v_1' + v_1'', v_2 \rangle = \langle v_1', v_2 \rangle + \langle v_1'', v_2 \rangle$

$$\langle v_1, v_2' + v_2'' \rangle = \langle v_1, v_2' \rangle + \langle v_1, v_2'' \rangle$$

Symmetry: $\langle v, u \rangle = \langle u, v \rangle$

also, $\langle \alpha v_1, v_2 \rangle = \langle v_1, \alpha v_2 \rangle = \alpha \langle v_1, v_2 \rangle$ (α - scalar)

Positive defiteness: $\langle v, v \rangle > 0$ for all v other than null vector

$$\langle 0, 0 \rangle = 0$$

inner product is basically generalised dot product in $\mathbb{R}^n$

$$\langle [x_1, x_2 .. x_n]^T, [y_1, y_2 ... y_n]^T \rangle$$

$$= x_1y_1 + x_2y_2 \dots x_ny_n$$

☆ if $\langle v_1, v_2 \rangle = 0$, they are orthogonal. $(v_1 \perp v_2)$ (aka perp.)

☆ $0 \perp v \quad \forall v$

☆ unit vector : $\langle v, v \rangle = 1$

In general, norm of a vector $\|v\| = \sqrt{\langle v, v \rangle}$

some random formulae:

$$\|v_1 + v_2\|^2 + \|v_1 - v_2\|^2 = 2(\|v_1\|^2 + \|v_2\|^2)$$

$$\|v_1 + v_2\|^2 = \|v_1\|^2 + \|v_2\|^2 \quad \text{when } v_1 \perp v_2$$

Cauchy-Schwarz Inequality:

if V is an inner product space & $v_1, v_2 \in V$ then

$$|\langle v_1, v_2 \rangle| \leq \|v_1\| \cdot \|v_2\|$$

equality holds iff v_1, v_2 are linearly dep (multiples of each other)

no one cares about proof.

Theorem:

suppose V is a vector space endowed w/ an inner product $\langle , \rangle$, then:

nonzero vectors $v_1, v_2 \dots v_k$ are orthogonal

$$\Downarrow$$

$v_1, v_2 \dots v_k$ are linearly independent

why?:
gram det $\neq 0$ (all diagonal elements non zero, and rest 0)

* The converse is false

lin. indep. $\not\Longrightarrow$ orthogonal

eg. $0\hat{i} + 1\hat{j}$
$1\hat{i} + 2\hat{j}$

Angle b/w two vectors:

let v, w be two non zero vectors in an inner product space

$$|\langle v, w \rangle| \leq \|v\| \, \|w\|$$

$$\Rightarrow \quad -1 \leq \left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle \leq 1$$

So $\exists$ a unique $\theta \in [0, \pi]$

$$\text{st} \quad \cos\theta = \left\langle \frac{v}{\|v\|}, \frac{w}{\|w\|} \right\rangle$$

This θ is the angle b/w vectors

$\theta = \pi/2$ when $v_1 \perp v_2$

The Gram Schmidt Process:

let's say you have vectors (nonzero) v_1, v_2, v_3

make $\hat{v}_1 = \dfrac{v_1}{\| v_1 \|}$

express v_2 as $\parallel$ to v_1 comp + $\perp$ to v_1 comp

$$\begin{array}{cc} \Downarrow & \Downarrow \\ (\bar{v}_2 \cdot \hat{v}_1)\, \hat{v}_1 & \bar{v}_2 - v_2^T \end{array}$$

v_2^T : tangential

v_2^N : normal

now, make $\hat{v}_2$

express v_3 using components in plane of v_1 and v_2 and normal to that plane

$$\bar{v}_3 = (\bar{v}_3 \cdot \hat{v}_2)\, \bar{v}_2 + (\bar{v}_3 \cdot \hat{v}_1)\, \hat{v}_1 + \bar{v}_3^N$$

at any pt. if $v_n^N = 0$ ⇒ no normal comp.

you can express the last vector as l.c. of the prev. ones ⇒ lin dep.

orthonormal systems & orthonormal basis:

A set of vectors $v_1 , v_2 .. v_k$ in a inner product space V is said to form an orthonormal system if

(i) $v_i \perp v_j$ for $i \neq j$

(ii) each v_i is a unit vector

* In G.S process:

the normal component extracted in every step is vasically another orthonormal vector which will make up the vasis. Once a normal comp. comes out as 0, we know our job is done and we have the vasis.

let there be vectors $\{ v_1 , v_2 , v_3 \dots v_i \}$

let the corr. normal components be :

$$\frac{v_1}{\| v_1 \|} = w_1$$

$$\frac{v_2^N}{\| v_2^N \|} = w_2 \quad \text{(normal to line } v_1)$$

$$\frac{v_3^N}{\| v_3^N \|} = w_3 \quad \text{(normal to plane of } v_1 \,\&\, v_2)$$

$$\vdots$$

$$\frac{v_i^N}{\| v_i \|^N} = w_N$$

then $\{w_1, w_2 \dots w_i\}$ is an orthonormal system

For ~~[illegible]~~ $1 \leq j \leq i$:

$$\text{Lin span } \{v_1 \dots v_j\} = \{w_1 \dots w_j\}$$

As long as process doesn't terminate, the vectors are L.I.

if it terminates at some pt, then those vectors up till that pt are L.D.

<u>Corollary:</u> Given a vector space V endowed w/ an inner product

(i) there exists an orthonormal basis for V

(ii) given any basis $B = \{v_1 \dots v_k\}$ we can find an orthonormal basis $\{w_1 \dots w_k\}$ using GSP such that

$$\text{span } \{v_1 \dots v_k\} = \text{span } \{w_1 \dots w_k\}$$

☆ A circle in $\mathbb{R}^n$ is a set of vectors v such that all of them have $\langle v, v \rangle = \|v\|^2 = \text{constant}$

eg. $\bar{u} \cos\theta + \bar{v} \sin\theta$ where $\bar{u}, \bar{v}$ are $\perp$ unit vectors

Scanned with CamScanner

if you have a $3 \cdot 3$ matrix where every row is ~~an~~ a unit vector & is orthonormal to other rows,

then $A A^T = I_3 = A^T A$

$$\begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}^{\rightarrow} \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}^{\downarrow} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

★ The columns will also be orthonormal as well as unit vectors too! (Rank = 3 = Row/ Column rank)

★ given $\bar{v}, \bar{w}$ in $\mathbb{R}^3$, find $\bar{z}$? such that $A = [v, w, z]$

$A^T A = I$ ($\bar{v}, \bar{w}$ are $\perp$ unit vecs)

$$\bar{z} = \frac{\bar{v} \times \bar{w}}{\|\bar{v} \times \bar{w}\|}$$

★ in $\mathbb{R}^n$, to find this n^{th} vector, find normal component up till last step using G.S process & normalise it to get ans.

Inner products & symmetric matrices

A - real symm · matrix

$u, v \in \mathbb{R}^n$ (column vectors)

$u^T A v$ is a real number = $(1\times n)(n\times n)(n\times 1)$

= (1×1) matrix

Define $\langle u, v \rangle = u^T A v$

... satisfies all inner product conditions except positive definiteness so we define that:

Def: A real symm matrix is said to be non negative if

$$u^T A u \geq 0 \quad \forall \; u \in \mathbb{R}^n$$

and is said to be +ve definite if

$$u^T A u > 0 \quad \forall \; u \in \mathbb{R}^n - \{0\}$$

$$u^T A u = 0 \quad \text{if } u = \{0\}$$

* if A & B are real & +ve definite

$\alpha A + \beta B$ are also " (α, β are scalars)

* So now if A is a +ve def. matrix,

$\langle u, v \rangle = u^T A v$ is an inner product on $\mathbb{R}^n$

. Each $n \times n$ matrix can be regarded as an element of $\mathbb{R}^{n^2}$

if its symm then you need $\frac{n(n+1)}{2}$ arguments to describe it (dimension)

skew symm " $\frac{n(n-1)}{2}$ "

* $Tr(A) = \Sigma$ diagonal elements in square matrix

Complex Vector Spaces:

* Everything is same as before, but replace $\mathbb{R}$ w/ $\mathbb{C}$

now, vectors have complex entries

$$z = \begin{bmatrix} z_1 \\ z_2 \\ \vdots \\ z_n \end{bmatrix} \qquad \bar{z} = \begin{bmatrix} \bar{z_1} \\ \bar{z_2} \\ \vdots \\ \bar{z_n} \end{bmatrix}$$

. scalars are now complex
· all theorems still hold.
· instead of base, dim we now say $\mathbb{C}$ base, $\mathbb{C}$ dim
. instead of inner product, we have sm new. called hermitian form. It spits out a complex number for each vector pair. Its rules are slightly diff. from inner product.

→ complex conjugate

(i) $\langle v_1, v_2 \rangle = \overline{\langle v_2, v_1 \rangle}$

(ii) $\langle v_1' + v_1'', v_2 \rangle = \langle v_1', v_2 \rangle + \langle v_1'', v_2 \rangle$

$\langle v_1, v_2' + v_2'' \rangle = \langle v_1, v_2' \rangle + \langle v_1, v_2'' \rangle$

$\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$ ... important!

$\langle v, \lambda w \rangle = \bar{\lambda} \langle v, w \rangle$ note the difference ... sesquilinearity

(iii) $\langle v, v \rangle > 0 \quad \forall \, v \in \mathbb{C}^n - \{0\}$

$\langle v, v \rangle = 0 \quad$ if $v = \{0\}$

★ Math sesquilinearity convention is opp. to that of QM

eg. usual hermitian product on $\mathbb{C}^n$

$z, w \in \mathbb{C}^n$

$$\langle z, w \rangle = \sum_{j=1}^{n} z_j \overline{w_j}$$

can be written as

$$\langle z, w \rangle = w^* z$$

$$w^* = \bar{w}^T$$

adjoint = conjugate, transposed

$$(AB)^* = B^* A^*$$

- if $A = A^*$ ⇒ self adjoint / hermitian matrix
- if A, B, commute & are both hermitian ⇒ AB is hermitian

* if $A = A^*$,

we can consider hermitian product as

$$\langle z, w \rangle = w^* A z$$

again, all conditions except positive definiteness can easily be proved

def: A hermitian matrix A is said to be +ve definite if

$$w^* A w > 0 \quad \forall \; w \neq 0$$

if A is +ve definite, then $\langle z, w \rangle = w^* A z$ defines a hermitian product on $\mathbb{C}^n$

* $(zA)^* = \bar{z} A^*$

* $A^* = -A$ ⇒ shew-hermitian

* let A be any $n \times n$ matrix

$$B = \frac{1}{2}(A + A^*)$$

$$C = \frac{1}{2i}(A - A^*)$$

$$\therefore A = B + iC$$

this helps when B & C commute

aka $BC = CB$

this gives the final condition

as : $A^* A = A A^*$

★ An $n \times n$ matrix is said to be normal if $A^* A = A A^*$

A is said to be normal iff $A = B + iC$

B, C — hermitian + commuting

★ Unitary : $A A^* = A^* A = I_n$

if A is real, $A^* = A^T$

$AA^T = A^T A = I$ → orthogonal

this gives the final condition

as: $$A^* A = A A^*$$

★ An $n \times n$ matrix is said to be normal if $A^* A = AA^*$

A is said to be normal iff $A = B + iC$

B, C – hermitian + commuting

★ Unitary: $AA^* = A^* A = I_n$

if A is real, $A^* = A^T$

$$AA^T = A^T A = I \text{ — orthogonal}$$

Eigenvalues and eigenvectors

$$Av = \lambda v$$

A : $n \times n$ matrix

λ : eigenvalue associated to eigenvector v

- A is a mapping from $R^n \to \mathbb{R}^n$ such that the output is collinear to the input

- $\begin{bmatrix} \cos \pi/3 & -\sin \pi/3 \\ \sin \pi/3 & \cos \pi/3 \end{bmatrix}$ → no eigenvector in R^3

  bc its a rotation transformation,

  collineation not possible

- if a non zero vector v lies in the null space of A:

$$Av = 0v$$

$0 \rightarrow$ eigenvalue

$\therefore$ if nullity $A > 0 \Rightarrow 0$ is an eigenvalue

- Given a $n \times n$ matrix A, suppose v is an eigenvector w/ eigenvalue λ

$$Av = \lambda v \quad ; \quad v \neq 0$$

$$(A - \lambda I)v = 0$$

$$\Rightarrow \quad \det(A - \lambda I) = 0$$

Conversely, if $\det(A - \lambda I) = 0$

nullity $(A - \lambda I) > 0$

$\exists \; v \neq 0$ s.t : $(A - \lambda I)v = 0$

$\Rightarrow$ A has eigenvector v with eigenvalue λ

Theorem: The eigenvalues of A are the roots of the polynomial

$\det(A - xI) = 0$ (the characteristic eqn)

- it is a polynomial of degree n
- may / may not have real roots
- may have repeated roots

- Given A, λ, to find v just solve $(A-\lambda I)v = 0$ using REF

<u>Theorem:</u> Let A be an $n \times n$ matrix and $\lambda_1, \lambda_2 \dots \lambda_k$ be distinct eigenvalues of A and $v_1, v_2 \dots v_k$ then $v_1 .. v_k$ are linearly independent

It follows that if a real matrix A has n distinct eigenvalues w/ corresponding eigenvectors $\{v_1 .. v_n\}$, it forms a basis of $\mathbb{R}^n$.

☆ A $(n \times n)$ real matrix A is said to be diagonalizable if there is a basis of $\mathbb{R}^n$ w/ eigenvectors of A

(n distinct eigenvalues $\Rightarrow$ diagonalisable)

But,

- diagonalisable $\not\Rightarrow$ n distinct ev's

  eg. $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\lambda = 1$

- Real matrix can have complex ev.

- eigenvector $\neq 0$

eigenvalue can be 0

<u>Nature of the characteristic eq^n:</u>

$$\det(xI - A) = 0$$

- leading coeff = 1
- lets see what we know ab. the coefficients

$$\det(xI - A) = x^n + c_1 x^{n-1} \dots c_n$$

- do some hax, you get:

$$(-1)^n \det A = c_n$$

$$c_{n-1} = \text{coeff of } x = (-1)^{n-1} \sum (n-1)\times(n-1) \text{ principal minors of } A$$

★ A $k \times k$ principal minor is a $(k \times k)$ subdet such that the diagonal of the subdet is a part of the main diagonal

$$c_{n-2} = \text{coeff of } x^2 = (-1)^{n-2} \sum (n-2)\times(n-2) \text{ pm's of } A$$

.. pattern goes on

- For a 3×3 matrix, CP is:

$$x^3 - (\text{Trace})x^2 + \left(\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} + \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} \right) x$$

$$-(\det A)$$

★ Sum of eigenvalues = Trace

Product of eigenvalues = Det

$\sum \lambda_i \lambda_j$ (sum taking 2 at a time) = $\sum$ (2×2) PM's

Imp: if A is a real matrix, complex ev's will appear in conjugate pairs!

Theorem: Suppose A is a matrix w/ ev's

$\lambda_1, \lambda_2 \ldots \lambda_n$ (counted w/ multiplicities) and

$g(A)$ is a polynomial in A

then $g(\lambda_1), g(\lambda_2) \ldots g(\lambda_n)$ is the list of ev's of $g(A)$

* ev's of a skew symm matrix are 0 or imaginary

Diagonalisability :

Let A, B be two $(n\times n)$ matrices

A, B are similar if $\exists$ a non singular matrix P such that $P^{-1}AP = B$

Suppose v is an eigenvector of A w/ ev λ

$$Av = APP^{-1}v$$

$$\parallel$$

$$\lambda v = \lambda PP^{-1}v$$

$$(P^{-1}AP)(P^{-1}v) = \lambda(P^{-1}v)$$

$$B(P^{-1}v) = \lambda(P^{-1}v)$$

$\Rightarrow$ $P^{-1}v$ is an eigenvector of B w/ same ev λ

Theorem: For a $(n\times n)$ matrix A, the foll. are equivalent

(i) $\exists$ P s.t

$$P^{-1}AP = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \ddots \lambda_n \end{bmatrix}$$ diagonal matrix

(ii) There is a basis of $\mathbb{R}^n$ consisting of eigenvectors of A

A matrix 'A' is said to be diagonalisable if it satisfies (one and hence both) of the following conditions

(i) A has n lin. indep eigenvectors.

(ii) ∃ a non sing. matrix s.t

$$P^{-1}AP = \text{diag}(\lambda_1, . \lambda_2 .. \lambda_n)$$

Corollary: If A has n distinct ev's, then it is diagonalisable

★ Eigenvecs corr. to diff ev's are LI.

★ The columns of P are eigenvecs of A. That is what results in the eigenvalue matrix. Note! the evectors must corr. to diff. evalues ..

Geometric & Algebraic multiplicity:

Let A be a (n×n) matrix

$$\det(xI - A) = (x-\lambda_1)^{k_1}(x-\lambda_2)^{k_2} \dots (x-\lambda_t)^{k_t}$$

$$k_j = \text{algebraic multiplicity of } \lambda_j$$

and nullity of $(A-\lambda_j I)$ is called the geometric multiplicity of λ_j

- ★ Nullity of A = GM of eigenvalue 0
- ★ AM of λ $\geq$ GM of λ
- ★ sum of AM's of all λ = n (degree of C·E)

Theorem:

Let A be an $(n \times n)$ matrix, the foll. are equivalent:

(i) A has n lin indep. eigenvectors

(ii) There exists non-sing P s.t $P^{-1}AP$ is diag.

(iii) For each e.v λ,

$$GM = AM$$

Cayley Hamilton Theorem:

Every matrix satisfies its own characteristic equation.

The Spectral Theorem:

Real case: A → real symm $(n \times n)$ matrix

$\exists$ orthonormal basis of $\mathbb{R}^n$ consisting of eigenvectors of A

equivalently, $\exists$ an orthogonal matrix O such that

$$O^T A O = \text{diag}(\lambda_1, \lambda_2 \ldots \lambda_n)$$

$\lambda_1, \lambda_2 .. \lambda_n$ are the eigenvalues of A and are all real.

Complex case : Let A be a $(n \times n)$ hermitian matrix

$\exists$ an orthonormal basis of $\mathbb{C}^n$

consisting of eigenvectors of A

equivalently, $\exists$ a unitary matrix U

such that

$$U^* A U = \text{diag}(\lambda_1, \lambda_2 \dots \lambda_n)$$

★ All the eigenvalues are real here as well.

orthonormality $\Rightarrow$ inner product = 0 ; $\mathbb{R}$

hermitian product = 0 ; $\mathbb{C}$

Lemma: The ev's of real symm matrix are real & its evecs lie in $\mathbb{R}^n$

" hermitian " are real & its evecs lie in $\mathbb{C}^n$

★ For a hermitian / real symm matrix A, eigenvectors corr. to different eigenvalues are perpendicular!

* Coda:

we found an orthonormal basis $v_1 \dots v_n$ of $\mathbb{C}^n$ consisting of eigenvectors of A

$$U = \left[\overset{v_1}{\left[\,\right]} \cdots \overset{v_n}{\left[\,\right]} \right]$$ is a unitary matrix

$$UU^* = U^*U = I$$

$$AU = \begin{bmatrix} Av_1 & , Av_2 & \cdots & Av_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= \begin{bmatrix} \lambda v_1 & , \lambda v_2 & \cdots & \lambda v_n \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

$$= [v_1 \;\; v_2 \; \cdots v_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}$$

$$AU = U \begin{bmatrix} \lambda_1 & \ddots & 0 \\ 0 & & \lambda_n \end{bmatrix}$$

$$U^*AU = \begin{bmatrix} \lambda_1 & \ddots & 0 \\ 0 & & \lambda_n \end{bmatrix}$$

Commuting matrices :

if A, B are $(n \times n)$ hermitian (or real symm) matrices

we know $\exists$ unitary matrices U and V s.t

$$U^*AU = \text{diag}(\lambda_1 \dots \lambda_n)$$

$$V^*BV = \text{diag}(\mu_1 \dots \mu_n) \qquad \lambda, \mu \text{ - ev's of } A \,\&\, B$$

if we want U and V to be the same matrix, then the necessary & sufficient condition is

$AB = BA$ .. they commute

Lemma 1: A, B are commuting hermitian matrices

λ is an ev. of A (w/ geom mult. k)

$\{v_1 .. v_k\}$ is an orthonormal set of eigenvectors of A all w/ value ev λ

then each of $Bv_1, Bv_2 \dots Bv_k$ is a linear combination of $v_1 \dots v_k$

Lemma 2: if A, B are commuting hermitian matrices

then $\exists$ a $v \neq 0$ s.t simultaneously

$$Av = \lambda_1 v$$

and

$$Bv = \mu_1 v$$

.. commuting hermitian matrices have a common eigenvector

quadrics & conics:

we have:

$$2x^2+2y^2-2xy=1$$

what shape is this? ellipse? hyperbola?

lets see how to find out:

it can be rewritten as:

$$[x\ \ y]\ \underbrace{\begin{bmatrix}2 & -1\\ -1 & 2\end{bmatrix}}_{A}\begin{bmatrix}x\\ y\end{bmatrix}=1$$

rewrite this equation now as:

$$[x\ \ y]\ P^T A P \begin{bmatrix}x\\ y\end{bmatrix}=1$$

what is P?

a matrix whose columns are eigenvectors of A

it rotates the conic, and now instead of the usual basis, we have a basis of eigenvectors which means we only have nice x^2, y^2, z^2 terms & no xy, yz stuff (xy, yz pops up when stuff other than elongation / shortening of basis happens but w/ eigenbasis – the basis vectors are only scaled by λ, making things simple)

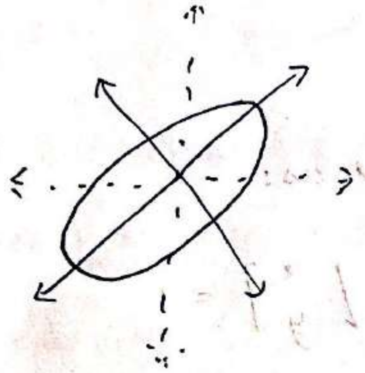
$P^T A P$ = diag (eigenvalues of A)

to form P, generate evecs of A and make them columns of a matrix

here, $P^T A P = \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}$

$[x\ y][P^T A P]\begin{bmatrix} x \\ y \end{bmatrix} \Rightarrow \boxed{x^2 + 3y^2 = 1}$

= ellipse



In new coord system: Major & Minor axes : $\hat{e}_1$ & $\hat{e}_2$

if you look at it w/ your head ~~straight~~ rotated

But, to find axes in unrotated XY plane, what you do is, (head straight)

apply P to $\hat{e}_1$, $\hat{e}_2$

$P\hat{e}_1$, $P\hat{e}_2$ are the axes

= v_1 , v_2 (eigenvectors) $\leftrightarrow$ (eigenbasis)

Any quadric

$$ax^2 + by^2 + cz^2 + 2hxy + 2gxz + 2fyz = 1$$

can be brought to the form

$$\lambda x^2 + \mu y^2 + \nu z^2 = 1$$

via an orthogonal change of coordinates

λ, μ, ν are ev's of $\begin{bmatrix} a & h & g \\ h & b & f \\ g & f & c \end{bmatrix}$

transformation matrix is O

$$O = \begin{bmatrix} v_1 & v_2 & v_3 \end{bmatrix}_{3\times 3}$$

v_1, v_2, v_3 - evecs of λ, μ, ν

$OO^* = O^*O = I_3$ (rotation of coordinates matrix)

Possible loci:

(i) All 3 are -ve

→ imaginary locus

→ no $\begin{bmatrix} x \\ y \\ z \end{bmatrix}$ in $\mathbb{R}^3$ satisfies $-x^2 - y^2 - z^2 = 1$ (eg.)

(ii) 3 positive .

→ Ellipsoid

$$\lambda x^2 + \mu y^2 + \nu z^2 \equiv \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2}$$

(iii) 2 positive , 1 negative

→ Hyperboloid of one sheet

(iii) 1 positive , 2 negative

→ Hyperboloid of two sheets

(iv) if one out of the three is zero , last 2 +ve

→ Cylinder

* Bessel's inequality & Parseval's formula:

. Let $S = \{w_1 .. w_k\}$ be an orthonormal set of vectors (unit) which span a k dim space in $\mathbb{R}^n$.

. now if you take a vector v, its components along each basis vectors of S, will account for v^T (the part of v which lies in that span)

$\therefore$ $\sum$ (component along w_i)^2

will be less than $\|v\|^2$

because the v_{normal} components were left out.

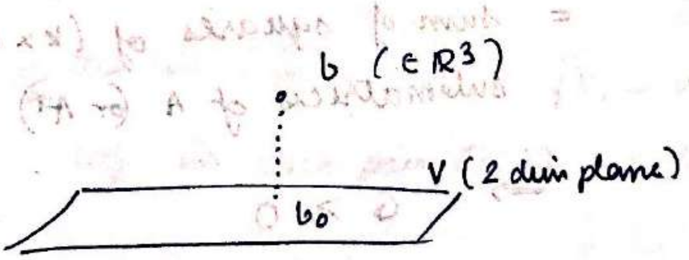
★ it is equal when v lies in the span of S. ($v^N = 0$)
⟨Parseval's formula⟩

★ Least square approximation :

V is a k dimensional subspace of $\mathbb{R}^n$; $b \in \mathbb{R}^n$

$v \in V$, find v_0 s.t $\|b - v\|^2$ is least when $v = v_0$

eg.



k = 2
n = 3

$b_0 = \sum \langle b, w_i \rangle w_i$ = the component of b in the span of $\{w_1, .. w_k\}$

| orthogonal unit vector basis of V

→ this works bc now,

$b - b_0 = b - b^T$
$= b_N \Rightarrow$ the shortest distance is along the ⊥ path

Cauchy - Binet Theorem:

- A is a $k \times n$ matrix
- B is a $n \times k$ matrix
- $\det(AB)$ = sum of $(k \times k)$ principal minors of BA

★ if AB is non diagonalisable / non invertible, perturb it (change the entries a bit) to make it diag. / invertible. Now make these perturbed matrices tend to the actual one, and apply limit to C-B formula, it will give the desired answer.

Cor: if $B = A^T$

$\det(AA^T)$ = grammian $(v_1 \cdots v_k)$

= sum of $(k \times k)$ principal minors of $A^T A$

= sum of squares of $(k \times k)$ submatrices of A (or A^T)

$\implies G \geq 0$

★ if $v_1 \dots v_k \in \mathbb{R}^n$

$G(v_1 \dots v_k) \leq \|v_1\|^2 \dots \|v_k\|^2$

equality holds <u>iff</u> either

- one of the v_j is 0

OR

- the vectors are mutually orthogonal

<u>Cor</u>: Hadamard's inequality

if A is a $(n \times n)$ matrix w/ real entries

$(\det A)^2 \leqslant \|v_1\|^2 \dots \|v_n\|^2$

$v_1 \dots v_n$ - columns of A

equality holds <u>iff</u> cols. of A are orthogonal

★ $G(v_1 \dots v_k \,;\, w_1 \dots w_l) \leq G(v_1 \dots v_k)\, G(w_1 \dots w_l)$

equality holds if either

(i) one of the sets $\{w_1 \dots w_l\}$ or $\{v_1 \dots v_k\}$ lin dep.

(ii) $v_i \perp w_j$ for each pair (i, j)

$1 \leq i \leq k$

$1 \leq j \leq l$