1. SVD Short Questions  Assume we have the compact form of the SVD of $A = U_1 S V_1^T = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$.

(a) Compute $AV_1 V_1^T$

**Solutions:** Recall that $V_1$ is an orthogonal matrix, so it has orthonormal columns, giving it the property $V_1^T V_1 = I$. Hence we can write:

$$AV_1 V_1^T = U_1 S V_1^T V_1 V_1^T = U_1 S V_1^T = A$$

(b) What is the subspace that spans the column space of $A$?

**Solutions:** Given a vector $\vec{x}$, the column space of $A$ is also the same as the space of all possible $A\vec{x}$.

$$A\vec{x} = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T \vec{x}$$

But, $\vec{v}_i^T \vec{x}$ is a scalar, hence,

$$A\vec{x} = \sum_{i=1}^{r} (\sigma_i \vec{v}_i^T \vec{x}) \vec{u}_i$$

From that decomposition, we can see that $A\vec{x}$ is a linear combination of $\vec{u}_1, \vec{u}_2, \ldots, \vec{u}_r$. Hence the span of columns of $A$ is the subspace spanned by the columns of $U_1$.

2. Frobenius Norm  In this problem we will investigate the properties of the Frobenius norm.

(a) The trace of a matrix is the sum of its diagonal entries. For example, let $Q \in \mathbb{R}^{N \times N}$, then,

$$Tr\{Q\} = \sum_{i=1}^{N} Q_{ii}$$

Much like the norm of a vector $\vec{x} \in \mathbb{R}^N$ is $\sqrt{\sum_{i=1}^{N} x_i^2}$, the Frobenius norm of a matrix $Q$ is defined as,

$$||Q||_F = \sqrt{\sum_{i=1}^{N} \sum_{j=1}^{N} |Q_{ij}|^2}$$

Note that matrices have other types of norms as well. With the above definitions, show that,

$$||A||_F = \sqrt{Tr\{A^T A\}}$$
Solutions:

\[ Tr\{A^T A\} = \sum_{i=1}^{N} (A^T A)_{ii} \]

\[ = \sum_{i=1}^{N} \left( \sum_{j=1}^{N} (A^T)_{ij} A_{ji} \right) \]

\[ = \sum_{i=1}^{N} \left( \sum_{j=1}^{N} A_{ji} A_{ji} \right) \]

\[ = \sum_{i=1}^{N} \sum_{j=1}^{N} (A^2)_{ji} \]

\[ = ||A||^2_F \]

(b) Show that if \( U \) and \( V \) are orthonormal, then

\[ ||UA||_F = ||AV||_F = ||A||_F \]

Solutions:

\[ ||UA||_F = \sqrt{Tr\{(UA)^T(UA)\}} = \sqrt{Tr\{A^T U^T U A\}} = \sqrt{Tr\{A^T A\}} = ||A||_F \]

To show the second set of equality, we must note that \( Tr\{A^T A\} = Tr\{AA^T\} \). Hence,

\[ ||AV||_F = \sqrt{Tr\{(AV)(AV)^T\}} = \sqrt{Tr\{AVV^T A^T\}} = \sqrt{Tr\{AA^T\}} = ||A||_F \]

(c) Show that \( ||A||_F = \sqrt{\sum_{i=1}^{N} \sigma_i^2} \)

Solutions:

\[ ||A||_F = ||U\Sigma V^T||_F = ||\Sigma V^T||_F = ||\Sigma||_F \]

\[ = \sqrt{Tr\{\Sigma^T \Sigma\}} = \sqrt{\sum_{i=1}^{N} \sigma_i^2} \]
Let $A \in \mathbb{R}^{M \times N}$ be a “fat” matrix, where $M < N$. $A$ is full rank, with $\text{Rank}(A) = M$.

a) $A = U S V^T$ is the SVD of $A$. What are the sizes of $U$, $S$, $V$?

**Solution:**

Since $\text{Rank}(A) = M$, $S \in \mathbb{R}^{M \times M}$.

\[ A = \begin{bmatrix} \cdot & \cdot \\ \cdot & \cdot \end{bmatrix} \quad M = \begin{bmatrix} U_1 \\ \cdot \\ \cdot \end{bmatrix} \quad S = \begin{bmatrix} \cdot \\ \cdot \end{bmatrix} \quad V^T = \begin{bmatrix} \cdot & \cdot & \cdot \\ \end{bmatrix} \]

So, $U_1 \in \mathbb{R}^{M \times M}$, $S \in \mathbb{R}^{M \times M}$, $U_1 \in \mathbb{R}^{M \times M}$.

b) You are given the following equation, where $\vec{x}$ is unknown:

\[ A \vec{x} = \vec{y} \]

$A$ is the same as above, and can represent some linear system. $\vec{y}$ is known and can represent a desired output of system $A$. We would like to design an input $\vec{x}$, which satisfies the above equality. Note, that since $A$ is fat, we can not just compute an inverse. In fact, there are infinite number of solutions to Eq. 1.

We define a pseudo-inverse $A^\dagger = V_1 S^{-1} U_1^T$.

Show that $\vec{x} = A^\dagger \vec{y}$ is a solution to Eq. 1.
Solution:

$U$ is a square orthonormal matrix. Hence, $U^T U = U U^T = I_{n \times n}$

$V$ is tall, and orthonormal. Hence, $V^T V = I_{m \times m}$

\[ A \hat{x} = A A^T \bar{y} = U S V_i^T V_i S^{-1} U^T \bar{y} = U S S^{-1} U^T \bar{y} = U U^T \bar{y} = \bar{y} \]

$c$) Show that $\hat{x} + \bar{x}$ is also a solution,

\[ A(\hat{x} + \bar{x}) = \bar{y} \]

only if $\bar{x}$ is spanned by the null-space of $V_i$

Solution:

\[ \bar{y} = A(\hat{x} + \bar{x}) = A\hat{x} + A\bar{x} = \bar{y} + A\bar{x} \Rightarrow \text{true only if } A\bar{x} = 0 \]

\[ A\bar{x} = U_i S V_i^T \bar{x} = 0 \]

since $S$ has non-zero diagonals, this is true only if $V_i^T \bar{x} = 0$
d) Show that when \( \hat{x} = A^T \hat{y} \), is a solution for Eq. 1. \( \hat{x} \) has the minimum norm among all solutions that satisfy Eq. 1.

In other words: let \( \hat{x} \) satisfy \( A \hat{x} = y \). If \( \hat{x} \neq x \), then \( \|\hat{x}\| < \|x\| \).

Solution:

Let \( A = U \Sigma V^T \) be the full SVD. \( V = [V_1, V_2] \)

If \( \hat{x} \neq x \), then \( \hat{x} = x + \tilde{x} \)

The norm does not change when multiplying by an orthonormal matrix. So,

\[
\|\hat{x}\| = \|VV^T \hat{x}\| = \|VV^T(x + \tilde{x})\| = \\
\|V_1V_1^T \hat{x} + V_2V_2^T \tilde{x}\|
\]

From part (C),

\[
\|V_1V_1^T \hat{x} + V_2V_2^T \tilde{x}\| < \|V_1V_1^T \hat{x}\| + \|V_2V_2^T \tilde{x}\|
\]

Thus,

\[
\|V_1V_1^T \hat{x}\| + \|V_2V_2^T \tilde{x}\| > \|\hat{x}\|
\]
e) From

Find the vector \( \mathbf{x} \) with the smallest norm, that satisfies,

\[
A \mathbf{x} = \begin{bmatrix} 2 \\ 0 \\ 0 \end{bmatrix}
\]

Solution:

\[
\mathbf{x}^* = A^+ y = V_1 S^{-1} U_1^T
\]

\[
S = \begin{bmatrix} 3 \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{bmatrix} = S^{-1} = \frac{1}{6} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}
\]

\[
S^{-1} U_1^T = \frac{1}{6} \begin{bmatrix} \frac{3}{2} & 0 \\ 0 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} -\frac{3}{2} & \frac{3}{2} \\ -1 & 1 \end{bmatrix}
\]

\[
V_1 \cdot S^{-1} U_1^T = \frac{1}{6} \begin{bmatrix} \frac{3}{2} & -\frac{1}{2} \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{5}{2} & \frac{1}{2} \end{bmatrix}
\]

\[
A^+ = \frac{1}{6} \begin{bmatrix} 2 & -1 \\ -1 & 2 \\ -1 & 2 \end{bmatrix} \quad \text{and} \quad \mathbf{x}^* = A^+ \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} \frac{2}{3} \\ \frac{2}{3} \\ -1 \end{bmatrix}
\]
f) Now, let $A \in \mathbb{R}^{M \times N}$ be a tall full rank matrix, $M>N$. Given a set of equations,

$$Ax = \mathbf{0}$$

there is generally no solution that satisfies all the equations exactly. However, we know that the least squares solution $x_{ls}$ minimizes the norm of the error $\|Ax_{ls} - y\|$

In 16A we learned that the solution has a closed form:

$$x_{ls} = (A^TA)^{-1}A^Ty$$

In that case, we can say that $(A^TA)^{-1}A^T$ is a pseudo-inverse of $A$.

Show that $(A^TA)^{-1}A^T = A^* = V^* S' U^T$

Solution:

Note that $A$ is tall, so

$$A = U_1 S V_1^T$$

Now $V_1 \in \mathbb{R}^{N \times N}$ is square and orthonormal. Also,

$U_1 \in \mathbb{R}^{M \times N}$ is tall and orthonormal so $U_1^T U_1 = I_{N \times N}$

So,

$$(A^TA) = V_1 S U_1^T U_1 S V_1^T = V_1 S^2 V_1^T$$

$$(A^TA)^{-1} = V_1 S^{-2} V_1^T$$

$$(A^TA)^{-1}A^T = V_1 S^{-2} V_1^T V_1^T S U_1^T = V_1 S^{-2} S^2 U_1^T = V_1 S^{-1} U_1^T$$
$A^+ = U S^{-1} U^T$ is also called the "Moore–Penrose Pseudo-Inverse"

The same equation using the SVD of $A$ can be used for both tall and fat matrices.

When $A$ is tall, $A^+ y$ will be the least squares solution.

When $A$ is fat, $A^+ y$ will be the minimum norm solution.

Same-same, but different!