Enduring Value

classmate

Nature's Wonder

ARCTIC FOX

soorajss1729@gmail.com +91-9400635788

INDEX

23

NAME: SOORAJ.S. STD.: SEC.: ROLL NO.: SUB.:

S. No.	Date	Title	Page No.	Teacher's Sign / Remarks
		OBJECTIVE MATHEMATICS		
		-O.P. ARORA		

$\boxed{19}$ MATRICES & DETERMINANTS

For real column vectors $a, b \in \mathbb{R}^n$, $\boxed{tr(ba^T) = tr(ab^T) = \vec{a}\cdot\vec{b}}$

Trace of a matrix

The trace of a square matrix A_{nn} is defined to be the sum of elements on the main diagonal of A.

Ex:-

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 5 & 7 \\ 9 & 4 & 3 \end{bmatrix} ; \; tr\,A = 1+5+3 = 9$$

- The trace is only defined for a square matrix.

$$\boxed{tr(A) = \sum_{i=1}^{n} a_{ii}}$$

$$tr(A+B) = tr(A) + tr(B)$$

$$tr(kA) = k \cdot tr(A)$$

$$tr(A^T) = tr(A)$$

$$\boxed{tr(ABC) = tr(BCA) = tr(CAB)}$$

cyclic property

$$tr(AB) = tr(BA)$$

$$tr(AB) \neq tr(A) \cdot tr(B)$$

* The determinant of A is equal to the product of its eigen values, ie.,

$$\boxed{\det A = |A| = \prod_{i=1}^{n} \lambda_i}$$

Proof

- Consider the linear transformation of n-D vectors defined by an $n \times n$ matrix A,

$$A\vec{v} = \vec{w}$$

(OR)

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

If v & w are scalar multiples,

i.e., if $A\vec{v} = \vec{w} = \lambda\vec{v}$; $\vec{v} \neq \vec{0}$ ——— (1)

$\rightarrow$ $\vec{v}$: eigen vector of the linear transformation A

λ: eigenvalue corresp. to that eigen vector, $\vec{v}$

Eqn (1): Eigen equation for the matrix A

$$(A-\lambda I)\vec{v} = \vec{0} \quad \text{——— (2)}$$

Eqn (2) has a non-zero solution $\vec{v}$ iff $|A-\lambda I| = 0$

ie, $(A-\lambda I)$ is singular

$(A-\lambda I)^{-1}$ does not exist

$\therefore$ The eigen values of A are values of λ that satisfy the equation

$$|A-\lambda I| = 0$$

(OR)

$$\begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix} = 0$$

Leibniz rule for determinants

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

$\Rightarrow$ $|A-\lambda I|$ is a polynomial function of the variable λ ~~and~~ of the degree n.

Its term of degree 'n' is always $(-1)^n \lambda^n$

- The fundamental theorem of algebra implies that the characteristic polynomial of an $n \times n$ matrix A, being a polynomial of degree n, can be factored into the product of n linear terms.

$$|A-\lambda I| = \boxed{(-1)^n} \times (\lambda-\lambda_1)(\lambda-\lambda_2) \cdots\cdots (\lambda-\lambda_n)$$

$$= \det(A-\lambda I)$$

$f(\lambda) = |A - \lambda I|$

$= \det(A-\lambda I) = (-1)^n * (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

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

$= f(0) = (-1)^n \times (0-\lambda_1)(0-\lambda_2)\cdots(0-\lambda_n)$

$= (-1)^n * (-1)^n \times \prod_{i=1}^{n} \lambda_i$

$= (-1)^{2n} * \prod_{i=1}^{n} \lambda_i = \prod_{i=1}^{n} \lambda_i$

$$\Longrightarrow \det A = \prod_{i=1}^{n} \lambda_i$$

$f(\lambda) = \det(A - \lambda I) = (-1)^n * (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots(\lambda-\lambda_n)$

$= (-1)^n * \Big[\lambda^n - (\lambda_1+\lambda_2+\cdots+\lambda_n)\lambda^{n-1} + (\lambda_1\lambda_2+\lambda_1\lambda_3+\cdots+\lambda_{n-1}\lambda_n)\lambda^{n-2}$

$\cdots\cdots (-1)^n \lambda_1\lambda_2\cdots\lambda_n\Big]$

$= (-1)^n * \Big[\lambda^n - \lambda^{n-1}\sum_i \lambda_i + \lambda^{n-2}\sum_{i<j} \lambda_i\lambda_j - \cdots (-1)^n \prod_{i=1}^{n} \lambda_i$

$= (-1)^n * \Big[\lambda^n - \lambda^{n-1} \cdot tr(A) + \cdots\cdots (-1)^n \cdot \det(A)$

* The trace of a matrix (square) A is equal to the sum of its eigen values

$$\boxed{\begin{aligned} tr(A) &= \sum_{i=1}^{n} a_{ii} \\ &= \sum_{i=1}^{n} \lambda_i \end{aligned}}$$

$$tr(A) = tr\left(\sum_i \lambda_i |\lambda_i\rangle\langle\lambda_i|\right)$$
$$= \sum_i \lambda_i \, tr(|\lambda_i\rangle\langle\lambda_i|)$$
$$= \sum_i \lambda_i \langle\lambda_i|\lambda_i\rangle$$

$$\boxed{\begin{aligned} tr(A) &= \sum_i \langle\lambda_i|A|\lambda_i\rangle \\ &= \sum_i \lambda_i \end{aligned}}$$

Proof

$$\det(A-\lambda I) = (-1)^n (\lambda-\lambda_1)(\lambda-\lambda_2)\cdots\cdots(\lambda-\lambda_n)$$

$$\sum_{i=1}^{n} \lambda_i = \text{coefficient of } (-1)^{n-1} \lambda^{n-1}$$

Leibniz rule for determinants

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

To get the power λ^{n-1} for a permutation, we need ~~atleast~~ $(n-1)$ diagonal elements.

ie; $\sigma(i) = i$ for $(n-1)$ values of i

But,
once we know the value of a permutation on $(n-1)$ inputs, we know the last one too.

$\therefore$ Only identity permutations are considered. to get the coeff. of λ^{n-1}.

ie., $f(\lambda) = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda)$

$f(\lambda) = (-1)^n * (\lambda - a_{11})(\lambda - a_{22})\cdots\cdots(\lambda - a_{nn})$

Coeff. of $(-1)^{n-1}\lambda^{n-1}$ is : $\sum_{i=1}^{n} a_{ii}$

$\Longrightarrow$

$$Tr(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$$

$$= \text{coeff. of } (-1)^{n-1}\lambda^{n-1}$$

$$\boxed{tr(AB) = tr(BA)}$$

Proof

$A = [a_{ij}]_{n\times n}$; $B = [b_{ij}]_{n\times n}$

$$(ab)_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}$$

$$(ab)_{ii} = \sum_{k=1}^{n} a_{ik} b_{ki} \; ; \; (ba)_{ii} = \sum b_{ik} a_{ki}$$

$$Tr(AB) = \sum_{i=1}^{n} (ab)_{ii}$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} a_{ik} b_{ki} = \sum_{m,n=1}^{n} a_{m,n} b_{n,m}$$

$$Tr(BA) = \sum_{i=1}^{n} (ba)_{ii}$$

$$= \sum_{i=1}^{n} \sum_{k=1}^{n} b_{ik} a_{ki} = \sum_{m,n=1}^{n} a_{m,n} b_{n,m}$$

$$\longrightarrow Tr(AB) = Tr(B, A)$$

The trace corresp. to the derivative of the determinant: it is the Lie algebra analog of the (Lie group) map of the determinant.

Jacobi's formula for the derivative of the determinant

$$\frac{d}{dt}\det A = \det A \cdot tr\left(A^{-1}\frac{dA}{dt}\right)$$

$$= tr\left(\operatorname{adj} A \cdot \frac{dA}{dt}\right)$$

17/6/2020

* the determinant is a multilinear function of its rows.

Proof

$$\frac{d}{dt}\det A = \frac{d}{dt}\begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}$$

$$= \begin{vmatrix} \dot{a}_{11} & \dot{a}_{12} & \cdots & \dot{a}_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \dot{a}_{21} & \dot{a}_{22} & \cdots & \dot{a}_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \dot{a}_{n1} & \dot{a}_{n2} & \cdots & \dot{a}_{nn} \end{vmatrix}$$

* This eq. requires the computation of n determinants for the computation of a single derivative

If $A(t) = I$,

$$\frac{d}{dt} A(t) = \begin{vmatrix} \dot a_{11} & \dot a_{12} & \cdots & \dot a_{1n} \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} + \begin{vmatrix} 1 & 0 & \cdots & 0 \\ \dot a_{21} & \dot a_{22} & \cdots & \dot a_{2n} \\ \vdots & & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{vmatrix} + \cdots + \begin{vmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ \dot a_{n1} & \dot a_{n2} & \cdots & \dot a_{nn} \end{vmatrix}$$

$$= \dot a_{11} + \dot a_{22} + \cdots + \dot a_{nn} = Tr(\dot A(t))$$

$$\left.\frac{d}{dt} \det A(t) \right| = tr\left(\frac{d}{dt} A(t)\right) \quad \text{when } A(t) = I$$

When $A(t)$ is invertible at t_0,

let $\psi(t) = A^{-1}(t_0) A(t)$, then $\psi(t_0) = I$.

$$\frac{d}{dt} \det A(t) \Big|_{t=t_0} = \frac{d}{dt} \det \left(A(t_0) \psi(t) \right) \Big|_{t=t_0}$$

$$= \frac{d}{dt} \left[\det(A(t_0)) \det \psi(t) \right] \Big|_{t=t_0}$$

$$= \underline{\det A(t_0) \frac{d}{dt} \det \psi(t) \Big|_{t=t_0}},$$

$$= \det(A(t_0)) \cdot tr \left(\frac{d}{dt} \psi(t) \right) \Big|_{t=t_0}$$

$$= \det(A(t_0)) \cdot tr \left(\frac{d}{dt} A^{-1}(t_0) A(t) \right) \Big|_{t=t_0}$$

$$= \det(A(t_0)) \cdot tr \left(A^{-1}(t_0) \frac{d}{dt} A(t) \Big|_{t=t_0} \right)$$

$$\frac{d}{dt} \det A(t) \Big|_{t=t_0} = \det A(t_0) \cdot tr \left(A^{-1}(t_0) \cdot \dot{A}(t_0) \right)$$

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

$$= tr \left(\operatorname{adj}(A(t_0)) \cdot \dot{A}(t_0) \right)$$

Crammer's rule

Consider a system of n linear equations for n unknowns

$$A\vec{x} = \vec{b} \quad \& \quad |A| = \det A \neq 0$$

$$x_i = \frac{\det(A_i)}{\det(A)} = \frac{\Delta_i}{\Delta} \; ; \quad i = 1, 2, \dots, n$$

where, $\det A = \Delta$: determinant of the matrix formed by replacing the i^{th} column of 'A' by the column vector $\vec{b}$.

Proof

$$A\vec{x} = \vec{b} \implies \vec{x} = A^{-1}\vec{b} = \frac{\operatorname{adj} A}{|A|}\vec{b}$$

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1n} \\ A_{21} & A_{22} & \cdots & A_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ A_{n1} & A_{n2} & \cdots & A_{nn} \end{bmatrix}\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = \frac{1}{|A|}\begin{bmatrix} b_1 A_{11} + b_2 A_{12} \cdots + b_n A_{1n} \\ b_1 A_{21} + b_2 A_{22} \cdots + b_n A_{2n} \\ \vdots \\ b_1 A_{n1} + b_2 A_{n2} \cdots + b_n A_{nn} \end{bmatrix}$$

$$x_k = \frac{\sum_{i=1}^{n} b_i A_{ki}}{|A|}$$

$$= \frac{b_1 A_{k1} + b_2 A_{k2} \cdots + b_n A_{kn}}{\det A} = \frac{\det A_i^T}{\det A}$$

$$= \frac{\det A_i}{\det A}$$

Nilpotent Matrix

In linear algebra, a nilpotent matrix is a square matrix A such that

$$A^k = 0 \text{ for some +ve integer } k.$$

If k is the least +ve integer for which $A^k = 0$, then 'A' is said to be nilpotent of index k.

More generally,
a nilpotent transformation is a linear transformation L of a vector space such that $L^k = 0$ for some +ve integer k (and thus, $L^j = 0$ for all $j \geq k$).

* any triangular matrix with zeros along the main diagonal is nilpotent with index $\leq n$

Ex.

$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$, $A^2 = 0$

$B = \begin{bmatrix} 0 & 2 & 1 & 6 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{bmatrix}$, $B^4 = 0$

$C = \begin{bmatrix} 5 & -3 & 2 \\ 15 & -9 & 6 \\ 10 & -6 & 4 \end{bmatrix}$; $C^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

* The determinant & trace of a nilpotent matrix are always zero.

Proof

(i) $A^k = 0$, $k \in \mathbb{Z}^+$

$$|A^k| = 0 = |A|^k \longrightarrow |A| = 0.$$

(ii) $tr(A) = \sum_{i=1}^{n} a_{ii} = \sum_{i=1}^{n} \lambda_i$

Let λ is an eigen value of 'A'. & let $\vec{v} \neq 0$ such that

$$A\vec{v} = \lambda\vec{v}$$

$$A^k = 0 \longrightarrow A^k\vec{v} = \lambda^k\vec{v} = \vec{0}$$

$$\vec{v} \neq 0 \longrightarrow \lambda^k = 0 \Rightarrow \lambda = 0$$

$\therefore$ all eigenvalues of 'A' are zero.

$$\therefore tr(A) = \sum_i \lambda_i = 0$$

* The index of an $n \times n$ nilpotent matrix is always less than or equal to n.

$$k \le n$$

proof

?

$$A^k = 0 \implies \lambda_i = 0$$

$$f(\lambda) = (-1)^n \lambda^n = 0$$

Cayley-Hamilton theorem $\implies f(A) = 0 = (-1)^n A^n$

$$\implies A^n = 0$$

$$\implies A^k = 0 \quad \text{for } k \le n.$$

Cayley-Hamilton theorem

In linear algebra, the Cayley-Hamilton theorem states that every square real or complex matrices satisfy its own characteristic equation.

If the characteristic polynomial (equation) of an $n \times n$ matrix 'A' is :

$$f(\lambda) = \det(A - \lambda I) = (-1)^n (\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n)$$

$$= a_n \lambda^n + a_{n-1} \lambda^{n-1} + \cdots + a_0 = 0$$

then,

$$\boxed{f(A) = a_n A^n + a_{n-1} A^{n-1} + \cdots + a_0 I = 0}$$

Proof - for diagonalizable matrices but holds for but theorem holds for all sq. matrices.

If a vector $\vec{v}$ of size n is an eigenvector of A with eigenvalue λ, $A\vec{v} = \lambda\vec{v}$, then

$$f(A) \cdot \vec{v} = a_n \cdot A^n \cdot \vec{v} + a_{n-1} A^{n-1} \cdot \vec{v} + \cdots + a_1 A \cdot \vec{v} + a_0 I_n \vec{v}$$

$$= a_n \cdot \lambda^n \vec{v} + a_{n-1} \lambda^{n-1} \vec{v} + \cdots + a_1 \lambda \vec{v} + a_0 \vec{v}$$

$$= f(\lambda) \cdot \vec{v} = 0$$

$P(A)\cdot\vec{v} = 0$ for every eigenvector $\vec{v}$

For a diagonal matrix, $A = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

$$f(A) = (-1)^n (A-\lambda_1 I)(A-\lambda_2 I)\cdots(A-\lambda_n I)$$

$$= (-1)^n \operatorname{dig}(0, \lambda_2-\lambda_1, \ldots, \lambda_n-\lambda_1)\operatorname{diag}(\lambda_1-\lambda_2, 0, \ldots, \lambda_n-\lambda_2)$$

$$\ldots\ldots \operatorname{diag}(\lambda_1-\lambda_n, \lambda_2-\lambda_n, \ldots, 0)$$

$$= 0$$

For a diagonalizable matrix A,

i.e., $A = XDX^{-1}$, $D = \operatorname{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$

$$f(A) = f(XDX^{-1})$$

$$= \sum_{k=0}^{n} a_k A^k = \sum_{k=0}^{n} a_k (XDX^{-1})^k$$

$$(XDX^{-1})^k = XD\cancel{X^{-1}}\cdot\cancel{X}D\cancel{X^{-1}}\cdot\cancel{X}D\cancel{X^{-1}}\cdots\cancel{X}D\cancel{X^{-1}}\cdot\cancel{X}DX^{-1}$$

$$= XD^kX$$

$$= \sum_{k=0}^{n} a_k \cdot XD^kX^{-1}$$

$$= X\left(\sum_{k=0}^{n} a_k D^k\right) X^{-1} = X \, 0 \, X^{-1} = 0.$$

Proof

Schur's theorem : Every square matrix is similar to a upper triangular matrix

ILA②

$$A = BTB^{-1}$$

$$T = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} = B^{-1}AB$$

$$\prod_{i=1}^{n}(T-\lambda_i I) = (T-\lambda_1 I)(T-\lambda_2 I)\cdots(T-\lambda_n I)$$

$$= \begin{bmatrix} 0 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \cdots \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix}$$

We can see that

$$T-\lambda_1 I = \begin{bmatrix} 0 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$ has 1^{st} one column all zero.

$$(T-\lambda_1 I)(T-\lambda_2 I) = \begin{bmatrix} 0 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix} = \begin{bmatrix} 0 & 0 & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

has 1^{st} two columns all zero

and so on. Thus we can prove that

$$\prod_{i=1}^{n}(T-\lambda I) = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 0 \end{bmatrix} = 0$$, has all n column zero

→ All triangular matrices satisfy Cayley-Hamilton theorem.

Schur's theorem ILA(8) → $A = BTB^{-1}$

$$\prod_{i=1}^{n}(A-\lambda I) = \prod_{i=1}^{n}\left(BTB^{-1} - \lambda BB^{-1}\right)$$

$$= B\left[\prod_{i=1}^{n}(T-\lambda I)\right]B^{-1} = B0B^{-1} = 0$$

→ All square matrices satisfy the Cayley-Hamilton theorem.

Diagonalization of Matrices

A square matrix A is called diagonizable or non-defective if it is similar to a diagonal matrix i.e., if there exists an invertible matrix P & a diagonal matrix D such that $P^{-1}AP = D$

(or) $$A = PDP^{-1}$$

An $n \times n$ square matrix A is diagonalizable iff A has n linearly independent eigenvectors.

In fact, $A = PDP^{-1}$, with D a diagonal matrix, iff the columns of P are n linearly independent eigenvectors of A. In this case, the diagonal entries of D are eigenvalues of A that corresp. respectively, to the eigenvectors in P.

* Diagonalization is the process of finding the above P and D.

$$A\left[\vec{v}_1\,\vec{v}_2\cdots\vec{v}_n\right]=\left[\vec{v}_1\,\vec{v}_2\cdots\vec{v}_n\right]\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}$$

If $\left[\vec{v}_1\,\vec{v}_2\cdots\vec{v}_n\right]$ is invertible,

$$A=\left[\vec{v}_1\,\vec{v}_2\cdots\vec{v}_n\right]\begin{bmatrix}\lambda_1 & 0 & \cdots & 0\\ 0 & \lambda_2 & \cdots & 0\\ \vdots & & \ddots & \vdots\\ 0 & 0 & \cdots & \lambda_n\end{bmatrix}\left[\vec{v}_1\,\vec{v}_2\cdots\vec{v}_n\right]^{-1}$$

$$\therefore\ A = PDP^{-1}$$

* Matrix 'A' is similar to matrix 'B' if there exists an invertible matrix P such that $A = P^{-1}BP$.

* A matrix is diagonalizable iff it is similar to a diagonal matrix.

An $n\times n$ matrix A is diagonalizable iff it has n linearly independent eigen vectors.

Ex:-

Q. Diagonalize $A = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}$

Ans:

$$AX = \lambda X \implies \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} \lambda x_1 \\ \lambda x_2 \end{bmatrix}$$

$$(A-\lambda I)X = \begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

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

$$(1-\lambda)x_1 + x_2 = 0$$
$$-2x_1 + (4-\lambda)x_2 = 0$$

$$|A-\lambda I| = 0 \implies (1-\lambda)(4-\lambda) + 2 = 0$$
$$\lambda^2 - 5\lambda + 6 = (\lambda-3)(\lambda-2) = 0$$

$\lambda = 3$ (or) 2

$\lambda = 2$ $\left\{ \begin{array}{l} -x_1 + x_2 = 0 \\ 2x_1 + 2x_2 = 0 \end{array} \right.$

$\implies x_2 = x_1$

$\vec{v}_1 = k\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \text{span}\left(\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)$

$\lambda = 3$ $\left\{ \begin{array}{l} -2x_1 + x_2 = 0 \\ -2x_1 + x_2 = 0 \end{array} \right.$

$\implies x_2 = 2x_1$

$\vec{v}_2 = k\begin{bmatrix} 1 \\ 2 \end{bmatrix} = \text{span}\left(\begin{bmatrix} 1 \\ 2 \end{bmatrix}\right)$

$$D = \begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$$

$$P = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\cancel{PP^{-1}P = PI} \Rightarrow \cancel{\begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}}$$

$$P = PI = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{array}\right] \longrightarrow \left[\begin{array}{cc|cc} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{array}\right]$$

$$\left[\begin{array}{cc|cc} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \end{array}\right] \Longrightarrow P^{-1} = \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 \\ -2 & 4 \end{bmatrix} = A = PDP^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix}$$

* Similarity is an equivalence relation on the space of square matrices. Say A, B, C are square matrices of order n.

(i) A is similar to A (reflexive)

(ii) If A is similar to B then B is similar to A (symmetric)

(iii) If A is similar to B & B is similar to C, then A is similar to C (transitive)

Proof

(i) I_n is a non-singular matrix such that

$$I_n^{-1} A I_n = I_n A I_n = A$$

$\therefore$ A is similar to A

(ii) Assume that A is similar to B, there is a non-singular matrix S so that

$A = S^{-1} B S$

S^{-1} is invertible

$$(S^{-1})^{-1} A (S^{-1}) = S A S^{-1} = S(S^{-1} B S) S^{-1}$$

$$= I_n B I_n = B$$

$\therefore$ B is similar to A

(ii) A is similar to B & B is similar to C

$A = S^{-1}BS$ & $B = R^{-1}CR$

$A = S^{-1}BS = S^{-1}(R^{-1}CR)S$

$= (S^{-1}R^{-1})C\,RS$

$= (RS)^{-1}C(RS)$

$\therefore$ A is similar to C. via the non-singular matrix RS.

---

* Similar matrices have equal eigenvalues.

ie;

If A & B are similar matrices, then the characteristic polynomials of A and B are equal.

$P_A(\lambda) = P_B(\lambda)$

Proof $P_A(\lambda) = \det(A - \lambda I_n) = \det(S^{-1}BS - \lambda S^{-1}I_nS)$

$= \det(S^{-1}(B-\lambda I_n)S) = \det(S^{-1})\det(B-\lambda I_n)\det(S)$

$= \det(S^{-1})\det(S)\det(B-\lambda I_n) = \det(S^{-1}S)\det(B-\lambda I_n)$

$= 1\cdot\det(B-\lambda I_n) = P_B(\lambda)$

* Every ...

(Same ...

Jord...

ILA (15)

8.3 (B)

* Every square matrix is similar to its transpose (same eigenvalues, same multiplicity, same Jordan form).

ILA (5)

8.3 (2)

* Eigenvectors corresponding to different eigenvalues must be linearly independent.

Proof.

By induction.

true for $k=1$

For $k=2$, $c_1\vec{v}_1 + c_2\vec{v}_2 = 0$

$$c_1A\vec{v}_1 + c_2A\vec{v}_2 = c_1\lambda_1\vec{v}_1 + c_2\lambda_2\vec{v}_2 = 0$$

$$c_1\lambda_1\vec{v}_1 + c_2\lambda_1\vec{v}_2 = 0$$

$$\cancel{c_2}\vec{v}_1 + (\lambda_2-\lambda_1)c_2\vec{v}_2 = 0$$

$\lambda_2 \neq \lambda_1$ ~~&~~ $\vec{v}_2 \neq 0 \longrightarrow c_2 = 0 \Rightarrow c_1 = 0.$

$\therefore$ $\vec{v}_1$ & $\vec{v}_2$ are linearly independent

For $k \geq 2$, $c_1\vec{v}_1 + \cdots + c_k\vec{v}_k = 0$

$$c_1A\vec{v}_1 + c_2A\vec{v}_2 + \cdots + c_kA\vec{v}_k = c_1\lambda_1\vec{v}_1 + \cdots + c_k\lambda_k\vec{v}_k = 0$$

$$c_1\lambda_1\vec{v}_1 + \cdots + c_k\lambda_1\vec{v}_k = 0$$

$$(\lambda_2-\lambda_1)c_2\vec{v}_2 + (\lambda_3-\lambda_1)c_3\vec{v}_3 + \cdots + (\lambda_k-\lambda_1)c_k\vec{v}_k = 0$$

By induction,

$\vec{v}_2, \vec{v}_3, \ldots, \vec{v}_k$ are linearly independent.

$\therefore$

$$(\lambda_j - \lambda_1) c_j = 0 \quad \text{for } j = 2, 3, \ldots, k.$$

$$\lambda_1 \neq \lambda_j \text{ for } j = 2, 3, \ldots, k \longrightarrow c_j = 0 \text{ for } j = 2, 3, \ldots, k$$

$$\therefore \ c_1 = 0$$

$\therefore \vec{v}_1, \vec{v}_2, \ldots, \vec{v}_k$ are linearly independent

* Orthogonal nonzero vectors are linearly independent whereas,

linearly independent vectors need not have to be orthogonal.

$$\boxed{\begin{array}{l} \text{orthogonal (non zero)} \Longrightarrow \text{linearly independent} \\ \text{Linearly independent} \not\Longrightarrow \text{orthogonal} \end{array}}$$

Proof

ANNWAYS ILA (6)

$$c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k = \vec{0}$$

$$0 = \vec{v}_i \cdot \vec{0} = \vec{v}_i \cdot (c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_k\vec{v}_k) = \vec{0}$$

$$\vec{v}_i \cdot \vec{v}_j = 0 \quad \text{if } i \neq j$$

$$c_1\vec{v}_i\cdot\vec{v}_1 + c_2\vec{v}_i\cdot\vec{v}_2 + \cdots + c_k\vec{v}_i\cdot\vec{v}_k = \vec{0}$$

all terms but the ith one are zero,

$$\therefore \quad c_i\vec{v}_i\cdot\vec{v}_i = c_i|\vec{v}_i|^2 = 0$$

$$\vec{v}_i \neq \vec{0} \Longrightarrow c_i = 0$$

This computation holds for every $i = 1, 2, \ldots, k$

$$\therefore \quad c_1 = c_2 = \cdots = c_k = 0$$

$\therefore$ $v_1, v_2, \ldots, v_k$ are linearly independent.

Zero vector is linearly dependent

Recall → • To be a basis of a vector space, all that is required is that the n vectors are independent.

Linearly independence of vectors need not imply orthogonality.

Ex:

1. $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \vec{u} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ ie. $\vec{v}$ is not a scalar multiple of $\vec{u}$, $\vec{v}$ is not ||el to $\vec{u}$

$\vec{v} \cdot \vec{u} \neq 0$ & $\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} = 1 \neq 0$

linearly independent but not orthogonal

2. $\vec{v} = \begin{bmatrix} 2 \\ 5 \end{bmatrix}, \vec{u} = \begin{bmatrix} -1 \\ 3 \end{bmatrix}$

$\vec{v} \cdot \vec{u} = -2 + 15 \neq 0$ & $\begin{vmatrix} 2 & -1 \\ 5 & 3 \end{vmatrix} = 6 + 5 = 11 \neq 0$

linearly independent but not orthogonal

3. $\vec{v} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, u = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\vec{v} \cdot \vec{u} = 2 \neq 0$ & $\begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = 1 \neq 0$

linearly independent but not orthogonal.

4. $\vec{u} =$

$\vec{u} \cdot \vec{w} =$
$\vec{u} \cdot \vec{w} = 1$
$\vec{v} \cdot \vec{w} = 0$

linearly

4. $\vec{u} = \begin{bmatrix}1\\0\\0\end{bmatrix}, \vec{v} = \begin{bmatrix}1\\1\\0\end{bmatrix}, \vec{w} = \begin{bmatrix}1\\1\\1\end{bmatrix}$

$$\left.\begin{array}{l}\vec{u}\cdot\vec{v} = 1 \neq 0 \\ \vec{u}\cdot\vec{w} = 1 \neq 0 \\ \vec{v}\cdot\vec{w} = 2 \neq 0\end{array}\right\} \text{linearly independent} \quad \& \quad \begin{vmatrix}1 & 1 & 1\\0 & 1 & 1\\0 & 0 & 1\end{vmatrix} = \begin{vmatrix}1 & 1\\0 & 1\end{vmatrix} = 1 \neq 0$$

linearly independent, but not orthogonal.

* Eigenvalues of A^T are the same as the eigenvalues of A

Proof

The characteristic polynomials

$$P_A(\lambda) = \det(A - \lambda I) \quad \& \quad P_{A^T}(\lambda) = \det(A^T - \lambda I)$$

$$P_{A^T}(\lambda) = \det(A^T - \lambda I) = \det(A^T - \lambda I^T)$$

$$= \det((A - \lambda I)^T) = \det(A - \lambda I) \quad [\because \det(A^T) = \det A]$$

$$= P_A(\lambda)$$

$\therefore$ eigenvalues of A & A^T are the same.

* If the eigenvalues of an $n \times n$ invertible matrix 'A' are $\lambda_1, \lambda_2, \dots, \lambda_n$, then the eigenvalues of A^{-1} are $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \dots, \frac{1}{\lambda_n}$.

Proof

eigenv...

Proof

If λ is an eigenvalue of A with eigenvector $\vec{v}$

$$A\vec{v} = \lambda\vec{v}$$

$$\underbrace{A^{-1}A}_{I}\vec{v} = \lambda A^{-1}\vec{v} \implies \underline{\underline{A^{-1}\vec{v} = \frac{1}{\lambda}\vec{v}}}$$

$\implies$ $\vec{v}$ is also an eigenvector of A^{-1} with eigenvalue $\frac{1}{\lambda}$.

*

AB & BA have the same eigenvalues

Proof Let A & B be square matrices of same order. If one of them is invertible, say A is invertible then

$$A^{-1}(AB)A = BA$$

$\therefore$ AB & BA are similar. $\Longrightarrow$ same eigenvalues.

Suppose none of them is invertible.

Define 2 matrices C & D of order $n \times n$ as follows

$$C = \left[\begin{array}{c|c} xI_n & A \\ \hline B & I_n \end{array}\right] \text{ and } D = \left[\begin{array}{c|c} I_n & 0 \\ \hline -B & xI_n \end{array}\right].$$

where x is an indeterminate.

$$\det\begin{bmatrix} A & B \\ C & D \end{bmatrix} = \det A \cdot \det S = \det(A) \cdot \det(D - CA^{-1}B)$$

$$\det\begin{bmatrix} A & 0 \\ C & D \end{bmatrix} = \det(A) \cdot \det(D)$$

Using that,

$$\cancel{\det(CD) = \det(C)\det(D) = \det(xI_n)\det(\ldots)}$$

$$\det(CD) = \det\begin{bmatrix} xI_n - AB & xA \\ 0 & xI_n \end{bmatrix}$$

$$= \det(xI_n - AB)\det(xI_n)$$

$$= x^n \det(xI_n - AB)$$

$$\det(DC) = \det\left(\begin{bmatrix} I_n & 0 \\ -B & xI_n \end{bmatrix}\begin{bmatrix} xI_n & A \\ B & I_n \end{bmatrix}\right)$$

$$= \det\begin{bmatrix} xI_n & A \\ 0 & xI_n - BA \end{bmatrix} = \det(xI_n)\det(xI_n - BA)$$

$$= x^n \det(xI_n - BA)$$

$$\det(CD) = \det(DC) \Longrightarrow \underline{\det(xI_n - AB) = \det(xI_n - BA)}$$

The characteristic polynomials of AB & BA are same.

Triangular Matrix.

A square matrix is called lower triangular if all the entries above the main diagonal are zero.

Similarly,
a square matrix is called upper triangular if all the entries below the main diagonal are zero.

* A matrix that is both upper and lower triangular is called a diagonal matrix.

Lower (left) triangular matrix,

$$L = \begin{bmatrix} l_{11} & 0 & 0 & \cdots & 0 \\ l_{21} & l_{22} & 0 & \cdots & 0 \\ l_{31} & l_{32} & l_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ l_{n1} & l_{n2} & l_{n3} & \cdots & l_{nn} \end{bmatrix}$$

Upper (Right) triangular matrix

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} & \cdots & u_{1n} \\ 0 & u_{22} & u_{23} & \cdots & u_{2n} \\ 0 & 0 & u_{33} & \cdots & u_{3n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & u_{nn} \end{bmatrix}$$

Upper triangularity is preserved by many operations:

(i) The sum of 2 upper triangular matrices is upper triangular

(ii) The product of 2 upper triangular matrices is upper triangular

(iii) The inverse of an upper triangular matrix

(iv) The product of an upper triangular matrix and a scalar is upper triangular

Similarly for lower triangular matrix.

* Eigenvalues of a triangular (upper or lower) matrix are the diagonal elements of the matrix

⟹ The determinant and permanent of a triangular matrix equals the product of the diagonal entries.

A is triangular ⟹ $\lambda_i = a_{ii} \quad , i = 1, 2, \dots n$

$$\det(A) = \text{perm}(A) = \prod_{i=1}^{n} a_{ii} = \prod_{i=1}^{n} \lambda_i$$

$$\text{perm}(A) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{\sigma \in S_n} a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

Proof

$$\det(A) = \sum_{\sigma \in S_n} \text{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} = \sum_{\sigma \in S_n} \text{sgn}(\sigma) * a_{1,\sigma(1)} a_{2,\sigma(2)} \cdots a_{n,\sigma(n)}$$

$$A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}$$

$$\det(A - \lambda I) = (a_{11} - \lambda)(a_{22} - \lambda) \cdots (a_{nn} - \lambda) = 0$$

$$\lambda \in \{a_{11}, a_{22}, \dots, a_{nn}\}.$$

* A matrix which is simultaneously triangular and normal is also diagonal.

<u>Proof</u>.

$$A = \begin{bmatrix} a_{11} & 0 & 0 & \cdots & 0 \\ a_{21} & a_{22} & 0 & \cdots & 0 \\ a_{31} & a_{32} & a_{33} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \cdots & a_{nn} \end{bmatrix}, \quad A^{\dagger} = \begin{bmatrix} \bar{a}_{11} & \bar{a}_{21} & \bar{a}_{31} & \cdots & \bar{a}_{n1} \\ 0 & \bar{a}_{22} & \bar{a}_{32} & \cdots & \bar{a}_{n2} \\ 0 & 0 & \bar{a}_{33} & \cdots & \bar{a}_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \bar{a}_{nn} \end{bmatrix}$$

$$AA^{\dagger} = \begin{bmatrix} |a_{11}|^2 & a_{11}\bar{a}_{21} & a_{11}\bar{a}_{31} & \cdots & a_{11}\bar{a}_{n1} \\ a_{21}\bar{a}_{11} & a_{21}\bar{a}_{21}+|a_{22}|^2 & a_{21}\bar{a}_{31}+a_{22}\bar{a}_{32} & \cdots & a_{21}\bar{a}_{n1}+a_{22}\bar{a}_{n2} \\ a_{31}\bar{a}_{11} & a_{31}\bar{a}_{21}+a_{32}\bar{a}_{22} & |a_{31}|^2+|a_{32}|^2+|a_{33}|^2 & \cdots & a_{31}\bar{a}_{n1}+a_{32}\bar{a}_{n2}+a_{33}\bar{a}_{n3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1}\bar{a}_{11} & a_{n1}\bar{a}_{21}+a_{n2}\bar{a}_{22} & a_{n1}\bar{a}_{31}+a_{n2}\bar{a}_{32}+a_{n3}\bar{a}_{33} & \cdots & |a_{n1}|^2+|a_{n2}|^2+\cdots+|a_{nn}|^2 \end{bmatrix}$$

$$A^{\dagger}A = \begin{bmatrix} |a_{11}|^2+|a_{21}|^2+\cdots+|a_{n1}|^2 & \bar{a}_{21}a_{22}+\bar{a}_{31}a_{32}+\cdots+\bar{a}_{n1}a_{n2} & \bar{a}_{31}a_{33}+\cdots+\bar{a}_{n1}a_{n3} & \cdots & \bar{a}_{n1}a_{nn} \\ \bar{a}_{22}a_{21}+\bar{a}_{32}a_{31}+\cdots+\bar{a}_{n2}a_{n1} & |a_{22}|^2+|a_{32}|^2+\cdots+|a_{n2}|^2 & & & \\ & & |a_{33}|^2+\cdots+|a_{n3}|^2 & & \\ & & & & \\ & & & & |a_{nn}|^2 \end{bmatrix}$$

$A^\dagger A = A A^\dagger$

$\Rightarrow |a_{11}|^2 + |a_{21}|^2 + \cdots + |a_{n1}|^2 = |a_{11}|^2$

$\cancel{|a_{22}|^2} + |a_{32}|^2 + \cdots |a_{n2}|^2 = \overset{=0}{|a_{21}|^2} + \cancel{|a_{22}|^2}$

$\cancel{|a_{33}|^2} + |a_{43}|^2 + \cdots + |a_{n3}|^2 = \overset{=0}{|a_{31}|^2} + \overset{=0}{|a_{32}|^2} + \cancel{|a_{33}|^2}$

$\vdots$

$\cancel{|a_{n-1,n-1}|^2} + |a_{n,n-1}|^2 =$

$\cancel{|a_{nn}|^2} = |a_{n1}|^2 + |a_{n2}|^2 + |a_{n3}|^2 + \cdots + \cancel{|a_{nn}|^2}$

$\Rightarrow a_{21} = a_{31} = a_{41} = \cdots = a_{n1} = 0$

$\cancel{a_{n1} = a_{n2} = \cdots = a_{n(n-1)} = 0}$

$a_{32} = a_{42} = \cdots = a_{n2} = 0$

$a_{43} = a_{53} = \cdots = a_{n3} = 0$

$\vdots$

$a_{n1} = a_{n2} = a_{n3} = \cdots = a_{n,n-1} = 0.$

$$\Longrightarrow A = \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_{nn} \end{bmatrix}$$

Diagonal matrix

A diagonal matrix is a matrix in which all off diagonal entries are zero.
i.e., the matrix $D = [d_{ij}]$ with n columns and n rows is diagonal if

$d_{ij} = 0 \;\forall\; i,j \in \{1, 2, \ldots, n\},\; i \neq j.$

$$D = \begin{bmatrix} d_{11} & 0 & \cdots & 0 \\ 0 & d_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{nn} \end{bmatrix}$$

- A matrix which is simultaneously triangular and normal is also diagonal.

- The term diagonal matrix may sometimes refer to a rectangular diagonal matrix, which is an m×n matrix with all the entries not of the form d_{ii} being zero.

Ex:- $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \\ 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & -3 & 0 & 0 \end{bmatrix}$

Vector operations

$$D = \mathrm{diag}(a_1, a_2, \ldots, a_n) \quad \& \quad v = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$Dv = \mathrm{diag}(a_1, \ldots, a_n)\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1x_1 \\ a_2x_2 \\ \vdots \\ a_nx_n \end{bmatrix}$$

This can be expressed more compactly by using a vector $d = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix}$ instead of a diagonal matrix $D = \mathrm{diag}(a_1, a_2, \ldots, a_n)$., and taking the Hadamard product of the vectors.

$$\boxed{\begin{aligned} Dv = d \odot v = d \circ v &= \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \odot \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} a_1x_1 \\ a_2x_2 \\ \vdots \\ a_nx_n \end{bmatrix} \\ &= \begin{bmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} \end{aligned}}$$

x2 (4) Hadamard product

This
storing
matrix

* $A = d$
a

This is mathematically equivalent, but avoids storing all the zeros terms of this sparse matrix.

* $A = \mathrm{diag}(a_1, a_2, \ldots, a_n) \Longrightarrow A^{-1} = \mathrm{diag}\left(\frac{1}{a_1}, \frac{1}{a_2}, \ldots, \frac{1}{a_n}\right)$

and $a_1 * a_2 * \cdots * a_n \neq 0$

Idempotent matrix

An idempotent matrix is a matrix which when multiplied by itself yields itself.

$$A^2 = A$$

* 'A' must be square matrix.

Ex:-

$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 3 & -6 \\ 1 & -2 \end{bmatrix}$, $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 2 & -2 & -4 \\ -1 & 3 & 4 \\ 1 & -2 & -3 \end{bmatrix}$

* An idempotent matrix is always diagonalizable & its eigenvalue are either 0 or 1.

Proof

$$A\vec{v} = \lambda\vec{v}$$

$$A^2\vec{v} = \lambda^2\vec{v} = \lambda\vec{v}$$

$$\vec{v} \neq \vec{0} \longrightarrow \lambda = 0 \text{ (or) } \lambda = 1$$

* $AB = A$ and $BA = B \Longrightarrow A$ & B are idempotent

Proof :

$$ABA = AA = A^2 = AB = A$$

$$A^2 = A$$

Orthogonal matrix

Ude IL A-G

An orthogonal matrix is a square matrix whose columns and rows are orthogonal unit vectors (Orthonormal vectors).

$$Q^TQ = QQ^T = I \Longleftrightarrow Q^T = Q^{-1}$$

(when Q is square)

- $Q^TQ = I$ even when Q is rectangular. In this case Q^T is only an inverse from the left.

Proof

Let,

$$Q = \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$$

where q_i : unit column vector

$$\text{orthogonality} \longrightarrow q_i^T \cdot q_j = \delta_{ij} = \begin{cases} 1 \,; & i = j \\ 0 \,; & i \neq j \end{cases}$$

$$Q^T = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix}$$

$$Q^T Q = \begin{bmatrix} q_1^T \\ q_2^T \\ \vdots \\ q_n^T \end{bmatrix} \cdot \begin{bmatrix} q_1 & q_2 & \cdots & q_n \end{bmatrix}$$

$$= \begin{bmatrix} q_1^T q_1 & q_1^T q_2 & \cdots & q_1^T q_n \\ q_2^T q_1 & q_2^T q_2 & \cdots & q_2^T q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_n^T q_1 & q_n^T q_2 & \cdots & q_n^T q_n \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

* The determinant of any orthogonal matrix is either +1 or -1

Proof

$\det(QQ^T) = \det(Q)\det(Q^T) = [\det Q]^2 = \det I = 1$

$|\det Q| = 1 \Longrightarrow \underline{\underline{\det Q = \pm 1}}$

Ex:-

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$: reflection across x-axis.

$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$: permutation of coordinate axes.

* An orthogonal matrix preserves the inner product of vectors, and therefore acts as a isometry of Euclidean space such as a rotation, reflection or rotoreflection

ie ⇒ it is a unitary transformation.

Proof

$$(Q\vec{u})\cdot(Qv) = (Qu)^T(Qv) = u^T Q^T Q v$$
$$= u^T I v = u^T v = u\cdot v$$

$$|Qu|^2 = (Qu)\cdot(Qu) = (Qu)^T(Qu)$$
$$= u^T Q^T Q u = u^T u = u\cdot u = |u|^2$$

* $|\lambda| = 1$ for every eigenvalue λ of an orthogonal matrix.

$$A^T = A^{-1} \longrightarrow AA^T = A^TA = I \qquad \begin{array}{l} A \text{ is real} \\ \bar{A} = A \end{array}$$

Proof

$$Av = \lambda v \implies v^\dagger A^\dagger = v^\dagger \bar{A}^T = v^\dagger A^T = \bar{\lambda} v^\dagger$$

$$|Av|^2 = (Av)\cdot(Av) = \langle Av|Av\rangle = (Av)^\dagger (Av)$$

$$= (v^\dagger A^\dagger)(Av)$$

$$= \bar{\lambda}\lambda v^\dagger v = |\lambda|^2 |v|^2 = |v|^2$$

$$\implies |\lambda|^2 = 1 \longrightarrow |\lambda| = 1$$

(OR) $\langle Av|v\rangle = (Av)^\dagger v = v^\dagger A^\dagger v = v^\dagger A^T v$ [A is real]

$$= v^\dagger A^{-1} v$$

[λ is an eigenvalue of A $\Rightarrow \frac{1}{\lambda}$ " A^{-1}.

$$\implies \bar{\lambda} v^\dagger v = \frac{1}{\lambda} v^\dagger v$$

$$\longrightarrow |\lambda|^2 = 1 \implies |\lambda| = 1$$

$$\boxed{\begin{array}{l} A \text{ is orthogonal matrix} \\ A^T = A^{-1} \\ \text{(or)} \\ AA^T = A^TA = I \end{array} \Bigg] \rightarrow \begin{array}{l} \det(A) = \pm 1 \\ |\lambda| = 1 \end{array}}$$

~~[illegible]~~

In 2D,

$$\det(A) = \begin{cases} +1 & \Longrightarrow \text{rotation} \\ -1 & \Longrightarrow \text{reflection} \end{cases}$$

$SO(n)$: proper rotations

~~im~~proper rotations.

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

$a^2 + c^2 = b^2 + d^2 = 1 \Longrightarrow$ (a,c) and (b,d) are on the unit circle

$ab + cd = 0$

$ad - bc = \pm 1$

$(a, c) = (\cos\theta, \sin\theta)$

$(b, d) = (\cos\phi, \sin\phi)$

$ab + cd = 0 \Longrightarrow \cos\theta\cos\phi + \sin\theta\sin\phi = \cos(\phi - \theta) = 0$

$|A| = 1 \Rightarrow ad - bc = 1 \Rightarrow \cos\theta\sin\phi - \sin\theta\cos\phi = \sin(\phi - \theta) = 1$

$\phi - \theta = (2n+1)\frac{\pi}{2} \Longrightarrow \underline{\phi = \theta + (2n+1)\frac{\pi}{2}}$

$b = \cos\left(\theta + (2n+1)\frac{\pi}{2}\right) = -\sin\theta$

$d = \sin\left(\theta + (2n+1)\frac{\pi}{2}\right) = \cos\theta$

$$\det(A) = 1 \Longrightarrow A = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} = Rot_{cw}(\theta)$$

$|A| = -1$: $ad - bc = -1 \longrightarrow \sin(\phi - \theta) = -1$

$\cos(\phi - \theta) = 0$

$\Rightarrow \phi - \theta = (2n+1)\frac{3\pi}{2} \longrightarrow \phi = \theta + (2n+1)\frac{3\pi}{2}$

$b = \cos\left[\theta + (2n+1)\frac{3\pi}{2}\right] = \sin\theta$

$d = \sin\left[\theta + (2n+1)\frac{\pi}{2}\right] = -\cos\theta$

$$\det(A) = -1 \longrightarrow A = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} = Ref(\theta/2)$$

* If $A = Rot(\theta)$,

eigenvalues of A are $e^{i\theta}, e^{-i\theta}$ for some $\theta \in (0, \pi/2)$

Proof - 2D

SO(2) group

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} = Rot_{ccw}(\theta) \qquad (A-\lambda I)\vec{v} = 0$$

$$P_A(\lambda) = \det[A - \lambda I] = \begin{vmatrix} \cos\theta - \lambda & -\sin\theta \\ \sin\theta & \cos\theta - \lambda \end{vmatrix} = 0$$

$$\cos^2\theta - 2\lambda\cos\theta + \lambda^2 + \sin^2\theta = 0$$

$$\boxed{P_A(\lambda) =} \boxed{\lambda^2 - (2\cos\theta)\lambda + 1 = 0}$$

$$\Delta = 4\cos^2\theta - 4 = 4(-\sin^2\theta) = -4\sin^2\theta$$

$$\lambda = \frac{2\cos\theta \pm 2i\sin\theta}{2} = \cos\theta \pm i\sin\theta = e^{\pm i\theta}$$

$\lambda_1 = e^{i\theta}$,

$$(A-\lambda I)\vec{v} = \begin{bmatrix} -i\sin\theta & -\sin\theta \\ \sin\theta & -i\sin\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

If $\lambda \neq 0, \pi \rightarrow \sin\theta \neq 0$

$$\left[\begin{array}{cc|c} -i\sin\theta & -\sin\theta & 0 \\ \sin\theta & -i\sin\theta & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} -i & -1 & 0 \\ 1 & -i & 0 \end{array}\right]$$

$$\left[\begin{array}{cc|c} 1 & -i & 0 \\ 1 & -i & 0 \end{array}\right] \rightarrow \left[\begin{array}{cc|c} 1 & -i & 0 \\ 0 & 0 & 0 \end{array}\right].$$

$x - iy = 0$

Set $y = t$, $x = it$

$$\vec{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix} t.$$

$\lambda_2 = e^{-i\theta} = \bar{\lambda}_1$, $\quad A v_1 = \lambda_1 v_1$

$\bar{A}\bar{v}_1 = A\bar{v}_1 = \bar{\lambda}_1 \bar{v}_1$ $\quad$ [A is real

$$v_2 = \bar{v}_1 = \begin{bmatrix} -i \\ 1 \end{bmatrix} t$$

$$\boxed{\begin{array}{l} \underset{2\times 2}{A} = Rot(\theta) \Longrightarrow \quad \lambda_1 = e^{i\theta}, \ \lambda_2 = e^{-i\theta} \\ \qquad\qquad\qquad\quad \vec{v}_1 = \begin{bmatrix} i \\ 1 \end{bmatrix} t, \ \vec{v}_2 = \begin{bmatrix} -i \\ 1 \end{bmatrix} t \\ \text{Normalization: } t = \frac{1}{\sqrt{2}} \end{array}}$$

$$AB = Rot(\theta)\, Rot(\phi) = Rot(\theta+\phi)$$

$$BA = Rot(\phi)\, Rot(\theta) = Rot(\theta+\phi) = AB$$

$\Longrightarrow$ $SO(2)$ is abelian.

A group is a pair $(G, \cdot)$ where G is a set and $\cdot : G \times G \to G$ is a binary operation on G with the following properties:

* A group is a set G, together with an operation (called the group law of G) that combines any two elements 'a' and 'b' to form another element, denoted $a \cdot b$.

To qualify as a group, the set and operation $(G, \cdot)$, must satisfy 4 requirements known as the group axioms.

i) Closure

For all a, b in G, the result of the operation $a \cdot b$ is also in G

ii) Associativity $\boxed{\forall a,b,c \in G, (a \cdot b) \cdot c = a \cdot (b \cdot c) = a \cdot b \cdot c}$

For all a, b, c in G,

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

iii) Identity element

There exist an element 'e' in G, such that, for every element 'a' in G, the equation $e \cdot a = a \cdot e = a$ holds.

~~[illegible]~~ $\boxed{\exists e \in G \text{ such that } e \cdot a = a \cdot e = a \ \forall a \in G}$

Such an element is unique & thus one speaks of the identity element.

iv) Inverse element

For each 'a' in G, there exists an element 'b' in G such that $a \cdot b = b \cdot a = e$, where e is the identity element.

$\boxed{\forall a \in G, \exists b \in G \text{ such that } a \cdot b = b \cdot a = e}$

A group in which the group operation is not commutative is called a non-abelian group/non-commutative group.

Commutivity

For all $a, b \in G$, $a \cdot b = b \cdot a$

* An isometry of $\mathbb{R}^n$ is a function $f: \mathbb{R}^n \longrightarrow \mathbb{R}^n$ such that, for any 2 vectors $x, y \in \mathbb{R}^n$, we have $|f(x) - f(y)| = |x - y|$.
i.e., f preserves distances b/w points in $\mathbb{R}^n$.

Let F be a field. Then the general linear group $GL_n(F)$ or $GL(n, F)$ is the group of invertible $n \times n$ matrices with entries in F under matrix multiplication.

- If $n \geq 2$, then the group $GL_n(F)$ is non-abelian.

Orthogonal Group.

For any field F, a $n \times n$ matrix with entries in F such that its inverse equals its transpose is called an orthogonal matrix over F. The $n \times n$ orthogonal matrices form a subgroup, denoted $O(n,F)$ of the general linear group $GL(n,F)$, i.e.,

$$O(n,F) = \{ A \in GL(n,F) \mid A^{-1} = A^{T} \}$$

- $O(n)$ is the set of $n \times n$ matrices that preserve inner product on $\mathbb{R}^n$.
- The set $O(n)$ is a group under matrix multiplication.
- If $A \in O(n)$, then 'A' is a linear isometry

* The subset $SO(n) = \{A \in O(n) \mid \det A = 1\}$ is a subgroup of $O(n)$, is called the Special orthogonal group, is the set of all proper rotations in $\mathbb{R}^n$.

$$\boxed{SO(n) = \{A \in O(n) \mid \det A = 1\} = \{A \in GL(n,\mathbb{R}) \mid A^{-1} = A^T \text{ \& } \det A = 1\}}$$

* The groups $O(n)$ & $SO(n)$ are real compact Lie groups of dimension: ${}^nC_2 = \frac{n(n-1)}{2}$

Lie group & Lie algebra

A **Lie group** is a group whose elements are organized continuously and smoothly, as opposed to discrete groups, where the elements are separated – this makes Lie groups differentiable manifolds.

The circle & sphere are examples of smooth manifolds.

Informally, a Lie group is a group (continuous group) of symmetries where the symmetries are continuous. A circle has a continuous group of symmetries ∴ you can rotate the circle an arbitrarily small amount and it looks the same. This is in contrast to the hexagon for example, if you rotate the hexagon by a small amount then it'll look different.

Lie groups have elements which are arbitrarily close to the identity transformation.

Vector Space

A vector space consists of a set V (elements of V are called vectors), a field F (elements of F are called scalars), and two operations:

- An operation called vector addition that takes 2 vectors $\vec{v}, \vec{w} \in V$, and produces a 3^{rd} vector, written as $\vec{v}+\vec{w} \in V$.

- An operation called scalar multiplication that takes a scalar $c \in F$ and a vector $\vec{v} \in V$, and produces a new vector, written $c\vec{v} \in V$

which satisfy the following conditions (called axioms):

(i) Associativity of vector addition:
$$(\vec{u}+\vec{v})+\vec{w} = \vec{u}+(\vec{v}+\vec{w}) \quad \forall\ \vec{u},\vec{v},\vec{w} \in V$$

(ii) Existence of a zero vector:
There exists a vector $\vec{0} \in V$ such that $\vec{u}+\vec{0} = \vec{u}$ $\forall\ \vec{u} \in V$, and $\vec{0}$ is called the zero vector

(iii) Existence of negatives:
(additive inverse)

For every $\vec{v} \in V$ ~~such that~~ there exists an additive inverse $\vec{w} \in V$ such that $\vec{v} + \vec{w} = \vec{0}$

(iv) Associativity of multiplication:

$(ab)\vec{u} = a(b\vec{u})$ for any $a, b \in \mathbb{F}$ & $u \in V$

(v) Distributivity:

$(a+b)\vec{u} = a\vec{u} + b\vec{u}$ and $a(\vec{u} + \vec{v}) = a\vec{u} + a\vec{v}$

for all $a, b \in F$ and $\vec{u}, \vec{v} \in V$.

(vi) Unitarity:

$1\vec{u} = \vec{u}$ for all $\vec{u} \in V$

Examples

1. $V = \{ n \times 1 \text{ column matrices of real } \# \}$

$F = \{ x \mid x \in \mathbb{R} \}$. and defining vector addition & scalar multiplication by

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} x_1+x_2 \\ y_1+y_2 \\ \vdots \\ x_n+y_n \end{bmatrix} ; \quad c \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} cx_1 \\ cx_2 \\ \vdots \\ cx_n \end{bmatrix}$$

2. $V = \{ \text{polynomials of degree} \le n \text{ with real coeff.} \}$

$F = \{ x \mid x \in \mathbb{R} \}$.

and define vector addition and scalar multiplication ~~in the following~~ as:

$$(a_0 + a_1 t + \cdots + a_n t^n) + (b_0 + b_1 t + \cdots + b_n t^n) = (a_0+b_0) + (a_1+b_1)t + \cdots + (a_n+b_n)t^n$$

and

$$c(a_0 + a_1 t + \cdots + a_n t^n) = ca_0 + ca_1 t + \cdots + ca_n t^n$$

* There is a sense in which we can think of P_3 as the same as $\mathbb{R}^4$.

$a_0 + a_1 t + a_2 t + a_3 t^3$ corresp. to $\begin{bmatrix} a_0 \\ a_1 \\ a_2 \\ a_4 \end{bmatrix}$

3. $V = \{$ infinite sequence of real # $(x_1, x_2, \dots)\}$

$F = \{\alpha \mid \alpha \in \mathbb{R}\}$ and vector addition and scalar multiplication is:

$(x_1, x_2, \dots) + (y_1, y_2, \dots) = (x_1 + y_1, x_2 + y_2, \dots)$

$c(x_1, x_2, \dots) = (cx_1, cx_2, \dots)$

4. $V = \{$ set of all continuous functions, $f : \mathbb{R} \to \mathbb{R}\}$.

$F = \{\alpha \mid \alpha \in \mathbb{R}\}$ and vector addition and scalar multiplication as:

$f+g$ is the continuous function defined by

$(f+g)(a) = f(a) + g(a)$

cf is the cont. function defined by $(cf)(x) = cf(x)$

Note: Sum of 2 cont. functions is cont.

Multiplying a cont. function by a real # gives another cont. function.

5. $V =$ …

$F =$ …

scalar …

$(f+g)$ …

…

set …

$y'' =$ …

$y =$ …

$(f+g)'' =$ …

$(cf)''$ …

5 $V = \{$ set of all functions $f : \mathbb{R} \to \mathbb{R}$ that satisfy the eq^n. $f'' = -f$

$F = \{x \mid x \in \mathbb{R}\}$ and vector addition & scalar multiplication is:

$(f+g)(a) = f(a) + f(g)$ and $(cf)(x) = c\, f(x)$

~~Wang~~

set $y = e^{ax}$

$y'' = a^2 e^{ax} = -e^{ax} \longrightarrow a^2 = -1$

$\Longrightarrow a = \pm i$

$y = Ae^{ix} + Be^{-ix} = (A+B)\cos x + i(A-B)\sin x$

$= C_1 \cos x + C_2 \sin x$

$(f+g)'' = f'' + g'' = -f - g = -(f+g)$ for $f, g \in V$

$(cf)'' = cf'' = c(-f) = -(cf)$ & $c \in \mathbb{R}$

6. $V = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \middle| \; x+y+z=0 \right.$, i.e., plane through origin $\left. \right\}$.

$$\begin{pmatrix} x_1 \\ y_1 \\ z_1 \end{pmatrix} + \begin{pmatrix} x_2 \\ y_2 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1+x_2 \\ y_1+y_2 \\ z_1+z_2 \end{pmatrix} \text{ and } c\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} cx \\ cy \\ cz \end{pmatrix}$$

7. The singleton set $\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$ is a vector space

under the operations

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } r\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

* (7) vector space must have atleast one element, its zero vector.

$\therefore$ One element vector space is the smallest one possible.

8. The set of 2x2 matrices with real # entries is a vector space under the natural entry by entry operations.

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} + \begin{bmatrix} w & x \\ y & z \end{bmatrix} = \begin{bmatrix} a+w & b+x \\ c+y & d+z \end{bmatrix} \quad \text{and}$$

$$r\begin{bmatrix} a & b \\ c & d \end{bmatrix} = \begin{bmatrix} ra & rb \\ rc & rd \end{bmatrix}$$

* we can think of this vector space as the same as $\mathbb{R}^4$.

* An *algebra* is simply a vector space V over a field F with an additional binary operation, $\cdot : V \times V \to V$, called algebra multiplication which is bilinear.

$$(x+y)\cdot z = x\cdot z + y\cdot z \quad \& \quad x\cdot(y+z) = x\cdot y + x\cdot z$$

$$(ax)\cdot y = a(x\cdot y) \quad \& \quad x\cdot(by) = b(x\cdot y)$$

Note: we don't require the multiplication '$\cdot$' to be commutative or even associative

$$x\cdot y \neq y\cdot x \quad \text{and} \quad x\cdot(y\cdot z) \neq (x\cdot y)\cdot z$$

* A Lie algebra is a vector space L over a field F with a bilinear operation $[\,,\,] : L \times L \to L$, called the Lie bracket such that:

(i) $[x, x] = 0 \quad \forall\, x \in L$ [anti-symmetry]

(or) $[x, y] = -[y, x] \quad \forall\, x, y \in L$

(ii) $[x, [y, z]] + [y, [z, x]] + [z, [x, y]] = 0 \quad \forall\, x, y, z \in L$

— Jacobi identity

(iii) $[ax + by, z] = a[x, z] + b[y, z]$

— Bilinearity.

Lie algebra

If we want to work with more complicated Lie groups, working directly with the transformation matrices becomes prohibitively difficult. Instead, most of the information we need to know about the group is already present in the infinitesimal transformations.

SO(2)

Consider those transformations that are close to the identity.

Since the identity is $A(0)$, these will be the transformations $A(\varepsilon)$ with $\varepsilon \ll 1$.

ie.,

an infinitesimal SO(2) rotation by the angle $\delta\phi = \varepsilon$.

$$\cos\varepsilon = 1 - \frac{\varepsilon^2}{2} + \cdots \approx 1$$

$$\sin\varepsilon = \varepsilon - \frac{\varepsilon^3}{3} + \cdots \approx \varepsilon$$

$$A(\varepsilon) = \begin{bmatrix} \cos\varepsilon & -\sin\varepsilon \\ \sin\varepsilon & \cos\varepsilon \end{bmatrix} \approx \begin{bmatrix} 1 & -\varepsilon \\ \varepsilon & 1 \end{bmatrix}$$

$$= I + \varepsilon \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

$$= I - i\varepsilon \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} = I - i\varepsilon\sigma_2$$

with $\sigma_2^\dagger = \sigma_2$ and $\sigma_2^2 = I_2$ & $-i\sigma_2$ is skew-symmetric

The only information here besides the identity is the matrix $\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}$, but this is enough to recover the whole group.

For general Lie groups, we get one generator for each continuous parameter labeling the group elements. The set of all linear combinations of these generators is a vector space called the Lie algebra of the group.

$$A(n\varepsilon) = [A(\varepsilon)]^n = \left[I + \varepsilon \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \right]^n$$

Binomial expansion $\Rightarrow$

$$A(n\varepsilon) \approx \sum_{k=0}^{n} {}^nC_k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k \varepsilon^k \cdot 1^{n-k}$$

$$A(\theta) = \lim_{\substack{\varepsilon \to 0 \\ n\varepsilon \to \theta \\ n \to \infty}} A(n\varepsilon) = \lim_{n \to \infty} \sum_{k=0}^{n} {}^nC_k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k \varepsilon^k$$

$$= \lim_{n \to \infty} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k \varepsilon^k$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!} \varepsilon^k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k$$

$$= \lim_{n\to\infty} \sum_{k=0}^{\infty} \frac{1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!} (n\varepsilon)^k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k$$

$$= \sum_{k=0}^{\infty} \frac{1}{k!} \theta^k \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k$$

$$e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!}$$

$$A(\theta) = \exp\left(\theta \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}\right)$$

$$= \exp\left(-i\theta \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}\right)$$

$$= e^{-i\theta\sigma_2} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$\Rightarrow$ The Pauli matrix σ_2 is the generator of the $SO(2)$ group.

$A(\theta) =$

$$A(\theta) = \sum_{k=0}^{\infty} \frac{1}{k!} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^k \theta^k$$

$$= \sum_{m=0}^{\infty} \frac{1}{(2m)!} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{2m} \theta^{2m} + \sum_{m=0}^{\infty} \frac{1}{(2m+1)!} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}^{2m+1} \theta^{2m+1}$$

$$= \sum_{m=0}^{\infty} \frac{(-1)^{2m}}{(2m)!} \sigma_2^{2m} \theta^{2m} + \sum_{m=0}^{\infty} \frac{(-i\sigma_2)^{2m}}{(2m+1)!} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \theta^{2m+1}$$

$$= I \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} \theta^{2m} + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} \theta^{2m+1}$$

$$= I \left(I - \frac{\theta^2}{2!} + \frac{\theta^4}{4!} - \cdots \right) + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \left(\theta - \frac{\theta^3}{3!} + \cdots \right)$$

$$= \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \cos\theta + \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix} \sin\theta$$

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

$$A(d\theta) = A(\varepsilon) = I - i\varepsilon\sigma_2$$

$$A(\theta + d\theta) = A(\theta + \varepsilon) = A(\theta)A(\varepsilon)$$

$$= A(\theta) - i\varepsilon A(\theta)\sigma_2 = A(\theta) - i d\theta A(\theta)\sigma_2$$

$$A(\theta + d\theta) = A(\theta) + d\theta \frac{dA(\theta)}{d\theta}$$

$$\frac{dA(\theta)}{d\theta} = -i\sigma_2 A(\theta)$$

with boundary condition $A(0) = I$.

$$\Longrightarrow A(\theta) = e^{-i\sigma_2\theta}$$

σ_2 : generator of the SO(2) group.

where, $\sigma_2 = i \left.\frac{dA(\theta)}{d\theta}\right|_{\theta=0}$

SO(3)

* 3D is a special case when
\# of axis = # of planes = 3.

	# of axis n	# of planes (# of rotation matrices), ${}^{n}C_2$
2D	2	${}^{2}C_2 = 1$
3D	3	${}^{3}C_2 = 3$
4D	4	${}^{4}C_2 = 6$
5D	5	${}^{5}C_2 = \frac{5\times 4\times 3}{3\times 2} = 10$

* Any arbitrary element of $SO(3)$ may be written as the composition of rotations in the planes generated by the 3 standard orthogonal basis vectors of $\mathbb{R}^3$.

$$\boxed{\begin{gathered} R_z(\theta)=\begin{bmatrix}\cos\theta & -\sin\theta & 0\\ \sin\theta & \cos\theta & 0\\ 0 & 0 & 1\end{bmatrix},\quad R_x(\theta)=\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & -\sin\theta\\ 0 & \sin\theta & \cos\theta\end{bmatrix} \\ R_y(\theta)=\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\ 0 & 1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix}\end{gathered}}$$

represent rotations by θ in the xy-, yz-, zx-planes respectively.

The rotation matrix for a rotation by an angle θ about an axis in the direction of the unit vector $\hat{n} = (n_1, n_2, n_3)$ is given by:

$$R_n(\theta) = \begin{bmatrix} \cos\theta + n_1^2(1-\cos\theta) & n_1 n_2(1-\cos\theta) - n_3\sin\theta & n_1 n_3(1-\cos\theta) + n_2\sin\theta \\ n_1 n_2(1-\cos\theta) + n_3\sin\theta & \cos\theta + n_2^2(1-\cos\theta) & n_2 n_3(1-\cos\theta) - n_1\sin\theta \\ n_1 n_3(1-\cos\theta) - n_2\sin\theta & n_2 n_3(1-\cos\theta) + n_1\sin\theta & \cos\theta + n_3^2(1-\cos\theta) \end{bmatrix}$$

(or)

$$R(\hat{n},\theta) = \cos\theta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (1-\cos\theta) \begin{bmatrix} n_1^2 & n_1 n_2 & n_1 n_3 \\ n_2 n_1 & n_2^2 & n_2 n_3 \\ n_3 n_1 & n_3 n_2 & n_3^2 \end{bmatrix}$$

$$+ \sin\theta \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}$$

$$= I\cos\theta + n \otimes n\,(1-\cos\theta) + [n]_\times \sin\theta$$

Rodrigue's rotation formula

where,

$$\boxed{R_{ij}(\hat{n},\theta) = \cos\theta\,\delta_{ij} + (1-\cos\theta)\,n_i n_j - \sin\theta\,\epsilon_{ijk}\,n_k}$$

Proof – Rodrigue's rotation formula

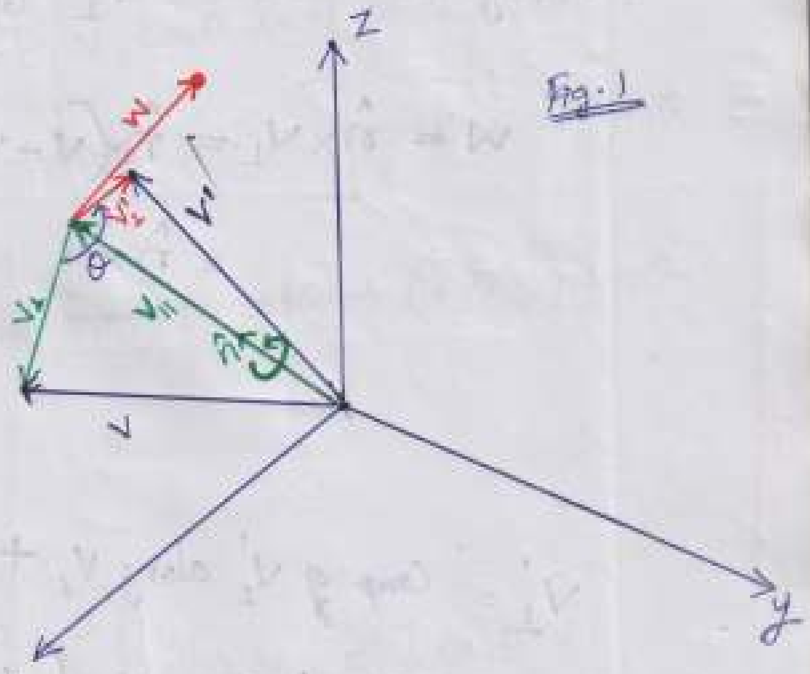


Fig. 1

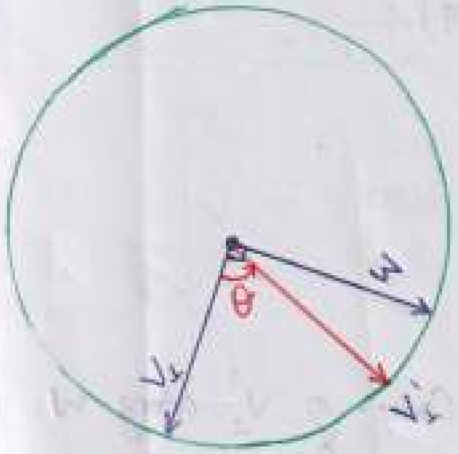


Fig. 2

$\hat{n} = (n_1, n_2, n_3)$ & $v' = R(\hat{n}, \theta)\, v$

$v_{\parallel} = (v \cdot \hat{n})\hat{n}$

$v_\perp = v - v_\parallel = v - (v \cdot \hat{n})\hat{n}$

Figure $\longrightarrow$ $v_\parallel$ is not affected by rotation.

The plane in Fig. 2 is spanned by two orthogonal vectors $v_\perp$ and w,

$$w = \hat{n}\times v_\perp = \hat{n}\times(v - v_\parallel)$$
$$= \hat{n}\times v$$

$v'_\perp =$ Comp. of $v'_\perp$ along $v_\perp$ + Comp. of $v'_\perp$ along w

$$= \left(v'_\perp\cdot\frac{v_\perp}{|v_\perp|}\right)\frac{v_\perp}{|v_\perp|} + \left(v'_\perp\cdot\frac{w}{|w|}\right)\frac{w}{|w|}$$

$$|v_\perp| = |v'_\perp| = |w|$$

$$v'_\perp = v_\perp\cos\theta\cdot\frac{\cancel{|v_\perp||v'_\perp|}}{\cancel{|v_\perp|^2}} + w\sin\theta\,\frac{\cancel{|v'_\perp||w|}}{\cancel{|w|^2}}$$

$$v'_\perp = v_\perp\cos\theta + w\sin\theta$$

$$v' = v'_{\parallel} + v'_{\perp} = v_{\parallel} + v'_{\perp}$$

$$= (v\cdot\hat{n})\hat{n} + v_{\perp}\cos\theta + w\sin\theta$$

$$= (v\cdot\hat{n})\hat{n} + \left[v - (v\cdot\hat{n})\hat{n}\right]\cos\theta + w\sin\theta$$

$$v' = v\cos\theta + \hat{n}(\hat{n}\cdot v)(1-\cos\theta) + (\hat{n}\times v)\sin\theta$$

$$v\cos\theta = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\cos\theta\, v = I\cos\theta\, v$$

$$\hat{n}(\hat{n}\cdot v) = \hat{n}\hat{n}^T v = \begin{bmatrix} n_1 \\ n_2 \\ n_3 \end{bmatrix}\begin{bmatrix} n_1 & n_2 & n_3 \end{bmatrix} v$$

$$= \begin{bmatrix} n_1^2 & n_1 n_2 & n_1 n_3 \\ n_2 n_1 & n_2^2 & n_2 n_3 \\ n_3 n_1 & n_3 n_2 & n_3^2 \end{bmatrix} v = \hat{n}\otimes\hat{n}\, v$$

$$\hat{n}\times v = \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix} v = [n]_{\times} v$$

$$V' = \left[\bar{I}\cos\theta + \hat{n} \otimes \hat{n}\,(1-\cos\theta) + [\hat{n}]_{\times} \sin\theta \right] V$$

$$= R(\hat{n}, \theta)\, V$$

$R_z(\theta) \cdot$

$R_x(\theta)\ R$

$R_z(\theta)\ R_z$

$\Longrightarrow$ S

* $$R_z(\theta) \cdot R_x(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta\cos\theta & \sin^2\theta \\ \sin\theta & \cos^2\theta & -\sin\theta\cos\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}$$

$$R_x(\theta) R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta\cos\theta & \cos^2\theta & -\sin\theta \\ \sin^2\theta & \sin\theta\cos\theta & \cos\theta \end{bmatrix}$$

$$R_z(\theta) R_x(\theta) \neq R_x(\theta) R_z(\theta)$$

$\Longrightarrow$ SO(3) is not abelian.

* Eigenvalues & Eigenvectors

$$R_z(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} R(\theta) & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det(R_z(\theta) - \lambda I_3) = \begin{vmatrix} R(\theta) - \lambda I_2 & 0 \\ 0 & 1-\lambda \end{vmatrix} = 0.$$

$$\det(R(\theta) - \lambda I_2) \cdot \det(1-\lambda) = 0.$$

$$\underline{\lambda = 1, e^{\pm\theta}}$$

$$\lambda = 1 : \left[\begin{array}{ccc|c} \cos\theta - 1 & -\sin\theta & 0 & 0 \\ \sin\theta & \cos\theta - 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array}\right] \rightarrow \begin{array}{l} \sin\theta(\cos\theta - 1)x - \sin^2\theta\, y = 0 \\ \sin\theta(\cos\theta - 1)x + (\cos\theta - 1)y = 0 \\ \hline x = 0, y = 0, z = t \end{array}$$

$|t^2| = 1$

$$\underline{v_1 = (0, 0, 1)}$$

$$\lambda = e^{i\theta} : \quad \cos\theta + i\sin\theta \quad \left[\begin{array}{ccc|c} -i\sin\theta & -\sin\theta & 0 & 0 \\ \sin\theta & -i\sin\theta & 0 & 0 \\ 0 & 0 & 1-e^{i\theta} & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} -i & -1 & 0 & 0 \\ 1 & -i & 0 & 0 \\ 0 & 0 & 1-e^{i\theta} & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & -i & 0 & 0 \\ 1 & -i & 0 & 0 \\ 0 & 0 & 1-e^{i\theta} & 0 \end{array}\right] \Longrightarrow \begin{array}{l} x = iy \;\&\; z = 0 \;\&\; y = -ix \\ x = t,\ y = -it,\ z = 0 \\ |t^2|[1+1] = 1 \Rightarrow t = 1/\sqrt{2} \end{array}$$

$$\underline{v_2 = \frac{1}{\sqrt{2}}(1, -i, 0)}$$

$\lambda = e^{-i\theta}$: $\cos\theta - i\sin\theta$

$R_x(\theta) =$

$\lambda = 1$:

$v_1 =$

$\lambda = e^{i\theta}$: $\cos\theta + i\sin\theta$

$\lambda = e^{-i\theta}$: $\cos\theta - i\sin\theta$

$$\left[\begin{array}{ccc|c} i\sin\theta & -\sin\theta & 0 & 0 \\ \sin\theta & i\sin\theta & 0 & 0 \\ 0 & 0 & 1-e^{-i\theta} & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} i & -1 & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & 1-e^{-i\theta} & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1 & i & 0 & 0 \\ 1 & i & 0 & 0 \\ 0 & 0 & 1-e^{i\theta} & 0 \end{array}\right] \longrightarrow x = -iy \ \& \ z = 0$$

$y = ix$

$t^2[1+1] = 1 \Rightarrow t = 1/\sqrt{2}$

$$v_3 = \frac{1}{\sqrt{2}}(1, i, 0)$$

$$R_x(\theta) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & R(\theta) \end{bmatrix}$$

$\lambda : 1, e^{\pm\theta}$

$\lambda = 1$: $\left[\begin{array}{ccc|c} 0 & 0 & 0 & 0 \\ 0 & \cos\theta - 1 & -\sin\theta & 0 \\ 0 & \sin\theta & \cos\theta - 1 & 0 \end{array}\right] \rightarrow$ $\sin\theta(\cos\theta - 1)y - \sin^2\theta\, z = 0$

$\sin\theta(\cos\theta - 1)y + (\cos\theta - 1)^2 z = 0$

$y = 0, z = 0, x = t$

$$v_1 = (1, 0, 0)$$

$\lambda = e^{i\theta}$: $\cos\theta + i\sin\theta$

$$\left[\begin{array}{ccc|c} 1-e^{i\theta} & 0 & 0 & 0 \\ 0 & -i\sin\theta & -\sin\theta & 0 \\ 0 & \sin\theta & -i\sin\theta & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1-e^{i\theta} & 0 & 0 & 0 \\ 0 & -i & -1 & 0 \\ 0 & 1 & -i & 0 \end{array}\right]$$

$$\left[\begin{array}{ccc|c} 1-e^{i\theta} & 0 & 0 & 0 \\ 0 & 1 & -i & 0 \\ 0 & 1 & -i & 0 \end{array}\right] \longrightarrow x = 0, \ y = iz \rightarrow z = -iy$$

$y = t, \ x = 0, \ z = -it$

$$v_2 = \frac{1}{\sqrt{2}}(0, 1, -i)$$

$\lambda = e^{-i\theta}$: $\begin{bmatrix} 1-e^{-i\theta} & 0 & 0 & | & 0 \\ 0 & i\sin\theta & -\sin\theta & | & 0 \\ 0 & \sin\theta & i\sin\theta & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1-e^{-i\theta} & 0 & 0 & | & 0 \\ 0 & i & -1 & | & 0 \\ 0 & 1 & i & | & 0 \end{bmatrix}$

$\cos\theta - i\sin\theta$

$\begin{bmatrix} 1-e^{-i\theta} & 0 & 0 & | & 0 \\ 0 & 1 & i & | & 0 \\ 0 & 1 & i & | & 0 \end{bmatrix} \Rightarrow x=0,\ y=-iz \Rightarrow z = iy$

$y = t,\ z = i$

$v_3 = \frac{1}{\sqrt{2}}(0, 1, i)$

$R_y(\theta) = \begin{bmatrix} \cos\theta & 0 & -\sin\theta \\ 0 & 1 & 0 \\ \sin\theta & 0 & \cos\theta \end{bmatrix}$

$\lambda : 1, e^{\mp i\theta}$

$\lambda = 1$: $\begin{bmatrix} \cos\theta - 1 & 0 & -\sin\theta & | & 0 \\ 0 & 0 & 0 & | & 0 \\ \sin\theta & 0 & \cos\theta - 1 & | & 0 \end{bmatrix} \rightarrow$ $(\cos\theta - 1)\sin\theta x - \sin^2\theta z = 0$

$(\cos\theta - 1)\sin\theta x + z(\cos\theta - 1)^2 = 0$

$x = 0,\ z = 0,\ y = t$

$v_1 = (0, 1, 0)$

$\lambda = e^{i\theta}$: $\begin{bmatrix} -i\sin\theta & 0 & -\sin\theta & | & 0 \\ 0 & 1-e^{i\theta} & 0 & | & 0 \\ \sin\theta & 0 & -i\sin\theta & | & 0 \end{bmatrix} \rightarrow \begin{bmatrix} -i & 0 & -1 & | & 0 \\ 0 & 1-e^{i\theta} & 0 & | & 0 \\ 1 & 0 & -i & | & 0 \end{bmatrix}$

$\cos\theta + i\sin\theta$

$\begin{bmatrix} 1 & 0 & -i & | & 0 \\ 0 & 1-e^{i\theta} & 0 & | & 0 \\ 1 & 0 & -i & | & 0 \end{bmatrix} \Longrightarrow x = i\cancel{y},\ y = 0$

$z = -ix$

$x = t,\ y = 0,\ z = -it$

$v_2 = \frac{1}{\sqrt{2}}(1, 0, -i)$

$\lambda = e^{-i\theta}$:

$\cos\theta - i\sin\theta$

$x =$

$x = t$

~~$R_z(\theta)$~~ : λ	
$R_z(\theta)$	$v_1 =$
$R_x(\theta)$	v_1
$R_y(\theta)$	v_1

$\lambda = e^{-i\theta}$
$\cos\theta - i\sin\theta$

$$\left[\begin{array}{ccc|c} i\sin\theta & 0 & -\sin\theta & 0 \\ 0 & 1-e^{-i\theta} & 0 & 0 \\ \sin\theta & 0 & i\sin\theta & 0 \end{array}\right] \rightarrow \left[\begin{array}{ccc|c} 1 & 0 & i & 0 \\ 0 & 1-e^{-i\theta} & 0 & 0 \\ 1 & 0 & i & 0 \end{array}\right]$$

$x = -iz$ & $y = 0. \Rightarrow z = ix$

$x = t,\ y = 0,\ z = it$

$$\underline{\underline{V_3 = t(1,0,i)}}$$

$\cancel{R_z(\theta)}$	$\lambda = 1$	$\lambda = e^{i\theta}$ $= \cos\theta + i\sin\theta$	$\lambda = e^{-i\theta}$ $= \cos\theta - i\sin\theta$
$R_z(\theta)$	$V_1 = t(0,0,1)$	$V_2 = t(1,-i,0)$	$V_3 = t(1,i,0)$
$R_x(\theta)$	$V_1 = t(1,0,0)$	$V_2 = t(0,1,-i)$	$V_3 = t(0,1,i)$
$R_y(\theta)$	$V_1 = t(0,1,0)$	$V_2 = t(1,0,-i)$	$V_3 = t(1,0,i)$

Lie algebra - ~~so(3)~~ so(3)

Rodrigue's formula

$$R_{\hat{n}}(\theta) = R(\hat{n},\theta) = I\cos\theta + n\otimes n\,(1-\cos\theta) + [\hat{n}]_{\times}\sin\theta$$

$$= \cos\theta\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix} + (1-\cos\theta)\begin{bmatrix}n_1^2 & n_1n_2 & n_1n_3\\ n_2n_1 & n_2^2 & n_2n_3\\ n_3n_1 & n_3n_2 & n_3^2\end{bmatrix}$$

$$+\sin\theta\begin{bmatrix}0 & -n_3 & n_2\\ n_3 & 0 & -n_1\\ -n_2 & n_1 & 0\end{bmatrix}$$

In the case of an infinitesimally small SO(3) rotation by the angle $|\delta\theta| = \varepsilon \ll 1$ about the direction along the unit vector $\hat{n}$, a Taylor expansion of the Rodrigue's formula to 1st order gives:

$$R_{\hat{n}}(\delta\theta) = R_{\hat{n}}(\varepsilon) \approx I_3 + \varepsilon\begin{bmatrix}0 & -n_3 & n_2\\ n_3 & 0 & -n_1\\ -n_2 & n_1 & 0\end{bmatrix}$$

$$= I_3 - i\varepsilon\begin{bmatrix}0 & -in_3 & in_2\\ in_3 & 0 & -in_1\\ -in_2 & in_1 & 0\end{bmatrix} = I_3 - i\varepsilon J_{\hat{n}}$$

with $-iJ_{\hat{n}}$ is real & skew symmetric.

$$R_{\hat n}(n\varepsilon) = \left[I_3 - i\varepsilon J_{\hat n}\right]^n = \sum_{k=0}^{n} {}^nC_k (-i\varepsilon)^k J_{\hat n}^k$$

$$R_{\hat n}(\theta) = \lim_{\substack{n\to\infty \\ n\varepsilon\to\theta}} R_{\hat n}(n\varepsilon)$$

$$= \lim_{\substack{n\to\infty \\ n\varepsilon\to\theta}} \sum_{k=0}^{n} {}^nC_k (-i)^k \varepsilon^k J_{\hat n}^k$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{n!}{(n-k)!\,k!} (-i)^k \varepsilon^k J_{\hat n}^k$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1)}{k!} (-i)^k \varepsilon^k J_{\hat n}^k$$

$$= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{1\left(1-\frac{1}{n}\right)\cdots\left(1-\frac{k-1}{n}\right)}{k!} (-i)^k (n\varepsilon)^k J_{\hat n}^k$$

$$= \sum_{k=0}^{n} \frac{(-i\theta J_{\hat n})^k}{k!} = e^{-i\theta J_{\hat n}}$$

$$R_{\hat{n}}(\theta) = e^{-i\theta J_{\hat{n}}}$$

$$= \exp\left(-i\theta \begin{bmatrix} 0 & -in_3 & in_2 \\ in_3 & 0 & -in_1 \\ -in_2 & in_1 & 0 \end{bmatrix}\right)$$

$$= \exp\left(\theta \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}\right)$$

$$= \cos\theta \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + (1-\cos\theta) \begin{bmatrix} n_1^2 & n_1n_2 & n_1n_3 \\ n_2n_1 & n_2^2 & n_2n_3 \\ n_3n_1 & n_3n_2 & n_3^2 \end{bmatrix}$$

$$+ \sin\theta \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}$$

$$= \cos\theta\, I_3 + n \otimes n\,(1-\cos\theta) + [n]_\times \sin\theta$$

Considering the infinitesimal rotations about the 3 axes of the coordinate system, one finds a basis of $so(3)$, namely

$$-iJ_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad -iJ_2 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ -1 & 0 & 0 \end{bmatrix},$$

$$-iJ_3 = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

The corresp. matrices J_k are the generators of the $SO(3)$ group

$$\boxed{J_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}$$

where, $J_k = i \left.\dfrac{dR_k(\theta)}{d\theta}\right|_{\theta=0}$ & $(J_k)_{ij} = -i\epsilon_{ijk}$

$$R_{\hat{n}}(\delta\theta) \sim I - i\,\delta\theta\, J_{\hat{n}}$$

If we multiply 2 rotations about the same axis.

$$R_{\hat{n}}(\theta_1+\theta_2) = R_{\hat{n}}(\theta_1)\, R_{\hat{n}}(\theta_2)$$

Taking $\theta_1 = \theta$ & $\theta_2 = \delta\theta$ to be infinitesimal

$$R(\theta+\delta\theta) \approx \left[1 - i\,\delta\theta\, J_{\hat{n}}\right] R_{\hat{n}}(\theta)$$

$$\frac{dR_{\hat{n}}(\theta)}{d\theta} = \lim_{\delta\theta \to 0} \frac{R(\theta+\delta\theta) - R(\theta)}{\delta\theta} = -i\, J_{\hat{n}}\, R_{\hat{n}}(\theta)$$

$$\Longrightarrow R_{\hat{n}}(\theta) = R_{\hat{n}}(0)\cdot e^{-iJ_{\hat{n}}\theta} = e^{-iJ_{\hat{n}}\theta} \quad \text{if } R_{\hat{n}}(0) = I$$

$$-iJ_{\hat{n}} = \begin{bmatrix} 0 & -n_3 & n_2 \\ n_3 & 0 & -n_1 \\ -n_2 & n_1 & 0 \end{bmatrix}$$

$$-i\left(J_{\hat{n}}\right)_{ij} = -\sum_{k=1}^{3} \epsilon_{ijk}\, n_k$$

$$\left(J_k\right)_{ij} = -i\epsilon_{ijk}$$

$$\Longrightarrow -i\left(J_{\hat{n}}\right)_{ij} = \sum_{k=1}^{3} \left[-i\left(J_k\right)_{ij}\right] n_k$$

$$\left(J_{\hat{n}}\right)_{ij} = \sum_{k=1}^{3} \left(J_k\right)_{ij}\, n_k$$

$$\Rightarrow \boxed{J_{\hat{n}} = \sum_{k=1}^{3} n_k J_k = \hat{n}\cdot\vec{J}}$$

where, $\vec{J} = (J_1, J_2, J_3)$ & $\hat{n} = (n_1, n_2, n_3)$

* The dot product is not a scalar product, but merely a convenient notation.

*

* J_1, J_2, J_3 form a basis of generators, $J_{\hat{n}}$ being the generator of rotations about the $\hat{n}$-direction

ie, the generators J_1, J_2, J_3 of the SO(3) group span a basis for the Lie algebra so(3).

$$J_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \quad J_2 = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}, \quad J_3 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$J_1 J_2 - J_2 J_1 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix}\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}$$

$$= \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} - \begin{bmatrix} 0 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = iJ_3.$$

$$\boxed{\begin{gathered} [J_1, J_2] = iJ_3, \quad [J_2, J_3] = iJ_1, \quad [J_3, J_1] = iJ_2 \\ \implies [J_i, J_j] = i\sum_{k=1}^{3} \epsilon_{ijk} J_k \quad \forall\ i,j \in \{1,2,3\} \end{gathered}}$$

- rotations about different axes do not commute.

$$\frac{dR_{\hat{n}}(\theta)}{d\theta} = -i J_{\hat{n}} R_{\hat{n}}(\theta)$$

with boundary condition $R_{\hat{n}}(0) = I$

$$\Rightarrow R_{\hat{n}}(\theta) = e^{-iJ_{\hat{n}}\theta}$$

$$= e^{-i(\hat{n}\cdot J)\theta}$$

$$= e^{-i(n_1J_1+n_2J_2+n_3J_3)\theta}$$

$$\neq e^{-in_1J_1\theta}\cdot e^{-in_2J_2\theta}\cdot e^{-in_3J_3\theta}$$

$$J_1^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$Tr(J_1^2) = 1+1 = 2.$$

$$J_1 J_2 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ -1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$Tr(J_1 J_2) = 0.$$

Normalization Condition on the generators

$$\boxed{Tr(J_i J_j) = 2\delta_{ij} \quad \forall\ i,j \in \{1,2,3\}.}$$

Involutory matrix

An *involutory matrix* is a matrix that is its own inverse.

$$A^{-1} = A \Longrightarrow A^2 = I$$

Ex:

The 2×2 real matrix $\begin{bmatrix} a & b \\ c & -a \end{bmatrix}$ is involutory provided that $a^2+bc=1$

The Pauli matrices

$$\sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

* The determinant of an involutory matrix is ± 1.

Proof.

$$\det(AA^{-1}) = \det(A)\cdot\det(A^{-1})$$

$$= (\det A)^2 = \det I = 1$$

$$|\det A| = 1 \implies \det A = \pm 1$$

* Matrix A is involutory iff $(I-A)(I+A) = 0$

$$A^2 = I \iff (I-A)(I+A) = 0$$

Rank of a matrix.

* The rank of a matrix 'A' is the dimension of the vector space generated (or spanned) by its columns.

→ this corresp. to the maximal # of linearly independent columns of 'A'.

⟹ is identical to the dimension of the vector space spanned by its rows.

ie, rank is thus a measure of the "non-degenerateness" of the system of linear equations.

* A non-zero matrix A is said to have rank r if there exists atleast one minor of order r which is non-zero & every minor of order (r+1) is zero.

⇕

* The rank of a matrix 'A' is the maximal order of a non-zero minor of 'A'.

• Idea of proof : If a minor of order 'k' is non-zero, then the corresp. columns of 'A' are linearly independent.

Ex:-

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

Minors of order 3:

$$\begin{vmatrix} -1 & 0 & 2 \\ 0 & 1 & 1 \\ 2 & 0 & -4 \end{vmatrix} = \begin{vmatrix} -1 & 0 & 1 \\ 0 & 1 & -1 \\ 2 & 0 & -2 \end{vmatrix} = \begin{vmatrix} -1 & 2 & 1 \\ 0 & 1 & -1 \\ 2 & -4 & -2 \end{vmatrix} = \begin{vmatrix} 0 & 2 & 1 \\ 1 & 1 & -1 \\ 0 & -4 & -2 \end{vmatrix} = 0.$$

$\begin{vmatrix} 2 & 1 \\ 1 & -1 \end{vmatrix} = -2 \neq 0$ $\Big\}$ 'A' contains atleast 1 non-vanishing 2 rowed minor

Hence,

Rank (A) = 2

* Let $\{v_1, v_2, \ldots, v_k\}$ be a set of 'k' different vectors in $\mathbb{R}^n$. Then $v_1, v_2, \ldots, v_k$ are linearly independent iff their span,
span$\{v_1, v_2, \ldots, v_k\}$ has dimension 'k'.

$\Rightarrow$ Vectors $\{v_1, v_2, \ldots, v_k\}$ are linearly independent iff they form a basis for the subspace that they span.

Ex:- 2 vectors are linearly independent iff they span a plane; 3 vectors are linearly independent iff they span a 3D subspace.

* Vectors $\{v_1, v_2, \ldots, v_k\}$ in $\mathbb{R}^n$ are said to be linearly dependent iff there exists a subspace of $\mathbb{R}^n$ with dimension less than 'k' that contains $v_1, v_2, \ldots, v_k$.

Check 14/7/2020

$\vec{A}_1 = (a_1, a_2, a_3)$, $\vec{A}_2 = (b_1, b_2, b_3)$

$$x_1 \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix} + x_2 \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \longrightarrow \begin{bmatrix} a_1 & b_1 \\ a_2 & b_2 \\ a_3 & b_3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

For $\vec{A}_1, \vec{A}_2$ to be independent,

$$\vec{A}_1 x_1 + \vec{A}_2 x_2 = 0 \quad \text{iff} \quad x_1 = x_2 = 0$$

$$a_1 x_1 + b_1 x_2 = 0$$
$$a_2 x_1 + b_2 x_2 = 0$$
$$a_3 x_3 + b_3 x_3 = 0$$

If we find the solutions to any 2 out of 3 equations to be $x_1 = x_2 = 0$, then it implies the solutions to all 3 equations must be $x_1 = x_2 = 0$

ie,

if $\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} \neq 0 \implies (a_1, a_2), (b_1, b_2)$ are linearly independent

$\implies \vec{A}_1, \vec{A}_2$ are linearly independent

$$x_1 \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} + x_2 \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} + \cdots + x_r \begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$AX = B \implies \begin{bmatrix} a_1 & b_1 & \cdots & c_1 \\ a_2 & b_2 & \cdots & c_2 \\ \vdots & \vdots & & \vdots \\ a_n & b_n & \cdots & c_n \end{bmatrix}_{n\times r} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_r \end{bmatrix}_{r\times 1} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}_{n\times 1}$$

If $\vec{A}_1, \vec{A}_2, \ldots, \vec{A}_r$ are linearly independent and $n > r$, rref(A) will be of the form

$$\text{rref}(A) = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix}$$

Columns with pivot 1 are linearly independent and eliminating non-pivot rows (with zeros) will not affect the linearity of the column vectors.

ie., eliminating rows of 'A' corresp. to non-pivot rows of rref(A) will not affect the linearity of corresp. column vectors

The column rank of 'A' is the dimension of the column space of 'A', while the row rank of 'A' is the dimension of the row space of 'A'.

$$\boxed{\text{Row rank}(A) = \text{Column rank}(A) = \text{rank}(A)}$$

Proof

Let 'A' be an $m \times n$ matrix, & let the column rank of 'A' be 'r', and $\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_r$ be any basis for the column space of 'A'. Place these as the columns of an $m \times r$ matrix X.

→ Every column of 'A' can be expressed as a linear combination of the 'r' columns in 'X'. i.e., there is an $r \times n$ matrix B such that

$$A_{m \times n} = X_{m \times r} \, B_{r \times n}$$

$$\begin{array}{cccccc}
 & [\vec{a}_1\ \vec{a}_2 \cdots \vec{a}_n] & & [\vec{v}_1\ \vec{v}_2 \cdots \vec{v}_r] & [\vec{b}_1\ \vec{b}_2 \cdots \vec{b}_n] & \\
\begin{bmatrix}\vec{\alpha}_1\\ \vec{\alpha}_2\\ \vdots\\ \vec{\alpha}_m\end{bmatrix} &
\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ a_{21} & a_{22} & \cdots & a_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mn}\end{bmatrix}
= \begin{bmatrix}\vec{w}_1\\ \vec{w}_2\\ \vdots\\ \vec{w}_m\end{bmatrix} &
\begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1r}\\ x_{21} & x_{22} & \cdots & x_{2r}\\ \vdots & \vdots & & \vdots\\ x_{m1} & x_{m2} & \cdots & x_{mr}\end{bmatrix} &
\begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & & \vdots\\ b_{r1} & b_{r2} & \cdots & b_{rn}\end{bmatrix} &
\begin{bmatrix}\vec{\beta}_1\\ \vec{\beta}_2\\ \vdots\\ \vec{\beta}_r\end{bmatrix}
\end{array}$$

$$= \begin{bmatrix}
x_{11}b_{11}+x_{12}b_{21}+\cdots+x_{1r}b_{r1} & x_{11}b_{12}+x_{12}b_{22}+\cdots+x_{1r}b_{r2} & \cdots & x_{11}b_{1n}+x_{12}b_{2n}+\cdots+x_{1r}b_{rn}\\
x_{21}b_{11}+x_{22}b_{21}+\cdots+x_{2r}b_{r1} & x_{21}b_{12}+x_{22}b_{22}+\cdots+x_{2r}b_{r2} & \cdots & x_{21}b_{1n}+x_{22}b_{2n}+\cdots+x_{2r}b_{rn}\\
\vdots & \vdots & & \vdots\\
x_{m1}b_{11}+x_{m2}b_{21}+\cdots+x_{mr}b_{r1} & x_{m1}b_{12}+x_{m2}b_{22}+\cdots+x_{mr}b_{r2} & \cdots & x_{m1}b_{1n}+x_{m2}b_{2n}+\cdots+x_{mr}b_{rn}
\end{bmatrix}$$

$$a_{ik} = (xb)_{ik} = \sum_{j=1}^{r} x_{ij}\, b_{jk}$$

$$\begin{bmatrix}\vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n\end{bmatrix}_{m\times n} = \begin{bmatrix}\vec{v}_1 & \vec{v}_2 & \cdots & \vec{v}_r\end{bmatrix}_{m\times r} \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ b_{r1} & b_{r2} & \cdots & b_{rn}\end{bmatrix}_{r\times n}$$

$$a_i = \vec{v}_1 b_{1i} + \vec{v}_2 b_{2i} + \cdots + \vec{v}_r b_{ri}$$

$\Rightarrow$ i^{th} column of 'A' is a linear combination of the 'r' columns of 'X' with coeff from the i^{th} column of 'B'.

$$\begin{bmatrix}\vec{\alpha}_1\\ \vec{\alpha}_2\\ \vdots\\ \vec{\alpha}_m\end{bmatrix}$$

$\vec{\alpha}_j =$

$\Rightarrow$ j^{th} ro… the … j^{th} row

Definit…

Let … linearly … V. … vectors

$$\begin{bmatrix} \vec{\alpha}_1 \\ \vec{\alpha}_2 \\ \vdots \\ \vec{\alpha}_m \end{bmatrix} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1r} \\ x_{21} & x_{22} & \cdots & x_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ x_{m1} & x_{m2} & \cdots & x_{mr} \end{bmatrix} \begin{bmatrix} \vec{\beta}_1 \\ \vec{\beta}_2 \\ \vdots \\ \vec{\beta}_r \end{bmatrix}$$

$$\vec{\alpha}_j = \vec{\beta}_1 x_{j1} + \vec{\beta}_2 x_{j2} + \cdots + \vec{\beta}_r x_{jr}$$

$\Rightarrow$ j^{th} row of 'A' is a linear combination of the 'r' rows of 'B' with coeff. from the j^{th} row of 'X'.

Steinitz Exchange Lemma

Let $v_1, v_2, \ldots, v_m$ be a collection of linearly independent vectors in a vector space V. Let $w_1, w_2, \ldots, w_n$ be a collection of vectors that span V. Then $m \le n$.

$\Rightarrow$ row rank$(A) \leq r$

$\Rightarrow$ row rank$(A) \leq$ column rank(A)

$\Rightarrow$ row rank$(A^T) \leq$ column rank(A^T)

$\Rightarrow$ column rank$(A) \leq$ row rank(A)

Row rank(A) = column rank(A) = rank(A)

* Rank of a null matrix is zero

* $\operatorname{rank}(A^T) = \operatorname{rank}(A)$

* $\operatorname{rank}(A_{m \times n}) \leq \min(m, n)$

* The rank of the null matrix is zero, and the rank of every non-null matrix is greater than or equal to 1.

* A square matrix of order 'n' is non-singular if its rank $r = n$.

ie., $\boxed{|A| \neq 0 \text{, then } \operatorname{rank}(A) = n}$

* $\text{rank}(AB) \leq \min(\text{rank}(A), \text{rank}(B))$

Proof

$$C_{m\times n} = A_{m\times r}\, B_{r\times n}$$

$$\begin{array}{cc} & [\vec{\mu}_1\ \vec{\mu}_2 \cdots \vec{\mu}_n] \\ \begin{bmatrix}\vec{c}_1\\ \vec{c}_2\\ \vdots\\ \vec{c}_m\end{bmatrix} & \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n}\\ c_{21} & c_{22} & \cdots & c_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ c_{m1} & c_{m2} & \cdots & c_{mn}\end{bmatrix}\end{array} = \begin{array}{cc} & [\vec{\alpha}_1\ \vec{\alpha}_2 \cdots \vec{\alpha}_r] \\ \begin{bmatrix}\vec{a}_1\\ \vec{a}_2\\ \vdots\\ \vec{a}_m\end{bmatrix} & \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1r}\\ a_{21} & a_{22} & \cdots & a_{2r}\\ \vdots & \vdots & \ddots & \vdots\\ a_{m1} & a_{m2} & \cdots & a_{mr}\end{bmatrix}\end{array} \begin{array}{cc} [\vec{\beta}_1\ \vec{\beta}_2 \cdots \vec{\beta}_n] & \\ \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ b_{r1} & b_{r2} & \cdots & b_{rn}\end{bmatrix} & \begin{bmatrix}\vec{b}_1\\ \vec{b}_2\\ \vdots\\ \vec{b}_r\end{bmatrix}\end{array}$$

$$[\vec{\mu}_1\ \vec{\mu}_2 \cdots \vec{\mu}_n] = [\vec{\alpha}_1\ \vec{\alpha}_2 \cdots \vec{\alpha}_r] \begin{bmatrix} b_{11} & b_{12} & \cdots & b_{1n}\\ b_{21} & b_{22} & \cdots & b_{2n}\\ \vdots & \vdots & \ddots & \vdots\\ b_{r1} & b_{r2} & \cdots & b_{rn}\end{bmatrix}$$

$$\mu_i = \vec{\alpha}_1 b_{1i} + \vec{\alpha}_2 b_{2i} + \cdots + \vec{\alpha}_r b_{ri}$$

$\Rightarrow$ Each column of $C = AB$ is a combination of the columns of A

$\Rightarrow$ Column space of $(AB) \subseteq$ column space (A)

$$\Longrightarrow \dim\left(\text{colum space}(AB)\right) \le \dim\left(\text{column space}(A)\right)$$

$$\Longrightarrow \text{rank}(AB) \le \text{rank}(A)$$

* If the matrix B is non-singular

$$\boxed{\begin{aligned} |B| \neq 0 \quad &\Longrightarrow \text{rank}(AB) = \text{rank}(A) \\ |A| \neq 0 \quad &\Longrightarrow \text{rank}(AB) = \text{rank}(B) \end{aligned}}$$

$\text{rank}(A) = r = \text{rank}(A^{-1})$

$A_{r \times r} \, B_{r \times m} = (AB)_{r \times m}$

$\text{rank}(AB) \le \text{rank}(A^{-1})$

Proof

$|A| \neq 0 \Rightarrow A^{-1}$ exists

$\min[\text{rank}(A^{-1}), \text{rank}(AB)]$

$\text{rank}(B) = \min[\text{rank}(A), \text{rank}(B)]$

$$\text{rank}\left(A^{-1}(AB)\right) \le \text{rank}(AB)$$

$$\text{rank}(B) = \text{rank}\left(A^{-1}(AB)\right) \le \text{rank}(AB) \le \text{rank}(B)$$

$$\Longrightarrow \underline{\text{rank}(AB) = \text{rank}(B)}$$

* If matrix A is real,

$$\operatorname{rank}(AA^T) = \operatorname{rank}(A^T A) = \operatorname{rank}(A)$$

Proof

Let $x \in N(A)$,

$$Ax = 0 \longrightarrow A^T A x = 0$$
$$\longrightarrow x \in N(A^T A)$$

$$\therefore N(A) \subset N(A^T A)$$

Let $x \in N(A^T A)$,

$$A^T A x = 0 \longrightarrow x^T A^T A x = 0$$
$$\longrightarrow (Ax)^T (Ax) = 0$$
$$\longrightarrow Ax = 0 \longrightarrow x \in N(A)$$

$$\therefore N(A^T A) \subset N(A)$$

$$\Longrightarrow N(A^T A) = N(A)$$

$$\Longrightarrow \dim(N(A^T A)) = \dim(N(A))$$

$$\Longrightarrow \operatorname{rank}(A^T A) = \operatorname{rank}(A)$$

$A^T A$

$Ax \in C$

$N(A^T)$

If A h

Ax

$\therefore A^T A$

[illegible] $A^TAx = 0$

$Ax \in C(A)$ & $Ax \in N(A^T)$

$N(A^T) \perp C(A) \implies Ax = 0$

If A has linearly independent columns

$Ax = 0 \longrightarrow x = 0$

$\therefore A^TAx = 0 \implies x = 0 \Big\}$ $N(A^TA) = \{0\}$

$\therefore A^TA$ is invertible.

* If matrices A & B are of the same order

$$\boxed{\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B)}$$

Proof

Let $a_1, a_2, \ldots, a_r$ and $b_1, b_2, \ldots, b_s$ denote linearly independent columns of A and B respectively.

$$\text{rank}(A) = \dim(\text{column space}(A)) = \dim(\text{span}(a_1, a_2, \ldots, a_r)) = r$$

$$\text{rank}(B) = \dim(\text{column space}(B)) = \dim(\text{span}(b_1, b_2, \ldots, b_s)) = s$$

then,

$\{a_1, a_2, \ldots, a_r, b_1, b_2, \ldots, b_s\}$ span the column space of $A+B$.

Any vector (column) of $A+B$ is a linear combination of vectors from the above set.

$c_1, c_2, \ldots, c_k$ be linear independent columns of $C = A+B$,

Hence,

$$k \leq r + s$$

$$\text{rank}(A+B) \leq \text{rank}(A) + \text{rank}(B).$$

(OR) Any vector $\vec{v} \in \text{span}(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)$

$$\vec{v} = r_1(a_1+b_1) + r_2(a_2+b_2) + \cdots + r_n(a_n+b_n)$$

$$= (r_1 a_1 + r_2 a_2 + \cdots + r_n a_n) + (r_1 b_1 + r_2 b_2 + \cdots + r_n b_n)$$

$$\in \text{span}(a_1, a_2, \ldots, a_n) + \text{span}(b_1, b_2, \ldots, b_n)$$

$\therefore$

$$\text{span}(a_1+b_1, a_2+b_2, \ldots, a_n+b_n) \subseteq \text{span}(a_1, a_2, \ldots, a_n) + \text{span}(b_1, b_2, \ldots, b_n)$$

For 2 subspaces U and V of a vector space, then

$$\dim(U+V) \leq \dim(U) + \dim(V)$$

$\therefore$

$$\begin{aligned} \text{rank}(A+B) &= \dim\left(\text{span}(a_1+b_1, a_2+b_2, \ldots, a_n+b_n)\right) \\ &\leq \dim\left(\text{span}(a_1, a_2, \ldots, a_n)\right) + \dim\left(\text{span}(b_1, b_2, \ldots, b_n)\right) \\ &= \text{rank}(A) + \text{rank}(B). \end{aligned}$$

Computing rank using Gaussian Elimination

* The rank of a matrix is unchanged under elementary operations on its rows and columns.

Idea of proof →

$|E_{ij}| = -1$; $|E_i(k)| = k$; $|E_{ij}(k)| = 1$

Elementary matrices responsible for elementary row & column operations are non-singular or invertible.

$|E| \neq 0 \longrightarrow \text{rank}(EA) = \text{rank}(AE) = \text{rank}(A)$

- switching a rows is just switching the order of the vectors, it doesn't change the span of the set similarly for multiplying by a scalar ($\neq 0$) one of the rows.

say, $E_{12}(k) A = A'$

Any vector $\vec{v} \in \text{span}(v_1 + kv_2, v_2, \ldots, v_n) = \text{rowspan}(A')$

$\vec{v} = a_1(v_1 + kv_2) + a_2 v_2 + \cdots + a_n v_n$

$= a_1 v_1 + (a_2 + k) v_2 + \cdots + a_n v_n \in \text{span}(v_1, v_2, \ldots, v_n)$

→ adding a scalar times one row to another doesn't change the row space.

Gaussian elimination - rank.

* Use elementary row operations to reduce 'A' to echelon form.

The rank of 'A' is the # of pivots or leading coefficients in the echelon form. In fact, the pivot columns (columns with pivots in them) are linearly independent.

- It is not necessary to find the reduced echelon form - any echelon form will do since only the pivots matter.

Echelo

- All non
all zero
- Each
entry)
of the
- All e
entry an

Ex:

Echelon form (or Row Echelon form) - REF

- All non-zero rows are above any rows of all zeros.
- Each leading entry (i.e., left most non-zero entry) of a row is in a column to the right of the leading entry of the row above it.
- All entries in a column below a leading entry are zero.

Ex: $\begin{bmatrix} \blacksquare & * & * & * & * \\ 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$, $\begin{bmatrix} \blacksquare & * & * \\ 0 & \blacksquare & * \\ 0 & 0 & \blacksquare \\ 0 & 0 & 0 \end{bmatrix}$

$$\begin{bmatrix} 0 & \blacksquare & * & * & * & * & * & * & * & * & * \\ 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & \blacksquare & * & * & * & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \blacksquare & * & * \end{bmatrix}$$

Reduced Row Echelon form (RREF)

The following conditions are added to the conditions for REF.

- The leading entry in each non-zero row is 1
- Each leading 1 is the only non-zero entry in its column.

Ex:

$$\begin{bmatrix} 0 & 1 & * & 0 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 1 & 0 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 1 & * & * & 0 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & * & * \end{bmatrix}$$

* Uniqueness of RREF : Each matrix is row-equivalent to one and only one RRE matrix.

i.e., no matter how one gets to it, the RREF of every matrix is unique.

Examples

1. $$\begin{bmatrix} 0 & 3 & -6 & 6 & 4 & -5 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 3 & -9 & 12 & -9 & 6 & 15 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 3 & -7 & 8 & -5 & 8 & 9 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 2 & -4 & 4 & 2 & -6 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix}$$

$$\rightarrow \begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 3 & -6 & 6 & 4 & -5 \end{bmatrix} \rightarrow \underbrace{\begin{bmatrix} 3 & -9 & 12 & -9 & 6 & 15 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}}_{\text{row echelon form (REF)}}$$

$$\rightarrow \begin{bmatrix} 3 & 0 & -6 & 9 & 15 & -12 \\ 0 & 1 & -2 & 2 & 1 & -3 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -2 & 3 & 5 & -4 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}$$

$$\rightarrow \underbrace{\begin{bmatrix} 1 & 0 & -2 & 3 & 0 & -24 \\ 0 & 1 & -2 & 2 & 0 & -7 \\ 0 & 0 & 0 & 0 & 1 & 4 \end{bmatrix}}_{\text{RREF}}$$

* Two rectangular $m \times n$ matrices A and B are called equivalent if $B = Q^{-1}AP$ for some invertible $n \times n$ matrix P and some invertible $m \times m$ matrix Q.

- Matrix equivalence is an equivalence relation on the space of rectangular matrices.

For 2 rectangular matrices of the same size, their equivalence can also be characterized by the following conditions:

(i) The matrices can be transformed into one another by a combination of elementary row & column operations.

(ii) Two matrices are equivalent iff they have the same rank.

If B is equivalent to A

Idea of proof

For some elementary matrices $E_1, E_2, \cdots, E_k, F_1, F_2, \cdots, F_l$ we have

$$B = (E_k E_{k-1} \cdots E_2 E_1) A (F_1 F_2 \cdots F_{l-1} F_l)$$

where,

$$Q^{-1} = E_k E_{k-1} \cdots E_2 E_1 \longrightarrow Q = E_1 E_2 \cdots E_k$$

and $P = E_1 E_2 \cdots E_l$ are invertible matrices.

* If A is a non-singular matrix, then there exists elementary matrices $E_1, E_2, \cdots, E_k$ such that $A = E_k E_{k-1} \cdots E_2 E_1 I$

$$= E_k E_{k-1} \cdots E_2 E_1$$

$$B = Q^{-1} A P$$

* Similar

Let

be simila

* Similar matrices are equivalent,

But equivalent square matrices need not be similar.

Rouché–Capelli theorem

Consider the system $Ax=b$ with coefficient matrix A and augmented matrix $[A|b]$. The sizes of b, A and $[A|b]$ are $m\times 1$, $m\times n$ and $m\times(n+1)$, respectively; in addition, the # of unknowns is n.

The possibilities for solving the system are:

(i) $Ax=b$ has a unique solution iff

$$\operatorname{rank}[A] = \operatorname{rank}[A|b] = n$$

(ii) $Ax=b$ is inconsistent (ie; no solution exists) iff $\operatorname{rank}[A] < \operatorname{rank}[A|b]$

(iii) $Ax=b$ has infinitely many solutions iff

$$\operatorname{rank}[A] = \operatorname{rank}[A|b] < n$$

Intuitive proof

$Ax = b$ has solution(s) iff there are $x_1, x_2, \ldots, x_n$ such that:

quick 27/6/20

$\vec{A}_1 x_1 + \vec{A}_2 x_2 + \cdots + \vec{A}_n x_n = b$

$\vec{b}$ is a linear combination of the column vectors.

$\Rightarrow$ rank$[A]$ = rank$[A|b]$

If the system is inconsistent
ie., $Ax = b$ has no solution

There are no x_i such that $\sum \vec{A}_i x_i = b$

$\therefore$ b is linearly independent

$\Rightarrow$ rank$[\vec{A}]$ < rank$[A|b]$

Ex:-

$[A|b] =$

$\rightarrow \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\rightarrow \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

$x +$

Take,

$x =$

$y =$

Ex:- $3x - y + 7z = 1$

$6x + z = 2$

$$[A|b] = \left[\begin{array}{ccc|c} 3 & -1 & 7 & 1 \\ 6 & 0 & 1 & 2 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & -1/3 & 7/3 & 1/3 \\ 1 & 0 & 1/6 & 1/3 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & -1/3 & 7/3 & 1/3 \\ 0 & 1/3 & -13/6 & 0 \end{array}\right] \to \left[\begin{array}{ccc|c} 1 & 0 & 1/6 & 1/3 \\ 0 & 1/3 & -13/6 & 0 \end{array}\right]$$

$$\to \left[\begin{array}{ccc|c} 1 & 0 & 1/6 & 1/3 \\ 0 & 1 & -13/2 & 0 \end{array}\right] = \text{rref}[A|b]$$

$x + \frac{1}{6} z = \frac{1}{3}$ & $y - \frac{13}{2} z = 0$

Take, $z = \lambda \in \mathbb{R}$

$$x = \frac{1}{3} - \frac{1}{6}\lambda = \frac{1}{3}\left(1 - \frac{\lambda}{2}\right)$$

$$y = \frac{13}{2}\lambda$$

$\text{rank}[A|b] = \text{rank}[A]$

Gaussian elimination $\Longrightarrow$

\# of pivots of rref(A) = r

$\therefore$ rank(A) = r

The "pivot columns" will corresponds to unknowns and "non-pivot columns" to parameters.

ie.,
the set of solutions will depend upon $n - r = n - \text{rank}(A)$ parameters.

The # of solutions are infinite iff there are parameters.

ie., $n - \text{rank}(A) > 0 \Rightarrow \text{rank}(A) < n$

Unitary matrix

A complex square matrix U is unitary if its conjugate transpose $U^{\dagger} = \bar{U}^{T}$ is also its inverse.

$$U^{\dagger} = \bar{U}^{T} = U^{-1} \iff U^{\dagger}U = UU^{\dagger} = I$$

* An $n \times n$ sq. matrix is unitary iff its columns form an orthonormal basis in $\mathbb{C}^n$.

whereas,

A real $n \times n$ sq. matrix is orthogonal iff its columns form an orthonormal basis in $\mathbb{R}^n$.

Proof Let $U = [u_1 \; u_2 \; \cdots \; u_n]$

$$U^{\dagger} = \begin{bmatrix} \bar{u}_1^{T} \\ \bar{u}_2^{T} \\ \vdots \\ \bar{u}_n^{T} \end{bmatrix} = \begin{bmatrix} u_1^{\dagger} \\ u_2^{\dagger} \\ \vdots \\ u_n^{\dagger} \end{bmatrix}$$

$$U^{\dagger}U = \begin{bmatrix} \bar{u}_1^{T} \\ \bar{u}_2^{T} \\ \vdots \\ \bar{u}_n^{T} \end{bmatrix} \begin{bmatrix} u_1 & u_2 & \cdots & u_n \end{bmatrix}$$

$$= \begin{bmatrix} \bar{u}_1^{T}u_1 & \bar{u}_1^{T}u_2 & \cdots & \bar{u}_1^{T}u_n \\ \bar{u}_2^{T}u_1 & \bar{u}_2^{T}u_2 & \cdots & \bar{u}_2^{T}u_n \\ \vdots & \vdots & \ddots & \vdots \\ \bar{u}_n^{T}u_1 & \bar{u}_n^{T}u_2 & \cdots & \bar{u}_n^{T}u_n \end{bmatrix}$$

$$= \begin{bmatrix} \langle u_1|u_1\rangle & \langle u_1|u_2\rangle & \cdots & \langle u_1|u_n\rangle \\ \langle u_2|u_1\rangle & \langle u_2|u_2\rangle & \cdots & \langle u_2|u_n\rangle \\ \vdots & \vdots & \ddots & \vdots \\ \langle u_n|u_1\rangle & \langle u_n|u_2\rangle & \cdots & \langle u_n|u_n\rangle \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix} = I$$

$U^{\dagger}U = I$

$$\bar{u}_i^T u_j = u_i \cdot u_j = \langle u_i | u_j \rangle = \delta_{ij} = \begin{cases} 1 ; & i = j \\ 0 ; & i \neq j \end{cases}$$

* Given 2 complex vectors x and y, multiplication by a Unitary matrix U preserves their inner product.

$$\boxed{\langle Ux | Uy \rangle = \langle x | y \rangle \; ; \; x, y \in \mathbb{C}^n}$$

* Unitary matrix preserves an inner product in the space $\mathbb{C}^n$.

$\Rightarrow$ Unitary matrices preserves the norms of the vectors (complex)

$$\boxed{|Ux| = |x| \quad \forall \; x \in \mathbb{C}^n}$$

Proof

① $\langle Ux | Uy \rangle = (Ux) \cdot (Uy) = \overline{(Ux)}^T Uy$

$= \bar{x}^T \bar{U}^T U y = \bar{x}^T U^\dagger U y$

$= \bar{x}^T I y = \bar{x}^T y = x \cdot y = \langle x | y \rangle$

* A square matrix A is normal if it commutes with its conjugate-transpose.

$$\boxed{A \text{ is normal} \Rightarrow \begin{array}{c} A^{\dagger}A = AA^{\dagger} \\ [A, A^{\dagger}] = 0 \end{array}}$$

Ex:- Unitary matrices ($A^{\dagger} = A^{-1}$ (or) $AA^{\dagger} = A^{\dagger}A = I$)

Hermitian matrices ($A^{\dagger} = A$)

Skew-Hermitian matrices ($A^{\dagger} = -A$)

*

* A matrix is normal iff it is unitarily diagonalizable

Proof for: If 'A' is unitarily diagonalizable, then 'A' must be normal.

$$U^{\dagger} = U^{-1} \Leftrightarrow UU^{\dagger} = U^{\dagger}U = I$$

$AU = UA \Longrightarrow$ An

$$A = UDU^{-1} = UDU^{\dagger}$$

$$AA^{\dagger} = (UDU^{\dagger})(UDU^{\dagger})^{\dagger} = (UDU^{\dagger})((U^{\dagger})^{\dagger}D^{\dagger}U^{\dagger})$$

$$= (UDU^{\dagger})(UD^{\dagger}U^{\dagger})$$

$$= UD(U^{\dagger}U)D^{\dagger}U^{\dagger} = UDID^{\dagger}U^{\dagger}$$

$$= UDD^{\dagger}U^{\dagger} = UD^{\dagger}DU^{\dagger}$$ [diagonal matrices commute each other

$$= UD^{\dagger}IDU^{\dagger} = UD^{\dagger}U^{\dagger}UDU^{\dagger}$$

$$= A^{\dagger}A$$

$$\boxed{A \text{ is unitarily diagonalizable} \Longrightarrow A \text{ is normal}}$$

Proof for

Any com

$A = \frac{A+}{}$

$A =$

$\frac{A-A^{\dagger}}{2}$: skew-he

$-i\left(\frac{A-A^{\dagger}}{2}\right) =$

A is norm

$A^{\dagger}A =$

$0 = AA^{\dagger} - A$

since 2

they can

simultaneously

$B = VD_B V^{\dagger}$

page 2

Proof for : If A is normal then it can be diagonalizable by a unitary similarity transformation.

Any complex matrix 'A' can be uniquely written as:

$$A = \frac{A+A^\dagger}{2} + \frac{A-A^\dagger}{2} = \frac{A+A^\dagger}{2} + i\left(-i\,\frac{A-A^\dagger}{2}\right)$$

$$A = B + iC \quad , \text{ where } B,C : \text{hermitian matrices}$$

$$\boxed{\begin{array}{l} \frac{A-A^\dagger}{2} : \text{skew-hermitian} \\ C = -i\left(\frac{A-A^\dagger}{2}\right) : \text{hermitian} \end{array}}$$

$$B = \frac{A+A^\dagger}{2} \; ; \; C = -i\,\frac{A-A^\dagger}{2}$$

$$B = B^\dagger \; ; \; C = C^\dagger$$

A is normal :

$$A^\dagger A = A A^\dagger$$

$$0 = AA^\dagger - A^\dagger A = (B+iC)(B-iC) - (B-iC)(B+iC)$$

$$= 2i\,(CB - BC) \implies \underline{\underline{CB = BC}}.$$

Since 2 hermitian matrices B and C commute, they can be simultaneously diagonalized by a ~~similarity~~ unitary similarity transformation.

$B = VD_BV^\dagger$ and $C = VD_CV^\dagger$ where D_B, D_C : diagonal matrices.

Next @ page 2

$$A = B + iC = VD_BV^{\dagger} + i\,VD_CV^{\dagger}$$

$$= V\underbrace{\left[D_B + iD_C\right]}_{\text{diagonal matrix}}V^{\dagger}$$

$\therefore$ Any normal matrix can be diagonalizable by a unitary transformation.

---

* $A = V\left[D_B + iD_C\right]V^{\dagger}$ $\Longrightarrow$ the eigen values of a normal matrix are in general complex. (in contrast to the real eigenvalues of a hermitian matrix).

* Two he...
diagonali...
iff th...

$A^{\dagger} = A$ & ...
(Hermitian...
&
$AB = BA$

Gi...

Proof

Part 1: ...

$AB = V$...

$= V$...

$= V^{\dagger}$...

$=$ ...

* Two hermitian matrices are simultaneously diagonalizable by a unitary similarity transformation iff they commute

$A^{\dagger}=A$ & $B^{\dagger}=B$ (Hermitian) & $AB=BA$ $\Longleftrightarrow$ $A=VD_AV^{\dagger}$, $B=VD_BV^{\dagger}$

where, D_A, D_B: diagonal matrices

$\Downarrow$

Eigenvectors of A, B coincide

Proof

(a) Part 1: $A=VD_AV^{\dagger}$ & $B=VD_BV^{\dagger}$

$$AB=VD_AV^{\dagger}VD_BV^{\dagger}=VD_AI\,D_BV^{\dagger}$$

$$=VD_AD_BV^{\dagger}=VD_BD_AV^{\dagger}$$

[diagonal matrices commute each other

$$=VD_BV^{\dagger}VD_AV^{\dagger}$$

$$=BA$$

Part 2:

If 2 hermitian matrices A & B commute, ~~then~~
~~A and B can be simultaneously diagonalizable~~

$AB = BA$

$A^\dagger = A \longrightarrow$ we can find a unitary matrix U such that $A = UD_AU^\dagger \Longleftrightarrow U^\dagger A U = D_A$

D_A : diagonal-matrix

Using the same matrix U, we shall find

$$B' = U^\dagger B U$$

$$U^\dagger A U = D_A = \begin{bmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & a_n \end{bmatrix}$$

$$U^\dagger B U = B' = \begin{bmatrix} b'_{11} & b'_{12} & \cdots & b'_{1n} \\ b'_{21} & b'_{22} & \cdots & b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b'_{n1} & b'_{n2} & \cdots & b'_{nn} \end{bmatrix}$$

$$(B')^{\dagger} = (U^{\dagger}BU)^{\dagger} = U^{\dagger}B^{\dagger}U = U^{\dagger}BU = B'$$

$\longrightarrow$ B' is hermitian

$$D_A B' = (U^{\dagger}AU)(U^{\dagger}BU) = U^{\dagger}ABU = U^{\dagger}BAU$$
$$= (U^{\dagger}BU)(U^{\dagger}AU) = B'D_A$$

$$D_A B' = \begin{bmatrix} a_1 b'_{11} & a_1 b'_{12} & \cdots & a_1 b'_{1n} \\ a_2 b'_{21} & a_2 b'_{22} & \cdots & a_2 b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_n b'_{n1} & a_n b'_{n2} & \cdots & a_n b'_{nn} \end{bmatrix} ; \quad B'D_A = \begin{bmatrix} a_1 b'_{11} & a_2 b'_{12} & \cdots & a_n b'_{1n} \\ a_1 b'_{21} & a_2 b'_{22} & \cdots & a_n b'_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_n b'_{n1} & a_2 b'_{n2} & \cdots & a_n b'_{nn} \end{bmatrix}$$

$$D_A B' - B'D_A = 0 \Longrightarrow (a_i - a_j)\, b'_{ij} = 0$$

If all the a_i were distinct,

$$b'_{ij} = 0 \quad \text{for} \quad i \neq j$$

$\Rightarrow$ B' is diagonal

Case 2: If some of the diagonal elements are equal.

i.e., some of the eigenvalues of A are degenerate

we can order the columns of U so that the degenerate eigenvalues are continuous along the diagonal.

Henceforth,

we assume this to be the case:

$\therefore \; b_{ij} = 0 \quad \text{for} \quad a_i \neq a_j$

D_A & B' in block matrix form:

$$D_A = \begin{bmatrix} \lambda_1 I_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 I_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_k I_k \end{bmatrix} ; \quad B' = \begin{bmatrix} B_1' & 0 & \cdots & 0 \\ 0 & B_2' & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_k' \end{bmatrix}$$

assuming that, A possesses 'k' distinct eigenvalues.

I_j : identity matrix whose dimension is equal to the multiplicity of the corresp. eigenvalue λ_j.

B' is hermitian $\longrightarrow$ B'_j is hermitian with the same dimension as I_j.

$\Rightarrow B'_j$ can be diagonalized by a unitary similarity transformation.

i.e., we can find a unitary matrix of the form

$$U' = \begin{bmatrix} U'_1 & 0 & \cdots & 0 \\ 0 & U'_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & U'_k \end{bmatrix}$$

such that $(U')^{\dagger} B' U' = D_B$; diagonal matrix

$$(U')^\dagger D_A U' = D_A = (U')^\dagger U^\dagger A U U'$$

$$\Longrightarrow (UU')^\dagger A (UU') = D_A$$

$$(U')^\dagger B' U' = (U')^\dagger U^\dagger B U U' = D_B.$$

$$\Longrightarrow (UU')^\dagger B (UU') = D_B.$$

$\therefore$ We have an invertible matrix $V = UU'$ such that

$$V^\dagger A V = D_A \quad \text{and} \quad V^\dagger B V = D_B$$

$\therefore$ A & B are simultaneously diagonalizable by a unitary similarity transformation.

$\therefore$ Columns of V are the simultaneous eigenvectors of A and B.

(*) 2 diagonalizable matrices are simultaneously diagonalizable iff they commute.

Proof

Suppose that A and B are $n \times n$ matrices, with complex entries, which are both diagonalizable.

Assume that A and B commute, ie.,

$[A, B] = AB - BA = 0 \longrightarrow AB = BA$.

We decompose $\mathbb{C}^n$ as a direct sum of eigenspaces of 'A', ie, $\mathbb{C}^n = E_{\lambda_1} \oplus E_{\lambda_2} \oplus \cdots \oplus E_{\lambda_m}$, where $\lambda_1, \lambda_2, \ldots, \lambda_m$ are the eigenvalues of A, and E_{λ_i} is the eigenspace for λ_i. (Here $m \leq n$, since some eigenspaces could be of dimension bigger than one).

* A square matrix 'A' is called diagonalizable if it is similar to a diagonal matrix
i.e., if there exists an invertible matrix P and a diagonal matrix D such that
$P^{-1}AP = D \Leftrightarrow A = PDP^{-1}$.

$\Downarrow$

An $n \times n$ matrix 'A' over a field $\mathbb{F}$ is diagonalizable if and only if the sum of the dimensions of its eigenspaces is equal to n, which is the case if and only if there exists a basis of $\mathbb{F}^{n \times n}$ consisting of eigenvectors of 'A'. If such a basis has been found, one can form the matrix P having these basis vectors as columns, and $P^{-1}AP$ will be a diagonal matrix whose diagonal entries are the eigenvalues of 'A'.

Choose a basis $X = \{x_1, x_2, \ldots, x_m\}$ for this eigenspace E_{λ_i}, of corresponding eigenvectors x_j such that $Ax_j = \lambda_i x_j$ for all $i = 1, 2, \ldots, m$.

$$A(Bx_j) = ABx_j = BAx_j = B\lambda_i x_j = \lambda_i (Bx_j)$$

$\Rightarrow$ x_j and Bx_j are eigenvectors of A with eigenvalue λ_i

ie, ~~S~~ince B commutes A,
A preserves each of the E_{λ_i}

ie,
B takes eigenvectors of A to new eigenvectors of A with same eigenvalue.

Consider the $m \times m$ matrix B', which the matrix B (restricted to E_{λ_i}) in terms of the basis $\{x_j\}_{j=1}^{m}$.

$$B' = P_{X \leftarrow E} \, B \, P_{X \leftarrow E}^{-1}$$

$$= X^{-1} B X$$

B' has 'm' eigenvectors

Let B' has an eigenvalue μ and also an eigenvector y' i.e.,

$$B'y' = \mu y'$$

$$X^{-1} B X y' = \mu y'$$

$$B X y' = \mu X y' \Rightarrow By = \mu y$$

where $y = Xy'$ is the same eigenvector in the standard basis.

$y, y' \in E_{\lambda_i}$

$$y = \sum_j a_j x_j$$

$$\Rightarrow Ay = A\left(\sum_j a_j x_j\right) = \sum_j a_j A x_j$$

$$= \sum_j a_j \lambda_i x_j = \lambda_i \left(\sum_j a_j x_j\right)$$

$$= \lambda_i y$$

$\Rightarrow y$ is an eigenvector of both A & B

$\therefore$ Diagonalizing B' obtains a new basis for E_{λ_i} consisting of simultaneous eigenvector for A and B.

* $|\lambda| = 1$ for every eigenvalue λ of a unitary matrix U.

ie.,

U is a normal matrix with eigenvalues lying on the unit circle.

Proof

$$U^{-1} = U^{\dagger} \Longrightarrow UU^{\dagger} = U^{\dagger}U = I$$

$$|Uv|^2 = |v|^2$$

$$|Uv|^2 = (Uv)^{\dagger} \cdot Uv = (Uv)^{\dagger}(Uv) = (U|v\rangle)^{\dagger} U|v\rangle$$

$$= \langle v|U^{\dagger}U|v\rangle = \langle v|v\rangle = |v|^2$$

$$U|v\rangle = \lambda|v\rangle \Longrightarrow \langle v|U^{\dagger} = \langle v|\lambda^*$$

$$\langle v|v\rangle = \langle v|U^{\dagger}U|v\rangle = \langle v|\lambda^* \lambda|v\rangle = |\lambda|^2 \langle v|v\rangle$$

$$\Longrightarrow |\lambda|^2 = 1 \longrightarrow |\lambda| = 1$$

- A normal matrix is unitary iff all its eigenvalues have absolute value one.

* $|\det U| = 1$

the determinant of a unitary matrix lie on the unit circle.

Proof $\quad U^{-1} = U^{\dagger} \Rightarrow UU^{\dagger} = 1$

~~det~~

$$1 = \det I = \det UU^{-1} = \det UU^{\dagger} = \det U \cdot \det U^{\dagger} = \det U \cdot \overline{\det U}$$

$$= |\det U|^2$$

$$\Longrightarrow |\det U| = 1$$

$$\det(U^{\dagger}) = \overline{\det U}$$

* For

a theor

Proof –

Let

$A = iH$

$A^{\dagger} =$

$\Longrightarrow$

e^{A}

$(e^{A})^{\dagger} =$

$e^{A}(e^{A})^{\dagger} =$

$= U$

$= U$

* For any unitary matrix U, there exist a hermitian matrix H such that $U = e^{iH}$

Proof – Partial : H hermitian $\longrightarrow iH$ skew-hermitian $\longrightarrow$ e^{iH} is unitary

Let

$A = iH$: skew-hermitian

$A^{\dagger} = -A$

$\Longrightarrow A = U D_A U^{-1}$ $\qquad U^{\dagger} = U^{-1}$

$e^{A} = U e^{D_A} U^{-1}$

$(e^{A})^{\dagger} = (U e^{D_A} U^{-1})^{\dagger} = \cancel{(\;)^{\dagger}} (U e^{D_A} U^{\dagger})^{\dagger}$

$= U (e^{D_A})^{\dagger} U^{\dagger} = U \, \mathrm{diag}(\overline{e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}}) U^{\dagger}$

$= U \, \mathrm{diag}(e^{-\lambda_1}, \ldots, e^{-\lambda_n}) U^{\dagger}$ $\quad$ [eigenvalues of skewhermitian matrix are purely imaginary

~~Now~~

$e^{A} (e^{A})^{\dagger} = U e^{D} U^{\dagger} U (e^{D_A})^{\dagger} U^{\dagger} = U e^{D} (e^{D_A})^{\dagger} U^{\dagger}$

$= U \, \mathrm{diag}(e^{\lambda_1} \ldots e^{\lambda_n}) \, \mathrm{diag}(e^{-\lambda_1} \ldots e^{-\lambda_n}) U^{\dagger}$

$= U \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{bmatrix} \begin{bmatrix} e^{-\lambda_1} & & \\ & \ddots & \\ & & e^{-\lambda_n} \end{bmatrix} U^{\dagger} = U U^{\dagger} = 1$

$\Longrightarrow e^{A}$ is unitary

SU(2) group

The special unitary group SU(2) is defined as:

$$\boxed{SU(2) := \left\{ g \in GL(2, \mathbb{C}) \,\middle|\, gg^{\dagger} = I_2, \det(g) = 1 \right\}}$$

Let, $U = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ with $a, b, c, d \in \mathbb{C}$

and $UU^{H} = I$ (or) $U^{H} = U^{-1}$

ILA (15)
9.2(22)

$$\begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix} = \frac{1}{\det(U)} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$\det(U) = 1$,

$$\begin{bmatrix} a^* & c^* \\ b^* & d^* \end{bmatrix} = \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

$d = a^*$, $c = -b^*$

$$U = \begin{bmatrix} a & b \\ -b^* & a^* \end{bmatrix} = \begin{bmatrix} \epsilon_0 - i\epsilon_3 & -\epsilon_2 - i\epsilon_1 \\ \epsilon_2 - i\epsilon_1 & \epsilon_0 + i\epsilon_3 \end{bmatrix}$$

where, $\epsilon_{0,1,2,3}$ are Cayley-Klein parameters

and $(\epsilon_0)^2 + (\epsilon_1)^2 + (\epsilon_2)^2 + (\epsilon_3)^2 = 1$

$\therefore$ we only have 3 free parameters to describe a 2×2 unimodular unitary matrix.

$\epsilon_0 I - i$

$=$

$=$

$\boxed{U = \epsilon_0 I \quad = \epsilon_0 I}$

where,

$\sigma_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$\sigma_i \cdot \sigma_j =$

$[\sigma_i, \sigma_j] =$

$\{\sigma_i, \sigma_j\} =$

$$\varphi_0 I - i\left(\varphi_1\sigma_1 + \varphi_2\sigma_2 + \varphi_3\sigma_3\right) =$$

$$= \varphi_0\begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix} - i\varphi_1\begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix} - i\varphi_2\begin{bmatrix}0 & -i\\ i & 0\end{bmatrix} - i\varphi_3\begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

$$= \begin{bmatrix}\varphi_0 - i\varphi_3 & -\varphi_2 - i\varphi_1\\ \varphi_2 - i\varphi_1 & \varphi_0 + i\varphi_3\end{bmatrix} = U$$

$$\boxed{\begin{aligned} U &= \varphi_0 I - i\left(\varphi_1\sigma_1 + \varphi_2\sigma_2 + \varphi_3\sigma_3\right)\\ &= \varphi_0 I - i\,\vec{\varphi}\cdot\vec{\sigma}\end{aligned}}$$

where, σ_i are the Pauli matrices

$$\sigma_1 = \begin{bmatrix}0 & 1\\ 1 & 0\end{bmatrix},\quad \sigma_2 = \begin{bmatrix}0 & -i\\ i & 0\end{bmatrix},\quad \sigma_3 = \begin{bmatrix}1 & 0\\ 0 & -1\end{bmatrix}$$

$$\sigma_i\sigma_j = \delta_{ij} I + i\epsilon_{ijk}\sigma_k$$

$$\left[\sigma_i, \sigma_j\right] = 2i\epsilon_{ijk}\sigma_k$$

$$\left\{\sigma_i, \sigma_j\right\} = \sigma_i\sigma_j + \sigma_j\sigma_i = 2\delta_{ij} I$$

$$\sigma_x^2 = \sigma_y^2 = \sigma_z^2 = I$$

$$[\sigma_x, \sigma_y] = \sigma_x\sigma_y - \sigma_y\sigma_x = 2i\sigma_z$$

$$\sigma_x\sigma_y = i\sigma_z ; \quad \sigma_y\sigma_x = -i\sigma_z$$

$$\sigma_x\sigma_y = -\sigma_y\sigma_x$$

$$|\epsilon|^2 = \epsilon_1^2 + \epsilon_2^2 + \epsilon_3^2 = 1 - \epsilon_0^2 \leq 1 \quad \rightarrow \quad |\epsilon| \in [1,1]$$

$$\epsilon_0^2 = 1 - |\epsilon|^2$$

Introducing the unit vector, $\hat{n} = \frac{\vec{\varepsilon_\varphi}}{|\vec{\varepsilon_\varphi}|}$ where

$\vec{n} = (n_1, n_2, n_3)$ · and the angle $\frac{\phi}{2} \in [0, \pi]$

defined by $\cos\frac{\phi}{2} = \varepsilon_{\varphi_0}$ and $\sin\frac{\phi}{2} = |\vec{\varepsilon_\varphi}|$

$\varepsilon_{\varphi_0} = \cos\frac{\phi}{2}$ and $\vec{\varepsilon_\varphi} = \hat{n}\sin\frac{\phi}{2}$.

$$\boxed{\begin{aligned} U &= \varepsilon_{\varphi_0} I - i\vec{\varepsilon_\varphi}\cdot\vec{\sigma} \\ &= \left(\cos\frac{\phi}{2}\right) I - i\left(\sin\frac{\phi}{2}\right)\hat{n}\cdot\vec{\sigma} \end{aligned}}$$

given that $\det(U) = 1$

$$= \cos\frac{\phi}{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - i\sin\frac{\phi}{2}\left\{ n_1\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} + n_2\begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix} + n_3\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \right\}$$

$$\boxed{U(\hat{n},\phi) = \begin{bmatrix} \cos\frac{\phi}{2} - i n_3\sin\frac{\phi}{2} & -(n_2 + i n_1)\sin\frac{\phi}{2} \\ (n_2 - i n_1)\sin\frac{\phi}{2} & \cos\frac{\phi}{2} + i n_3\sin\frac{\phi}{2} \end{bmatrix}}$$

Spinor

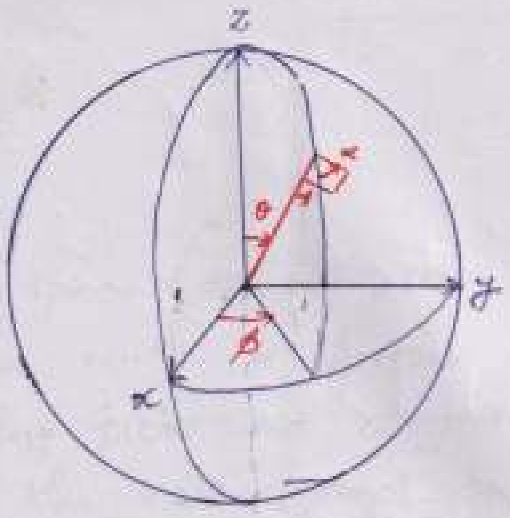
A vector is an arrow in space.

A 4-vector, as an arrow in space-time.

⇒ 2 component complex vector

To specify a spinor state one must furnish 4 real parameters and a sign - r, θ, ϕ, α.

~~A~~ rank-1 spinor (spinor), can be pictured as a vector with 2 further features: a flag that picks out a plane in space containing the vector, and an overall sign. The crucial property is that under the action of a rotation, the direction of the spinor changes just as a vector would, and the flag is carried along in the same way as if it were rigidly attached to the 'flagpole'. ~~A~~ rotation about the axis picked out by the flagpole would have no effect on a vector pointing in that direction, but it does affect the spinor because it rotates the flag.



* The spinor has a direction in space (flagpole), an orientation about this axis (flag) and an overall sign (not shown).

A suitable set of parameters to describe the spinor state, up to a sign, is $(r, \theta, \phi, \alpha)$ as shown. The 1st three fix the length & direction of the flagpole by using standard spherical coordinates, the last gives the orientation of the flag.

A spinor, like a vector, can be rotated. Under the action of a rotation, the spinor magnitude is fixed while the angles θ, ϕ, α change.

We shall refer to the two-component complex vector (below) as a spinor.

$$s = \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} \sqrt{r}\cos(\theta/2)\, e^{i\frac{(-\alpha-\phi)}{2}} \\ \sqrt{r}\sin(\theta/2)\, e^{i\frac{(-\alpha+\phi)}{2}} \end{bmatrix}$$

$$= \sqrt{r}\, e^{-\frac{i\alpha}{2}} \begin{bmatrix} \cos(\theta/2)\, e^{-\frac{i\phi}{2}} \\ \sin(\theta/2)\, e^{\frac{i\phi}{2}} \end{bmatrix}$$

For any given spinor s,

The components (r_x, r_y, r_z) of the flagpole vector are given by:

$$r_x = ab^* + ba^* = s^\dagger \sigma_x s = \langle s|\sigma_x|s\rangle$$

$$r_y = i(ab^* - ba^*) = s^\dagger \sigma_y s = \langle s|\sigma_y|s\rangle$$

$$r_z = |a|^2 - |b|^2 = s^\dagger \sigma_z s = \langle s|\sigma_z|s\rangle$$

$\Longrightarrow$

$$r_x = a^* b^* + b a^* = r \sin\theta/2 \cos\theta/2 \left[e^{i\phi} + e^{i\phi} \right]$$

$$= \frac{r}{2} \sin\theta \times 2 \cos\phi = r \sin\theta \cos\phi$$

$$r_y = i(ab^* - ba^*) = i\, r \sin\theta/2 \cos\theta/2 \left[e^{-i\phi} - e^{i\phi} \right]$$

$$= r \frac{i}{2} \sin\theta \times -2i \sin\phi$$

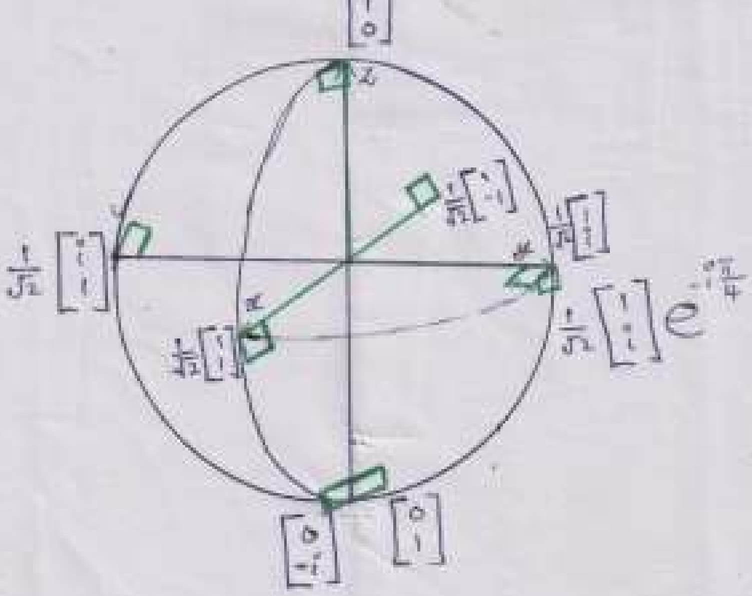
$$= r \sin\theta \sin\phi$$

$$r_z = |a|^2 - |b|^2 = r\left[\cos^2\frac{\theta}{2} - \sin^2\frac{\theta}{2}\right] = r\cos\theta$$

Spinor : a vector in a 2-D complex vector space

A pair of opposite flagpole states such as "straight up along z" and "straight down along z" are orthogonal to one another in the complex vector space.

ie, a rotation thro' an angle θ_0 in the complex 'spin space' corresp. to a rotation thro' an angle $2\theta_0$ in the 3D real space. This is called 'angle doubling'.



1. $\theta_z = 0,\ \phi_z = 0,\ \alpha_z = 0$

$$S_z = \begin{bmatrix} \sqrt{r}\, e^{-i\alpha_z/2} \\ 0 \end{bmatrix} = \sqrt{r} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

2. $\theta_{-z} = 180,\ \phi_{-z} = 0,\ \alpha_{-z} = 0$

$$S_{-z} = \begin{bmatrix} 0 \\ \sqrt{r}\, e^{-i0} \end{bmatrix} = \sqrt{r} \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

3. $\theta'_{-z} = 180,\ \phi'_z = 0,\ \alpha'_{-z} = 180$

$$S'_{-z} = \sqrt{r} \begin{bmatrix} 0 \\ e^{-i(\pi/2)} \end{bmatrix} = \sqrt{r} \begin{bmatrix} 0 \\ -i \end{bmatrix}$$

4. $\theta'_z = 0,\ \phi'_z = 0,\ \alpha'_z = 180$

$$S'_z = \sqrt{r} \begin{bmatrix} e^{-i\pi/2} \\ 0 \end{bmatrix} = \sqrt{r} \begin{bmatrix} -i \\ 0 \end{bmatrix}$$

5. $\theta_y = 90^\circ,\ \phi_y = 90^\circ,\ \alpha_y = 0$

$$S_y = \frac{\sqrt{r}}{\sqrt{2}} \begin{bmatrix} e^{-i\pi/4} \\ e^{i\pi/4} \end{bmatrix} = \frac{\sqrt{r}}{\sqrt{2}} \begin{bmatrix} e^{-i\pi/4} \\ i e^{-i\pi/4} \end{bmatrix} = \sqrt{r} \times \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix} e^{-i\pi/4}$$

6. $\theta_y' = 90, \ \phi_y' = 90, \ \alpha_y' = 180$

$$S_y' = \frac{\sqrt{7}}{\sqrt{2}} \begin{bmatrix} e^{-i3\pi/4} \\ e^{i\pi/4} \end{bmatrix} = \begin{bmatrix} \ \\ \ \end{bmatrix}$$

7. $\theta_y'' = 90, \ \phi_y'' = 90, \ \alpha_y'' = -90$

$$S_y = \frac{\sqrt{7}}{\sqrt{2}} \begin{bmatrix} e^{i0} \\ e^{i\pi/2} \end{bmatrix} = \sqrt{7} \times \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$U(\hat{n},\phi) = \exp\left(-\frac{i\phi}{2}\hat{n}\cdot\vec{\sigma}\right)$$

$$(\hat{n}\cdot\vec{\sigma})^2 = \sum n_i\sigma_i n_j\sigma_j$$

$$= \sum n_i n_j(\sigma_i\sigma_j)$$

$$= \sum n_i n_j(\delta_{ij} I + i\epsilon_{ijk}\sigma_k)$$

$$= \hat{n}\cdot\hat{n}\, I + i(\hat{n}\times\hat{n})\cdot\vec{\sigma} = I$$

$$(\hat{n}\cdot\vec{\sigma})^3 = (\hat{n}\cdot\vec{\sigma})^2(\hat{n}\cdot\sigma) = \hat{n}\cdot\vec{\sigma}$$

$$\Longrightarrow (\hat{n}\cdot\vec{\sigma})^{2n} = I \quad \& \quad (\hat{n}\cdot\vec{\sigma})^{2n+1} = \hat{n}\cdot\vec{\sigma}$$

$$U(\hat{n},\phi) = \exp\left(-\frac{i\phi}{2}\hat{n}\cdot\vec{\sigma}\right)$$

$$= \sum_{k=0}^{\infty}\frac{1}{k!}\left(\frac{-i\phi}{2}\right)^k(\hat{n}\cdot\vec{\sigma})^k$$

$$= \sum_{m=0}^{\infty}\frac{1}{(2m)!}\left(\frac{-i\phi}{2}\right)^{2m}(\hat{n}\cdot\vec{\sigma})^{2m} + \sum_{m=0}^{\infty}\frac{1}{(2m+1)!}\left(\frac{-i\phi}{2}\right)^{2m+1}(\hat{n}\cdot\vec{\sigma})^{2m+1}$$

$=$

$$= I \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m)!} \left(\frac{\phi}{2}\right)^{2m} - i(\hat{n}\cdot\vec{\sigma}) \sum_{m=0}^{\infty} \frac{(-1)^m}{(2m+1)!} \left(\frac{\phi}{2}\right)^{2m+1}$$

$$= I \cos\left(\frac{\phi}{2}\right) - i(\hat{n}\cdot\vec{\sigma}) \sin\left(\frac{\phi}{2}\right)$$

* For any operator A such that $A^2 = I$, then

$$e^{i\phi A} = \cos(\phi) I + i \sin(\phi) A$$

$$e^{-i\phi A} = \cos(\phi) I - i \sin(\phi) A$$

$U(\hat{x}$

$U(\hat{y},$

$U(\hat{z},$

→ spin

$e^{i\frac{\phi}{2}\sigma}$

$$U(\hat{x},\phi) = e^{i\frac{\phi}{2}\sigma_x} = \begin{bmatrix} \cos(\phi/2) & i\sin(\phi/2) \\ i\sin(\phi/2) & \cos(\phi/2) \end{bmatrix}$$

$$U(\hat{y},\phi) = e^{i\frac{\phi}{2}\sigma_y} = \begin{bmatrix} \cos(\phi/2) & \sin(\phi/2) \\ -\sin(\phi/2) & \cos(\phi/2) \end{bmatrix}$$

$$U(\hat{z},\phi) = e^{i\frac{\phi}{2}\sigma_z} = \begin{bmatrix} e^{i\phi/2} & 0 \\ 0 & e^{-i\phi/2} \end{bmatrix}$$

→ spin rotation matrices

$$e^{i\frac{\phi}{2}\sigma_j} = I\cos\frac{\phi}{2} + i\sin(\phi/2)\,\sigma_j$$

$$s' = e^{i\frac{\phi}{2}\sigma_x} \cdot s$$

$$r' = \langle s'|\sigma|s'\rangle = \langle s|e^{-i\frac{\phi}{2}\sigma_x}\,\sigma\, e^{i\frac{\phi}{2}\sigma_x}|s\rangle$$

σ_x is Hermitian

$$x' = \langle s|e^{-i\frac{\phi}{2}\sigma_x}\,\sigma_x\, e^{i\frac{\phi}{2}\sigma_x}|s\rangle$$

$$= \langle s|\sigma_x\, e^{-i\frac{\phi}{2}\sigma_x}\, e^{i\frac{\phi}{2}\sigma_x}|s\rangle$$

since σ_x commutes with I and itself.

$$e^{-i\frac{\phi}{2}\sigma_x}\, e^{i\frac{\phi}{2}\sigma_x} = I$$

$$= \langle s|\sigma_x|s\rangle = x$$

$y' =$

$\alpha = \frac{\phi}{2}$

$(\cos\alpha - i s$

$y' =$

$=$

$=$

Similarly,

$z' =$

$\Longrightarrow$

$$y' = \langle s| e^{-i\frac{\phi}{2}\sigma_y} \sigma_y e^{i\frac{\phi}{2}\sigma_y} |s\rangle$$

$\alpha = \frac{\phi}{2}$, $\quad \sigma_x \sigma_y = -\sigma_y \sigma_x$

$$(\cos\alpha - i\sin\alpha\, \sigma_x)\sigma_y(\cos\alpha + i\sin\alpha\, \sigma_x) =$$

$$= \sigma_y(\cos\alpha + i\sin\alpha\,\sigma_x)(\cos\alpha + i\sin\alpha\,\sigma_x)$$

$$= \sigma_y(\cos^2\alpha - \sin^2\alpha + 2i\sin\alpha\cos\alpha\,\sigma_x)$$

$$= \sigma_y(\cos\phi + i\sin\phi\,\sigma_x)$$

$$y' = \langle s|\sigma_y|s\rangle\cos\phi + i\sin\phi\,\langle s|\sigma_y\sigma_x|s\rangle$$

$$= y\cos\phi + \langle s|\sigma_z|s\rangle\sin\phi \qquad [\sigma_y\sigma_x = -i\sigma_z$$

$$= y\cos\phi + z\sin\phi$$

Similarly,

$$z' = z\cos\phi - y\sin\phi$$

$$\Longrightarrow \quad r' = R_x r$$

$\therefore$

Multiplying a spinor by each of the spin rotation matrices results in a rotation of the flagpole by the corresp. matrix for a rotation in 3D:

$$R_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & \sin\phi \\ 0 & -\sin\phi & \cos\phi \end{bmatrix}, \quad R_y = \begin{bmatrix} \cos\phi & 0 & -\sin\phi \\ 0 & 1 & 0 \\ \sin\phi & 0 & \cos\phi \end{bmatrix}$$

$$R_z = \begin{bmatrix} \cos\phi & \sin\phi & 0 \\ -\sin\phi & \cos\phi & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The rotation angle ϕ is twice the angle $\frac{\phi}{2}$ which appears in the 2×2 spin rotation matrices.

Any
as :

and
written

where
in 3D

$J_x = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & i \end{bmatrix}$

$\Rightarrow$ the g
homomorph

14

Any member of SO(2) can be written as :

$$U = e^{i\sigma \cdot \theta}$$

and any member of SO(3) can be written as :

$$R = e^{iJ \cdot \theta}$$

where J are the generators of rotations in 3D :

$$J_x = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix} ; J_y = \begin{bmatrix} 0 & 0 & i \\ 0 & 0 & 0 \\ -i & 0 & 0 \end{bmatrix} , J_z = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\Rightarrow$ the groups SO(3) and SO(2) are homomorphic.

LA 14

Qubit

A qubit is the fundamental quantum state representing the smallest unit of quantum information containing one bit of classical information accessible by measurement

$$|\psi\rangle = c_0|0\rangle + c_1|1\rangle = c_0\begin{bmatrix}1\\0\end{bmatrix} + c_1\begin{bmatrix}0\\1\end{bmatrix}$$

$$c_0 = r_0 e^{i\phi_0}, \quad c_1 = r_1 e^{i\phi_1}$$

We have 4 unknowns (2 phases & 2 amplitudes) that uniquely determine components it'd seem that a qubit should have 4 free real-valued parameters.

$$|\psi\rangle = r_0 e^{i\phi_0}|0\rangle + r_1 e^{i\phi_1}|1\rangle$$

In the case of quantum bits, we know that a quantum state does not change if we multiply it with any # of unit norm,

ie :- $|\psi\rangle = e^{i\phi}|\psi\rangle$

This is because the quantum property of measurement follows from identifying the moduli squared of the amplitude as an occupation probability $|c_1|^2$ and $1-|c_0|^2$ for the qubit to occupy its logical states $|1\rangle$ and $|0\rangle$, respectively.

$$e^{-i\phi_0}|\psi\rangle = e^{-i\phi_0}\left(r_0 e^{i\phi_0}|0\rangle + r_1 e^{i\phi_1}|1\rangle\right)$$

$$= r_0|0\rangle + r_1 e^{i(\phi_1-\phi_0)}|1\rangle$$

From 4 parameters, we end up with 3 parameters r_0, r_1 and $\phi =: \phi_1 - \phi_0$.

$$|c_0|^2 + |c_1|^2 = 1$$

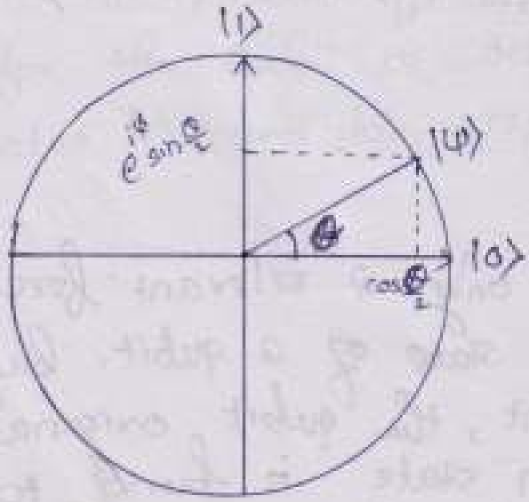
$$\rightarrow \quad r_0^2 + r_1^2 = 1$$

$$|\psi\rangle = \sqrt{1-r_1^2}\,|0\rangle + r_1 e^{i\phi}|1\rangle$$

There are only 2 relevant free parameters to specify the state of a qubit. but upon measurement, the qubit originally in the superposition state is found to occupy only one of its logical states.

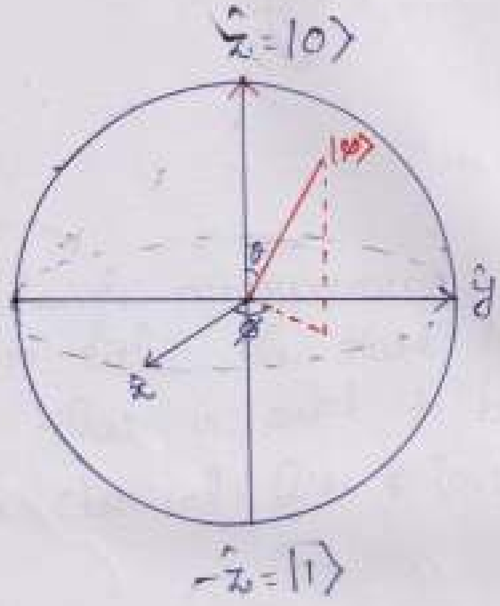
$$|\psi\rangle \xrightarrow{\text{measure}} \begin{cases} |1\rangle & \text{with probability } r_1^2 \\ |0\rangle & \text{"} \quad 1-r_1^2 \end{cases}$$

Upon a single measurement, $|\psi\rangle$ is found to be in either the state $|0\rangle$ or $|1\rangle$, an outcome that is said to be specified by a single classical bit $\in \{0,1\}$.



* A qubit in Hilbert space in its SU(2) representation

Bloch sphere representation



A
along
the
pure

For
(in a
qubit

$|\psi\rangle$

The
so $|\psi$
physica

* The
is tha
phases
i.e., a
and in
Multip
spinor

A qubit can be represented by any point along the ~~surface~~ surface of the Bloch sphere, with the north and south poles corresp. to the pure states $|0\rangle$ and $|1\rangle$ resp.

For a given point (θ, ϕ) on the sphere (in usual spherical coordinates) the corresp. qubit is defined by:

$$|\psi\rangle = e^{-i\phi/2} \cos(\theta/2) |0\rangle + e^{i\phi/2} \sin(\theta/2) |1\rangle$$

The global phase is not encoded in a qubit, so $|\psi\rangle$ and $e^{i\alpha}|\psi\rangle$ corresp. to the same physical state.

* The distinction b/w a spinor and a qubit is that, spinors distinguish b/w such global phases.
ie., a spinor defined as $|\chi\rangle = \begin{bmatrix} a \\ b \end{bmatrix}$ with $a, b \in \mathbb{C}$ and imposed normalization $\langle\chi|\chi\rangle = 1$.
Multiplying $|\chi\rangle$ by $e^{i\alpha}$ results in a different spinor, one that corresp. to the same qubit.

Matrix exponential, e^{A}

$$\frac{dx}{dt} = ax \xrightarrow{\text{solution}} x(t) = x(0)\cdot e^{at}$$

If we have an $n\times n$ system of ODEs,

$$\boxed{\frac{d\vec{x}}{dt} = A\vec{x}}$$

we know that,

$A = VDV^{-1}$ is diagonalizable with

eigensolutions $A\vec{v}_k = \lambda_k \vec{v}_k \quad (k = 1, 2, \dots, n)$

$$A = VD_A V^{-1} \Longrightarrow D_A = V^{-1}AV$$

$V = [\vec{v}_1 \cdots \vec{v}_n]$; V_i : eigenvectors of A

Change of variables,

$\vec{x} = V\vec{y}$, where V has to be determined

$V = ?$

$$\frac{d\vec{x}}{dt} = A\vec{x} \Longrightarrow V\frac{d\vec{y}}{dt} = AV\vec{y}$$

$$\frac{d\vec{y}}{dt} = V^{-1}AV\vec{y} = D_A\vec{y}$$

$$\begin{bmatrix} \frac{d}{dt}y_1 \\ \vdots \\ \frac{d}{dt}y_n \end{bmatrix} = \begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix} \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \begin{matrix} \longrightarrow y_1(t) = C_1 e^{\lambda_1 t} \\ \vdots \\ \longrightarrow y_n(t) = C_n e^{\lambda_n t} \end{matrix}$$

We obtain the decoupled system, $\frac{dy}{dt} = D_A y$

which has the solution

$$\vec{y}_i(t) = C_i e^{\lambda_i t}$$

$$\vec{x}(t) = V\,\vec{y}(t) = \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix}$$

$$\boxed{\vec{x}(t) = c_1 v_1 e^{\lambda_1 t} + \cdots + c_n v_n e^{\lambda_n t}}$$

$$\vec{x}(0) = c_1 v_1 + \cdots + c_n v_n$$

$$\vec{x}(0) = V\vec{c} \Longrightarrow \vec{c} = V^{-1}\vec{x}(0)$$

$$V = \begin{bmatrix} v_1 \cdots v_n \end{bmatrix}, \quad \vec{c} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix} = (c_1, \cdots, c_n)$$

We would like to write the solution

$\vec{x}(t) = (\text{some matrix}) \times \vec{x}(0)$

$$x(t) = V \begin{bmatrix} c_1 e^{\lambda_1 t} \\ \vdots \\ c_n e^{\lambda_n t} \end{bmatrix} = V \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}$$

$$= V \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} \vec{c} \quad ; \quad \vec{c} = V^{-1} \vec{x}(0)$$

$$x(t) = V \begin{bmatrix} e^{\lambda_1 t} & & \\ & \ddots & \\ & & e^{\lambda_n t} \end{bmatrix} V^{-1} \vec{x}(0) = e^{At} \vec{x}(0)$$

where, we have defined the "matrix exponential" of a diagonalizable matrix as:

$$e^{At} = V e^{D_\lambda t} V^{-1}$$

$$\boxed{\frac{d\vec{x}}{dt} = A\vec{x} \implies \vec{x}(t) = e^{At} \cdot \vec{x}(0)}$$

Ex:-

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$\lambda_1 = 1, \ \lambda_2 = -1$

$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \ \vec{v}_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$e^{At} = V e^{D_A t} V^{-1} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}^{-1}$$

$$= \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} e^t & 0 \\ 0 & e^{-t} \end{bmatrix} \frac{1}{2} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} e^t & e^{-t} \\ e^t & -e^{-t} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} e^t + e^{-t} & e^t - e^{-t} \\ e^t - e^{-t} & e^t + e^{-t} \end{bmatrix}$$

$$= \frac{1}{2} \begin{bmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{bmatrix}$$

$$\lim_{t \to \infty} \sinh t = \frac{e^t}{2} ; \quad \lim_{t \to \infty} \cosh t = \frac{e^t}{2}$$

$$\approx \frac{e^t}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$x(t) = e^{At} \vec{x}(0) \approx \frac{e^t}{2} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \vec{x}(0)$$

From the eigenvector expansion,

$$\vec{x}(t) = c_1 e^{t}\begin{bmatrix}1\\1\end{bmatrix} + c_2 e^{-t}\begin{bmatrix}1\\-1\end{bmatrix} \approx c_1 e^{t}\begin{bmatrix}1\\1\end{bmatrix}$$

$t \to \infty$

eigenvectors are orthogonal,

$$\vec{x}(0) \approx c_1 \vec{v}_1 \longrightarrow \vec{v}_1^T \vec{x}(0) \approx c_1 \vec{v}_1^T v_1$$

$$c_1 = \frac{\vec{v}_1^T \vec{x}(0)}{\vec{v}_1^T \vec{v}_1} = \frac{\vec{v}_1^T \vec{x}(0)}{2}$$

$$\vec{x}(t) \approx c_1 e^{t} \vec{v}_1 = \frac{e^t}{2} \vec{v}_1 \vec{v}_1^T \vec{x}(0) = \frac{e^t}{2} v_1 \vec{v}_1^T$$

$$= \frac{e^t}{2}\begin{bmatrix}1 & 1\\1 & 1\end{bmatrix}\vec{x}(0)$$

When $t = 1$,

$$\boxed{\begin{aligned} e^{A} &= V e^{D_A} V^{-1} \\ e^{V D_A V^{-1}} &= \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix} \begin{bmatrix} e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n} \end{bmatrix} \begin{bmatrix} \vec{v}_1 & \cdots & \vec{v}_n \end{bmatrix}^{-1} \end{aligned}}$$

- e^{A} is the matrix with the same eigenvectors as A, but with eigenvalues λ replaced by e^{λ}.

$$A\vec{v}_k = \lambda\vec{v}_k \implies e^{A}\vec{v}_k = e^{\lambda_k}\vec{v}_k$$

If A is an $n \times n$ matrix, then matrix exponential e^A is defined as:

$$\boxed{e^A = \sum_{k=0}^{\infty} \frac{A^k}{k!} = I + A + \frac{A^2}{2!} + \cdots\cdots}$$

$$\boxed{= \lim_{n \to \infty} \left(I + \frac{A}{n}\right)^n}$$

For a

$e^D = I + D$

$=$

$=$

For a diagonal matrix D,

$$e^D = 1 + D + \frac{D^2}{2} + \cdots$$

$$= \begin{bmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{bmatrix} + \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} + \begin{bmatrix} \frac{d_1^2}{2!} & & \\ & \ddots & \\ & & \frac{d_n^2}{2!} \end{bmatrix} + \cdots$$

$$= \begin{bmatrix} e^{d_1} & & \\ & \ddots & \\ & & e^{d_n} \end{bmatrix}$$

$$A = VD_AV^{-1} \Longrightarrow e^{A} = V e^{D_A} V^{-1}$$

Proof

$$e^{A} = \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{(VD_AV^{-1})^k}{k!}$$

$$= \sum_{k=0}^{\infty} \frac{VD_AV^{-1} \times VD_AV^{-1} \times \cdots\cdots (VD_AV^{-1})}{k!} \quad [k \text{ times}]$$

$$= \sum_{k=0}^{\infty} \frac{V(D_A)^k V^{-1}}{k!} = V\left(\sum_{k=0}^{\infty} \frac{(D_A)^k}{k!}\right)V^{-1}$$

$$= V e^{D_A} V^{-1}$$

* $\boxed{\frac{d}{dt} e^{At} =}$

* $e^{A+B} \neq$

$e^{A+B} = e$

* $(e^{A})^{-1} =$

* $(e^{A})^{T} = e$

A is symm

A is skew

* $(e^{A})^{\dagger} = e^{(A}$

A is hermitia

A is skew-hermi

* $\boxed{\frac{d}{dt} e^{At} = Ae^{At}}$ using Taylor's expansion to prove

* $e^{A+B} \neq e^{A} e^{B}$.

$e^{A+B} = e^{A} \cdot e^{B}$ iff $AB = BA$

* $(e^{A})^{-1} = e^{-A}$

* $(e^{A})^{T} = e^{(A^{T})}$

A is symmetric $\Longrightarrow$ e^{A} is symmetric

A is skew symmetric $\longrightarrow$ e^{A} is orthogonal

* $(e^{A})^{\dagger} = e^{(A^{\dagger})}$

A is hermitian $\Longrightarrow$ e^{A} is hermitian

A is skew-hermitian $\Longrightarrow$ e^{A} is unitary

* $\boxed{\det(e^{A}) = e^{tr(A)}}$

$$\det(e^{A}) = \prod_{i=1}^{n} e^{\lambda_i}$$

$$e^{tr(A)} = e^{\sum \lambda_i} = \prod_{i=1}^{n} e^{\lambda_i} = \det(e^{A})$$

*

$\cos A = I$

$\sin A = A$

$e^{iA} = \cos$

$\cos A = Re$

$\sin A = Im(e$

$\cos^2 A + \sin^2 A$

*

$$\cos A = I - \frac{A^2}{2!} + \frac{A^4}{4!} - \cdots$$

$$\sin A = A - \frac{A^3}{3!} + \frac{A^5}{5!} - \cdots$$

$$e^{iA} = \cos A + i \sin A$$

$$\cos A = \mathbb{R}(e^{iA}) = \frac{e^{iA} + e^{-iA}}{2}$$

$$\sin A = \operatorname{Im}(e^{iA}) = \frac{e^{iA} - e^{-iA}}{2i}$$

$$\cos^2 A + \sin^2 A = I$$

Hermitian Matrix

A Hermitian matrix (or self-adjoint matrix) is a complex square matrix that is equal to its own conjugate transpose.

A hermitian $A = [a_{ij}]$ $\Longrightarrow$	$a_{ij} = \overline{a_{ji}}$ $A = \overline{A}^T = A^H = A^\dagger$
$\Rightarrow \langle w \| Av \rangle = \langle Aw \| v \rangle$	
$\Rightarrow \langle v \| Av \rangle \in \mathbb{R}, \; v \in V$	

* A square matrix 'A' is Hermitian iff it is unitarily diagonalizable with real eigenvalues.

* The entries on the main diagonal of any Hermitian matrix are real

Proof

$$a_{ij} = \overline{a}_{ji}$$

If $i = j$, $a_{ii} = \overline{a}_{ii} \longrightarrow a_{ii} \in \{\text{purely real } \cancel{\text{imaginary}}\}$.

$$A = \begin{bmatrix} 2 & 2+i & 4 \\ 2-i & 3 & i \\ 4 & -i & 1 \end{bmatrix}$$

* A square matrix 'A' is Hermitian iff it is unitarily diagonalizable with real eigenvalues.

$$\boxed{A^\dagger = A \longrightarrow \lambda \in \mathbb{R}}$$

- A normal matrix is Hermitian iff all its eigenvalues are real.

Proof

$$\langle Av | Av \rangle = (Av)^\dagger (Av) = v^\dagger A^\dagger A v = v^\dagger A A v$$

$$= v^\dagger A^2 v = v^\dagger \lambda^2 v = \lambda^2 v^\dagger v = \lambda^2 \langle v | v \rangle$$

$$= \lambda^2 |v|^2$$

$$\longrightarrow \lambda^2 = \frac{\langle Av | Av \rangle}{|v|^2} \in \mathbb{R}^+ \cup \{0\}.$$

$$\longrightarrow \lambda \in \mathbb{R}$$

(OR)

$$A\vec{v} = \lambda \vec{v} \quad \& \quad (A\vec{v})^\dagger = \bar{\lambda} \vec{v}^\dagger$$

$$v^\dagger A^\dagger = v^\dagger A = \bar{\lambda} v^\dagger$$

$$v^\dagger A v = \lambda v^\dagger v = \lambda \langle v | v \rangle = \lambda |v|^2$$

$$v^\dagger A v = \bar{\lambda} v^\dagger v = \bar{\lambda} \langle v | v \rangle = \bar{\lambda} |v|^2$$

$$(\lambda - \bar{\lambda}) |v|^2 = 0 \xrightarrow[|v| \neq 0]{} \lambda \in \mathbb{R}$$

* Every Hermitian matrix is a normal matrix.

$$AA^\dagger = A^\dagger A$$

$\Longrightarrow$ unitary diagonalizable

$$A^\dagger = A \Longrightarrow A = UD_AU^\dagger,$$

U: unitary

* The hermitian matrix with dimension n has n linearly independent eigenvectors.

* A Hermitian matrix has orthogonal eigenvectors for distinct eigenvalues.

Even if there are degenerate eigenvalues, it is always possible to find an orthogonal basis of $\mathbb{C}^n$ consisting of n eigenvector of A.

ie.,
The degenerate eigenvectors are not necessarily orthogonal to each other. One can always convert them (by Gram-Schmidt) into a set of degenerate eigenvectors that are orthogonal to each other and to all other eigenvectors.

Proof.

$A\vec{v}_i = \lambda$

$(Av_i)^\dagger = \lambda$

$v_i^\dagger A = \bar{\lambda}$

$v_i^\dagger A v_j =$

$v_i^\dagger A v_j =$

$(\bar{\lambda}_i - \lambda_j$

$\lambda_i \neq \lambda_j$

Say,

λ_i is

$A\vec{v}_i = \lambda$

that is diff

Proof.

$$A\vec{v}_i = \lambda_i v_i \quad \& \quad A v_j = \lambda_j v_j$$

$$(Av_i)^\dagger = v_i^\dagger A^\dagger = \bar{\lambda}_i v_i^\dagger \qquad \left[\ \lambda \in \mathbb{R}\right.$$

$$v_i^\dagger A = \bar{\lambda}_i v_i^\dagger = \lambda_i v_i^\dagger$$

$$v_i^\dagger A v_j = \cancel{\bar{\lambda}}_i v_i^\dagger v_j = \cancel{\bar{\lambda}}_i \langle v_i | v_j \rangle$$

$$v_i^\dagger A v_j = \lambda_j v_i^\dagger v_j = \lambda_j \langle v_i | v_j \rangle$$

$$(\cancel{\bar{\lambda}}_i - \lambda_j)\langle v_i | v_j \rangle = 0$$

$$\lambda_i \neq \lambda_j \implies \langle v_i | v_j \rangle = v_i^\dagger v_j = 0$$

$\therefore$ eigenvectors v_i, v_j are mutually orthogonal.

Case.

Say, λ_1 is k-fold degenerate

$$A\vec{v}_i = \lambda_1 \vec{v}_i \quad \text{for} \quad i = 1, 2, \dots, k$$

But that is different from any of $\lambda_{k+1}, \lambda_{k+2}$ etc.

Any linear combination of these V_i is also an eigenvector with eigenvalue λ_1.

since for, $$\vec{z} = \sum_{i=1}^{k} c_i \vec{V}_i$$

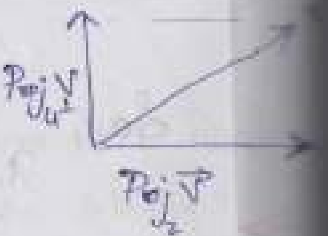
$$A\vec{z} = A\sum_{i=1}^{k} c_i \vec{V}_i = \sum_{i=1}^{k} c_i A \vec{v}_i$$

$$= \sum_{i=1}^{k} c_i \lambda_1 \vec{V}_i = \lambda_1 \vec{z}$$

If $\vec{V}_i$ defined are not already mutually orthogonal, then we can construct new eigenvectors $\vec{z}_i$ that are orthogonal by the Gram-Schmidt orthogonalization.

$$\mathrm{Proj}_{z}(V) = (\vec{V}\cdot\hat{z})\,\hat{z} = (\vec{V}^{\dagger}\hat{z})\,\hat{z} = \frac{\vec{V}^{\dagger}\vec{z}}{|z|^2}\,\hat{z}$$

=



$$\mathrm{Proj}_{z^{\perp}} V = \vec{V} - \mathrm{Proj}_{z} V$$

$\vec{z}_1 = \vec{v}_1$ $\qquad\qquad\qquad \hat{z}_1 = \frac{\vec{z}_1}{|\vec{z}_1|}$

$\vec{z}_2 =$ comp. of $\vec{v}_2 \perp$ to $\vec{z}_1$

$= \vec{v}_2 - \mathrm{Proj}_{z_1} \vec{v}_2$ $\qquad\qquad \hat{z}_2 = \frac{\vec{z}_2}{|\vec{z}_2|}$

$= \vec{v}_2 - (\vec{v}^{\dagger} \hat{z}_1) \hat{z}_1$

$\vec{z}_3 =$ comp. of $\vec{v}_3 \perp$ to span$(\vec{z}_1, \vec{z}_2)$

$= \vec{v}_3 - \left[\text{Projection of } \vec{v}_3 \text{ on span}(\hat{z}_1, \hat{z}_2)\right]$

$= v_3 - \left[\mathrm{Proj}_{z_1} v_3 + \mathrm{Proj}_{z_2} v_3\right]$ $\qquad ; \hat{z}_3 = \frac{\vec{z}_3}{|\vec{z}_3|}$

$= v_3 - (\vec{v}_3^{\dagger} \hat{z}_1) \hat{z}_1 - (\vec{v}_3^{\dagger} \hat{z}_2) \hat{z}_2$

$\vdots$

$\vec{z}_k = \vec{v}_k - \sum_{j=1}^{k-1} \mathrm{Proj}_{z_j} \vec{v}_k$

$= \vec{v}_k - \sum_{j=1}^{k-1} (\vec{v}_k^{\dagger} \hat{z}_j) \hat{z}_j$ $\qquad ; \hat{z}_k = \frac{\vec{z}_k}{|\vec{z}_k|}$

∴ Even if 'A' has some degenerate eigenvalues, we can by construction obtain a set of 'n' mutually orthogonal eigenvectors. & these eigenvectors are complete in that they form a basis for the n-dimensional vector space.

* The inverse of an invertible Hermitian matrix is Hermitian.

$$(A^{-1})^\dagger = A^{-1}$$

Proof

$$AA^{-1} = I \implies I = I^\dagger = (AA^{-1})^\dagger = (A^{-1})^\dagger A^\dagger = (A^{-1})^\dagger A$$

$$I = (A^{-1})^\dagger A \implies (A^{-1})^\dagger = A^{-1}$$

* The product of 2 ~~matrices~~ Hermitian matrices A & B is Hermitian iff $AB = BA$

~~$(AB)^\dagger$~~

Proof

$$(AB)^\dagger = \overline{(AB)}^T = \overline{B^T A^T} = \overline{B}^T\, \overline{A}^T = B^\dagger A^\dagger$$

$$= BA \quad \left[B^\dagger = B \ \& \ A^\dagger = A \right.$$

* A^n is Hermitian if A is Hermitian & $n \in \mathbb{Z}$

* If n orthonormal eigenvectors $u_1, u_2, \ldots, u_n$ of a Hermitian matrix are chosen & written as the columns of the matrix U, then the eigendecomposition of 'A' is

$$A = U D_A U^{\dagger} \quad , \text{ where } \quad UU^{\dagger} = I = U^{\dagger} U$$

$A^{\dagger} = A$
A is Hermitian

$$\therefore \quad A = \sum_j \lambda_j u_j u_j^{\dagger} = \sum_j \lambda_j |u_j\rangle\langle u_j|$$

Proof

$$A = U D_A U^{\dagger} = \begin{bmatrix} \vec{u}_1 \vec{u}_2 \cdots \vec{u}_n \end{bmatrix}_{1\times n} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}_{n\times n} \begin{bmatrix} \vec{u}_1^{\dagger} \\ \vec{u}_2^{\dagger} \\ \vdots \\ \vec{u}_n^{\dagger} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 \vec{u}_1 \; \lambda_2 \vec{u}_2 \cdots \lambda_n \vec{u}_n \end{bmatrix}_{1\times n} \begin{bmatrix} \vec{u}_1^{\dagger} \\ \vec{u}_2^{\dagger} \\ \vdots \\ \vec{u}_n^{\dagger} \end{bmatrix}_{n\times 1}$$

$$= \lambda_1 \vec{u}_1 \vec{u}_1^{\dagger} + \lambda_2 \vec{u}_2 \vec{u}_2^{\dagger} + \cdots\cdots + \lambda_n \vec{u}_n \vec{u}_n^{\dagger}$$

$$A = \sum_{j=1}^{n} \lambda_j u_j u_j^{\dagger}$$

* The determinant of a Hermitian matrix is real.

$$\boxed{A^{\dagger} = A \Longrightarrow \det(A) \in \mathbb{R}}$$

$$\overline{(z+w)} = \bar{z} + \bar{w}$$

Proof.

$\det(A^{T}) = \det(A)$

$\det(A^{\dagger}) = \overline{\det(A)}$

$A^{\dagger} = A \longrightarrow \det(A) = \overline{\det(A^{\dagger})}$

$\therefore \ \underline{\underline{\det(A) \in \mathbb{R}}}$

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma(i)} \right)$$

S_n: symmetric group on n-elements
– is the set of all permutations of $(1,2,\ldots,n)$

$$\det(A) = \epsilon_{i_1 i_2 \cdots i_n} \times a_{1i_1} a_{2i_2} \cdots a_{ni_n}$$

$\epsilon_{i_1 i_2 \cdots i_n}$: Levi-Civita symbol

* The sum of a square matrix and its conjugate transpose $(A + A^{\dagger})$ is Hermitian

The difference of a square matrix & its conjugate transpose $(A - A^{\dagger})$ is skew-Hermitian (anti-hermitian)

* Any square matrix can be written as the sum of a Hermitian matrix and a ~~[illegible]~~ skew-Hermitian matrix.

$$\boxed{\begin{array}{l}\text{Toeplitz decomposition of } A \\ A = \left(\dfrac{A+A^{+}}{2}\right) + \left(\dfrac{A-A^{+}}{2}\right)\end{array}}$$

Ex: of Hermitian matrices

1. Pauli matrices

$$\sigma_1 = \sigma_x = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \sigma_2 = \sigma_y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}; \quad \sigma_3 = \sigma_z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$\sigma_1^2 = \sigma_2^2 = \sigma_3^2 = i\sigma_1\sigma_2\sigma_3 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = I$$

$\det \sigma_i = -1$ & $tr(\sigma_i) = 0$

$\lambda_i = \pm 1$

Eigenvectors: $\psi_{x^+} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \psi_{x^-} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$$\psi_{y^+} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ i \end{bmatrix}, \quad \psi_{y^-} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -i \end{bmatrix}$$

$$\psi_{z^+} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \psi_{z^-} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$\psi_{x^+} = i\sigma_y \psi_{x^-}; \quad \psi_{y^+} = \sigma_x \psi_{y^-}; \quad \psi_{z^+} = \sigma_x \psi_{z^-}$$

$$\sigma_a = \begin{bmatrix} \delta_{a3} & \delta_{a1} - i\delta_{a2} \\ \delta_{a1} + i\delta_{a2} & -\delta_{a3} \end{bmatrix};$$

where, $\delta_{ab} = \begin{cases} 1, & a = b \\ 0, & a \neq b \end{cases}$

2. Gell-Mann matrices

$$\lambda_1 = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \lambda_2 = \begin{bmatrix} 0 & -i & 0 \\ i & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \lambda_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\lambda_4 = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{bmatrix}, \lambda_5 = \begin{bmatrix} 0 & 0 & -i \\ 0 & 0 & 0 \\ i & 0 & 0 \end{bmatrix}, \lambda_6 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\lambda_7 = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & -i \\ 0 & i & 0 \end{bmatrix}, \lambda_8 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -2 \end{bmatrix}$$

* generalize the Pauli matrices for SU(2) to SU(3).

Skew-Hermitian matrix

A square matrix with complex entries is said to be skew-Hermitian or antihermitian if its conjugate transpose is the -ve of the original matrix.

$A = [a_{ij}]$ skew-hermitian (antihermitian) $\Longrightarrow$ $A^{\dagger} = -A$

$a_{ij} = -\overline{a}_{ji}$

* All entries on the main diagonal of a skew-Hermitian matrix have to be purely imaginary

$$\overline{a_{ii}} = -a_{ii}$$

Ex: $\begin{bmatrix} -i & 2+i \\ -2+i & 0 \end{bmatrix}$

Proof.

$a_{ij} = -\overline{a_{ji}}$

$a_{ii} = -\overline{a_{ii}} \implies \overline{a_{ii}} = -a_{ii}$

* The eigenvalues of a skew-Hermitian matrix are all purely imaginary (and possibly zero).

$$\boxed{A^{\dagger} = -A \longrightarrow \bar{\lambda} = -\lambda}$$

Proof.

$$\langle Av|Av\rangle = (Av)^{\dagger} Av = v^{\dagger} A^{\dagger} A v = -v^{\dagger} A^{2} v$$

$$= -\lambda^{2} v^{\dagger} v = -\lambda^{2} |v|^{2}$$

$$-\lambda^{2} = \frac{\langle Av|Av\rangle}{|v|^{2}}$$

$$\lambda^{2} = \frac{-\langle Av|Av\rangle}{|v|^{2}} \in \mathbb{R}^{-}$$

$$\cancel{\lambda = \frac{\sqrt{\langle Av|Av\rangle}}{|v|}} \; \& \; \lambda = \frac{\sqrt{-\langle Av|Av\rangle}}{|v|} = i\frac{\sqrt{\langle Av|Av\rangle}}{|v|}$$

$$\bar{\lambda} = \frac{-i\sqrt{\langle Av|Av\rangle}}{|v|} = -\lambda$$

$\Longrightarrow$ λ is purely imaginary

(OR) $Av = \lambda v$ & $(Av)^{\dagger} = (\lambda v)^{\dagger}$

$$v^{\dagger}A^{\dagger} = \bar{\lambda}\, v^{\dagger}$$

$$\boxed{-v^{\dagger}A = \bar{\lambda}\, v^{\dagger}}$$

$$v^{\dagger}Av = \lambda v^{\dagger}v = \lambda |v|^2$$

$$v^{\dagger}Av = \bar{\lambda} v^{\dagger}v = -\lambda |v|^2$$

$$(\lambda + \bar{\lambda})|v|^2 = 0$$

$|v| \neq 0 \Rightarrow \underline{\bar{\lambda} = -\lambda}$

$\therefore \lambda$ is purely imaginary

* Every skew-symmetric matrix is a normal matrix.

$$AA^{\dagger} = -A^2 = A^{\dagger}A$$

$\longrightarrow$ unitarily diagonalizable

$\therefore \quad A = U D_A U^{\dagger}$,

U: unitary matrix

* 'A' is skew-Hermitian iff iA (or equivalently, $-iA$) is Hermitian.

* 'A' is skew-Hermitian iff $R(A)$ is skew-symmetric & $I(A)$ is symmetric.

$$\boxed{\begin{array}{ccc} A \text{ is skew-Hermitian} & & R(A) \text{ is skew-symmetric} \\ A^{\dagger} = -A & \Longleftrightarrow & \\ & & I(A) \text{ is symmetric} \end{array}}$$

<u>Proof</u>

A is skew-symmetric : $A^{\dagger} = -A$

$$a_{ji} = -\bar{a}_{ij} \Rightarrow$$

$$Re\{a_{ji}\} + i\, Im\{a_{ji}\} = -Re\{a_{ij}\} + i\, I_m\{a_{ij}\}$$

$$Re\{a_{ji}\} = -Re\{a_{ij}\} \Rightarrow Re(A)^{T} = -R(A)$$

$\therefore$ $Re(A)$ is skew symmetric

$$Im\{a_{ji}\} = I_m\{a_{ij}\} \Rightarrow I_m(A)^{T} = I_m(A)$$

$\therefore$ $I_m(A)$ is symmetric

* If A is skew-Hermitian, then A^k is Hermitian if k is an even integer & skew-Hermitian if k is an odd integer.

$$\boxed{A^\dagger = -A \Longrightarrow A^k \text{ is } \begin{cases} \text{Hermitian} & \text{if } k = 2n \\ \text{skew-Hermitian} & \text{if } k = 2n+1 \end{cases}}$$

* If 'A' is skew-Hermitian, then the matrix exponential e^A is unitary.

$A^\dagger = -A$ & $A = U D_A U^{-1} = U D_A U^\dagger$

$$e^A = U e^{D_A} U^{-1} = U e^{D_A} U^\dagger$$

$$(e^A)^\dagger = (U e^{D_A} U^\dagger)^\dagger = U (e^{D_A})^\dagger U^\dagger$$

$$= U \, \mathrm{diag}\left(\overline{e^{\lambda_1}, e^{\lambda_2}, \ldots, e^{\lambda_n}}\right) U^\dagger$$

$$= U \, \mathrm{diag}\left(e^{-\lambda_1}, e^{-\lambda_2}, \ldots, e^{-\lambda_n}\right) U^\dagger$$

$$e^A (e^A)^\dagger = U e^{D_A} U^\dagger U (e^{D_A})^\dagger U^\dagger$$

$$= U e^{D_A} (e^{D_A})^\dagger U^\dagger$$

$$= U \, \mathrm{diag}(e^{\lambda_1}, \ldots, e^{\lambda_n}) \, \mathrm{diag}(e^{-\lambda_1}, \ldots, e^{-\lambda_n}) U^\dagger$$

$$= U U^\dagger = 1$$

$\Longrightarrow e^A$ is unitary