Inclass 15: Singular Value Decomposition

[AIX7021] Computer Vision

Seongsik Park (s.park@dgu.edu)

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

행렬 = linear transformation (선형 사상)

A = 벡터 ⟶ 벡터 변환. linear

입력 x 출력 y

linear transform

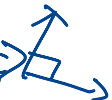
$y = A x$

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

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

U, V = orthogonal matrix 직교행렬

두 벡터 u, v

- orthogonal : 두 벡터가 직교, $\Leftrightarrow$ $u \cdot v = 0$ $\Leftrightarrow$ 
- orthonormal : 두 벡터가 orthogonal & 둘다 unit
  $\Leftrightarrow$ $u \cdot v = 0$, $\|u\|_2 = \|v\|_2 = 1$

한 행렬 - orthogonal : 모든 column vector들이
서로 orthonormal 할 때.

= 모든 row vector들이
서로 orthonormal.

$$\mathbf{u} = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \qquad \mathbf{v} = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}$$

$\mathbf{u}, \mathbf{v}$ orthogonal? O

orthonormal? O

unit 직교 unit

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

U orthogonal matrix? O

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

orthogonal matrix O

U : orthogonal $\iff$ $U^T U = U U^T = I$

diagonal matrix 대각행렬

= 대각선에만 값이 존재, 나머지위치는 0

$$\begin{bmatrix} 10 & 0 \\ 0 & 9 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \quad \cancel{\begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}}$$

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

나머지 0으로 채움

diagonal matrix $N \times N$ $N^2 > n$

square matrix

$$A = \mathrm{diag}(d_1, d_2, \cdots, d_N)$$

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

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

$$A = \mathrm{diag}\,(d_1, d_2, \cdots, d_N)$$

$$A^{-1} = \mathrm{diag}(d_1^{-1}, d_2^{-1}, \cdots, d_N^{-1})$$

Singular value decomposition

Singular value decomposition (SVD) of a given matrix A

$$\mathbf{A} = \underline{\mathbf{U}\boldsymbol{\Sigma}}\mathbf{V}^T \tag{1}$$

$\underline{U}\ \underline{\Sigma}\ \underline{V}^T$

$$= \begin{bmatrix} | & | & & | \\ \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_r \\ | & | & & | \end{bmatrix} \begin{bmatrix} \underline{\sigma_1} & 0 & \cdots & 0 \\ 0 & \underline{\sigma_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \underline{\sigma_r} \end{bmatrix} \begin{bmatrix} — & \mathbf{v}_1^T & — \\ — & \mathbf{v}_2^T & — \\ & \vdots & \\ — & \mathbf{v}_r^T & — \end{bmatrix} \tag{2}$$

σ_k = singular value

내림차순 (큰값부터)

where

- $\mathbf{A} \in \Re^{m\times n}$, $\operatorname{rank}(\mathbf{A}) = r$
- $\mathbf{U} \in \Re^{m\times r}$, $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ ←
- $\mathbf{V} \in \Re^{n\times r}$, $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ ←
- $\boldsymbol{\Sigma} = \operatorname{diag}(\sigma_1, ..., \sigma_r)$ where $\sigma_1 \geq \cdots \geq \sigma_r > 0$

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

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

$$V = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \cdots & v_N \\ | & | & & | \end{bmatrix}$$

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_N)$$

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

$$\begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & & 0 & \sigma_N \end{bmatrix} \underbrace{\begin{bmatrix} - & v_1^T & - \\ - & v_2^T & - \\ & \vdots & \\ - & v_N^T & - \end{bmatrix}}_{N \times N} \underbrace{x}_{N \times 1}$$

$$\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_N \\ | & | & & | \end{bmatrix} \underbrace{\begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_N \end{bmatrix}}_{N \times N} \underbrace{\begin{bmatrix} v_1^T x \\ v_2^T x \\ \vdots \\ v_N^T x \end{bmatrix}}_{N \times 1}$$

$$\begin{bmatrix} | & | & & | \\ u_1 & u_2 & \cdots & u_N \\ | & | & & | \end{bmatrix} \begin{bmatrix} \sigma_1 v_1^T x \\ \sigma_2 v_2^T x \\ \vdots \\ \sigma_N v_N^T x \end{bmatrix}$$

$$y = \begin{matrix} | \\ u_1 \\ | \end{matrix} \underline{(\sigma_1 v_1^T x)} + \begin{matrix} | \\ u_2 \\ | \end{matrix} (\sigma_2 v_2^T x) + \cdots \begin{matrix} | \\ u_N \\ | \end{matrix} (\sigma_N v_N^T x)$$

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

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

$$A x = \underset{u_1 \quad\quad u_2}{\begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix}} \begin{bmatrix} 10 & 0 \\ 0 & 5 \end{bmatrix} \underbrace{\underset{v_1 \quad v_2}{\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}^T \begin{bmatrix} 3 \\ 2 \end{bmatrix}}$$

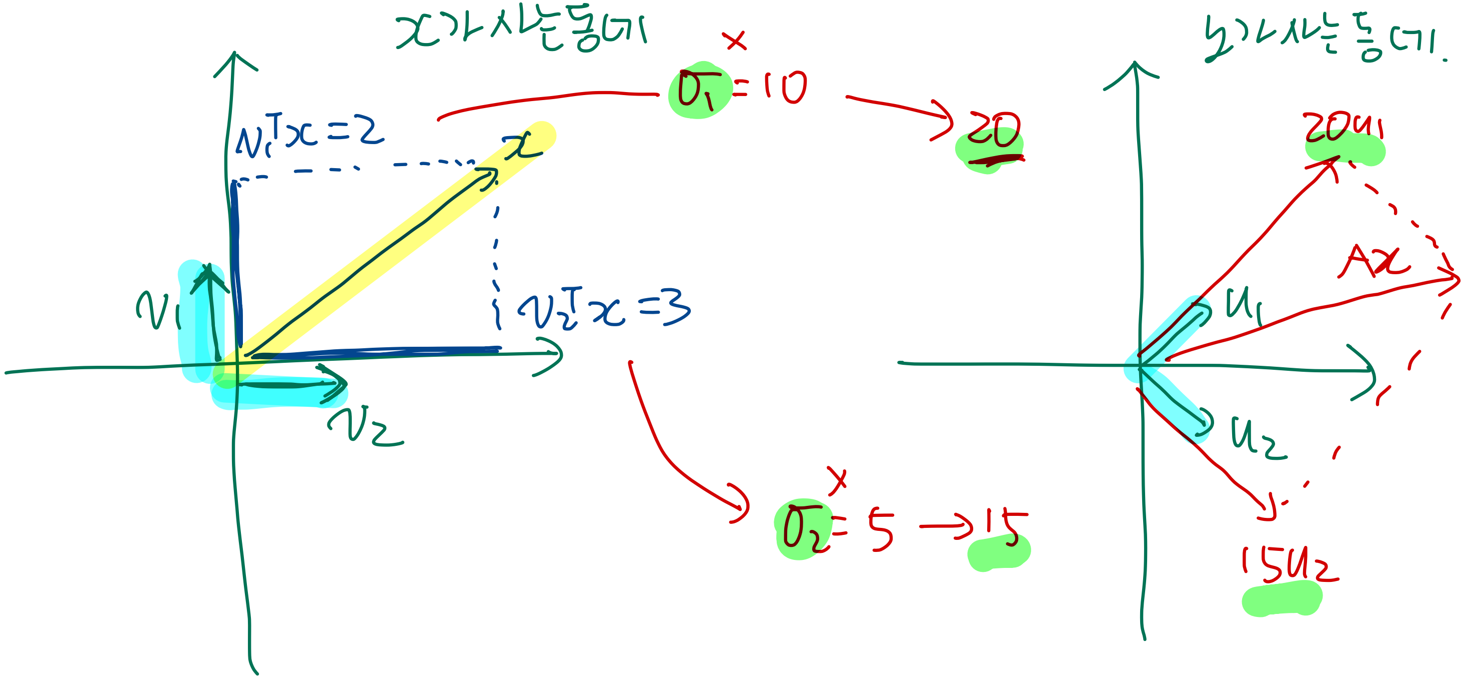
σ_1 σ_2

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

$$x \to \boxed{A} \to y$$

$$x \to \underbrace{\boxed{V^T}}_{\text{inner product}} \to \underbrace{\boxed{\Sigma}}_{\text{scaling}} \to \underbrace{\boxed{U}}_{\text{span}} \to \underline{y}$$

$$x \begin{array}{l} \nearrow \; v_1^T x \to \sigma_1 v_1^T x \to \\ \to \; v_2^T x \to \sigma_2 v_2^T x \to \\ \searrow \; \cdots \end{array} \left. \begin{array}{c} u_1 \sigma_1 v_1^T x \\ + \\ u_2 \sigma_2 v_2^T x \\ + \\ + \end{array} \right\} \Rightarrow y$$



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

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

행렬 벡터

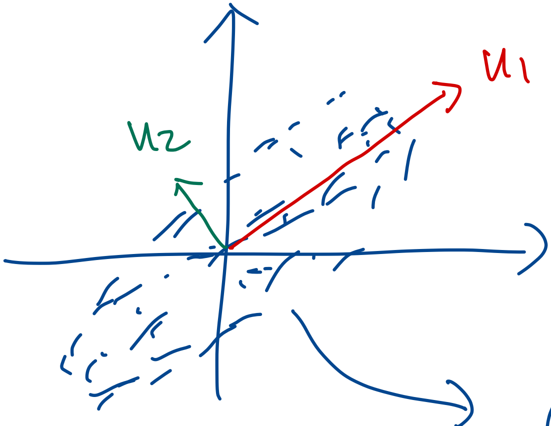
$\underset{N\times N}{A}$ $\underset{N\times 1}{x}$

①

$$\underbrace{\begin{bmatrix} | & | & & | \\ a_1 & a_2 & \cdots & a_N \\ | & | & & | \end{bmatrix}}\underline{\begin{bmatrix} x_1 \\ \vdots \\ x_N \end{bmatrix}} = \begin{matrix} | \\ a_1 \\ | \end{matrix} x_1 + \begin{matrix} | \\ a_2 \\ | \end{matrix} x_2 + \cdots \begin{matrix} | \\ a_N \\ | \end{matrix} x_N$$

②

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



$$\mathrm{Cov} = U \Sigma V^T$$

Singular value decomposition

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

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

where

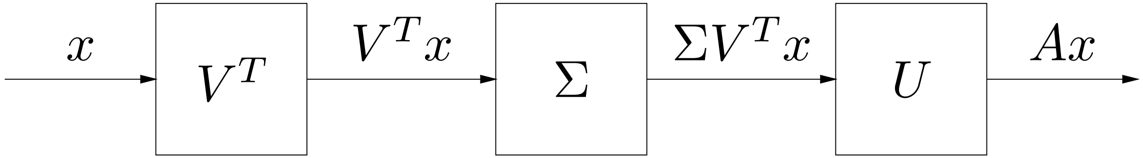
- σ_i are the nonzero *singular values* of $\mathbf{A}$
- $\mathbf{v}_i$ are the *right* or *input singular vectors* of $\mathbf{A}$
- $\mathbf{u}_i$ are the *left* or *output singular vectors* of $\mathbf{A}$

Interpretations

SVD:

V가 먼저

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



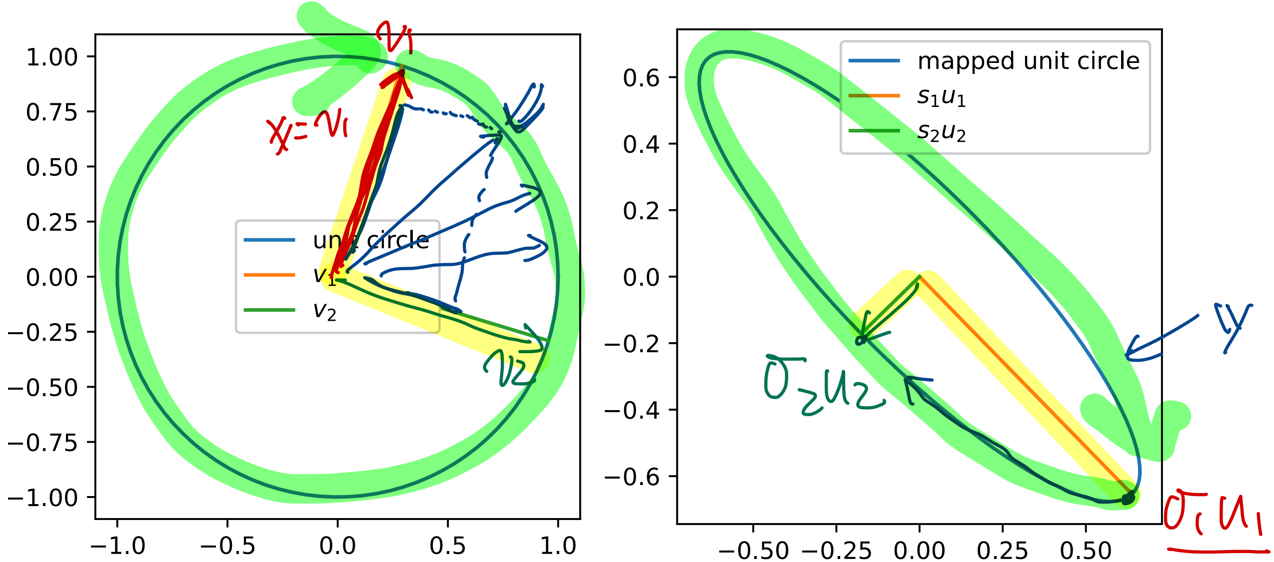
Linear mapping $\mathbf{y} = \mathbf{A}\mathbf{x}$ can be decomposed as

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

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

Geometric interpretation

x가 사는 동네 $y = Ax$ 가 사는 동네.



$$A x = A v_1 = \sigma_1 u_1 \qquad y = A \underline{x}$$

$$A \underline{x} = A \underline{v_2} = \sigma_2 u_2 \qquad \|x\|_2 = 1$$