1.7 Symmetric matrices, Hermitian matrices, positive definite matrices

Theorem (a) Every real symmetric matrix A can be written as QDQ^T for a real diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and an orthogonal matrix Q.

(b) Every complex Hermitian matrix A can be written as UDU^* for a real diagonal matrix $D = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$ and a unitary matrix U.

Proof. If $A \in M_2(\mathbb{R})$ be symmetric, let $v \in= [v_1, v_2]^T \in \mathbb{R}^2$ be such that $v^T v = 1$ and

$$v^T A v = \lambda_1 = \max\{x^T A x : x \in \mathbb{R}^2, x^T x = 1\}.$$

Then $Q = \begin{pmatrix} v_1 & v_2 \\ v_2 & -v_1 \end{pmatrix}$ is orthogonal, and we claim that $Q^T A Q = \operatorname{diag}(\lambda_1, \lambda_2)$. If not, assume $Q^T A Q = \begin{pmatrix} \lambda_1 & b \\ b & c \end{pmatrix}$. Then

$$f(\theta) = [\cos\theta, \sin\theta] Q^T A Q [\cos\theta, \sin\theta]^T = \cos^2\theta\lambda_1 + \cos\theta\sin\theta b + \sin^2\theta c,$$

and $f'(0) = 2b \neq 0$. So, we can find θ near 0 such that $u^T A u > \lambda_1$ with $u = Q[\cos\theta, \sin\theta]^T$, which is a contradiction.

Now, for $A \in M_n(\mathbb{R})$. Let $v \in \mathbb{R}^n$ be a unit vector such that

$$v^T A v = \lambda_1 = \max\{x^T A x : x \in \mathbb{R}^n, x^T x = 1\}.$$

Suppose Q is an orthogonal matrix such that $Q^T A Q = A_1$. Then $A_1 = [\lambda_1] \oplus A_2$ ....

For a Hermitian matrix $A \in M_n(\mathbb{C})$, if $v \in \mathbb{C}^n$ and $\mu = v^* AV$, then $\bar{\mu} = (v^* A v)^* = v^* A^* v = v^* A v = \mu$. So, μ is real and we can consider a unit vector $v \in \mathbb{C}^n$ such that

$$v^* A v = \lambda_1 = \max\{u^* A u : u \in \mathbb{C}^n, u^* u = 1\}.$$

In practice, we can do the proof and diagonalization as follows. First, we can treat a real symmetric matrix as a complex Hermitian matrix A. Then there is a possible complex eigenvalue λ and a unit eigenvector x. We have $Ax = \lambda x$ and thus $x^* A x = x^* \lambda x = \lambda$. Now, $\bar{\lambda} = *x^* A x)^* = x^* A^* x x^* A x = \lambda$. So, λ is real.

Then we compute an eigenvalue λ and an unit eigenvector x, which is real in the real case, so that $Ax = \lambda x$, and let U_1 be a unitary matrix (an orthogonal matrix in the real case) with x as the first column. Then $U_1^* A U_1$ has the first column equal to $[\lambda, 0, \ldots, 0]^T$. But $U_1^* A U_1$ is also Hermitian (real symmetric). So $U_1^* A U_1 = [\lambda] \oplus A_1$ where A_1 is also Hermitina (real symmetric). By induction, $U_2^T A_1 U_2 = D_2$, a real diagonal matrix. Thus, $U = U_1([1] \oplus U_2)$ is unitary and $U^* A U = D$.

$$U_1^* A U_1 = \left[\begin{array}{c|c} \lambda_1 & 0 \\ \hline 0 & A_1 \end{array}\right]$$

$[x \cdots]$

$Ax = \lambda x$

Applications

1. Maximum and minimum of real valued function $f(x) = f(x_1, \ldots, x_n)$.

Then $\underline{f(x_1 + t_1, \ldots, x_n + t_n)} \to 0$

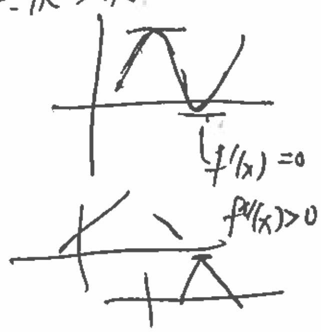
$= f(x_1, \ldots, x_n) + \underbrace{(f_{x_1}(x), \ldots, f_{x_n}(x))}_{\phi \quad \phi}(t_1, \ldots, t_n)^T + (t_1, \ldots, t_n)\boxed{J_f(x)}(t_1, \ldots, t_n)^T + O(t^3)$,

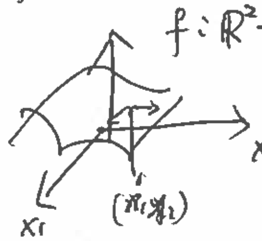
where

$$J_f(x) = (f_{x_i, x_j}(x)) = \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right).$$

Thus, $f(x)$ is minimum if $f_{x_j}(x) = 0$ for all $j = 1, \ldots, n$, and $J_f(x)$ is positive semi-definite.

$f: \mathbb{R} \to \mathbb{R}$ $\qquad$ $f: \mathbb{R}^2 \to \mathbb{R}$ (continuously differentiable)





$f_1(x, y)$
$f_2(x, y)$

$(x, y) J_f \binom{x}{y}$

min. if $\begin{cases} f_1(x, y) = 0 \\ f_2(x, y) = 0 \\ J_f(x, y) \text{ p definite} \end{cases}$ $\quad \phi \quad A$

2. Major and minor axes of elliptical disk/ellipsoid.

Suppose the ellipse equation $1 = 5x^2 + 8xy + 5y^2$ is written as $1 = (x, y)A(x, y)^T$ with $A = \left[\frac{\partial^2 f}{\partial x_1 \partial x_2}\right]$ $\begin{pmatrix} 5 & 4 \\ 4 & 5 \end{pmatrix}$. Then A is positive definite, and $A = QDQ^T$ with $D = \text{diag}(9, 1)$ and $Q = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$. Then the ellipse equation becomes $1 = 9X^2 + Y^2$ with $X = (x + y)\sqrt{2}$ and $Y = (-x + y)/\sqrt{2}$. Geometrically, we apply a rotation of $-\pi/4$, we get a vertical ellipse.

$\cancel{0}$

In general, If $f(x_1 + t_1, x_2 + t_2) = f(x_1, x_2) + (f_1(x_1, x_2)\ f_2(x_1, x_2))\begin{pmatrix} t_1 \\ t_2 \end{pmatrix}$

$+ \boxed{(t_1\ t_2)(J_f)\begin{pmatrix} t_1 \\ t_2 \end{pmatrix} + O(t^3)}$

Given $(f_1(x_1, x_2), f_2(x_1, x_2)) = (u_1, u_2) \neq 0$

Choose (t_1, t_2) s.t. $(t_1\ t_2) = \gamma(u_1, u_2)$, $\gamma < 0$

If $(f_1(x_1, x_2), f_2(x_1, y_2)) = (0, 0)$

$J_f = Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^T$

$(t_1\ t_2) J_f \binom{t_1}{t_2} = (t_1\ t_2) Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^T \binom{t_1}{t_2}$

$= (y_1\ y_2)\begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}\binom{y_1}{y_2}$

$= |y_1|^2 \lambda_1 + |y_2|^2 \lambda_2$

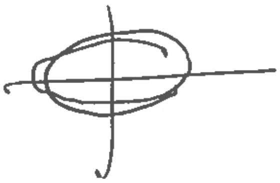
(2) In general, for the expression

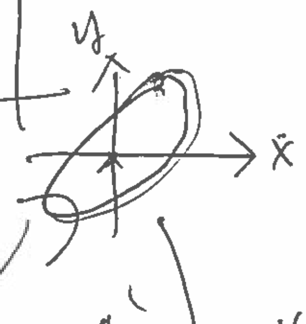
$$(x_1\ x_2)\, A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix},$$

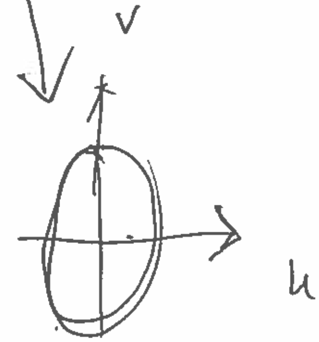
we can relate it to the ellipse equation

$$\frac{x^2}{a}+\frac{y^2}{b}=1$$

$$\boxed{ax^2+bxy+cy^2=1}$$

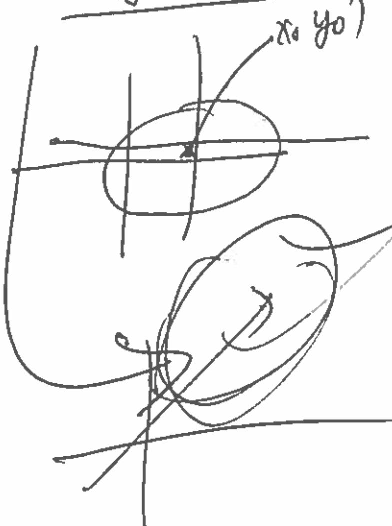






$$ax^2+bx+cy^2+dy+exy=f$$

$$a(x-x_0)^2+c(y-y_0)^2 + e(x-x_0)(y-y_0) = \hat{f}$$



$$(x,y)\begin{bmatrix} a & b/2 \\ b/2 & c \end{bmatrix}\begin{pmatrix} x \\ y \end{pmatrix}=1$$

$$\|$$

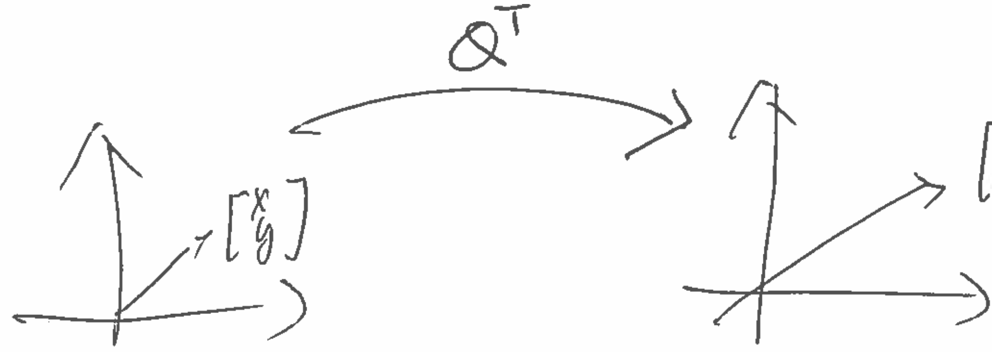
$$(x,y)\, Q \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} Q^T \begin{pmatrix} x \\ y \end{pmatrix} = 1$$

$$[x,y]\begin{bmatrix} \cos\theta & \mp\sin\theta \\ \pm\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} \cos\theta & \pm\sin\theta \\ \mp\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix}=$$

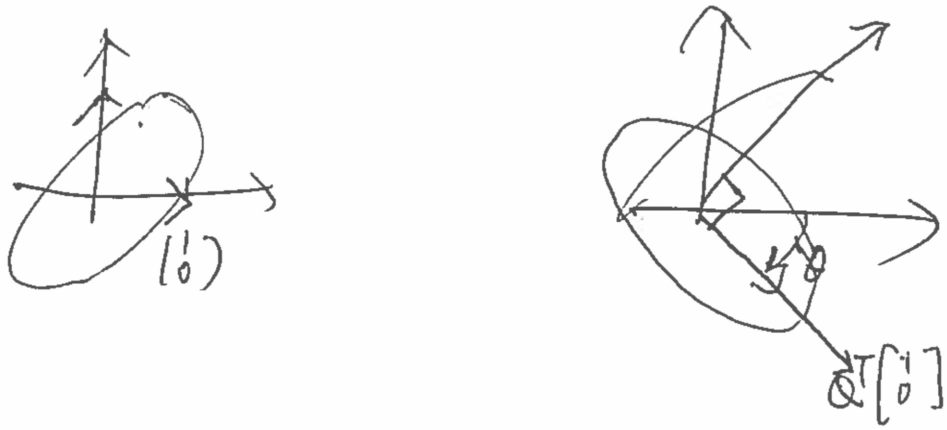
$$[u,v]\begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}\begin{bmatrix} u \\ v \end{bmatrix}=1$$

$$\begin{bmatrix} -\sin\theta \\ \cos\theta \end{bmatrix} \text{ or } \begin{bmatrix} \sin\theta \\ -\cos\theta \end{bmatrix}$$

$$\begin{bmatrix} x \\ y \end{bmatrix} \mapsto \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \cos\theta\, x + \sin\theta\, y \\ -\sin\theta\, x + \cos\theta\, y \end{bmatrix}$$



$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} \cos\theta\, x + \sin\theta\, y \\ -\sin\theta\, x + \cos\theta\, y \end{bmatrix}$$



See p. 50 – 51. to see actual example.

In matlab. $\cancel{Q = Q^T}$ $\cdot\ A = A^T$

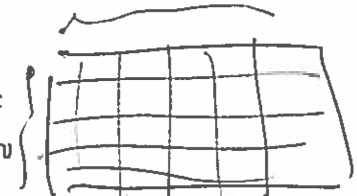
$$[U, D] = \mathrm{eig}(A) \quad \rightarrow \quad [u_1 | u_2]$$

$$\rightarrow \quad D = \begin{bmatrix} d_1 & 0 \\ 0 & d_2 \end{bmatrix}$$

s.t. $\cancel{Q = U}$

$$A = UDU^T$$

200

$A =$  $\equiv$ picture

100

1.8 Singular value decomposition, polar decomposition.

Theorem Let $A \in M_{m,n}$. There exists unitary $U \in M_{m,k}$ and $V \in M_{n,k}$ such that $U^*U = I_k = V^*V$ and $\Sigma = \mathrm{diag}\,(s_1, \dots, s_k)$ with $s_1 \geq \cdots \geq s_k > 0$ such that

$$A = s_1 u_1 v_1^* + \cdots + s_k u_k v_k^*$$

$$\underbrace{A}_{m\times n} = \overset{m\times k}{U}\underbrace{\Sigma}_{k\times k}\overset{k\times n}{V^*} = \sum_{j=1}^{k} s_j u_j v_j^*.$$

$$= [u_1 | \cdots | u_k] \begin{bmatrix} s_1 & 0 \\ 0 & s_k \end{bmatrix} \begin{bmatrix} v_1^* \\ \vdots \\ v_k^* \end{bmatrix}_{k\times n}$$

Proof. If $V^*A^*AV = \mathrm{diag}\,(s_1^2, \dots, s_n^2)$, then AV has orthogonal columns of lengths $s_1, \dots, s_n$. So, there is a unitary U such that $AV = U\Sigma$, where $\Sigma = \mathrm{diag}\,(s_1, \dots, s_n)$. □

Definition The values $s_1 \geq \cdots \geq s_k$ are known as the (nonzero) singular values of A. The produce $U\Sigma V^*$ is called the singular decomposition of A.

Note that AA^* and A^*A have eigenvalues $s_1^2 \geq \cdots \geq s_k^*$ and zoeros.

$$A^*A = \underset{n\times k}{V} \begin{bmatrix} s_1 & 0 \\ 0 & s_k \end{bmatrix} U^* U \begin{bmatrix} s_1 & \\ & s_k \end{bmatrix} \underset{k\times n}{V^*} = V \begin{bmatrix} s_1^2 & 0 \\ 0 & s_k^2 \end{bmatrix} V^* = \sum_{i=1}^{k} s_i^2 v_i v_i^* \in M_n$$

$\therefore A^*A$ has non-zero e.v. $s_1^2, \dots, s_k^2$

$$AA^* = U \begin{bmatrix} s_1 & \\ & s_k \end{bmatrix} V^* V \begin{bmatrix} s_1 & 0 \\ 0 & s_k \end{bmatrix} U^* = \sum_{i=1}^{k} s_i^2 u_i u_i^* \in M_m$$

has non-zero e.v. $s_1^2, \dots, s_k^2$

Also, the Wielandt matrix $S = \begin{bmatrix} 0 & A \\ A^* & 0 \end{bmatrix}$ has eigenvalues $s_1, \dots, s_k, -s_k, \dots, s_1$ and zeros.

Theorem Let $A \in M_n$. Then $A = UP = QV$ for some unitary U, V and positive semidefinite P, Q with eigenvalues $s_1 \geq \cdots \geq s_n$.

Example

$$V^*A^*AV = \begin{bmatrix} s_1^2 & 0 \\ 0 & s_n^2 \end{bmatrix}$$

$$\begin{bmatrix} y_1^* \\ y_n^* \end{bmatrix} [y_1 \cdots y_n]$$

$y_i^* y_i = s_i^2$, $i \neq j$: $y_i^* y_j = 0$

If $s_i \neq 0$, then $\dfrac{y_i^* y_i}{s_i^2} = 1$, $\dfrac{y_i}{s_i} = u_i$ has length 1

$$[y_1 \cdots y_k] = [u_1 \cdots u_k] \begin{bmatrix} s_1 & \\ & s_k \end{bmatrix}$$

$$A[y_1 \cdots y_k] = [u_1 \cdots u_k] \begin{bmatrix} s_1 & 0 \\ 0 & s_k \end{bmatrix}$$

ϕ

So $A = [u_1 \cdots u_k] \begin{bmatrix} s_1 & 0 \\ 0 & s_k \end{bmatrix} \begin{bmatrix} y_1^* \\ y_k^* \end{bmatrix}$.

A is $m \times n$

$$S = \underbrace{\left[\begin{array}{c|c} O_m & A \\ \hline A^* & O_n \end{array}\right]}_{m+n} \Big\} m+n$$

has eigenvalues $s_1, \dots, s_k \; -s_k \dots -s_1$ $0, \dots, 0$.

$$= \left[\begin{array}{c|c} O_m & U\begin{bmatrix} s_1 & \\ & s_k\end{bmatrix}V^* \\ \hline V\begin{bmatrix} s_1 & \\ & s_k\end{bmatrix}U^* & O_n \end{array}\right]$$

$$= \left[\begin{array}{c|c} \hat{U} & O \\ \hline O & \hat{V} \end{array}\right] \left[\begin{array}{c|c} O_m & \begin{array}{c|c} s_1 \ddots s_k & 0 \\ \hline 0 & 0 \end{array} \\ \hline \begin{array}{c|c} s_1 \ddots s_k & 0 \\ \hline 0 & 0 \end{array} & O_n \end{array}\right] \left[\begin{array}{c|c} \hat{U}^* & O \\ \hline O & \hat{V}^* \end{array}\right]$$

$$\begin{bmatrix} 0 & abc^* \\ c^*b^*a^* & 0 \end{bmatrix} = \begin{bmatrix} a & 0 \\ 0 & c \end{bmatrix}\begin{bmatrix} 0 & b \\ b^* & 0 \end{bmatrix}\begin{bmatrix} a^* & 0 \\ 0 & c^* \end{bmatrix}$$

$\hat{U} \in M_m$ is unitary $= [U \mid U_2]$ $\rightarrow$ $m-k$ columns.

$\hat{V} \in M_n$ is unitary $= [V \mid U_2]$ $n-k$ columns

Note e.v. of SAS^{-1} are the same as those of A

Note $\underset{m \times m}{\hat{U}} \underset{m \times n}{\left[\begin{array}{c|c} s_1 \ddots s_k & 0 \\ \hline 0 & 0 \end{array}\right]} \underset{n \times n}{V^*} = \sum_{i=1}^{k} s_i u_i v_i^*$

$\det(\lambda I - SAS^{-1})$

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

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

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

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

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

$\Rightarrow S$ has ~~the same~~ eigenvalues as

$$\begin{array}{c} \begin{array}{cc} 1 \dots m & m+1 \dots m+k \end{array} \\ \left[\begin{array}{c|c} O & \begin{array}{c|c} s_1 \ddots s_k & \\ \hline & 0 \end{array} \\ \hline \begin{array}{c|c} s_1 \ddots s_k & 0 \\ \hline 0 & 0 \end{array} & O \end{array}\right] \end{array}$$

$$\begin{array}{c|cc} & 1\; m+1 & 2\; m+2 \\ \hline 1 & 0 \;\; s_1 & 0 \\ m+1 & s_1 \;\; 0 & \\ & 0 & 0 \;\; s_2 \\ & & s_2 \;\; 0 \end{array}$$

Up to a permutation similarity, $P(S)$

$$P^t S P \text{ is } = \begin{pmatrix} \begin{matrix} 0 & s_1 \\ s_1 & 0 \end{matrix} & 0 & & 0 & \\ 0 & \begin{matrix} & s_2 \\ s_2 & \end{matrix} & & & \\ & & \ddots & & \\ & & & \begin{matrix} 0 & s_k \\ s_k & 0 \end{matrix} & 0 \\ & 0 & & & 0 \end{pmatrix}$$

$$\det(\lambda I - P^t S P) = (\lambda^2 - s_1^2)(\lambda^2 - s_2^2) \cdots (\lambda^2 - \lambda_k^2)\, \lambda^{m+n-2k}$$

$$= (\lambda - s_1)(\lambda + s_1) \cdots (\lambda - \lambda_k)(\lambda + \lambda_k)\, \lambda^{m+n-2k}$$