CSE 840: Computational Foundations of Artificial Intelligence	Feb 17, 2025

Singular Value Decomposition

Instructor: Vishnu Boddeti — *Scribe: Fangwei Zhang*

1 Singular Value Decomposition

Proposition 1 *Consider $A \in \mathbb{R}^{m\times n}$ of rank r. Then we can write A in the form*

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

where $U \in \mathbb{R}^{m\times m}$, $V \in \mathbb{R}^{n\times n}$ are orthogonal matrices and $\Sigma \in \mathbb{R}^{m\times n}$ is "diagonal" and exactly r of the diagonal values $\sigma_1, \sigma_2, \cdots$ are non-zero.

$$\underbrace{A}_{m\times n} = \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_m \\ | & | & & | \end{bmatrix}}_{m\times m} \underbrace{\begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_r \\ 0 & 0 & \cdots & 0 \end{bmatrix}}_{m\times n} \underbrace{\begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_n^T & - \end{bmatrix}}_{n\times n}$$

Proof: Construct U, V, Σ, such that $A = U\Sigma V^T$.

Given $A \in \mathbb{R}^{m\times n}$, we consider

$$B := A^T A \in \mathbb{R}^{n\times n}$$

Observe: - B is symmetric:

$$(A^T A)^T = A^T (A^T)^T = A^T A$$

- B is positive semi-definite:

$$x^T B x = \langle x, Bx\rangle = \langle x, A^T A x\rangle$$

$$= \langle Ax, Ax\rangle = \|Ax\|^2 \geq 0$$

So there exists an orthonormal basis of eigenvectors $x_1, x_2, \ldots, x_n$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n \geq 0$.

Define:

- $\Sigma =$ "diag(σ_i)" $\in \mathbb{R}^{m\times n}$
  where $\sigma_i = \sqrt{\lambda_i}$.

- U is defined as:

$$U = \begin{pmatrix} 1 \\ r_i \end{pmatrix} \quad \text{matrix with columns}$$

  where

$$r_i := \frac{Ax_i}{\sigma_i}$$

- V is defined as:

$$V = \begin{pmatrix} 1 \\ x_i \end{pmatrix} \quad \text{matrix with } x_i \text{ as columns.}$$

Now we need to show that with these definitions, we have:

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

<u>**Sketch:**</u>

- Columns of $U \cdot \Sigma$ are given as:

$$\sigma_i r_i = \sigma_i \cdot \frac{Ax_i}{\sigma_i} = Ax_i$$

- Now multiply with V^T:
  - Rows of V^T are the x_i.
  - Exploit that:
    * If $i \neq j$, then $x_i \perp x_j$.
    * $\|x_i\| = 1$.
  - The terms consisting of i, j with $i \neq j$ cancel, while the terms with $i = j$ will be 1.

Thus, we will be left with the matrix A.

$\square$

Example:

To perform Singular Value Decomposition (SVD) for the matrix

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix},$$

let's break it down step by step.

<u>**Step 1: Compute AA^T**</u>

First, we need to calculate the matrix AA^T (where A^T is the transpose of matrix A):

$$A^T = \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix}$$

Now, compute AA^T:

$$AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 17 & 8 \\ 8 & 17 \end{bmatrix}$$

Step 2: Find the Eigenvalues of AA^T

To find the eigenvalues of AA^T, we solve the characteristic equation:

$$\det(AA^T - \lambda I) = 0$$

$$\det \begin{bmatrix} 17-\lambda & 8 \\ 8 & 17-\lambda \end{bmatrix} = 0$$

$$(\lambda - 25)(\lambda - 9) = 0$$

Thus, the eigenvalues are $\lambda_1 = 25$ and $\lambda_2 = 9$. These eigenvalues correspond to the singular values $\sigma_1 = 5$ and $\sigma_2 = 3$, since the singular values are the square roots of the eigenvalues.

Step 3: Find the Right Singular Vectors (Eigenvectors of A^TA)

Next, we find the eigenvectors of A^TA for $\lambda = 25$ and $\lambda = 9$.

For $\lambda = 25$:

Solve $(A^TA - 25I)v = 0$:

$$A^TA - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{bmatrix}$$

Row-reducing this matrix:

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The eigenvector corresponding to $\lambda = 25$ is:

$$v_1 = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}$$

For $\lambda = 9$, solving $(A^TA - 9I)v = 0$:

$$v_2 = \begin{bmatrix} \frac{1}{\sqrt{18}} \\ -\frac{1}{\sqrt{18}} \\ \frac{4}{\sqrt{18}} \end{bmatrix}$$

For the third eigenvector v_3, since v_3 must be perpendicular to v_1 and v_2:

$$v_3 = \begin{bmatrix} \frac{2}{3} \\ -\frac{2}{3} \\ \frac{1}{3} \end{bmatrix}$$

Step 4: Compute the Left Singular Vectors (Matrix U)

To compute the left singular vectors U, we use the formula $u_i = \frac{1}{\sigma_i} A v_i$. This results in:

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

Step 5: Final SVD Equation

Finally, the Singular Value Decomposition of matrix A is:

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

Where:

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix}$$

$$V = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

Thus, the SVD of matrix A is:

$$A = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ \frac{1}{\sqrt{18}} & -\frac{1}{\sqrt{18}} & \frac{4}{\sqrt{18}} \\ \frac{2}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

This is the result of the Singular Value Decomposition of matrix A.

2 Key Differences between SVD & Eigenvalue Decomposition

- **SVD always exists**, no matter how A looks like.
- U, V are **orthogonal** (not true for eigenvectors in general).
- Singular values are always **real** and **non-negative**.
- If $A \in \mathbb{R}^{n \times n}$ is **symmetric**, then the SVD is "nearly the same" as the eigenvalue decomposition.
- If (λ_i, v_i) are the eigenvalue/eigenvector pairs of A, then $(|\lambda_i|, v_i)$ are the singular value/singular vector pairs of A.
- In particular, **left- and right-singular vectors** are the same.
- Left-singular vectors of A are the eigenvectors of AA^T.
- Right-singular vectors of A are the eigenvectors of $A^T A$.
- If $\lambda_i \neq 0$ is an eigenvalue of $A^T A$ (or equivalently, AA^T), then:

  $$\sqrt{\lambda_i} \neq 0$$

  is a singular value of A.

3 Matrix Norms

Given a matrix $A \in \mathbb{R}^{m \times n}$, we define the following norms:

- **Maximum norm (Infinity norm):**

  $$\|A\|_{\max} = \|A\|_\infty = \max_{i,j} |a_{ij}|$$

- **One norm (Absolute sum norm):**

  $$\|A\|_1 = \sum_{i,j} |a_{ij}|$$

- **Frobenius norm:**

  $$\|A\|_F = \sqrt{\sum_{i,j} a_{ij}^2} = \sqrt{\operatorname{tr}(A^T A)}$$

  $$= \sqrt{\sum \sigma_i^2}, \quad \text{where } \sigma_i \text{ are the singular values of } A.$$

- **Spectral norm (Operator norm):**

  $$\|A\|_2 = \sigma_{\max}(A), \quad \text{where } \sigma_{\max} \text{ is the largest singular value.}$$

  $$= \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}$$

  where the denominator uses the **Euclidean norm** on vectors in $\mathbb{R}^m$. The spectral norm is also known as the **operator norm** or **spectral norm**.

4 Rank-k Approximation of Matrices

Given a matrix $A = U\Sigma V^T$, where the singular values $\sigma_1, \sigma_2, \ldots$ are sorted in descending order. Now we define a new matrix A_k as follows:

$$A_k = U_k \Sigma_k V_k^T$$

where: - We take the first k columns of U. - The first k entries of Σ. - The first k rows of V^T.

More formally:

$$A_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$$

where each term $\sigma_i u_i v_i^T$ is a **rank-1** matrix.

Proposition 2 *Let B be any rank-k matrix $B \in \mathbb{R}^{m \times n}$. Then:*

$$\|A - A_k\|_F \leq \|A - B\|_F$$

"A_k is the best rank-k approximation (in Frobenius norm)."

Proposition 3 *For any matrix B of rank-k,*

$$B \in \mathbb{R}^{m \times n}, \quad \|A - A_k\|_2 \leq \|A - B\|_2.$$

*where $\|\cdot\|_2$ denotes the **operator norm**.*

"A_k is the best rank-k approximation (in operator norm)."

Figure 1: Rank-k Approximation example

5 Pseudo-Inverse of Matrix

Define: For $A \in \mathbb{R}^{m \times n}$, a **pseudo-inverse** of A is defined as the matrix $A^\dagger \in \mathbb{R}^{n \times m}$ which satisfies the following properties:

- $AA^\dagger A = A$
- $A^\dagger AA^\dagger = A^\dagger$
- $(AA^\dagger)^T = AA^\dagger$
- $(A^\dagger A)^T = A^\dagger A$

Intuition:

- A is a projection from $\mathbb{R}^3 \to \mathbb{R}^2$:

$$A \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- It cannot be inverted, obviously. (Inverting means reconstructing the original)
- But we can "make up" a reconstruction:

$$R : \mathbb{R}^2 \to \mathbb{R}^3$$

$$R \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \\ 5 \end{pmatrix}$$

- Now we have:

$$ARA = A$$

which implies:

$$AA^\dagger A = A$$

Proposition 4 *Let $A \in \mathbb{R}^{m \times n}$, and let $A = U\Sigma V^T$ be its SVD. Then:*

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

where $\Sigma^\dagger \in \mathbb{R}^{n \times m}$ and is defined as:

$$\Sigma^\dagger_{ii} = \begin{cases} \frac{1}{\Sigma_{ii}}, & \text{if } \Sigma_{ii} \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

$$\Sigma = \begin{pmatrix} \sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \sigma_r \\ 0 & \dots & 0 \end{pmatrix}, \quad \Sigma^\dagger = \begin{pmatrix} 1/\sigma_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/\sigma_r \\ 0 & \dots & 0 \end{pmatrix}$$

Intuition:

Assume $A \in \mathbb{R}^{n \times n}$ is **invertible**, and assume it has an **eigendecomposition**:

$$A = U \Lambda U^T$$

- All entries of $\mathrm{diag}(\Lambda)$ are nonzero (eigenvalues are nonzero).
- The inverse of A is given by:

  $$A^{-1} = U \Lambda^{-1} U^T$$

  where:

  $$\Lambda^{-1} = \begin{pmatrix} 1/\lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & 1/\lambda_n \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix}$$