SDS 387
Linear Models

Fall 2024

Lecture 12 - Tue, Oct 3, 2024

Instructor: Prof. Ale Rinaldo

- Project proposal: extension to Friday, Oct 11 (next week)
- Also, a correction: over the space of square matrices
$$\langle A, B \rangle = tr(A^T B)$$ defines an innerproduct

- Spectral properties of matrices. Last time we learned about eigenvalues and eigenvectors. For each eigenvalues there exists one or more eigenvectors spanning eigenspaces (linear subspaces).

- If $\underset{n \times n}{A}$ is symmetric, we obtain a convenient variational characterization of eigenvalues. Let $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ be the eigenvalues ordered in a decreasing manner. Then the Courant-Fischer-Weyl min-max theorem says that

$$\lambda_i = \max_{\substack{S \subseteq \mathbb{R}^n \\ \mathrm{rank}(S) = i}} \min_{\substack{x \in S \\ \|x\|=1}} x^T A x$$

linear subspace

$$= \min_{\substack{T \subseteq \mathbb{R}^n \\ \mathrm{rank}(T) = n-i+1}} \max_{\substack{x \in T \\ \|x\|=1}} x^T A x \quad \mapsto \sum_{i,j} x_i x_j A_{ij}$$

linear subspace

As a result:

assuming all eigenvalues are distinct

$$\lambda_n = \lambda_{min}(A) = \min_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} x^T A x$$

$\lambda_{min}(A)$: smallest eigenvalue

$$\lambda_{n-1} = \min_{\substack{x \in \mathbb{R}^n \\ \|x\|=1 \\ x \perp u_n}} x^T A x$$

u_n: eigenvector corresponding to λ_n

⋮

$$\lambda_1 = \lambda_{max}(A) = \max_{\substack{x \in \mathbb{R}^n \\ \|x\|=1}} x^T A x$$

- $\mathrm{Rank}(A) = \#$ non zero eigenvalues of $\underset{n \times n}{A}$ symmetric

- <u>Spectral Theorem</u>: $\underset{n \times n}{A}$ symmetric and has rank $r \leq n$

Then

$$\underset{n \times n}{A} = \underset{n \times r}{U} \; \underset{r \times r}{\Lambda} \; \underset{r \times n}{U^T}$$

Λ: diagonal

$\underset{n \times r}{U}$ orthonormal columns spanning $C(A)$ $\quad (U^T U = I_r)$

$$= \sum_{i=1}^{r} \lambda_i u_i u_i^T$$

linear combination of rank 1 matrices

λ_i: i^{th} eigenvalue

u_i: ith column of U, eigenvector of λ_i

quadratic form

spsd

- $\underset{n\times n}{A}$ symmetric is positive semidefinite if $x^T A x \geq 0$ $\forall x \in \mathbb{R}^n$

  ↳ matrix equivalent of non-negative number

  negative semidefinite if $x^T A x \leq 0$ $\forall x \in \mathbb{R}^n$

- Covariance matrices are psd. Why? $X \in \mathbb{R}^n$ has

  variance covariance $\Sigma = \mathbb{E}\left[(X - \underset{\text{mean}}{\mu})(X-\mu)^T\right]$

  $= \mathbb{E}[XX^T] - \mu\mu^T$

  For any $c \in \mathbb{R}^n$ $c^T X$ is a r.v. with variance

  $c^T \Sigma c \geq 0$ HW

- If $\underset{n\times n}{A}$ psd has rank $r \leq n$ then

  $\lambda_{r+1}, \ldots, \lambda_n = 0$

- $\text{tr}(A) = \sum_i \lambda_i$ $\det(A) = \prod_i \lambda_i$

  if $A = \Sigma$ covariance matrix

  this is called total variance

- A psd iff $A = QQ^T$

- Singular Value Decomposition

  $\forall x \in \mathbb{R}^n$ $x^T A^T A x = (Ax)^T Ax$
  $= \|Ax\|^2 \geq 0$

  Let $\underset{m\times n}{A}$. Then $\underset{n\times n}{A^T A}$ and $\underset{m\times n}{AA^T}$ are both psd

  and have same non-zero eigenvalues. The singular values of A are positive squared roots of these eigenvalues.

SVD

Let $\underset{m\times n}{A}$ have rank $r \leq \min\{m,n\} = q$

There exist orthogonal matrices $\underset{m\times m}{U}$ and $\underset{n\times n}{V}$ and a diagonal matrix

$$\underset{q\times q}{\Sigma} = \begin{bmatrix} \sigma_1 & & 0 \\ & \ddots & \\ 0 & & \sigma_q \end{bmatrix}$$

with $\underbrace{\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r}_{\text{positive}} > \underbrace{\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_q = 0}_{0}$

and such that

$$A = U \underset{m\times n}{\Lambda} V^T$$

where

$$\Lambda = \begin{cases} \Sigma & \text{if } m = n \\ \begin{bmatrix} \Sigma & \underset{n-m}{0} \end{bmatrix} & q = m \\ \begin{bmatrix} \Sigma \\ 0 \end{bmatrix}_{m-n} & q = n \end{cases}$$

In other words:

$$\underset{m\times n}{A} = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

singular values

ith column of V

ith column of U

the u_i's are called left singular vectors. The first n columns of U spans $C(A^T)$

column range

the v_i's are the right singular vectors. The first n columns of V spans $C(A)$

- Remarks:
  - if $\underset{n\times n}{A}$ symmetric is psd then singular values = eigenvalues and right singular vectors = left singular vectors = eigenvectors

    Also, in this case,

    $$\lambda_{max}(A) = \lambda_1 = \sigma_1 = \sigma_{max}(A) = \max_{\|x\|=1} x^T A x$$

  - If $\underset{n\times n}{A}$ is not symmetric then

    $$\sigma_1 = \max_{\|x\|=1} |x^T A x|$$

  - if $\underset{m\times n}{A}$ general

    $$\sigma_1 = \max_{x \in S^{m-1}} \max_{y \in S^{n-1}} x^T A y$$

  - In general eigenvalues $\neq$ singular values

- Powers of $\underset{n\times n}{A}$:

  $$A^k = \underbrace{A \cdot A \cdot \ldots \cdot A}_{k \text{ times}} \longrightarrow \Lambda^k = \begin{bmatrix} \lambda_1^k & & 0 \\ & \lambda_2^k & \\ 0 & & \ddots \\ & & & \lambda_n^k \end{bmatrix}$$

  $$= U \Lambda^k U^T$$

  where $A = U \Lambda U^T$ by spectral theorem

- Square root: if $\underset{n\times n}{A}$ is pd then $\exists!\ \underset{n\times n}{Q}$ pd s.t.

  $$A = Q^2$$

  $$Q = A^{1/2}$$

- Also if $\underset{n\times n}{A} = U \Lambda U^T$ $\quad A^{-1} = U \Lambda^{-1} U^T$

PROJECTION (ONTO A LINEAR SUBSPACE)

In $\mathbb{R}^d$ let S be a linear subspace. For a vector $x \notin S$ the orthogonal projection of x onto S is the (unique) vector $y \in S$ s.t. $y = \underset{z \in S}{\operatorname{argmin}} \|x - z\|$ and $x-y$ is orthogonal to S (i.e. $x-y \in S^\perp$)

Here we are using $\langle x, y \rangle = x^T y$ that induces the Euclidean norm (i.e. $\langle x, x \rangle = \|x\|^2$)

Remark: We can more generally define projections onto a closed convex set S. In this case y (the projection of x onto S) is unique and

$$\langle x-y, y'-y \rangle \leq 0 \quad \forall y' \in S$$

HW