for all $t \in (0, \delta)$. Note that if $\|x\|^m = 1$, then the maximum possible value of $|x_{i_1} \cdots x_{i_m}|$ is 1. So for any $x \neq 0$, we have $|x_{i_1} \cdots x_{i_m}| \leq \|x\|^m$. Hence

$$\left|\frac{r(x)}{\|x\|^m}\right| \leq \frac{\epsilon}{m!} \sum_{i_1,\ldots,i_m} \left|\frac{x_{i_1} \cdots x_{i_m}}{\|x\|^m}\right| \leq \frac{\epsilon n^m}{m!} \leq \epsilon \times \text{bounded quantity}.$$

This completes the proof. ∎

**[26.18]** **Corollary : Taylor-I** Let $E \subseteq \mathbb{R}^n$ be open, $f : E \to \mathbb{R}$ be in $\mathcal{C}^2(E)$ and $[a, a+x] \subseteq E$. Then

$$f(a+x) = f(a) + \langle \nabla f(a), x \rangle + \frac{1}{2} x^t H(a) x + r(x),$$

where $\lim\limits_{\|x\| \to 0} \frac{r(x)}{\|x\|^2} = 0$.

**[26.19]** **Corollary: Taylor-II** Let $E \subseteq \mathbb{R}^n$ be open, $f : E \to \mathbb{R}$ be in $\mathcal{C}^2(E)$ and $[a, a+x] \subseteq E$. Then $\exists t \in (0,1)$ such that

$$f(a+x) = f(a) + \langle \nabla f(a), x \rangle + \frac{1}{2} x^t H(a+tx) x.$$

**[26.20]** **Taylor series** Let $f : \mathbb{R}^n \to \mathbb{R}$ be infinitely differentiable. Then the TAYLOR SERIES $\boldsymbol{T_f(x;a)}$ of $f$ about the point $a$ is defined as

$$f(a) + \sum_i D_i f(a)(x-a)_i + \frac{1}{2!} \sum_{i,j} D_{ij} f(a)(x-a)_i (x-a)_j + \frac{1}{3!} \sum_{i,j,k} D_{ijk} f(a)(x-a)_i (x-a)_j (x-a)_k + \cdots$$

**[26.21]** **Example** Take $f(y,z) = y^2 z^4 + yz^3 - 5yz + 6$ and $a = (1,2)$. Find the coefficient of $(y-1)^2(z-2)^2$ in $T_f((y,z);a)$ in two different ways.

$f(x) = x^9 - 7x^6 + 8x^3$

$T_f(x;2) = - - - - -$

sum of all coefficients ?

*Answer.* To apply Taylor's theorem, put $w = (y,z) - (1,2)$. Terms with degree 4 can only occur in

$$\frac{1}{4!} \sum_{i,j,k,l=1}^{2} D_{ijkl} f(a)\; w_i w_j w_k w_l.$$

We want $w_i w_j w_k w_l = (y-1)^2(z-2)^2$, which can be done in $\frac{4!}{2!2!}$ ways. Hence, the coefficient is

$$\frac{1}{4!}\binom{4}{2} D_{1,1,2,2} f(a) = \frac{1}{2!2!}(2.4.3.2^2) = 24.$$

∘ Alternately, note that

$$f(y,z) = (y-1+1)^2(z-2+2)^4 - (y-1+1)(z-2+2)^3 - 5(y-1+1)(z-2+2) + 6.$$

The coefficient for $(y-1)^2(z-2)^2$ can only come from the first term. When expanded using binomial expansion, it will look like $(y-1)^2\binom{4}{2}(z-2)^2 2^2$. So the required coefficient is 24.

| To remember the coefficients of Taylor series of $f : \mathbb{R}^n \to \mathbb{R}$ at $a$. |
|---|
| a) Take $\alpha = (\alpha_1, \cdots, \alpha_n)$ where $\alpha_i$ are nonnegative integers and take $a = (a_1, \cdots, a_n)^t$. |
| b) Use the notations $\boldsymbol{D}^{\boldsymbol{\alpha}} := D_1^{\alpha_1} D_2^{\alpha_2} \cdots D_n^{\alpha_n}$, $\boldsymbol{x}^{\boldsymbol{\alpha}} := x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, $\alpha! := \alpha_1! \alpha_2! \cdots \alpha_n!$. |
| c) Then the coefficient of $x^\alpha$ in $T_f(x; a)$ is $\frac{1}{\alpha!} D^\alpha f(a)$. |

## 27 Lecture 27

### Positive definite matrices

matrices

**[27.1]** **Definition** Let $A \in M_n(\mathbb{C})$. $\longrightarrow$ $n \times n$ complex

$\begin{bmatrix} 1 \\ i \end{bmatrix}^* = \begin{bmatrix} 1 & -i \end{bmatrix}$

a) It is called POSITIVE DEFINITE(pd) if $x^*Ax > 0$ holds $\forall x \in \mathbb{C}^n \setminus \{0\}$.

b) It is called POSITIVE SEMIDEFINITE(psd) if $x^*Ax \geq 0$ holds $\forall x \in \mathbb{C}^n$.

pd if $x^*Ax > 0$ $\forall x \neq 0$

psd if $x^*Ax \geq 0$ $\forall x \in \mathbb{C}^n$

$\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$ $\to 7$

**[27.2]** **Facts and examples**

$\begin{bmatrix} 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} = 8$

$(3,3)$

a) Let $A \in M_n(\mathbb{C})$ be pd. Then $a_{ii} = e_i^t A e_i > 0$.

$\forall x \neq 0$ $\overline{x^* A x > 0}$

a') Let $A \in M_n(\mathbb{C})$ be psd. Then $a_{ii} = e_i^t A e_i \geq 0$.

b) Take any matrix $A \in M_n(\mathbb{C})$. Then $A^*A$ and $AA^*$ are psd. If $A$ is nonsingular, then they are pd.

*Proof.* (Self) For each $x$, we have $x^*A^*Ax = (Ax)^*Ax = \|Ax\|^2 \geq 0$. If $x \neq 0$ and $A$ is nonsingular, then $Ax \neq 0$ and so $\|Ax\|^2 > 0$. ■

$x^*(A^*A)x = (Ax)^* Ax = \|Ax\|^2 \geq 0$ $\qquad e_i^T A e_j = a_{ij}$

Let $x \neq 0$. Then $Ax \neq 0$, so $x^*A^*Ax = \|Ax\|^2 > 0$.

c) The matrix $I$ is pd. The zero matrix is a psd matrix.

$x^* I x = \|x\|^2 > 0$

d) A pd matrix is by definition a psd matrix.

e) A singular matrix cannot be pd. Then $x^*Ax = 0$, so it cannot be pd.

$\exists x \neq 0$, $Ax = 0$.

$\to$ Hermitian

**[27.3]** **Fact** `(Psd implies Hermitian)` Let $A \in M_n(\mathbb{C})$ be a psd matrix. Then $A^* = A$.

*Proof.* (Self) Fix $i, j$, $i \neq j$. Let $v = e_i + e_j$. As $v^*Av \geq 0$, we have $a_{ii} + a_{ij} + a_{ji} + a_{jj} \geq 0$. But we already know that $a_{ii} + a_{jj} \geq 0$. So, we get $a_{ij} + a_{ji} \in \mathbb{R}$. That is, $Im(a_{ij}) = -Im(a_{ji})$.

Now take $v = e_i + \mathrm{i}e_j$. As $v^*Av \geq 0$, we have $a_{ii} + \mathrm{i}a_{ij} - \mathrm{i}a_{ji} + a_{jj} \geq 0$. But we already know that $a_{ii} + a_{jj} \geq 0$. So, we get $a_{ij} - a_{ji} \in \mathrm{i}\mathbb{R}$. That is, $Re(a_{ij}) = -Re(a_{ji})$.

So $\overline{a_{ij}} = a_{ji}$. ■

$\boxed{a_{12} = \overline{a_{21}}}$ $\quad$ ~~$a_{12}$~~ $\quad$ A psd $\Rightarrow$ diag entries $\geq 0$.

$a_{12}, a_{21}$

$\begin{bmatrix} 1 & 1 & 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = a_{11} + a_{12} + a_{21} + a_{22} \geq 0$ $\Rightarrow \boxed{a_{12} + a_{21} \in \mathbb{R}}$

$\begin{bmatrix} 1 & -i & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \begin{bmatrix} 1 \\ i \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = a_{11} + i a_{12} - i a_{21} + a_{22} \geq 0 \Rightarrow i(a_{12} - a_{21}) \in \mathbb{R}$

$\boxed{a_{12} - a_{21} \in i\mathbb{R}}$

**[27.4]** **Lemma** Let $A \in M_n(\mathbb{R})$.

$(u^t A v)^t = v^t A^t u = v^t A u$

a) Then $A$ is pd iff $A^t = A$ and $x^t A x > 0$ holds $\forall x \in \mathbb{R}^n$, $x \neq 0$.

b) Then $A$ is psd iff $A^t = A$ and $x^t A x \geq 0$ holds $\forall x \in \mathbb{R}^n$.

*Proof.* (Self) a) Let $A$ be a real pd matrix. As $A$ is pd, it is Hermitian, and in our case it is symmetric. and $x^t A x > 0$ for all real $x \neq 0$.

Conversely, suppose that $A$ is real symmetric and $x^t A x > 0$ for all real $x \neq 0$. Note that, for a real symmetric matrix $A$, we have $x^t A y = y^t A x$. Hence

$$(x + \mathrm{i}y)^* \mathrm{A}(\mathrm{x} + \mathrm{i}\mathrm{y}) = \mathrm{x}^t\mathrm{A}\mathrm{x} + \mathrm{y}^t\mathrm{A}\mathrm{y} + \mathrm{i}(\mathrm{x}^t\mathrm{A}\mathrm{y} - \mathrm{y}^t\mathrm{A}\mathrm{x}) = \mathrm{x}^t\mathrm{A}\mathrm{x} + \mathrm{y}^t\mathrm{A}\mathrm{y} > 0,$$

if $x + \mathrm{i}y \neq 0$. The proof for b) is similar. ∎

$A$ pd $\Rightarrow A^* = A$ and $x^* A x > 0$ $\forall x \neq 0$

A real: $A^t = A$, $\boxed{x^t A x > 0}$ $\forall x \neq 0$ real $\Rightarrow$ $A$ is pd, $x^* A x > 0$ $\forall x \neq 0$.

$x = u + iv \neq 0$

$x^* A x = u^* A u + u^* A(iv) - i v^* A u + v^* A v > 0$ | $u^* A v - v^* A u = 0$

**[27.5]** **Theorem** `(Equivalent conditions for pd/psd)` Let $A \in M_n(\mathbb{C})$. TFAE.

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

a) $A$ is pd (resp. psd).

b) $A^* = A$ and the eigenvalues of $A$ are positive (resp. nonnegative).

c) $A = B^* B$ for some nonsingular (resp. not necessarily nonsingular) matrix $B$.

*Proof.* (Self) a)$\Rightarrow$b). Let $A$ be pd. We already know that $A^* = A$. Let $(\lambda, x)$ be an eigenpair. Then $x^* A x = x^* x \lambda \Rightarrow \lambda = \frac{x^* A x}{x^* x} > 0$, as $A$ is pd.

b)$\Rightarrow$c). Suppose that $A^* = A$ and all eigenvalues of $A$ are positive. Then, by spectral theorem (for Hermitian matrices), there exists a unitary matrix $U$ and a diagonal matrix $\Lambda = diag(\lambda_1, \ldots, \lambda_n)$, $\lambda_i > 0$ such that $A = U^* \Lambda U$. So

$$A = U^* \Lambda^{1/2} \Lambda^{1/2} U = U^* \Lambda^{1/2} U U^* \Lambda^{1/2} U = B^* B.$$

Notice that $B$ is nonsingular as $\Lambda$ has positive diagonal entries.

c)$\Rightarrow$a). Already done.

$A$ pd $\Rightarrow$ eig values $> 0$

The proof for the psd part is similar. ∎

let $(\lambda, x) \to$ eigenpair

$A = U^* \Lambda U$ → diagonal matrix of eigenvalues

$x^* A x > 0$

$x^* \lambda x = \lambda x^* x = \lambda \|x\|^2$

$= U^* \Lambda^{1/2} \Lambda^{1/2} U$

$= U^* \Lambda^{1/2} U (U^* \Lambda^{1/2} U)$

$= B^* B$

so $\lambda = \frac{x^* A x}{\|x\|^2} > 0$

We will need the following notations.

$\{1,2,\dots,n\}$

**[27.6] Definition**

a) Let $A \in M_n(\mathbb{C})$ and $S \subseteq [n]$. By $\boldsymbol{A(S,S)}$, we denote the submatrix of $A$ whose entries are indexed by $S$. This is called the PRINCIPAL SUBMATRIX of $A$ indexed by $S$.

$A = (\{1,4\},\{1,2\})$

b) For $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 2 & 4 & 3 \\ 0 & 2 & 3 & 1 \\ 9 & 8 & 7 & 6 \end{bmatrix}$ and $S = \{1,4\}$, we have $A(S,S) = \begin{bmatrix} 1 & 4 \\ 9 & 6 \end{bmatrix}$.

$\begin{bmatrix} 1 & 2 \\ 9 & 8 \end{bmatrix}$

c) A LEADING PRINCIPAL SUBMATRIX is one for which the subset $S$ is $\{1,2,\dots,k\}$ for some $k$.

$S=\{1\}$ $\quad [1]$

$S=\{1,2\}$ $\quad \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$

$S=\{1,2,3\}$ $\quad \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 4 \\ 0 & 2 & 3 \end{bmatrix}$

d) For the matrix $A$ in b), and $S = \{1,2\}$, we have $A(S,S) = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix}$. It is a leading principal submatrix.

e) The matrix $A$ in b), has only four leading principal submatrices.

f) The determinant of a square submatrix is called a MINOR. The determinant of a leading principal submatrix is called a 'leading principal minor'.

g) Sometimes we use $\boldsymbol{A(S|S)}$ to mean the matrix $A(S^c, S^c)$.

$A(S,S)$

$A(S|S) := A(S^c,S^c)$

We will need the following two results from advanced linear algebra.

**[27.7] Cauchy interlacing theorem** Let $A$ and $B = \left[\begin{array}{c|c} A & y \\ \hline y^* & b \end{array}\right]$ be Hermitian. Let $\lambda_1 \le \cdots \le \lambda_n$ be the eigenvalues of $A$ and $\mu_1 \le \cdots \le \mu_{n+1}$ be the eigenvalues of $B$. Then

$$\mu_1 \le \lambda_1 \le \mu_2 \le \lambda_2 \le \cdots \le \lambda_n \le \mu_{n+1}.$$

That is, the eigenvalues of $A$ and $B$ interlace.

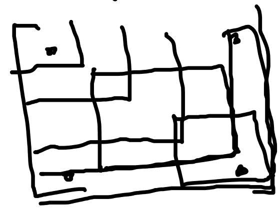
$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$

**[27.8] Theorem** The eigenvalues of a matrix $A$ are continuous functions of its entries.

Now we can give two useful conditions to test whether a matrix is pd/psd.

Hermitian

**[27.9] Theorem** (Test for pd/psd) Let $A \in M_n(\mathbb{C})$.



$>0$

a) Then $A$ is pd iff $A^* = A$ and the leading principal minors of $A$ are positive.

b) Then $A$ is psd iff $A^* = A$ and all principal minors are nonnegative.

$A^*=A$ , Leading principal minors $>0$

$a_{11} > 0$

$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix}$

$\lambda_1 \quad \overset{>0}{a_{11}} \le \lambda_2 > 0 \Rightarrow \lambda_2 > 0$

$0 < \lambda_1 \quad \lambda_2$

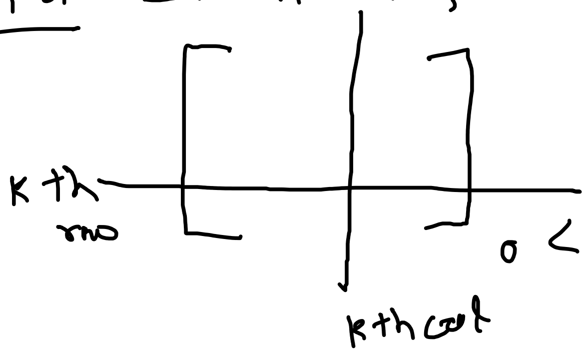
$\mu_1 \le \quad \le \mu_2 \le \quad \le \mu_3 \Rightarrow \mu_1 > 0$

*Proof.* (Not for exam) a) Suppose that $A$ is pd. Then $A^* = A$ and the eigenvalues $\lambda_i > 0$. By Cauchy interlacing theorem, the eigenvalues of the leading principal matrices are $> 0$. Hence their determinant is positive.

Conversely, suppose that $A^* = A$ and the leading principal minors are positive. By use of interlacing, we see that eigenvalues of $A$ are positive. So $A$ is pd.

---

$A$ pd $\Rightarrow A^* = A$, eigenvalues of $A > 0$




$A(k|k) \longrightarrow$ pd

$0 < \lambda_1 \quad \lambda_2 \quad \lambda_{n-1} \quad \lambda_n$

$\nu_1 \quad \nu_2 \quad \nu_{n-1}$

b) If $A$ is psd, then showing all principal minors $\geq 0$ is similar to that of a). To prove the converse, let $A^* = A$ and assume that all principal minors are $\geq 0$. Note that

$$|A + \epsilon e_1 e_1^t| = (a_{11} + \epsilon)|A(1|1)| - a_{12}|A(1|2)| + a_{13}|A(1|3)| - \cdots = |A| + \epsilon|A(1|1)| \geq |A|,$$

so that adding $\epsilon > 0$ to a diagonal entry does not decrease the determinant. Notice that we get $|A+\epsilon e_1 e_1^t| > |A|$, if $|A(1|1)| > 0$ and $\epsilon > 0$.

Next we show that $|A + \epsilon I| > 0$ whenever $A$ has all minors nonnegative. We proceed by induction. Assume that it is true for $n-1$. Consider the case for $n$ and let $B$ be obtained by adding $\epsilon$ to all diagonal entries of $A$ except the first. Notice that $A(1|1)$ is Hermitian and it has all principal minor $\geq 0$. By induction hypothesis, we have that $|B(1|1)| > 0$. Hence by earlier argument $|B| > 0$.

It now follows that $A + \epsilon I$ is pd. This is true for each $\epsilon > 0$. As eigenvalues are continuous functions of the entries, it follows that eigenvalues of $A$ are nonnegative. Hence $A$ is psd.

**[27.10]** **NoPen**

a) If we say 'a matrix $A \in M_n(\mathbb{C})$ is psd iff $A^* = A$ and all leading principal minors are nonnegative', what would go wrong?

b) Let $A \in M_n(\mathbb{C})$ be pd. Should every (nonempty) principal submatrix of $A$ be pd?

c) Let $A \in M_n(\mathbb{C})$ have positive eigenvalues. Must it be similar to a positive definite matrix?

d) If $A$ is pd and $S$ is a nonsingular matrix, is $S^*AS$ necessarily pd?

e) How do I create a pd matrix $A \in M_{20}(\mathbb{R})$ which does not have a zero entry?

f) Let $f \in \mathcal{C}^2(E)$, $E \subseteq \mathbb{R}^n$ be open and $a \in E$. Suppose that $B_\delta(a) \subseteq E$ and for each $x \in B_\delta(a)$, $x \neq a$, we see that $H(x)$ is a pd matrix. Can we conclude that $H(a)$ is a psd matrix?

**[27.11]** **Exercise(E+)** Let $A \in M_n(\mathbb{C})$ be a positive definite matrix and $\lambda > 0$ be the smallest eigenvalue of $A$. Show that, for each $x \in \mathbb{C}^n$, we have $x^*Ax \geq \lambda x^*x$.