

# Reduksi Dimensi

# Bab 30 Reduksi Dimensi SVD

# Setelah menyelesaikan bab ini Anda akan mengetahui:

- Reduksi dimensi merupakan proses mengurangi jumlah variabel input atau kolom di dalam proses membuat model
- SVD merupakan teknik aljabar linear untuk reduksi dimensi
- Bagaimana mengevaluasi model prediksi dengan input proyeksi SVD

# Singular Value Decomposition

# Singular Value Decomposition



0	9	0	0	0	4	0	0
0	0	6	0	0	0	1	0
0	0	0	5	0	0	1	0
0	0	0	0	0	0	3	0
0	0	6	0	0	0	0	0

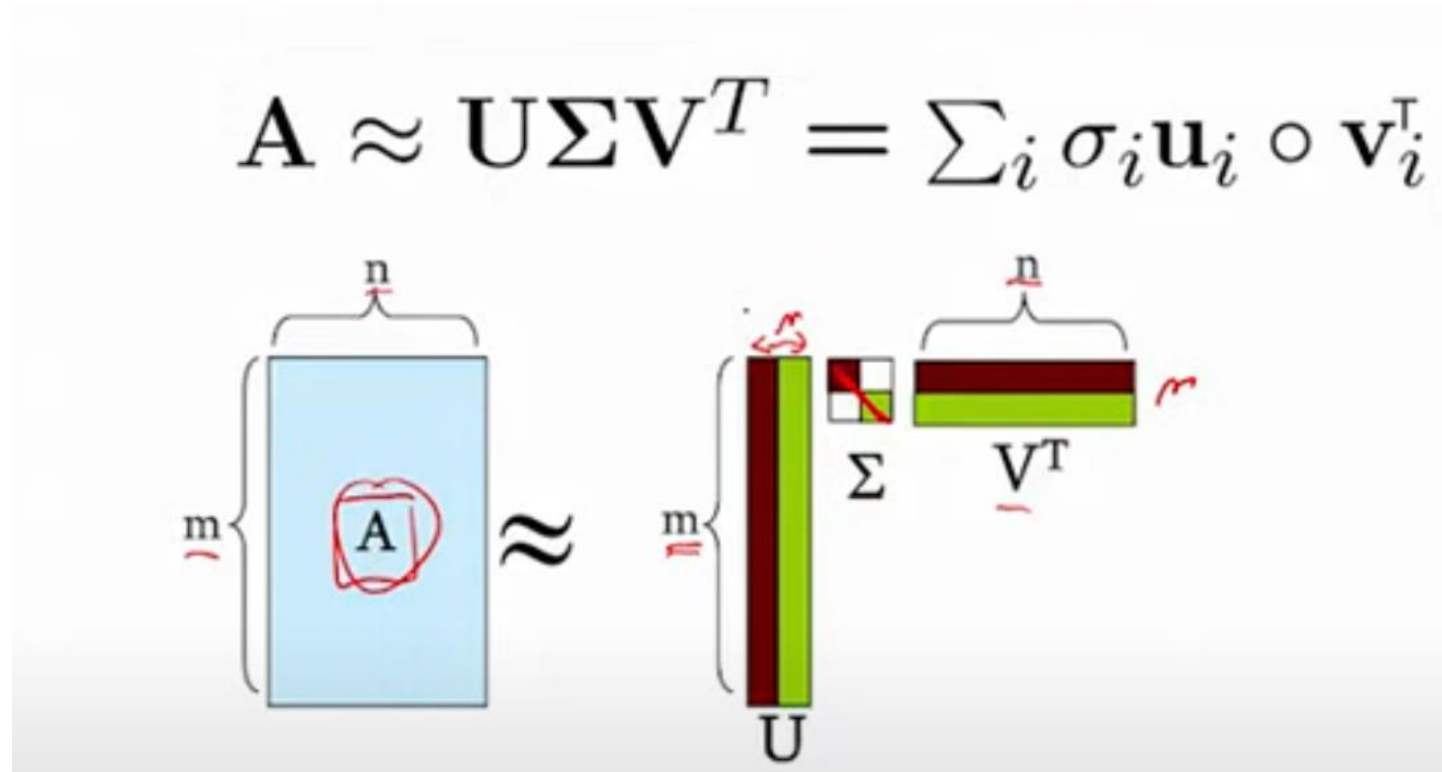
- SVD biasanya digunakan pada **sparse data**
- **Sparse data** adalah **jenis data yang memiliki banyak nilai kosong, misalnya recommender system** dimana user hanya merating beberapa barang atau **bag of words** dimana sebuah teks hanya memiliki beberapa kata saja
- Contoh sparse data lainnya adalah:
  - Recommender systems
  - Customer-product purchases
  - User-song listen counts
  - User-movie ratings
  - Text classification
  - One-hot encoding
  - Bag-of-words counts
- SVD memproyeksikan data dengan m-fitur menuju subspace dengan ruang lebih kecil namun masih menjaga esensi data asal

# Intuition: Singular Value Decomposition (SVD)

$$\mathbf{A}_{[m \times n]} = \mathbf{U}_{[m \times r]} \Sigma_{[r \times r]} (\mathbf{V}_{[n \times r]})^T$$

- **A: Input data matrix**
  - $m \times n$  matrix (e.g.,  $m$  documents,  $n$  terms)
- **U: Left singular vectors**
  - $m \times r$  matrix ( $m$  documents,  $r$  concepts)
- **$\Sigma$ : Singular values**
  - $r \times r$  diagonal matrix (strength of each 'concept')  
( $r$  : rank of the matrix **A**)
- **V: Right singular vectors**
  - $n \times \underline{r}$  matrix ( $n$  terms,  $r$  concepts)

# Intuition: Singular Value Decomposition (SVD)



# Intuition: Singular Value Decomposition (SVD)

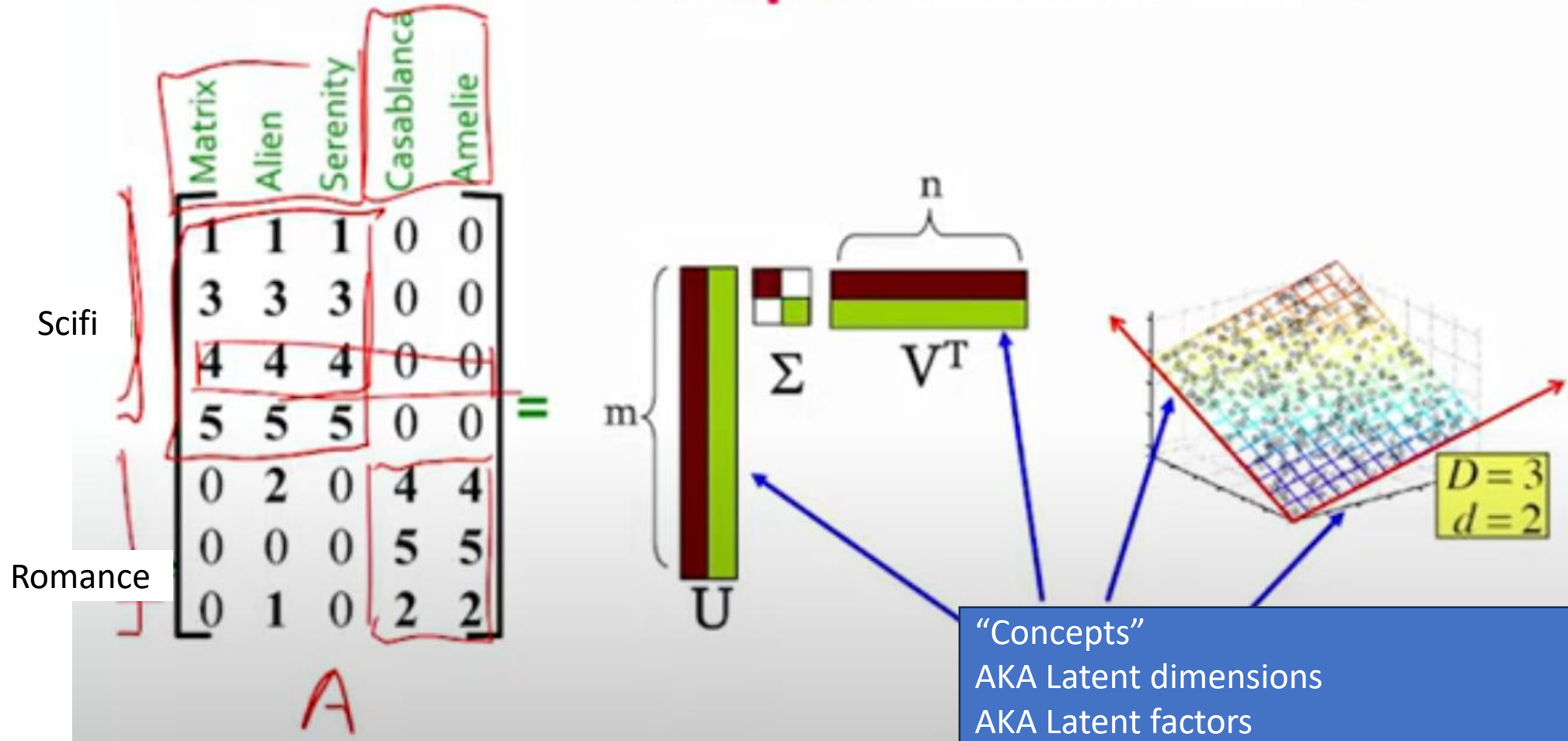
It is **always** possible to decompose a real matrix  $A$  into  $A = U \Sigma V^T$ , where

- $U, \Sigma, V$ : **unique**
- $U, V$ : **column orthonormal**
  - $U^T U = I; V^T V = I$  ( $I$ : identity matrix)
  - (Columns are orthogonal unit vectors)
- $\Sigma$ : **diagonal**
  - Entries (**singular values**) are **positive**, and sorted in decreasing order ( $\sigma_1 \geq \sigma_2 \geq \dots \geq 0$ )



# Intuition: Singular Value Decomposition (SVD)

■  $A = U \Sigma V^T$  - example: Users to Movies



# Intuition: Singular Value Decomposition (SVD)

■  $A = U \Sigma V^T$  - example: Users to Movies

		Matrix	Alien	Serenity	Casablanca	Amelie				
		1	1	1	0	0				
		3	3	3	0	0				
		4	4	4	0	0				
		5	5	5	0	0				
SciFi		0	2	0	4	4				
		0	0	0	5	5				
Romance		0	1	0	2	2				

$$=$$

		0.13	0.02	-0.01
		0.41	0.07	-0.03
		0.55	0.09	-0.04
		0.68	0.11	-0.05
		0.15	-0.59	0.65
		0.07	-0.73	-0.67
		0.07	-0.29	0.32

$$\times$$

		12.4	0	0
		0	9.5	0
		0	0	1.3

$$\times$$

		0.56	0.59	0.56	0.09	0.09
		0.12	-0.02	0.12	-0.69	-0.69
		0.40	-0.80	0.40	0.09	0.09

# Intuition: Singular Value Decomposition (SVD)

## ■ $A = U \Sigma V^T$ - example: Users to Movies

Matrix Alien Serenity Casablanca Amelie

SciFi

Romance

SciFi-concept

Romance-concept

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \times \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$

# Intuition: Singular Value Decomposition (SVD)

■  $A = U \Sigma V^T$  - example:

$U$  is "user-to-concept" similarity matrix

Matrix Alien Serenity Casablanca Amelie

SciFi

Romance

SciFi-concept Romance-concept

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \times \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$



# Intuition: Singular Value Decomposition (SVD)

■  $A = U \Sigma V^T$  - example:

Matrix    Alien    Serenity    Casablanca    Amelie

SciFi

Romance

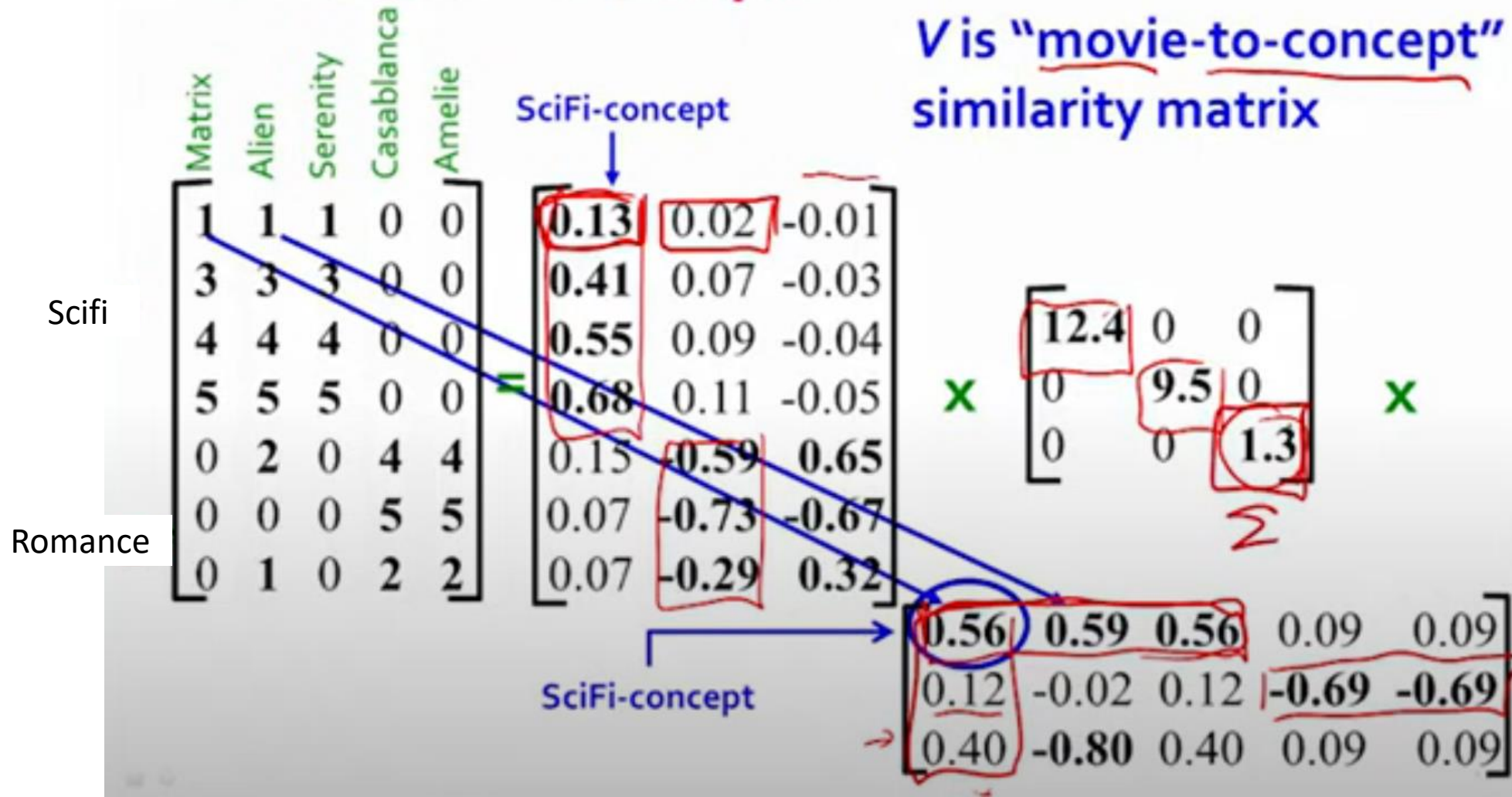
$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \times \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$

SciFi-concept

"strength" of the SciFi-concept

# Intuition: Singular Value Decomposition (SVD)

## ■ $A = U \Sigma V^T$ - example:



# Intuition: Singular Value Decomposition (SVD)

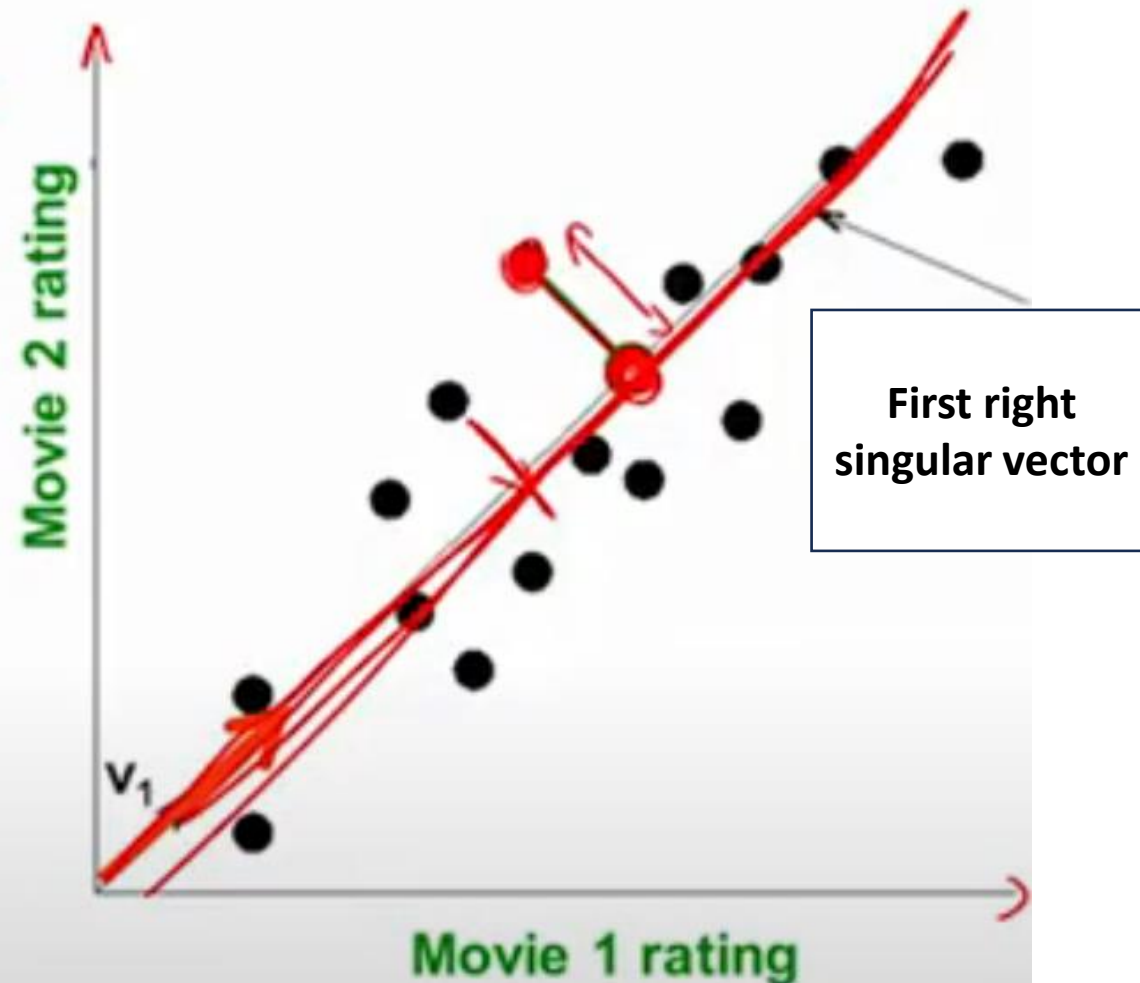
## SVD - Interpretation #1

‘**movies**’, ‘**users**’ and ‘**concepts**’:

- $U$ : user-to-concept similarity matrix
- $V$ : movie-to-concept similarity matrix
- $\Sigma$ : its diagonal elements:  
‘strength’ of each concept

# Intuition: Singular Value Decomposition (SVD)

- SVD gives 'best' axis to project on:
  - 'best' = min sum of squares of projection errors
- In other words, minimum reconstruction error



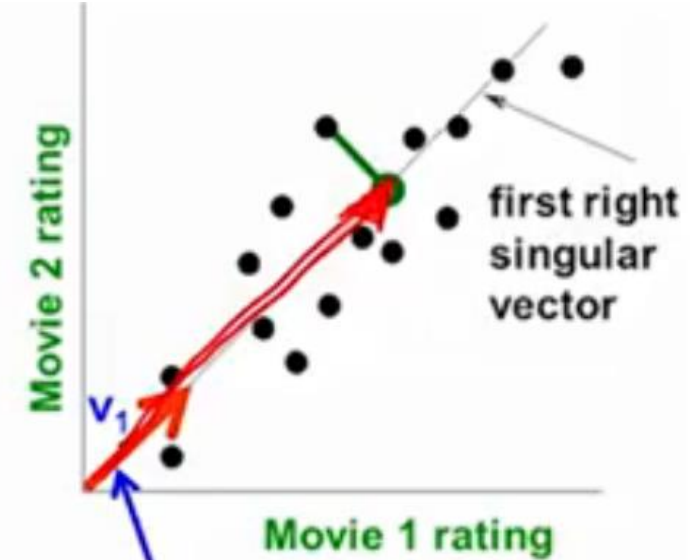


# Intuition: Singular Value Decomposition (SVD)

## ■ $A = U \Sigma V^T$ - example:

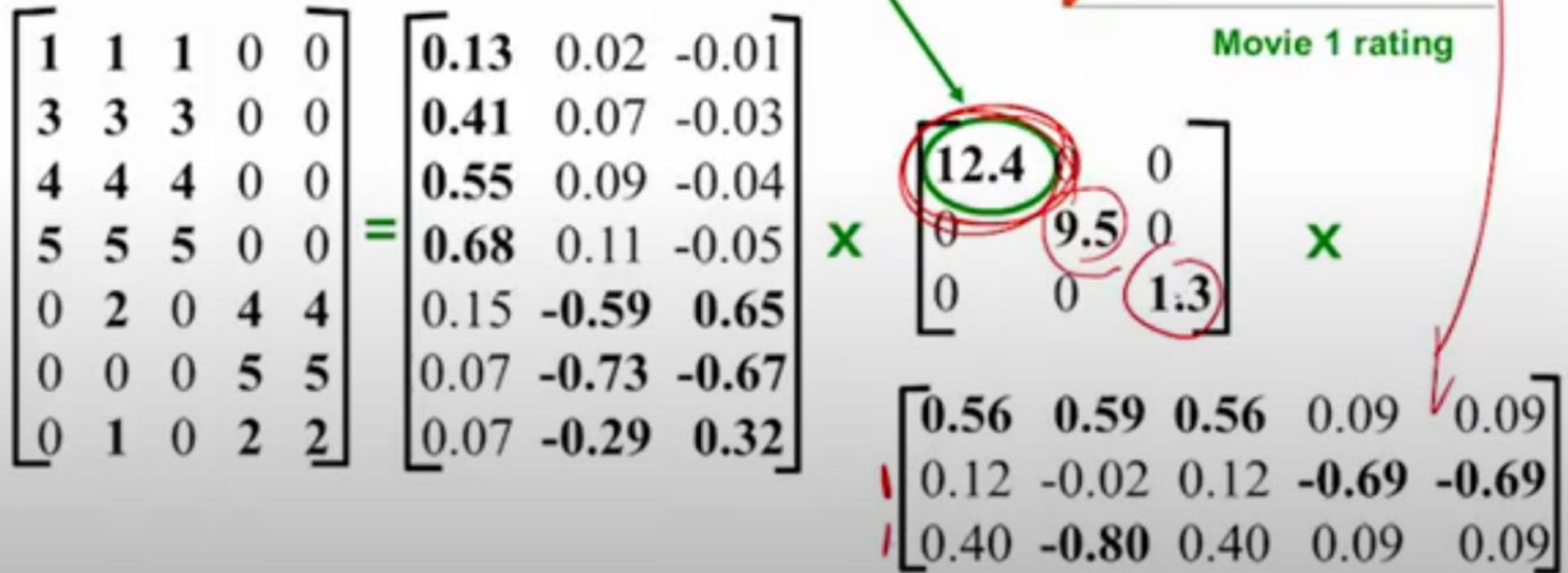
- $V$ : “movie-to-concept” matrix
- $U$ : “user-to-concept” matrix

$$\begin{array}{c} \textcolor{red}{U} \end{array}
 \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}
 =
 \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix}
 \begin{array}{c} \textcolor{green}{\Sigma} \end{array}
 \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix}
 \begin{array}{c} \textcolor{green}{V^T} \end{array}
 \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$



# Intuition: Singular Value Decomposition (SVD)

■  $A = U \Sigma V^T$  - example:



# Intuition: Singular Value Decomposition (SVD)

## More details

- **Q:** How exactly is dim. reduction done?

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \times \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$

# Intuition: Singular Value Decomposition (SVD)

## More details

- **Q:** How exactly is dim. reduction done?
- **A:** Set smallest singular values to zero

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} \approx \overset{U}{\begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix}} \times \overset{\Sigma}{\begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & \cancel{1.3} \end{bmatrix}} \times \overset{V^T}{\begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ \cancel{0.40} & \cancel{-0.80} & \cancel{0.40} & \cancel{0.09} & \cancel{0.09} \end{bmatrix}}$$



# Intuition: Singular Value Decomposition (SVD)

## More details

- **Q:** How exactly is dim. reduction done?
- **A:** Set smallest singular values to zero

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} \approx \begin{bmatrix} 0.13 & 0.02 \\ 0.41 & 0.07 \\ 0.55 & 0.09 \\ 0.68 & 0.11 \\ 0.15 & -0.59 \\ 0.07 & -0.73 \\ 0.07 & -0.29 \end{bmatrix} \times \begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix} \times \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \end{bmatrix}$$

# Intuition: Singular Value Decomposition (SVD)

## More details

- **Q:** How exactly is dim. reduction done?
- **A:** Set smallest singular values to zero

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} \approx \begin{bmatrix} 0.92 & 0.95 & 0.92 & 0.01 & 0.01 \\ 2.91 & 3.01 & 2.91 & -0.01 & -0.01 \\ 3.90 & 4.04 & 3.90 & 0.01 & 0.01 \\ 4.82 & 5.00 & 4.82 & 0.03 & 0.03 \\ 0.70 & 0.53 & 0.70 & 4.11 & 4.11 \\ -0.69 & 1.34 & -0.69 & 4.78 & 4.78 \\ 0.32 & 0.23 & 0.32 & 2.01 & 2.01 \end{bmatrix}$$

$A$   $B$

• **Frobenius norm:**

$$\|M\|_F = \sqrt{\sum_{ij} M_{ij}^2}$$

•  $\|A-B\|_F = \sqrt{\sum_{ij} (A_{ij}-B_{ij})^2}$   
is "small"

# Melihat Contoh SVD Di Jupyter Lab