Example 113. Show that the eigenvalues of A^TA are all nonnegative.

Proof. Suppose that λ is an eigenvalue of A^TA. Then $A^TA\boldsymbol{v} = \lambda\boldsymbol{v}$ (where $\boldsymbol{v}$ is a λ-eigenvector).
It follows that $\underbrace{\boldsymbol{v}^TA^TA\boldsymbol{v}}_{=\|A\boldsymbol{v}\|^2 \geqslant 0} = \lambda\boldsymbol{v}^T\boldsymbol{v} = \lambda\|\boldsymbol{v}\|^2$. Finally, $\lambda\|\boldsymbol{v}\|^2 \geqslant 0$ implies that $\lambda \geqslant 0$. □

Example 114. Determine the SVD of $A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Solution. $A^TA = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}$ has 3-eigenvector $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and 1-eigenvector $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Since $A^TA = V\Sigma^T\Sigma V^T$, we conclude that $V = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$ and $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$.

$\boldsymbol{u}_1 = \frac{1}{\sigma_1}A\boldsymbol{v}_1 = \frac{1}{\sqrt{3}}\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{6}}\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}$

$\boldsymbol{u}_2 = \frac{1}{\sigma_2}A\boldsymbol{v}_2 = \frac{1}{1}\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \frac{1}{\sqrt{2}}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}$

$\boldsymbol{u}_3$ is chosen so that the matrix U is orthogonal. For instance, $\boldsymbol{u}_3 = \frac{1}{\sqrt{3}}\begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix}$.

Hence, $U = \begin{bmatrix} -2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$.

In summary, $A = U\Sigma V^T$ with $U = \begin{bmatrix} -2/\sqrt{6} & 0 & -1/\sqrt{3} \\ 1/\sqrt{6} & 1/\sqrt{2} & -1/\sqrt{3} \\ -1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \end{bmatrix}$, $\Sigma = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$, $V = \frac{1}{\sqrt{2}}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}$.

How did we find $\boldsymbol{u}_3$? We already have the vectors $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$, and need a vector orthogonal to both.

That is, we need to find the vector spanning $\operatorname{span}\left\{\begin{bmatrix} -2 \\ 1 \\ -1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\right\}^\perp = \operatorname{col}\left(\begin{bmatrix} -2 & 0 \\ 1 & 1 \\ -1 & 1 \end{bmatrix}\right)^\perp = \operatorname{null}\left(\begin{bmatrix} -2 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}\right)$.

[Without the intermediate steps, can you see why the null space consists of precisely the vectors orthogonal to both $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$?]

More generally, proceeding like this, we can always fill in "missing" vectors $\boldsymbol{u}_i$ to obtain an orthonormal basis $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_m$ that we can use as the columns of U.

Example 115. (least squares) Recall that if $A\boldsymbol{x} = \boldsymbol{b}$ is inconsistent, it is often useful to determine a least squares solution by solving $A^TA\boldsymbol{x} = A^T\boldsymbol{b}$.

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

Otherwise, there may be several least squares solutions. The one of smallest norm $\boldsymbol{x}^+$ is called the optimal least squares solution (and there is indeed only one such optimal solution).

The **pseudoinverse** of A is the matrix A^+ such that $\boldsymbol{x}^+ = A^+\boldsymbol{b}$. It turns out that it is easily obtained from the SVD of A:

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.

- If A is invertible, then $A^+ = A^{-1}$.

  Why? A is invertible if and only if Σ is invertible. Clearly, $\Sigma^{-1} = \Sigma^+$.
  Hence, $A^{-1} = (U\Sigma V^T)^{-1} = V\Sigma^{-1}U^T = V\Sigma^+U^T = A^+$.

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

  Why? If $A = U\Sigma V^T$, then $(A^TA)^{-1}A^T = (V\Sigma^T\Sigma V^T)^{-1}V\Sigma^TU^T = V(\Sigma^T\Sigma)^{-1}\Sigma^TU^T$. Hence, it only remains to see why $(\Sigma^T\Sigma)^{-1}\Sigma^T = \Sigma^+$. See the first comment in Example 116.

  So what? As recalled above, $\boldsymbol{x} = (A^TA)^{-1}A^T\boldsymbol{b}$ is the unique least squares solution to $A\boldsymbol{x} = \boldsymbol{b}$. So, we have shown that $A\boldsymbol{x} = \boldsymbol{b}$ has least squares solution $\boldsymbol{x} = A^+\boldsymbol{b}$.

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

  Why? This is easy to see from $A^+ = V\Sigma^+U^T$ and $\Sigma^{++} = \Sigma$. (Do you see it? If not quite, spell out the second comment in Example 116.)

Example 116. What is the pseudoinverse of $\Sigma = \begin{bmatrix} 2 & 0 \\ 0 & 3 \\ 0 & 0 \end{bmatrix}$?

Solution. Inverting the nonzero diagonal elements, then transposing, we find $\Sigma^+ = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$.

Comment. $\Sigma^T\Sigma = \begin{bmatrix} 4 & \\ & 9 \end{bmatrix}$, so that $(\Sigma^T\Sigma)^{-1}\Sigma^T = \begin{bmatrix} 1/4 & \\ & 1/9 \end{bmatrix}\begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} = \begin{bmatrix} 1/2 & 0 & 0 \\ 0 & 1/3 & 0 \end{bmatrix}$. That's indeed Σ^+.

Comment. Observe that, obviously, $\Sigma^{++} = \Sigma$.

Example 117. 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 118. (extra) 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.