**L05 Moore-Penrose inverses**

1. Relations of inverses

   (1) Inverses
   Three types of inverses: $A^L$, $A^R$, $A^{-1}$. They may not exist. When they are existent, $A^{-1}$ is unique, but the other two may not be unique.
   Two types of generalized inverses, $A^-$ and $A^+$. They all exist. $A^+$ is unique. $A^+ \in A^-$.

   (2) Three cases

   (i) If $A^L \neq \emptyset$, $A^+ \in A^- = A^L$. (No assumption on the existence for $A^R$ and $A^{-1}$.)
   (ii) If $A^R \neq \emptyset$, $A^+ \in A^- = A^R$.
   (iii) If $A^{-1}$ exists, $A^{-1} = A^+ = A^- = A^L = A^R$.

   **Pf:** (i) $A^+ \in A^-$: trivial.
   $A^- \subset A^L$: $B \in A^- \Longrightarrow ABA = A \Longrightarrow A^L ABA = A^T A \Longrightarrow BA = I$
   $\Longrightarrow B \in A^L$.
   $A^- \supset A^L$: $B \in A^L \Longrightarrow BA = I \Longrightarrow ABA = A \Longrightarrow B \in A^-$.
   (iii) By (i) and (ii) $A^- = A^L = A^R$. Now we show $A^+ = A^- = A^{-1}$.
   $A^+ \in A^-$: trivial.

   $$A^- \subset A^{-1}: B \in A^- \Longrightarrow ABA = A \Longrightarrow \begin{cases} A^{-1}ABA = A^{-1}A \\ ABAA^{-1} = AA^{-1} \end{cases}$$

   $$\Longrightarrow \begin{cases} BA = I \\ AB = I \end{cases} \Longrightarrow B = A^{-1}.$$

   $A^{-1} \in A^+$: $AA^{-1}A = A$, $A^{-1}AA^{-1} = A^{-1}$, $AA^{-1} = I$ is symmetric and $A^{-1}A = I$ is symmetric. So $A^{-1} \in A^+$.

   **Ex1:** If $A \in R^{m\times n}$ has full column rank, then $A^+A = A^-A = A^LA = I_n$.
   If $A \in R^{m\times n}$ has full row rank, then $AA^+ = AA^- = AA^R = I_m$.

   (3) Recall that $\mathrm{rank}(AA') = \mathrm{rank}(A'A) = \mathrm{rank}(A)$.

   (i) If $A$ has full column rank, then $A'A$ is full rank square matrix. So $(A'A)^{-1}$ exists and $(A'A)^{-1}A' = A^+$.
   (ii) If $A$ has full row rank, then $AA'$ is full rank square matrix. So $(AA')^{-1}$ exists and $A'(AA')^{-1} = A^+$.

   **Ex2:** In regression you probably see the claim that $\widehat{\beta} = (X'X)^{-1}X'y$ is a least square estimator for $\beta$. Now we know that $\widehat{\beta} = X^+y$ and we assume $X$ has full column rank so that $\widehat{\beta} = (X'X)^{-1}X'y$.

   (4) If $A$ is idempotent, i.e., $A^2 = A$, then $\mathrm{rank}(A) = \mathrm{tr}(A)$.

   **Pf:** By the compact SVD for $A$ with $\mathrm{rank}(A) = r$, $A = U_I\Delta_rV_I'$. Then

   $$A^2 = A \Longleftrightarrow U_I\Delta_rV_I'U_I\Delta_rV_I' = U_I\Delta_rV_I' \Longleftrightarrow \Delta_rV_I'U_I = I_r.$$

   So $\mathrm{tr}(A) = \mathrm{tr}(U_I\Delta_rV_I') = \mathrm{tr}(\Delta_rV_I'U_I) = \mathrm{tr}(I_r) = r = \mathrm{rank}(A)$.

   **Ex3:** $AA^-$, $A^-A$, $I-AA^-$ and $I-A^-A$ are all idempotent. So, $\mathrm{rank}(A^-A) = \mathrm{tr}(A^-A)$, $\mathrm{rank}(AA^-) = \mathrm{tr}(AA^-)$, $\mathrm{rank}(I-A^-A) = \mathrm{tr}(I-A^-A)$ and $\mathrm{rank}(I-AA^-) = \mathrm{tr}(I-AA^-)$.

2. Simple $A^+$

(1) If $A$ is symmetric and idempotent, i.e., $A' = A = A^2$, then $A^+ = A = AA^+ = A^+A$

**Pf:** With $G = A$, $AGA = AAA = A$; $GAG = AAA = A = G$; $AG = AA = A$ is symmetric; $GA = AA = A$ is symmetric. Thus $A^+ = G = A$ and $AA^+ = A = A^+A$.

**Ex4:** For $A \in R^{m\times n}$, $AA^+$, $A^+A$, $I - AA^+$ and $I - A^+A$ are symmetric and idempotent. So $(AA^+)^+ = AA^+$, $(A^+A)^+ = A^+A$, $(I - AA^+)^+ = I - AA^+$ and $(I - A^+A)^+ = I - A^+A$.

(2) If $A$ has orthonormal columns, i.e., $A'A = I$, then $A^+ = A' \in A^- = A^L$;
If $A$ has orthonormal rows, i.e., $AA' = I$, then $A^+ = A' \in A^- = A^R$.

**Pf:** For the 1st one check condition (ii): Let $B = A'$, then $BAB = A'AA' = A' = B$.

(3) If the columns of $A$ are perpendicular to the columns of $B$, i.e., $A'B = 0$, then
$(A,\, B)^+ = \begin{pmatrix} A^+ \\ B^+ \end{pmatrix}$ so $(A,\, B)(A,\, B)^+ = AA^+ + BB^+$.
If the rows of $A$ are perpendicular to the rows of $B$, i.e., $AB' = 0$, then
$\begin{pmatrix} A \\ B \end{pmatrix}^+ = (A^+,\, B^+)$ so $\begin{pmatrix} A \\ B \end{pmatrix}^+ \begin{pmatrix} A \\ B \end{pmatrix} = A^+A + B^+B$.

**Pf:** For the 1st one show (i). With $A'B = 0$,
$A^+B = (A^+AA^+)B = A^+(AA^+)'B = A^+(A^+)'A'B = 0$. Similarly $B^+A = 0$.
So $(A,\, B)\begin{pmatrix} A^+ \\ B^+ \end{pmatrix}(A,\, B) = (A,\, B)\begin{pmatrix} A^+A & 0 \\ 0 & B^+B \end{pmatrix} = (A,\, B)$.

**Ex5:** $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}^+ = ((A,0)^+,\, (0,B)^+) = \begin{pmatrix} A^+ & 0 \\ 0 & B^+ \end{pmatrix}$.

3. Cases of $(AB)^+ = B^+A^+$

(1) Recall

(i) If $A = 0$, then $(AB)^+ = B^+A^+ = 0$. If $B = 0$, then $(AB)^+ = B^+A^+ = 0$.

(ii) If $A = B'$, i.e., $B = A'$, then $(AB)^+ = B^+A^+ = (A')^+A^+ = B^+(B')^+$.

(2) If $A$ has orthonormal columns, then $(AB)^+ = B^+A^+$.
If $B$ has orthonormal rows, then $(AB)^+ = B^+A^+$.

**Pf:** For 1st one show (iii).
$A$ has orthonormal columns. So $A^+ = A'$ and $A'A = I$.
(iii) $(AB)(B^+A^+) = A(BB^+)A'$ is symmetric.

(3) If $A$ has full column rank and $B$ has full row rank, then $(AB)^+ = B^+A^+$.

**Pf:** For 1st one show (ii).
$A$ has full column rank and $B$ has full row rank. So $A^+A = I$ and $BB^+ = I$.
(ii) $(B^+A^+)(AB)(B^+A^+) = B^+IIA^+ = B^+A^+$.

**L06: Orthogonal complement of space**

1. Two types of spaces

   (1) Two spaces

   For $A \in R^{m\times n}$, $y = Ax$ is a linear transformation of $x \in R^n$ to $y \in R^m$ with range

   $$\text{Range}(A) = \text{R}(A) = \{y = Ax \in R^m : x \in R^n\}.$$

   This range is span of the columns of $A$ also called the column space of $A$ denoted as $\text{Span}(A) = C(A)$ with $\dim[\mathcal{R}(A)] = \text{rank}(A)$.
   The Kernel of the transformation also called the null space of $A$,

   $$\mathcal{N}(A) = \{x \in R^n : Ax = 0\},$$

   is a subspace of $R^n$ with $\dim[\mathcal{N}(A)] = n - \text{rank}(A)$.

   **Comment:** For liner transformation $f$,

   $$\dim[\text{domain}(f)] = \dim[\text{Kernel}(f)] + \dim[\text{Range}(f)].$$

   With $y = Ax$: $n = [n - \text{rank}(A)] + \text{rank}(A)$.

   (2) Expressions of $\mathcal{R}(A)$
   For $A \in R^{m\times n}$, (i) $\mathcal{R}(A) = \mathcal{R}(AA^-)$ (ii) $\mathcal{R}(A) = \mathcal{R}(AA')$ (iii) $\mathcal{R}(A) = \mathcal{R}((A')^+)$

   **Proof.** Note that $\mathcal{R}(AB) \subset \mathcal{R}(A)$ since $y \in \mathcal{R}(AB) \Longrightarrow y = ABx = A(Bx) \in \mathcal{R}(A)$.
   (i) $\mathcal{R}(A) = \mathcal{R}(AA^-A) \subset \mathcal{R}(AA^-) \subset \mathcal{R}(A)$.
   (ii)$\mathcal{R}(A) = \mathcal{R}(AA^+A) = \mathcal{R}\left(A(A^+A)'\right) = \mathcal{R}\left(AA'(A^+)'\right) \subset \mathcal{R}(AA') \subset \mathcal{R}(A)$.
   (iii) $\mathcal{R}(A) = \mathcal{R}\left((AA^+)'\right) \subset \mathcal{R}\left((A^+)'\right) = \mathcal{R}((A^+AA^+)') = \mathcal{R}(AA^+(A^+)') \subset \mathcal{R}(A)$.

   **Ex1:** For $A \in R^{m\times n}$, $\mathcal{R}(A) = \mathcal{R}(AA^+)$ where $AA^+$ is symmetric and idempotent.

   (3) Expressions of $\mathcal{N}(A)$

   For $A \in R^{m\times n}$, (i) $\mathcal{N}(A) = \mathcal{N}(A^-A)$ (ii) $\mathcal{N}(A) = \mathcal{N}(A'A)$ (iii) $\mathcal{N}(A) = \mathcal{N}\left((A^+)'\right)$.

   **Proof.** Note that $\mathcal{N}(A) \subset \mathcal{N}(BA)$ since
   $x \in \mathcal{N}(A) \Longrightarrow Ax = 0 \Longrightarrow BAx = 0 \Longrightarrow x \in \mathcal{N}(BA)$.

   (i) $\mathcal{N}(A) \subset \mathcal{N}(A^-A) \subset \mathcal{N}(AA^-A) = \mathcal{N}(A)$.
   (ii) $\mathcal{N}(A) \subset \mathcal{N}(A'A) \subset \mathcal{N}\left((A^+)'A'A\right) = \mathcal{N}((AA^+)'A) = \mathcal{N}(A)$.
   (iii) $\mathcal{N}(A) = \mathcal{N}((A^+A)') \subset \mathcal{N}\left((A^+)'A'(A^+)'\right) = \mathcal{N}\left((A^+)'\right) \subset \mathcal{N}\left(A'(A^+)'\right)$
   $= \mathcal{N}(A^+A) = \mathcal{N}(A)$.

   **Ex2:** For $A \in R^{m\times n}$, $\mathcal{N}(A) = \mathcal{N}(A^+A)$ where $A^+A$ is symmetric itempotent.

2. Cross expressions

   (1) Condition for cross expression
   If $D$ is idempotent, i.e., $D^2 = D$, then $\mathcal{R}(D) = \mathcal{N}(I - D)$ and $\mathcal{N}(D) = \mathcal{R}(I - D)$.

**Proof** Only show $\mathcal{R}(D) = \mathcal{N}(I - D)$

$\subset$: $y \in \mathcal{R}(D) \Longrightarrow y = Dx \Longrightarrow (I - D)y = Dx - D^2x = 0 \Longrightarrow y \in \mathcal{N}(I - D)$.

$\supset$: $y \in \mathcal{N}(I - D) \Longrightarrow (I - D)y = 0 \Longrightarrow y = Dy \in \mathcal{R}(D)$.

(2) Expressing $\mathcal{R}(A)$ by null space

For $A \in R^{m\times n}$, $\mathcal{R}(A) = \mathcal{R}(AA^-) = \mathcal{N}(I_m - AA^-)$.

**Ex3:** $\mathcal{R}(A) = \mathcal{R}(AA^+) = \mathcal{N}(I_m - AA^+)$.

(3) Expressing $\mathcal{N}(A)$ by range

$\mathcal{N}(A) = \mathcal{N}(A^-A) = \mathcal{R}(I_n - A^-A)$.

**Ex4:** $\mathcal{N}(A) = \mathcal{N}(A^+A) = \mathcal{R}(I_n - A^+A)$.

3. Orthogonal complements

(1) Orthogonal complements
For $A \in R^{m\times n}$, the orthogonal complement of $\mathcal{R}(A)$,

$$\mathcal{R}^{\perp}(A) = \{y \in R^m : \langle y, z\rangle = 0 \text{ for all } z \in \mathcal{R}(A)\}$$

is also a space in $R^m$. The orthogonal complement of $\mathcal{N}(A)$,

$$\mathcal{N}^{\perp}(A) = \{y \in R^n : \langle y, z\rangle = 0 \text{ for all } z \in \mathcal{N}(A)\}$$

is a space in $R^n$.

(2) Expressing $\mathcal{R}^{\perp}(A)$ by null space: $\qquad \mathcal{R}^{\perp}(A) = \mathcal{N}(A')$.

**Proof.** $\mathcal{R}^{\perp}(A) = \mathcal{N}(A')$ since

$$\begin{aligned} y \in \mathcal{R}^{\perp}(A) &\iff \langle y, z\rangle = 0 \text{ for all } z \in \mathcal{R}(A) \iff \langle y, Ax\rangle = 0 \text{ for all } x \in R^n \\ &\iff x'A'y = 0 \text{ for all } x \in R^n \iff A'y = 0 \iff y \in \mathcal{N}(A'). \end{aligned}$$

So $\mathcal{R}^{\perp}(A) = \mathcal{N}(A')$.

(3) Expressing $\mathcal{N}^{\perp}(A)$ by range: $\qquad \mathcal{N}^{\perp}(A) = \mathcal{R}(A')$.

**Proof.** Note that $A^+A$ is idempotent, $I - A^+A$ is symmetric and idempotent.

$$\begin{aligned} \mathcal{N}^{\perp}(A) &= [\mathcal{N}(A)]^{\perp} = [\mathcal{N}(A^+A)]^{\perp} = [\mathcal{R}(I - A^+A)]^{\perp} = \mathcal{N}(I - A^+A) \\ &= \mathcal{R}\left(I - (I - A^+A)\right) = \mathcal{R}(A^+A) = \mathcal{R}\left((A')(A')^+\right) = \mathcal{R}(A'). \end{aligned}$$

**Ex5:** $\mathcal{R}^{\perp}(A) = \mathcal{N}(A') = \mathcal{N}\left((A')^+A'\right) = \mathcal{N}(AA^+) = \mathcal{R}(I - AA^+)$.

**Ex6:** $\mathcal{N}^{\perp}(A) = \mathcal{R}(A') = \mathcal{R}(A^+) = \mathcal{R}(A^+A)$.