

Singular Value Decomposition

An application to Big Data

Davide Sferrazza

January 8, 2023

Università degli Studi di Palermo

1 Singular Value Decomposition

- What is it?
- How can singular values be computed?

2 References

Singular Value Decomposition

Definition of SVD

Theorem

Given a matrix $A \in \mathbb{R}^{m \times n}$, it can always be found a decomposition such that

$$A = U\Sigma V^T$$

where $U \in \mathbb{R}^{m \times m}$, $V \in \mathbb{R}^{n \times n}$ and $\Sigma \in \mathbb{R}^{m \times n}$.

U and V are two orthogonal matrices and Σ is a diagonal matrix, namely:

$$(\Sigma)_{ij} = \begin{cases} 0, & i \neq j \\ \sigma_i, & i = j \end{cases}$$

where $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_p \geq 0$, $p = \min\{m, n\}$.

Definition of SVD

The non-zero entries of Σ , denoted by σ_i , are called *singular values*.

They are arranged in a nonincreasing order by convention.

The column vectors \mathbf{u}_i of U are called *left singular vectors* and those \mathbf{v}_i of V are called *right singular vectors*.

Since in general $m \neq n$, we have:

$$A = \sum_{i=1}^p \mathbf{u}_i \sigma_i \mathbf{v}_i^T$$

Theorem

If for some r such that $1 \leq r < p$ we have

$$\sigma_1 \geq \dots \geq \sigma_r > \sigma_{r+1} = \dots = \sigma_p = 0$$

then

- $\text{rank}(A) = r$
- $A = \sum_{i=1}^r \mathbf{u}_i \sigma_i \mathbf{v}_i^T$

This means that all other $p - r$ dimensions of matrix A are linear combinations of the first r .

Definition of SVD

Lower rank approximation

Let $A \in \mathbb{R}^{m \times n}$ be a matrix whose rank is $\text{rank}(A) = r$.

If for a fixed integer value $k < r$ we define

$$A_k = \sum_{i=1}^k \sigma_i \mathbf{u}_i \mathbf{v}_i^T \quad (1)$$

and

$$\mathcal{B} = \{B \in \mathbb{R}^{m \times n} : \text{rank}(B) = k\}$$

then

$$\min_{B \in \mathcal{B}} \|A - B\|_2 = \|A - A_k\|_2 = \sigma_{k+1}$$

This result tell us that A_k represents the best approximation (considering the *spectral norm*) of rank k of matrix A .

Singular values computation

To compute the singular values, consider the transpose of A given its decomposition:

$$A^T = (U\Sigma V^T)^T = V\Sigma^T U^T$$

The symmetric matrix $A^T A$ is equal to:

$$A^T A = (V\Sigma^T U^T)(U\Sigma V^T) = V\Sigma^T \Sigma V^T$$

Furthermore, this equation can be written as:

$$A^T A V = V \Sigma^T \Sigma$$

This means that the diagonal entries of the square matrix $\Sigma^T \Sigma$, which are the square of the singular values, are the eigenvalues of matrix $A^T A$ and V is the matrix of eigenvectors.

Singular values computation

Similarly, consider the product of AA^T . It is equal to:

$$AA^T = (U\Sigma V^T)(V\Sigma^T U^T) = U\Sigma\Sigma^T U^T$$

Which means that:






$$AA^T U = U\Sigma\Sigma^T$$

Hence U is the matrix of eigenvectors of AA^T .

Since $\text{rank}(A) = r$, only the first r eigenvalues of AA^T and $A^T A$ are non-zero.

References

References

-  Aggarwal, Charu C et al. *Data mining: the textbook*. Vol. 1. Springer, 2015.
-  Dua, Dheeru and Casey Graff. *UCI Machine Learning Repository*. 2017.
URL: <http://archive.ics.uci.edu/ml>.
-  Leskovec, Jure, Anand Rajaraman, and Jeffrey David Ullman. *Mining of massive data sets*. Cambridge university press, 2020.
-  Monegato, Giovanni. *Metodi e algoritmi per il calcolo numerico*. Clut, 2008.
-  SYNAPSE. *gene expression cancer RNA-Seq Data Set*. URL: <https://www.synapse.org/#!/Synapse:syn4301332>.



Van Loan, Charles F and G Golub. “Matrix computations (Johns Hopkins studies in mathematical sciences)”. In: *Matrix Computations* (1996).