# CS 357 - 18 Singular Value Decompositions

Boyang Li (boyangl3)

**Overview of SVD:** A more general factorization for any $m \times n$ matrix, there exists a Singular Value Decompsition $\mathbf{A} = \mathbf{U\Sigma V^T}$.

- $\mathbf{U}$ is composed of the eigenvectors of $\mathbf{AA}^T$ as its columns, the vectors are called left singular vectors of $\mathbf{A}$. $\mathbf{U}$ is an orthogonal basis of $\mathbb{R}^m$. $\mathbf{U}$ is an $m \times m$ orthogonal matrix.

$$\mathbf{AA}^T = (\mathbf{U\Sigma V}^T)(\mathbf{U\Sigma V}^T)^T$$
$$(\mathbf{U\Sigma V}^T)(\mathbf{V}^T)^T\mathbf{\Sigma}^T\mathbf{U}^T = \mathbf{U\Sigma}(\mathbf{V}^T\mathbf{V})\mathbf{\Sigma}^T\mathbf{U}^T = \mathbf{U}(\mathbf{\Sigma\Sigma}^T)\mathbf{U}^T$$

- $\mathbf{V}$ is composed of the eigenvectors of $\mathbf{A}^T\mathbf{A}$ as its columns, the vectors arecalled right singular vectors of $\mathbf{A}$. $\mathbf{V}$ is an orthogonal basis of $\mathbb{R}^n$. $\mathbf{V}$ is an $n \times n$ orthogonal matrix.

$$\begin{aligned} \mathbf{A}^T\mathbf{A} &= (\mathbf{U\Sigma V}^T)^T(\mathbf{U\Sigma V}^T) \\ &= \mathbf{V}(\mathbf{\Sigma}^T\mathbf{\Sigma})\mathbf{V}^T \\ &= \mathbf{V\Sigma^2 V}^T \end{aligned}$$

- $\mathbf{\Sigma}$ is a $m \times n$ diagonal matrix composed of square roots of the eigenvalues of $\mathbf{A}^T\mathbf{A}$, called singular values. The diagonal of $\mathbf{\Sigma}$ is ordered by non-increasing singular values and the columns of $\mathbf{U}$ and $\mathbf{V}$ are ordered respectively.

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_s \\ 0 & & 0 \\ \vdots & \ddots & \vdots \\ 0 & & 0 \end{bmatrix} \text{when } m > n, \text{ and } \mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & & & 0 & \dots & 0 \\ & \ddots & & & \ddots & \\ & & \sigma_s & 0 & \dots & 0 \end{bmatrix} \text{when } m < n.$$

Note: $\sigma_1 \geq \sigma_2 \cdots \geq \sigma_s \geq 0$.

**Time complexity:** All of the computation cost of SVD, matrix-matrix production and LU decomposition is $\mathcal{O}(n^3)$. But the cost ranking is SVD > Matrix-Matrix Production > LU.

**Reduced SVD:** SVD of a non-square matrix $\mathbf{A}$ of size $m \times n$ can be represented in a reduced format:

- For $m \geq n$: $\mathbf{U}$ is $m \times n$, $\mathbf{\Sigma}$ is $n \times n$, and $\mathbf{V}$ is $n \times n$.
- For $m \leq n$: $\mathbf{U}$ is $m \times m$, $\mathbf{\Sigma}$ is $m \times m$, and $\mathbf{V}$ is $n \times m$. ($\mathbf{V}^T$ is $m \times n$)

In general we will represent the reduced SVD as: $\mathbf{A} = \mathbf{U}_R\mathbf{\Sigma}_R\mathbf{V}_R^T$, where $\mathbf{U}_R$ is a $m \times k$ matrix, $\mathbf{V}_R$ is a $n \times k$ matrix, $\mathbf{\Sigma}_R$ is a $k \times k$ matrix, $k = \min(m, n)$.

**Important concepts about SVD:** For a $m \times n$ matrix $\mathbf{A} = \mathbf{U\Sigma V}^T$:

- **Rank of a matrix:** Let $s = \min(m, n)$, if there are $r$ non-zero singular singular values, $\text{rank}(\mathbf{A}) = r$, if $r = s$, the matrix is full rank, if $r < s$, the matrix is rank deficient.
- **Nullspace:** $\text{null}(\mathbf{A}) = \text{span}(\mathbf{v_{r+1}}, \mathbf{v_{r+2}}, \ldots, \mathbf{v_n})$; solve space for $\mathbf{A}$ when $\mathbf{Ax} = \vec{0}$.
- **Range:** $\text{range}(\mathbf{A}) = \text{span}(\mathbf{u_1}, \mathbf{u_2}, \ldots, \mathbf{u_r})$
- **Rank-nullity theorem:** $\text{rank}(\mathbf{A}) + \text{null}(\mathbf{A}) = n$.
- **Pseudoinverse:** If the matrix $\mathbf{\Sigma}$ is rank deficient, we cannot get its inverse. We define instead the pseudoinverse: $(\mathbf{\Sigma}^+)_{ii} = \begin{cases} 1/\sigma_i & \sigma_i \neq 0 \\ 0 & \sigma_i = 0 \end{cases}$. So the pseudoinverse of $\mathbf{A}$ is defined as $\mathbf{A}^+ = \mathbf{V\Sigma}^+\mathbf{U}^T$. Here are some true concepts that might be useful for the quiz:
  - ✓ If $\mathbf{A}$ is orthogonal, then $\mathbf{A}^{-1}$ and $\mathbf{A}^+$ exist and qre equal.
  - ✓ If $\mathbf{A}^{-1}$ exists, then $\mathbf{A}^+$ also exists and they are equal.
  - ✓ It is possible for $\mathbf{A}^+$ to exist when $\mathbf{A}^{-1}$ does not exist.
  - ✓ For any matrix, the pseudoinverse always exists.
- **Euclidean Norm of a matrix / inverse:**
  - $\|\mathbf{A}\|_2 = \sigma_1$ - The largest singular value.
  - If the matrix is full-rank, $\|\mathbf{A}^{-1}\|_2 = 1/\sigma_n$ - Inverse of the smallest singular value.
  - If the rank is rank-deficient, $\|\mathbf{A}^+\|_2 = 1/\sigma_r$ - Inverse of the smallest non-zero singular value.
  - For a zero matrix, $\|\mathbf{A}^+\|_2 = 0$
  - For a full-rank matrix, the 2-norm condition number is $\sigma_{\max}/\sigma_{\min}$. If the matrix is rank-deficient, the 2-norm condition number if $\infty$.
- **Low-rank approximation:** $\|\mathbf{A} - \mathbf{A}_k\|_2 = \left\| \sum_{i=k+1}^{n} \sigma_i \mathbf{u_i} \mathbf{v_i^T} \right\|_2 = \sigma_{k+1}$
- **Using SVD to solve system of equation:** Cost to solve $\mathcal{O}(n^2)$
  1. $\mathbf{\Sigma y} = \mathbf{U}^T\mathbf{b}$
  2. $\mathbf{x} = \mathbf{Vy}$

```python
import numpy as np
import numpy.linalg as la
U1, S1, VT1 = np.linalg.svd(A, full_matrices=True)     # Full SVD
U2, S2, VT2 = np.linalg.svd(B, full_matrices=False)    # Reduced SVD
```

Link to course textbook for more detailed information.