

Preclass 07: Singular Value Decomposition

[SCS4049] Machine Learning and Data Science

Seongsik Park (s.park@dgu.edu)

School of AI Convergence & Department of Artificial Intelligence, Dongguk University

Singular value decomposition

Singular value decomposition (SVD) of a given matrix A

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T \quad (1)$$

$$= \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k \end{bmatrix} \begin{bmatrix} \text{---} & \mathbf{v}_1^T & \text{---} \\ \text{---} & \mathbf{v}_2^T & \text{---} \\ & \vdots & \\ \text{---} & \mathbf{v}_r^T & \text{---} \end{bmatrix} \quad (2)$$

where

- $\mathbf{A} \in \mathbb{R}^{m \times n}$, $\text{rank}(\mathbf{A}) = r$
- $\mathbf{U} \in \mathbb{R}^{m \times r}$, $\mathbf{U}^T \mathbf{U} = \mathbf{I}$
- $\mathbf{V} \in \mathbb{R}^{n \times r}$, $\mathbf{V}^T \mathbf{V} = \mathbf{I}$
- $\mathbf{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_r)$ where $\sigma_1 \geq \dots \geq \sigma_r > 0$

Singular value decomposition

$$\text{with } \mathbf{U} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \\ | & | & & | \end{bmatrix} \text{ and } \mathbf{V} = \begin{bmatrix} | & | & \cdots & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \cdots & \mathbf{v}_r \\ | & | & & | \end{bmatrix},$$

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^T \quad (3)$$

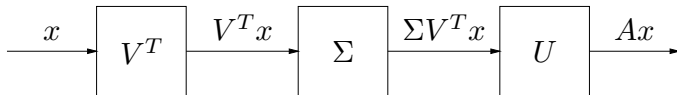
where

- σ_i are the nonzero *singular values* of $\mathbf{A}$
- $\mathbf{v}_i$ are the *right* or *input singular vectors* of $\mathbf{A}$
- $\mathbf{u}_i$ are the *left* or *output singular vectors* of $\mathbf{A}$

Interpretations

SVD:

$$\mathbf{A} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_{k=1}^r \sigma_k \mathbf{u}_k \mathbf{v}_k^T \quad (4)$$



Linear mapping $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be decomposed as

- compute coefficients of $\mathbf{x}$ along input directions $\mathbf{v}_1, \dots, \mathbf{v}_r$
- scale coefficients by σ_i
- reconstitute along output directions $\mathbf{u}_1, \dots, \mathbf{u}_r$

difference with eigenvalue decomposition for symmetric $\mathbf{A}$: input and output directions are *different*

Geometric interpretation

