EECS 127 Lecture 5

Feb 4, 2020.

Admin

- Travel Thursday.
- HW due Friday.

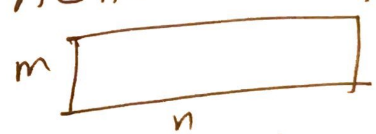
Today:

- Minimum-norm problem.
  - ↳ Pseudo-inverse.
- Matrix norms
- Low-Rank approximation.
- Perturbation analysis / Condition number.

---

① Minimum-Norm.

$A \in \mathbb{R}^{m \times n}$. $\quad m < m$ $\quad A\vec{x} = \vec{b}$



Full row rank.

$$\min \|\vec{x}\|_2$$
$$\text{s.t. } A\vec{x} = \vec{b}$$

$N(A) \oplus R(A^T) = \mathbb{R}^n$

Consider $\vec{x} = \vec{x}_n + \vec{y}$.

$= \vec{x}_n + A^T\vec{z}$

$\vec{x}_n \in N(A)$. $\rightarrow A\vec{x}_n = 0$.

$\vec{y} \in R(A^T) \Rightarrow A^T\vec{z} = \vec{y}$

$$A\vec{x} = A(\vec{x}_n + A^T \cdot \vec{z}) = A \cdot A^T \vec{z}$$

$$\|\vec{x}\|_2^2 = (\vec{x}_n + A^T\vec{z})^T (\vec{x}_n + A^T\vec{z})$$

$$= \|\vec{x}_n\|_2^2 + 2 \underbrace{\langle \vec{x}_n, A^T\vec{z} \rangle}_{\substack{=0 \\ \text{by FTLA.}}} + \|A^T\vec{z}\|_2^2$$

Choose $\vec{x}_n = 0$.

For chosen solution: $A\vec{x} = AA^T\vec{z} = \vec{b}$

$$\Rightarrow \boxed{\vec{z} = (AA^T)^{-1} \cdot \vec{b}}$$

$$\vec{x} = \underline{\underline{\vec{x}_n}} + A^T\vec{z}$$

Choose $\vec{x}_n = 0$ to minimize norm of $\|\vec{x}\|$.

$$\Rightarrow \vec{x} = A^T \cdot \vec{z} = A^T (AA^T)^{-1} \cdot \vec{b}$$

A is full row rank

$\vec{x}_*$ = $A^T (AA^T)^{-1} \cdot \vec{b}$

minimum norn sol^n

$(AB)^T = B^T A^T$

$$\underset{m\times r}{A = U_r}\ \underset{r\times r}{\Sigma_r}\ \underset{r\times n}{V_r^T}$$

$$A^T(AA^T)^{-1} = (U_r \Sigma V_r^T)^T\left(U_r \Sigma_r V_r^T (U_r \Sigma V_r^T)^T\right)^{-1}$$

$$= V_r \Sigma U_r^T \left(U_r \Sigma_r \underbrace{V_r^T V_r} \Sigma U_r^T\right)^{-1}$$

$$= V_r \Sigma U_r^T \left(U_r \Sigma_r^2 U_r^T\right)^{-1}$$

$$= V_r \Sigma \underbrace{U_r^T U_r} \Sigma^{-2} U_r^T$$

$$= V_r \Sigma^{-1} U_r^T$$

$V_r^T V_r$

$\begin{bmatrix} | & | & | & | \end{bmatrix} = V_r$

n

$V_r = \begin{bmatrix} \vec{v}_1^T \\ \text{---} \\ \text{---} \\ \text{---} \end{bmatrix} \begin{bmatrix} \vec{v}_1 & | & | & | & | \end{bmatrix}$

$U_r =$
$m \times r$

$$\underbrace{\left(U_r \Sigma_r^2 U_r^T\right)^{-1}}$$

$$= U_r \Sigma^{-1} U_r$$

$$\underset{m\times r}{U_r}\ \underset{r\times r}{\Sigma_r^2}\ \underset{r\times m}{U_r^T}\ \underset{m\times r}{U_r}\ \underset{r\times r}{\Sigma^{-2}}\ \underset{r\times m}{U_r^T}$$

$$= U_r \Sigma_r^2 I_r \Sigma^{-2} U_r^T$$

$$= U_r \Sigma U_r^T = \cancel{U_m}$$

$$= U_m U_m^T$$

because $r = m$ (full row rank)

$$= I_m.$$

- $V_r \Sigma^{-1} U_r^T$ is called Pseudo-inverse of A.

$A^\dagger$ "dagger"

- <u>Properties</u>:

$$A A^\dagger A = A$$
$$A A^\dagger = U_r U_r^T$$
$$A^\dagger A = V_r V_r^T$$
$$A^\dagger A A^\dagger = A^\dagger$$

- If A is invertible: $A^{-1} = A^\dagger$
- If A is full row rank: $A^\dagger = A^T (A A^T)^{-1}$ (Min-norm solution) "right inverse"
- If A is full col. rank: $A^\dagger = (A^T A)^{-1} A^T$. (Least squares sol^n) "left inverse"

• Low Rank Approximation.

$A \in \mathbb{R}^{m \times n}$

$\underset{m\times m}{A = U}\ \underset{m \times n}{\Sigma}\ \underset{n \times n}{V^T}$ full SVD. $\quad$ $\text{Rank}(A) = r$

Can we find a low-rank q approximation?

◦ Matrix Norms

(1) Giant block of data $\qquad$ (2) Operator.

(1) Frobenius norm.

$$\|A\|_F = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} = \sqrt{\text{trace}(A^T A)}$$

Property: $\|AU\|_F = \|UA\|_F = \|A\|_F$ $\quad$ U is an orthonormal / orthogonal / unitary.

(2) "Spectral" norm, Operator norm, l-2 norm.

$$\|A\|_2 = \max_{\|\vec{x}\|_2 = 1} \|A\vec{x}\|_2 = \max_{\|\vec{x}\|=1} \sqrt{\vec{x}^T A^T A \vec{x}} = \sqrt{\lambda_{max}(A^T A)} = \sigma_{max}(A)$$

Thm: $A \in \mathbb{R}^{m \times n}$, $A = U \Sigma V^T. = \sum_{i=1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T$

$$A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^T \qquad \sigma_1 \geq \sigma_2 \geq \cdots \sigma_k \geq \cdots \geq \sigma_n \geq 0$$

Eckart-Yang-Mirsky Thm

① $$A_k = \underset{\substack{B \in \mathbb{R}^{m \times n} \\ \text{Rank}(B) = k}}{\operatorname{argmin}} \|A - B\|_2$$

② $$A_k = \underset{\substack{B \in \mathbb{R}^{m \times n} \\ \text{Rank}(B) = k}}{\operatorname{argmin}} \|A - B\|_F$$

Proof ①: Consider $\|A - A_k\|_2 = \|\sum_{i=k+1}^{n} \sigma_i \vec{u}_i \vec{v}_i^T\|_2$

$$= \sigma_{k+1}.$$

For any other matrix B, $\text{rank}(B) = k$.

$\|A - B\|_2 \geq \sigma_{k+1}$. $\qquad \text{Rank}(B) = k$, $\dim(\text{Null}(B)) = n - k$

$$\boxed{\|A\|_2 \geq \|A\vec{x}\|_2 \quad \forall \|\vec{x}\|_2 = 1}$$

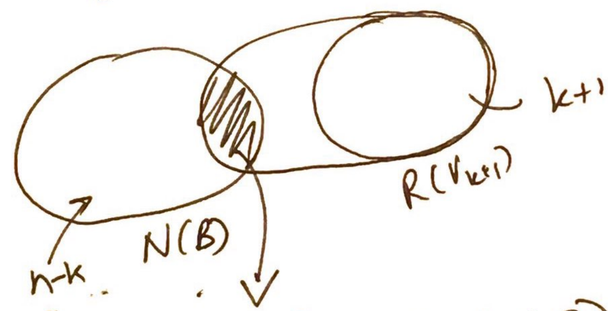
Consider: $\|A - B\|_2^2$

$$\geq \|(A-B)\vec{w}\|_2^2 \qquad \|\vec{w}\|_2 = 1$$

want $\vec{w} \in \text{Null}(B)$.

$$V = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix} \qquad V_{k+1} = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_{k+1} \\ | & & | \end{bmatrix}$$

$$\text{Dim}(\text{Range}(V_{k+1})) = k+1$$



$$\vec{w} \in \text{Null}(B)$$
$$\in \text{Range}(V_{k+1})$$

- Consider: $\|A-B\|_2^2$

$$\geq \|(A-B)\vec{w}\|_2^2$$

$$= \|A\vec{w}\|_2^2 \qquad \vec{w} \in N(B).$$

$$\vec{w} = \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_{k+1} \vec{v}_{k+1} = V \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$= \|U \Sigma V^T (\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \cdots + \alpha_{k+1} \vec{v}_{k+1})\|_2^2$$

$$= \left\| U \Sigma V^T V \cdot \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{k+1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2^2$$

$$= \sigma_1^2 \alpha_1^2 + \sigma_2^2 \alpha_2^2 + \cdots + \sigma_{k+1}^2 \alpha_{k+1}^2 \qquad \|\vec{w}\| = 1$$

$$\Rightarrow \sum \alpha_i^2 = 1$$

$$\geq \sigma_{k+1}^2 (\alpha_1^2 + \cdots + \alpha_{k+1}^2)$$

$$= \sigma_{k+1}^2$$