Calculating the Singular Value Decomposition

In lecture, we learned how to find the SVD of a wide matrix, \( A \) of dimension \( m \times n \) \((n > m)\). To decompose \( A \) into \( U \Sigma V^T \) we took the following steps:

(a) Compute the symmetric matrix \( A^T A \) with dimension \( n \times n \).

(b) Find the eigenvalues \( (\lambda_1, \lambda_2, \ldots, \lambda_m) \) and eigenvectors \( (\vec{v}_1, \vec{v}_2, \ldots, \vec{v}_n) \) of \( A^T A \). By the spectral theorem for real symmetric matrices, these eigenvectors are orthonormal.

(c) \( \sigma_i = \sqrt{\lambda_i} \) where \( \lambda_i \) are the sorted in descending order eigenvalues of \( A^T A \). We know these are all non-negative because \( (A\vec{v}_i)^T (A\vec{v}_i) = \|A\vec{v}_i\|^2 \) and \( (A\vec{v}_i)^T (A\vec{v}_j) = \vec{v}_i^T (A^T A) \vec{v}_j = \lambda_i \vec{v}_i^T \vec{v}_j = \lambda_i \). The corresponding normalized eigenvectors \( \vec{v}_i \) form the \( V \) matrix.

(d) Using \( \sigma_i \) and \( \vec{v}_i \) we can find the corresponding vectors of the \( U \) matrix, \( \vec{u}_i \) by computing \( \vec{u}_i = \frac{A\vec{v}_i}{\sigma_i} \).

These are normalized since \( \sigma_i = \|A\vec{v}_i\| \) by the argument above, and orthogonal since \( (A\vec{v}_i)^T (A\vec{v}_j) = \vec{v}_i^T (A^T A) \vec{v}_j = \lambda_i \vec{v}_i^T \vec{v}_j = 0 \) if \( i \neq j \), since \( V \) is an orthonormal matrix.

Using this knowledge, if we were asked find the SVD of a tall matrix \( A \) of dimension \( m \times n \) such that \( m > n \), one option would be to take the transpose of this matrix to get a wide matrix, proceed as described above, and transpose the result.

In this discussion we will better understand how to compute the SVD for any matrix.

Questions

1. Understanding the SVD

   We can compute the SVD for a wide matrix \( A \) with dimension \( m \times n \) where \( n > m \) using \( A^T A \) with the method described above. However, when doing so you may realize that \( A^T A \) is much larger than \( AA^T \) for such wide matrices. This makes it more efficient to find the eigenvalues for \( AA^T \). In this question we will explore how to compute the SVD using \( AA^T \) instead of \( A^T A \).

   (a) What are the dimensions of \( AA^T \) and \( A^T A \).

      \[ \text{Answer: } \dim(AA^T) = m \times m \]
      \[ \dim(A^T A) = n \times n \]

   (b) Given that the \( A = U\Sigma V^T \), find a symbolic expression for \( AA^T \).

      \[ \text{Answer: } AA^T = U\Sigma V^T V \Sigma^T U^T \]
      \[ V^T V = I \]
\[ AA^T = U \Sigma \Sigma^T U^T \]

(c) Using the solution to the previous part explain how to find \( U \) and \( \Sigma \) from \( AA^T \).

**Answer:** Knowing that \( AA^T \) is a symmetric matrix, we know that its normalized eigenvectors will be orthonormal.

From the properties of the SVD we know that \( U \) is an orthonormal matrix of dimension \( m \times m \) and \( \Sigma \Sigma^T \) is an \( m \times m \) diagonal matrix with the entries on the diagonal being \( \sigma_i^2 \).

Using the above information we can see that we can calculate \( U \) by diagonalizing the symmetric matrix \( AA^T \). By the spectral theorem for real symmetric matrices, we will get an orthonormal basis of eigenvectors. The square root of the corresponding eigenvalues of \( AA^T \) will give us the singular values \( \sigma_i \) (You can construct \( \Sigma \) by putting these on the diagonal of a zero matrix with the same dimensions as \( A \)) and the corresponding eigenvectors will form the \( U \) matrix.

(d) Now that we have found the singular values \( \sigma_i \) and the corresponding vectors \( \vec{u}_i \) in the matrix \( U \), devise a way to find the corresponding vectors \( \vec{v}_i \) in matrix \( V \).

**Answer:**

\[
\vec{v}_i = \frac{V \Sigma^T U^T \vec{u}_i}{\sigma_i} = \frac{A^T \vec{u}_i}{\sigma_i}
\]

(e) Now we have a way to find the vectors \( \vec{v}_i \) in matrix \( V \), verify that they are orthonormal.

**Answer:** To verify that \( \vec{v}_i \) in matrix \( V \) are orthonormal we show that:

i. \( \vec{v}_i \) are orthogonal to one another

ii. each \( \vec{v}_i \) has norm 1.

**Orthogonality:**

To show orthogonality we must show that any two vectors \( \vec{v}_i = \frac{A^T \vec{u}_i}{\sigma_i} \) and \( \vec{v}_j = \frac{A^T \vec{u}_j}{\sigma_j} \) with \( i \neq j \) have an inner product of zero.

\[
\vec{v}_i^T \vec{v}_j = \frac{u_i^T A A^T u_j}{\sigma_i \sigma_j}
\]

\[
= \frac{u_i^T A A^T u_j}{\sigma_i \sigma_j}
\]

\[
= \frac{u_i^T u_j}{\sigma_i \sigma_j}
\]

\[
= \frac{\sigma_j^2 u_i^T u_j}{\sigma_i \sigma_j}
\]

\[
= 0
\]

This is because we know that \( \vec{u}_i \) and \( \vec{u}_j \) are orthonormal as they are eigenvectors of a symmetric matrix \( AA^T \).

Thus for all \( i \neq j \)

\[
\vec{v}_i^T \vec{v}_j = 0
\]
Norm of 1: Follow the steps above with \( i = j \) to see

\[
\vec{v}_i^T \vec{v}_j = \vec{v}_i^T \vec{v}_i \quad (5)
\]

\[
= \frac{(\sigma_i)^2 \vec{u}_i^T \vec{u}_i}{\sigma_i \sigma_j} \quad (6)
\]

\[
= \frac{(\sigma_i)^2}{(\sigma_i)^2} \vec{v}_i^T \vec{v}_i \quad (7)
\]

\[
= 1 \quad (8)
\]

(f) Now that we have found \( \vec{v}_i \) you may notice that we only have \( m < n \) vectors of dimension \( n \). This is not enough for a basis. How would you complete the \( m \) vectors to form an orthonormal basis?

**Answer:** Gram Schmidt.

Just add in the standard basis for \( n \)-dimensional space, and orthonormalize. We know that this collection of \( n + m \) vectors spans the whole space, and so after orthonormalization, we will have a collection of orthonormal vectors that spans the whole space. Along the way, some vectors will be found to be linearly dependent on those that came before — that is fine, discard these. At the end, we will have \( n \) orthonormal vectors, the first set of which are the original \( \vec{v}_i \).

(g) Given that \( A = U \Sigma V^T \) verify that the vectors you found to extend the \( \vec{v}_i \) into a basis are in the nullspace of \( A \).

**Answer:** Let \( V = \begin{bmatrix} V_s & R \end{bmatrix} \) where \( V_s \) are the \( \{ \vec{v}_i \} \) we found using the \( \{ \vec{u}_i \} \) and \( R \) is composed of the remaining vectors found using gram schmidt. Let \( S \) be an \( m \times m \) diagonal square matrix with \( \sigma_i \) on the diagonal (\( \text{sigma}_i \) is allowed to be zero) such that \( \Sigma = \begin{bmatrix} S & 0 \end{bmatrix} \) where 0 denotes filling in the remaining matrix dimensions with zeros.

\[
A = U \Sigma V^T = U \begin{bmatrix} S & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} V_s^T \\ R^T \end{bmatrix}
\]

\[
AR = U \begin{bmatrix} S & 0 \\ 0 & R \end{bmatrix} \begin{bmatrix} V_s^T \\ R^T \end{bmatrix} R
\]

\[
= U \begin{bmatrix} S & 0 \\ 0 & R^T R \end{bmatrix}
\]

\[
= U \begin{bmatrix} 0 \\ R \end{bmatrix} = 0
\]

Thus everything in the subspace spanned by \( R \) maps to \( \vec{0} \), showing that the subspace is in the nullspace of \( A \).

(h) Using the previous parts of this question and what you learned from lecture write out a procedure on how to find the SVD for any matrix.

**Answer:** We calculate the SVD of matrix \( A \) as follows.

i. Pick \( A^T A \) or \( AA^T \) — whichever one is smaller.
If using $A^TA$, find the eigenvalues $\lambda_i$ of $A^TA$ and order them, so that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_n = 0$.

If using $AA^T$, find its eigenvalues $\lambda_1, \ldots, \lambda_m$ and order them the same way.

ii. If using $A^TA$, find orthonormal eigenvectors $\vec{v}_i$ such that $A^T A \vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, \ldots, r$

If using $AA^T$, find orthonormal eigenvectors $\vec{u}_i$ such that $AA^T \vec{u}_i = \lambda_i \vec{u}_i, \quad i = 1, \ldots, r$

iii. Set $\sigma_i = \sqrt{\lambda_i}$.

If using $A^TA$, obtain $\vec{u}_i$ from $\vec{u}_i = \frac{1}{\sigma_i} A \vec{v}_i, \quad i = 1, \ldots, r$.

If using $AA^T$, obtain $\vec{v}_i$ from $\vec{v}_i = \frac{1}{\sigma_i} A^T \vec{u}_i, \quad i = 1, \ldots, r$.

iii. If you want to completely construct the $U$ or $V$ matrix, complete the basis (or columns of the appropriate matrix) using Gram-Schmidt to get a full orthonormal matrix.

The full matrix form of SVD is taken to better understand the matrix $A$ in terms of the 3 nice matrices $U, \Sigma, V$. Often in practice, we do not completely construct the $U$ and $V$ matrices. After all, in many applications, we don’t need all the vectors.

2. SVD Example

Define the matrix

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

(a) Find the SVD of $A$ (compact form is fine).

**Answer:** First, compute $A^T A = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}$. The eigenvalues of $A^T A$ are 18 and 0, with corresponding unit eigenvectors

$$\vec{v}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \vec{v}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Therefore, $A$ has one singular value $\sqrt{18} = 3\sqrt{2}$

We obtain

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

and $A$ can be decomposed as

$$A = 3\sqrt{2} \begin{bmatrix} 1/3 & 1/\sqrt{2} \\ -2/3 & -1/\sqrt{2} \\ 2/3 \end{bmatrix}$$
(b) Find the rank of \( A \).

**Answer:** 1

(c) Find a basis for the nullspace of \( A \).

**Answer:**

\[
\text{span}\left\{ \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \right\}
\]

(d) Find a basis for the range (or columnspace) of \( A \).

**Answer:**

\[
\text{range}(A) = \text{span}\left\{ \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix} \right\}
\]

(e) Repeat parts (a) - (d), but instead, create the SVD of \( A^T \). What are the relationships between the answers for \( A \) and the answers for \( A^T \)?

**Answer:**

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

\[
AA^T = \begin{bmatrix} 9 & -9 \\ -9 & 9 \end{bmatrix}
\]

\[
\lambda = 18, 0
\]

\[
\vec{u}_1 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}, \vec{u}_2 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.
\]

\[
\vec{v}_1 = \frac{1}{\sigma_1} A^T \vec{u}_1 = \frac{1}{3\sqrt{2}} \begin{bmatrix} 1/\sqrt{2} \\ -2/3 \\ -2/3 \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} = \begin{bmatrix} 1/3 \\ -2/3 \\ 2/3 \end{bmatrix}
\]

\[
A = 3\sqrt{2} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/3 & -2/3 & 2/3 \end{bmatrix}
\]

At this point, we already know the rank is 1. The column space is also formed by the \( \vec{u} \) vector.

\[
\text{range}(A) = \text{span}\left\{ \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \right\}
\]

Two vectors in the nullspace are

\[
\begin{bmatrix} 2 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 2 \\ 0 \\ -1 \end{bmatrix}
\]

Note, if we had just noticed

\[
A^T = (U\Sigma V^T)^T = V\Sigma^T U^T
\]

We could’ve skippepd many steps for SVD calculation.
3. Eigenvalue Decomposition and Singular Value Decomposition

We define the Eigenvalue Decomposition as follows:

If a matrix $A \in \mathbb{R}^{n \times n}$ has $n$ linearly independent eigenvectors $\vec{p}_1, \ldots, \vec{p}_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$, then we can write:

$$A = P\Lambda P^{-1}$$

Where the columns of $P$ consist of $\vec{p}_1, \ldots, \vec{p}_n$, and $\Lambda$ is a diagonal matrix with diagonal entries $\lambda_1, \ldots, \lambda_n$.

For the sake of convenience, assume that these are sorted by their absolute values in descending order.

Consider a symmetric matrix $A = A^T \in \mathbb{R}^{n \times n}$. This is a symmetric matrix and thus has orthogonal eigenvectors and real eigenvalues (not necessarily non-negative).

Therefore its eigenvalue decomposition can be written as,

$$A = P\Lambda P^T$$

(a) First, assume $\lambda_i \geq 0, \forall i$. Find the SVD of $A$. **Answer:** Observe that,

$$A^TA = PA_2^T P^T$$

This means that,

$$\sigma_i = \lambda_i$$

We have,

$$Av_i = \lambda_i v_i$$

Plugging into our SVD condition $Av_i = \sigma_i u_i$:

$$\sigma_i v_i = \sigma_i u_i$$

This means that,

$$U = V = P$$

Therefore, in this case, the eigenvalue decomposition is the same as the singular value decompositions.

(b) Let one particular eigenvalue $\lambda_j$ be negative, with the associated eigenvector being $\vec{p}_j$. Succinctly,

$$A\vec{p}_j = \lambda_j \vec{p}_j$$

with $\lambda_j < 0$

We are still assuming that,

$$A = P\Lambda P^T$$

i. What is the singular value $\sigma_j$ associated to $\lambda_j$?

ii. What is the relationship between the left singular vector $u_j$, the right singular vector $v_j$ and the eigenvector $\vec{p}_j$?

**Answer:**
\[ \sigma_j = |\lambda_j| \]

ii. Either,
\[ \bar{u}_j = \bar{p}_j \text{ and } v_j = -p_j \]
or,
\[ u_j = -p_j \text{ and } v_j = p_j \]

This is because the diagonal entries of \( \Sigma \) MUST be non-negative.

**Extra Practice**

1. **More SVD**

Define the matrix
\[ A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}. \]

(a) Find the SVD of \( A \) (compact form is fine).

**Answer:** For a wide matrix, it is easier to pick \( AA^T \) to work with.

\[ AA^T = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix} \]

Next, we find the eigenvalues of the above matrix.

\[ \det(A - \lambda I) = \lambda^2 - 34\lambda + 225 = (\lambda - 25)(\lambda - 9) = 0 \]

Hence, the eigenvalues are \( \lambda_1 = 25 \) and \( \lambda_2 = 9 \), and the singular values are \( \sigma_1 = \sqrt{25} = 5 \) and \( \sigma_2 = \sqrt{9} = 3 \).

Next we find the left singular vectors (i.e. the columns of U). Finding the null\( (A - \lambda_1 I) \) and null\( (A - \lambda_2 I) \) will give us \( u_1 \) and \( u_2 \) respectively.

Hence, \( u_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \) and \( u_2 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{-1}{\sqrt{2}} \end{bmatrix} \) (don’t forget to normalize the vectors).

Lastly, we find the right singular vectors (or the columns of V)

\[ v_1 = \frac{1}{\sigma_1} A^T u_1 \]
\[ = \frac{1}{5} \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix} \]
\[ = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} \]

Similarly, we get \( v_2 = \begin{bmatrix} \frac{1}{\sqrt{18}} \\ \frac{1}{\sqrt{18}} \\ \frac{-1}{\sqrt{18}} \end{bmatrix} \).
(b) Find the rank of \( A \).

**Answer:** The rank of \( A \) is 2.

(c) Find a basis for the nullspace of \( A \).

**Answer:** The nullspace is 
\[
\begin{pmatrix}
-2 \\
2 \\
1
\end{pmatrix}
\]

Note: The nullspace of \( A \) is given by the orthogonal complement of space spanned by the columns of the \( V \) matrix. Hence, we can find the nullspace by finding the eigenvector(s) corresponding to the zero eigenvalue of the \( A^T A \) matrix as well.

(d) Find a basis for the range (or columnspace) of \( A \).

**Answer:** The columns of \( U \) give us the basis for range(\( A \))