L21 Least square solutions to $AXB = D$

1. Least square solutions

   (1) Least square solutions
   For equation $AXB = D$, $\widehat{X}$ is a least square solution (LSS) if

   $$\|D - A\widehat{X}B\|^2 \leq \|D - AXB\|^2 \text{ for all } X \in C^{p\times q}.$$

   (2) Relation to ordinary solutions
   If $AXB = D$ is consistent, then

   $$\widehat{X} \text{ is a LSS} \iff \widehat{X} \text{ is an ordinary solution.}$$

   Proof If $AXB = D$ is consistent, then there exists X_0 such that $\|AX_0B - D\|^2 = 0$. So $\widehat{X}$ is a LSS if and only if $\|A\widehat{X}B - D\|^2 \leq 0$, i.e., $\widehat{X}$ is an ordinary solution.

   (3) Collection of all LSSs
   The collection of all LSSs to $AXB = D$ is $A^+DB^+ + \mathcal{N}(A, B)$.

   Proof
   $$\begin{aligned} & \widehat{X} \text{ is a LSS to } AXB = D \\ \iff & \|D - A\widehat{X}B\|^2 \leq \|D - AXB\|^2 \text{ for all } X \in C^{p\times q} \\ \iff & A\widehat{X}B = \pi(D|\mathcal{R}(A, B)) = AA^+DB^+B \\ \iff & A(\widehat{X} - A^+DB^+)B = 0 \iff \widehat{X} - A^+DB^+ \in \mathcal{N}(A, B) \\ \iff & \widehat{X} \in A^+DB^+ + \mathcal{N}(A, B). \end{aligned}$$

   Ex1: For vector equation $Ax = b$,
   $$\begin{aligned} \widehat{x} \text{ is a LSS to } Ax = b \iff & \widehat{x} \text{ is a LSS to } Ax1 = b \\ \iff & \widehat{x} \in A^+b1^+ + \mathcal{N}(A, 1) = A^+b + \mathcal{N}(A). \end{aligned}$$

   Ex2: $A^+XB^+ = D$ is consistent. Find the collection of all its solutions.
   The collection of all its solutions is

   $$(A^+)^+D(B^+)^+ + \mathcal{N}(A^+, B^+) = ADB + \mathcal{N}(A^*, B^*).$$

2. Minimum norm LSS

   (1) A^+DB^+ is minimum norm LSS to $AXB = D$

   Proof $A^+DB^+ = A^+DB^+ + 0 \in A^+DB^+ + \mathcal{N}(A, B)$. So A^+DB^+ is a LSS.
   But $A^+DB^+ \in \mathcal{R}(A^+, B^+) = \mathcal{R}(A^*, B^*) = \mathcal{N}^\perp(A, B)$.
   So for LSS $\widehat{X} = A^+DB^+ + Z$ where $Z \in \mathcal{N}(A, B)$, $A^+DB^+ \perp Z$. Hence

   $$\|\widehat{X}\|^2 = \|A^+DB^+ + Z\|^2 = \|A^+DB^+\|^2 + \|Z\|^2 \geq \|A^+DB^+\|^2.$$

   So A^+DB^+ is the minimum norm LSS.

   (2) Lemma
   If X_0 is a LSS to $AXB = D$, then the collection of all LSSs is $X_0 + \mathcal{N}(A, B)$.

   Proof X_0 as a LSS by (3) of 1 satisfies $AX_0B = AA^+DB^+B$. We show

   $$X_0 + \mathcal{N}(A, B) = A^+DB^+ + \mathcal{N}(A, B).$$

$\subset$**:** If $\widehat{X} \in X_0 + \mathcal{N}(A, B)$, then $\widehat{X} = X_0 + Z_1$ where $Z_1 \in \mathcal{N}(A, B)$.
So $\widehat{X} = A^+DB^+ + (X_0 - A^+DB^+ + Z_1)$.
But $A(X_0 - A^+DB^+ + Z_1)B = 0$.
Hence $\widehat{X} \in A^+DB^+ + \mathcal{N}(A, B)$.

$\supset$**:** If $\widehat{X} \in A^+DB^+ + \mathcal{N}(A, B)$, then $\widehat{X} = A^+DB^+ + Z_2$ where $Z_2 \in \mathcal{N}(A, B)$.
So $\widehat{X} = X_0 + (A^+DB^+ - X_0 + Z_2)$.
But $A(A^+DB^+ - X_0 + Z_2)B = 0$.
Hence $\widehat{X} \in X_0 + \mathcal{N}(A, B)$.

Comment: In terms of expressing the collection of all LSSs, the role of A^+DB^+ can be replaced by any LSS X_0.
Every matrix X in $\mathcal{N}(A, B)$ is moved along X_0 to form $X_0 + \mathcal{N}(A, B)$. But only when $X_0 = A^+DB^+$ the moving direction is perpendicular to $\mathcal{N}(A, B)$.

Ex3: $\pi(A^+DB^+ \mid \mathcal{N}(A, B)) = (A^+DB^+) - (A^+A)(A^+DB^+)(BB^+) = 0$.

3. An application in Statistics

(1) A linear model for random matrix
For random $Y \in R^{m\times n}$ the model

$$Y = A\Theta B + \mathcal{E} \text{ with } E(\mathcal{E}) = 0$$

is a linear model since the linear transformation $E(Y) = A\Theta B$ specifies $E(Y)$ in linear space $\mathcal{R}(A, B)$.

(2) Least square estimators
For parameter matrix $\Theta \in R^{p\times q}$, $\widehat{\Theta}$ is called a least square estimator (LSE) if

$$\|Y - A\widehat{\Theta}B\|^2 \leq \|Y - A\Theta B\|^2 \text{ for all } \Theta \in R^{p\times q}.$$

Clearly by the definition

$$\begin{aligned}
\widehat{\Theta} \text{ is a LSE for } \Theta \quad &\Longleftrightarrow \quad \widehat{\Theta} \text{ is LSS to } A\Theta B = Y \\
&\Longleftrightarrow \quad A\widehat{\Theta}B = \pi(Y \mid \mathcal{R}(A, B)) \\
&\Longleftrightarrow \quad \widehat{\Theta} \in A^+YB^+ + \mathcal{N}(A, B).
\end{aligned}$$