

Why Reduce Dimensions?

Why reduce dimensions?

- **Discover hidden correlations/topics**
 - Words that occur commonly together
- **Remove redundant and noisy features**
 - Not all words are useful
- **Interpretation and visualization**
- **Easier storage and processing of the data**

SVD - Definition

$$A_{[m \times n]} = U_{[m \times r]} \Sigma_{[r \times r]} (V_{[n \times r]})^T$$

- **A: Input data matrix**

- m x n matrix (e.g., m documents, n terms)

- **U: Left singular vectors**

- m x r matrix (m documents, r concepts)

- **Σ : Singular values**

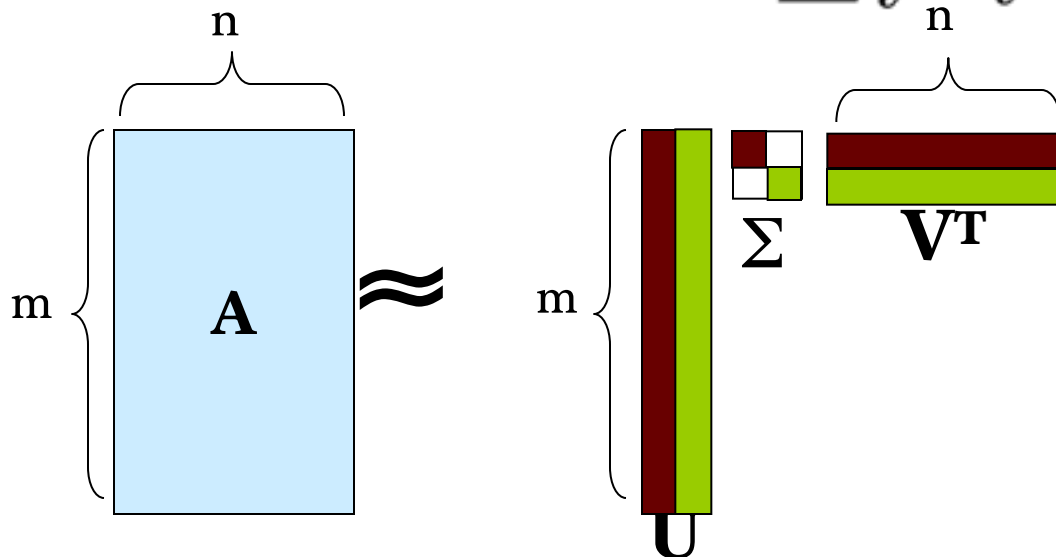
- r x r diagonal matrix (strength of each 'concept')
(r : rank of the matrix A)

- **V: Right singular vectors**

- n x r matrix (n terms, r concepts)

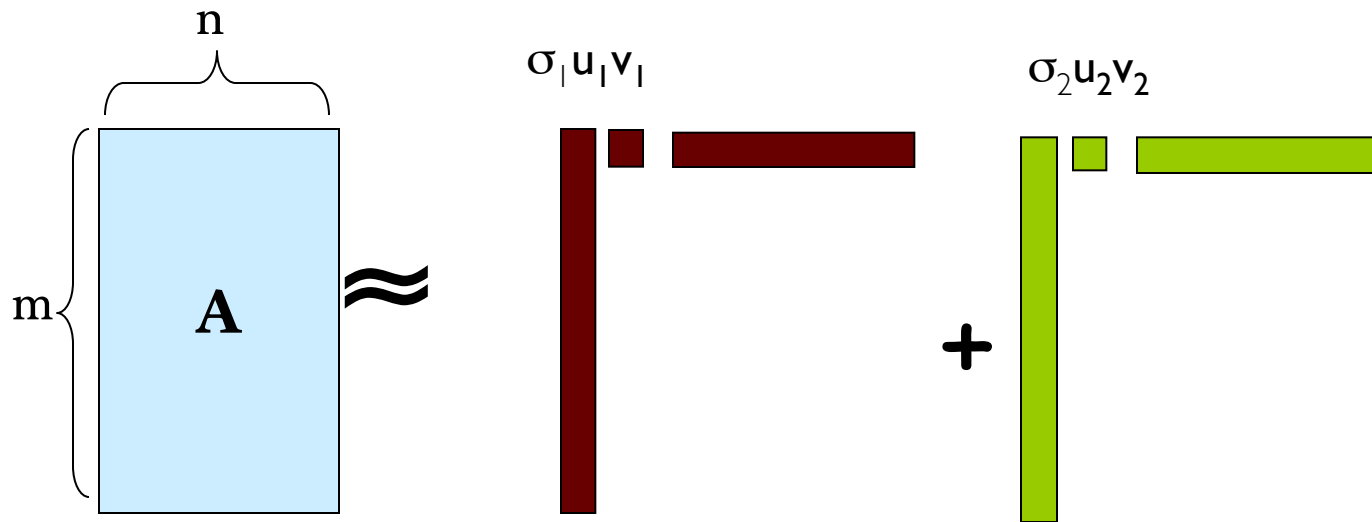
SVD

$$\mathbf{A} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_i \sigma_i \mathbf{u}_i \circ \mathbf{v}_i^T$$



SVD

$$\mathbf{A} \approx \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T = \sum_i \sigma_i \mathbf{u}_i \circ \mathbf{v}_i^T$$



σ_i ... scalar

\mathbf{u}_i ... vector

\mathbf{v}_i ... vector

SVD - Properties

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

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

Nice proof of uniqueness: <http://www.mpi-inf.mpg.de/~bast/ir-seminar-ws04/lecture2.pdf>

SVD – Example: Users-to-Movies

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

$$\begin{array}{c}
 \begin{array}{c} \uparrow \\ \text{SciFi} \\ \downarrow \\ \uparrow \\ \text{Romnce} \\ \downarrow \end{array}
 \begin{array}{c} \text{Matrix} \\ \text{Alien} \\ \text{Serenity} \\ \text{Casablanca} \\ \text{Amelie} \end{array}
 \begin{bmatrix}
 1 & 1 & 1 & 0 & 0 \\
 3 & 3 & 3 & 0 & 0 \\
 4 & 4 & 4 & 0 & 0 \\
 5 & 5 & 5 & 0 & 0 \\
 0 & 2 & 0 & 4 & 4 \\
 0 & 0 & 0 & 5 & 5 \\
 0 & 1 & 0 & 2 & 2
 \end{bmatrix}
 =
 \begin{bmatrix}
 0.13 & 0.02 & -0.01 \\
 0.41 & 0.07 & -0.03 \\
 0.55 & 0.09 & -0.04 \\
 0.68 & 0.11 & -0.05 \\
 0.15 & -0.59 & 0.65 \\
 0.07 & -0.73 & -0.67 \\
 0.07 & -0.29 & 0.32
 \end{bmatrix}
 \times
 \begin{bmatrix}
 12.4 & 0 & 0 \\
 0 & 9.5 & 0 \\
 0 & 0 & 1.3
 \end{bmatrix}
 \times
 \begin{bmatrix}
 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\
 0.40 & -0.80 & 0.40 & 0.09 & 0.09
 \end{bmatrix}
 \end{array}$$

SVD – Example: Users-to-Movies

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

Diagram illustrating the SVD decomposition of a user-movie rating matrix A into three matrices: U , Σ , and V^T .

Matrix A (User-Movie Ratings):

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

Matrix U (User-Concepts):

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

Matrix Σ (Singular Values):

12.4	0	0
0	9.5	0
0	0	1.3

Matrix V^T (Concept-Movie):

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

The decomposition is shown as:

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

SVD – Example: Users-to-Movies

■ $A = U \Sigma V^T$ - example:

U is “user-to-concept” similarity matrix

$$\begin{array}{c} \uparrow \\ \text{SciFi} \\ \downarrow \\ \uparrow \\ \text{Romnce} \\ \downarrow \end{array}
 \begin{array}{c} \text{Matrix} \\ \text{Alien} \\ \text{Serenity} \\ \text{Casablanca} \\ \text{Amelie} \end{array}
 \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}
 =
 \begin{array}{c} \text{SciFi-concept} \\ \text{Romance-concept} \end{array}
 \begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix}
 \times
 \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix}
 \times
 \begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$

SVD – Example: Users-to-Movies

■ $A = U \Sigma V^T$ - example:

Diagram illustrating the SVD decomposition of a user-movie rating matrix A into three matrices: U , Σ , and V^T .

Matrix A (User-Movie Ratings):

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

Matrix U (User-Concept):

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

Matrix Σ (Concept Strengths):

12.4	0	0
0	9.5	0
0	0	1.3

Matrix V^T (Concept-Movie):

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

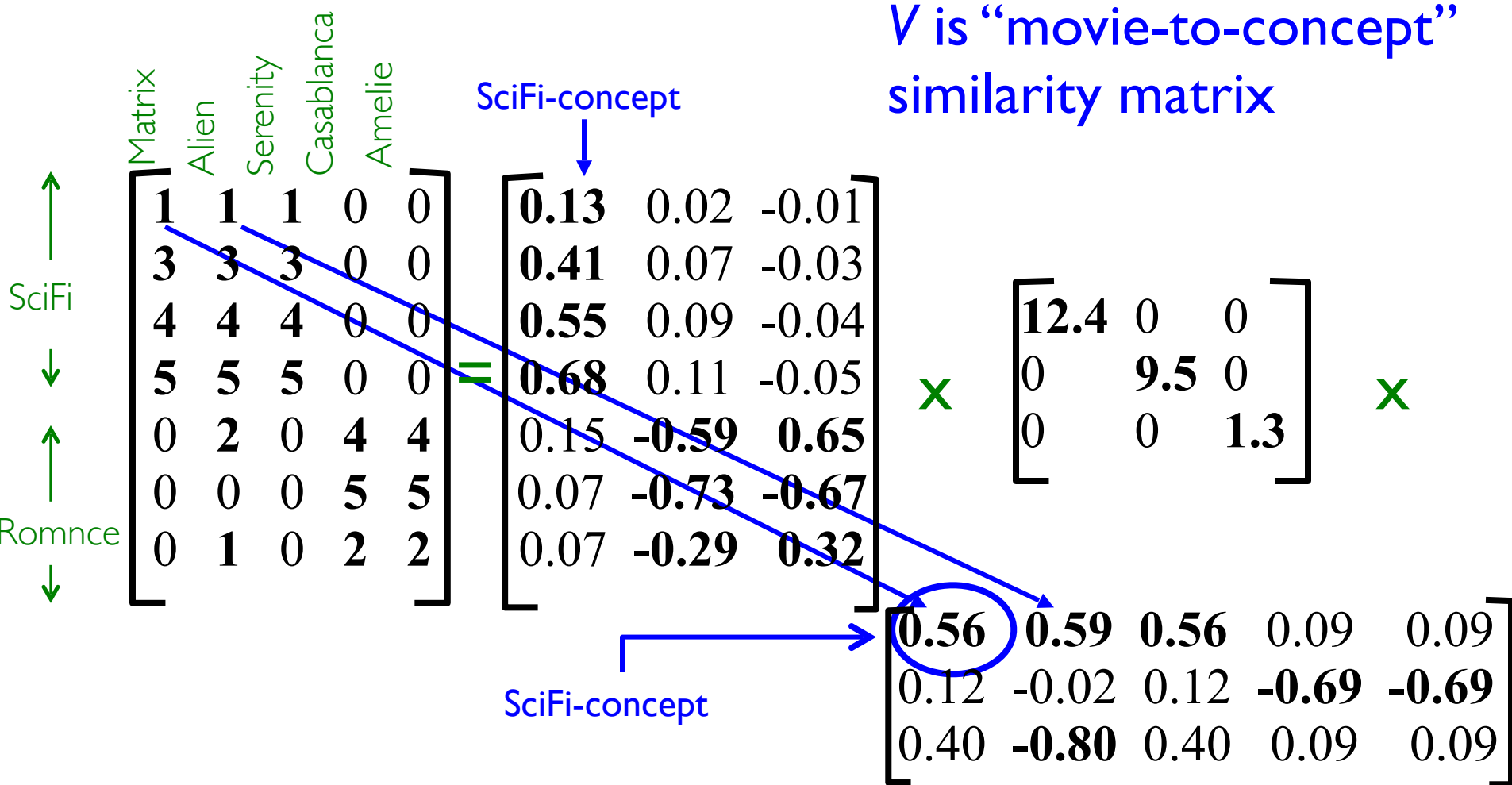
The equation is represented as: $A = U \Sigma V^T$.

Annotations:

- Green arrows on the left indicate the "SciFi" and "Romnce" (Romance) dimensions for the first matrix.
- A blue arrow points to the first column of the second matrix, labeled "SciFi-concept".
- A blue circle highlights the value 12.4 in the first row of the third matrix, with a label "strength" of the SciFi-concept.
- Green 'x' symbols indicate matrix multiplication between the second and third matrices, and between the result and the fourth matrix.

SVD – Example: Users-to-Movies

■ $A = U \Sigma V^T$ - example:



SVD - Interpretation #1

‘movies’, ‘users’ and ‘concepts’:

- **U: user-to-concept similarity matrix**
- **V: movie-to-concept similarity matrix**
- **Σ : its diagonal elements:
‘strength’ of each concept**