**Review.** Let $A$ be $m \times n$. The SVD $A = U\Sigma V^T$ is equivalent to $AV = U\Sigma$. That is, $A\boldsymbol{v}_i = \sigma_i \boldsymbol{u}_i$.

Here, $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ (the columns of $U$ and $V$) are orthonormal bases of the output and input spaces ($\mathbb{R}^m$ and $\mathbb{R}^n$).

Make sure you see what happens to the input vector $\boldsymbol{v}_i$ (from $\mathbb{R}^n$) when $A = U\Sigma V^T$ is applied:

- $V^T = V^{-1}$ sends $\boldsymbol{v}_i$ to $\boldsymbol{e}_i$ ($\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n$ are the standard basis vectors in $\mathbb{R}^n$),
- $\Sigma$ sends $\boldsymbol{e}_i$ to $\sigma_i \tilde{\boldsymbol{e}}_i$ ($\tilde{\boldsymbol{e}}_1, \ldots, \tilde{\boldsymbol{e}}_m$ are the standard basis vectors in $\mathbb{R}^m$),
- $U$ sends $\tilde{\boldsymbol{e}}_i$ to $\boldsymbol{u}_i$ and, hence, $\sigma_i \tilde{\boldsymbol{e}}_i$ to $\sigma_i \boldsymbol{u}_i$ (which is in $\mathbb{R}^m$).

If we choose $\boldsymbol{u}_i$ and $\boldsymbol{v}_i$ as the new standard basis, then the linear map $\boldsymbol{x} \mapsto A\boldsymbol{x}$ has matrix $\Sigma$ (diagonal). In that basis, $A$ is just scaling the basis vectors!

**Example 175.** Show that the eigenvalues of $A^T A$ are all nonnegative.

**Proof.** Suppose that $\lambda$ is an eigenvalue of $A^T A$. Then $A^T A \boldsymbol{v} = \lambda \boldsymbol{v}$ (where $\boldsymbol{v}$ is a $\lambda$-eigenvector).

It follows that $\underbrace{\boldsymbol{v}^T A^T A \boldsymbol{v}}_{= \|A\boldsymbol{v}\|^2 \geqslant 0} = \lambda \boldsymbol{v}^T \boldsymbol{v} = \lambda \|\boldsymbol{v}\|^2$. Finally, $\lambda \|\boldsymbol{v}\|^2 \geqslant 0$ implies that $\lambda \geqslant 0$. □

The **pseudoinverse** of an $m \times n$ matrix $A$ is the matrix $A^+$ such that 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.

In particular:

- If $A\boldsymbol{x} = \boldsymbol{b}$ has a unique solution, then $\boldsymbol{x} = A^+ \boldsymbol{b}$ is that solution.
- If $A\boldsymbol{x} = \boldsymbol{b}$ has many solutions, then $\boldsymbol{x} = A^+ \boldsymbol{b}$ is the one of smallest norm (the "optimal" one; and there is indeed only one such optimal solution).
- If $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent but has a unique least squares solution, then $\boldsymbol{x} = A^+ \boldsymbol{b}$ is that least squares solution.
- If $A\boldsymbol{x} = \boldsymbol{b}$ has many least squares solutions, then $\boldsymbol{x} = A^+ \boldsymbol{b}$ is the one with smallest norm.

When there is a unique (least squares) solution, we know how to find the pseudoinverse:

- If $A$ is invertible, then $A^+ = A^{-1}$.
- If $A$ has full column rank, then $A^+ = (A^T A)^{-1} A^T$.

  **Recall.** If $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, a least squares solution can be determined by solving $A^T A \boldsymbol{x} = A^T \boldsymbol{b}$. If $A$ has full column rank (i.e. the columns of $A$ are independent; in this context, the typical case), then $\boldsymbol{x} = (A^T A)^{-1} A^T \boldsymbol{b}$ is the **unique** least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$.

**Example 176.**

(a) What is the pseudoinverse of $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$?

(b) What is the pseudoinverse of $\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$?

(c) What is the pseudoinverse of $\Sigma = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$?

(d) In each case, compute $\Sigma^+ \Sigma$ and $\Sigma \Sigma^+$.

Armin Straub
straub@southalabama.edu

**Solution.**

(a) Recall that, if $A$ has full column rank, then $A^+ = (A^T A)^{-1} A^T$.

Here, $\Sigma^T \Sigma = \begin{bmatrix} 4 & 0 \\ 0 & 9 \end{bmatrix}$, so that $\Sigma^+ = (\Sigma^T \Sigma)^{-1} \Sigma^T = \begin{bmatrix} 1/4 & \\ & 1/9 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$.

**Alternative.** Let us think about the optimal solution to $\Sigma \boldsymbol{x} = \boldsymbol{b}$, that is, $\begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$.

The (unique) least squares solution is $\boldsymbol{x} = \begin{bmatrix} b_1/2 \\ b_2/3 \end{bmatrix}$. (Review if this is not obvious!)

Since $\begin{bmatrix} b_1/2 \\ b_2/3 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} \boldsymbol{b}$, we conclude that $\Sigma^+ = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$.

(b) Let us think about the smallest norm ("optimal") solution to $\Sigma \boldsymbol{x} = \boldsymbol{b}$, that is, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$.

The general solution is $\boldsymbol{x} = \begin{bmatrix} b_1/2 \\ b_2/3 \\ t \end{bmatrix}$, where $t$ is a free parameter.

Clearly, the smallest norm solution is $\begin{bmatrix} b_1/2 \\ b_2/3 \\ 0 \end{bmatrix}$.

Since $\begin{bmatrix} b_1/2 \\ b_2/3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{bmatrix} \boldsymbol{b}$, we conclude that $\Sigma^+ = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{bmatrix}$.

(c) Now, $\Sigma \boldsymbol{x} = \boldsymbol{b}$, that is, $\begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ has no solution (unless $b_2 = 0$).

We therefore need to think about least squares solutions.

The general least squares solution (why?!) is $\boldsymbol{x} = \begin{bmatrix} b_1/2 \\ s \\ t \end{bmatrix}$, where $s, t$ are free parameters.

Clearly, the smallest norm least squares solution is $\begin{bmatrix} b_1/2 \\ 0 \\ 0 \end{bmatrix}$.

Since $\begin{bmatrix} b_1/2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \boldsymbol{b}$, we conclude that $\Sigma^+ = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}$.

(d) Firstly, $\Sigma^+ \Sigma = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\Sigma \Sigma^+ = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Secondly, $\Sigma^+ \Sigma = \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\Sigma \Sigma^+ = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 1/3 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$.

[Note how the pseudoinverse tries to behave like the regular inverse. But since $\Sigma$ has only $2$ columns, $\Sigma^+ \Sigma$ and $\Sigma \Sigma^+$ can have rank at most $2$ (so cannot be the full $3 \times 3$ identity).]

Thirdly, $\Sigma^+ \Sigma = \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $\Sigma \Sigma^+ = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1/2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$.

[Here, $\Sigma$ has rank $1$, so that $\Sigma^+ \Sigma$ and $\Sigma \Sigma^+$ can have rank at most $1$.]

**In general.** Proceeding, as in this example, we find that the pseudoinverse of any $m \times n$ diagonal matrix $\Sigma$ is the $n \times m$ (transposed dimensions!) diagonal matrix whose nonzero entries are the inverses of the entries of $\Sigma$.

**Comment.** Observe that, in all three cases, $\Sigma^{++} = \Sigma$.

**Comment.** Note that $\begin{bmatrix} 1 & 0 \\ 0 & \varepsilon \end{bmatrix}^+ = \begin{bmatrix} 1 & 0 \\ 0 & \varepsilon^{-1} \end{bmatrix}$ for small $\varepsilon \neq 0$, while $\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$. This shows that the pseudoinverse is not a continuous operation.