It turns out that the pseudoinverse A^+ can be easily obtained from the SVD of A:

Theorem 177. The **pseudoinverse** of an $m\times n$ matrix A with SVD $A=U\Sigma V^T$ is

$$A^+ = V\Sigma^+ U^T,$$

where Σ^+, the pseudoinverse of Σ, is the $n\times m$ diagonal matrix, whose nonzero entries are the inverses of the entries of Σ.

Proof. The equation $A\boldsymbol{x}=\boldsymbol{b}$ is equivalent to $U\Sigma V^T\boldsymbol{x}=\boldsymbol{b}$ and, thus, $\Sigma V^T\boldsymbol{x}=U^T\boldsymbol{b}$.
Write $\boldsymbol{y}=V^T\boldsymbol{x}$ and note that $\boldsymbol{y}$ and $\boldsymbol{x}$ have the same norm (why?!).
We already know that the equation $\Sigma\boldsymbol{y}=U^T\boldsymbol{b}$ has optimal solution $\boldsymbol{y}=\Sigma^+U^T\boldsymbol{b}$.
Since $\boldsymbol{y}$ and $\boldsymbol{x}$ have the same norm, it follows that $\boldsymbol{x}=V\boldsymbol{y}=V\Sigma^+U^T\boldsymbol{b}$ is the optimal solution to $A\boldsymbol{x}=\boldsymbol{b}$.
Hence, $A^+=V\Sigma^+U^T$. □

Lemma 178. The pseudoinverse of A^+ is $A^{++}=A$.

Proof. Starting with the SVD $A=U\Sigma V^T$, we have $A^+=V\Sigma^+U^T$, which is the SVD of A^+.
Therefore, $A^{++}=U\Sigma^{++}V^T$. The claim thus follows from $\Sigma^{++}=\Sigma$. □

Example 179. Determine the pseudoinverse of $A=\begin{bmatrix}1&-1\\0&1\\1&0\end{bmatrix}$ in two ways.

First, using the SVD and, second, using the fact that A has full column rank.

Solution. (SVD) We have computed the SVD of this matrix before.
Since $A=U\Sigma V^T$ with $U=\begin{bmatrix}-2/\sqrt{6}&0&-1/\sqrt{3}\\1/\sqrt{6}&1/\sqrt{2}&-1/\sqrt{3}\\-1/\sqrt{6}&1/\sqrt{2}&1/\sqrt{3}\end{bmatrix}$, $\Sigma=\begin{bmatrix}\sqrt{3}&0\\0&1\\0&0\end{bmatrix}$, $V=\frac{1}{\sqrt{2}}\begin{bmatrix}-1&1\\1&1\end{bmatrix}$,
the pseudoinverse is $A^+=V\Sigma^+U^T$ where $\Sigma^+=\begin{bmatrix}1/\sqrt{3}&0&0\\0&1&0\end{bmatrix}$.
Multiplying these matrices, $A^+=\frac{1}{3}\begin{bmatrix}1&1&2\\-1&2&1\end{bmatrix}$.
Comment. For many applications, it may be neither necessary nor helpful to multiply V,Σ^+,U^T.

Solution. (full column rank) Since A clearly has full column rank, we also have $A^+=(A^TA)^{-1}A^T$.
Indeed, $A^+=(A^TA)^{-1}A^T=\begin{bmatrix}2&-1\\-1&2\end{bmatrix}^{-1}\begin{bmatrix}1&0&1\\-1&1&0\end{bmatrix}=\frac{1}{3}\begin{bmatrix}2&1\\1&2\end{bmatrix}\begin{bmatrix}1&0&1\\-1&1&0\end{bmatrix}=\frac{1}{3}\begin{bmatrix}1&1&2\\-1&2&1\end{bmatrix}$.

Example 180. What is the pseudoinverse of $A=\begin{bmatrix}2&2\\1&1\end{bmatrix}$?

Solution. Recall (or compute) that $A=U\Sigma V^T$ with $U=\frac{1}{\sqrt{5}}\begin{bmatrix}2&-1\\1&2\end{bmatrix}$, $\Sigma=\begin{bmatrix}\sqrt{10}&\\&0\end{bmatrix}$, $V=\frac{1}{\sqrt{2}}\begin{bmatrix}1&-1\\1&1\end{bmatrix}$.
Hence, $A^+=V\Sigma^+U^T$ where $\Sigma^+=\begin{bmatrix}1/\sqrt{10}&0\\0&0\end{bmatrix}$.
Multiplying these matrices (which may not be necessary or helpful for applications), $A^+=\frac{1}{10}\begin{bmatrix}2&1\\2&1\end{bmatrix}$.

Note. Since A does not have full column rank, $A^+=(A^TA)^{-1}A^T$ cannot be used. That's because A^TA is not invertible.

Comment. Here, $A^+A=\boldsymbol{v}_1\boldsymbol{v}_1^T=\frac{1}{2}\begin{bmatrix}1&1\\1&1\end{bmatrix}$ and $AA^+=\boldsymbol{u}_1\boldsymbol{u}_1^T=\frac{1}{5}\begin{bmatrix}4&2\\2&1\end{bmatrix}$ are not visually like the identity. However, note that these are the (orthogonal) projections onto $\boldsymbol{v}_1$ and $\boldsymbol{u}_1$ respectively (in particular, the eigenvalues are $1,0$).

Armin Straub
straub@southalabama.edu

Review.

- If the $m \times n$ matrix A has SVD $A = U\Sigma V^T$, then its pseudoinverse is $A^+ = V\Sigma^+ U^T$.

  Here, Σ^+, the pseudoinverse of Σ, is the $n \times m$ diagonal matrix, whose nonzero entries are the inverses of the entries of Σ.

- The system $A\boldsymbol{x} = \boldsymbol{b}$ has "optimal" solution $\boldsymbol{x} = A^+\boldsymbol{b}$.

  Here, "optimal" means that $\boldsymbol{x}$ is the smallest least squares solution.

Example 181.

(a) Find the pseudoinverse of $A = [\ 1 \ \ 2 \ \ 3\ ]$.

(b) Find the smallest solution to $x_1 + 2x_2 + 3x_3 = 6$.

As before, smallest solutions means the solution $\boldsymbol{x}$ such that $\|\boldsymbol{x}\|$ is as small as possible. One obvious solution is $[\ 1, 1, 1\ ]^T$, but is it the smallest?

Solution.

(a) $A^TA = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} [\ 1 \ \ 2 \ \ 3\ ] = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$ has 14-eigenvector $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$ and 0-eigenvectors $\begin{bmatrix} -2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} -3 \\ 0 \\ 1 \end{bmatrix}$.

$\boldsymbol{u}_1 = \frac{1}{\sigma_1} A\boldsymbol{v}_1 = \frac{1}{\sqrt{14}} [\ 1 \ \ 2 \ \ 3\ ] \frac{1}{\sqrt{14}} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = 1$

Hence, $A = U\Sigma V^T$ with $U = [\ 1\ ]$, $\Sigma = [\ \sqrt{14} \ \ 0 \ \ 0\ ]$, $V = \begin{bmatrix} 1/\sqrt{14} & * & * \\ 2/\sqrt{14} & * & * \\ 3/\sqrt{14} & * & * \end{bmatrix}$.

$$A^+ = V\Sigma^+ U^T = \begin{bmatrix} 1/\sqrt{14} & * & * \\ 2/\sqrt{14} & * & * \\ 3/\sqrt{14} & * & * \end{bmatrix} \begin{bmatrix} 1/\sqrt{14} \\ 0 \\ 0 \end{bmatrix} [\ 1\ ] = \frac{1}{14} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Comment. No surprise on U. The only options for U are $U = [\ 1\ ]$ and $U = [\ -1\ ]$.

Comment. Realizing what we did here allows us to write down A^+ immediately for all $1 \times n$ matrices A. See Example 182.

Homework. Complete the SVD of A. That is, find an option for the two missing columns of V, so that V is an orthogonal matrix. In other words, find an orthonormal basis for the 0-eigenspace.

Comment. An even better approach would be to compute AA^T first (instead of A^TA) which would allow us to compute U first (rather than V first). Can you fill in the blanks?

(b) We are solving $A\boldsymbol{x} = [\ 6\ ]$ with $A = [\ 1 \ \ 2 \ \ 3\ ]$ as in the previous example.

We conclude that the smallest solution is $\boldsymbol{x} = A^+[\ 6\ ] = \frac{3}{7} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$.

Compare. $\left\| \frac{3}{7} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \right\| = \frac{3}{7}\sqrt{14} \approx 1.604$ is indeed smaller than, say, $\left\| \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \right\| = \sqrt{3} \approx 1.732$.

Geometric picture. The equation $x_1 + 2x_2 + 3x_3 = 6$ describes a plane (not through the origin), and we are asking for the point on that plane which is closest to the origin. That's a typical question in Calculus III. Note that $[\ 1 \ \ 2 \ \ 3\ ]^T$ is the normal vector of the plane. Explain why the answer had to be a multiple of that normal vector!

Example 182. More generally, find the pseudoinverse of $A = [\ a_1 \ \ a_2 \ \ a_3\ ]$.

Solution. As in the previous example, we see that the answer will be $A^+ = \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|^2}$ with $\boldsymbol{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}$.

Comment. Likewise for $A = [\ a_1 \ \ a_2 \ \ \ldots \ \ a_n\ ]$.