1. SVD

Suppose we have a matrix $A \in \mathbb{R}^{m \times n}$ with rank $r$.

We define the compact SVD as follows:

$$A_{m \times n} = U_r \Sigma_r V_r^T_{r \times r \times r \times n}.$$

Here, $\Sigma_r \in \mathbb{R}^{r \times r}$ is a diagonal matrix containing non-zero singular values of $A$.

$$\Sigma_r = \begin{bmatrix} \sigma_1 & \ldots & \ldots & \sigma_r \\ \end{bmatrix},$$

with $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$.

Next, $U_r \in \mathbb{R}^{m \times r}$ is given by,

$$U_r = [u_1, u_2, \ldots, u_r],$$

where $u_i$ is a left singular vector corresponding to non-zero singular value, $\sigma_i$, for $i = 1, 2, \ldots, r$. The columns of $U_r$ are orthonormal and together they span the columnspace of $A$.

Finally, $V_r^T \in \mathbb{R}^{r \times n}$ is given by,

$$V_r^T = \begin{bmatrix} v_1^T \\ v_2^T \\ \vdots \\ v_r^T \end{bmatrix},$$

where $v_j$ is a right singular vector corresponding to non-zero singular value, $\sigma_j$ for $j = 1, 2, \ldots, r$. The rows of $V_r^T$ are orthonormal and span the rowspace of $A$. Equivalently the columns of $V_r$ span the column space of $A^T$.

The matrix $A$ can be expressed as,

$$A = \sigma_1 u_1 v_1^T + \sigma_2 u_2 v_2^T + \ldots + \sigma_r u_r v_r^T.$$

Assume now that $m \geq n$.

Another type of SVD which might be more familiar is the full SVD of $A$ which is defined as follows:

$$A_{m \times n} = U_{m \times m} \Sigma_{m \times n} V_{n \times n}^T.$$

Here, $\Sigma \in \mathbb{R}^{m \times n}$ has non-diagonal entries as zero. The diagonal entries of $\Sigma$ contain the singular values and we can write $\Sigma$ in terms of $\Sigma_r$ as,

$$\Sigma = \begin{bmatrix} \Sigma_r & 0_{r \times (n-r)} \\ 0_{(m-r) \times r} & 0_{(m-r) \times (n-r)} \end{bmatrix}.$$
Next, \( U \in \mathbb{R}^{m \times m} \) is an orthogonal matrix. \( U \) can be expressed in terms of \( U_r \) as,

\[
U = \begin{bmatrix} U_r & u_{r+1} & \cdots & u_m \end{bmatrix}_{m \times r} \begin{bmatrix} u_{r+1} & \cdots & u_m \end{bmatrix}_{m \times (m-r)}
\]

The columns \( u_{r+1}, u_{r+2}, \ldots, u_n \) are left singular vectors corresponding to singular value 0, and together span the nullspace of \( A^\top \).

Finally, \( V^\top \) is an orthogonal matrix and can be expressed in terms of \( V_r^\top \) as,

\[
V^\top = \begin{bmatrix} V_r^\top & v_{r+1}^\top & \cdots & v_n^\top \end{bmatrix}_{r \times n} \begin{bmatrix} v_{r+1}^\top & \cdots & v_n^\top \end{bmatrix}_{(n-r) \times n}
\]

The rows \( v_{r+1}^\top, v_{r+2}^\top, \ldots, v_n^\top \) when transposed are the right singular vectors corresponding to singular value of 0 and together span the nullspace of \( A \).

(a) For this problem assume that \( m > n > r \). Which of the following are True:

i. \( UU^\top = I \)
   
   **Solution:** True. \( UU^\top = I_m \) because \( U \) is an orthogonal matrix.

ii. \( U^\top U = I \)
   
   **Solution:** True. \( U^\top U = I_m \) because \( U \) is an orthogonal matrix.

iii. \( V^\top V = I \)
   
   **Solution:** True. \( V^\top V = I_n \) because \( V \) is an orthogonal matrix.

iv. \( VV^\top = I \)
   
   **Solution:** True. \( VV^\top = I_n \) because \( V \) is an orthogonal matrix.

v. \( U_r^\top U_r = I \)
   
   **Solution:** True. \( U_r^\top U_r = I_r \) because the columns of \( U_r \) are orthonormal.

vi. \( U_r^\top U_r^\top = I \)
   
   **Solution:** False. \( U_r U_r^\top \) is a \( m \times m \) matrix but has rank less than or equal to \( r \) (since \( U_r \) has rank \( r \) and product of matrices has rank less than or equal to minimum of individual ranks).

vii. \( V_r V_r^\top = I \)
   
   **Solution:** False. \( V_r V_r^\top \) is a \( n \times n \) matrix but has rank less than or equal to \( r \) (since \( V_r \) has rank \( r \) and product of matrices has rank less than or equal to minimum of individual ranks).

viii. \( V_r^\top V_r = I \)
   
   **Solution:** True. \( V_r^\top V_r = I_r \) because the columns of \( V_r \) are orthonormal.

(b) Going from the full SVD to compact SVD. Find the compact SVD of \( A \) which has the full SVD:

\[
A = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.
\]
Solution: The compact SVD of $A$ is given by:

$$A = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 0 \end{bmatrix}.$$

(c) Going from compact SVD to full SVD: Find the full SVD of $A$ which has the compact SVD:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

Solution: Observe that in this case, the full SVD of $A$ has $\Sigma$ and $V^\top$ as those in the compact SVD but $U \in \mathbb{R}^{3 \times 3}$. Thus we need to find a unit-norm column $u_3$ orthogonal to columns of $U_r$. We can use a system of linear equations to solve this. That is we want $u_3 = [x, y, z]$ so we must have

- $[1/\sqrt{2}, 1/\sqrt{2}, 0]^\top u_3 = 0$
- $[0, 0, 1]^\top u_3 = 0$
- $\|u_3\|_2 = 1$

Check that $u_3 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$ satisfies our requirements.

Thus the full SVD of $A$ is given by:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \begin{bmatrix} 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 2 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}.$$

Alternatively we can use Gram-Schmidt procedure to find $u_3$. This has the added advantage of being useful when we want to find the full SVD when more than one singular vector is missing.

2. SVD Part 2

Consider $A$ to be the $4 \times 3$ matrix

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix} \tag{1}$$

where $a_i$ for $i \in \{1, 2, 3\}$ form a set of orthogonal vectors satisfying $\|a_1\|_2 = 3$, $\|a_2\|_2 = 2$, $\|a_3\|_2 = 1$.

(a) What is the SVD of $A$? Express it as $A = U\Sigma V^\top$, with $\Sigma$ the diagonal matrix of singular values ordered in decreasing fashion, and explicitly describe $U$ and $V$.

Solution: The SVD of $A = U\Sigma V^\top$. Due to the orthogonality of the $a_i$ we have that

$$A^\top A = V\Sigma^2 V^\top = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{2}$$
Thus \( V = I \) and \( \Sigma = \text{diag}(3, 2, 1) \). Finally we have that \( U = A\Sigma^{-1} \) which becomes

\[
U = \begin{bmatrix}
\alpha_1 & \sqrt{2} & \alpha_2 \\
\alpha_3 & \alpha_4 & \alpha_5 \\
\end{bmatrix}
\] (3)

(b) What is the dimension of the null space, \( \text{dim}(\text{null}(A)) \)?

**Solution:** From part (a) all of the singular values of the \( A \) are non-zero. So the dimension of the null space is 0. Alternatively, all the columns of \( A \) are orthogonal – so no (non-zero) linear combination of them can equal zero.

(c) What is the rank of \( A \), \( \text{rank}(A) \)? Provide an orthonormal basis for the range of \( A \).

**Solution:** The rank of \( A \) is simply the number of non-zero singular values. So \( \text{rank}(A) = 3 \). The columns of \( U \) (defined above) provide an orthonormal basis for the range of \( A \).

(d) Let \( I_3 \) denote the \( 3 \times 3 \) identity matrix. Consider the matrix \( \tilde{A} = \begin{bmatrix} A \\ I_3 \end{bmatrix} \in \mathbb{R}^{7 \times 3} \). What are the singular values of \( \tilde{A} \) (in terms of the singular values of \( A \))?

**Solution:** We have that \( \tilde{A}^\top \tilde{A} = A^\top A + I_3 = V(\Sigma^2 + I_3)V^\top \). Hence if we denote \( \sigma_i \) as the singular values of \( A \) then the singular values of \( \tilde{A} \) are \( \tilde{\sigma}_i = \sqrt{\sigma_i^2 + 1} \) which are \( \sqrt{10}, \sqrt{5}, \sqrt{2} \).

(e) (Optional) Find an SVD of the matrix \( \tilde{A} \).

**Solution:** The SVD of \( \tilde{A} = \tilde{U}\tilde{\Sigma}\tilde{V}^\top \) has \( \tilde{\Sigma} = \text{diag}(\tilde{\sigma}_1, \ldots, \tilde{\sigma}_m) \); with \( \tilde{\sigma}_i = \sqrt{\sigma_i^2 + 1} \) where \( \sigma_i \) are the singular values of \( A \).

The eigenvectors of \( \tilde{A}\tilde{A}^\top \) form the columns of matrix \( \tilde{U} \), while the eigenvectors of \( \tilde{A}^\top \tilde{A} \) form the columns of \( \tilde{V} \). We can see this by writing out \( \tilde{A} \) in terms of the SVD of \( A \) and the identity matrix \( I \).

Since \( \tilde{A}^\top \tilde{A} = A^\top A + I \), we can choose \( \tilde{V} = V \). The condition for \( (\tilde{u}, \tilde{v}) \) to be a pair of left- and right-singular vectors of \( \tilde{A} \) are that both vectors must be unit-norm, and

\[
\tilde{A}\tilde{v} = \tilde{\sigma}\tilde{u}, \quad \tilde{A}^\top \tilde{u} = \tilde{\sigma}\tilde{v}.
\]

We have seen that we can choose \( \tilde{v} = v \) to be an eigenvector of \( A^\top A \) (that is, a right singular vector of \( A \)). Further, decomposing \( \tilde{u} = \begin{bmatrix} \tilde{u}_1 \\ \tilde{u}_2 \end{bmatrix} \), with \( \tilde{u}_1 \in \mathbb{R}^4 \) and \( \tilde{u}_2 \in \mathbb{R}^3 \), we obtain

\[
Av = \tilde{\sigma}\tilde{u}_1, \quad v = \tilde{\sigma}\tilde{u}_2, \quad A^\top \tilde{u}_1 + \tilde{u}_2 = \tilde{\sigma}v.
\]

Solving for the second equation: \( \tilde{u}_2 = v/\tilde{\sigma} \), we obtain from the third \( A^\top \tilde{u}_1 = (\tilde{\sigma} - 1/\tilde{\sigma})v \).

Multiplying by \( A \), and with the first equation, we then obtain

\[
AA^\top \tilde{u}_1 = \tilde{\sigma}(\tilde{\sigma} - 1/\tilde{\sigma})\tilde{u}_1 = \sigma^2\tilde{u}_1.
\]

This shows that we can set \( \tilde{u}_1 \) to be proportional to a left singular vector \( u \) of \( A \), and \( \tilde{u}_2 = v/\tilde{\sigma} \) proportional to \( v \). We have

\[
\tilde{u} = \begin{bmatrix} \alpha u \\ \frac{1}{\tilde{\sigma}}v \end{bmatrix},
\]

where \( \alpha \) must be chosen so that the above has unit Euclidean norm, that is:

\[
\alpha = \frac{\sqrt{\tilde{\sigma}^2 - 1}}{\tilde{\sigma}} = \frac{\sigma}{\sqrt{\sigma^2 + 1}}.
\]
We have obtained that a generic pair of left- and right singular vectors \((\tilde{u}, \tilde{v})\) of \(\tilde{A}\) corresponding to the singular value \(\sqrt{\sigma^2 + 1}\), can be constructed from a generic pair of left- and right singular vectors \((u, v)\) of \(A\) corresponding to the singular value \(\sigma\), with the choice

\[
\tilde{u} = \begin{bmatrix}
\frac{\sigma}{\sqrt{\sigma^2 + 1}} u \\
\frac{1}{\sqrt{\sigma^2 + 1}} v
\end{bmatrix}, \quad \tilde{v} = v.
\]