Chapter 6: Eigenvalues

Ex: $A=\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$ $\quad A: \mathbb{R}^n \to \mathbb{R}^2$

square only $\quad$ Linear operator

Let $x=\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $\quad Ax=y$

$y = Ax = \begin{bmatrix} 4 \\ 4 \end{bmatrix} = 4x$

$x \to 4x$

Given $A_{m\times n}$ Find $x \in \mathbb{R}^n$

$Ax = \alpha x$

① Def

Definition Let A be an $n \times n$ matrix. A scalar λ is said to be an **eigenvalue** or a **characteristic value** of A if there exists a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$. The vector $\mathbf{x}$ is said to be an **eigenvector** or a **characteristic vector** belonging to λ.

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

If (λ, x) are eigen pair, then $\det(A-\lambda I)=0$.

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

$A-\lambda I = \begin{bmatrix} 3-\lambda & 1 \\ 1 & 3-\lambda \end{bmatrix}$

$\det(A-\lambda I) = (3-\lambda)^2 - 1 = 0$

$3-\lambda = \pm 1, \quad \lambda = 4 \text{ or } 2$

For eigen vectors of $\lambda_1 = 4$

$(A-4I)x=0$

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

For $\lambda_2 = 2$

$(A-2I)x=0$

$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} x = 0 \Rightarrow x = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$

② Calculation

$A_{n\times n}$ λ is a scalar. The following statement are equivalent

(i) λ is a eigenvalue of A

(ii) $(A-\lambda I)x=0$ has nonzero sol

(iii) $N(A-\lambda I) \neq \{0\}$

(iv) $A-\lambda I$ is singular

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

How to find λ?

$\det(A-\lambda I)$

polynomial of degree n

$$\det \begin{bmatrix} a_{11}-\lambda & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn}-\lambda \end{bmatrix}$$

$= (a_{11}-\lambda) M_{11} + a_{12}(-1)^{1+2} M_{12} + \cdots \; a_{1n}(-1)^{1+n} M_{1n}$

$=$ polynomial of degree n by math induction

$\det(A-\lambda x) = p(x)$: characteristic polynomial

$p(x)=0$

$p(x) = \alpha(x-\lambda_1)(x-\lambda_2)\cdots(x-\lambda_n)$

for $u \in \mathbb{R}^n$, $A = uu^T_{n \times n}$

if $u \perp v$,

$Av = uu^T v = 0$

$Au = uu^T u = \underbrace{(u^T u)}_{\lambda} u$

$A = \begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$ $A_i : k \times k$

$\det(A - \lambda I) = \det \begin{bmatrix} A_1 - \lambda I_k & 0 \\ 0 & A_2 - \lambda I_k \end{bmatrix}$

$= \det(A_1 - \lambda I_k) \cdot \det(A_2 - \lambda I_k) = 0$

If $A_1 u_1 = \lambda_1 u_1$

$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} \begin{bmatrix} u_1 \\ 0 \end{bmatrix} = \lambda \begin{bmatrix} u_1 \\ 0 \end{bmatrix}$

eigenvalue of A

Corresponding eigenvectors

③ properites

$A_{n\times n}$, $\lambda_1 , \lambda_2 \cdots \lambda_n$ are eigenvalues.

(i) $\lambda_1 \cdot \lambda_2 \cdots \lambda_n = \det(A) = p(0)$

(ii) $\lambda_1 + \lambda_2 + \cdots \lambda_n = \sum_{i=1}^{n} a_{ii}$

(iii)

Theorem 6.1.1 *Let A and B be n × n matrices. If B is similar to A, then the two matrices have the same characteristic polynomial and, consequently, the same eigenvalues.*

proof of (i): by math induction

$B(\lambda)$: $n\times n$ matrix

no more that n entries are given by $\alpha + \beta\lambda$

other entries are independent λ

Then

$$\det(B) = P_k(\lambda), \quad k \le n$$

proof of (ii): $p(\lambda) = \det(A - \lambda I) = \det \begin{bmatrix} a_{11}-\lambda & \cdots & & a_{1n} \\ a_{21} & a_{22}-\lambda & & \\ \vdots & & \ddots & \\ a_{n1} & \cdots & & a_{nn}-\lambda \end{bmatrix}$

$$= (a_{11}-\lambda)(-1)^{1+1}M_{11} + \underbrace{a_{12}(-1)^{1+2}M_{12} + \cdots}_{P_{k-1}(\lambda)} a_{1n}(-1)^{1+n}M_{1n}$$

⊕ Complex eigenvalues

$A \in \mathbb{R}^{n\times n}$

$\downarrow$

$P(\lambda)$ with real coefficents

Ex $A = \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix} \Rightarrow \lambda_{1,2} = 1 \pm 2i$

$A \in \mathbb{C}^{n\times n}$

Ex $A = \begin{bmatrix} 1 & i \\ i & 1 \end{bmatrix}$

$\det(A-\lambda I) = \det\begin{bmatrix} 1-\lambda & i \\ i & 1-\lambda \end{bmatrix} = (1-\lambda)^2 + 1 = 0$

$(1-\lambda)^2 = -1 \qquad 1-\lambda = \pm i \qquad \lambda_{1,2} = 1 \pm i$

$Ax = \lambda x, \quad AAx = \lambda Ax \qquad A^2 x = \lambda^2 x$

$f(x) = b_0 + b_1 x + b_2 x^2$

$f(A) = b_0 + b_1 A + b_2 A^2$

$f(A)x = (b_0 + b_1 A + b_2 A^2)x = b_0 + b_1 Ax + b_2 A^2 x$

$= (b_0 + b_1\lambda + b_2\lambda^2)x = f(\lambda)x$

6.3 Diagonalization

Ex $A = \begin{bmatrix} 3 & 2 \\ 3 & -2 \end{bmatrix}$

eigenvalues: $\lambda_1 = 4$ $\quad$ $\lambda_2 = -3$

$x_1 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ $\quad$ $x_2 = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$

$$X = (x_1, x_2)$$

$$AX = (Ax_1, Ax_2) = (\lambda_1 x_1, \lambda_2 x_2) = (x_1, x_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}$$

$$= X\Lambda \qquad \Lambda: \text{diagonal}$$

$$AX = X\Lambda$$

for the example X: nonsingular

$$A = X\Lambda X^{-1}$$

or

$$X^{-1}AX = \Lambda$$

A is similar to a diagonal matrix

<i> diagonalizable

<ii> defective

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

$\lambda_1 = \lambda_2 = 1$ $\qquad$ $(A - \lambda I) = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

$x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

① Definition

Definition An $n \times n$ matrix A is said to be **diagonalizable** if there exists a nonsingular matrix X and a diagonal matrix D such that

$$X^{-1}AX = D$$

We say that X **diagonalizes** A.

② Theorem

Theorem 6.3.1 *If $\lambda_1, \lambda_2, \ldots, \lambda_k$ are distinct eigenvalues of an $n \times n$ matrix A with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$, then $\mathbf{x}_1, \ldots, \mathbf{x}_k$ are linearly independent.*

proof: Assume that

$C_1 x_1 + C_2 x_2 + \cdots C_k x_k = 0$

$A(C_1 x_1 + C_2 x_2 + \cdots C_k x_k) = 0$

$C_1 A x_1 + C_2 A x_2 + \cdots C_k A x_k = 0$

$C_1 \lambda_1 x + C_2 \lambda x_2 + \cdots C_k \lambda_k x_k = 0$ *① !

$\lambda_k (C_1 x_1 + \cdots C_k x_k) = 0$

$C_1 \lambda_k x_1 + \cdots C_k \lambda_k x_k = 0$ *②

①-② $C_1(\lambda_1 - \lambda_k) X + \cdots C_{k-1} (\lambda_{k-1} - \lambda_k) x_k = 0$

if $x_1, x_2 \cdots x_{k-1}$ are linearly independent

then $C_1 = \cdots = C_{k-1} = 0$

Theorem 6.3.2 *An $n \times n$ matrix A is diagonalizable if and only if A has n linearly independent eigenvectors.*

proof: From theorem 6.3.1

if $k = n$, there are n distinct eigenvectors

$x_1 \cdots x_n$

let $X = (x_1 \cdots x_n)$

$X^{-1} A X = \Lambda$ $\quad\quad \Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix}$

A: diagonalizable

$X^{-1} A^k X = \Lambda^k$

② Jordan form

For any $A_{n \times n}$ matrix, there is a nonsingular matrix X such that

$$X^{-1}AX = J$$

$$J = \begin{bmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{bmatrix} \qquad J_i = \begin{bmatrix} \lambda_j & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_j \end{bmatrix}_{k_j \times k_j}$$

Ex: $$J = \left[\begin{array}{c|cc} \lambda_1 & & \\ \hline & \lambda_2 & 1 \\ & & \lambda_2 \end{array}\right]$$

6.4 Hermitian Matrices

① Definition

Definition A matrix M is said to be **Hermitian** if $M = M^H$.

Hermitian transpose

$A \in \mathbb{C}^{m\times n}$

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

if $A \in \mathbb{R}^{m\times n}$, $A^H = A^T$

Inner product space in $\mathbb{C}$

$x, y \in \mathbb{C}^n$ $\quad \langle x, y\rangle = x^H y$

$\|x\|^2 = \langle x, x\rangle = x^H x$

② Eigenvalues and Eigenvectors of Hermitian matrix

Theorem 6.4.1 *The eigenvalues of a Hermitian matrix are all real. Furthermore, eigenvectors belonging to distinct eigenvalues are orthogonal.*

special case: $A \in \mathbb{R}^{n\times n}$, symmetric

proof: let $Au_1 = \lambda u_1$ where $\langle \lambda_1, u_1\rangle$ are eigenpair of A

$x_1 = \alpha + \beta i$

$u_1^H A u_1 = u_1^H \lambda_1 u_1$ $\qquad$ $u_1^H A u_1 = \bar{\lambda}_1 u_1^H u_1$

$(Au_1)^H = (\lambda_1 u_1)^H$ $\qquad$ $\lambda u_1^H u_1 = \bar{\lambda}_1 u_1^H u_1$

$u_1^H A^H = \bar{\lambda}_1 u_1^H$ $\qquad$ $(\lambda_1 - \bar{\lambda}_1) u_1^H u_1 = 0$

$\lambda_1 = \bar{\lambda}_1 \Rightarrow \lambda_1$ is real number

(ii) $Ax_1 = \lambda_1 x_1$ $\quad Ax_2 = \lambda x_2$, $\lambda_1 \neq \lambda_2$: real

$$\begin{cases} x_2^H A x_1 = \lambda_1 x_2^H x_1 \\ x_2^H A x_1 = \lambda_2 x_2^H x_1 \end{cases}$$

$\Rightarrow (\lambda_1 - \lambda_2) x_2^H x_1 = 0$ since $\lambda_1 \neq \lambda_2$, so $x_2^H x_1 = 0$

Definition An $n \times n$ matrix U is said to be **unitary** if its column vectors form an orthonormal set in $\mathbb{C}^n$.

$$U^H U = I$$

Corollary 6.4.2 *If the eigenvalues of a Hermitian matrix A are distinct, then there exists a unitary matrix U that diagonalizes A.*

$$Q^T A Q = D$$

③ Schur's theorem

Theorem 6.4.3 Schur's Theorem

For each $n \times n$ matrix A, there exists a unitary matrix U such that $U^H AU$ is upper triangular. (T)

proof: If it's true for Hermitian matrix A

$$U^H A U = T$$

$$(U^H A U)^H = T^H$$

$$U^H A^H U = T^H$$

since $A^H = A$, $U^H A U = \boxed{T^H = T = D}$

Using math induction

(i) For $n=1$, it holds

(ii) Assume that theorem holds for any $k \times k$ matrix

we consider $A_{(k+1)\times(k+1)}$

Let (λ_1, w_1) be eigen-pairs.

$A w_1 = \lambda_1 w_1 \qquad \|w_1\| = 1$

$\mathrm{span}\{w_1\}$, $\{w_1, w_2 \cdots w_n\}$ be an orthonormal basis

$\mathbb{R}^n = S \oplus S^\perp$

let $W = (w_1, w_2 \cdots w_n)$

$$W^H A w_1 = \lambda_1 W^H w_1 = \lambda_1 \begin{pmatrix} w_1^H \\ \vdots \\ w_n^H \end{pmatrix} w_1 = \lambda \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda e_1$$

$$W^H A W = W^H A (w_1 \sim w_n)$$

$$= \begin{bmatrix} \lambda & \times & \times & \cdots & \times \\ 0 & & & & \\ \vdots & & & M_{k \times k} & \\ 0 & & & & \end{bmatrix}$$

by assumption, there exist a unitary matrix V_1

$V_1^H M V_1 = T_1$ (upper traingle)

$M = V_1 T_1 V_1^H$

$V^H W^H A W V$ where $V = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & V_1 & \end{bmatrix}$

$$\begin{bmatrix} 1 & 0 \\ 0 & V^H \end{bmatrix} \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & M & \\ 0 & & & \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V_1 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & & & \\ \vdots & & T_1 & \\ 0 & & & \end{bmatrix}$$

Theorem 6.4.4 Spectral Theorem

If A is Hermitian, then there exists a unitary matrix U that diagonalizes A.

Theorem 6.4.6 The Real Schur Decomposition

If A is an $n \times n$ matrix with real entries, then A can be factored into a product QTQ^T, where Q is an orthogonal matrix and T is in Schur form (2).

④ Remarks

$Ax = \lambda x$

λ: Algebric multiplicity k

$\dim N(A - \lambda I)$: Geometric multiplicity

Assignment: $\sum k_j \geq \sum \dim N(A - \lambda I)$

(i) Algebric multiplicity = Geometric multiplicity

(for symetric and Hermiton matrix)

(ii) Rank(A) = No. of linealy independent eigen vectors
(corresponding nonzero eigen values)

+

dim N(A) = No of linealy independen eigenvectors
(Corresponding zero eigen values)

(iii) Invariant

Definition A subspace S of $\mathbb{R}^n$ is said to be **invariant** under a matrix A if, for each $\mathbf{x} \in S$, $A\mathbf{x} \in S$.

(iv) Generalized eigenvalues

For $A, B \in \mathbb{R}^{n\times n}$

$Ax = \lambda Bx$ $\quad \lambda$: eigen value $\quad x$: eigenvectors

if $B = I$ $\quad Ax = \lambda x$

if B invertible $\quad Ax = \lambda Bx \quad B^{-1}Ax = \lambda x$

Let A and B be symetric

and B be nonsingular

$B^{-1}Ax = \lambda x$

$$A_{n\times n} = \frac{A+A^T}{2} + \frac{A-A^T}{2}$$

$$A_{n\times n} = \frac{A+A^H}{2} + \frac{A-A^H}{2}$$

6.5 singular value decompostion (SVD)

$A \in \mathbb{R}^{m \times n}$, not square

① Def:

The following decompostion is called SVD

$A = U \Sigma V^T$ $\quad A \in \mathbb{R}^{m \times n}$

where $U_{m \times m}$ orthogonal matrix

$V_{n \times n}$ orthogonal matrix

$\Sigma_{m \times n}$: $\begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$ $(m \geq n)$

σ_j: single value

$\sigma_1 \geq \sigma_2 \cdots \sigma_n \geq 0$.

② Theorem

Theorem 6.5.1 The SVD Theorem

If A is an m × n matrix, then A has a singular value decomposition.

Proof: • $N(A^TA) = N(A)$ • $\text{Rank}(A^TA) = \text{rank}(A)$

• A: symmetric, $\text{Rank}(A)$ = no of nonzero eigenvalue

(i) $A^TA = (U\Sigma V^T)^T(U\Sigma V^T)$

$= V\Sigma^T U^T U \Sigma V^T$

since $U^TU = I$, $= V\Sigma^T\Sigma V^T$

$= \boxed{V \Sigma_n^2 V^T}$, where $\Sigma_n = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \end{bmatrix}$

eigen-decomposition

V: eigenvector

Σ_n^2: eigenvalue

$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T$

$$= U \Sigma V^T V \Sigma U^T$$

$$= \underbrace{U}_{\text{eigenvector}} \underbrace{\Sigma_n^2}_{\text{eigenvalue}} U^T$$

Assume: $A^TAx = \lambda x$

$$x^TA^TAx = \lambda x^Tx$$

$$\lambda = \frac{x^TA^TAx}{x^Tx} \geq 0$$

Let $\lambda_1 \geq \lambda_2 \cdots \geq \lambda_r > \lambda_{r+1} \cdots = 0$

be eigenvalues of A^TA

where $\text{Rank}(A) = \text{Rank}(A^TA) = r$.

(ii) Take

$\sigma_j = \sqrt{\lambda_j}$, $\sigma_1 \geq \sigma_2 \geq \cdots \sigma_r > 0$. $\sigma_j = 0$ for $j \geq r+1$

$A^TAV = V\Lambda$, where $\Lambda = \Sigma_n^2$

Let $V = (V_1 \cdots V_n)$

$$\boxed{A^TAV_j = \lambda_j V_j}$$

(λ_j, V_j) can be found from the last equation.

$\Rightarrow (\sigma_j, V_j)$

$A^TAV_j = 0$ for $j \geq r+1$

$V_1 = (v_1 \cdots v_r)$

$V_2 = (V_{r+1} \cdots V_n)$, so $A^TAV_2 = 0$

$V = (V_1, V_2)$

$VV^T = I \quad (V_1, V_2)(V_1, V_2)^T = I$

$$V_1V_1^T + V_2V_2^T = I$$

$$A = AV_1V_1^T + AV_2V_2^T \; = 0$$

$$\Rightarrow A = A v_1 v_1^T$$

(iii) Find U

$$AV = U\Sigma, \quad U = (u_1, u_2 \cdots u_m)$$

$$\Updownarrow$$

$$\underline{Av_j = \sigma_j u_j} \quad \Rightarrow A v_1 = \Sigma_1 u_1$$

For $j = 1, 2, \cdots r$

$$u_j = \frac{1}{\sigma_j} A v_j \in R(A)$$

$$u_i^T u_j = \left(\frac{1}{\sigma_i} A v_i\right)^T \frac{1}{\sigma_j} A v_j$$

$$= \frac{1}{\sigma_i \sigma_j} v_i^T A^T A v_j$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j} v_i^T v_j$$

$$= \frac{\lambda_j}{\sigma_i \sigma_j} \delta_{ij}$$

$$u_i^T u_j = \delta_{ij}$$

So $\{u_1 \cdots u_r\}$ is an orthogonal basis for $R(A)$

$\{u_{r+1} \cdots u_m\} \subset R(A)^\perp = N(A^T)$

and $U = (u_1, u_2 \cdots u_m)$

Finally: $U\Sigma V^T = (u_1, u_2) \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} (v_1, v_2)^T$

$$= u_1 \Sigma_1 v_1^T = A v_1 v_1^T = A$$

where $\Sigma_1 = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots \; \sigma_r & \\ 0 & & 0 \end{bmatrix}$

Remarks: (i) SVD ~~are~~ is not unique

Singular values are unique

$\sigma_j = \sqrt{\lambda_j}$ λ_j: eigenvalues of A^TA

(ii) $\{v_1 \cdots v_n\}$ are eigenvectors of A^TA

$\{u_1 \cdots u_m\}$ are eigenvectors of AA^T

Thereover $Av_j = \sigma_j u_j$

$A^T u_j = \sigma_j v_j$

$\begin{cases} v_1 \cdots v_m \text{ are right singular values of } A \\ u_1 \cdots u_m \text{ are left singular values of } A \end{cases}$

(iii) $\{v_1 \cdots v_r\} \subset R(A^TA) = R(A^T)$

$\hookrightarrow$ orthonormal basis for $R(A^T)$

$\{v_{r+1} \cdots v_n\}$: orthonormal basis for $N(A^TA) = N(A)$

$\{u_1 \cdots u_r\}$: orthonormal basis for $R(A)$

$\{u_{r+1} \cdots u_m\}$: orthonormal basis for $N(A^T)$

(iv) $A = U_1 \Sigma_1 V_1^T = \begin{bmatrix} U_1 \end{bmatrix}_{m\times r} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix}_{r\times r} \begin{bmatrix} V_1 \end{bmatrix}_{r\times n}$

⊕ Applications

$\min \| A_{m\times n} - B_r \|_*$ $B_r \in R_r^{m\times n}$

if $A = U\Sigma V^T$

$B_r = U_1 \Sigma_r V_1^T$

$\|A\|_F = \Sigma a_{ij}^2$

$A: m \times n \qquad \operatorname{rank}(A) = r$

$\mathbb{R}_r^{m\times n} = \{ A \in \mathbb{R}^{m\times n} , \operatorname{rank}(A) = r \}$

For given A

$$\min \| A - B_r \|_F$$

$$B_r \in \mathbb{R}_r^{m\times n}$$

where $\|A\|_F = \sum_{i,j=1}^{m\ n} a_{ij}^2 = tr(A^T A)$

Lemma 6.5.2 *If A is an $m \times n$ matrix and Q is an $m \times m$ orthogonal matrix, then*

$$\|QA\|_F = \|A\|_F$$

proof: $\|QA\|_F^2 = tr((QA)^T QA) = tr(A^T Q^T Q A) = tr(A^T A) = \|A\|_F^2$

Theorem 6.5.3 *Let $A = U\Sigma V^T$ be an $m \times n$ matrix, and let $\mathcal{M}$ denote the set of all $m \times n$ matrices of rank k or less, where $0 < k < \operatorname{rank}(A)$. If X is a matrix in $\mathcal{M}$ satisfying (7), then*

$$\|A - X\|_F = \left(\sigma_{k+1}^2 + \sigma_{k+2}^2 + \cdots + \sigma_n^2\right)^{1/2}$$

$X = U_r \Sigma_r V_r^T = U \hat{\Sigma} V^T$

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

where $U_r = (u_1, u_2 \cdots u_r)$

$V_r = (v_1, v_2 \cdots v_r)$

$\Sigma_r = \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{pmatrix}$

(Compact form)

6.6 Quadratic Form

① Def

Definition

A **quadratic equation** in two variables x and y is an equation of the form

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0 \quad (1)$$

Equation (1) may be rewritten in the form

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + \begin{bmatrix} d & e \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} + f = 0 \quad (2)$$

Let

$$\mathbf{x} = \begin{bmatrix} x \\ y \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$$

The term

$$\mathbf{x}^T A \mathbf{x} = ax^2 + 2bxy + cy^2$$

is called the **quadratic form** associated with (1).

② principal axes theorem

$x^T A x$

A: symmetric

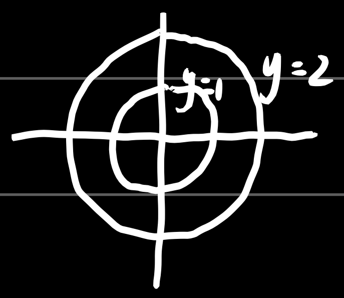
$Q^T A Q = \Lambda \qquad A = Q \Lambda Q^T$

$x^T A x = x^T Q \Lambda Q^T x$. let $Q^T x = y$

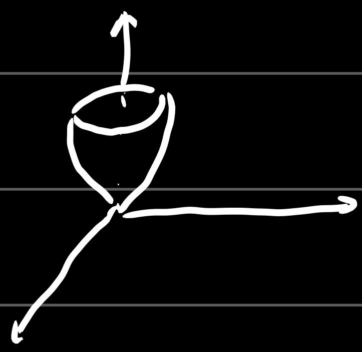
$x^T A x = y^T \Lambda y = \lambda_1 y_1^2 + \cdots \lambda_n y_n^2$

③ optimazation

$y = f(x) = x_1^2 + x_2^2$



$z = x_1^2 + x_2^2$



x^TAx is called postive definite if $x^TAx > 0$ for all $x \neq 0$

x^TAx is called negative definite if $x^TAx < 0$ for all $x \neq 0$

x^TAx is called postive semidefinite if $x^TAx \geq 0$

⊕ Theorem

Theorem 6.6.2 *Let A be a real symmetric $n \times n$ matrix. Then A is positive definite if and only if all its eigenvalues are positive.*

proof: $x^TAx > 0$ for all $x \neq 0$

<i> let $Au = \lambda u$, $u^TAu = \lambda u^Tu$

$$\lambda = \frac{u^TAu}{u^Tu} > 0$$

<ii> $Au_j = \lambda_j u_j$ $\lambda_j > 0$. For any $x \in \mathbb{R}^n$

$(u_1, u_2 \cdots u_n)$: orthonormal vectors

$$x^TAx = (\Sigma a_j u_j)^T A \Sigma a_j u_j$$

$$= (\Sigma \alpha_i u_i)(\Sigma \alpha_j \lambda_j u_j)$$

$$= \alpha_i^2 \lambda_i u_i^T u_i + \cdots \alpha_n^2 \lambda_i u_n^T u_n$$

$$= \lambda_1 \alpha_1^2 + \lambda_2 \alpha_2^2 + \cdots \lambda_n \alpha_n^2 > 0$$

From the proof we know that

$$x^TAx = \lambda_1 \alpha_1^2 + \cdots \lambda_n \alpha_n^2$$

where $x = \Sigma \alpha_j u_j$

$$\|x\|^2 = (\Sigma \alpha_j u_j)^T \Sigma \alpha_j u_j$$

$$= \alpha_1^2 + \cdots \alpha_n^2$$

$$\lambda_n \leq \frac{x^T A x}{\|x\|^2} = \frac{\lambda_1 \alpha_n^2 + \cdots \lambda_n \alpha_n^2}{\alpha_1^2 + \cdots \alpha_n^2} \leq \lambda_1$$

6.7 Symmetric positive definite matrix

Property I If A is a symmetric positive definite matrix, then A is nonsingular.

Property II If A is a symmetric positive definite matrix, then $\det(A) > 0$.

Property III If A is a symmetric positive definite matrix, then the leading principal submatrices $A_1, A_2, \ldots, A_n$ of A are all positive definite.

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \\ \vdots & & & \\ a_{n1} & \cdots & \cdots & a_{nn} \end{bmatrix}$$

eigenvalues of $A > 0$

(iv) for a positive definite matrix

$A = LU$, $A_k = L_k U_k$, $\det(A_k) = \det(L_k)\det(U_k)$

$= \det(U_k)$

$\Rightarrow u_{kk} > 0$