EE 203001 SES #36

Diagonalization & pseudoinverse

Transformation matrices

Given a lin. transf. T

A : corr. matrix with basis $\underline{v}_1, \cdots, \underline{v}_n$

B : corr. matrix with basis $\underline{w}_1, \cdots, \underline{w}_n$

Q : Is there a relation between A & B?

$$\underset{\underline{v}\to\underline{v}}{A} = \underset{\underline{w}\to\underline{v}}{M^{-1}}\,\underset{\underline{w}\to\underline{w}}{B}\,\underset{\underline{v}\to\underline{w}}{M}$$

$\Rightarrow$ A & B are similar !

Basis of eigenvectors

If $\underline{v}_i$ are eigenvectors of T

$$\Rightarrow \left.\begin{aligned} T(\underline{v}_1) &= \lambda_1 \underline{v}_1 \\ T(\underline{v}_2) &= \lambda_2 \underline{v}_2 \\ &\vdots \\ T(\underline{v}_n) &= \lambda_n \underline{v}_n \end{aligned}\right) \Rightarrow A = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} = \Lambda$$

or

$$\underset{\underline{v}\to\underline{v}}{\Lambda} = \underset{std\to\underline{v}}{S^{-1}}\;\underset{std\to std}{A}\;\underset{\underline{v}\to std}{S}$$

Basis of singular vectors (SVD)

Input basis : $\underline{v}_1, \cdots, \underline{v}_n \subseteq \mathbb{R}^n$

Output basis : $\underline{u}_1, \cdots, \underline{u}_m \subseteq \mathbb{R}^m$

$$\underset{\underline{v} \to \underline{u}}{\Sigma} = \underset{std \to \underline{u}}{U^{-1}} \quad \underset{std \to std}{A} \quad \underset{\underline{v} \to std}{V}$$

(The SVD chooses the orthonormal bases, $U^{-1} = U^T$, $V^{-1} = V^T$ that diagonalize A)

(The two orthonormal basis are eigenvectors of A^TA, $\underline{v}$'s &
AA^T, $\underline{u}$'s)

Left & right inverses ; pseudoinverse

Two sided inverse

If $m = n = r$, A has full rank

two-sided inverse A^{-1} exists

$\Rightarrow A^{-1}A = I$, $AA^{-1} = I$

We called it "inverse" of A

$A\underline{x} = \underline{b}$ has 1 sol. (unique)

Left inverse ($r=n$)

A has full col. rank, col.s are indep.

$\Rightarrow N(A)$ only contains $\{\underline{0}\}$

$\Rightarrow A\underline{x}=\underline{b}$ has 0 or 1 sol.

A^TA is an invertible PD matrix

$$\Rightarrow \underbrace{(A^TA)^{-1}A^T}A = I$$

left inverse: $A^{-1}_{left} = (A^TA)^{-1}A^T$

Note: left inverse is NOT unique but this is our favorite

Note:

$$AA^{-1}_{left} = I \text{ only when } m=n$$

Otherwise,

$$AA^{-1}_{left} = A(A^TA)^{-1}A^T = P$$

(projection from R^m to $C(A)$)

Right inverse ($r=m$)

A has full row rank

$\Rightarrow$ All rows have pivots, $C(A)=R^m$

$\Rightarrow \dim N(A) = n-m$

$\Rightarrow$ $n-m$ free variables (if $n>m$)

$\Rightarrow A\underline{x} = \underline{b}$ has ∞ sol.s

AA^T is an invertible PD matrix

$$\Rightarrow A\underbrace{A^T(AA^T)^{-1}}_{\text{right inverse}} = I$$

right inverse $A_{right}^{-1} = A^T(AA^T)^{-1}$

Note: right inverse is NOT unique
but this is our favorite

Note:

$$A_{right}^{-1}A = I \text{ only when } m=n$$

Otherwise,

$$A_{right}^{-1}A = A^T(AA^T)^{-1}A = P$$

(projection from R^n to $C(A^T)$)

<u>Pseudoinverse</u> $(r<n, r<m)$

Recall:

$$AA_{left}^{-1} = A(A^TA)^{-1}A^T = P$$

(projection from R^m to $C(A)$)

(trouble: nontrivial $N(A^T)$)

$$A_{right}^{-1}A = A^T(AA^T)^{-1}A = P$$

(projection from R^n to $C(A^T)$)

(trouble: nontrivial $N(A)$)

Q: Can we avoid this?

Yes! Focus on $C(A)$ & $C(A^T)$

$\boxed{\text{Fact}}$ If $\underline{x} \neq \underline{y}$ are vectors in $C(A^T)$

then $A\underline{x} \neq A\underline{y}$ in $C(A)$

($\underline{x} \in C(A^T) \xrightarrow{A} A\underline{x} \in C(A)$ is

a one-to-one mapping

$\Rightarrow$ inverse exists! $\Rightarrow$ pseudoinverse)

Proof: (by contradiction)

If $\exists\ \underline{x} \neq \underline{y} \in C(A^T)$ & $A\underline{x} = A\underline{y}$

$\Rightarrow A(\underline{x} - \underline{y}) = \underline{0} \Rightarrow \underline{x} - \underline{y} \in N(A)$

But $\underline{x}, \underline{y} \in C(A^T) \Rightarrow \underline{x} - \underline{y} \in C(A^T)$

Since $N(A) \perp C(A^T) \Rightarrow \underline{x} - \underline{y} = 0$

$\Rightarrow \underline{x} = \underline{y}$ (contradiction!)

Finding pseudoinverse A^+

$\boxed{\text{Def}}$ Pseudoinverse A^+ of A is a matrix

for which $\underline{x} = A^+ A \underline{x} \quad \forall\, \underline{x} \in C(A^T)$

Note: $N(A^+) = N(A^T)$ —— (*)

(show later)

Q: How to find A^+?

Recall from SVD:

$$A\underline{v_i} = \sigma_i \underline{u_i} \quad , i=1,\cdots,r$$

(basis for $C(A^T)$) (basis for $C(A)$)

$$\Rightarrow \quad A^+\underline{u_i} = 1/\sigma_i \, \underline{v_i} \quad , i=1,\cdots,r$$

(basis for $C(A)$) (basis for $C(A^T)$)

& $A^+\underline{u_i} = 0$ for $i > r$ (basis for $N(A^T)$)

(we know what happens to each (*)
basis vector $\underline{u_i} \in \mathbb{R}^m$, we know A^+)

In summary:

If $A = U\Sigma V^T \Rightarrow A^+ = V\Sigma^+U^T$

$$= \begin{bmatrix} \underline{v_1} \cdots \underline{v_r} \cdots \underline{v_n} \end{bmatrix} \begin{bmatrix} \sigma_1^{-1} & & & & \\ & \ddots & & & \\ & & \sigma_r^{-1} & & \\ & & & 0 & \\ & & & & \ddots \; 0 \end{bmatrix} \begin{bmatrix} \underline{u_1} \cdots \underline{u_r} \cdots \underline{u_m} \end{bmatrix}$$

$n\times n \qquad n\times m \qquad m\times m$

(If $n=m$, $\Sigma^+ = \Sigma^{-1}$, $A^+ = A^{-1}$)

Ex: (p.404) Find pseudoinverse for

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

$$A^TA = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$$

$\Rightarrow \lambda_1 = 10,\ \lambda_2 = 0$

$A^TA - 10I = \begin{bmatrix} -5 & 5 \\ 5 & -5 \end{bmatrix} \Rightarrow \underline{v_1} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$A^TA = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix} \Rightarrow \underline{v_2} = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$

$\Rightarrow A\underline{v_1} = \sigma_1 \underline{u_1} \Rightarrow \underline{u_1} = \frac{1}{\sqrt{10}}\begin{bmatrix} 2\sqrt{2} \\ \sqrt{2} \end{bmatrix} = \frac{1}{\sqrt{5}}\begin{bmatrix} 2 \\ 1 \end{bmatrix}$

$\Rightarrow A^+ = V\Sigma^+U^T = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}\begin{bmatrix} \frac{1}{\sqrt{10}} & \\ & 0 \end{bmatrix}\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$

$= \frac{1}{10}\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$

A^+ has rank 1

A takes $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in $C(A^T)$ to $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ in $C(A)$

A^+ takes $\begin{bmatrix} 4 \\ 2 \end{bmatrix}$ in $C(A)$ to $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ in $C(A^T)$

Application to least square sol.

Recall: To find least square sol. in ch. 4., we have

$$A^TA\hat{x} = A^T\underline{b}$$

with assumpt. A has full col. rank

(so A^TA is invertible)

Now, if A has dependent col.s there are many sol.s to

$$A^T A \hat{\underline{x}} = A^T \underline{b}$$

One sol.: $\underline{x}^+ = A^+ \underline{b}$

(chk: $A^T A A^+ \underline{b} = A^T \underline{b}$

since $\underline{e} = \underline{b} - AA^+\underline{b} \in N(A^T)$)

(see Fig below)

Note: Any vector in $N(A)$ can be added to $\underline{x}^+$ to give another sol. $\hat{\underline{x}}$ but $\underline{x}^+$ is the shortest least square sol. (one with shortest length)

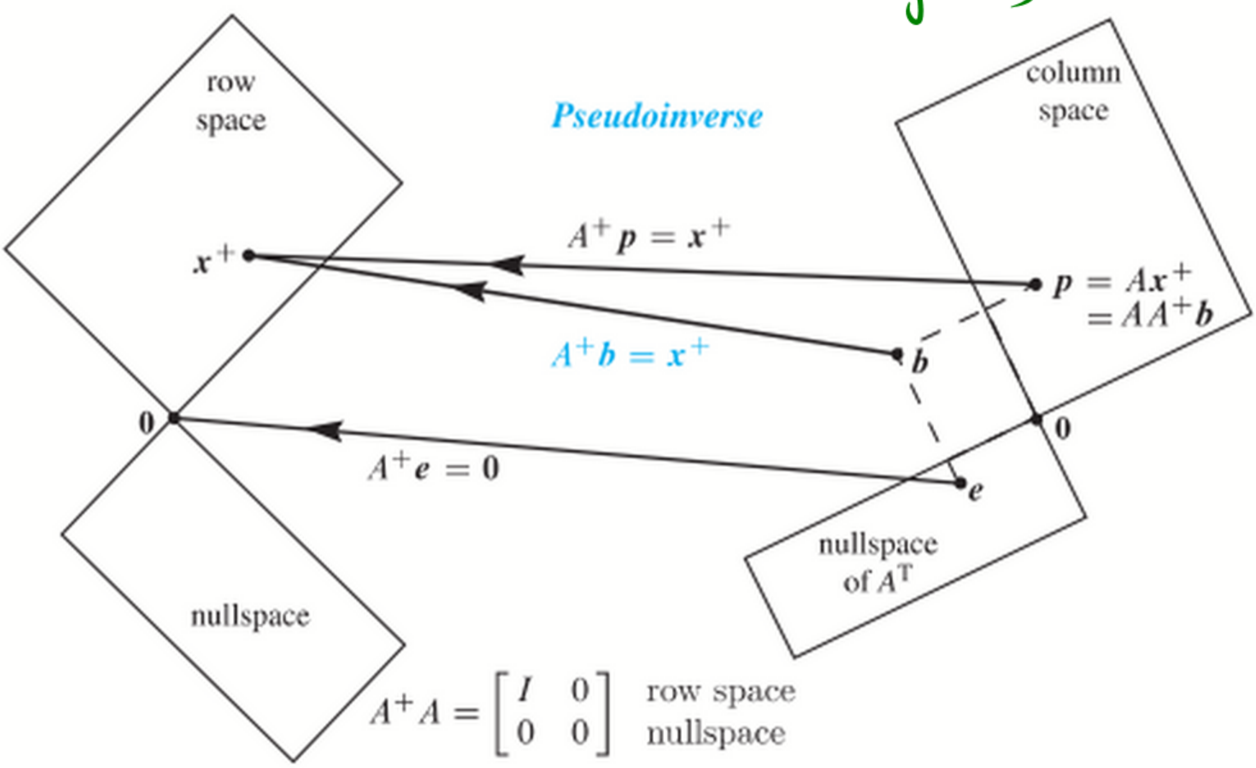

Figure 50: Ax^+ in the column space goes back to $A^+Ax^+ = x^+$ in the row space.