Using SVD to make sense of Defn. 1.13

Recall the Geometry of SVD

$A = U \Sigma V^T$, $A^T = V \Sigma^T U^T$, $\text{rank}(A) = r$

$$\begin{array}{ccccc}
 & A & & A^T & \\
v_1 & \xrightarrow{\sigma_1} & u_1 & \xrightarrow{\sigma_1} & v_1 \\
v_2 & \xrightarrow{\sigma_2} & u_2 & \xrightarrow{\sigma_2} & v_2 \\
\vdots & \vdots & \vdots & \vdots & \\
v_r & \xrightarrow{\sigma_r} & u_r & \xrightarrow{\sigma_r} & v_r
\end{array}$$

$$\left.\begin{array}{c} v_{r+1} \\ \vdots \\ v_m \end{array}\right\} \to 0 \qquad \left.\begin{array}{c} u_{r+1} \\ \vdots \\ \vdots \\ u_n \end{array}\right\} \to 0$$

Geometry for Pseudo Inverse A':

$$\begin{array}{ccccc}
 & A & & A' & \\
v_1 & \xrightarrow{\sigma_1} & u_1 & \xrightarrow{1/\sigma_1} & v_1 \\
v_2 & \xrightarrow{\sigma_2} & u_2 & \xrightarrow{1/\sigma_2} & v_2 \\
\vdots & & \vdots & & \vdots \\
v_r & \xrightarrow{\sigma_r} & u_r & \xrightarrow{1/\sigma_r} & v_r
\end{array}$$

$$\left.\begin{array}{c} v_{r+1} \\ \vdots \\ v_m \end{array}\right\} \to 0 \qquad \left.\begin{array}{c} u_{r+1} \\ \vdots \\ \vdots \\ u_n \end{array}\right\} \to 0$$

When we are given a Linear System (OR a Least Squares problem) to solve; i.e.

$$Ax = b$$

where $A \in \mathbb{R}^{n \times m}$, $\text{rank}(A) = r$, AND we want to apply the Pseudo inverse to obtain the minimum norm "solution", we

Ⓐ Project b into $R(A)$ - if needed

Ⓑ Apply A' to get x.

So, by the "geometry" sketches on previous page, we see that the $x = A'b$ will come from $\text{span}\{v_1, v_2, \dots, v_r\}$

But $R(A^T) = \text{span}\{v_1, v_2, \dots, v_r\}$, and we know that

$$R(A^T) = N(A)^\perp$$

$\underbrace{\text{So, } x \in N(A)^\perp \Rightarrow \langle x, w \rangle = 0 \text{ for all } w \text{ satisfying } Aw = 0}_{\text{part 2 of defn 1.13}}$

<u>Defn 1.13</u>

The least squares pseudo-inverse of A is the matrix, denoted A' for which the vector $x = A'b$ satisfies

1. $A^* A x = A^* b$ (satisfies Normal EQNS)
2. $\langle x, w \rangle = 0$ for all w satisfying $Aw = 0$
   ↳ i.e. $x \in N(A)^\perp = R(A^T)$.

The above discussion & defn lead us to a different approach for computing A'.

→

Gaussian Elimination for Computing Pseudoinverse A' (pg. 30)

$\underbrace{\text{Gaussian Elimination}}$

To use G.E. to find a typical inverse, we augment the matrix with the Identity matrix and Eliminate. That is, apply GE to

$$[A \mid I] \quad \text{row reduce ALOT} \longrightarrow [I \mid A^{-1}]$$

↳ for Invertible A case

Q: How do we use this idea to find the pseudoinverse in a case where $A \in \mathbb{R}^{n \times m}$?

A: We need to resolve how to deal with Ⓐ and Ⓑ on previous page.

Ⓐ If $b \notin R(A)$, then we compute $b_r = Pb \in R(A)$ where P is a projection matrix. What does P look like?

Important!
Note that we already know what it looks like if we express it in terms of a basis of $R(A)$. BUT for this case, we want to express P in term of a basis for $N(A^*)$.

Let $\{u_1, u_2, \ldots, u_k\}$ be an ONB of $N(A^*)$. Since $\mathbb{R}^n = R(A) \oplus N(A^*)$, then we can express

$$b = b_r + e \quad \text{with} \quad b_r \in R(A),\ e \in N(A^*)$$

Note that $e \in N(A^*) \Rightarrow \langle e, u \rangle = 0 \ \forall u \in R(A)$ (Since $R(A) = N(A^*)^{\perp}$)

Since, $e \in N(A^*)$, then in terms of the ONB, we write

$$e = \sum_{j=1}^{k} \alpha_j u_j = \alpha_1 u_1 + \alpha_2 u_2 + \cdots + \alpha_k u_k$$

To compute the α_j's, we note that

$$\langle e, u_i \rangle = \langle \sum_{j=1}^{k} \alpha_j u_j, u_i \rangle = \alpha_i \langle u_i, u_i \rangle = \alpha_i \quad \text{since } \{u_j\}_{j=1}^{k} \text{ are an ONB.}$$

Using $b = b_r + e$, we have th

$$\begin{aligned} \alpha_i &= \langle b - b_r, u_i \rangle \\ &= \langle b, u_i \rangle - \underbrace{\langle b_r, u_i \rangle}_{=0 \text{ forall } i=1,2,\ldots k} \end{aligned}$$

since $b_r \in R(A) = N(A^*)^\perp$

Hence, $\alpha_i = \langle b, u_i \rangle$ for $i = 1, 2, \ldots k$
Then

$$\begin{aligned} b_r &= b - e = b - \sum_{j=1}^{k} \alpha_j u_j \\ &= b - \sum_{j=1}^{k} \langle b, u_j \rangle u_j \\ &= b - \sum_{j=1}^{k} \langle u_j, b \rangle u_j \\ &= b - \sum_{j=1}^{k} [u_j u_j^T] b \\ &= \left[I - \sum_{j=1}^{k} u_j u_j^T \right] b \end{aligned}$$

Hence, in terms of the ONB of $N(A^*)$, the projection matrix P has the form

$$P = I - \sum_{j=1}^{k} u_j u_j^T$$

<u>Note</u>: If $N(A^*)=\{0\}$, then $k=0$, $b_r=b$ and $P=I$

- If $N(A)=\{0\}$, then we row reduce
$$[A|P]$$
to compute A^{-1} (the real A^{-1})

- If $N(A)=\text{span}\{w_1,w_2,\dots,w_\ell\}$, then we compute the Pseudoinverse A' by row reducing

$$\left[\begin{array}{c|c} A & P \\ w_1^T & \\ w_2^T & \\ \vdots & \\ w_\ell^T & \end{array}\right]$$

These rows impose the conditions that $x \perp N(A)$. That is, $\langle x, w_i\rangle = 0 \;\forall\, w_i \in N(A)$, $i=1,2,\dots,\ell$

This is simply imposing the conditions of Defn. 1.13
Once we row reduce, this leads us to the system

$$[I|A']$$

(if we discard all the zero rows)

The following Example comes from the problems at the end of the Chapter.

Ex: $A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \\ -1 & 3 \end{bmatrix}$ Compute $A^{\dagger}$ via G.E.

$R(A) = \text{span}\left\{ \begin{bmatrix} 3 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} -1 \\ 1 \\ 3 \end{bmatrix} \right\}$ and $N(A) = \{0\}$

$N(A^*) = \text{span}\left\{ \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\} = \text{span}\underbrace{\left\{ \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \right\}}_{\text{O.N.B.}}$

(Note: I got this by row-reducing $\begin{bmatrix} 3 & 1 & -1 \\ -1 & 1 & 3 \end{bmatrix} = A^*$)

Projection Matrix:

$$P = I - \frac{1}{\sqrt{6}} \begin{bmatrix} 1 \\ -2 \\ 1 \end{bmatrix} \cdot \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & -2 & 1 \end{bmatrix}$$

$$= \frac{1}{6} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 6 \end{bmatrix} - \frac{1}{6} \begin{bmatrix} 1 & -2 & 1 \\ -2 & 4 & -2 \\ 1 & -2 & 1 \end{bmatrix} = \frac{1}{6} \begin{bmatrix} 5 & 2 & -1 \\ 2 & 2 & 2 \\ -1 & 2 & 5 \end{bmatrix}$$

$$= \begin{bmatrix} \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{bmatrix}$$

Row Reduce:

Lots of Scratch Work Here!

$$\left[\begin{array}{cc|ccc} 3 & -1 & \frac{5}{6} & \frac{1}{3} & -\frac{1}{6} \\ 1 & 1 & \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ -1 & 3 & -\frac{1}{6} & \frac{1}{3} & \frac{5}{6} \end{array}\right] \longrightarrow \begin{bmatrix} 1 & 0 & \frac{7}{24} & \frac{1}{6} & \frac{1}{24} \\ 0 & 1 & \frac{1}{24} & \frac{1}{6} & \frac{7}{24} \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$$A' = \begin{bmatrix} \frac{7}{24} & \frac{1}{6} & \frac{1}{24} \\ \frac{1}{24} & \frac{1}{6} & \frac{7}{24} \end{bmatrix} = \frac{1}{24}\begin{bmatrix} 7 & 4 & 1 \\ 1 & 4 & 7 \end{bmatrix}$$

The example in your textbook also has the case where $N(A) \neq \{0\}$. You might want to take a look at that one.