Linear Algebra for Computer Science

Lecture 26

Positive definiteness

Remember: Eigendecomposition of Real Symmetric and Hermitian matrices

LA2

$$\left.\begin{array}{l} A \in \mathbb{R}^{n\times n} \\ A^T = A \end{array}\right\} \quad A = V\Lambda V^T$$

V^T → orthogonal ($V^TV = VV^T = I$)

Λ → diagonal

$\Lambda, V \in \mathbb{R}^{n\times n}$

$$\left.\begin{array}{l} A \in \mathbb{C}^{n\times n} \\ A^* = A \end{array}\right\} \quad A = V\Lambda V^*$$

$\cancel{A}\ \Lambda \in \mathbb{R}^{n\times n}$ diagonal

$V^*V = VV^* = I$

Positive Definite (PD) matrices

K. N. Toosi
University of Technology

Positive definite مثبت معین

$$\boxed{A \succ 0}$$

Symmetric $A^T = A$

$A \in \mathbb{R}^{n \times n}$ positive definite $\begin{cases} A^T = A \\ \forall x \in \mathbb{R}^n,\ x \neq 0 \quad x^T A x > 0 \end{cases}$

Positive Semi-definite (PSD) matrices

K. N. Toosi
University of Technology

Positive Semi-definite

$A^T = A$

$\forall x \in \mathbb{R}^n \quad x^T A x \succcurlyeq 0$

$\boxed{A \succcurlyeq 0}$

$A \succ 0 \Longrightarrow A \succcurlyeq 0$

Positive definite matrices $\subset$ positive semi-definite matrices

Positive negative and semi-negative matrices

K. N. Toosi
University of Technology

Negative definite $A = A^T$

$\forall x \in \mathbb{R}^n \quad x^T A x < 0$

$x \neq 0$

Negativ semi-definite $A = A^T$

$\forall x \in \mathbb{R}^n \quad x^T A x \leq 0$

Definiteness for complex matrices

K. N. Toosi
University of Technology

For $A \in \mathbb{C}^{n \times n}$ positive-difinite $\begin{cases} A^* = A \\ \forall x \in \mathbb{C}^n,\ x \neq 0 \quad x^* A x > 0 \end{cases}$

for complex matrices

Positive definite

Note: Here, by positive-definite we mean symmetric positive definite

PD matrices might have negative (off-diagonal) elements

K. N. Toosi
University of Technology

LA

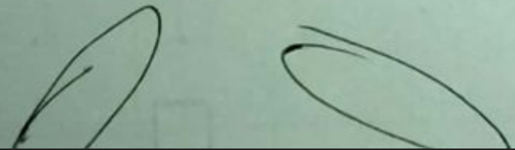
$$A = \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix}$$

$$x^T A x = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 1 & -\varepsilon \\ -\varepsilon & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= x_1^2 + x_2^2 \pm 2\varepsilon x_1 x_2$$

$$x_1^2 + x_2^2 - 2x_1x_2 = (x_1 \mp x_2)^2$$

$$x_1^2 + x_2^2 + 2x_1x_2 = (x_1 + x_2)^2$$



Positive Definite matrices are non-singular

K. N. Toosi
University of Technology

LA26

$A \succ 0$ (A is PD)

Can A be singular?

A singular $\Rightarrow \exists\, x \neq 0$, $x \in \mathbb{R}^n$ $\quad Ax = 0 \Rightarrow x^T A x = 0$

$\Rightarrow$ A is not PD

تناقض!

All PD matrices are non-singular

Does $x^T A x = 0$ imply singularity?
Case1: A symmetric

for some $x \neq 0$ we have $x^T A x = 0$
& ($A = A^T$) A symmetric. Is A singular.

$$A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \quad x = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad x^T A x = 0$$

Does $x^T A x = 0$ imply singularity?
Case 2: A is PSD

A: positive-semidefinite, $x^T A x = 0$ for some $x \neq 0$

is A singular?

$x^T V \Lambda V^T x = 0 \Rightarrow y \neq 0 \quad y^T \Lambda y = 0$

Eigenvalues of PD and PSD matrices

A symmetric $\Rightarrow A = V\Lambda V^T \Rightarrow Av_i = \lambda_i v_i$

A is PD $\Rightarrow$ What about it's eigenvalues?

Let λ be an eigenvalue of A. & $v \neq 0$ be ~~the~~ a corresponding eigenvector.

$$\left.\begin{aligned} Av=\lambda v &\Rightarrow v^TAv = \lambda v^Tv \\ A \text{ is PD} &\Longrightarrow v^TAv>0 \\ v\neq 0 &\Longrightarrow v^Tv = \|v\|^2>0\end{aligned}\right\} \lambda>0$$

A is PD $\Longrightarrow$ All eigenvalues (>0) are positive.

A is PSD $\Longrightarrow$ All eigenvalues (≥ 0) are nonnegative.

PD ⇔ Positive Eigenvalues

LA26 III

$$\left.\begin{array}{l} A \text{ real \& symmetric} \\ \text{all eigenvalues are positive} \end{array}\right\} \Longrightarrow A = V\Lambda V^T$$

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \qquad \lambda_i > 0 \quad i = 1 - n$$

For any $x \neq 0 \in \mathbb{R}^n$

$$x^T A x = x^T V \Lambda V^T x = (V^T x)^T \Lambda (V^T x)$$

Let $y = V^T x \Rightarrow y \neq 0$

$x = Vy$

$x = 0$

$$x^T A x = y^T \Lambda y \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$y_1^2 \lambda_1 + y_2^2 \lambda_2 + \cdots + y_n^2 \lambda_n > 0$$

$\Rightarrow A$ positive definite

PD ⇔ Positive Eigenvalues

K. N. Toosi
University of Technology

LA 26 (III)

A real & symmetric, all eigenvalues are positive $\Longrightarrow A = V \Lambda V^T$

$$\Lambda = \begin{bmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{bmatrix} \qquad \lambda_i > 0 \quad i = 1 - n$$

For any $x \neq \emptyset \in \mathbb{R}^n$

$$x^T A x = x^T V \Lambda V^T x = (V^T x)^T \Lambda (V^T x)$$

let $y = V^T x \Rightarrow y \neq 0$

$x = Vy$

$x = 0$

$$x^T A x = y^T \Lambda y \qquad y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$$

$$y_1^2 \lambda_1 + y_2^2 \lambda_2 + \cdots + y_n^2 \lambda_n > 0$$

$\Longrightarrow A$ positive definite

PSD ⇔ Nonnegative Eigenvalues

A symmetric matrix A is positive definite $\iff$ All eigenvalues are positiv ~~p~~

~~symmetric~~

$A=A^T$ is PSD $\iff$ all eigenvalues are nonnegative.

When is A^TA PD?

LA 26 IV

$C = A^TA \in \mathbb{R}^{n\times n}$, $A \in \mathbb{R}^{m\times n}$, A has full column rank.

$\Downarrow \qquad \Downarrow$

~~rank~~

$\mathrm{rank}(A) = n, \quad m \geq n$

$\Downarrow$

there are n independent rows

$$A = \begin{bmatrix} a_1^T \\ a_2^T \\ \vdots \\ a_m^T \end{bmatrix} \qquad A^T = \begin{bmatrix} a_1 & a_2 & \cdots & a_m \end{bmatrix}$$

$C = A^TA$ A has full column rank $\Rightarrow \boxed{Ax \neq 0}$ for all $x \neq 0$

Choose any $x \in \mathbb{R}^n, x \neq 0 \Rightarrow$

$$x^TCx = x^TA^TAx \Rightarrow (Ax)^T(Ax) = \|Ax\| > 0$$

A^TA is positive definite.

The Correlation Matrix is PSD

Assume $\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_m \in \mathbb{R}^n$, $\text{span}(a_1, - , a_m) = \mathbb{R}^n$

(there are n independent vector among $\vec{a}_1 - \vec{a}_m$)

Let $A^T = [a_1 \ a_2 - a_m]$.

$C = A^T A = \sum a_i a_i^T \in \mathbb{R}^{n \times n}$ is positive-definite

C is called the correlation matrix made of $a_1, a_2, \ldots, a_n$.

The Covariance Matrix is PSD

Let $a_1, a_2, \ldots, a_m \in \mathbb{R}^n$, $\vec{\mu} = \frac{1}{m}\sum_{i=1}^{m} a_i$,

$\bar{a}_i = a_i - \mu$

The correlation matrix of $\bar{a}_1, \bar{a}_2, \ldots, \bar{a}_n$ is called the covariance matrix.

Covariance Matrix and Distribution of Data

K. N. Toosi
University of Technology

LA 26 (V)

$a_1, a_2, \ldots, a_m \qquad \mu = \frac{1}{m}\sum_{i=1}^{m} a_i$

$\bar{a}_i = a_i - \mu$

$$\Sigma = \sum_{i=1}^{m} \bar{a}_i \cdot \bar{a}_i^T = \sum_{i=1}^{m} (a_i - \mu)(a_i - \mu)^T$$

$\downarrow$

covariance matrix is positive semidefinite

$\text{span}(\bar{a}_1, \ldots, \bar{a}_m) = \mathbb{R}^n \Rightarrow \Sigma$ is PD

Covariance Matrix and Distribution of Data

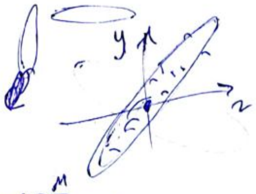
K. N. Toosi
University of Technology

$$\text{data} = D = \begin{bmatrix} d_1^T \\ d_2^T \\ d_3^T \\ \vdots \\ d_M^T \end{bmatrix}$$

Correlation Matrix

$$C = \frac{1}{M}\sum_{i=1}^{M} \underbrace{d_i\, d_i^T}_{n\times n} = \frac{1}{M} D^T D$$

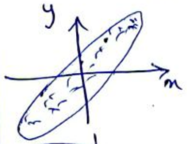
$$D^T D = \begin{bmatrix} d_1 & d_2 & \cdots & d_M \end{bmatrix} \begin{bmatrix} d_1^T \\ d_2^T \\ \vdots \\ d_M^T \end{bmatrix}$$



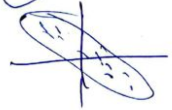
$$d_i^T = \begin{bmatrix} x_i & y_i & z_i \end{bmatrix}$$

$$\mathrm{Corr}(x,y) = \frac{1}{M}\sum_{i=1}^{M} x_i\, y_i$$

$$\begin{bmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ & \vdots & \\ x_M & y_M & z_M \end{bmatrix}$$



$\mathrm{Corr}(x,y) > 0$



$\mathrm{Corr}(x,y) < 0$

$\mathrm{Corr}(x,y)\ \cancel{D} = 0$

Decomposition of PSD matrices

Any positive semi-definite matrix P can be decomposed as $P = A^T A$. $A \in \mathbb{R}^{m \times n}$

Orthogonal Abmiguity in the Decomposition of PSD matrices

K. N. Toosi
University of Technology

$$P = A^T \underline{A} \qquad A \in \mathbb{R}^{n\times n}$$

Let $H \in \mathbb{R}^{n\times n}$ is orthogonal

$$P = A^T A = A^T \underline{I} A = A^T H^{-1} H A = A^T \underline{H^T H} A$$

$$= (\underline{HA})^T (\underline{HA})$$

$$= A'^T A'$$

Square root of a PSD matrix

For a (symmetric) positive semi-definite matrix A there is a unique positive semi-definite matrix P such that A = P P (= P^H P). P is called the square root of A and is denoted by $A^{-\frac{1}{2}}$.

Cholskey Decomposition

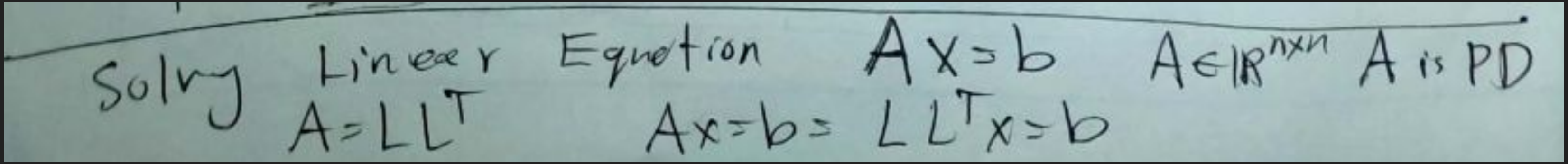
K. N. Toosi
University of Technology

Cholskey Decomposition

Every positive semi definite matrix $A \in \mathbb{R}^{n \times n}$ can be decomposed as $A = LL^T$ where L is ~~lower-~~ lower-triangular.

$A \in \mathbb{C}^{n \times n}$ $\quad A = LL^H = LL^*$.

Solving A x = b with Cholesky Decompotion When A is PD



The Cholesky decomposition can be computed much faster than the LU decomposition for a PSD matrix.

To Solve $A\, x = L\, L^T\, x = b$, let $y = L^T x$. First solve for $L\, y = x$; Then solve for $L^T x = y$ (similarly to LU decomposition).