

## Diagonalization of Real Symmetric Matrices

A real symmetric matrix  $A_{n \times n}$  can be written as:

$$A = \sum_{i=1}^n \lambda_i u_i u_i^T \quad \begin{matrix} \leftarrow \text{orthonormal e.vectors} \\ \leftarrow \text{eigenvalues (real)} \in \mathbb{R} \end{matrix}$$

$$= U \Lambda U^T$$

where

$$U = [u_1 \ u_2 \ \dots \ u_n], \quad \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$$

Note  $U^T U = I \Rightarrow U^T = U^{-1}$   $U$  is an orthogonal matrix

Result For any  $x \in \mathbb{R}^n$ ,

$$\lambda_{\min} \|x\|^2 \leq x^T A x \leq \lambda_{\max} \|x\|^2$$

Proof Since  $u_i$ 's are orthonormal,

$$x = \sum_{i=1}^n \alpha_i u_i \text{ for some } \alpha_i \in \mathbb{R}.$$

$$\Rightarrow x^T A x = \left( \sum_{i=1}^n \alpha_i u_i \right)^T A \left( \sum_{j=1}^n \alpha_j u_j \right)$$

$$= \sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j u_i^T A u_j = \lambda_i \alpha_i$$

$$= \sum_{i=1}^n \alpha_i^2 \lambda_i$$

$$\Rightarrow \lambda_{\min} \underbrace{\sum_{i=1}^n \alpha_i^2}_{\|x\|^2} \leq x^T A x \leq \lambda_{\max} \underbrace{\sum_{i=1}^n \alpha_i^2}_{\|x\|^2}$$

"=" achieved

if  $x = \text{e.vector}$   
for  $\lambda_{\min}$

"=" achieved

if  $x = \text{e. vector}$   
for  $\lambda_{\max}$

$$\underline{\text{Trace}} \quad A_{n \times n}, \quad \text{Tr}(A) = \sum_{i=1}^n A_{kk}$$

Invariance property :  $A_{m \times n}, B_{n \times k}, C_{k \times m}$

$$\text{Tr}(ABC) = \text{Tr}(CAB) = \text{Tr}(BCA)$$

Determinant and Trace in terms of eigenvalues:  $A_{n \times n}$

$$\det(A) = \prod_{i=1}^n \lambda_i, \quad \text{Tr}(A) = \sum_{i=1}^n \lambda_i$$

### Definitions (Definiteness)

Let  $A_{n \times n}$  be a symmetric matrix. Then  $A$  is:

- Positive definite (PD):  $x^T A x > 0, \forall x \in \mathbb{R}^n, x \neq 0$
- Positive semi-definite (PSD):  $x^T A x \geq 0, \forall x \in \mathbb{R}^n$
- Negative definite (ND):  $x^T A x < 0, \forall x \in \mathbb{R}^n, x \neq 0$
- Negative semi-definite (NSD):  $x^T A x \leq 0, \forall x \in \mathbb{R}^n$
- Indefinite:  $x^T A x < 0$  for some  $x$ ,  $> 0$  for some other  $x$

$$A \text{ is ND} \Leftrightarrow -A \text{ is PD} \\ (\text{NSD}) \qquad \qquad \qquad (\text{PSD})$$

Note : We can extend definitions to non-symmetric  $n \times n$  matrices, but

$$x^T A x = (x^T A x)^T = x^T A^T x \Rightarrow x^T A x = x^T \underbrace{\left( \frac{A + A^T}{2} \right)}_{\text{Symmetric part of } A} x$$

Tests for PD Symmetric matrix A is PD

$\Leftrightarrow$  All eigenvalues of A are  $> 0$

$\Leftrightarrow$  All leading principal minors of A are  $> 0$

Example

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

We saw in lec 1 that  $\lambda_1 = 3$ ,  $\lambda_2 = \lambda_3 = 1$

$\Rightarrow$  A is PD

Leading p.m. test:

$$\det([1]) = 1 > 0 \quad \det\left(\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\right) = 2 > 0$$

$$\det(A) = 3 > 0 \quad \Rightarrow \text{A is PD}$$

Tests for PSD Symmetric matrix A is PSD

$\Leftrightarrow$  All eigenvalues of A are  $\geq 0$

$\Leftrightarrow$  All principal minors of A are  $\geq 0$

Leading not enough)

Examples

$$\begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \cdot \lambda_1 = 1, \lambda_2 = 0 \Rightarrow \text{PSD}$$

• also by p.m. test

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix} \stackrel{?}{=} \text{PSD}$$

$$\begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \cdot \lambda_1 = 0, \lambda_2 = -1 \text{ not PSD}$$

• leading p.m.'s all  $\geq 0$

• include all p.m.'s:

$$\det(I - A) < 0 \Rightarrow \text{not PSD}$$

## Principal Minor test for ND, NSD

Test  $-A$  for PD, PSD.

### Examples

$$1. \quad A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix} \quad -A = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \text{ is PSD}$$

$\Rightarrow A \text{ is NSD}$

$$2. \quad B = \begin{bmatrix} -3 & -3 & 0 \\ -3 & -10 & -1 \\ 0 & -1 & -8 \end{bmatrix} \quad \det([-3]) < 0$$

$\Rightarrow B \text{ is not PSD}$

$$\begin{aligned} -B &= \begin{bmatrix} 3 & 3 & 0 \\ 3 & 10 & 1 \\ 0 & 1 & 8 \end{bmatrix} \quad \det([3]) > 0 \\ &\quad \det\left(\begin{bmatrix} 3 & 3 \\ 3 & 10 \end{bmatrix}\right) > 0 \\ \det(-B) &= 3(79) - 3(24) > 0 \\ \Rightarrow -B &\text{ is PD} \Rightarrow B \text{ is ND}. \end{aligned}$$

$$3. \quad C = \begin{bmatrix} 1 & 3 \\ 3 & 1 \end{bmatrix} \quad \text{leading p.m.'s: } \det([1]) > 0$$

$\det(C) = -8 < 0$

$\Rightarrow C \text{ is not PSD}$

$$\begin{aligned} -C &= \begin{bmatrix} -1 & -3 \\ -3 & -1 \end{bmatrix} \quad \det([-1]) < 0 \\ &\Rightarrow -C \text{ is not PSD} \\ &\Rightarrow C \text{ is not NSD} \end{aligned}$$

$C$  is neither PSD nor NSD

$C$  is indefinite

Matrix Norm (Induced Norm) Given any vector norm in  $\mathbb{R}^n$   
 We have the following induced norm on  $A_{n \times n}$ :

$$\|A\| = \max_{\{x \in \mathbb{R}^n : \|x\|=1\}} \|Ax\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

↑ set is compact  $\Rightarrow$  max attained  
 $\|Ax\|$  continuous

Spectral Radius For  $n \times n$  matrix  $A$ ,

$$S(A) = \max_{i=1, \dots, n} |\lambda_i|$$

Result  $S(A) \leq \|A\|$

Proof Let  $\lambda$  be an eigenvalue of  $A$  and  $u$  corresponding unit eigenvector

$$\text{Then } \|Au\| = |\lambda| \|u\| = |\lambda|$$

$$\max_{\{x : \|x\|=1\}} \|Ax\| \geq \|Au\| = |\lambda| \quad \text{f.e.v.'s}(u, \lambda)$$

$$\Rightarrow \|A\| \geq \max_i |\lambda_i| = S(A)$$

Result For symmetric  $A_{n \times n}$ ,  $\|A\| = S(A)$

Proof Since  $S(A) \leq \|A\|$  for all  $A_{n \times n}$ , we need

to show  $S(A) \geq \|A\|$  for symmetric  $A$

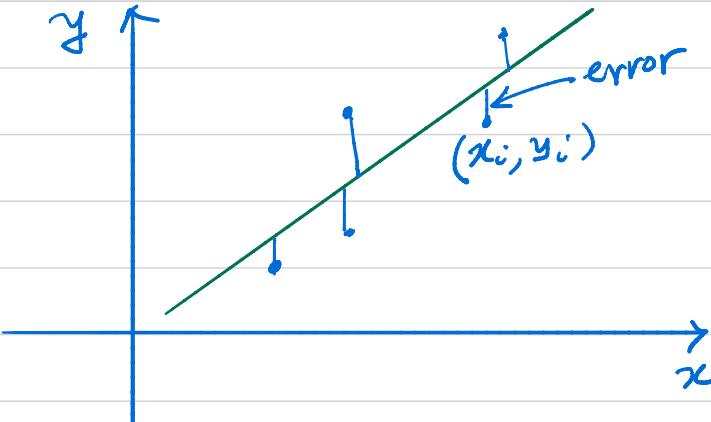
For  $x \in \mathbb{R}^n$ ,  $x = \sum_{i=1}^n \alpha_i u_i$  ← orthonormal e. vectors

$$\|Ax\|^2 = \left\| \sum_{i=1}^n \alpha_i A u_i \right\|^2 = \left\| \sum_{i=1}^n \alpha_i \lambda_i u_i \right\|^2$$

$$= \sum_{i=1}^n (\alpha_i)^2 |\lambda_i|^2 \leq \underbrace{\left( \sum_{i=1}^n (\alpha_i)^2 \right)}_{\|x\|^2} S(A)^2$$

## Examples of Optimization Problems

### 1. Least Squares (Linear Regression) :



Need to find "best" linear fit for data of form:

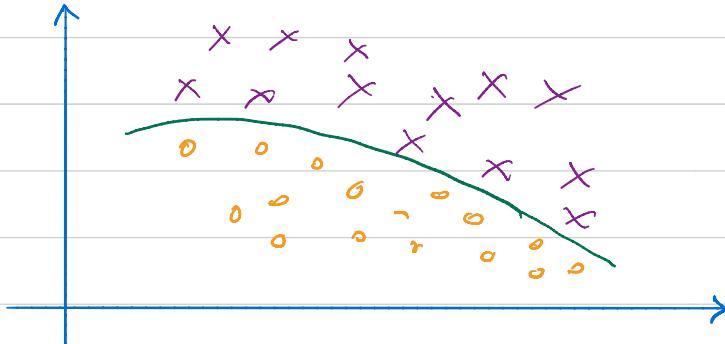
$$y_i = \underline{a}^T x_i + b$$

$\uparrow$                      $\uparrow$   
scalar                vector

Residual Sum of Squares (RSS) : 
$$\sum_{i=1}^N (y_i - (\underline{a}^T x_i + b))^2$$

Optimization Problem: 
$$\min_{\underline{a}, b} \text{RSS}(\underline{a}, b)$$

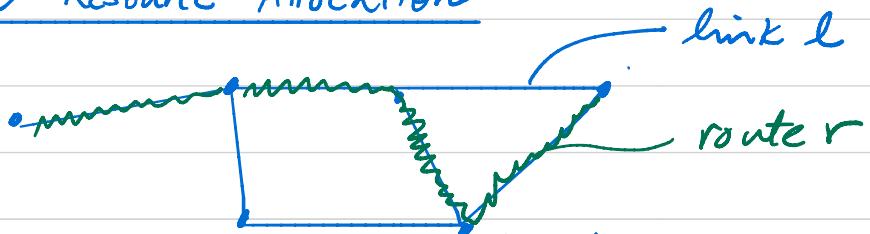
### 2. Supervised Learning - Binary classification



Goal: Learn "best" boundary separating classes

Examples: Support Vector Machine, Neural Network  
posed as optimization problems

### 3. Internet Resource Allocation



$C_l$  : Capacity of link  $l$  bits/s

$l \in r$  if link  $l$  belongs to route  $r$

$x_r$  : data rate (bits/s) for files on route  $r$

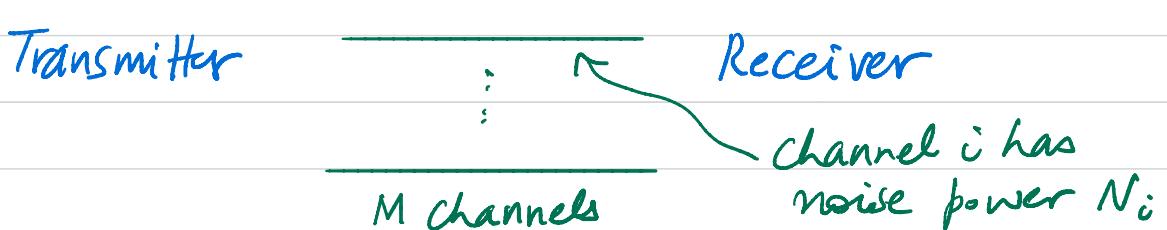
$U_r(x_r)$  : Utility to user on route  $r$  due to transmission at rate  $x_r$

Optimization Problem :  $\max_{\{x_r\}} \sum_r U_r(x_r)$

$$\text{s.t. } \sum_{r: l \in r} x_r \leq C_l + l$$

$$x_r \geq 0 \quad \forall r$$

### 4. Power Allocation in Wireless Networks



Total power  $P$  needs to be allocated on  $M$  channels

$$P_i = \text{Power on channel } i \Rightarrow \text{Rate on } i = \log\left(1 + \frac{P_i}{N_i}\right)$$

$$\text{Optimization Problem: } \max \sum_{i=1}^M \log\left(1 + \frac{P_i}{N_i}\right)$$

$$\text{s.t. } \sum P_i \leq P, \quad P_i \geq 0 \quad \forall i$$

## Conditions for Optimality

Consider  $\min_{x \in S}$  or  $\max_{x \in S} f(x)$   $x \in \mathbb{R}^n$   $S \subseteq \mathbb{R}^n$

$f$ : real-valued objective function

$S$ : constraint set

$x^* \in S$  is called a global minimum if

$$f(x^*) \leq f(x) \quad \forall x \in S$$

$x^* \in S$  is a unique global min. if

$$f(x^*) < f(x) \quad \forall x \neq x^*$$

$x^* \in S$  is a local min. if  $\exists \epsilon > 0$  s.t.

$$f(x^*) \leq f(x) \quad \forall x \in S: \|x - x^*\| < \epsilon$$

(Inequalities reversed for max)

All global min. are local min. but not all local min. are global min.

