Singular Value Decomposition

The definition

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

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

where

- \( \bar{u}_1, \ldots, \bar{u}_r \) are orthonormal vectors in \( \mathbb{R}^m \);
- \( \bar{v}_1, \ldots, \bar{v}_r \) 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 \( \bar{u}_1, \ldots, \bar{u}_r \). Consequently,
  \[
  U_1^T U_1 = I_{r \times r}
  \]

- \( V_1 \) is an \([n \times r]\) matrix whose columns consist of \( \bar{v}_1, \ldots, \bar{v}_r \). 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, \( AA^T \)).

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 $A^TA$, find orthonormal eigenvectors $\vec{v}_i$ such that

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

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

(c) This is not in scope but 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 normalize 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.

## 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 column space) of $A$.

e) Repeat parts (a) - (d) for $A^T$ instead. 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 a SVD of $A$. 


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

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

We are still assuming that,

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

a) What is the singular value $\sigma_j$ associated to $\lambda_j$?

b) What is the relationship between the left singular vector $u_j$, the right singular vector $v_j$ and the eigenvector $p_j$?