Linear transformation

Let V and W be vector spaces over the field F.

Linear transformation $T: V \to W$: $T(\alpha|v_1\rangle + |v_2\rangle) = \alpha T|v_1\rangle + T|v_2\rangle$.

Examples : V : n-tuple ; T : m×n matrix then W : m-tuple

V : Polynomials of degree n ; T : derivative, W : degree(n−1)

Note : to specify T, enough to specify $T|e_i\rangle\ \forall i$, where $\{|e_i\rangle\}$ is an ordered basis of V.

Let $|\varepsilon_i\rangle$ be a basis in W. Then $T|e_i\rangle$ can be expanded :

$$T|e_j\rangle = \sum_{i=1}^{m} A_{ij} |\varepsilon_i\rangle, \quad j = 1 \cdots n \qquad [\dim V = n, \dim W = m]$$

Then for arbitrary vector $|v\rangle = \sum_{i=1}^{n} v_i |e_i\rangle \in V$

$$T|v\rangle = \sum_{i=1}^{n} v_i \sum_{j=1}^{m} A_{ji} |\varepsilon_j\rangle = \sum_{j=1}^{m} \omega_j |\varepsilon_j\rangle$$

⇒ transformation T can be represented by the matrix A :

$$\omega_j = A_{ji} v_i$$

Example: $\frac{d}{dx}$ on space of cubic polynomials.

i th Column of A : $T|e_i\rangle$ in the basis $\{|\varepsilon\rangle\}$.

Linearity : $T|0\rangle$ is the null vector in W

$$T|0\rangle_V = T(|v\rangle + (-1)|v\rangle) = T|v\rangle + (-1) \cdot T|v\rangle = |0\rangle_W$$

$T|v\rangle$ is a subspace of W, called range of T.

rank of T : dimension of range of T.

Set of all vectors in $|v\rangle$ in V such that $T|v\rangle = 0$: subspace of V, called null space of T. Its dimension: nullity of T.

(for finite dimensional V) $\text{rank}(T) + \text{nullity}(T) = \dim(V)$.

<u>Proof</u> Say nullity $= k$ and choose a basis $\{|e_i\rangle, i=1\cdots n\}$ in V such that $\{|e_i\rangle, i=1\cdots k\}$ span the null space.

Clearly, $T|e_i\rangle, i=k+1\cdots n$ span the range of T:

for any $|v\rangle$, $T|v\rangle = T\left\{\sum_{i=1}^{k} v_i|e_i\rangle\right\} + T\left\{\sum_{i=k+1}^{n} v_i|e_i\rangle\right\} = \sum_{i=k+1}^{n} v_i T|e_i\rangle$

The $T|e_i\rangle$ are linearly independent: if $\sum_{i=k+1}^{n} c_i T|e_i\rangle = 0$

$$\Rightarrow T\sum_{i=k+1}^{n} c_i|e_i\rangle = 0$$

$\Rightarrow \sum_{i=k+1}^{n} c_i|e_i\rangle$ is in null space of T $\Rightarrow$ only soln. $c_i = 0$

$\Rightarrow T|e_i\rangle, i=k+1\cdots n$ form a basis of range of T $\Rightarrow \text{rank}(T) = n-k$.

In the matrix representation of T, the range is the space spanned by the columns of the matrix A (we will also call it column space).

Take, eg., $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$ Column space: $\mathbb{R}^3$

null space: ϕ

Eq. $A\vec{x} = \vec{b}$: Soln. if $\vec{b}$ is in column space of A

$\Rightarrow$ Soln. for any 3-component $\vec{b}$

for $D = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 0 \\ -2 & -1 & 2 \end{pmatrix}$: Column space: $\mathbb{R}^2$: $a_1\begin{pmatrix}1\\2\\1\end{pmatrix} + a_2\begin{pmatrix}1\\0\\2\end{pmatrix}$: $\vec{b}$ must be of this form

Vectors $\gamma\begin{pmatrix}-1/2\\1\\0\end{pmatrix}$ solve $A\vec{x} = 0$: $\mathbb{R}^1$

Example: $\frac{d}{dx}$ on cubic polynomials.

Set of all linear transformations from V to W from a vector space $L(V,W)$ with the definitions

$$(T_1+T_2)\,|v\rangle = T_1|v\rangle + T_2|v\rangle$$

$$(cT)\,|v\rangle = c\,(T|v\rangle)$$

If $\dim V = n$ and $\dim W = m$, $\dim L(V,W) = mn$.

If W is the field of scalars F, T is called a linear functional.

Example: If A is $n\times n$ matrix, $\mathrm{tr}\,A$ forms a linear functional.

Example: $V = \mathbb{R}^n$, $f : |x\rangle \to \sum_i a_i x_i$ is a linear functional.

Set of all linear functionals on V : vector space of dimn n (dual, V^*)

If V is an inner product space, this space is same as the space of the bra vectors: any functional can be written as an inner product with some bra.

Given any $f(V)$, can find a unique bra $\langle f|$ such that $f|v\rangle = \langle f|v\rangle$

Proof: f determined by its action on an orthonormal basis $|e_i\rangle$.

Let $f(|e_i\rangle) = a_i \;\forall\; i = 1\ldots n$. Then $\langle f| = \langle \sum a_i^* e_i|$.

To show that $\langle f|$ is unique; if $\langle g|$ has same action on all $|v\rangle$: $\langle g-f|v\rangle = 0 \;\forall\; |v\rangle \Rightarrow \langle g-f| = \langle 0|$.

—o—

In general, given a basis $\{|e_i\rangle\}$ in V, $\exists$ set of functionals (unique) $f_i(|e_j\rangle) = \delta_{ij}$ which form a basis of V^*.

Define product of linear transformations: if $T: V \to W$ and $U: W \to Z$

then product matrix $UT : V \to Z$ defined by

$$(UT)\,|v\rangle = U\left(T|v\rangle\right).$$

UT is a linear transformation. Easy to see that matrix representations of U & T $\Rightarrow$ matrix represn. of UT.

Associativity: $U(TS) = (UT)S$; distributive: $U(T+S) = UT + US$

For finding column space: row-reduced echelon form.

(1) For the matrix eqn. $A\vec{x} = \vec{b}$, $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix}$

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 8 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix}$$

$$= LU$$

Solve: $LU\vec{x} = \vec{b}$: $L\vec{c} = \vec{b}$, $U\vec{x} = \vec{c}$

$$\begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -2 \\ 9 \end{pmatrix} \Rightarrow \vec{c} = \begin{pmatrix} 5 \\ -12 \\ 2 \end{pmatrix} \qquad \begin{pmatrix} 2 & 1 & 1 \\ 0 & -8 & -2 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 5 \\ -12 \\ 2 \end{pmatrix} \Rightarrow \vec{x} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$$

[Similar to Gaussian elimination, but better for multiple rhs.]

LDU: $$A = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & -1 & 1 \end{pmatrix}\begin{pmatrix} 2 & & \\ & -8 & \\ & & 1 \end{pmatrix}\begin{pmatrix} 1 & 1/2 & 1/2 \\ 0 & 1 & 1/4 \\ 0 & 0 & 1 \end{pmatrix}.$$

$$B = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 0 \\ -2 & 7 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 8 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}}_{P}\begin{pmatrix} 2 & 1 & 1 \\ 0 & 8 & 3 \\ 0 & 0 & -2 \end{pmatrix}$$

$$PB = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}\begin{pmatrix} 2 & 1 & 1 \\ 0 & 8 & 3 \\ 0 & 0 & -2 \end{pmatrix}$$

(Pivoting)

Take $S = \begin{pmatrix} 2 & 1 & 1 \\ 4 & 2 & 0 \\ -2 & -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 3 \end{pmatrix}$ ⇒ singular case

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

⇒ For null space we solve $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = 0$ ⇒ $\vec{x} = y \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}$

Take another example: $A' = \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix}$

$$A' = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 3 & 2 \\ 2 & 6 & 9 & 5 \\ -1 & -3 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

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

Null space: $U\vec{x} = 0$ ⇒ $(1\ 3)\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + (3\ 2)\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = 0$

$(3\ 1)\begin{pmatrix} x_3 \\ x_4 \end{pmatrix} = 0$

we call x_1, x_3 ⇒ variables for columns with pivots ⇒ basic and

x_2, x_4: free. $x_3 = -\frac{1}{3}x_4$, $x_1 = -3x_2 - x_4$

general soln for $U\vec{x} = 0$: $x_2 \begin{pmatrix} -3 \\ 1 \\ 0 \\ 0 \end{pmatrix} + x_4 \begin{pmatrix} -1 \\ 0 \\ -1/3 \\ 1 \end{pmatrix}$

If n eqns for m unknowns, $n < m$: A' has m columns and n rows ⇒ at most n pivots ⇒ at least $(m-n)$ dimensional nullspace.

$S\vec{x} = \vec{b} \Rightarrow U\vec{x} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 3/2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ -2b_1 + b_2 \\ -2b_1 + \frac{3}{2}b_2 + b_3 \end{pmatrix}$

$U = \begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \Rightarrow -2b_1 + \frac{3}{2}b_2 + b_3 = 0$ one condition $\Rightarrow$ set of all $\vec{b}$ span $\mathbb{R}^2$

$\Rightarrow$ Soln $\begin{pmatrix} 2 & 1 & 1 \\ 0 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} b_1 \\ -2b_1 + b_2 \\ 0 \end{pmatrix}$

Take, eg., $(b_1, b_2, b_3) = (1, 2, -1) \Rightarrow 2x + y + z = 1, \quad -2z = 0$

$\Rightarrow$ soln. $= \begin{pmatrix} 1/2 \\ 0 \\ 0 \end{pmatrix} + y \begin{pmatrix} -1/2 \\ 1 \\ 0 \end{pmatrix}$

for A': we get $\begin{pmatrix} 1 & 3 & 3 & 2 \\ 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 - 2b_1 \\ b_3 - 2b_2 + 5b_1 \end{pmatrix}$

$\Rightarrow b_3 - 2b_2 + 5b_1 = 0 \Rightarrow \{\vec{b}\}$ span $\mathbb{R}^2$.

Row-reduced echelon form for A': $\begin{pmatrix} 1 & 3 & 0 & 1 \\ 0 & 0 & 1 & 1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$ [additional operation: multiplication of row by a scalar]

for S: $\begin{pmatrix} 1 & 1/2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$

① all the pivots $= 1$

② each column with pivot has all other elements $= 0$

③ higher row has pivot towards left

Can show ① Any matrix can be brought to (unique) row-reduced echelon form

② Non-zero rows form a basis of rowspace. dim = # pivots = r

③ null space of U: dim: # columns without pivots $= n-r$

$U\vec{x} = 0 \Leftrightarrow LU\vec{x} = 0 \Rightarrow$ nullity $(A) = n-r$

A basis can be formed by setting all but one free variables to 0 and solving for the others.

④ Columns of U with pivots ⇒ linearly independent ⇒ form a basis for range of U : $U\vec{x} = \vec{c}$ so rank(U) $= r$

⇒ dimensionality of column space of A $= r$ $[\vec{b} = L\vec{c}]$

Columns of A that correspond to columns of U with pivot ⇒ form a basis of range (A)

Proof: need to show that $\sum_{i=1}^{r} x_i$ "i th" column(A) $= 0 \Rightarrow x_i = 0$

where these columns correspond to those with pivots in U

i.e. looking for $A\vec{x} = 0$ where in $\vec{x}$, the free variables $= 0$

Look for soln $U\vec{x} = 0$ with the free variables $= 0$

$\Rightarrow \vec{x} = 0$.

Some further properties:

① If A has right inverse, $AC = I$

$\Rightarrow$ for any $\vec{b}$, $A\vec{x} = \vec{b}$ with $\vec{x} = C\vec{b}$

$\Rightarrow$ range (A) $= m$ Possible only if $m \leq n$

② If A has left inverse, $BA = I$

$\Rightarrow$ for $A\vec{x} = \vec{b}$, $\vec{x} = B\vec{b}$ i.e. soln unique

In general, if $A\vec{x} = \vec{b}$ has a soln., can add any vector from null space to $\vec{x}$

$\Rightarrow$ if A has left inverse, null space $= \{\phi\}$ $\Rightarrow r = n$
possible only if $m \geq n$

$\Rightarrow$ If A has ~~both~~ left and right inverse, $m = n$ and rank $= m = n$

$\Rightarrow$ $A\vec{x} = \vec{b}$ has unique soln $\forall\ \vec{b}$

$\Rightarrow$ $C = (BA)C = B(AC) = B$ $\Rightarrow$ inverse of A (written A^{-1})

$\Rightarrow$ A has LDU decomposition with all diagonal elements of D nonzero.

To find A^{-1} : $AA^{-1} = I \Rightarrow A.(A^{-1})_{i\text{th column}} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \\ 0 \\ \vdots \end{pmatrix} \rightarrow i$

$\Rightarrow$ do Gauss elimination with all i rhs vectors

Example. $A = \begin{pmatrix} 2 & 1 & 1 \\ 4 & -6 & 0 \\ -2 & 7 & 2 \end{pmatrix}$ $\begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 4 & -6 & 0 & 0 & 1 & 0 \\ -2 & 7 & 2 & 0 & 0 & 1 \end{pmatrix} \xrightarrow[L^{-1}]{\text{row reduction}} \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & -8 & -2 & -2 & 1 & 0 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix}$ (DU | L^{-1})

$\Rightarrow$ LDU

$\downarrow$ row reduction $U'^{-1} = DU^{-1}D^{-1}$

$\begin{pmatrix} 1 & 0 & 0 & 3/4 & -5/16 & -3/8 \\ 0 & 1 & 0 & 1/2 & -3/8 & -1/4 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix} \longleftarrow \begin{pmatrix} 2 & 0 & 0 & 3/2 & -5/8 & -3 \\ 0 & -8 & 0 & -4 & 3 & 2 \\ 0 & 0 & 1 & -1 & 1 & 1 \end{pmatrix}$

(right block: $U^{-1}D^{-1}L^{-1}$; D | $DU^{-1}D^{-1}L^{-1}$)

Gauss - Jordan method.

Works if all the pivots nonzero.

For a $n \times n$ matrix, define determinant:

$$\det A = \varepsilon_{i_1 i_2 \cdots i_n} \underset{\text{1st column}}{A_{i_1 1}} \cdots A_{i_n n}$$

equivalently, $\varepsilon_{i_1 i_2 \cdots i_n} A_{i_1 j_1} \cdots A_{i_n j_n} = \varepsilon_{j_1 \cdots j_n} \det A$.

e.g., $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \Rightarrow \det A = a_{11} a_{22} - a_{12} a_{21}$

Properties:

1. Changes sign for each permutn. of columns or rows

2. $\det A^T = \det A \Rightarrow$ could write formula for rows too $\left((A^T)_{ij} = A_{ji} \right)$

3. det linear in any single row: $\varepsilon_{i_1 \cdots i_n} (A_{i_1 1} + B_{i_1 1}) A_{i_2 2} \cdots A_{i_n n}$

4. ~~[scribbled out]~~ multiplying a column/row by $c \Rightarrow c \times \det$

5. Subtracting $c \cdot i$th row from another row does not change det

~~[scribbled out]~~

6. Say $A = CD \Rightarrow A_{ij} = C_{ik} D_{kj}$

$$\Rightarrow \varepsilon_{i_1 \cdots i_n} A_{i_1 1} \cdots A_{i_n n} = \varepsilon_{i_1 \cdots i_n} C_{i_1 k_1} D_{k_1 1} C_{i_2 k_2} D_{k_2 2} \cdots C_{i_n k_n} D_{k_n n}$$

$$= \det C \cdot \varepsilon_{k_1 \cdots k_n} D_{k_1 1} \cdots D_{k_n n}$$

$$= \det C \cdot \det D.$$

7. Triangular matrix: $\det = \prod(\text{diagonal elements})$

$\Rightarrow$ (after elementary row operations) if # pivots $\neq n$, $\det = 0$

$\Rightarrow$ if $\det \neq 0$, matrix invertible.

8. define $\tilde{A}_{ij} = \text{cofactor}(A_{ij}) = (-1)^{i+j} \det M_{ij}$ where

M_{ij} is formed after deleting ith row and jth column of A

Then $\det A = A_{i1} \tilde{A}_{i1} + A_{i2} \tilde{A}_{i2} + \cdots + A_{in} \tilde{A}_{in}$.

Now form $\mathrm{cofactor}(A)_{ij} = \tilde{A}_{ji}$

$\Rightarrow [AC]_{ij} = A_{ik}\tilde{A}_{jk} = \det A$ if $i=j$

If $i \neq j$, $A_{ik}\tilde{A}_{jk} \Rightarrow$ determinant of a matrix with ith row repeated $\Rightarrow 0$

$\Rightarrow \quad A^{-1}_{ij} = (\det A)^{-1}\, \tilde{A}_{ji}$.

If A is nonsingular, $A\vec{x} = \vec{b}$ is solved by $\vec{x} = A^{-1}\vec{b}$

$\Rightarrow x_i = A^{-1}_{ij} b_j = (\det A)^{-1}\, b_j \tilde{A}_{ji}$

$= (\det A)^{-1} \det B_j$, $B_j =$ Matrix A with jth column replaced by b_j.

Eigenvalues and eigenvectors

If $A|x\rangle = \lambda|x\rangle \Rightarrow |x\rangle$ eigenvector, λ eigenvalue

eg. σ_1 as example, eigenvectors : $\begin{pmatrix}1\\1\end{pmatrix}, \begin{pmatrix}1\\-1\end{pmatrix}$ eigenvalues $1, -1$

If all vectors are eigenvectors of A, $A = cI$

If $A|x\rangle = \lambda|x\rangle$, $A^m|x\rangle = \lambda^m|x\rangle$ Converse not true : all vectors eigenvect of σ

If A : eigenvalues $\lambda_1 \cdots \lambda_n$, A^m : $\lambda_1^m \cdots \lambda_n^m$

Any $n \times n$ matrix has at least one eigenvector.

Secular / characteristic eqn : If $A|x\rangle = \lambda|x\rangle \Rightarrow (A - \lambda I)|x\rangle = 0$

$\Rightarrow \det(A - \lambda I) = 0$: Secular eqn.

This eqn. has at least one root [n roots but some degenerate]

$\Rightarrow \text{rank}(A - \lambda I) < n \Rightarrow$ null space of $(A - \lambda I)$ is non-null

$\Rightarrow$ at least one eigenvector corresponding to each distinct eigenvalue.

Eigenvectors for distinct eigenvalues are linearly independent.

Proof : by induction

$|x_1\rangle$: eigenvalue λ_1 $|x_2\rangle \neq c|x_1\rangle$ if $\lambda_2 \neq \lambda_1$

$|x_3\rangle$: if not LI, $|x_3\rangle = c_1|x_1\rangle + c_2|x_2\rangle$

$A|x_3\rangle = \lambda_3|x_3\rangle = (\lambda_3 c_1)|x_1\rangle + \lambda_3 c_2|x_2\rangle = \lambda_1 c_1|x_1\rangle + \lambda_2 c_2|x_2\rangle$

$\Rightarrow (\lambda_3 - \lambda_1) c_1|x_1\rangle + (\lambda_3 - \lambda_2) c_2|x_2\rangle = 0$

Since $|x_1\rangle, |x_2\rangle$ LI $\Rightarrow (\lambda_3 - \lambda_1)c_1 = 0, (\lambda_3 - \lambda_2) c_2 = 0$

Now say first $|x_1\rangle \cdots |x_j\rangle$ LI, $|x_{j+1}\rangle = \sum_{i=1}^{j} c_i|x_i\rangle$ with some $c_i \neq$

$A|x_{j+1}\rangle = \lambda_{j+1}\sum c_i|x_i\rangle = \sum c_i\lambda_i|x_i\rangle \Rightarrow c_i(\lambda_i - \lambda_{j+1}) = 0 \;\forall i = 1$

$\Rightarrow$ If all n eigenvalues distinct, n LI eigenvectors $\Rightarrow$ basis.

Secular eqn. $\det(A-\lambda I)=0 \Rightarrow$ polynomial in $\lambda \Rightarrow (\lambda-a_1)\cdots(\lambda-a_n)=0$

$$= \lambda^n - (a_1+\cdots+a_n)\lambda^{n-1} + \cdots + (-1)^n a_1\cdots a_n = 0$$

$$\det(A-\lambda I) = \varepsilon_{i_1\cdots i_n}(a_{i_1 1}-\lambda\delta_{i_1 1})\cdots(a_{i_n n}-\lambda\delta_{i_n n})$$

$$= (-1)^n\lambda^n + (-1)^{n-1}\lambda^{n-1}(a_{11}+\cdots+a_{nn}) + \cdots + \varepsilon_{i_1\cdots i_n} a_{i_1 1}\cdots a_{i_n n}$$

$\Rightarrow$ $\sum$ eigenvalues $=$ tr A, Π eigenvalues $= \det A$

Special matrices:

if H hermitian: $H|x\rangle = \lambda|x\rangle \Rightarrow \langle x|H|x\rangle = \lambda\langle x|x\rangle = \langle x|H|x\rangle^* = \lambda^*\langle x|x\rangle$

$\Rightarrow$ λ real

if $\lambda_x \neq \lambda_y$: $\langle x|H|y\rangle = \lambda_y\langle x|y\rangle = \langle y|H|x\rangle^* = \lambda_x\langle x|y\rangle$

$\Rightarrow \langle x|y\rangle = 0$

if U unitary, $UU^+ = \mathbb{1}$ $\quad U|x\rangle = \lambda|x\rangle \Rightarrow |\lambda|^2\langle x|x\rangle = \langle x|\underbrace{U^+U}_{1}|x\rangle$

$\Rightarrow |\lambda|^2 = 1 \Rightarrow \lambda = e^{i\theta}$

if $\lambda_x \neq \lambda_y$, $\langle y|\underbrace{U^+U}_{1}|x\rangle = \lambda_x\lambda_y^*\langle y|x\rangle$

$= \lambda_y\lambda_y^*\langle y|x\rangle \quad \Rightarrow \langle y|x\rangle = 0$

if normal operator, $AA^+ = A^+A$

if $A|x\rangle = \lambda|x\rangle \Rightarrow A^+|x\rangle = \lambda^*|x\rangle$

proof: $\| A^+|x\rangle - \lambda^*|x\rangle \|^2 = (\langle x|\lambda - \langle x|A)(A^+|x\rangle - \lambda^*|x\rangle) = 0$

if $\lambda_y \neq \lambda_x$: $\langle y|A|x\rangle = \lambda_x\langle y|x\rangle = \lambda_y\langle y|x\rangle \Rightarrow \langle y|x\rangle = 0$

If A is real, eigenvalues either real or come in conjugate pairs

if λ real, can choose $|x\rangle$ real

$A|x\rangle = \lambda|x\rangle \quad |x\rangle = |x\rangle_R + i|x\rangle_I \Rightarrow A|x_R\rangle = |x_R\rangle$

If λ complex, $|x\rangle$ complex and $|x^*\rangle$: eigenvalue λ^*

real symmetric matrix : all real eigenvalues

eigenvectors : can be chosen to be orthogonal

real orthogonal matrix: eigenvalues unimodular so $1, -1, (e^{i\theta}, e^{-i\theta})$

$\Rightarrow$ rotation in 3d: $\exists |x\rangle$ such that $O|x\rangle = |x\rangle$.
(det=1)

Matrix representation of A : $A|e_i\rangle = a_{ji}|e_j\rangle$

If A has n l.i. eigenvectors $|x_i\rangle \Rightarrow$ form a basis

in this basis : $a'_{ij} = \lambda_i \delta_{ij}$ diagonalizes

transforms of basis : $|x_i\rangle = S_{ji}|e_j\rangle \quad |e_j\rangle = S^{-1}_{kj}|x_k\rangle$

for arbitrary matrix A : transform to $|x\rangle$ basis :

$$A|x_i\rangle = a'_{ji}|x_j\rangle$$

$$\text{lhs} = A S_{ji}|e_j\rangle = S_{ji} a_{lj}|e_l\rangle = S_{ji} a_{lj} S^{-1}_{kl}|x_k\rangle$$

$$\Rightarrow a'_{ki} = S^{-1}_{kl} a_{lj} S_{ji} \quad \Rightarrow A' = S^{-1}AS$$

if $|x\rangle$ basis of eigenvectors : $A' = S^{-1}AS$ diagonal

transforms of form $S^{-1}AS$: similarity transforms

if $A|x\rangle = \lambda|x\rangle \Rightarrow A' S^{-1}|x\rangle = \lambda S^{-1}|x\rangle$

$\Rightarrow$ eigenvalues and # of eigenvectors unchanged

when S : matrix of eigenvectors : $S = (x_1 \; x_2 \cdots x_n)$

$$S^{-1}AS = S^{-1}(\lambda_1 x_1 \; \lambda_2 x_2 \cdots \lambda_n x_n) = \begin{pmatrix} \lambda_1 & & \\ & \lambda_2 & \\ & & \ddots \end{pmatrix}$$

$\Rightarrow$ if A diagonalized by S, columns of S eigenvectors of A.

If $|e_i\rangle, |x_i\rangle$ orthonormal basis, S unitary

$|x_i\rangle = S_{ji}|e_j\rangle \quad \langle x_i| = \langle e_j| S^*_{ji}$

$\Rightarrow \langle x_i|x_l\rangle = \langle e_j| S^*_{ji} S_{kl}|e_k\rangle = S^*_{ji} S_{jl} = \delta_{il}$

If $|e_i\rangle$ orthonormal and S unitary, $|x_i\rangle$ orthonormal.

$\Rightarrow$ Column vectors of unitary matrix form an orthonormal basis.

If $|x\rangle$ orthonormal basis, $U = (x_1 \; x_2 \cdots x_n)$ unitary.

$\Rightarrow$ A matrix can be diagonalized by unitary transform iff it has a complete set of mutually orthogonal eigenvectors.

For any $n \times n$ matrix A, $\exists$ unitary U such that $U^{-1}AU = T$ is upper triangular

Proof by construction: $\exists$ at least one eigenvector $|x_1\rangle$, eigenvalue λ_1

construct orthonormal basis $|x_1\rangle, |b\rangle, |c\rangle, \ldots$

$$U_1^{-1} A U_1 = U^{-1} \left(\lambda |x\rangle \; (\cdots) \right)$$

$$= \begin{pmatrix} \lambda_1 & \\ 0 & \\ 0 & M \\ 0 & \\ \vdots & \end{pmatrix}$$

Now take $(n-1)\times(n-1)$ M : form $U_2 = \begin{pmatrix} 1 & 0 & 0 & \cdots \\ 0 & |x_2\rangle & \cdots & \\ 0 & & & \\ 0 & & & \\ \vdots & & & \end{pmatrix}$

$$U_2^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots \\ 0 & & & & \\ 0 & & (\cdot & \cdot) & \\ 0 & & & & \\ \vdots & & & & \end{pmatrix} \qquad (U_2^{-1} U_1^{-1}) A (U_1 U_2) = \begin{pmatrix} \lambda_1 & \cdots & \cdots \\ 0 & \lambda_2 & \cdots \\ 0 & 0 & \ddots \\ 0 & 0 & \\ \vdots & 0 & \end{pmatrix}$$

Repeating, we get $T^{-1} A T = \begin{pmatrix} \lambda_1 & \cdots & \cdots \\ 0 & \lambda_2 & \cdots \\ 0 & 0 & \ddots \\ 0 & 0 & \\ 0 & 0 & \end{pmatrix}$ where $T = U_1 U_2 \cdots$

Since T unitary, $T^{-1} H T$ hermitian if H hermitian

$\Rightarrow$ H can be diagonalized by a unitary transformation

$\Rightarrow$ H has a complete set of orthonormal eigenvectors

Similarly, real symmetric matrix can be diagonalized by an orthogonal matrix.

Example: nondiagonal matrix: $\begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

nondiagonal matrix: $\begin{pmatrix} 2 & -1 \\ 1 & 0 \end{pmatrix}$ eigenvalue: 1, eigenvector: $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$

construct $U = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ $\quad U^{-1} A U = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$

Take $\begin{pmatrix} 1 & \sin\theta \\ \sin\theta & 1 \end{pmatrix}$ eigenvalues $1 \pm s$, orthonormal eigenvectors: $\frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \end{pmatrix}$

as $s \to 0$, eigenvalues collapse but eigenvectors remain orthogonal

$\begin{pmatrix} 1 & \cos\theta \\ \sin\theta & 1 \end{pmatrix}$ eigenvalues: $(1 \pm \sqrt{cs})$ eigenvectors: $\begin{pmatrix} \sqrt{c} \\ \sqrt{s} \end{pmatrix}, \begin{pmatrix} \sqrt{c} \\ -\sqrt{s} \end{pmatrix}$

as $\theta \to 0$, eigenvectors collapse to $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$.

If A is normal: $U_1^{-1} A U_1 = \begin{pmatrix} \lambda_1 & 0 \; 0 \cdots \\ 0 & \\ 0 & (\quad) \\ 0 & \\ \vdots & \end{pmatrix}$

$U_1 = (X_1, Y_2 \cdots Y_n)$ $U_1^+ A U_1 = \begin{pmatrix} X_1^+ \\ Y_2^+ \\ \vdots \\ Y_n^+ \end{pmatrix} A (X_1\, Y_2 \cdots Y_n) = \begin{pmatrix} X_1^+AX_1 & X_1^+AY_2 & \cdots & X_1^+AY_n \\ Y_2^+AX_1 & Y_2^+AY_2 & \cdots & Y_2^+AY_n \\ \vdots & \vdots & & \\ Y_n^+AX_1 & Y_n^+AY_2 & \cdots & Y_n^+AY_n \end{pmatrix}$

Since $X_1^+ A = \lambda_1 X_1^+$ $\Rightarrow U_1^+ A U_1 = \begin{pmatrix} \lambda_1 & 0 \; 0 \cdots \\ 0 & \\ 0 & B \\ \vdots & \end{pmatrix}$

Since U normal, B normal $\Rightarrow$ repeat.

Also if $U^+ A U = D \Rightarrow A = U D U^+ \Rightarrow A A^+ = U D D^+ U^+ = A^+ A$

$\Rightarrow A$ normal

$\Rightarrow$ matrix A can be diagonalized by unitary transformation,

and therefore has complete orthonormal eigenvectors, iff A normal.

<u>Cayley - Hamilton theorem</u> A matrix satisfies its own ~~eigenvalue~~ secular eqn. : if $\det(\lambda I - A) = f(\lambda)$, then $f(A) = 0$.

<u>Proof</u> Form the matrix K_{ij} = (ji) th cofactor of $M = \lambda I - A$

$= (-1)^{i+j} \det$ (matrix with ith row, jth column of M omitted)

$\Rightarrow M_{ij} \cdot K_{jk} = \delta_{ik} \det M \Rightarrow M \cdot K = f(\lambda) I$.

The elements of K : polynomials in λ

$$\Rightarrow K = \sum_{i=0}^{n-1} \lambda^i K_i$$

Now $f(\lambda) \cdot I = M \cdot K = (\lambda I - A) \cdot \sum_{i=0}^{n-1} \lambda^i K_i = \sum_{i=0}^{n-1} \lambda^{i+1} K_i - \sum_{i=0}^{n-1} \lambda^i A \cdot K_i$

$$= \lambda^n K_{n-1} + \sum_{i=1}^{n-1} \lambda^i (K_{i-1} - A \cdot K_i) - A \cdot K_0$$

If $f(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \dots + c_1 \lambda + c_0$,
Equating coefficients of λ^i : $I = K_{n-1}$, $c_i I = K_{i-1} - A \cdot K_i$ for $1 \le i \le n-1$

$$c_0 I = -A \cdot K_0$$

$$\Rightarrow f(A) = A^n + c_{n-1} A^{n-1} + \dots + c_1 A + c_0 = A^n K_{n-1} + \sum_{i=1}^{n-1} A^i (K_{i-1} - A \cdot K_i) - A \cdot K_0$$

$$= 0.$$

— —

Write the secular eqn. as $f(\lambda) = (\lambda - \lambda_1)^{m_1} (\lambda - \lambda_2)^{m_2} \dots (\lambda - \lambda_p)^{m_p}$

$\lambda_1, \dots \lambda_p$ distinct eigenvalues with degeneracies $m_1 \dots m_p$,
$\sum m_i = n$.

If matrix diagonalizable : A satisfies $p(A) = 0$

where $p(A) = (\lambda-\lambda_1) \cdots (\lambda-\lambda_p)$

(Clear if expand arbitrary vector in basis of eigenvectors).

Call minimal polynomial $p(A)$: polynomial $(A-\lambda_1)^{i_1} \cdots (A-\lambda_p)^{i_p} = 0$

for the smallest set $i_1 \cdots i_p$

If $i_1 \cdots i_p \neq 1 \; \forall p$, matrix not diagonalizable.

Proof: If matrix diagonalizable, eigenvectors form a basis

~~What is the simplest form a non-diagonalizable matrix takes~~

Expand a vector $|V\rangle = \sum v_i |x_i\rangle$, $A|x_i\rangle = \lambda_i |x_i\rangle$

Then $(A-\lambda_1) \cdots (A-\lambda_p) |V\rangle = \sum v_i (A-\lambda_1) \cdots (A-\lambda_p) |x_i\rangle$

$= 0.$

What is the simplest form a non-diagonalizable matrix takes under similarity transformation?

First, some formalism.

Say we have an eigenvalue λ_i with degeneracy m_i and only one eigenvector.

We define generalized eigenvectors : vectors $|y_i\rangle$ such that $(A-\lambda_i I)^k |y_i\rangle = 0$ but $(A-\lambda_i I)^{k-1} \neq 0$ [$|y_i\rangle$: rank k generalized eigenvector]

Then $|y_{m_i}\rangle, (A-\lambda_i I)|y_{m_i}\rangle, (A-\lambda_i I)^2 |y_{m_i}\rangle, \cdots, (A-\lambda_i I)^{m-1} |y_{m_i}\rangle$

Constitute the null space of $(A-\lambda_i I)^{m_i}$

We will show that the set of all the generalized eigenvectors of a matrix form a basis of the vector space, and the space spanned by generalized eigenvectors of two eigenvalues are orthogonal to each other.

Steps ①. Take the characteristic eqn. $f(\lambda) = \prod_{i=1}^{p} (\lambda-\lambda_i)^{m_i}$, $\sum_{i=1}^{p} m_i = N$

$$\frac{1}{f(\lambda)} = \sum_{i=1}^{p} \frac{a_i(\lambda)}{(\lambda-\lambda_i)^{m_i}}, \quad a_i(\lambda) : \text{polynomial of degree} \leq m_i$$

$$\Rightarrow 1 = \sum_{i=1}^{p} \underbrace{a_i(\lambda) \prod_{j\neq i} (\lambda-\lambda_j)^{m_j}}_{F_i(\lambda)} = \sum_{i=1}^{p} F_i(\lambda)$$

Writing the rhs as a polynomial in λ, $a_i(\lambda)$ such that coefficient of all $\lambda^i = 0$ except $\lambda^0 \Rightarrow 1$

therefore $\sum_{i=1}^{p} F_i(A) = 1$

$\Rightarrow$ any vector can be written $|v\rangle = \sum_{i=1}^{p} F_i(A)|v\rangle$

② The $F_i(A)$ project to orthogonal subspaces.

if $i \neq j$, $F_i(A) F_j(A)|v\rangle$ has the product $\prod (\lambda-\lambda_j)^{m_j} = f(\lambda)$

$\Rightarrow f(A) = 0$ imply $F_i(A) F_j(A)|v\rangle = 0$

Now using $1 = F_i(A) + \sum_{j\neq i} F_j(A) \Rightarrow F_i(A)|v\rangle = F_i(A) F_i(A)|v\rangle + \underbrace{\sum_{j\neq i} F_j(A) F_i(A)|v\rangle}_{0}$

$\Rightarrow F_i(A) F_j(A)|v\rangle = \delta_{ij} F_j(A)|v\rangle$

$\Rightarrow$ vectors $F_i(A)|v\rangle$ linearly independent : $\sum c_i F_i(A)|v\rangle = 0 \Rightarrow c_i = 0$ (for all $|v\rangle$)

Also the space $F_i(A)|v\rangle$ ($\forall |v\rangle$) is an invariant subspace of A:

$$A\, F_i(A)\,|v\rangle = F_i(A)\, A|v\rangle .$$

$\Rightarrow$ the sets $\phi_i(A)|v\rangle$ break V into orthogonal subspaces V_i:

$$V = \bigoplus_{i=1}^{p} V_i$$

③ Easy to check that V_i is the space spanned by generalized eigenvectors corresponding to λ_i.

$$\text{if } |x_i\rangle \in \phi_i(A)|v\rangle, \quad (A-\lambda_i I)^{m_i}|x_i\rangle = 0 .$$

Also from direct sum decomposition, if $(A-\lambda_i I)^{m_i}|x_i\rangle = 0$, $|x_i\rangle \in \phi_i(A)|v\rangle$

④ Let us start with a generalized eigenvector of rank m_i. Then the set $|y_{m_i}\rangle, (A-\lambda_i I)|y_{m_i}\rangle, \ldots, (A-\lambda_i I)^{m_i-1}|y_{m_i}\rangle$ are linearly independent.

For, if $\sum_{i=0}^{m_i-1} c_i (A-\lambda_i I)^i |y_{m_i}\rangle = 0$

multiplying by $(A-\lambda_i I)^{m_i-1}$ $\Rightarrow$ $c_0 = 0$

now " " $(A-\lambda_i I)^{m_i-2}$ $\Rightarrow$ $c_1 = 0$ etc. $\Rightarrow$ LI

$\Rightarrow$ These m_i vectors form a basis of nullspace of $(A-\lambda_i I)^{m_i}$ and, the nullspaces for different i break the vector space V into orthogonal subspaces.

The whole set of generalized eigenvectors form a basis for V.

Action of A: within V_i, Calling

$$(A-\lambda_i I)^{m_i-1}|y_{m_i}\rangle = |z_1\rangle$$
$$(A-\lambda_i I)^{m_i-2}|y_{m_i}\rangle = |z_2\rangle$$
$$\ldots\ldots$$
$$|y_{m_i}\rangle = |z_{m_i}\rangle$$

$$(A-\lambda_i I)|z_1\rangle = 0 \Rightarrow A|z_1\rangle = \lambda_i|z_1\rangle$$

$$(A-\lambda_i I)|z_2\rangle = |z_1\rangle \Rightarrow A|z_2\rangle = \lambda_i|z_2\rangle + |z_1\rangle$$

$$(A-\lambda_i I)|z_3\rangle = |z_2\rangle \Rightarrow A|z_3\rangle = \lambda_i|z_3\rangle + |z_2\rangle$$

and so on.

So in this subspace, using the basis $\{|z_1\rangle |z_2\rangle \cdots |z_{m_i}\rangle\}$ A looks like

$$\begin{pmatrix} \lambda_i & 1 & 0 & \cdots & \\ 0 & \lambda_i & 1 & 0 & \cdots \\ 0 & 0 & \lambda_i & 1 & 0 \cdots \\ \vdots & \vdots & 0 & \lambda_i & \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

When I take the basis consisting of all the generalized eigenvectors, A looks like

$$\begin{pmatrix} \lambda_1 & & & & & & \\ & \lambda_2 & 1 & & & & \\ & & \lambda_2 & & & & \\ & & & \lambda_3 & 1 & & \\ & & & & \lambda_3 & 1 & \\ & & & & & \lambda_3 & \\ & & & & & & \ddots \end{pmatrix} \quad \text{etc}$$

$\Rightarrow$ Jordan Canonical form. Diagonal elements = eigenvalues, off-diagonal elements = 0 except in the subspaces corresponding to generalized eigenvector of rank > 1, elements above diagonal = 1.

most simple form an arbitrary matrix can take by similarity transform.

Diagonal matrix $\Rightarrow$ Jordan form = diagonal form.