Chapter 4. Matrix Decompositions

4.5 Singular Value Decomposition 특잇값 분해

4.6 Matrix Approximation 행렬 근사

(4.7 Matrix Phylogeny 행렬 계통)

한지수

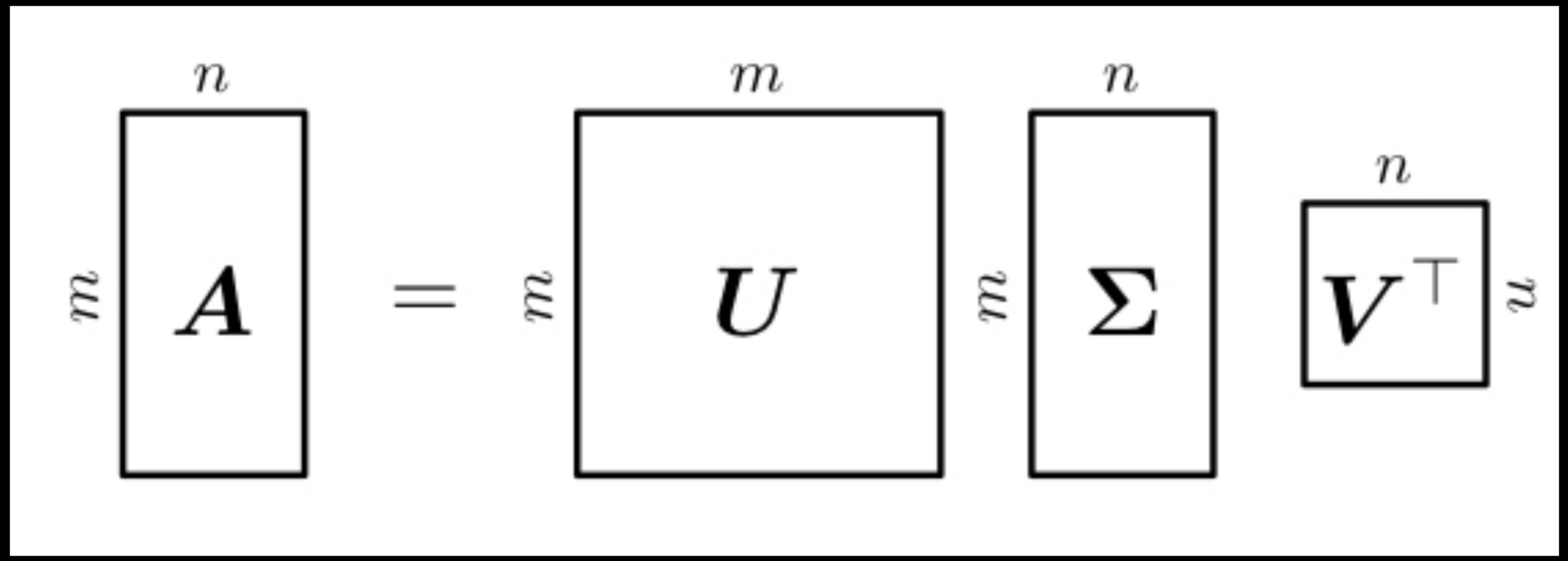
4.5 Single Matrix Decomposition 특잇값 분해

- 1 Geometric Intuitions for the SVD
- 2 Construction of the SVD
- 3 Eigenvalue Decomposition고윳값 분해 vs. Single Value Decomposition특잇값 분해

4.6 Matrix Approximation 행렬 근사

Single Matrix Decomposition (SVD)

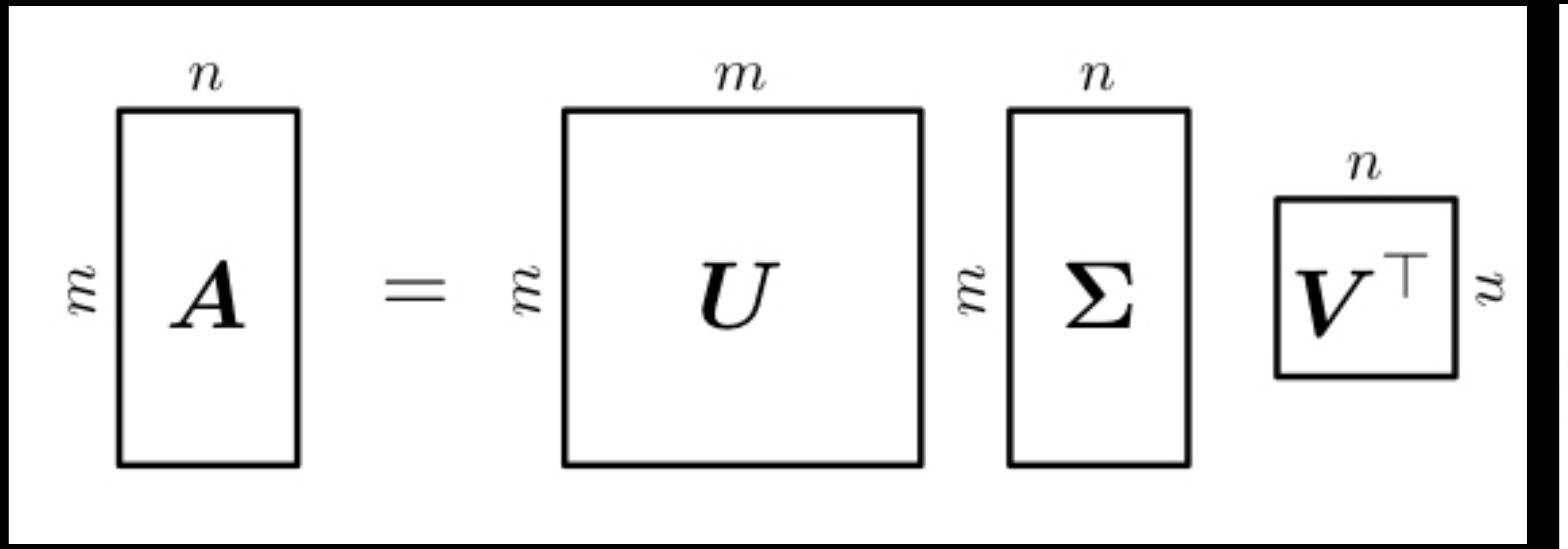
- The SVD of A is a decomposition of the form (Full SVD)



Rank $r \in [0, min(m, n)]$

  - Orthogonal Matrix $U \in R^{m \times m}$ (길이 m의 정사각행렬)
  - Orthogonal Matrix $V \in R^{n \times n}$ (길이 n의 정사각행렬)
  - Single Value Matrix $\Sigma \in R^{m \times n}$ ($\Sigma_{ii} = \sigma_i \geq 0, \Sigma_{ij} = 0, i \neq j$)
- The SVD exists for any Matrix $A \in R^{m \times n}$
- Used in a variety of applications in machine learning from least-squares problem in curve fitting

Singular Value, Singular Value Matrix

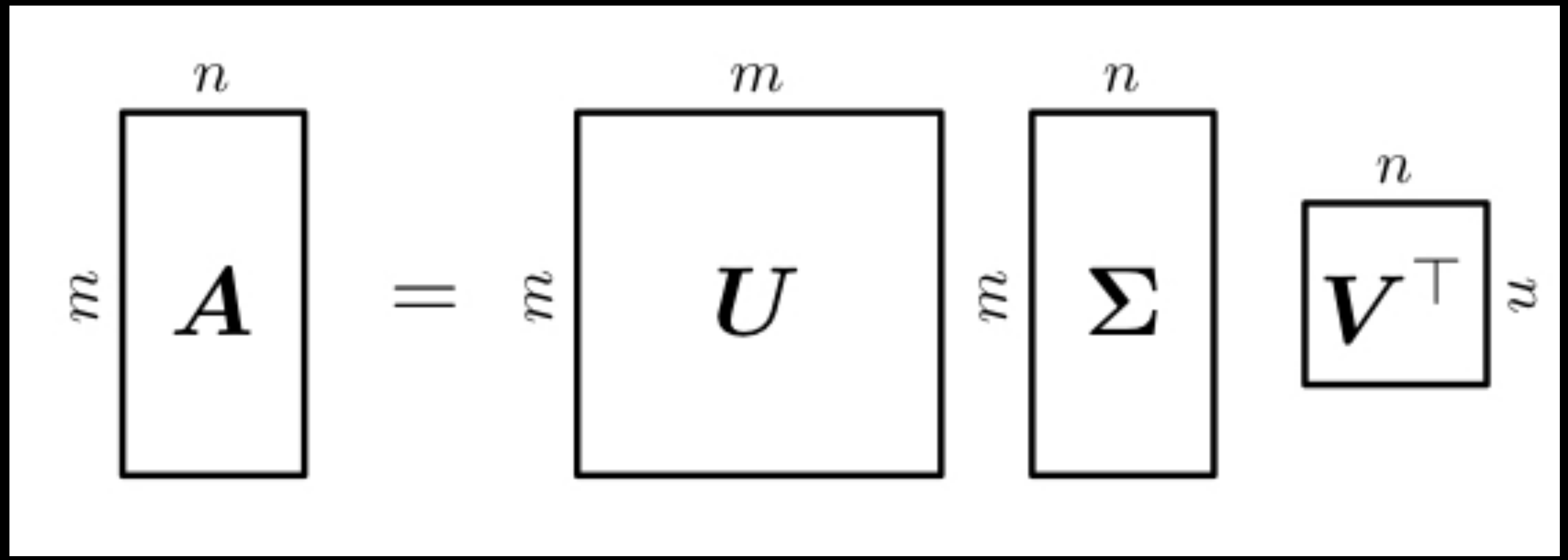
- Σ with Singular Values $\sigma_i, \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r \geq 0$
- U with column vectors: Left-Singular Vectors u_i
- V with column vectors: Right-Singular Vectors v_j
- 

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{bmatrix}$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ddots & 0 & \vdots & & \vdots \\ 0 & 0 & \sigma_m & 0 & \ldots & 0 \end{bmatrix}$$

Other SVD

- Reduced SVD (change U, Σ)



여기서 $U \in R^{m \times n}, \Sigma \in R^{n \times n}$

- Truncated SVD

: On Matrix approximation techniques using the SVD

Construction of the SVD (그냥 공식)

Using Eigenbasis of A^TA

- $A^TA = PDP^T = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^TU^TU\Sigma V^T = V\Sigma^T\Sigma V^T$

$$\boldsymbol{P}\begin{bmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{bmatrix}\boldsymbol{P}^\top$$

$$\boldsymbol{V}\begin{bmatrix}\sigma_1^2 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \sigma_n^2\end{bmatrix}\boldsymbol{V}^\top$$

- $V^T = P^T$
- $\gg \sigma_i^2 = \lambda_i$

Computing the SVD

- Step 1: get Right-Singular Vectors as eigenbasis of $A^T A$
  - Using the eigenvalue decomposition of $A^T A$
- Step 2: Singular-value Matrix
  - Using $\sigma_i^2 = \lambda_i$
- Step 3: Left-Singular vectors as the normalized Image of the right-singular vectors.
  - Using $u_i := \frac{Av_i}{\| Av_i \|} = \frac{1}{\sqrt{\lambda_i}} Av_i = \frac{1}{\sigma_i} Av_i$

Eigenvalue Decomposition vs. Single Value Decomposition

- $A = PDP^{-1}$
- Only defined for square matrix $R^{n \times n}$
- P is not necessarily orthogonal, inverse with P^{-1}
- Domain and codomain can't be vector spaces of different dimensions.

- $A = U\Sigma V^T$
- Always exists for any matrix $R^{m \times n}$
- U, V are orthonormal, not inverse of each other
- Domain and codomain can be vector spaces of different dimensions.

Eigenvalue Decomposition & Single Value Decomposition

- For Symmetric Matrices $A \in R^{n \times n}$, ED and SVD are the one and the same
- Both compositions of three linear mappings

  Change of basis in the domain, independent scaling of each new basis vector and mapping from domain to codomain, Change of basis in the codomain

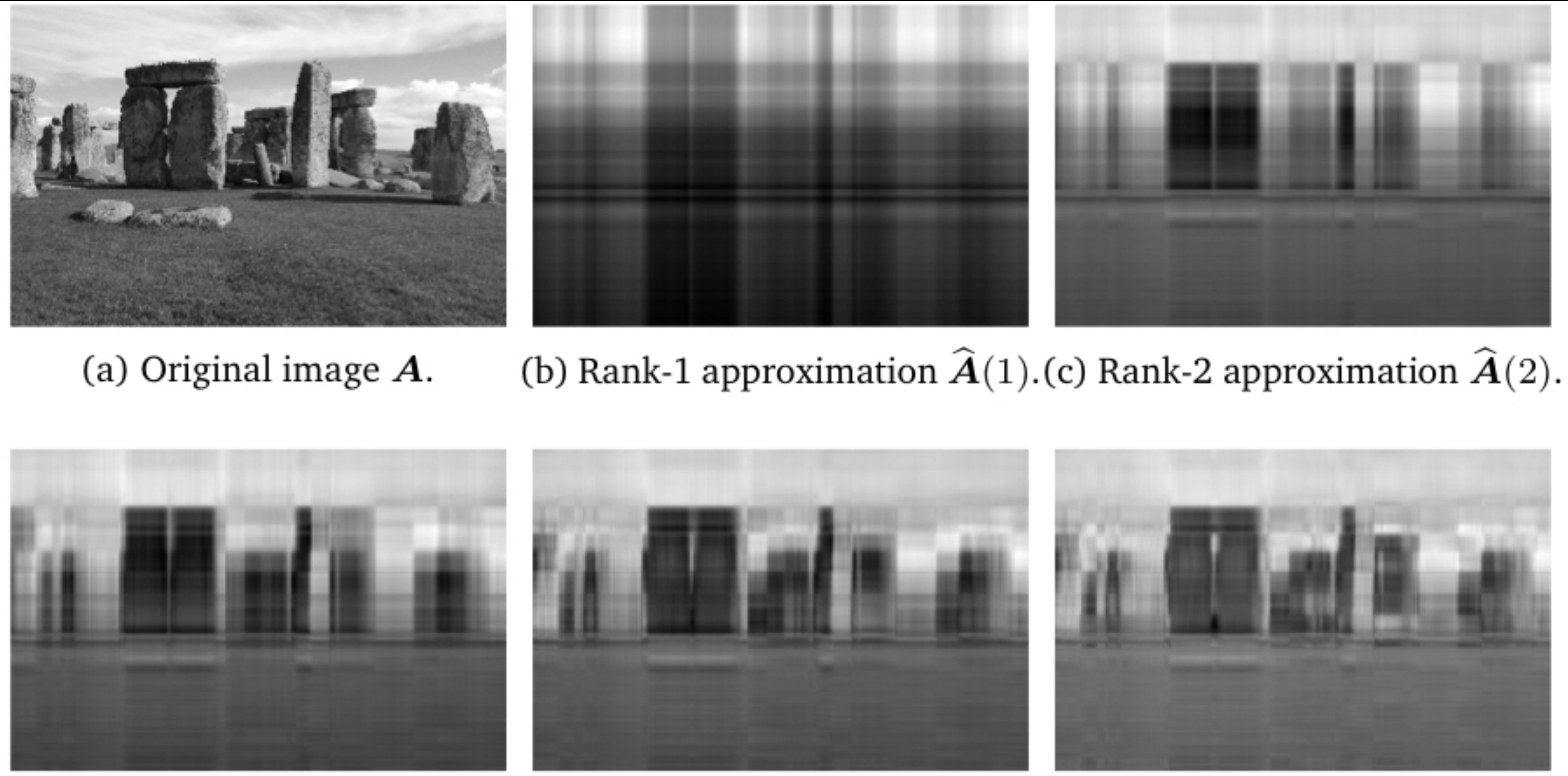
- Both closely related through their projections
  - The left-singular vector A
  - The right-singular vector A
  - The nonzero singular values of A

4.6 Matrix Approximation

- How the SVD allows us to **represent a matrix** A **as a simpler (low-rank) matrices** A_i
  - Rank-r matrix A to rank-k matrix $\widehat{A}$ in a spec
- Rank-k approximation: $\widehat{A}(k) := \sum_{i=1}^{k} \sigma_i A_i = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ (rank-r A, $k = 1...r$)
  - A_i : outer product of the ith orthogonal column vector of U and V

4.6 Matrix Approximation

- Rank-r matrix A to rank-k matrix $\widehat{A}$ in a principled, optimal manner



(a) Original image $\boldsymbol{A}$. (b) Rank-1 approximation $\widehat{\boldsymbol{A}}(1)$. (c) Rank-2 approximation $\widehat{\boldsymbol{A}}(2)$.

(d) Rank-3 approximation $\widehat{\boldsymbol{A}}(3)$. (e) Rank-4 approximation $\widehat{\boldsymbol{A}}(4)$. (f) Rank-5 approximation $\widehat{\boldsymbol{A}}(5)$.

- Image Processing, noise filtering, and regularization etc.