1. **SVD Short Questions**  Assume we have the compact form of the SVD of $A = U_1S_1V_1^T = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^T$.

(a) Compute $AV_1V_1^T$

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

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,

$$||A||_F = \sqrt{Tr\{A^TA\}}$$

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

$$||UA||_F = ||AV||_F = ||A||_F$$

(c) Show that $||A||_F = \sqrt{\sum_{i=1}^{N} \sigma_i^2}$
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). $A = U_1 S V_1^T$ is the SVD of $A$. What are the sizes of $U_1$, $S$, $V_1$?

**Solution:**

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

$$A = \begin{bmatrix} U_1 & \tilde{S} & V_1^T \end{bmatrix}$$

So, $U_1 \in \mathbb{R}^{M \times M}$, $S \in \mathbb{R}^{M \times M}$, $V_1 \in \mathbb{R}^{M \times M}$.

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 $A^+ = V_1 \tilde{S}^{-1} U_1^T$.

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

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

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

$$A \hat{x} = A A^T \hat{y} = U S V_s^T V_s S^{-1} U_s^T \hat{y} = U S S^{-1} U_s^T \hat{y} = U U_s^T \hat{y} = \hat{y}$$

$$V_s$$

$$\left( V_s V_s^T \neq I_{n \times n} \right)$$

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

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

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

Solution:

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

$$A \bar{x} = U_s S V_s^T \bar{x} = 0$$

Since $S$ has non-zero diagonals, this is true only if $V_s^T \bar{x} = 0$.
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 $\hat{x} \mid A \hat{x} = y$. If $\hat{x} \neq \check{x}$, then $\|\hat{x}\| > \|\check{x}\|$. 

**Solution:**

Let $A = U \Sigma V^T$ be the full SVD. $V = \begin{bmatrix} V_1 & V_\perp \end{bmatrix}$

If $\hat{x} \neq \check{x}$, then $\hat{x} = \check{x} + \tilde{x}$

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

$$\|\hat{x}\|^2 = \|VV^T \tilde{x}\|^2 = \|VV^T (\check{x} + \tilde{x})\|^2 = \|V_1 V_1^T \check{x} + V_\perp V_\perp^T \tilde{x}\|^2$$

From part (C),

$$= \|V_1 V_1^T \check{x} + V_\perp V_\perp^T \tilde{x}\|^2$$

V. $\|V_\perp\|$ is so

$$= \|V_1 V_1^T \check{x}\|^2 + \|V_\perp V_\perp^T \tilde{x}\|^2 = \|\check{x}\|^2 + \|\tilde{x}\|^2 > \|\hat{x}\|^2$$
e) From

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

Solution:

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

Let $\mathbf{A} = \begin{bmatrix} 2 & 2 & 1 & 1 \\ 1 & 1 & 2 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 3\sqrt{2} & 0 & 0 & 0 \\ 0 & \sqrt{2} & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & -\frac{1}{2} & -\frac{1}{\sqrt{2}} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & -\frac{1}{\sqrt{2}} \end{bmatrix}^T
\]

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

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

Solution:

\[
\mathbf{x} = \mathbf{A}^T \mathbf{y} = \mathbf{V}_i \mathbf{S}^{-1} \mathbf{U}_i^T
\]

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

\[
\mathbf{S}^{-1} \mathbf{U}_i^T = \frac{1}{\sqrt{2}} \begin{bmatrix} \frac{1}{3} & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{2} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -1 & 1 \end{bmatrix}
\]

\[
\mathbf{V}_i \cdot \mathbf{S}^{-1} \mathbf{U}_i^T = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \cdot \frac{1}{2} \begin{bmatrix} \frac{1}{3} & \frac{1}{3} \\ -1 & 1 \end{bmatrix} = \frac{1}{4} \begin{bmatrix} \frac{4}{3} & \frac{4}{3} \\ -\frac{2}{3} & -\frac{2}{3} \end{bmatrix}
\]

\[
\mathbf{A}^T = \frac{1}{6} \begin{bmatrix} 2 & -1 \\ -1 & 2 \\ -1 & 2 \\ -1 & 2 \end{bmatrix}
\]

and

\[
\mathbf{x} = \mathbf{A}^T \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \frac{1}{3} \begin{bmatrix} 2 \\ 2 \\ 2 \\ -1 \end{bmatrix}
\]
f) Now, let \( A \in \mathbb{R}^{M \times N} \) be a tall full rank matrix, \( M > N \). Given a set of equations,
\[
Ax = \hat{y}
\]
there is generally no solution that satisfies all the equations exactly. However, we know that the least squares solution \( \hat{x}_{LS} \) minimizes the norm of the error \( \|A \hat{x}_{LS} - \hat{y}\| \).

In 16A we learned that the solution has a closed form:
\[
\hat{x}_{LS} = (A^TA)^{-1}A^T\hat{y}
\]

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

Show that \((A^TA)^{-1}A^T = A^T = V_i S^t U_i^T\)

Solution:

Note that \( A \) is tall, so,
\[
A = U_1 \begin{bmatrix} S \end{bmatrix} V_i^T
\]

Now \( V_i \in \mathbb{R}^{N \times N} \) is square and orthonormal. Also,
\( U_i \in \mathbb{R}^{M \times N} \) is tall and orthonormal so \( U_i^T U_i = I_{N \times N} \).

So,
\[
(A^TA) = V_i S U_i^T U_i S V_i^T = V_i S^2 U_i^T
\]

\[
(A^TA)^{-1} = V_i S^{-2} U_i^T
\]

\[
(A^TA)^{-1}A^T = V_i S^{-2} U_i^T V_i^T S U_i^T = V_i S^{-2} S^t U_i^T = U_i S^{-t} U_i^T
\]
\[ A^T = U^T S^{-1} U \] 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, \( A^T y \) will be the least squares solution.

When A is fat, \( A^T y \) will be the minimum norm solution.

Same-same, but different!