1. **Rules of vector SVD**

   a) \( A = 1(1, 0)(1, 0)^T + 3(0, 1)(0, 1)^T \) is not an SVD since \( 1 < 3 \), but singular values must be in decreasing order.

   b) \( A = 4(1, 0)(0, 1)^T + 3(0, 1)(1, 0)^T \) is an SVD.

   c) \( A = 3(1, -1)(1, 0)^T + 2(1, 1)(0, 1)^T \) is not an SVD, since \((1, -1)\) and \((1, 1)\) are not unit vectors.

   d) \( A = -3(1/\sqrt{2}, -1/\sqrt{2}, 0)(1, 0)^T + 2(0, 0, 1)(0, 1)^T \) is not an SVD since \(-3 < 0\), but singular values must be positive.

   e) \( A = 3(-1/\sqrt{2}, 1/\sqrt{2}, 0)(1, 0)^T + 2(0, 0, 1)(0, 1)^T \) is an SVD.

   f) \( A = 5(1, 0, 0)(0, 1)^T + 3(0, 1, 0)(1, 0)^T + 2(0, 0, 1)(0, 1)^T \) is not an SVD, since the vectors \((0, 1), (1, 0), (0, 1)\) are not orthogonal.
2. The matrix SVD Suppose that $A$ is an $m \times n$ matrix of rank $r$, with SVD $A = U\Sigma V^T$.

   a) $U$ is an $m \times m$ matrix, $\Sigma$ is a $m \times n$ matrix, and $V$ is a $n \times n$ matrix. The matrices $U$ and $V$ are orthogonal matrices. The first $r$ diagonal entries of $\Sigma$ are $> 0$.

   b) $A^T A = (U\Sigma V^T)^T U\Sigma V^T = V\Sigma^T U^T U\Sigma V^T = V(\Sigma^T \Sigma)V^T$. Therefore $Q_1 = V$ and $D_1 = \Sigma^T \Sigma$. The columns of $V$ are eigenvectors of $A^T A$, and the eigenvalues are the diagonal entries of the $n \times n$ matrix $\Sigma^T \Sigma$, which are $\sigma^2_1, \ldots, \sigma^2_r, 0, \ldots, 0$.

   c) $AA^T = U\Sigma V^T (U\Sigma V^T)^T = U\Sigma V^T V\Sigma^T U^T = U(\Sigma\Sigma^T)U^T$. Therefore $Q_2 = U$ and $D_2 = \Sigma\Sigma^T$. The columns of $U$ are eigenvectors of $AA^T$, and the eigenvalues are the diagonal entries of the $n \times n$ matrix $\Sigma\Sigma^T$, which are $\sigma^2_1, \ldots, \sigma^2_r, 0, \ldots, 0$.

   d) $V^T v_i = (v_i^1 \cdot v_i, \ldots, v_i^j \cdot v_i, \ldots, v_i^n \cdot v_i) = (0, \ldots, 1, \ldots, 0), \Sigma V^T v_i = \Sigma(0, \ldots, 1, \ldots, 0) = (0, \ldots, \sigma_i, \ldots, 0), Av_i = U\Sigma V^T v_i = U(0, \ldots, \sigma_i, \ldots, 0) = \sigma_i U e_i = \sigma_i u_i$. 
3. Computing the vector SVD

a) \( A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} \). The matrix \( A^T A = \begin{pmatrix} 9 & 0 \\ 0 & 1 \end{pmatrix} \) has eigenvalues \( \lambda_1 = 9, \lambda_2 = 1 \), with eigenvectors \( v_1 = (1, 0) \) and \( v_2 = (0, 1) \). The singular values are \( \sigma_1 = \sqrt{9} = 3 \) and \( \sigma_2 = \sqrt{1} = 1 \). The left singular vectors are \( u_1 = \frac{1}{3} A v_1 = \frac{1}{3}(0, 3) = (0, 1) \) and \( u_2 = \frac{A v_2}{\sigma_2} = (-1, 0) \). The vector SVD is
\[
A = 3(0, 1)(1, 0)^T + 1(-1, 0)(0, 1)^T.
\]

b) \( A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix} \). The matrix \( A^T A = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 1 \end{pmatrix} \) has eigenvalues \( \lambda_1 = 9, \lambda_2 = 4, \lambda_3 = 1, \lambda_4 = 0 \). The orthonormal eigenvectors for the non-zero eigenvalues are \( v_1 = (0, 1, 0, 0) \), \( v_2 = (1, 0, 0, 0) \), and \( v_3 = (0, 0, 0, 1) \). The singular values are \( \sigma_1 = 3, \sigma_2 = 2, \sigma_3 = 1 \). The left singular vectors are \( u_1 = \frac{1}{2} A v_1 = (0, 0, -1), u_2 = \frac{1}{2} A v_2 = (1, 0, 0), u_3 = \frac{1}{1} A v_3 = (0, 1, 0) \). The vector SVD is
\[
A = 3(0, 0, -1)(0, 1, 0, 0)^T + 2(1, 0, 0)(1, 0, 0, 0)^T + 1(0, 1, 0)(0, 0, 0, 1)^T.
\]

c) \( A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} \). The matrix \( A^T A = \begin{pmatrix} 4 & 2 \\ 2 & 5 \end{pmatrix} \) has characteristic polynomial \( \lambda^2 - 9\lambda + 16 \), with eigenvalues \( \lambda_1 = \frac{9 + \sqrt{17}}{2}, \lambda_2 = \frac{9 - \sqrt{17}}{2} \). We then have eigenvectors
\[
v_1 = \frac{(-2.4 + \lambda_1)}{\sqrt{\lambda_1}} \approx (-0.615, -0.788), \quad v_2 = \frac{(-2.4 - \lambda_2)}{\sqrt{\lambda_2}} \approx (-0.788, 0.615).
\]
The singular values are \( \sigma_1 = \sqrt{\lambda_1} \approx 2.562 \) and \( \sigma_2 = \sqrt{\lambda_2} \approx 1.562 \).
The left singular vectors are \( u_1 = \frac{A v_1}{\sigma_1} \approx (-0.788, -0.615) \) and \( u_2 = \frac{A v_2}{\sigma_2} \approx (-0.615, 0.788) \).
The vector SVD is, approximately,
\[
A = 2.562(-0.788, -0.615)(-0.615, -0.788)^T + 1.562(-0.615, 0.788)(-0.788, 0.615)^T.
\]
(The fact that the \( u \) and \( v \) vectors look so similar seems to be a coincidence.)
4. Computing the matrix SVD

Compute the matrix SVD of each of the following matrices:

a) \( A = \begin{pmatrix} 0 & -1 \\ 3 & 0 \end{pmatrix} \). Since \( m = n = r = 2 \), we can just use the singular vectors and values found in problem 3a) (no need to find ONB for Nul(A) or Nul(\( A^T \)). We have

\[
A = U\Sigma V^T = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \end{pmatrix}^T.
\]

b) \( A = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -3 & 0 & 0 \end{pmatrix} \). We found the vector SVD of this in 3b). Since \( m = 3, n = 4, r = 3 \), we need to find the additional vector \( v_4 \), an ONB of Nul(A). Since the matrix \( A^T A \) was rather simple, and \( \text{Nul}(A^T A) = \text{Nul}(A) \), we can use that \( A^T A \) had unit eigenvector \( v_4 = (0, 0, 1, 0) \) for the eigenvalue \( \lambda_4 = 0 \). Then

\[
A = U\Sigma V^T = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 1 & 2 \\ -1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}^T.
\]

c) \( A = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} \). The matrix \( A^T A = \begin{pmatrix} 6 & 12 \\ 12 & 24 \end{pmatrix} \) has characteristic polynomial

\[
\lambda^2 - 30\lambda, \text{ with eigenvalues } \lambda_1 = 30 \text{ and } \lambda_2 = 0. \text{ The } \lambda_1\text{-eigenvector equals } v_1 = \frac{(12, 24)}{||12, 24||} = \frac{(1, 2)}{||1, 2||} = \frac{1}{\sqrt{5}}(1, 2), \text{ while the } \lambda_2\text{-eigenvector equals } v_2 = \frac{(12, -6)}{||12, -6||} = \frac{1}{\sqrt{5}}(2, -1).
\]

The only singular value is \( \sigma_1 = \sqrt{\lambda_1} = \sqrt{30} \). We find the left singular vector \( u_1 = \frac{1}{\sqrt{30}} A v_1 = \frac{1}{5\sqrt{6}}(5, 5, 10) = \frac{1}{\sqrt{6}}(1, 1, 2). \)

We already found the vector \( v_2 \) spanning Nul(\( A^T A \)) = Nul(A). It remains to find an ONB \( u_2, u_3 \) of Nul(\( A^T \)). It is not hard to see that \( (1, -1, 0) \) and \( (2, 0, -1) \) are a basis of Nul(\( A^T \)), but they are not orthonormal.

Doing Gram-Schmidt, we first replace \( (1, -1, 0) \) with \( (1, -1, 0) \) and replace \( (2, 0, -1) \) with \( (2, 0, -1) - \frac{(2, 0, -1)(1, -1, 0)}{(1, -1, 0)(1, -1, 0)}(1, -1, 0) = (2, 0, -1) - (1, -1, 0) = (1, 1, -1). \)

(I avoided using the usual names for vectors in Gram-Schmidt, since it would be easy to confuse with the \( u \) and \( v \) vectors of SVD, which have a totally different meaning).

Making these unit vectors, we find that \( u_2 = \frac{1}{\sqrt{2}}(1, -1, 0) \) and \( u_3 = \frac{1}{\sqrt{3}} (1, 1, -1) \) form an ONB of Nul(\( A^T \)).

We conclude that

\[
U = \begin{pmatrix} 1/\sqrt{6} & 1/\sqrt{2} & 1/\sqrt{3} \\ 1/\sqrt{6} & -1/\sqrt{2} & 1/\sqrt{3} \\ 2/\sqrt{6} & 0 & -1/\sqrt{3} \end{pmatrix}, V = \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}, \Sigma = \begin{pmatrix} \sqrt{30} & 0 \\ 0 & 0 \end{pmatrix}.
\]
Warning: Many other answers are possible for \( U \) and \( V \). Your columns of \( V \) might be off by a sign, your first column of \( U \) might be off by a sign, and the final two columns of \( U \) can look quite different.
5. **Sums of rank 1 matrices**

This final problem is not about SVDs, but just about sums of rank one matrices.

a) Without computing \( A \), we will explain why
\[
A = (1, 2, 1)(1, 1)^T + (1, -1, 1)(-1, 1)^T
\]

is a rank 2 matrix.

Since \( A(1, 1) = (1, 2, 1)((1, 1)(1, 1)) + (1, -1, 1)((-1, 1)(1, 1)) = 2(1, 2, 1) + 0 \), the vector \( (1, 2, 1) \) is in the column space of \( A \). Similarly, \( A(-1, 1) = (1, 2, 1)((-1, 1)(1, 1)) + (1, -1, 1)((-1, 1)(-1, 1)) = 0 + 2(-1, -1, 1) \), so \( 2(-1, -1, 1) \) is also in the column space of \( A \) since these two column space vectors are linearly independent, the rank of \( A \) is at least 2. Since \( A \) is a \( 3 \times 2 \) matrix, its rank is at most 2. Therefore the rank of \( A \) equals 2.

b) If \( A = u_1 v_1^T + \cdots + u_r v_r^T \) for some vectors \( u_i \in \mathbb{R}^m \) and \( v_j \in \mathbb{R}^n \), we will explain why the rank of \( A \) is at most \( r \).

For any vector \( x \), \( Ax = (v_1 \cdot x)u_1 + \cdots + (v_r \cdot x)u_r \). Therefore any vector \( b \) for which \( Ax = b \) is consistent must be a linear combination of \( u_1, \ldots, u_r \). In other words, \( \text{Col}(A) \subseteq \text{Span}\{u_1, \ldots, u_r\} \). Since \( \dim \text{Span}\{u_1, \ldots, u_r\} \leq r \), and \( \dim \text{Col}(A) \leq \dim \text{Span}\{u_1, \ldots, u_r\} \), we conclude that \( \text{rank}(A) = \dim \text{Col}(A) \leq r \).

c) Suppose that the vectors \( u_1, \ldots, u_r \in \mathbb{R}^m \) are a linearly independent set of vectors, and the vectors \( v_1, \ldots, v_r \in \mathbb{R}^n \) are also linearly independent. We consider the matrix \( A = u_1 v_1^T + \cdots + u_r v_r^T \).

Since the \( v_i \) vectors are linearly independent, \( \text{Span}\{v_1, \ldots, v_r\} \) is \( r \)-dimensional, while \( \text{Span}\{v_2, \ldots, v_r\} \) is \( (r - 1) \)-dimensional. Using Gram-Schmidt, we can find a vector \( v \in \text{Span}\{v_1, \ldots, v_r\} \) which is orthogonal to \( \text{Span}\{v_2, \ldots, v_r\} \) but not orthogonal to \( v_1 \) (i.e. we can project \( v_1 \) onto the orthogonal complement of \( \text{Span}\{v_2, \ldots, v_r\} \)).

Using this vector \( v \), \( Av = (v_1 \cdot v)u_1 + \cdots + (v_r \cdot v)u_r = (v_1 \cdot v)u_1 \), since \( v \cdot v_2 = 0, v \cdot v_3 = 0, \ldots \). In other words, \( A \left( \frac{v}{v} \right) = u_1 \), which verifies that \( u_1 \) is in \( \text{Col}(A) \).

A similar argument shows that each of the vectors \( u_i \) is in \( \text{Col}(A) \). Therefore \( \text{Span}\{u_1, \ldots, u_r\} \subseteq \text{Col}(A) \). Since the \( u_i \) vectors are linearly independent, \( r = \text{Span}\{u_1, \ldots, u_r\} \leq \dim \text{Col}(A) = \text{rank}(A) \). On the other hand, by b), \( \text{rank}(A) \leq r \). Therefore \( \text{rank}(A) = r \).