Eigenvalues and eigenvectors:

Motivation: Ordinary differential equations

$$\begin{cases} \dfrac{dv}{dt} = 4v - 5w, & v = 8 \text{ at } t=0 \\[2ex] \dfrac{dw}{dt} = 2v - 3w, & w = 5 \text{ at } t=0 \end{cases}$$

$$u(t) = \begin{bmatrix} v(t) \\ w(t) \end{bmatrix}$$

$$\frac{du}{dt} = Au, \quad u = u(0) \text{ at } t=0$$

Simpler case:- Single equation

$$\frac{du}{dt} = au \qquad \text{———} \quad (*)$$

Soln: $u(t) = e^{at}\, u(0)$

(*) is unstable	if $a>0$	e^{at} goes to ∞
neutrally stable	if $a=0$	bounded
stable	if $a<0$	zero

Case of a 2-d parameter:

Want solutions of the form

(**) $$v(t) = e^{\lambda t} y, \quad w(t) = e^{\lambda t} z$$

(or) $$u(t) = e^{\lambda t} x, \text{ where } x = \begin{bmatrix} y \\ z \end{bmatrix}$$

Substituting the solution above into ODE:-

$$\frac{du}{dt} = Au$$

$$\lambda e^{\lambda t} y = 4 e^{\lambda t} y - 5 e^{\lambda t} z$$

$$\lambda e^{\lambda t} z = 2 e^{\lambda t} y - 3 e^{\lambda t} z$$

Taking out $e^{\lambda t}$, we obtain

$$4y - 5z = \lambda y$$

$$2y - 3z = \lambda z$$

$$\boxed{Ax = \lambda x}$$ eigenvalue equation

λ = eigenvalue & x = eigenvector

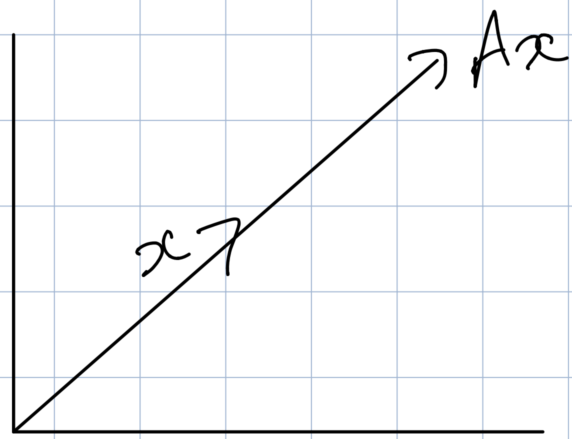
Can solve $\frac{du}{dt} = Au$ using solutions of the form ( ) if $Ax = \lambda x$ can be solved.

Need a λ & $x \neq 0$ so that $Ax = \lambda x$

Lecture-18

Eigenvalues and eigenvectors

$Ax = \lambda x$



Stretch x or shrink x but no change in direction

Case $\lambda = 0$: $Ax = 0$

$x \in N(A)$

Example : 1 Projection matrix P

Project onto a plane

$Px = x$ for x in plane.

$\lambda = 1$ is the eigenvalue & x is the eigenvector

$Px = 0 \quad \forall x$ perpendicular to the plane

$\lambda = 0$ and x is the eigenvector

(2) Permutation matrix

$$B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$Bx = x$ for $x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$Bx = -x$ for $x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

Finding the eigenvalues:

$$Ax = \lambda x$$

i.e., $(A - \lambda I)x = 0$

i.e., $(A - \lambda I)$ is singular

i.e., $\underbrace{\det(A - \lambda I) = 0}$

$\downarrow$

Characteristic polynomial of A which is of degree n

$(a_{11} - \lambda) \text{ - - - - - } (a_{nn} - \lambda)$

"n" roots = eigenvalues

For a given λ, $N(A - \lambda I)$ has to be found to obtain the eigenvectors.

Procedure to use for finding $N(A-\lambda I)$:
Elimination.

<u>Example:-</u>

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$$

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

$$\lambda^2 - 6\lambda + 8 = 0$$

$$\lambda_1 + \lambda_2 = 6, \quad \lambda_1 \lambda_2 = 8$$

Trace = sum of eigenvalues

Determinant = product —— " ——

Eigenvalues are $\lambda_1 = 4, \ \lambda_2 = 2$

$$A - 4I = \begin{bmatrix} -1 & 1 \\ 1 & -1 \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$A - 2I = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Letting $\quad B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

$$B + 3I = A$$

$$Ax = (B+3I)x = \lambda x + 3x$$
$$= (\lambda+3)x$$

Eigen vecs of B = eigenvecs of A
(why?)

Notes:

① If $Ax = \lambda_1 x$ and $Bx = \lambda_2 x$

then $\boxed{(A+B)x \overset{?}{=} (\lambda_1 + \lambda_2)x}$

does not always hold since
x's need not be the same
for A & B.

② Symmetric matrices
↳ real eigenvalues

An example where eigenvalues aren't real.

$A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ rotate by $90°$

$$\det(A - \lambda I) = \lambda^2 + 1$$

$$\lambda_1 = i \quad , \quad \lambda_2 = -i$$

An example where we don't get two independent eigenvecs?

$$A = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix}$$

$$\lambda_1 = \lambda_2 = 3$$

$$(A - \lambda I) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$$

$x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is a eigenvec

& there is no other indep. eigenvec.

Similarity & diagonalization

Suppose A has n linearly independent eigenvectors $\{x_1, \ldots, x_n\}$

$$S = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}$$

Then, $$S^{-1}AS = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix},$$

and $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of A.

Proof:

$$AS = A\begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ \lambda_1 x_1 & \cdots & \lambda_n x_n \\ | & & | \end{bmatrix}$$

$$= \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \underbrace{\begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix}}_{\text{defined to be } \Lambda}$$

$$= S\Lambda$$

So, $AS = S\Lambda$

or, $S^{-1}AS = \Lambda$ [S invertible (Why?)]

<u>Claim</u>:- If $\lambda_1, \ldots \lambda_n$ are distinct, then their eigenvectors $\{x_1, \ldots x_n\}$ are linearly independent.

<u>Proof for $n=2$</u>:

$$c_1 x_1 + c_2 x_2 = 0 \quad — (1)$$

$$A(c_1 x_1 + c_2 x_2) = 0$$

$$c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 = 0 \quad — (2)$$

$$c_1 (\lambda_1 - \lambda_2) x_1 = 0$$

$$\lambda_1 \neq \lambda_2, \quad x_1 \neq 0 \Rightarrow c_1 = 0$$

Similarly $c_2 = 0$

So, $\{x_1, x_2\}$ is a linearly indep. set

<u>H.W</u>:- Extend to the general case in "n" dimensions

Example:

$$\begin{bmatrix} 1 & 4 \\ 2 & 3 \end{bmatrix}$$

Find $\lambda_1, \lambda_2, x_1, x_2$.

Lecture-19 9/9/19

Diagonalization & Similarity

Recap:- If "A" has n independent eigenvectors $\{x_1, \ldots, x_n\}$, then

$$S^{-1}AS = \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix},$$

where $S = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}$.

Remarks:

(I) $A = S \Lambda S^{-1}$

S is not unique

x_1 is an eigenvec $\Rightarrow$ $c x_1$ is an eigenvec $\forall c \in \mathbb{R}$

So, each col. of S can be scaled to produce a new S.

(II) $A = S \Lambda S^{-1}$

Suppose col. 1. of S is y.

Col 1. of $S\Lambda = \lambda_1 y$

Col 1. of $AS = Ay$

Given $AS = S\Lambda$

$\Rightarrow Ay = \lambda_1 y \Rightarrow y$ is an eigenvec & λ_1 is an eigenval

(III) Powers of A.

Suppose λ is an eigenvalue & x an eigenvec of A.

$$A^2 x = A \lambda x = \lambda A x = \lambda^2 x$$

Suppose $S^{-1} A S = \Lambda$.

Is $S^{-1} A^2 S = \Lambda^2$?

$$(S^{-1} A S)(S^{-1} A S) = \Lambda^2$$

Works for a general k, i.e.,

$$S^{-1} A^k S = \Lambda^k.$$

(IV) Not all matrices are diagonalizable

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

Diagonalizability depends on "enough" eigenvecs

Invertibility —"— non-zero eigenvals

EXAMPLE

Fibonacci sequence

$0, 1, 1, 2, 3, 5, \ldots$

$$F_{k+2} = F_{k+1} + F_k$$

What is F_{100}?

$$F_{k+2} = F_{k+1} + F_k$$

$$F_{k+1} = F_{k+1}$$

$$u_k = \begin{bmatrix} F_{k+1} \\ F_k \end{bmatrix}, \quad u_{k+1} = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} u_k$$

Start with u_0

$$u_0 \xrightarrow{A} u_1 \xrightarrow{A} u_2 \longrightarrow \text{-----}$$

$u_k = A^k u_0$ is a solution to

$$u_{k+1} = A u_k.$$

Suppose A has n independent eigenvecs

$$u_0 = c_1 x_1 + c_2 x_2 + \text{---} + c_n x_n$$

$$A u_0 = c_1 \lambda_1 x_1 + c_2 \lambda_2 x_2 + \text{---} + c_n \lambda_n x_n$$

$$u_k = A^k u_0 = c_1 \lambda_1^k x_1 + \text{-----} + c_n \lambda_n^k x_n$$

Back to Fibonacci!

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 0 \end{bmatrix} \qquad p(\lambda) = \lambda^2 - \lambda - 1$$

$$\lambda_1 = \frac{1+\sqrt{5}}{2} \quad , \quad \lambda_2 = \frac{1-\sqrt{5}}{2}$$

$$u_k = c_1 \lambda_1^k x_1 + c_2 \lambda_2^k x_2$$

$$u_0 = \begin{bmatrix} F_1 \\ F_0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = c_1 x_1 + c_2 x_2$$

Need to find eigenvecs x_1, x_2 of A:

$$(A - \lambda I)(x) = \begin{bmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{bmatrix} x$$

$$x_1 = \begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} \quad x_2 = \begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$$

$$(A - \lambda_1 I)\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} = \begin{bmatrix} \lambda_1^2 - \lambda_1 - 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \left(\frac{1}{\sqrt{5}}\right)\begin{bmatrix} \lambda_1 \\ 1 \end{bmatrix} + \left(-\frac{1}{\sqrt{5}}\right)\begin{bmatrix} \lambda_2 \\ 1 \end{bmatrix}$$

$$F_k = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^k - \frac{1}{\sqrt{5}}\underbrace{\left(\frac{1-\sqrt{5}}{2}\right)^k}_{<1}$$

negligible contribution for large k

$$F_{100} \approx \frac{1}{\sqrt{5}}(1.618)^{100}$$

$\boxed{\text{Lecture-20}}$ 10/9/19

Big facts → coming next

A is a real symmetric matrix

1. Eigenvalues of A are real
2. Eigenvecs corresponding to different eigenvalues are independent
3. A is orthogonally diagonalizable

i.e., $A = Q \Lambda Q^T$, where $Q^T Q = I$

$$A = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \begin{bmatrix} - x_1^T - \\ \vdots \\ - x_n^T - \end{bmatrix}$$

Example:

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

Eigenvalue $= -3, 2$

$$x_1 = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \qquad x_2 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$$

$$q_1 = \begin{bmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{bmatrix} \qquad q_2 = \begin{bmatrix} -2/\sqrt{5} \\ 1/\sqrt{5} \end{bmatrix}$$

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

Check: $Q^T Q = I$

$$Q \Lambda Q^T = \begin{bmatrix} 1/\sqrt{5} & -2/\sqrt{5} \\ 2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix} \begin{bmatrix} -3 & 0 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ -2/\sqrt{5} & 1/\sqrt{5} \end{bmatrix}$$

$$= A$$

Background on complex matrices:

$\mathbb{C}^n$: complex counterpart of $\mathbb{R}^n$

$(x_1, \ldots, x_n) \in \mathbb{C}^n$

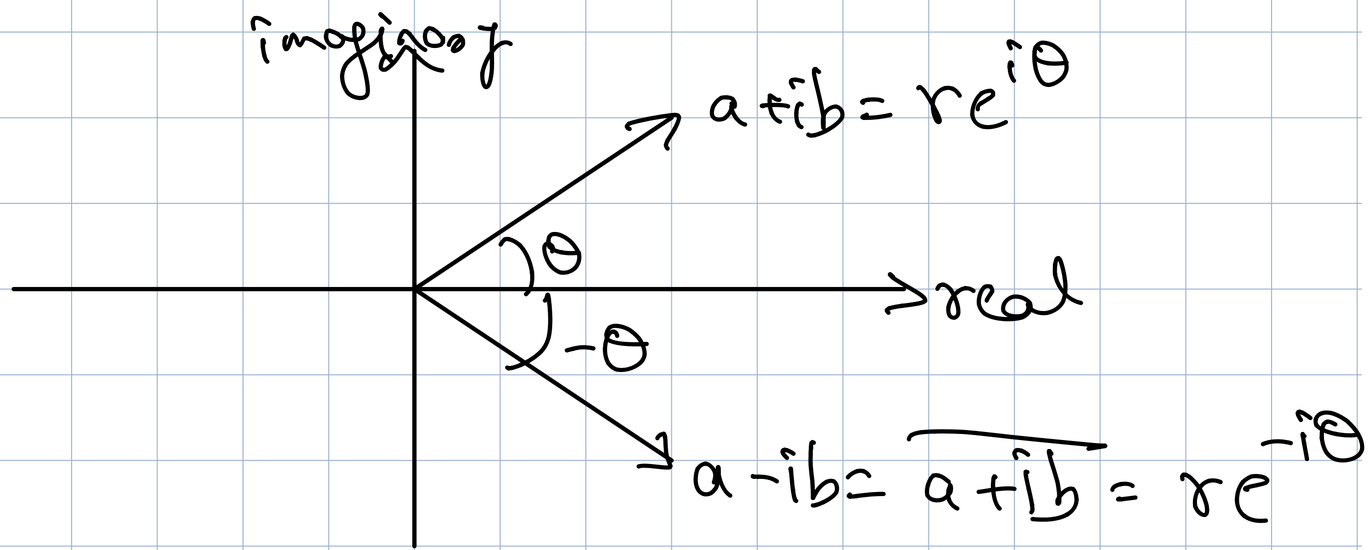
$x_i \to$ complex number

Addition & matrix multiplication as before:

$(a+ib) + (c+id) = (a+c) + i(b+d)$

$(a+ib)(c+id) = (ac-bd) + i(bc+ad)$

Complex conjugate of $(a+ib)$ is $(a-ib)$



$r^2 = a^2 + b^2 = |a+ib|$

$v_1, \ldots, v_k$ ← vectors in complex vector space $\mathbb{C}^n$

Linear combination: $c_1 v_1 + \ldots + c_k v_k = 0$

↑ complex numbers

Inner product & length:

In $\mathbb{R}^n$, $\|x\|^2 = x^T x$

Apply this definition to $\begin{bmatrix} 1 \\ i \end{bmatrix}$

$$\left\| \begin{bmatrix} 1 \\ i \end{bmatrix} \right\| = 0.$$

So, define the inner-product as

$$x \cdot y = \bar{x}^T y = \bar{x}_1 y_1 + \ldots + \bar{x}_n y_n$$

Note: $\bar{x}^T y \neq \bar{y}^T x$

Example: $x = \begin{bmatrix} 2-i \\ 1+i \end{bmatrix}$ $y = \begin{bmatrix} 6+8i \\ 9-13i \end{bmatrix}$

Check if $\quad \bar{x}^T y = \bar{y}^T x.$

Length of a complex vector:

$$\|x\|^2 = \bar{x}^T x$$

$$\left\| \begin{bmatrix} 1 \\ i \end{bmatrix} \right\| = \sqrt{2} \neq 0$$

$$\|x\| = 0 \quad \text{iff} \quad x = 0.$$

Check: (i) $\quad x \cdot y = \overline{y \cdot x}$

(ii) $\quad x \cdot (cy) = c\,(x \cdot y)$

(iii) $\quad (cx) \cdot y = \bar{c}\,(x \cdot y)$

$A^* =$ conjugate transpose of A

$$= \bar{A}^T = \overline{A^T}$$

$$A = \begin{bmatrix} 1+i & 3-2i \\ 2-4i & i \end{bmatrix} \qquad A^* = \begin{bmatrix} 1-i & 2+4i \\ 3+2i & -i \end{bmatrix}$$

Real matrix $A^* = A^T$

<u>Check:</u> (i) $(A^*)^* = A$

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

Real case

$(A^T)^T = A$

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

Inner-product:

$$x \cdot y = \bar{x}_1 y_1 + \cdots + \bar{x}_n y_n$$

$$= [\bar{x}_1 \cdots \bar{x}_n] \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$= x^* y$$

<u>Definition:</u> A is Hermitian if $A^* = A$

<u>Theorem:</u> A is Hermitian. Then,

(i) Eigenvalues of A are real

Proof: Suppose $Ax = \lambda x$, $x \neq 0, \lambda \neq 0$

$$(Ax)^* = (\lambda x)^* = \bar{\lambda} x^*$$

$$x^* A^* = \bar{\lambda} x^*$$

$$x^* A^* x = \bar{\lambda} x^* x$$

$$x^* A x = \bar{\lambda} x^* x$$

$$x^* \lambda x = \bar{\lambda} x^* x$$

$$\lambda x^* x = \bar{\lambda} x^* x$$

$\Rightarrow$ $\lambda = \bar{\lambda}$ since $x^* x \neq 0$

$\Rightarrow$ λ is real

Lecture - 21

12/9/19

(ii) Eigenvectors corresponding to different eigenvalues are orthogonal

Proof: $Ax = \lambda_1 x$, $Ay = \lambda_2 y$, $\lambda_1 \neq \lambda_2$

$$A = A^*$$

To show: $x \cdot y = 0$

$$x \cdot Ay = x \cdot \lambda_2 y = \lambda_2 (x \cdot y)$$

$$x \cdot Ay = x^* Ay = x^* A^* y = (Ax)^* y$$
$$= (\lambda_1 x)^* y = \bar{\lambda}_1 (x \cdot y) = \lambda_1 (x \cdot y)$$

So, $\lambda_2 (x \cdot y) = \lambda_1 (x \cdot y)$

$$\lambda_1 \neq \lambda_2 \Rightarrow x \cdot y = 0$$

$\blacksquare$

Remark:- If no eigenvalue is repeated for a real symmetric matrix, then we get n independent eigenvectors & hence,

$$A = Q \Lambda Q^T$$

Unitary matrices:

A matrix is unitary if it is square, and has orthonormal columns.

Real case: $Q^T Q = I \iff Q$ is orthogonal

Complex case: $U^* U = I \iff U$ is unitary

$$U^* = U^{-1}$$

Example: $U = \begin{bmatrix} 1/\sqrt{2} & i/\sqrt{2} \\ i/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}$

Check: U is unitary

Properties of unitary matrices:

(i) Length unchanged!

$$Ux \cdot Uy = (Ux)^* Uy$$
$$= x^* U^* U y$$
$$= x^* y = x \cdot y$$

$$\Rightarrow \quad \|Ux\|^2 = \|x\|^2$$

(ii) Eigenvalues of U satisfy
$$|\lambda| = 1$$

$$Ux = \lambda x, \quad x \neq 0$$

$$\|Ux\| = \|x\|$$

$$\|\lambda x\| = |\lambda| \, \|x\| = \|x\|$$

$$\Rightarrow |\lambda| = 1$$

(Or) $Ux \cdot Ux = \lambda x \cdot \lambda x$

$$= \lambda \bar{\lambda} \, (x \cdot x)$$

Using $\|Ux\| = \|x\|$, we get

$$Ux \cdot Ux = x \cdot x = \lambda \bar{\lambda} \, (x \cdot x)$$

$$\Rightarrow \lambda \bar{\lambda} = 1 \Rightarrow |\lambda| = 1$$

(iii) Eigenvectors corresponding to different eigenvalues are orthogonal.

Pf: $Ux = \lambda_1 x$, $Uy = \lambda_2 y$

$$x \cdot y = Ux \cdot Uy$$

$$= (\lambda_1 x) \cdot (\lambda_2 y)$$

$$= \bar{\lambda}_1 \lambda_2 \, (x \cdot y)$$

$$(\bar{\lambda}_1 \lambda_2 - 1)(x \cdot y) = 0$$

Suppose $\bar{\lambda}_1 \lambda_2 = 1$

$$\lambda_1 \bar{\lambda}_1 \lambda_2 = \lambda_1$$

Usig $\lambda_1 \bar{\lambda}_1 = 1$, $\lambda_2 = \lambda_1$, a contradiction.

So, $(x \cdot y) = 0$ $\blacksquare$

Next stop:

For a Hermitian matrix A,

one can find a unitary U s.t.

$$A = U \Lambda U^*$$

Lecture-22 16/9/19

Recap: $A^* = \bar{A}^T$

"A" Hermitian if $A^* = A$

"U" unitary if $U^*U = I$ & U is square

Schur's theorem:

Any $n \times n$ matrix is similar to a upper-triangular matrix via a unitary matrix, i.e.,

$$U^* A U = T$$

upper triangular

Proof of Schur's theorem for a 3×3 matrix A.

Let $p(\lambda)$ be the characteristic polynomial of A, and λ_1 be a root.

Let z_1 be the corresponding eigenvector.

Extend z_1 to a basis, and make it orthonormal.

Let $\{z_1, u, v\}$ be the orthonormal basis.

$$U_1 = \begin{bmatrix} | & | & | \\ z_1 & u & v \\ | & | & | \end{bmatrix}$$

$$AU_1 = A\begin{bmatrix} | & | & | \\ z_1 & u & v \\ | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} | & | & | \\ \lambda_1 z_1 & Au & Av \\ | & | & | \end{bmatrix}$$

$$U_1^* A U_1 = \begin{bmatrix} - & \bar{z}_1 & - \\ - & \bar{u} & - \\ - & \bar{v} & - \end{bmatrix} \begin{bmatrix} | & | & | \\ \lambda_1 z_1 & Au & Av \\ | & | & | \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & * & * \\ 0 & & \\ 0 & & B \end{bmatrix}$$

$$\bar{z}_1^T \lambda_1 z_1 = \lambda_1 \bar{z}_1^T z_1 = \lambda_1$$

$$\bar{u}^T z_1 = 0, \quad \bar{v}^T z_1 = 0$$

Repeat the procedure with B to get eigenvalue λ_2 of B & a unitary P s.t.

$$P^* B P = \begin{bmatrix} \lambda_2 & * \\ 0 & \lambda_3 \end{bmatrix}$$

Let $U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & & P \end{bmatrix}$

$$U_2^* = \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & & P^* \end{bmatrix}$$

Check: U_2 is unitary

$$U_2^*(U_1^* A U_1) U_2$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{P^*} \end{bmatrix} \begin{bmatrix} \lambda_1 & * & * \\ 0 & & \\ 0 & \multicolumn{2}{c}{B} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & \multicolumn{2}{c}{P} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & * & * \\ 0 & & \\ 0 & \multicolumn{2}{c}{P^* B P} \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & * \\ 0 & 0 & \lambda_3 \end{bmatrix} =: T$$

Set $U = U_1 U_2$

$$U^* = U_2^* U_1^* = U_2^{-1} U_1^{-1}$$

$$= (U_1 U_2)^{-1} = U^{-1}$$

$$U^* A U = T$$

$\blacksquare$

<u>EXAMPLE</u>

$$A = \begin{bmatrix} 5 & 8 & 16 \\ 5 & 0 & 9 \\ -3 & -5 & -10 \end{bmatrix}$$

$$p(\lambda) = -(\lambda - 1)(\lambda + 3)^2$$

$$\lambda_1 = 1, \quad \lambda_2 = -3$$

$$z_1 = \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix}$$

$$\{z_1, e_1, e_2\} \xrightarrow{\text{Gram-Schmidt}}$$

$$U_1 = \begin{bmatrix} -2/\sqrt{6} & 1/\sqrt{3} & 0 \\ -1/\sqrt{6} & -1/\sqrt{3} & 1/\sqrt{2} \\ 1/\sqrt{6} & 1/\sqrt{3} & 1/\sqrt{2} \end{bmatrix}$$

$$U_1^* A U_1 = \begin{bmatrix} 1 & -8\sqrt{2} & -12\sqrt{3} \\ 0 & -3 & 0 \\ 0 & \sqrt{6} & -3 \end{bmatrix}$$

B

$\lambda_2 = -3 \qquad e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$ eiginvec

$\{e_2, e_1\}$ orthonormal basis

$$P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$P^* B P = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} -3 & 0 \\ \sqrt{6} & -3 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$= \begin{bmatrix} -3 & \sqrt{6} \\ 0 & -3 \end{bmatrix}$$

$$U_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & & \\ 0 & & P \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$U_2^* U_1^* A U_1 U_2$$

$$= \begin{bmatrix} 1 & -8\sqrt{2} & -12\sqrt{3} \\ 0 & -3 & \sqrt{6} \\ 0 & 0 & -3 \end{bmatrix}$$

(A) <u>Theorem</u>: Any Hermitian matrix is unitarily diagonalizable. Moreover,

$$U^{-1} A U = D$$

$\hookrightarrow$ diagonal matrix with real numbers.

Proof:-

From Schur's theorem,

$$U^* A U = T$$

$$T^* = U^* A^* U$$

$$= U^* A U \text{, since } A \text{ is Hermitian}$$

$$= T$$

T is upper triangular, &

T^* is lower —"—.

$T = T^* \Rightarrow T$ is diagonal.

Lecture - 23

17/9/19

Given : $A^* = A$ and

$$U^* A U = D$$

To prove : The diagonal entries are real.

<u>Proof</u>:- We know $T^* = T$

$$(T^*)_{ii} = \overline{T_{ii}}$$

$$= T_{ii}$$

So, T_{ii} is real

$\blacksquare$

EXAMPLE

$$A = \begin{bmatrix} 2 & 1-i \\ 1+i & 3 \end{bmatrix}$$

$$p(\lambda) = (\lambda-1)(\lambda-4)$$

$$z_1 = \begin{bmatrix} -1+i \\ 1 \end{bmatrix} \qquad z_2 = \begin{bmatrix} 1-i \\ 2 \end{bmatrix}$$

$$z_1 \cdot z_2 = 0$$

$$u_1 = \begin{bmatrix} -1+i/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \qquad u_2 = \begin{bmatrix} 1-i/\sqrt{6} \\ 2/\sqrt{6} \end{bmatrix}$$

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix}$$

$$U^* A U = D = \begin{bmatrix} 1 & 0 \\ 0 & 4 \end{bmatrix}$$

Corollary:

A real symmetric matrix A is orthogonally diagonalizable, i.e., $\exists\, Q$ s.t.

$$Q^T A Q = D, \qquad Q^T Q = I$$

Pf: From Theorem Ⓐ, we get

$$U^{-1} A U = D$$

Cols. of U = eigenvectors of A

A is real symmetric,

so λ's are real

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

$$\downarrow$$

solution to this is a real "x"

$\Rightarrow$ U is real

$\Rightarrow$ $U^T = \overline{U^T} = U^* = U^{-1}$

$\blacksquare$

Remark:

Hermitian $\Rightarrow$ unitarily diagonalizable

$$\not\Leftarrow$$

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

$$A^* \neq A$$

$$\lambda_1 = i, \quad \lambda_2 = -i$$

$$z_1 = \begin{bmatrix} 1 \\ -i \end{bmatrix} \quad z_2 = \begin{bmatrix} 1 \\ i \end{bmatrix}$$

$$u_1 = \begin{bmatrix} 1/\sqrt{2} \\ -i/\sqrt{2} \end{bmatrix} \quad u_2 = \begin{bmatrix} 1/\sqrt{2} \\ i/\sqrt{2} \end{bmatrix}$$

$$U = \begin{bmatrix} u_1 & u_2 \end{bmatrix} \quad U^* A U = D$$

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

Fact: A matrix, say "A", is unitarily diagonalizable if and only if

$$AA^* = A^*A \quad \text{——— } (*)$$

Matrices that satisfy $(*)$ are "normal".

So, a Hermitian matrix is normal, and a Unitary ——— " ——— " ———

Lecture-24 23/9/19

Singular value decomposition

Every matrix cannot be diagonalized.

But, any "real" $m \times n$ matrix A can be written as

$$A = Q_1 \Sigma Q_2^T,$$

where $Q_1 \to m \times m$ matrix

$Q_2 \to n \times n$ matrix

Q_1, Q_2 orthogonal

$$\Sigma = \begin{bmatrix} D & 0 \\ 0 & 0 \end{bmatrix},$$ where

$$D = \begin{bmatrix} \sigma_1 & & & 0 \\ & \ddots & & \\ & & \ddots & \\ 0 & & & \sigma_r \end{bmatrix}, \quad \sigma_i > 0$$

Why does this decomposition hold for any matrix A?

A is $m \times n$

$A^T A$ is $n \times n$

$A^T A$ is symmetric and real

$\Rightarrow \exists$ a basis of orthonormal eigenvectors, say $\{x_1, \dots, x_n\}$

corresponding to eigenvalues $\{\lambda_1 \sim \lambda_n\}$

$$A^T A x_i = \lambda_i x_i, \quad i = 1 \sim n$$

$$\|x_i\|^2 = 1, \quad x_i \cdot x_j = 0, \; i \neq j$$

So,

$$(A^T A x_i) \cdot x_i = (\lambda_i x_i) \cdot x_i$$

$$= \lambda_i$$

$$(A^T A x_i) \cdot x_i = (A^T A x_i)^T x_i$$

$$= x_i^T A^T A x_i$$

$$= (A x_i)^T A x_i$$

$$= \|A x_i\|^2$$

$$\geq 0$$

So, $\lambda_i \geq 0$

Next Stop: $\sigma_i = \sqrt{\lambda_i}$

<u>Order the λ's:</u>

$\lambda_1, \cdots \lambda_r, \lambda_{r+1}, \cdots \lambda_n$

$\lambda_1 > 0 \cdots \lambda_r > 0, \lambda_{r+1} = \cdots = \lambda_n = 0$

Let $\sigma_i = \sqrt{\lambda_i}$ and

$$y_i = \frac{1}{\sigma_i} A x_i, \quad i = 1 \cdots r$$

$(y_i \in \mathbb{R}^m)$

$$\|y_i\| = \frac{1}{\sigma_i} \|A x_i\| = \frac{\sqrt{\lambda_i}}{\sigma_i} = 1$$

$$y_i \cdot y_j = \frac{1}{\sigma_i \sigma_j} Ax_i \cdot Ax_j$$

$$= \frac{1}{\sigma_i \sigma_j} x_i^T A^T A x_j$$

$$= \frac{1}{\sigma_i \sigma_j} x_i^T \lambda_j x_j$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j} x_i^T x_j = 0$$

So, $\{y_1 \ldots y_r\}$ is a orthonormal set of vectors

Extend this set to a basis & make it orthonormal

Let $\{y_1 \cdots y_m\}$ be that set.

$$Q_1 = \begin{bmatrix} | & & | \\ y_1 & \cdots & y_m \\ | & & | \end{bmatrix}$$

$$Q_2 = \begin{bmatrix} | & & | \\ x_1 & \cdots & x_n \\ | & & | \end{bmatrix}$$

$$\Sigma = Q_1^T A Q_2$$

$$= \begin{bmatrix} - & y_1^T & - \\ & \vdots & \\ - & y_m^T & - \end{bmatrix} \begin{bmatrix} | & & | \\ Ax_1 & \cdots & Ax_n \\ | & & | \end{bmatrix}$$

$$(\Sigma)_{ij} = y_i^T (A x_j)$$

If $j \le r$, $\quad y_j = \frac{1}{\sigma_j} A x_j$

$$\begin{aligned} y_i^T (A x_j) &= y_i^T (\sigma_j y_j) \\ &= \sigma_j \, y_i^T y_j \\ &= \begin{cases} \sigma_j & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \end{aligned}$$

If $j > r$, then (see next page)

$$\|Ax_j\|^2 = \lambda_j = 0$$

So, $Ax_j = 0$

$$\Rightarrow \quad y_i^T A x_j = 0$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & 0 & 0 \\ & \ddots & & \\ 0 & & \sigma_r & \\ & 0 & & 0 \end{bmatrix}$$

$$\Sigma = Q_1^T A Q_2$$

$$\Leftrightarrow \quad Q_1 \Sigma Q_2^T = A \quad \blacksquare$$

Remark:

$$AA^T = Q_1 \Sigma \Sigma^T Q_1^T$$

So, eigenvectors of AA^T go into Q_1

and,

$$A^T A = Q_2 \Sigma^T \Sigma Q_2^T$$

implying eigenvectors of $A^T A$ go into Q_2.

EXAMPLE:

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

Find SVD of A.

Is A diagonalizable?

No. A has $\sqrt{2}$ eigenvalue repeated twice.

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

So, $\begin{bmatrix} 1 \\ 0 \end{bmatrix}$ is an eigenvector of A

& there aren't any more independent ones.

Finding the SVD

(see next page)

$$A^T A = \begin{bmatrix} 2 & \sqrt{2} \\ \sqrt{2} & 3 \end{bmatrix}$$

Eigenvalues: 4, 1

$$\sigma_1 = 2, \quad \sigma_2 = 1$$

$$\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$$

Finding eigenvectors of $A^T A$:

$$A^T A - 4I = \begin{bmatrix} -2 & \sqrt{2} \\ \sqrt{2} & -1 \end{bmatrix}$$

So, $\begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$ is an eigenvector of $A^T A$

$$x_1 = \frac{1}{\sqrt{3}} \begin{bmatrix} 1 \\ \sqrt{2} \end{bmatrix}$$

$$A^T A - I = \begin{bmatrix} \text{Fill} \\ \text{this} \end{bmatrix}$$

So, $\begin{bmatrix} \sqrt{2} \\ -1 \end{bmatrix}$ is another eigenvector of $A^T A$

$$x_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} \\ -1 \end{bmatrix}$$

So, $Q_2 = \frac{1}{\sqrt{3}} \begin{bmatrix} \sqrt{2} & 1 \\ 1 & -\sqrt{2} \end{bmatrix}$

Using,

$$\sigma_1 y_1 = A x_1 \ \& \ \sigma_2 y_2 = A x_2,$$

we obtain

$$y_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} \sqrt{2} \\ 1 \end{bmatrix}, \quad y_2 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 \\ -\sqrt{2} \end{bmatrix}$$

$$Q_1 = \begin{bmatrix} | & | \\ y_1 & y_2 \\ | & | \end{bmatrix}$$

Check, $A = Q_1 \Sigma Q_2^T$

Lecture 25 - Positive definiteness

Consider

$$F(x,y) = 7 + 2(x+y)^2 - y\sin y - x^3$$

$$f(x,y) = 2x^2 + 4xy + y^2$$

At a stationary point, the first derivatives vanish

$$\frac{\partial F}{\partial x} = 4(x+y) - 3x^2 = 0$$

$$\frac{\partial F}{\partial y} = 4(x+y) - y\cos y - \sin y = 0$$

$$\frac{\partial f}{\partial x} = 4x + 4y = 0$$

$$\frac{\partial f}{\partial y} = 4x + 2y = 0$$

So $(x,y)=(0,0)$ is a "stationary point" for both F and f.

Question: Whether $(0,0)$ is a minima/maxima/saddle?

Answer: Find the second derivatives at $(0,0)$

$$\frac{\partial^2 F}{\partial x^2} = 4 = \frac{\partial^2 f}{\partial x^2}$$

$$\frac{\partial^2 F}{\partial x \partial y} = 4 = \frac{\partial^2 f}{\partial x \partial y}$$

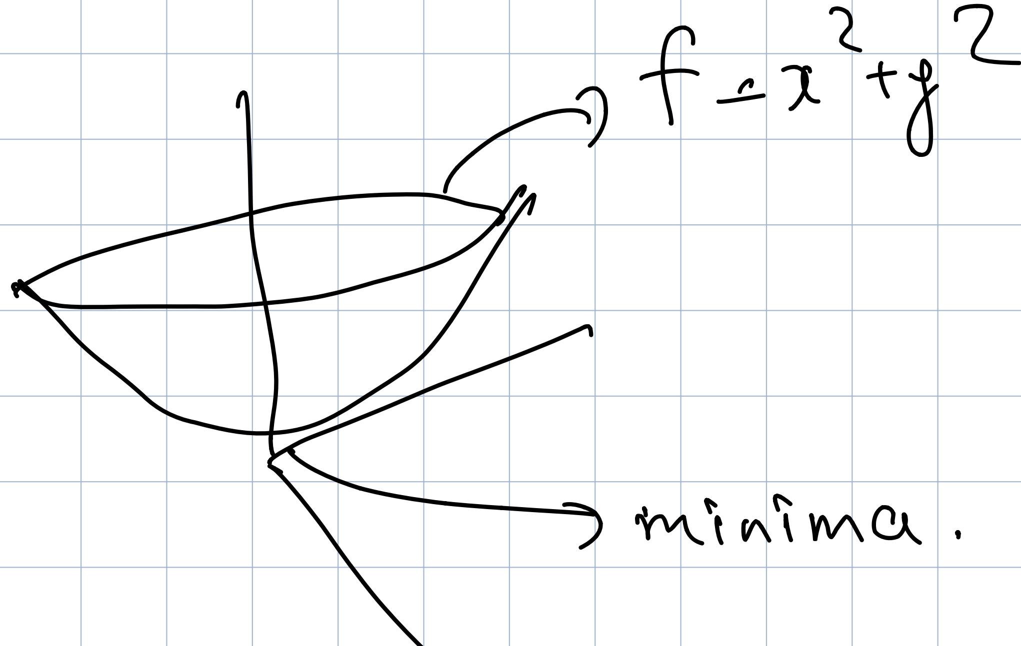
$$\frac{\partial^2 F}{\partial y^2} = 2 = \frac{\partial^2 f}{\partial y^2}$$

So, F and f behave identically near origin, i.e.,

F has a minimum iff f has a minimum

Remark:

Every quadratic $ax^2 + 2bxy + cy^2$ has a stationary point at $(0,0)$.



A function f that vanishes at $(0,0)$ & is strictly positive at other points is "positive definite".

Question: What conditions on a, b, c ensure $f = ax^2 + 2bxy + cy^2$ is p.d.?

Necessary conditions:

(I) If $f > 0$, then $a > 0$ [Look at $(1,0)$]

(II) If $f > 0$, then $c > 0$ [Look at f value at $(0,1)$]

But $a > 0$, $c > 0$ is not enough to ensure $f > 0$

e.g. $f = x^2 - 10xy + y^2$

Trick (Complete the square):

$$f = ax^2 + 2bxy + cy^2$$

$$= a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

(III) If $f > 0$, then

$$ac > b^2$$

So, $f = ax^2 + 2bxy + cy^2$ is p.d.

if and only if

$a > 0$ and $ac > b^2$

See next page

Any function $F(x,y)$ has a minimum at a point (x,y)

where $\dfrac{\partial F}{\partial x} = 0 = \dfrac{\partial F}{\partial y}$

$$\frac{\partial^2 F}{\partial x^2} > 0 \quad \text{and}$$

$$\left(\frac{\partial^2 F}{\partial x^2}\right)\left(\frac{\partial^2 F}{\partial y^2}\right) > \left(\frac{\partial^2 F}{\partial x \partial y}\right)^2$$

Note! Quadratic part of F:

$$\frac{x^2}{2} \frac{\partial^2 F}{\partial x^2}(\alpha,\beta) + xy \frac{\partial^2 F}{\partial x \partial y}(\alpha,\beta) + \frac{y^2}{2} \frac{\partial^2 F}{\partial y^2}(\alpha,\beta)$$

Remarks:-

1. If $ac = b^2$, then

f is positive semidefinite if $a > 0$ &

negative semi-definite if $a < 0$

2. Saddle point if $ac < b^2$

Connection to linear algebra:

$$ax^2 + 2bxy + cy^2 = [x \;\; y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$v = [x \ \ y]^T$

$ax^2 + 2bxy + cy^2$ is $v^T A v$

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

$v^T A v$ in $\mathbb{R}^n$

$$[x_1 \ - - \ x_n] \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$$

$v = (x_1, \dots x_n)^T$

$$f(v) = a_{11} x_1^2 + 2a_{12} x_1 x_2 + \dots + a_{nn} x_n^2$$

At $v = (0, \dots 0)$, $f = 0$

So, $(0, \dots 0)$ is a stationary point

Next question is to check if f has a minima/maxima/saddle at origin.

Examples:-

(1) $f = 2x^2 + 4xy + y^2$

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

saddle at origin

(2) $f = 2xy$

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

$\Big\}$ Saddle

(3) A is 3×3

$$f = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2$$

$$f = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix} \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & -1 & 2 \end{pmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

Check that f has a minimum at origin

Def: Matrix A is p.d.

if $v^T A v > 0 \quad \forall v \in \mathbb{R}^n$.

Test for positive-definiteness

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$ is p.d. if

$a > 0, \; ac - b^2 > 0$

$\Rightarrow$ both eigenvalues are > 0.

Fact:-

Each of the following tests are both necessary & sufficient for p.d.

① $v^T A v > 0 \quad \forall v$

② All eigenvalues of A are > 0

③ All the pivots (without row exchanges) are > 0.

Lecture-26

$$f(x,y) = ax^2 + 2bxy + cy^2$$

$$= a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

f is p.d if $f > 0$
$\forall\ (x,y) \neq (0,0)$ & 0 at $(0,0)$

f is p.d

$$\Updownarrow$$

$a > 0$ and $ac > b^2$

Recall the example
$$f(x,y) = 2x^2 + 4xy + y^2$$

This function has a saddle at origin because $ac < b^2$

brief remark on saddle:-

$f_1 = 2xy$ and $f_2 = x^2 - y^2$

$ac - b^2 = -1$ for both

Connection to linear algebra:-

$v^T A v$ in $\mathbb{R}^2$ $\quad ax^2 + 2bxy + cy^2$

$$= \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

$v^T A v$ in $\mathbb{R}^n$ $\quad \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j$

"real-symmetric"

A matrix A is positive definite

if (i) $v^T A v > 0 \quad \forall v \in \mathbb{R}^n, v \neq 0$.

or

(ii) All eigenvalues of A are > 0

or

(iii) All upper left submatrices have positive determinant

or

(iv) All the pivots are > 0

Suppose (i) holds

$$Ax = \lambda x$$

$$x^T A x = x^T \lambda x = \lambda \|x\|^2$$

$x \neq 0$, $x^T A x > 0 \Rightarrow$

$\lambda > 0$ i.e., (ii) holds.

Suppose (ii) holds.

Using spectral theorem, we obtain a orthonormal basis of eigenvectors.

Let $\{x_1 \, -- \, x_n\}$ be that basis

Any x, $\quad x = c_1 x_1 + \cdots + c_n x_n$

$$Ax = c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n$$

$$x^T A x = (c_1 x_1^T + \cdots + c_n x_n^T)$$
$$\times (c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n)$$
$$= c_1^2 \lambda_1 + \cdots + c_n^2 \lambda_n$$
$$> 0 \text{ since } \lambda_i > 0 \ \forall i$$

So, (i) holds.

Brief notes on (iii) & (iv)

In 2×2 case, $\quad ac > b^2$
$$\Leftrightarrow \det(A) > 0$$

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

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

$ac > b^2$ but

$A = -I$ not p.d.

infact, A is negative definite

Condition (iii) requires

$A_1 = [a_{11}]$ $A_2 = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

$$A_3 = \begin{bmatrix} a_{11} & & \\ & \ddots & \\ & & a_{33} \end{bmatrix} \cdots A_n = A$$

$$\det(A_i) > 0 \quad i = 1 \ldots n$$

On condition (iv)

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

LDL^T

$$= \begin{bmatrix} 1 & 0 & 0 \\ -1/2 & 1 & 0 \\ 0 & -2/3 & 1 \end{bmatrix} \begin{bmatrix} 2 & & 0 \\ & 3/2 & \\ 0 & & 4/3 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -\frac{2}{3} \\ 0 & 0 & 1 \end{bmatrix}$$

If $x = \begin{bmatrix} u \\ v \\ w \end{bmatrix}$, then

$$L^T x = \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & -2/3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \end{bmatrix}$$

$$= \begin{bmatrix} u - v/2 \\ v - 2w/3 \\ w \end{bmatrix}$$

$$x^T A x = x^T L D L^T x$$

$$= (L^T x)^T D \, L^T x$$

$$= 2\left(u - \frac{v}{2}\right)^2 + \frac{3}{2}\left(v - \frac{2w}{3}\right)^2 + \frac{4}{3} w^2 > 0$$

So, in this example, (iv) $\Rightarrow$ (i)

Check:

$$ax^2 + 2bxy + cy^2$$

$$= a\left(x + \frac{b}{a}y\right)^2 + \frac{ac-b^2}{a}y^2$$

$a, \ \frac{ac-b^2}{a}$ are the pivots

for the 2×2 matrix A.