LECTURE 14

PSEUDOINVERSE (PART II)

$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$ has infinitely many generalized inverses.

e.g.

$$\frac{1}{2}\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{and} \quad \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$$

are generalized inverses.

Recall AA^I and A^IA are projectors. If we require these projectors to be orthogonal, that is

$$(AA^I)^* = AA^I$$
$$(A^IA)^* = A^IA,$$

then the generalized inverse becomes unique.

DEFN (Moore-Penrose Pseudoinverse)

Let $A \in \mathbb{C}^{m \times n}$. The matrix X satisfying

(1) $AXA = A$

(2) $XAX = X$

(3) $(AX)^* = AX$

(4) $(XA)^* = XA$

is called the (Moore-Penrose) pseudoinverse of A.

Any matrix of the form

$$A^{I} = V \begin{bmatrix} \Sigma_r^{-1} & \sqrt{\Sigma_r^{-1}} B_1 \\ B_2 \sqrt{\Sigma_r^{-1}} & B_2 B_1 \end{bmatrix} U^*$$

is a generalized inverse.

(where U, V, Σ_r are s.t. $A = U \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix} V^*$ is an SVD for A)

If we further set $B_1 = B_2 = 0$, that is

$$A^{+} = V \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix} U^*,$$

then

$$A^+A = V\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} V^*$$

$$AA^+ = U\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix} U^*$$

are Hermitian.

NOTATION

A^+ – Pseudoinverse

SUMMARY

Let $A \in \mathbb{C}^{m\times n}$ with SVD

$$A = \underbrace{U}_{m\times m} \underbrace{\begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}}_{m\times n} \underbrace{V^*}_{n\times n}$$

($\Sigma_r \in \mathbb{R}^{r\times r}$ consists of positive singular values

Pseudoinverse

$$A^+ = \underbrace{V}_{n\times n} \underbrace{\begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}}_{n\times m} \underbrace{U^*}_{m\times m}$$

satisfies properties (1) – (4) on the previous page.

EXAMPLE

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

$\boxed{\text{SVD of } A}$

$$= \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)\begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & +1 \end{bmatrix}\right)$$

Pseudoinverse

$$A^+ = \left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & +1 \end{bmatrix}\right)\begin{bmatrix} 1/2 & 0 \\ 0 & 0 \end{bmatrix}\left(\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\right)$$

$$= \frac{1}{4}\begin{bmatrix} 1 & 1 \\ -1 & -1 \end{bmatrix}$$

Orthogonal projectors

$$A A^+ = \frac{1}{4}\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \quad \Big(\text{onto } \underbrace{\operatorname{span}\left\{\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right\}}_{\text{Range}(A)}\Big)$$

$$A^+ A = \frac{1}{4}\begin{bmatrix} 2 & -2 \\ -2 & 2 \end{bmatrix} \quad \Big(\text{onto } \underbrace{\operatorname{span}\left\{\begin{bmatrix} 1 \\ -1 \end{bmatrix}\right\}}_{\text{Range}(A^+)}\Big)$$

THM (Uniqueness)
Every matrix $A \in \mathbb{C}^{m \times n}$ has a unique pseudoinverse $A^+ \in \mathbb{C}^{n \times m}$.

PROOF

Suppose X_1 and X_2 are pseudoinverses.

(1) $\text{Range}(X_2 - X_1) \subseteq \text{Range}(A^*)$

$$X_i A X_i = A^* X_i^* X_i = X_i \qquad i=1,2$$

$$\Longrightarrow$$

$$A^*(X_2^* X_2 - X_1^* X_1) = X_2 - X_1$$

$$\Longrightarrow$$

$$\text{Range}(A^*) \supseteq \text{Range}(X_2 - X_1)$$

(2) $\text{Range}(X_2 - X_1) \subseteq \text{Null}(A)$

$$A X_i A = A \qquad i=1,2$$

$$\Longrightarrow$$

$$A^* X_i^* A^* = A^* \qquad i=1,2$$

$$\Longrightarrow$$

$$A^* A X_i = A^* \qquad i=1,2$$

$$\Longrightarrow$$

$$A^* A (X_2 - X_1) = 0$$

$$\Longrightarrow$$

$$\text{Null}(A^* A) = \text{Null}(A)$$

$$\supseteq$$

$$\text{Range}(X_2 - X_1)$$

NOTE Let $A = U\Sigma V^*$

$\text{Null}(A^* A)$

$=$

$\text{Null}(V \Sigma^2 V^*)$

$=$

$\text{Null}(A)$

But

$$\mathrm{Range}(A^*) \perp \mathrm{Null}(A) \quad \left(\begin{smallmatrix}HW\ 2\\ Q1\end{smallmatrix}\right)$$

Consequently

$$\mathrm{Range}(A^*) \cap \mathrm{Null}(A) = \{0\}$$

$$\supseteq$$

$$\mathrm{Range}(X_2 - X_1)$$

$$\Longrightarrow$$

$$X_2 = X_1$$

$\square$

PROPERTIES OF PSEUDOINVERSE

(1) $(A^+)^+ = A$

(2) $(A^+)^* = (A^*)^+$

(3) $\mathrm{Range}(A^+) = \mathrm{Range}(A^*)$

(4) $\mathrm{Null}(A^+) = \mathrm{Null}(A^*)$

(5) If $U \in \mathbb{C}^{m \times n}$ has orthonormal columns or rows

$$U^+ = U^*$$

(6) If $A \in \mathbb{C}^{m \times n}$ with $\mathrm{rank}(A) = n$

$$A^+ = (A^*A)^{-1}A^*$$

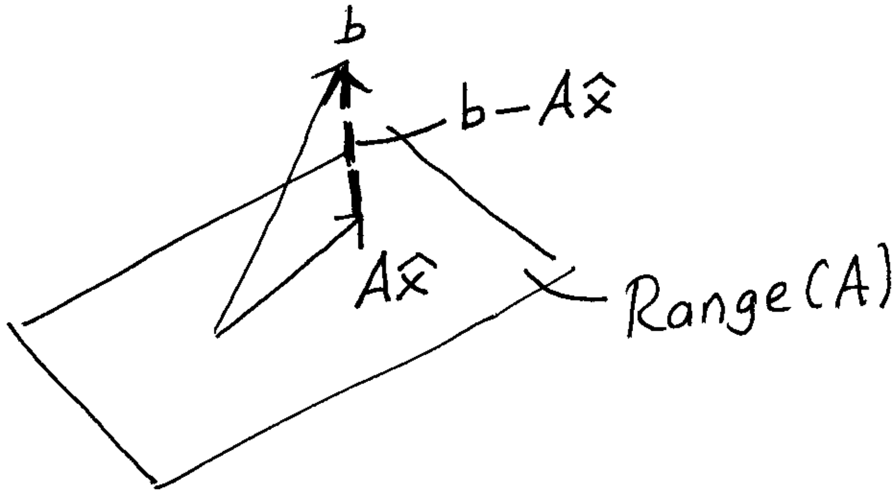
(7) If $A \in \mathbb{C}^{m \times n}$ with $\mathrm{rank}(A) = m$

$$A^+ = A^*(AA^*)^{-1}$$

LEAST SQUARES PROBLEM (REVISITED)

$$\|A\hat{x} - b\|_2 = \underset{x \in \mathbb{C}^n}{\text{minimize}} \|Ax - b\|_2$$

where $A \in \mathbb{C}^{m \times n}$, $b \in \mathbb{C}^m$ with $m > n$.



* $b - A\hat{x} \perp \text{Range}(A)$
* $A\hat{x}$ is the orthogonal projection of b onto Range(A)

$\hat{x}$ is a solution of LSP

$$\Longleftrightarrow$$

$$A\hat{x} = AA^{+}b$$

In particular

$$\hat{x} = A^{+}b$$

is a solution for LSP.

EXAMPLE

Find a solution for

$$\underset{x \in \mathbb{C}^n}{\text{minimize}} \left\| \begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix} - \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} \right\|_2$$

using pseudoinverse.

$\boxed{\text{FULL SVD}}$

$$\underbrace{\begin{bmatrix} 1 & -1 \\ 1 & -1 \\ 1 & -1 \end{bmatrix}}_{A} = \underbrace{\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & 1/\sqrt{6} \\ 1/\sqrt{3} & 0 & -2/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & 1/\sqrt{6} \end{bmatrix}}_{U} \underbrace{\begin{bmatrix} \sqrt{6} & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}}_{\Sigma} \underbrace{\begin{bmatrix} 1/\sqrt{2} & -1/\sqrt{2} \\ 1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix}}_{V^*}$$

$$A^+ = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ -1/\sqrt{2} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{6} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \\ 1/\sqrt{2} & 0 & 1/\sqrt{2} \\ 1/\sqrt{6} & -2/\sqrt{6} & 1/\sqrt{6} \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix}$$

A solution for LSP

$$\hat{x} = A^+ b$$

$$= \frac{1}{6} \begin{bmatrix} 1 & 1 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 2/3 \\ -2/3 \end{bmatrix}$$

THM

Let $\hat{x} = A^{+}b$. Then

$$\|\hat{x}\|_2 \leqslant \|\tilde{x}\|_2 \quad \forall \tilde{x} \quad \text{s.t}$$

$$\|A\tilde{x} - b\| = \min_{x \in \mathbb{C}^n} \|Ax - b\|$$

PROOF

Any solution $\tilde{x}$ of LSP must satisfy

$$(A\tilde{x} - b) \perp A\tilde{x}$$

$$\Longrightarrow$$

$$A\tilde{x} = AA^{+}b$$

Consequently

$$A(\hat{x} - \tilde{x}) = 0$$

$$\Longrightarrow$$

$$(\hat{x} - \tilde{x}) \in Null(A).$$

But

$$\hat{x} = A^{+}b$$

$$\Longrightarrow$$

$$\hat{x} \in Range(A^{+}) = Range(A^{*})$$

meaning $(\tilde{x} - \hat{x}) \perp \hat{x}$ (i.e. $Null(A) \perp Range(A^{*})$).

We deduce

$$\|\tilde{x}\|_2^2 = \|(\tilde{x} - \hat{x}) + \hat{x}\|_2^2 \underset{\text{(PYTHAGOREAN THM)}}{=} \|\tilde{x} - \hat{x}\|_2^2 + \|\hat{x}\|_2^2 \Longrightarrow \|\tilde{x}\|_2 \geqslant \|\hat{x}\|_2 \quad \square$$