

Preclass 07: Singular Value Decomposition

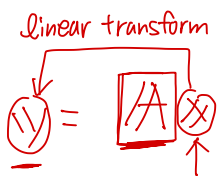
[SCS4049] Machine Learning and Data Science

Seongsik Park (s.park@dgu.edu)

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

행렬 = linear transformation (선형 사상)

A = 벡터 $\longrightarrow$ 벡터 변환, linear
입력 $\times$ 출력



$$\underline{A} = \underline{U} \cdot \underline{\Sigma} \cdot \underline{V}^T$$

$\uparrow$ $\uparrow$ $\uparrow$

$$= \boxed{U \Sigma V^T} \times$$

$U, V = \underline{\text{orthogonal matrix}}$ 직교행렬

$\frac{\text{두 벡터}}{u, v} < \begin{cases} \underline{\text{orthogonal}} : \text{두 벡터가 직교, } \Leftrightarrow \underline{u \cdot v = 0} \Leftrightarrow \text{직각} \\ \underline{\text{orthonormal}} : \text{두 벡터가 orthogonal \& 둘 다 unit} \\ \Leftrightarrow u \cdot v = 0, \|u\|_2 = \|v\|_2 = 1 \end{cases}$

직교행렬 - orthogonal : 모든 column vector들이
서로 orthonormal 할 때.

= 모든 row vector들이
서로 orthonormal.

$$|u\rangle = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad |v\rangle = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$|u\rangle, |v\rangle$ orthogonal? $\bigcirc$
orthonormal? $\bigcirc$

$$U = \begin{bmatrix} \text{unit } 1/\sqrt{2} & \text{unit } 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}$$

U orthogonal? $\bigcirc$
 matrix

$$U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

orthogonal $\bigcirc$
 matrix

U : orthogonal

$$\Leftrightarrow \underline{\underline{U^T U = U U^T = I}}$$

diagonal matrix 대각행렬

= 대각선에만 값이 존재, 나머지는 0

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

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

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

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

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

diagonal matrix $N \times N$ N^2 elements
square matrix

$$A = \text{diag}(d_1, d_2, \dots, d_n)$$

$$\text{diag}(10, 1) = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\underline{\text{diag}}(\underline{9}, \underline{4}, \underline{-3}) = \begin{bmatrix} 9 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

$$A = \text{diag}(d_1, d_2, \dots, d_n)$$
$$A^{-1} = \text{diag}(d_1^{-1}, d_2^{-1}, \dots, d_n^{-1})$$

Singular value decomposition

Singular value decomposition (SVD) of a given matrix A

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

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

$\sigma_k = \text{Singular value}$
 특이값 (큰 값부터)

where

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

$$y = Ax = \left(\underline{\underline{U \left(\Sigma (V^T x) \right)}} \right)$$

$$U = \begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_N \\ | & | & & | \end{bmatrix}$$

$$V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_N \\ | & | & & | \end{bmatrix}^{\leftarrow}$$

$$\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_N)$$

$$= \underbrace{\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \dots & u_N \\ | & | & & | \end{bmatrix}} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_N \end{bmatrix}} \underbrace{\begin{bmatrix} -v_1^T- \\ -v_2^T- \\ \vdots \\ -v_N^T- \end{bmatrix}} x$$

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & \sigma_N \end{bmatrix} \begin{bmatrix} -\nu_1^T \\ -\nu_2^T \\ \vdots \\ -\nu_N^T \end{bmatrix} \begin{bmatrix} x \\ \vdots \\ x \end{bmatrix}$$

$N \times N$

$$\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_N \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_N \end{bmatrix} \begin{bmatrix} \nu_1^T x \\ \nu_2^T x \\ \vdots \\ \nu_N^T x \end{bmatrix}$$

$N \times N$ $N \times 1$

$$\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_N \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 \nu_1^T x \\ \sigma_2 \nu_2^T x \\ \vdots \\ \sigma_N \nu_N^T x \end{bmatrix}$$

$$y = \underbrace{u_1 (\sigma_1 v_1^T x)} + u_2 (\sigma_2 v_2^T x) + \dots + u_N (\sigma_N v_N^T x)$$

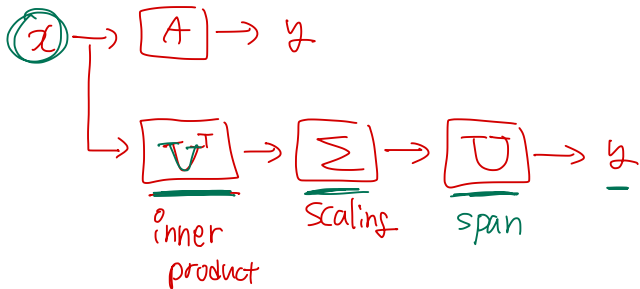
$$= \sum_{n=1}^N (\sigma_n \underline{v_n^T} x) u_n$$

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

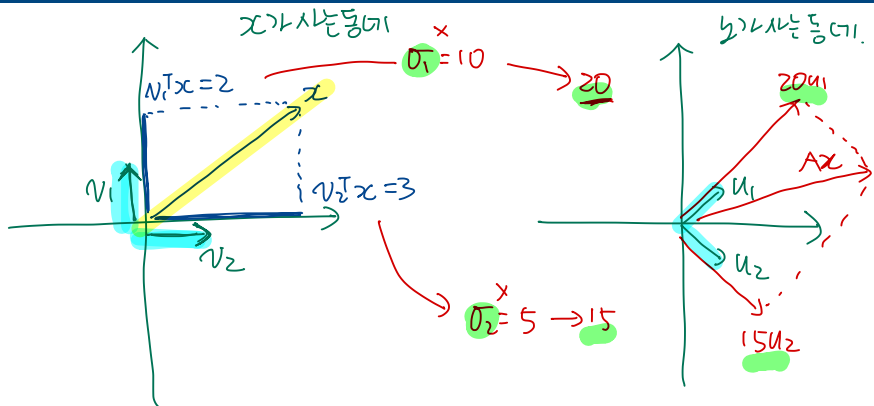
$$A(x) = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \begin{bmatrix} \overset{\sigma_1}{10} & 0 \\ 0 & \underset{\sigma_2}{5} \end{bmatrix} \begin{bmatrix} \downarrow 0 & \downarrow 1 \\ 1 & 0 \end{bmatrix}^T \begin{bmatrix} 3 \\ 2 \end{bmatrix}$$

u_1 u_2 v_1 v_2

$$= \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \cdot 10 \cdot 2 + \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix} \cdot 5 \cdot 3$$



$$\begin{aligned}
 x \begin{cases} \nearrow v_1^T x \rightarrow \sigma_1 v_1^T x \rightarrow u_1 \sigma_1 v_1^T x \\ \rightarrow v_2^T x \rightarrow \sigma_2 v_2^T x \rightarrow u_2 \sigma_2 v_2^T x \\ \searrow \dots \end{cases} \left. \begin{array}{c} + \\ + \\ + \end{array} \right\} \Rightarrow y
 \end{aligned}$$



$$V = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

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

$\vec{b} \in \mathbb{R}^N$

A
 $N \times N$

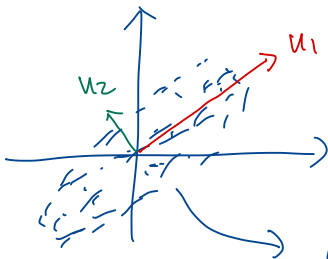
x
 $N \times 1$

①

$$\begin{bmatrix} | & | & \dots & | \\ a_1 & a_2 & \dots & a_N \\ | & | & \dots & | \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix} = \begin{bmatrix} | \\ a_1 x_1 + a_2 x_2 + \dots + a_N x_N \\ | \end{bmatrix}$$

②

$$\begin{bmatrix} - \tilde{a}_1^T - \\ - \tilde{a}_2^T - \\ \vdots \\ - \tilde{a}_N^T - \end{bmatrix} x = \begin{bmatrix} \tilde{a}_1^T x \\ \tilde{a}_2^T x \\ \vdots \\ \tilde{a}_N^T x \end{bmatrix}$$



$$\underline{\text{Cov}} = \underline{U} \Sigma \underline{V}^T$$

Singular value decomposition

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

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

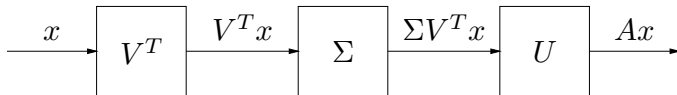
where

- 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:

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



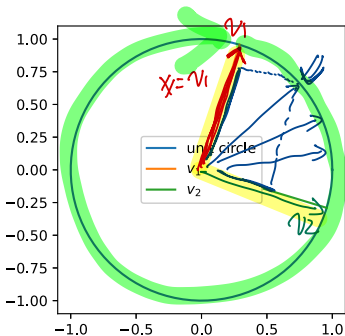
Linear mapping $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be decomposed as

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

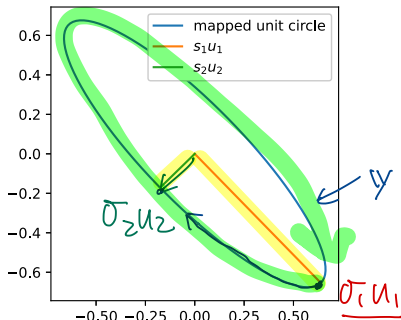
difference with eigenvalue decomposition for symmetric $\mathbf{A}$: input and output directions are *different*

Geometric interpretation

x 가 v_1 을 뜻함



$y = Ax$ 가 v_2 를 뜻함.



$$A \otimes v_1 = \sigma_1 u_1$$

$$y = Ax$$

$$Ax = \sigma_2 u_2$$

$$\|x\|_2 = 1$$