Math Refresher for DS

Lecture 5

girafe
ai

Last Time

- Eigenvalues & eigenvectors
- Eigendecomposition
  - Matrix diagonalization;
  - PCA.

Today

- Singular Value Decomposition

Reminder: Eigenvalues & Eigenvectors

- Let A be an $n \times n$ matrix.
- $v \in \mathbb{R}^n$ – eigenvector with eigenvalue λ if

$$Av = \lambda v$$

Reminder: Eigenvalues & Eigenvectors

- Let A be an $n \times n$ matrix.
- $v \in \mathbb{R}^n$ – eigenvector with eigenvalue λ if

$$Av = \lambda v$$

- To find all λs, we need to find roots of $\det(A - \lambda E)$.
  There are n roots in total (incl. repetitions).

girafe ai

Reminder: Eigenvalues & Eigenvectors

- Let A be an $n \times n$ matrix.
- $v \in \mathbb{R}^n$ – eigenvector with eigenvalue λ if

$$Av = \lambda v$$

- To find all λs, we need to find roots of $\det(A - \lambda E)$.
  There are n roots in total (incl. repetitions).

$$\det(A - \lambda E) \sim (\lambda - \lambda_1)^{n_1} \cdot \ldots \cdot (\lambda - \lambda_k)^{n_k}, \qquad n_1 + \cdots + n_k = n,$$

n_i – algebraic multiplicity of λ_i.

Reminder: Eigenvalues & Eigenvectors

- Let A be an $n \times n$ matrix.
- $v \in \mathbb{R}^n$ – eigenvector with eigenvalue λ if

$$Av = \lambda v$$

- To find all λs, we need to find roots of $\det(A - \lambda E)$.
  There are n roots in total (incl. repetitions).

$$\det(A - \lambda E) \sim (\lambda - \lambda_1)^{n_1} \cdot \ldots \cdot (\lambda - \lambda_k)^{n_k}, \qquad n_1 + \cdots + n_k = n,$$

n_i – algebraic multiplicity of λ_i.

- $E_{\lambda_i} = span\,\{v : Av = \lambda_i v\}, \;\; dimE_{\lambda_i} \leq n_i$ – geometric multiplicity.

Reminder: Eigendecomposition

- Let A be an $n \times n$ matrix.

Reminder: Eigendecomposition

- Let A be an $n \times n$ matrix.
- If A has n linearly independent eigenvectors $v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$, then A can be decomposed as follows:

$$A = V \Lambda V^{-1},$$

where $V = [v_1 \mid v_2 \mid \dots \mid v_n]$, $\quad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$

girafe ai

Reminder: Eigendecomposition

- Let A be an $n \times n$ matrix.
- If A has n linearly independent eigenvectors $v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$, then A can be decomposed as follows:

$$A = V\Lambda V^{-1},$$

$$\text{where } V = [v_1 \mid v_2 \mid \dots \mid v_n], \qquad \Lambda = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$$

- If A is symmetric, then V is orthogonal and $A = V\Lambda V^T$.

girafe ai

Reminder: PCA

$X - m \times n$ data matrix (m features, n examples)

$S = \frac{1}{n-1} X X^T$ – data covariance matrix ($m \times m$)

$$S = V \Lambda V^{-1} = V \Lambda V^T$$

$$\begin{bmatrix} \mathbf{s_{11}} & s_{12} & \cdots & s_{1m} \\ s_{21} & \mathbf{s_{22}} & \cdots & s_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ s_{n1} & s_{n2} & \cdots & \mathbf{s_{mm}} \end{bmatrix} = \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{mm} \end{bmatrix} \begin{bmatrix} \boldsymbol{\lambda_1} & 0 & \cdots & 0 \\ 0 & \boldsymbol{\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \boldsymbol{\lambda_m} \end{bmatrix} \begin{bmatrix} v_{11} & v_{12} & \cdots & v_{1m} \\ v_{21} & v_{22} & \cdots & v_{2m} \\ \vdots & \vdots & \cdots & \vdots \\ v_{n1} & v_{n2} & \cdots & v_{mm} \end{bmatrix}^T$$

Total variance of the data $\mathrm{T} = tr(S) = s_{11} + \cdots + s_{nn} = \lambda_1 + \cdots \lambda_m$

Orthogonal eigenvectors $v_1, \dots, v_n$ – principal components of the data

Direction of v_i describes λ_i out of the total variance T.

Eigenvalues of $A^T A$

- Let A be an $m \times n$ matrix.
- $A^T A$ is an $n \times n$ symmetric matrix.

Eigenvalues of $A^T A$

- Let A be an $m \times n$ matrix.

- $A^T A$ is an $n \times n$ symmetric matrix. Therefore, $A^T A$ has n linearly independent eigenvectors $v_1, \ldots, v_n$ with eigenvalues $\lambda_1, \ldots, \lambda_n$.

Eigenvalues of $A^T A$

- Let A be an $m \times n$ matrix.

- $A^T A$ is an $n \times n$ symmetric matrix. Therefore, $A^T A$ has n linearly independent eigenvectors $v_1, \dots, v_n$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

- We can show that $\boldsymbol{\lambda_i \geq 0}$:

$$0 \leq \|Av_i\|^2 = (Av_i, Av_i) = (Av_i)^T Av_i = v_i^T A^T A v_i = v_i^T \lambda_i v_i = \lambda_i \|v_i\|^2$$

$$\Leftrightarrow$$

$$\lambda_i \geq 0.$$

Eigenvalues of AA^T

- Let A be an $m \times n$ matrix.

- AA^T is an $m \times m$ symmetric matrix. Therefore, AA^T has m linearly independent eigenvectors $u_1, \dots, u_m$ with eigenvalues $\lambda_1, \dots, \lambda_n$.

- We can show that $\lambda_i \geq 0$:

$$0 \leq \|A^T u_i\|^2 = (A^T u_i, A^T u_i) = (A^T u_i)^T A^T u_i = u_i^T A A^T u_i = v_i^T \lambda_i u_i = \lambda_i \|u_i\|^2$$

$$\Leftrightarrow$$

$$\lambda_i \geq 0.$$

girafe ai

Positive Definite Matrices

- Square matrices with non-negative eigenvalues $\lambda_i \geq 0$ are called positive semi-definite.

$$A \text{ is positive definite} \Leftrightarrow x^T A x \geq 0 \ \forall x \in \mathbb{R}^n$$

- Square matrices with positive eigenvalues $\lambda_i > 0$ are called positive definite.

$$A \text{ is positive definite} \Leftrightarrow x^T A x > 0 \ \forall x \neq 0 \in \mathbb{R}^n$$

girafe ai

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Matrices $\boldsymbol{A^T A}$ and $\boldsymbol{AA^T}$ share the same set of non-zero eigenvalues!

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Matrices $\boldsymbol{A^T A}$ and $\boldsymbol{AA^T}$ share the same set of non-zero eigenvalues!

$$(A^T A)v = \lambda v, \qquad \lambda \neq 0$$

girafe ai

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Matrices $\boldsymbol{A^T A}$ and $\boldsymbol{AA^T}$ share the same set of non-zero eigenvalues!

$$(A^T A)v = \lambda v, \qquad \lambda \neq 0$$

$$\color{green}{A}A^T Av = \lambda \color{green}{A}v$$

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Matrices $\boldsymbol{A^T A}$ and $\boldsymbol{AA^T}$ share the same set of non-zero eigenvalues!

$$(A^T A)v = \lambda v, \qquad \lambda \neq 0$$

$$AA^T Av = \lambda Av$$

$v \neq 0$ is an eigenvector of $A^T A$ with $\lambda \neq 0 \iff$
$Av \neq 0$ is an eigenvalue of AA^T with the same eigenvalue $\lambda \neq 0$.

girafe ai

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.
- What is the relationship between eigenvalues of $A^T A$ and AA^T?

Matrices $\boldsymbol{A^T A}$ and $\boldsymbol{AA^T}$ share the same set of non-zero eigenvalues!

$$(AA^T)u = \lambda u, \qquad \lambda \neq 0$$

$$A^T AA^T u = \lambda A^T u$$

$u \neq 0$ is an eigenvector of AA^T with $\lambda \neq 0 \iff$
$A^T u \neq 0$ is an eigenvalue of $A^T A$ with the same eigenvalue $\lambda \neq 0$.

girafe ai

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.

$$v \neq 0 \text{ is an eigenvector of } A^T A \text{ with } \lambda \neq 0 \iff$$
$$Av \neq 0 \text{ is an eigenvalue of } AA^T \text{ with the same eigenvalue } \lambda \neq 0.$$

$$u \neq 0 \text{ is an eigenvector of } AA^T \text{ with } \lambda \neq 0 \iff$$
$$A^T u \neq 0 \text{ is an eigenvalue of } A^T A \text{ with the same eigenvalue } \lambda \neq 0.$$

- Why this is useful? Imagine that A is a 1000×2 matrix.

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.

$$v \neq 0 \text{ is an eigenvector of } A^T A \text{ with } \lambda \neq 0 \iff$$
$$Av \neq 0 \text{ is an eigenvalue of } AA^T \text{ with the same eigenvalue } \lambda \neq 0.$$

$$u \neq 0 \text{ is an eigenvector of } AA^T \text{ with } \lambda \neq 0 \iff$$
$$A^T u \neq 0 \text{ is an eigenvalue of } A^T A \text{ with the same eigenvalue } \lambda \neq 0.$$

- Why this is useful? Imagine that A is a 1000×2 matrix.

- You need to find eigenvalues of AA^T

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.

$v \neq 0$ is an eigenvector of $A^T A$ with $\lambda \neq 0$ $\iff$
$Av \neq 0$ is an eigenvalue of AA^T with the same eigenvalue $\lambda \neq 0$.

$u \neq 0$ is an eigenvector of AA^T with $\lambda \neq 0$ $\iff$
$A^T u \neq 0$ is an eigenvalue of $A^T A$ with the same eigenvalue $\lambda \neq 0$.

- Why this is useful? Imagine that A is a 1000×2 matrix.

- You need to find eigenvalues of AA^T which is a 1000×1000 matrix.

Eigenvalues of $A^T A$ and AA^T

- Let A be an $m \times n$ matrix.

$$v \neq 0 \text{ is an eigenvector of } A^T A \text{ with } \lambda \neq 0 \iff$$
$$Av \neq 0 \text{ is an eigenvalue of } AA^T \text{ with the same eigenvalue } \lambda \neq 0.$$

$$u \neq 0 \text{ is an eigenvector of } AA^T \text{ with } \lambda \neq 0 \iff$$
$$A^T u \neq 0 \text{ is an eigenvalue of } A^T A \text{ with the same eigenvalue } \lambda \neq 0.$$

- Why this is useful? Imagine that A is a 1000×2 matrix.

- You need to find eigenvalues of AA^T which is a 1000×1000 matrix.

Trick: compute eigenvalues of $A^T A$ instead!

SVD: Motivation

Eigendecomposition: Problems

- Only symmetric matrices are guaranteed to have n linearly independent eigenvectors *(The Spectral Theorem)*.

Eigendecomposition: Problems

- Only symmetric matrices are guaranteed to have n linearly independent eigenvectors *(The Spectral Theorem)*.

- Non-symmetric matrices may or may not have eigenbasis.

Eigendecomposition: Problems

- Only symmetric matrices are guaranteed to have n linearly independent eigenvectors *(The Spectral Theorem)*.

- Non-symmetric matrices may or may not have eigenbasis.

- And what about non-square matrices?

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} = ?$$

Eigendecomposition: Problems

- Only symmetric matrices are guaranteed to have n linearly independent eigenvectors *(The Spectral Theorem)*.

- Non-symmetric matrices may or may not have eigenbasis.

- And what about non-square matrices?

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix} = ?$$

Singular Value Decomposition : generalization of eigendecomposition for all matrices.

girafe ai

SVD:
Main Idea

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \;\; V = [v_1 \,|\ldots|\, v_n] - \text{orthogonal matrices},$$

Σ – "diagonal matrix"

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \quad V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

Σ – “diagonal matrix”

$$m \geq n: \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ & & 0 & \end{bmatrix}$$

girafe ai

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T\text{, where}$$

$$U = [u_1 \,|\ldots|\, u_m],\;\; V = [v_1 \,|\ldots|\, v_n] - \text{orthogonal matrices,}$$

Σ – “diagonal matrix”

$$m \geq n:\quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \boxed{\quad} & 0 & & \end{bmatrix}$$

$\boldsymbol{m - n}$ zero rows

girafe ai

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m],\;\; V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

Σ – “diagonal matrix”

$$m \geq n:\;\; \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \hline \multicolumn{4}{c}{0} \end{bmatrix}, \qquad m < n:\;\; \Sigma = \left[\begin{matrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_m \end{matrix}\;\; 0\right]$$

$\boldsymbol{m-n}$ zero rows

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1\,|\ldots|u_m],\;\; V = [v_1\,|\ldots|v_n] \text{ – orthogonal matrices,}$$

Σ – "diagonal matrix"

$$m \geq n:\;\; \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \hline & & \boxed{0} & \end{bmatrix}, \qquad m < n:\;\; \Sigma = \left[\begin{array}{cccc|c} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & \boxed{0} \\ 0 & 0 & \cdots & \sigma_m & \end{array}\right]$$

$\boldsymbol{m-n}$ zero rows

$\boldsymbol{n-m}$ zero columns

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\dots|\, u_m],\;\; V = [v_1 \,|\dots|\, v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,\;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

$$m \geq n:\quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ \hline & & 0 & \end{bmatrix}, \qquad m < n:\quad \Sigma = \left[\begin{array}{cccc|c} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \sigma_m & \end{array}\right]$$

$\boldsymbol{m-n}$ zero rows

$\boldsymbol{n-m}$ zero columns

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \;\; V = [v_1 \,|\ldots|\, v_n] - \text{orthogonal matrices,}$$

$$\Sigma - \text{“diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

girafe ai

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \;\; V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

- $u_1, \ldots, u_m$ – left singular vectors of A.
  $v_1, \ldots, v_n$ – right singular vectors of A.

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m],\ \ V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,\ \ \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

- $u_1, \ldots, u_m$ – left singular vectors of A.
  $v_1, \ldots, v_n$ – right singular vectors of A.
- $\sigma_1, \ldots, \sigma_{\max(m,n)}$ – singular values of A.

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|u_m],\ \ V = [v_1\,|\ldots|v_n] - \text{orthogonal matrices,}$$

$$\Sigma - \text{“diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,\ \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

- $u_1, \ldots, u_m$ – left singular vectors of A.
  $v_1, \ldots, v_n$ – right singular vectors of A.
- $\sigma_1, \ldots, \sigma_{\max(m,n)}$ – singular values of A.

- Unlike in eigendecomposition, U and V are (generally) not the same.

SVD: Main Idea

$$m \geq n: \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ & & 0 & \end{bmatrix}, \qquad m < n: \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \sigma_m & \end{bmatrix}$$

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \;\; V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

$$\Leftrightarrow$$

girafe ai

SVD: Main Idea

$$m \geq n: \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ & & 0 & \end{bmatrix}, \qquad m < n: \quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \\ 0 & 0 & \cdots & \sigma_m & \end{bmatrix}$$

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m], \;\; V = [v_1 \,|\ldots|\, v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

$$\Leftrightarrow$$

$$A =$$

SVD: Main Idea

$$m \geq n: \quad \Sigma = \begin{bmatrix} \begin{matrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{matrix} \\ 0 \end{bmatrix}, \qquad m < n: \quad \Sigma = \begin{bmatrix} \begin{matrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_m \end{matrix} & 0 \end{bmatrix}$$

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 | \ldots | u_m], \;\; V = [v_1 | \ldots | v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – “diagonal matrix” with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

$$\Leftrightarrow$$

$$A = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_r\sigma_r v_r^T.$$

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T \iff A = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_{\boldsymbol{r}}\sigma_{\boldsymbol{r}} v_{\boldsymbol{r}}^T.$$

- *(Reduced SVD)*:

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T \iff A = u_1\sigma_1 v_1^T + u_2\sigma_2 v_2^T + \cdots + u_r\sigma_r v_r^T.$$

- *(Reduced SVD)*: A can be decomposed as

$$A_{m\times n} = U^r_{m\times r}\Sigma^r_{r\times r}(V^r_{n\times r})^T, \text{ where}$$

$$U^r = [u_1 | \ldots | u_r],\ \ V^r = [v_1 | \ldots | v_r] - \text{orthogonal matrices,}$$

$$\Sigma^r - \text{diagonal matrix with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$$

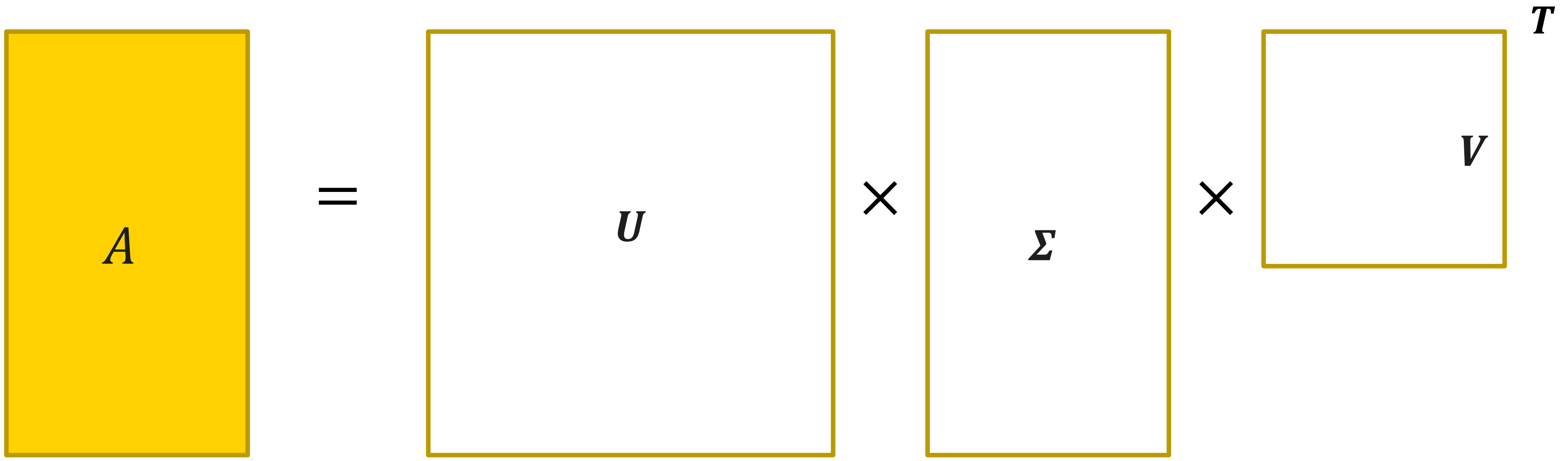
Reduced SVD: Main Idea

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$U = [u_1 \,|\ldots|\, u_m]$, $V = [v_1 \,|\ldots|\, v_n]$ – orthogonal matrices,

Σ – "diagonal matrix" with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$,

$\sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$

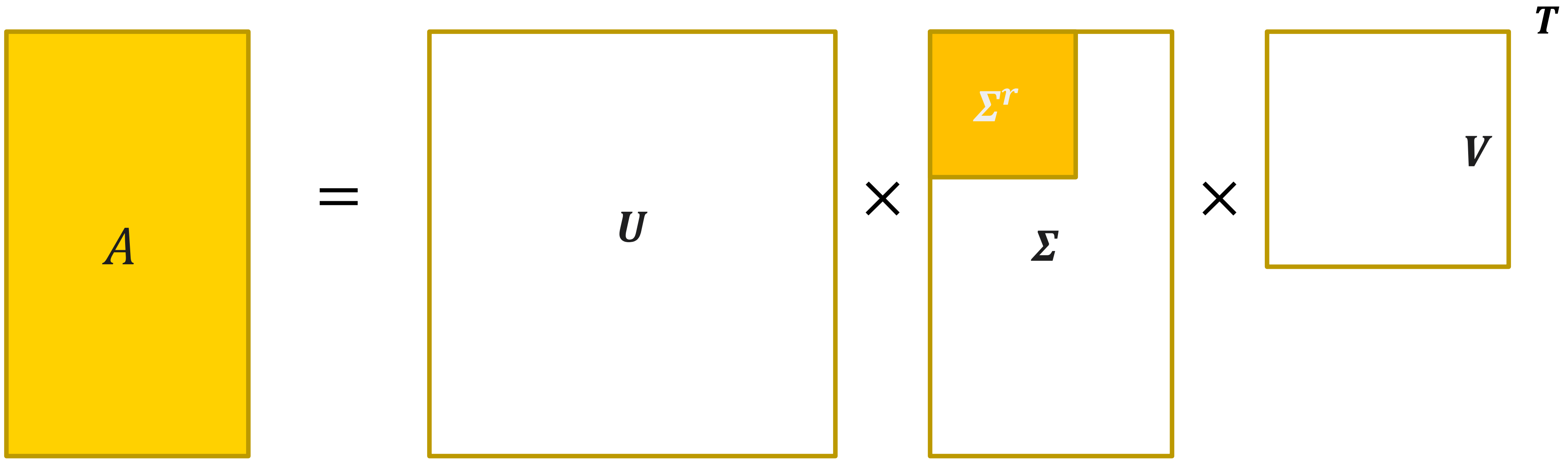


Reduced SVD: Main Idea

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$U = [u_1 \,|\ldots|\, u_m]$, $V = [v_1 \,|\ldots|\, v_n]$ – orthogonal matrices,

Σ – "diagonal matrix" with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$,
$\sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$



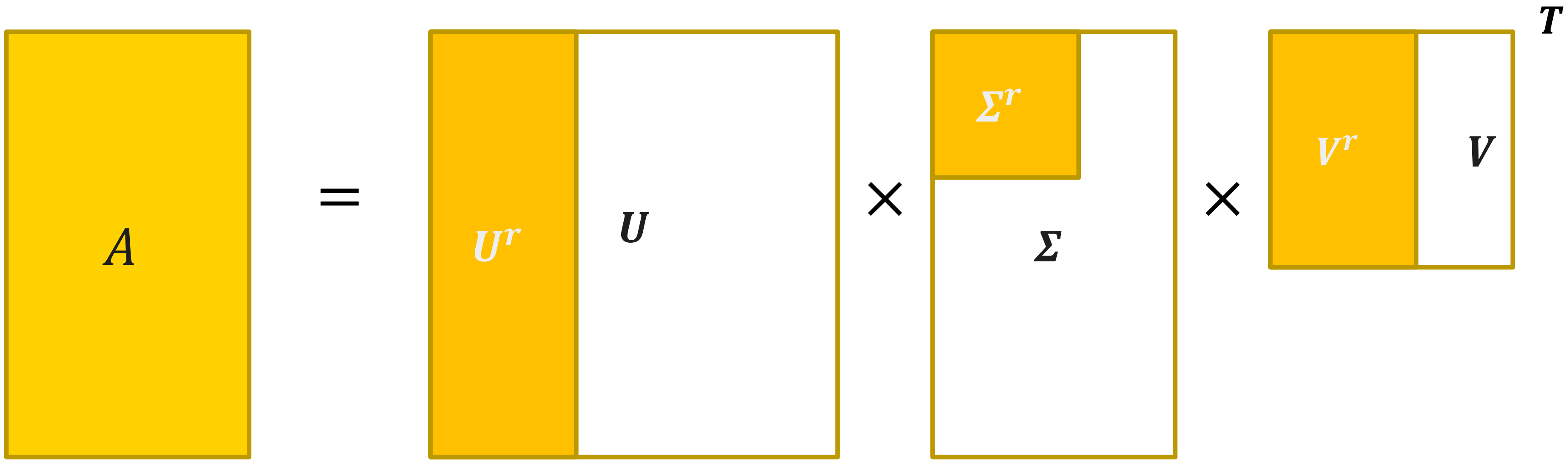
girafe ai

Reduced SVD: Main Idea

$$A_{m\times n} = U^r_{m\times r}\Sigma^r_{r\times r}(V^r_{n\times r})^T,$$ where

$U^r = [u_1 | \ldots | u_r]$, $V^r = [v_1 | \ldots | v_r]$ – orthogonal matrices,

Σ^r – diagonal matrix with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.



SVD: Derivation

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|\, u_m],\ \ V = [v_1 \,|\ldots|\, v_n] - \text{orthogonal matrices,}$$

Σ – "diagonal matrix" with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0,\ \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$

SVD: Main Idea

- Let A be an $m \times n$ matrix.
- *(SVD)*: A can be decomposed as

$$A_{m\times n} = U_{m\times m}\Sigma_{m\times n}(V_{n\times n})^T, \text{ where}$$

$$U = [u_1 \,|\ldots|u_m], \;\; V = [v_1 \,|\ldots|v_n] \text{ – orthogonal matrices,}$$

$$\Sigma \text{ – "diagonal matrix" with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0, \;\; \sigma_{r+1} = \cdots = \sigma_{\max(m,n)} = 0$$

How do we arrive to this?

SVD

- Let A be an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

SVD

- Let A be an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$row(A) = span\{A_1, \dots, A_m\} \subseteq \mathbb{R}^n$, $\{v_1, \dots, v_r\}$ – orthonormal basis

SVD

- Let A be an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$row(A) = span\{A_1, \dots, A_m\} \subseteq \mathbb{R}^n,$ $\{v_1, \dots, v_r\}$ – orthonormal basis

$col(A) = span\{A^1, \dots, A^n\} \subseteq \mathbb{R}^m,$ $\{u_1, \dots, u_r\}$ – orthonormal basis

SVD

- Let A be an $m \times n$ matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

$row(A) = span\{A_1, \dots, A_m\} \subseteq \mathbb{R}^n$, $\{v_1, \dots, v_r\}$ – orthonormal basis

$col(A) = span\{A^1, \dots, A^n\} \subseteq \mathbb{R}^m$, $\{u_1, \dots, u_r\}$ – orthonormal basis

Key idea: let's find $\boldsymbol{v_i}, \boldsymbol{u_i}$ such that $\boldsymbol{Av_i} = \boldsymbol{\sigma_i u_i}$.

girafe ai

SVD

- Let A be an $m \times n$ matrix.

- $\{v_1, \dots, v_r\}$ – orthonormal basis of $row(A)$,
  $\{u_1, \dots, u_r\}$ – orthonormal basis of $col(A)$

SVD

- Let A be an $m \times n$ matrix.

- $\{v_1, \dots, v_r\}$ – orthonormal basis of $row(A)$, $v_{r+1}, \dots, v_n \in null(A)$,
$\{u_1, \dots, u_r\}$ – orthonormal basis of $col(A)$

- $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ – orthonormal basis of $\mathbb{R}^n$,

SVD

- Let A be an $m \times n$ matrix.

- $\{v_1, \ldots, v_r\}$ – orthonormal basis of $row(A)$, $v_{r+1}, \ldots, v_n \in null(A)$,
  $\{u_1, \ldots, u_r\}$ – orthonormal basis of $col(A)$ $u_{r+1}, \ldots, u_m \in null(A^T)$.

- $\{v_1, \ldots, v_r, v_{r+1}, \ldots, v_n\}$ – orthonormal basis of $\mathbb{R}^n$,
  $\{u_1, \ldots, u_r, u_{r+1}, \ldots, u_m\}$ – orthonormal basis of $\mathbb{R}^m$:

SVD

- Let A be an $m \times n$ matrix.
- $\{v_1, \dots, v_r\}$ – orthonormal basis of $row(A)$, $v_{r+1}, \dots, v_n \in null(A)$,
  $\{u_1, \dots, u_r\}$ – orthonormal basis of $col(A)$ $u_{r+1}, \dots, u_m \in null(A^T)$.
- $\{v_1, \dots, v_r, v_{r+1}, \dots, v_n\}$ – orthonormal basis of $\mathbb{R}^n$,
  $\{u_1, \dots, u_r, u_{r+1}, \dots, u_m\}$ – orthonormal basis of $\mathbb{R}^m$:

Let's find v_i, u_i such that $Av_i = \sigma_i u_i$

$$\Leftrightarrow$$

$A_{m\times n} V_{n\times n} = U_{m\times m} \Sigma_{m\times n}$, where

$U = [u_1 | \dots | u_m]$, $V = [v_1 | \dots | v_n]$ – orthogonal matrices,

Σ – "diagonal matrix" with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$, $\sigma_{r+1} = \cdots = 0$

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma\mathrm{V}^{-1} = U\Sigma\mathrm{V}^{\mathrm{T}}$$ (V is orthogonal)

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma\mathrm{V}^{-1} = U\Sigma\mathrm{V}^{\mathrm{T}} \text{ (V is orthogonal)}$$

$$A^T A =$$

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma\mathrm{V}^{-1} = U\Sigma\mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

$$A^{T}A = \left(U\Sigma\mathrm{V}^{\mathrm{T}}\right)^{T} U\Sigma\mathrm{V}^{\mathrm{T}} =$$

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma \mathrm{V}^{-1} = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

$$A^T A = \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T U\Sigma \mathrm{V}^{\mathrm{T}} = V\Sigma^T \Sigma U^T U V^T =$$

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma \mathrm{V}^{-1} = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

$$A^T A = \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T U\Sigma \mathrm{V}^{\mathrm{T}} = V\Sigma^T\Sigma U^T U V^T = V_{n\times n}(\Sigma^T\Sigma)_{n\times n}(V_{n\times n})^T$$

girafe ai

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma \mathrm{V}^{-1} = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ } (V \text{ is orthogonal})$$

$$A^T A = \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T U\Sigma \mathrm{V}^{\mathrm{T}} = V\Sigma^T \Sigma U^T U V^T = V_{n\times n}(\Sigma^T\Sigma)_{n\times n}(V_{n\times n})^T$$

$$V = [v_1\,|\ldots|v_n] - \text{eigenvectors of } A^T A$$

SVD

$$A_{m\times n}V_{n\times n} = U_{m\times m}\,\Sigma_{m\times n}$$

$$A = U\Sigma \mathrm{V}^{-1} = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

$$A^T A = \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T U\Sigma \mathrm{V}^{\mathrm{T}} = V\Sigma^T\Sigma U^T U V^T = V_{n\times n}(\Sigma^T\Sigma)_{n\times n}(V_{n\times n})^T$$

$$V = [v_1\,|\ldots|v_n] - \text{eigenvectors of } A^T A$$

$$m \geq n:\quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n \\ & & 0 & \end{bmatrix},\; m < n:\quad \Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & \\ 0 & \sigma_2 & \cdots & 0 & \\ \vdots & \vdots & \vdots & \vdots & 0 \\ 0 & 0 & \cdots & \sigma_m & \end{bmatrix}$$

$$\Sigma^T\Sigma = \begin{bmatrix} \sigma_1^2 & 0 & \cdots & 0 \\ 0 & \sigma_2^2 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & \sigma_n^2 \end{bmatrix},\quad \sigma_i^2 = \lambda_i - \text{eigenvalues of } A^T A.$$

girafe ai

SVD

$A = U\Sigma V^{\mathrm{T}}$ (V is orthogonal)

By multiplying by A^T on the left we got that
$V = [v_1 | \dots | v_n]$ – eigenvectors of $A^T A$,
$\sigma_1^2, \dots, \sigma_n^2$ – corresponding eigenvalues *(some of them possibly 0s).*

SVD

$A = U\Sigma V^{T}$ (V is orthogonal)

By multiplying by A^T on the left we got that
$V = [v_1 \,|\ldots|\, v_n]$ – eigenvectors of $A^T A$,
$\sigma_1^2, \ldots, \sigma_n^2$ – corresponding eigenvalues *(some of them possibly 0s).*

Similarly, by multiplying by A^T on the right we get

SVD

$$A = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

By multiplying by A^T on the left we got that
$V = [v_1 | \ldots | v_n]$ – eigenvectors of $A^T A$,
$\sigma_1^2, \ldots, \sigma_n^2$ – corresponding eigenvalues *(some of them possibly 0s).*

Similarly, by multiplying by A^T on the right we get

$$AA^T = U\Sigma \mathrm{V}^{\mathrm{T}} \cdot \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T = U\Sigma V^T V \Sigma^T U^T = U_{m\times m}(\Sigma\Sigma^T)_{m\times m}(U_{m\times m})^T$$

girafe ai

SVD

$$A = U\Sigma \mathrm{V}^{\mathrm{T}} \text{ (}V\text{ is orthogonal)}$$

By multiplying by A^T on the left we got that

$V = [v_1 \,|\ldots|\, v_n]$ – eigenvectors of $A^T A$,

$\sigma_1^2, \ldots, \sigma_n^2$ – corresponding eigenvalues *(some of them possibly 0s).*

Similarly, by multiplying by A^T on the right we get

$$AA^T = U\Sigma \mathrm{V}^{\mathrm{T}} \cdot \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T = U\Sigma V^T V \Sigma^T U^T = U_{m\times m}(\Sigma\Sigma^T)_{m\times m}(U_{m\times m})^T$$

$U = [u_1 \,|\ldots|\, u_m]$ – eigenvectors of AA^T,

SVD

$A = U\Sigma \mathrm{V}^{\mathrm{T}}$ (V is orthogonal)

By multiplying by A^T on the left we got that
$V = [v_1 \,|\ldots|\, v_n]$ – eigenvectors of $A^T A$,
$\sigma_1^2, \ldots, \sigma_n^2$ – corresponding eigenvalues *(some of them possibly 0s).*

Similarly, by multiplying by A^T on the right we get

$$AA^T = U\Sigma \mathrm{V}^{\mathrm{T}} \cdot \left(U\Sigma \mathrm{V}^{\mathrm{T}}\right)^T = U\Sigma V^T V \Sigma^T U^T = U_{m\times m}(\Sigma\Sigma^T)_{m\times m}(U_{m\times m})^T$$

$U = [u_1 \,|\ldots|\, u_m]$ – eigenvectors of AA^T,
$\sigma_1^2, \ldots, \sigma_m^2$ – corresponding eigenvalues *(some of them possibly 0s).*

girafe ai

SVD: Example

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times3} = U_{2\times2}\,\Sigma_{2\times3}(V_{3\times3})^T$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\,\Sigma_{2\times 3}(V_{3\times 3})^T$$

Singular values = (non-zero) eigenvalues of AA^T:

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\,\Sigma_{2\times 3}(V_{3\times 3})^T$$

Singular values = (non-zero) eigenvalues of AA^T:

$$AA^T = \begin{pmatrix} 17 & 8 \\ 8 & 17 \end{pmatrix}, \qquad \det(AA^T - \lambda E) = (\lambda - 25)(\lambda - 9) = 0 \Leftrightarrow$$

$$\sigma_1 = \sqrt{25} = 5, \qquad \sigma_2 = \sqrt{9} = 3$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\,\Sigma_{2\times 3}(V_{3\times 3})^T$$

$$\boldsymbol{\Sigma} = \begin{pmatrix} \mathbf{5} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{3} & \mathbf{0} \end{pmatrix}, \qquad V = ?, \qquad U = ?$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T, \qquad \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

Columns of V are eigenvectors of $A^T A$.
Eigenvalues of $A^T A$ are 25, 9 and 0.

$$A^T A - 25E = \begin{pmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -17 \end{pmatrix} \sim \ldots \to v_1 = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \end{pmatrix}^T$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T, \qquad \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

Columns of V are eigenvectors of $A^T A$.
Eigenvalues of $A^T A$ are 25, 9 and 0.

$$A^T A - 9E = \begin{pmatrix} 4 & 12 & 2 \\ 12 & 4 & -2 \\ 2 & -2 & -1 \end{pmatrix} \sim \ldots \to v_2 = \left(\frac{1}{3\sqrt{2}} \quad \frac{-1}{3\sqrt{2}} \quad \frac{4}{3\sqrt{2}}\right)^T$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times3} = U_{2\times2}\Sigma_{2\times3}(V_{3\times3})^T, \qquad \Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}$$

Columns of V are eigenvectors of A^TA.
Eigenvalues of A^TA are 25, 9 and 0.

$$A^TA - 0E = \begin{pmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 10 \end{pmatrix} \sim \dots \to v_3 = \begin{pmatrix} \frac{2}{3} & \frac{-2}{3} & \frac{-1}{3} \end{pmatrix}^T$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad \boldsymbol{V} = \begin{pmatrix} \mathbf{1/\sqrt{2}} & \mathbf{1/3\sqrt{2}} & \mathbf{2/3} \\ \mathbf{-1/\sqrt{2}} & \mathbf{-1/3\sqrt{2}} & \mathbf{-2/3} \\ \mathbf{0} & \mathbf{4/3\sqrt{2}} & \mathbf{-1/3} \end{pmatrix}, \qquad U = ?$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}, \qquad U =?$$

Remember: $\boldsymbol{Av_i = \sigma_i u_i}$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}, \qquad U = ?$$

$$\begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}\begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{pmatrix} = 5u_1 \Rightarrow u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}, \qquad U = ?$$

$$u_1 = \begin{pmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{pmatrix}, \qquad \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}\begin{pmatrix} 1/3\sqrt{2} \\ -1/3\sqrt{2} \\ 4/3\sqrt{2} \end{pmatrix} = 3u_1 \Rightarrow u_2 = \begin{pmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{pmatrix}$$

Example

- Let's find SVD and reduced SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\Sigma = \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix}, \qquad V = \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}, \qquad \boldsymbol{U} = \begin{pmatrix} \mathbf{1/\sqrt{2}} & \mathbf{1/\sqrt{2}} \\ \mathbf{1/\sqrt{2}} & \mathbf{-1/\sqrt{2}} \end{pmatrix}$$

Example

- SVD of

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$\boldsymbol{A} = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}^{T}$$

From SVD to Reduced SVD

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}^T$$

girafe ai

From SVD to Reduced SVD

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times3} = U_{2\times2}\Sigma_{2\times3}(V_{3\times3})^T,$$

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}^T$$

From SVD to Reduced SVD

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times3} = U_{2\times2}\Sigma_{2\times3}(V_{3\times3})^T,$$

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}^T$$

$$A_{2\times3} = U_{2\times2}\Sigma_{2\times\mathbf{2}}(V_{3\times\mathbf{2}})^T,$$

girafe ai

From SVD to Reduced SVD

$$A = \begin{pmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{pmatrix}$$

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times 3}(V_{3\times 3})^T,$$

$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 & 0 \\ 0 & 3 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} & 2/3 \\ -1/\sqrt{2} & -1/3\sqrt{2} & -2/3 \\ 0 & 4/3\sqrt{2} & -1/3 \end{pmatrix}^T$$

Reduced SVD:

$$A_{2\times 3} = U_{2\times 2}\Sigma_{2\times \mathbf{2}}(V_{3\times \mathbf{2}})^T,$$

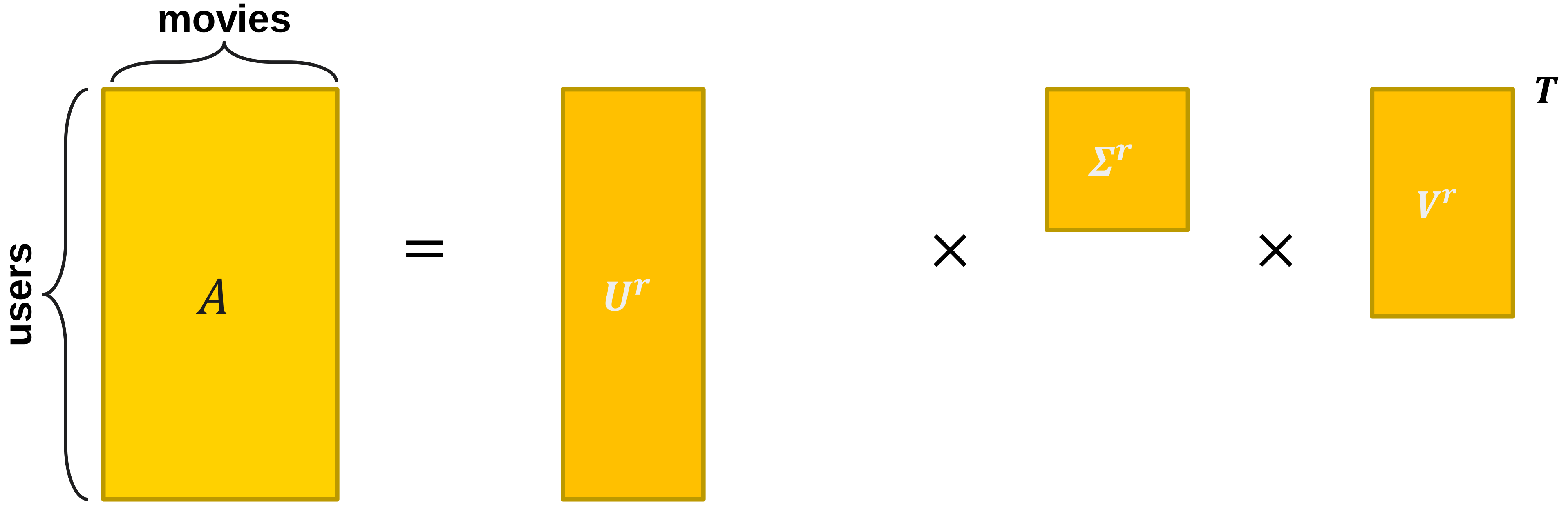
$$A = \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{pmatrix} \cdot \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1/\sqrt{2} & 1/3\sqrt{2} \\ -1/\sqrt{2} & -1/3\sqrt{2} \\ 0 & 4/3\sqrt{2} \end{pmatrix}^T$$

Reduced SVD: Main Idea

$$A_{m\times n} = U^r_{m\times \mathbf{r}}\Sigma^r_{\mathbf{r}\times \mathbf{r}}(V^r_{n\times \mathbf{r}})^T,$$ where

$U^r = [u_1 \,|\ldots|\, u_r],\ \ V^r = [v_1 \,|\ldots|\, v_r]$ – orthogonal matrices,

Σ^r – diagonal matrix with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$.



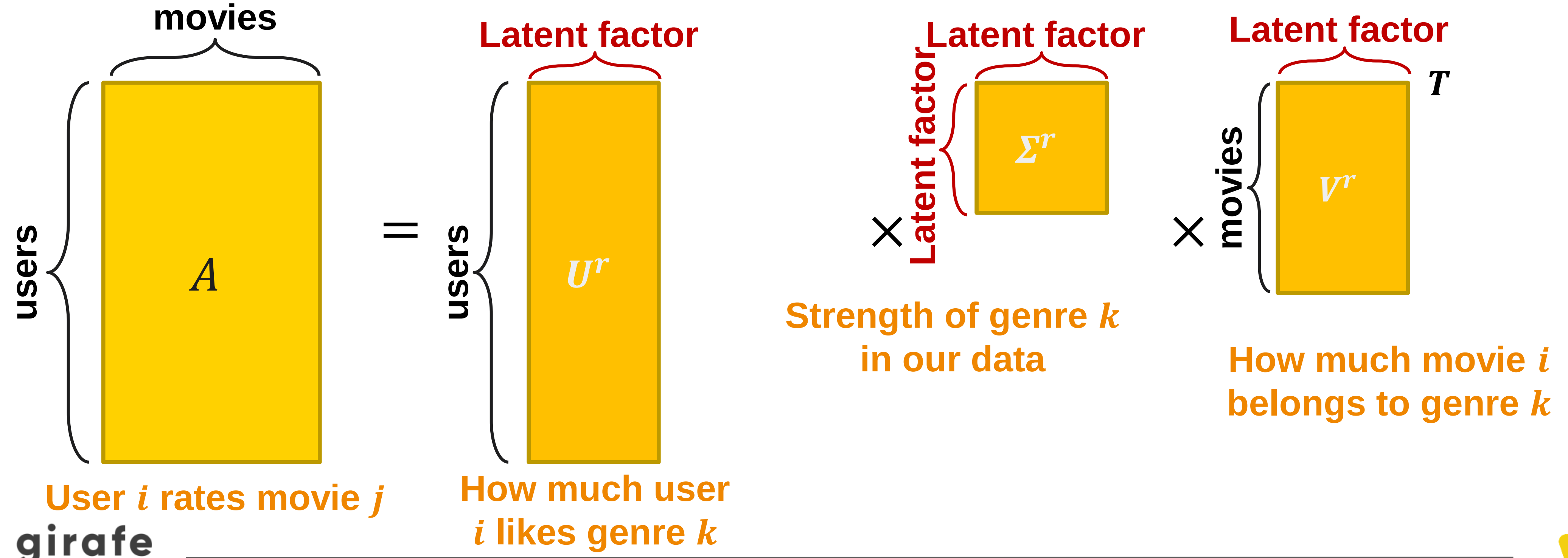
User i rates movie j

Reduced SVD: Main Idea

$$A_{m\times n} = U^r_{m\times r}\Sigma^r_{r\times r}(V^r_{n\times r})^T, \text{ where}$$

$$U^r = [u_1\,|\ldots|u_r],\ \ V^r = [v_1\,|\ldots|v_r] - \text{orthogonal matrices},$$

$$\Sigma^r - \text{diagonal matrix with } \sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0.$$



To sum up

- SVD: a generalization of eigendecomposition.
- Computing SVD: an example.
- Application: recommender systems.