Singular Value Decomposition

The definition

The SVD is a useful way to characterize a matrix. Let \( A \) be a matrix from \( \mathbb{R}^n \) to \( \mathbb{R}^m \) (or \( A \in \mathbb{R}^{m \times n} \)) of rank \( r \). It can be decomposed into a sum of \( r \) rank-1 matrices:

\[
A = \sum_{i=1}^{r} \sigma_i \mathbf{u}_i \mathbf{v}_i^T
\]

where

- \( \mathbf{u}_1, \ldots, \mathbf{u}_r \) are orthonormal vectors in \( \mathbb{R}^m \); \( \mathbf{v}_1, \ldots, \mathbf{v}_r \) are orthonormal vectors in \( \mathbb{R}^n \).
- the singular values \( \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0 \) are always real and positive.

We can also re-write the decomposition in matrix form:

\[
A = U_1 S V_1^T
\]

The properties of \( U_1, S \) and \( V_1 \) are,

- \( U_1 \) is an \([m \times r]\) matrix whose columns consist of \( \mathbf{u}_1, \ldots, \mathbf{u}_r \). Consequently,
  \[
  U_1^T U_1 = I_{r \times r}
  \]
- \( V_1 \) is an \([n \times r]\) matrix whose columns consist of \( \mathbf{v}_1, \ldots, \mathbf{v}_r \). Consequently,
  \[
  V_1^T V_1 = I_{r \times r}
  \]
- \( U_1 \) characterizes the column space of \( A \) and \( V_1 \) characterizes the row space of \( A \).
- \( S \) is an \([r \times r]\) matrix whose diagonal entries are the singular values of \( A \) arranged in descending order. The singular values are the square roots of the nonzero eigenvalues of \( A^T A \) (or, identically, \( A A^T \)).

The full matrix form of SVD is

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

where \( U^T U = I_{m \times m}, V^T V = I_{n \times n}, \Sigma \in \mathbb{R}^{m \times n} \), which contains \( S \) and elsewhere zero.
The calculation

We calculate the SVD of matrix $A$ as follows.

(a) Pick $A^T A$ or $AA^T$.

(b) i. If using $A^T A$, find the eigenvalues $\lambda_i$ of $A^T A$ and order them, so that $\lambda_1 \geq \cdots \geq \lambda_r > 0$ and $\lambda_{r+1} = \cdots = \lambda_m = 0$.

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

ii. If using $A^T A$, find orthonormal eigenvectors $\vec{v}_i$ such that

$$A^T \vec{v}_i = \lambda_i \vec{v}_i, \quad i = 1, \ldots, r$$

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

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

(c) If you want to completely construct the $U$ or $V$ matrix, complete the basis (or columns of the appropriate matrix) using Gram-Schmidt. Remember to orthonormalize afterwards.

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, we do not completely construct the $U$ and $V$ matrices.

Questions

1. SVD and Fundamental Subspaces

Define the matrix

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

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

(b) Find the rank of $A$.

(c) Find a basis for the kernel (or nullspace) of $A$.

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

(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 $A^T$?
2. Eigenvalue Decomposition and Singular Value Decomposition

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

Consider a matrix $A \in \mathbb{S}^n$, that is, $A = A^T \in \mathbb{R}^{n \times n}$. This is a symmetric matrix and has orthogonal eigenvectors. 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 $A$.

(b) Let one particular eigenvalue $\lambda_j$ be negative, with the associated eigenvector being $p_j$. Succinctly,

$$Ap_j = \lambda_j p_j \text{ with } \lambda_j < 0$$

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 $p_j$?

3. SVD and Induced 2-Norm

(a) Show that if $U$ is an orthogonal matrix then for any $\vec{x}$

$$\|U\vec{x}\| = \|\vec{x}\|.$$  

(b) Find the maximum

$$\max_{\{\vec{x}, \|\vec{x}\|=1\}} \|A\vec{x}\|$$

in terms of the singular values of $A$.

(c) Find the $\vec{x}$ that maximizes the expression above.

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).
(b) Find the rank of $A$.
(c) Find a basis for the kernel (or nullspace) of $A$.
(d) Find a basis for the range (or columnspace) of $A$. 