

# **Singular Value Decomposition**

# 수업 목표

## 이번 수업의 핵심:

- Singular value decomposition의 정의와 그 성질
- Singular value decomposition의 계산 방법
- Singular value decomposition을 활용한 Low-rank approximation

## 핵심 개념

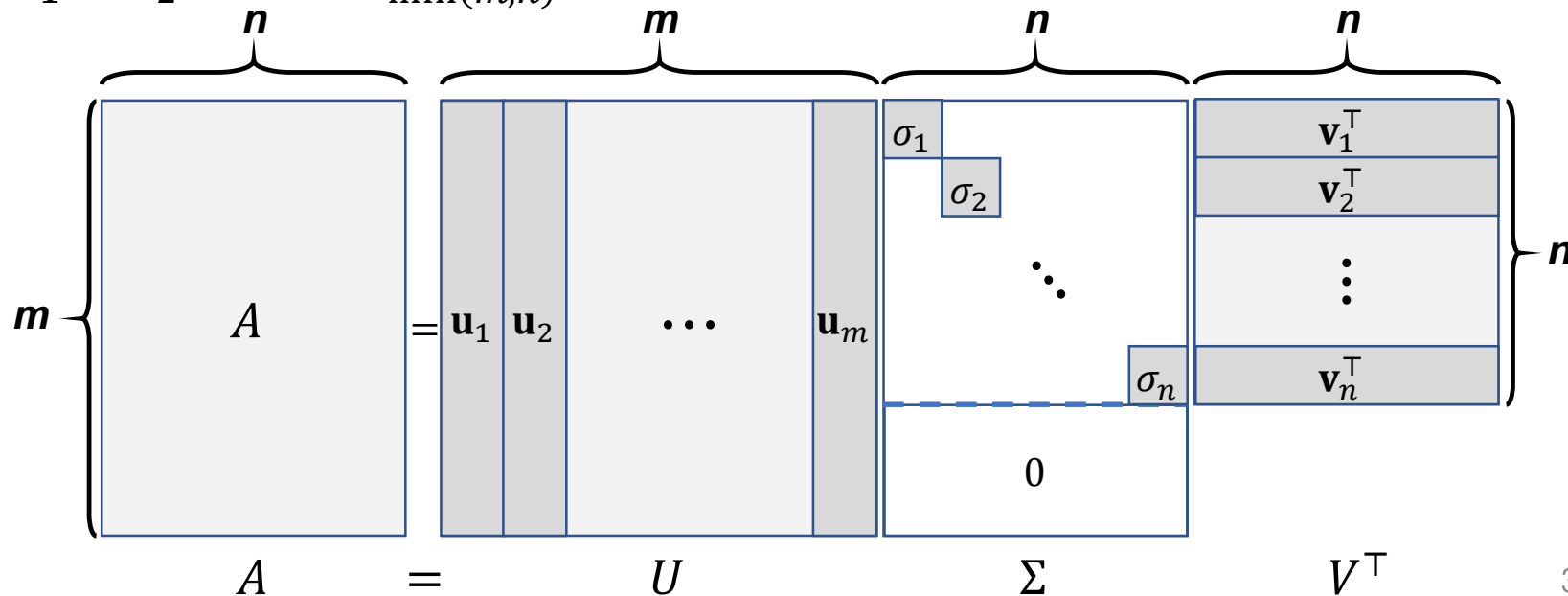
- Singular value decomposition
- Low-rank approximation

# Singular Value Decomposition (SVD, 특이값 분해)

직사각행렬  $A \in \mathbb{R}^{m \times n}$ 의 Singular value decomposition:

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

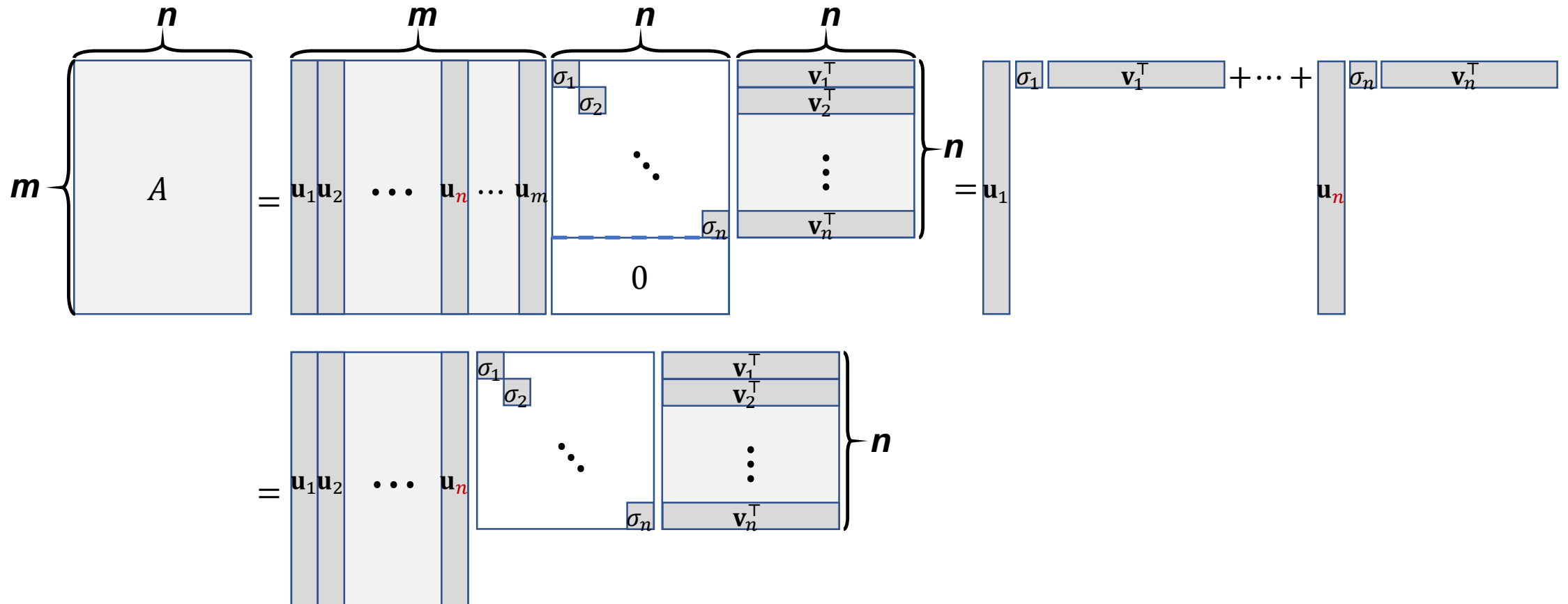
- $U \in \mathbb{R}^{m \times m}$ ,  $V \in \mathbb{R}^{n \times n}$ : Orthonormal한 열을 가지는 행렬
  - 각각  $A$ 의 Col  $A$ 와 Row  $A$ 의 Orthonormal basis
- $\Sigma \in \mathbb{R}^{m \times n}$ : 대각 원소들만 nonzero 값을 가지며, 내림차순으로 정렬된 행렬
  - i.e.,  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min(m,n)}$



# Reduced Form of SVD

- $A$ 의 SVD는  $\mathbf{u}_i$ 와  $\mathbf{v}_i$  Outer product의 합으로 볼 수 있음

$$A = U\Sigma V^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{where } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$$



## SVD의 또다른 관점

$A$ 의 Col  $A$ 와 Row  $A$ 의 Orthonormal Basis  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}, \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 는 Gram-Schmidt orthogonalization 를 통해 쉽게 찾을 수 있음

← 그러나, 이 경우 **Orthonormal basis는 유일하지 않고**,  $\mathbf{u}_i$ 와  $\mathbf{v}_i$  간의 어떤 연관관계가 존재하지 않음

반면, SVD에서는 다음 관계를 만족하는  $\{\mathbf{u}_1, \dots, \mathbf{u}_n\}, \{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ 를 Joint하게 찾아야 함:

$$A\mathbf{v}_i = \sigma_i \mathbf{u}_i, \quad \forall i = 1, \dots, n$$

## SVD의 또다른 관점

$A = U\Sigma V^T$ 에서 다음과 같이 가정:

$$U = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \in \mathbb{R}^{m \times n}, V = [\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] \in \mathbb{R}^{n \times n}, \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \end{bmatrix} \in \mathbb{R}^{n \times n}$$

- $AV = A[\mathbf{v}_1 \quad \mathbf{v}_2 \quad \cdots \quad \mathbf{v}_n] = [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n]$
- $U\Sigma = [\mathbf{u}_1 \quad \mathbf{u}_2 \quad \cdots \quad \mathbf{u}_n] \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \sigma_n \end{bmatrix} = [\sigma_1 \mathbf{u}_1 \quad \sigma_2 \mathbf{u}_2 \quad \cdots \quad \sigma_n \mathbf{u}_n]$
- $AV = U\Sigma \Leftrightarrow [A\mathbf{v}_1 \quad A\mathbf{v}_2 \quad \cdots \quad A\mathbf{v}_n] = [\sigma_1 \mathbf{u}_1 \quad \sigma_2 \mathbf{u}_2 \quad \cdots \quad \sigma_n \mathbf{u}_n]$
- $V \in \mathbb{R}^{n \times n}$ 가 Orthonormal한 열을 가지기 때문에  $V^{-1} = V^T$

따라서  $AV = U\Sigma \Leftrightarrow A = U\Sigma V^T$

# SVD 계산하기

- 주어진 행렬  $A$ 에 대해 먼저  $AA^T$ 와  $A^T A$ 를 계산

$$AA^T = U\Sigma V^T V \Sigma^T U^T = U\Sigma \Sigma^T U^T = U\Sigma^2 U^T$$

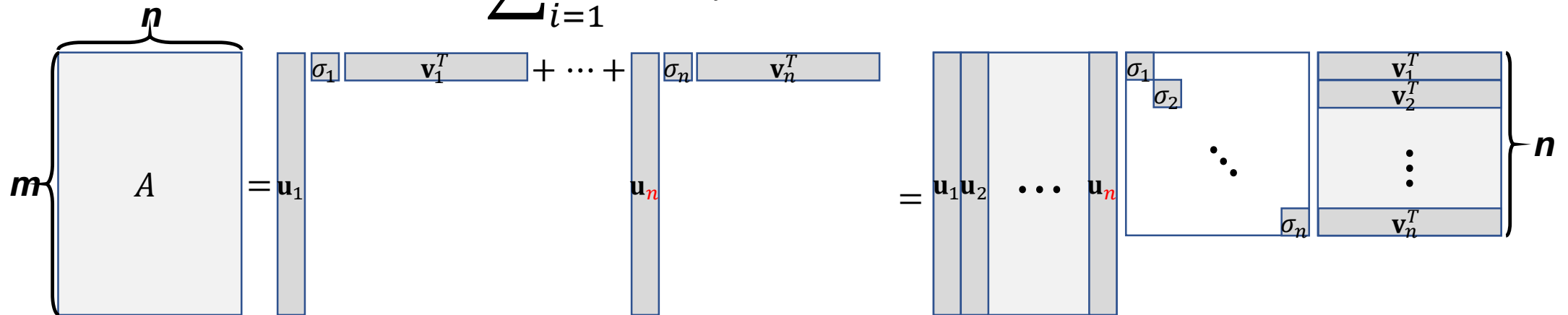
$$A^T A = V\Sigma^T U^T U \Sigma V^T = V\Sigma^T \Sigma V^T = V\Sigma^2 V^T$$

- 이 때,  $AA^T$ 와  $A^T A$ 는 Symmetric positive semi-definite
  - $(AA^T)^T = AA^T$ ,  $(A^T A)^T = A^T A$
  - $\mathbf{x}^T AA^T \mathbf{x} = (A^T \mathbf{x})^T (A^T \mathbf{x}) = \|A^T \mathbf{x}\|^2 \geq 0$ ,  $\mathbf{x}^T A^T A \mathbf{x} = (A \mathbf{x})^T (A \mathbf{x}) = \|A \mathbf{x}\|^2 \geq 0$
- 따라서  $AA^T$ 와  $A^T A$ 는 항상 orthogonally diagonalizable 하기 때문에, Spectral decomposition이 가능
  - Symmetric  $\rightarrow$  Eigenvector  $U$ 와  $V$ 가 Orthogonal:  $U^{-1} = U^T, V^{-1} = V^T$
  - Positive semi-definite  $\rightarrow \Sigma^2$ 가 음수가 아님, 즉 항상 실수  $\Sigma$ 를 구할 수 있음

# SVD 활용: Low-Rank Approximation

- 직사각행렬  $A \in \mathbb{R}^{m \times n}$ 의 SVD는 Outer product의 합으로 표현될 수 있음

$$A = U\Sigma V^T = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{where } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_n$$



- $A$ 를 근사하는 최적의 Low-rank 행렬을 찾는 문제

$$\hat{A}_r = \arg \min_{A_r} \|A - A_r\|_F \quad \text{subject to } \text{rank}(A_r) \leq r$$

- 위 문제, Low-rank approximation의 최적해:

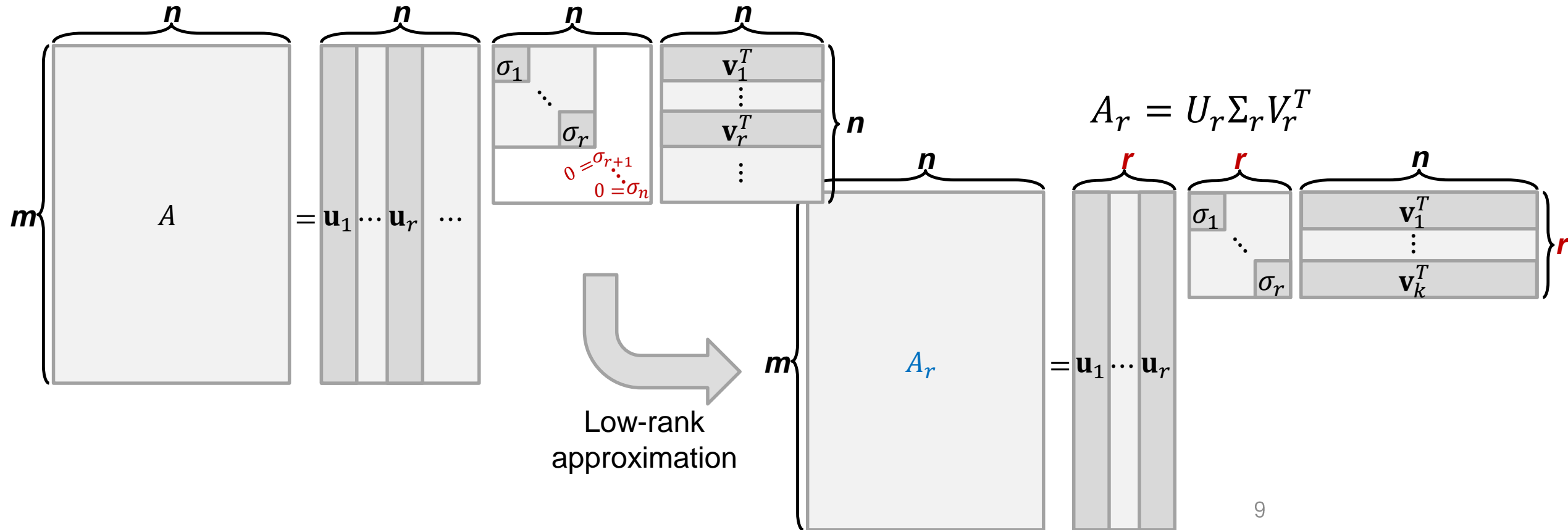
$$\hat{A}_r = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T, \quad \text{where } \sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r$$



# SVD 활용: Low-Rank Approximation

- $\forall i \geq (r + 1)$ 에 대해서  $\sigma_i = 0$ 로 설정한  $A_r$ 로  $A$ 를 근사

$$A = \sum_{i=1}^n \sigma_i \mathbf{u}_i \mathbf{v}_i^T \simeq A_r = \sum_{i=1}^r \sigma_i \mathbf{u}_i \mathbf{v}_i^T = U_r \Sigma_r V_r^T$$



# SVD와 Eigendecomposition 관계 및 활용

- 직사각행렬  $A \in \mathbb{R}^{m \times n}$ 는 SVD가 항상 가능
- 정사각행렬  $A \in \mathbb{R}^{n \times n}$ 가 Eigendecomposition이 안 되도 SVD는 항상 가능
- Positive positive (semi-)definite한 대칭 행렬  $S \in \mathbb{R}^{n \times n}$ 의 Eigendecomposition과 SVD는 사실상 동일

기계학습에선 Symmetric positive (semi-)definite matrix를 자주 다룸

- Item별 Feature를 나타내는 행렬  $A \in \mathbb{R}^{m \times n}$ 를 가정
- $A^T A$ 는 Data item간의 유사도 행렬을 나타냄
  - Data item 간의 내적을 일종의 유사도로 사용
- $AA^T$ 는 Feature간의 유사도를 나타냄
  - Feature간의 일종의 Correlation을 측정

$$A = \begin{array}{c|c|c|c} & \text{Item 1} & \text{Item 2} & \dots & \text{Item } n \end{array}$$

# 요약

- Singular value decomposition (SVD)의 정의
- Outer product들의 합으로 표현한 SVD의 표현
- Symmetric positive semi-definite 특성을 활용한 SVD 계산
- SVD를 활용한 행렬의 Low-rank approximation

The diagram illustrates the Singular Value Decomposition (SVD) of a matrix  $A$ . Matrix  $A$  is shown as a light gray rectangle with dimensions  $m$  (rows) and  $n$  (columns). It is equated to the product of three matrices:  $U$ ,  $\Sigma$ , and  $V^T$ .

- Matrix  $U$  is an  $m \times m$  matrix. It is composed of columns  $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_m$ . The first two columns are shaded gray, and the last column is also shaded gray.
- Matrix  $\Sigma$  is an  $m \times n$  matrix. It contains the singular values  $\sigma_1, \sigma_2, \dots, \sigma_n$  along its main diagonal, each in a shaded gray box. The rest of the matrix is white and labeled with a '0' at the bottom right, indicating zero values. A dashed line separates the diagonal elements from the zero block.
- Matrix  $V^T$  is an  $n \times n$  matrix. It is composed of rows  $\mathbf{v}_1^T, \mathbf{v}_2^T, \dots, \mathbf{v}_n^T$ . The first two rows are shaded gray, and the last row is also shaded gray.

Brackets above the matrices indicate their dimensions:  $n$  for  $A$ ,  $m$  for  $U$ ,  $n$  for  $\Sigma$ , and  $n$  for  $V^T$ . A bracket to the left of  $A$  indicates its dimension  $m$ . A bracket to the right of  $V^T$  indicates its dimension  $n$ .