# 大数据计算及应用

## Dimensionality Reduction: SVD

Slides adapted from http://www.mmds.org

# Agenda

| High dim. data | Graph data | Infinite data | Machine learning | Apps |
|---|---|---|---|---|
| Locality sensitive hashing | PageRank, SimRank | Filtering data streams | SVM | Recommender systems |
| Clustering | Community Detection | Web advertising | Decision Trees | Association Rules |
| Dimensionality reduction | Spam Detection | Queries on streams | Perceptron, kNN | Duplicate document detection |

# Dimensionality Reduction



$D = 2$
$d = 1$

$D = 3$
$d = 2$

☐ **Assumption:** Data lies on or near a low *d*-dimensional subspace

☐ **Axes of this subspace are effective representation of the data**

# Dimensionality Reduction

☐ **Compress / reduce dimensionality:**

■ $10^6$ rows; $10^3$ columns; no updates

■ Random access to any cell(s); **small error: OK**

| day customer | Wc 7/10/96 | Th 7/11/96 | Fr 7/12/96 | Sa 7/13/96 | Su 7/14/96 |
|---|---|---|---|---|---|
| ABC Inc. | 1 | 1 | 1 | 0 | 0 |
| DEF Ltd. | 2 | 2 | 2 | 0 | 0 |
| GHI Inc. | 1 | 1 | 1 | 0 | 0 |
| KLM Co. | 5 | 5 | 5 | 0 | 0 |
| Smith | 0 | 0 | 0 | 2 | 2 |
| Johnson | 0 | 0 | 0 | 3 | 3 |
| Thompson | 0 | 0 | 0 | 1 | 1 |

**The above matrix is really "2-dimensional." All rows can be reconstructed by scaling [1 1 1 0 0] or [0 0 0 1 1]**

# Rank of a Matrix

☐ **Q:** What is **rank** of a matrix **A**?

☐ **A:** Number of **linearly independent** rows (columns) of **A**

☐ **For example:**

■ Matrix **A** = $\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix}$ has rank **r=2**

　☐ **Why?** The first two rows are linearly independent, so the rank is at least 2, but all three rows are linearly dependent (the first is equal to the sum of the second and third) so the rank must be less than 3.

☐ **Why do we care about low rank?**

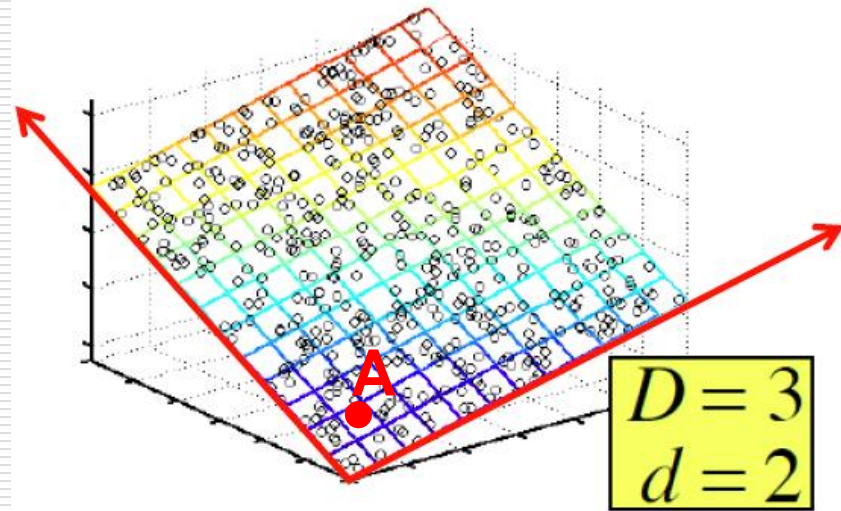■ We can write **A** as two "basis" vectors: [-2 -3 1] [3 5 0]

■ And new coordinates of : [1 1] [1 0] [0 1]

# Rank is "Dimensionality"

□ **Cloud of points 3D space:**

  ■ Think of point positions
    as a matrix:

  **1 row per point:** $\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix}$ **A**
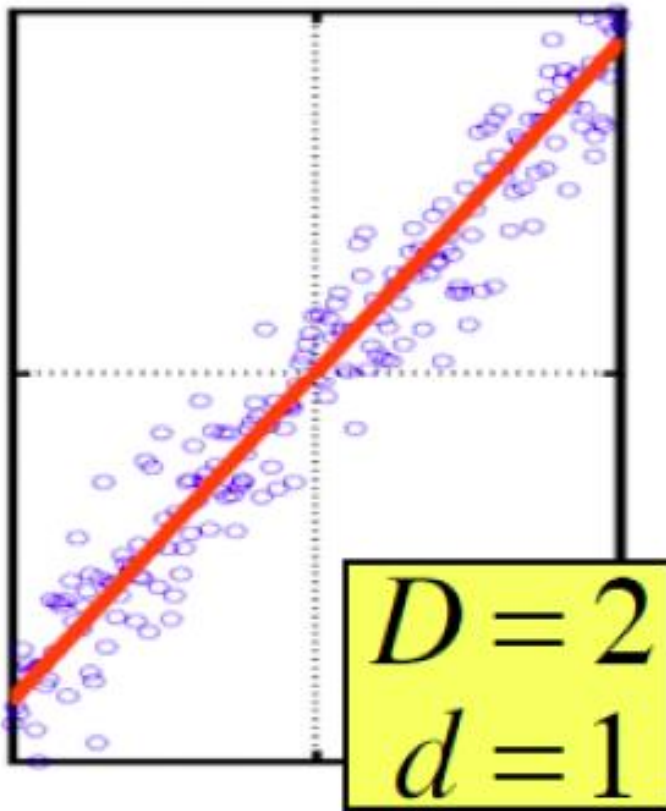  **B**
  **C**



$$D = 3$$
$$d = 2$$

□ **We can rewrite coordinates more efficiently!**

  ■ Old basis vectors: [1 0 0] [0 1 0] [0 0 1]

  ■ **New basis vectors: [-2 -3 1] [3 5 0]**

  ■ Then **A** has new coordinates: [1 1]. **B**: [1 0], **C**: [0 1]

    □ **Notice: We reduced the number of coordinates!**

# Dimensionality Reduction

☐ **Goal of dimensionality reduction is to discover the axis of data!**



$$D = 2$$
$$d = 1$$

Rather than representing every point with 2 coordinates we represent each point with 1 coordinate (corresponding to the position of the point on the red line).

By doing this we incur a bit of error as the points do not exactly lie on the line

# 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**



$D = 3$
$d = 2$
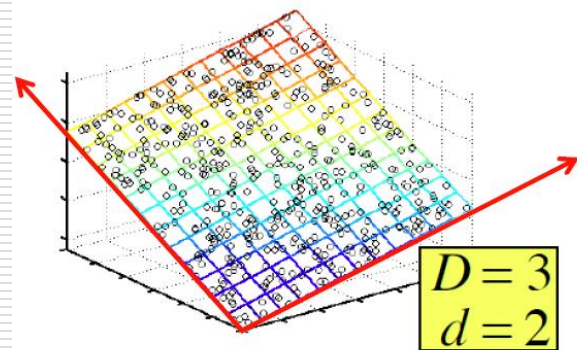
# Singular Value Decomposition (SVD)

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

- ☐ **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**
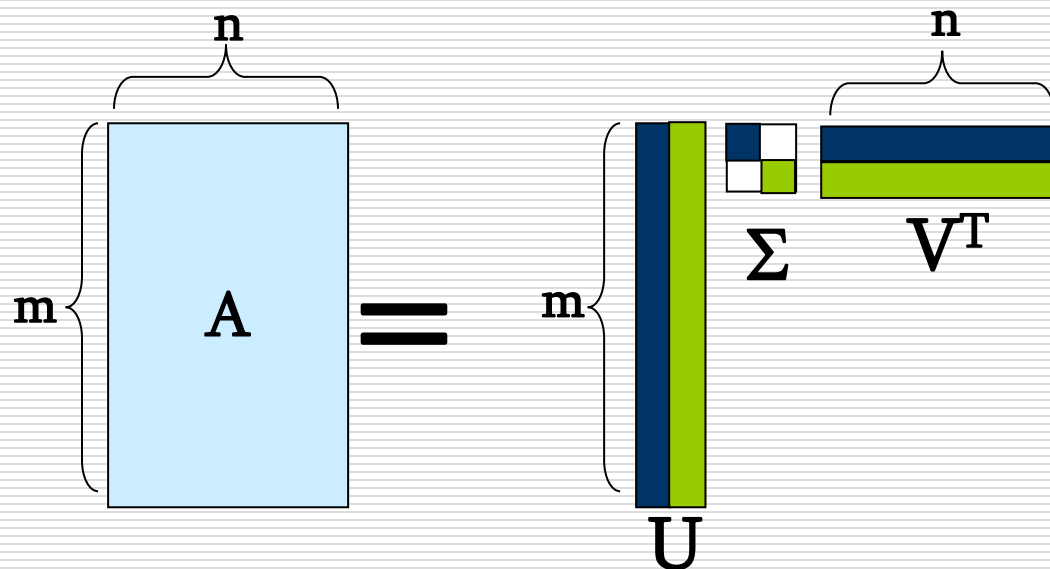  - ■ *r* x *r* diagonal matrix (strength of each 'concept') (*r* : rank of the matrix **A**)
- ☐ **V**: **Right singular vectors**
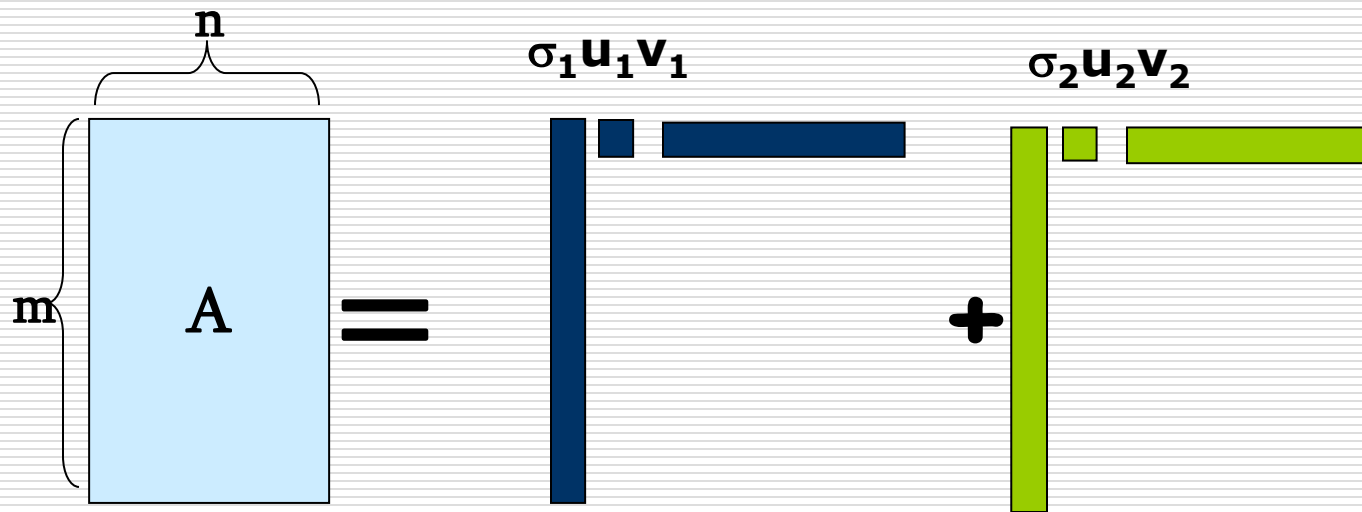  - ■ *n* x *r* matrix (*n* terms, *r* concepts)

# SVD

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

# SVD

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

$\sigma_i$ ... 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

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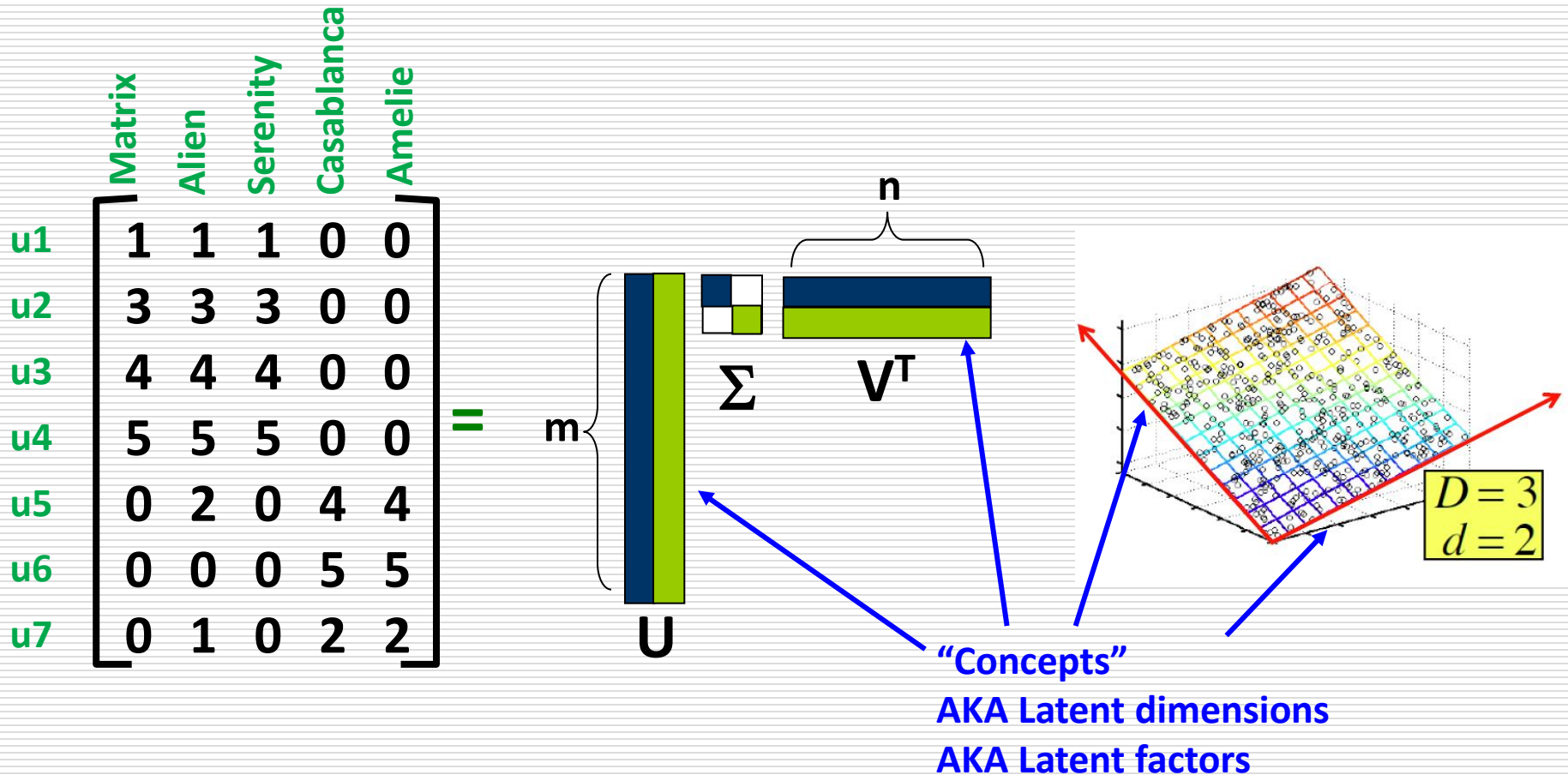  - (Columns are orthogonal unit vectors)
- $\Sigma$: diagonal
  - Entries (**singular values**) are positive, and sorted in decreasing order ($\sigma_1 \geq \sigma_2 \geq ... \geq 0$)

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

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



|  | Matrix | Alien | Serenity | Casablanca | Amelie |
|---|---|---|---|---|---|
| u1 | 1 | 1 | 1 | 0 | 0 |
| u2 | 3 | 3 | 3 | 0 | 0 |
| u3 | 4 | 4 | 4 | 0 | 0 |
| u4 | 5 | 5 | 5 | 0 | 0 |
| u5 | 0 | 2 | 0 | 4 | 4 |
| u6 | 0 | 0 | 0 | 5 | 5 |
| u7 | 0 | 1 | 0 | 2 | 2 |

= 

U   $\Sigma$   V$^T$

$D = 3$
$d = 2$

**"Concepts"**
**AKA Latent dimensions**
**AKA Latent factors**

# SVD – Example: Users-to-Movies

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

|  | Matrix | Alien | Serenity | Casablanca | Amelie |
|---|---|---|---|---|---|
| u1 | 1 | 1 | 1 | 0 | 0 |
| u2 | 3 | 3 | 3 | 0 | 0 |
| u3 | 4 | 4 | 4 | 0 | 0 |
| u4 | 5 | 5 | 5 | 0 | 0 |
| u5 | 0 | 2 | 0 | 4 | 4 |
| u6 | 0 | 0 | 0 | 5 | 5 |
| u7 | 0 | 1 | 0 | 2 | 2 |

=

$$\begin{bmatrix} 0.13 & 0.02 & -0.01 \\ 0.41 & 0.07 & -0.03 \\ 0.55 & 0.09 & -0.04 \\ 0.68 & 0.11 & -0.05 \\ 0.15 & -0.59 & 0.65 \\ 0.07 & -0.73 & -0.67 \\ 0.07 & -0.29 & 0.32 \end{bmatrix}$$

X

$$\begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix}$$

X

$$\begin{bmatrix} 0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\ 0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\ 0.40 & -0.80 & 0.40 & 0.09 & 0.09 \end{bmatrix}$$

# SVD – Example: Users-to-Movies

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

|  | Matrix | Alien | Serenity | Casablanca | Amelie |
|---|---|---|---|---|---|
| u1 | 1 | 1 | 1 | 0 | 0 |
| u2 | 3 | 3 | 3 | 0 | 0 |
| u3 | 4 | 4 | 4 | 0 | 0 |
| u4 | 5 | 5 | 5 | 0 | 0 |
| u5 | 0 | 2 | 0 | 4 | 4 |
| u6 | 0 | 0 | 0 | 5 | 5 |
| u7 | 0 | 1 | 0 | 2 | 2 |

**=**

SciFi-concept    Romance-concept

| | | |
|---|---|---|
| 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 |

**X**

| | | |
|---|---|---|
| 12.4 | 0 | 0 |
| 0 | 9.5 | 0 |
| 0 | 0 | 1.3 |

**X**

| | | | | |
|---|---|---|---|---|
| 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 |

15

# SVD – Example: Users-to-Movies

**A = U $\Sigma$ V$^T$ - example:**    *U* is "user-to-concept" similarity matrix

|        | Matrix | Alien | Serenity | Casablanca | Amelie |
|--------|--------|-------|----------|------------|--------|
| **u1** | 1      | 1     | 1        | 0          | 0      |
| **u2** | 3      | 3     | 3        | 0          | 0      |
| **u3** | 4      | 4     | 4        | 0          | 0      |
| **u4** | 5      | 5     | 5        | 0          | 0      |
| **u5** | 0      | 2     | 0        | 4          | 4      |
| **u6** | 0      | 0     | 0        | 5          | 5      |
| **u7** | 0      | 1     | 0        | 2          | 2      |

=

**SciFi-concept**   **Romance-concept**

| 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 |

X

| 12.4 | 0   | 0   |
|------|-----|-----|
| 0    | 9.5 | 0   |
| 0    | 0   | 1.3 |

X

| 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  |

# SVD – Example: Users-to-Movies

## $A = U \Sigma V^T$ - example:



|  | Matrix | Alien | Serenity | Casablanca | Amelie |
|---|---|---|---|---|---|
| u1 | 1 | 1 | 1 | 0 | 0 |
| u2 | 3 | 3 | 3 | 0 | 0 |
| u3 | 4 | 4 | 4 | 0 | 0 |
| u4 | 5 | 5 | 5 | 0 | 0 |
| u5 | 0 | 2 | 0 | 4 | 4 |
| u6 | 0 | 0 | 0 | 5 | 5 |
| u7 | 0 | 1 | 0 | 2 | 2 |

=

**SciFi-concept**

| 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 |

X

**"strength" of the SciFi-concept**

| 12.4 | 0 | 0 |
|---|---|---|
| 0 | 9.5 | 0 |
| 0 | 0 | 1.3 |

X

| 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 |

# SVD – Example: Users-to-Movies

## $A = U \Sigma V^T$ - example:

*V* is "movie-to-concept" similarity matrix

|  | Matrix | Alien | Serenity | Casablanca | Amelie |
|---|---|---|---|---|---|
| u1 | 1 | 1 | 1 | 0 | 0 |
| u2 | 3 | 3 | 3 | 0 | 0 |
| u3 | 4 | 4 | 4 | 0 | 0 |
| u4 | 5 | 5 | 5 | 0 | 0 |
| u5 | 0 | 2 | 0 | 4 | 4 |
| u6 | 0 | 0 | 0 | 5 | 5 |
| u7 | 0 | 1 | 0 | 2 | 2 |

**SciFi-concept**

$=$

| | | |
|---|---|---|
| 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 |

**X**

| | | |
|---|---|---|
| 12.4 | 0 | 0 |
| 0 | 9.5 | 0 |
| 0 | 0 | 1.3 |

**X**

**SciFi-concept**

| | | | | |
|---|---|---|---|---|
| 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 |

# 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

# Dimensionality Reduction with SVD
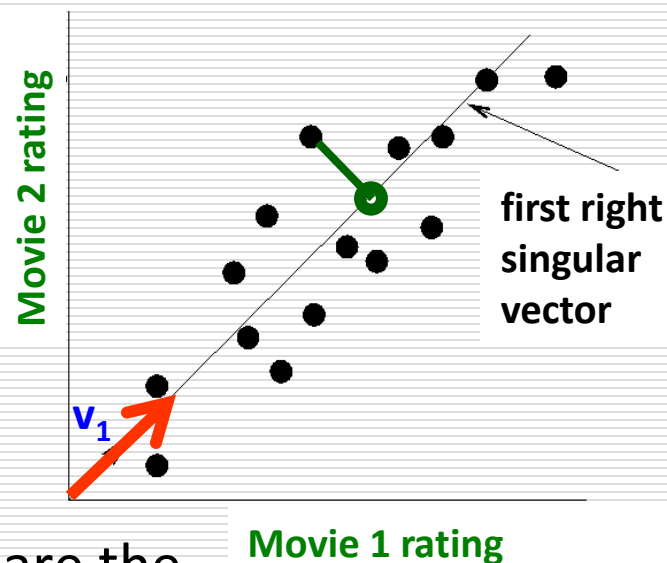
# SVD – Dimensionality Reduction



Movie 2 rating

first right
singular vector

$v_1$

Movie 1 rating
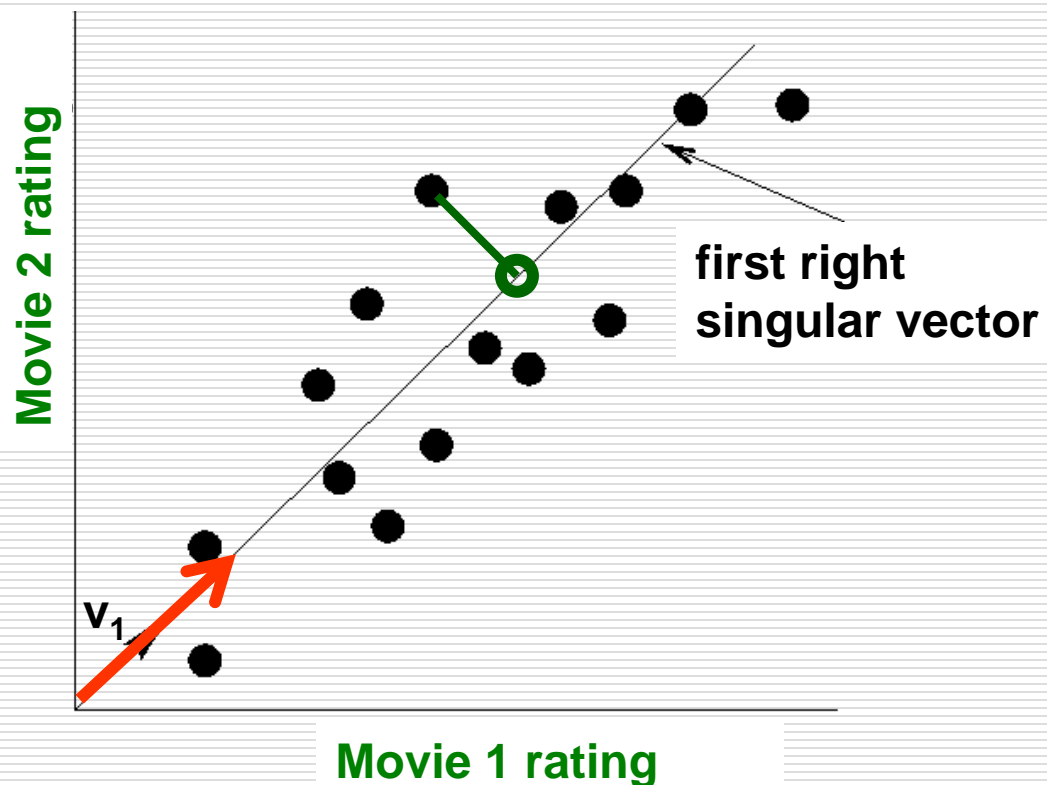
- ☐ Instead of using two coordinates $(x, y)$ to describe point locations, let's use only one coordinate $(z)$

- ☐ Point's position is its location along vector $v_1$

- ☐ **How to choose $v_1$?** **Minimize reconstruction error**

# SVD – Dimensionality Reduction

☐ **Goal: Minimize the sum of reconstruction errors:**

$$\sum_{i=1}^{N}\sum_{j=1}^{D}\left\|x_{ij}-z_{ij}\right\|^{2}$$



Movie 2 rating

Movie 1 rating

first right singular vector

$v_1$

☐ where $x_{ij}$ are the "old" and $z_{ij}$ are the "new" coordinates

☐ **SVD gives 'best' axis to project on:**

■ '**best**' = minimizing the reconstruction errors

☐ **In other words, minimum reconstruction error**

# SVD - Interpretation #2

☐ **A = U Σ V$^T$ - example:**

■ **V**: "movie-to-concept" matrix

■ **U**: "user-to-concept" matrix



Movie 2 rating — Movie 1 rating — **first right singular vector** — v$_1$
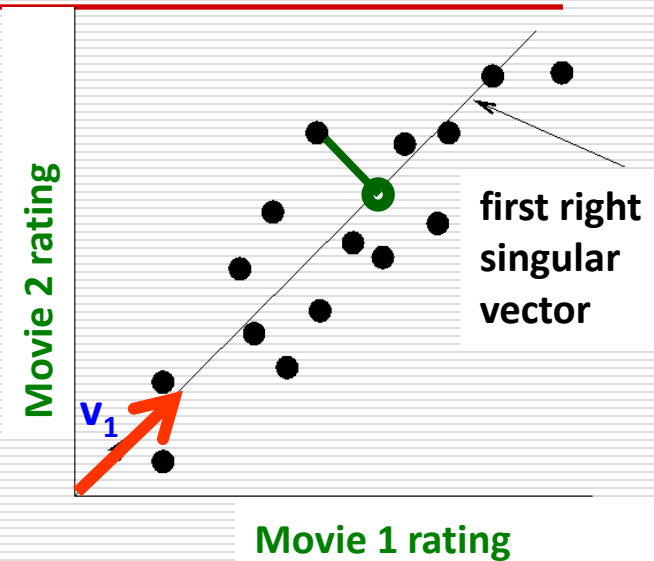
$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 2 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 1 & 0 & 2 & 2
\end{bmatrix}
=
\begin{bmatrix}
0.13 & 0.02 & -0.01 \\
0.41 & 0.07 & -0.03 \\
0.55 & 0.09 & -0.04 \\
0.68 & 0.11 & -0.05 \\
0.15 & -0.59 & 0.65 \\
0.07 & -0.73 & -0.67 \\
0.07 & -0.29 & 0.32
\end{bmatrix}
\times
\begin{bmatrix}
12.4 & 0 & 0 \\
0 & 9.5 & 0 \\
0 & 0 & 1.3
\end{bmatrix}
\times
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\
0.40 & -0.80 & 0.40 & 0.09 & 0.09
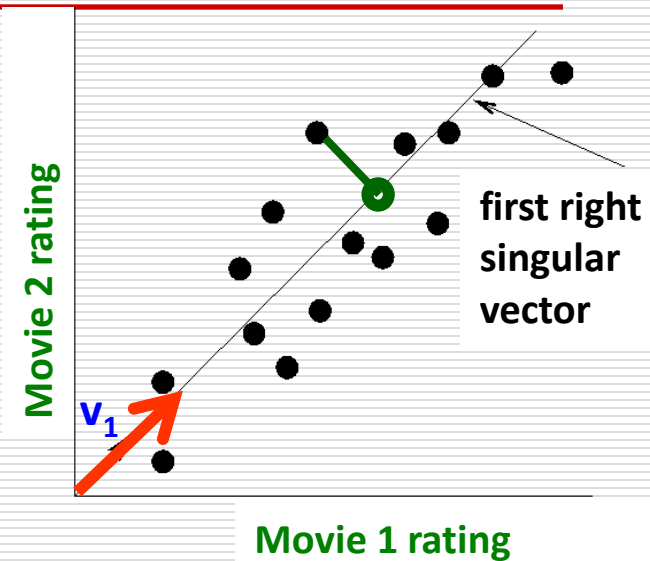\end{bmatrix}
$$

# SVD - Interpretation #2

□ **A = U Σ V^T - example:**



Movie 2 rating

first right singular vector

$v_1$

Movie 1 rating

variance ('spread')
on the $v_1$ axis

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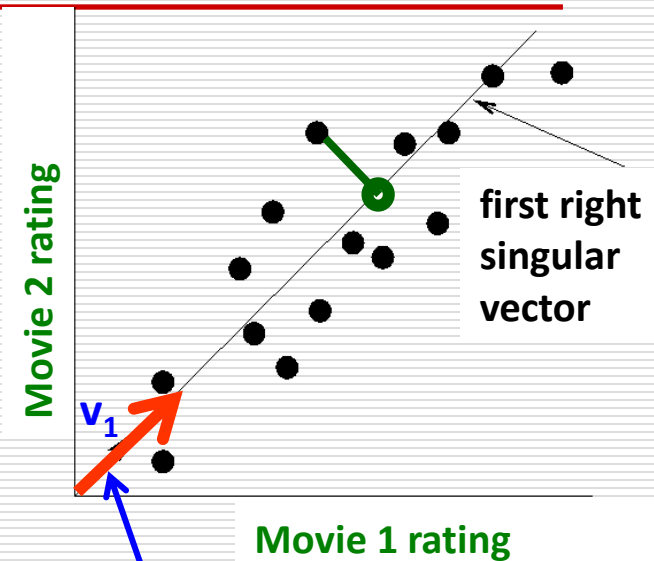
# SVD - Interpretation #2

## $A = U \Sigma V^T$ - example:

☐ **U $\Sigma$**: Gives the coordinates of the points in the projection axis



**Movie 2 rating**

**first right singular vector**

$v_1$

**Movie 1 rating**

$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 2 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 1 & 0 & 2 & 2
\end{bmatrix}
$$

**Projection of users on the "Sci-Fi" axis $(U \Sigma)^T$:**

$$
\begin{bmatrix}
1.61 & 0.19 & -0.01 \\
5.08 & 0.66 & -0.03 \\
6.82 & 0.85 & -0.05 \\
8.43 & 1.04 & -0.06 \\
1.86 & -5.60 & 0.84 \\
0.86 & -6.93 & -0.87 \\
0.86 & -2.75 & 0.41
\end{bmatrix}
$$

**More details**

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

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

# SVD - Interpretation #2

**More details**

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

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

$$
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\
0.40 & -0.80 & 0.40 & 0.09 & 0.09
\end{bmatrix}
$$

# SVD – Interpretation #2

**More details**

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

☐ **A: Set smallest singular values to zero**

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

$$
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\
0.40 & -0.80 & 0.40 & 0.09 & 0.09
\end{bmatrix}
$$

# SVD - Interpretation #2

**More details**

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

☐ **A: Set smallest singular values to zero**

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

$$
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
0.12 & -0.02 & 0.12 & -0.69 & -0.69 \\
0.40 & -0.80 & 0.40 & 0.09 & 0.09
\end{bmatrix}
$$

# SVD - Interpretation #2

**More details**

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

☐ **A: Set smallest singular values to zero**

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

$$
\begin{bmatrix}
0.56 & 0.59 & 0.56 & 0.09 & 0.09 \\
0.12 & -0.02 & 0.12 & -0.69 & -0.69
\end{bmatrix}
$$

# SVD – Interpretation #2

**More details**

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

☐ **A: Set smallest singular values to zero**

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

**Frobenius norm:**
$$\|M\|_F = \sqrt{\Sigma_{ij}\ M_{ij}^2}$$

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

# SVD – Best Low Rank Approx.

$$A = U \quad \text{Sigma} \quad V^T$$
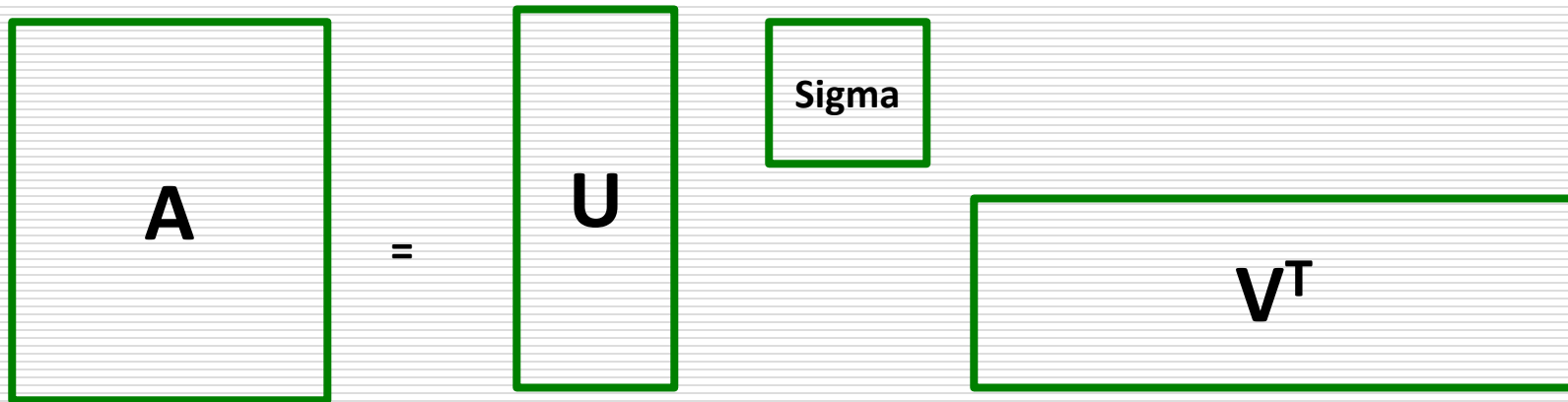
**B is best approximation of A**

$$B = U \quad \text{Sigma} \quad V^T$$
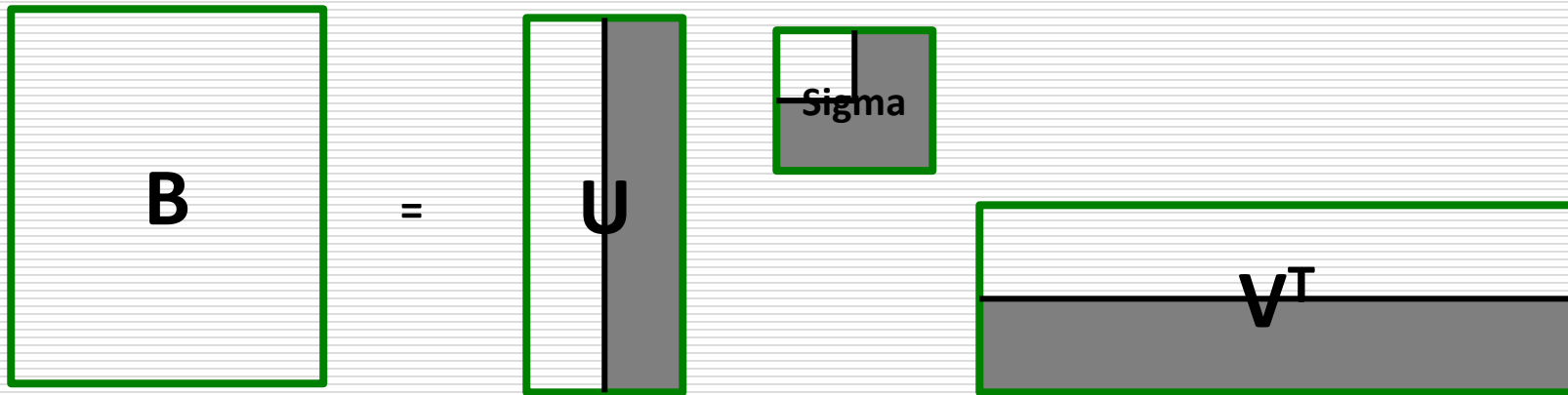
# SVD – Best Low Rank Approx.

☐ **Theorem:**
Let **A** = **U** $\sum$ **V**$^\mathsf{T}$ and **B** = **U S V**$^\mathsf{T}$ where
**S** = **diagonal *r*x*r* matrix** with $s_i = \sigma_i$ (*i=1…k*) else $s_i = 0$
then **B** is a **best** rank(**B**)=*k* approx. to ***A***

What do we mean by "best":

■ ***B* is a solution to min$_B$ ‖*A-B*‖$_F$ where rank(*B*)=*k***

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & \\ \vdots & \vdots & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix}_{m \times n} = \begin{pmatrix} u_{11} & \dots & \\ \vdots & \ddots & \\ u_{m1} & & \end{pmatrix}_{m \times r}^{U} \begin{pmatrix} \sigma_{11} & 0 & \dots \\ 0 & \ddots & \\ \vdots & & \end{pmatrix}_{r \times r}^{\Sigma} \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \\ & & \end{pmatrix}_{r \times n}^{V^\mathsf{T}}$$

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

# SVD – Best Low Rank Approx.

☐ **Theorem:** Let **A = U Σ V$^T$**  (σ$_1 \geq$σ$_2 \geq$…, rank(**A**)=**r**)

  **then B = U S V$^T$**

   ■ **S = diagonal *r*x*r* matrix** where $s_i = \sigma_i$ (*i=1…k*) else $s_i$=0

  is a best rank-**k** approximation to **A**:

   ■ ***B* is a solution to min$_B$ ||*A-B*||$_F$  where rank(*B*)=*k***

$$\begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & \\ \vdots & \vdots & \ddots & \\ x_{m1} & & & x_{mn} \end{pmatrix}_{m \times n} = \begin{pmatrix} u_{11} & \dots & \\ \vdots & \ddots & \\ u_{m1} & & \end{pmatrix}_{m \times r}^{U} \begin{pmatrix} \sigma_{11} & 0 & \dots \\ 0 & & \\ \vdots & & \end{pmatrix}_{r \times r}^{\Sigma} \begin{pmatrix} v_{11} & \dots & v_{1n} \\ \vdots & \ddots & \\ & & \end{pmatrix}_{r \times n}^{V^T}$$

☐ **We will need:**

   ■ **U Σ V$^T$ - U S V$^T$ = U (Σ - S) V$^T$**

# SVD – Best Low Rank Approx.

☐ **A** = **U** $\Sigma$ **V**$^\top$ , **B** = **U S V**$^\top$  ($\sigma_1 \geq \sigma_2 \geq \ldots \geq 0$, rank(**A**)=**r**)

■ **S** = diagonal $n_x n$ matrix where $s_i = \sigma_i$ (*i=1…k*) else $s_i = 0$

**then B** is solution to $\min_B \|A\text{-}B\|_F$ , rank(**B**)=**k**

■ **Why?**

$$\min_{B, rank(B)=k} \|A - B\|_F = \min \|\Sigma - S\|_F = \min_{s_i} \sum_{i=1}^{r} (\sigma_i - s_i)^2$$

**We used: U** $\Sigma$ **V$^\top$ - U S V$^\top$ = U ($\Sigma$ - S) V$^\top$**

☐ We want to choose $s_i$ to minimize $\sum_i (\sigma_i - s_i)^2$

☐ Solution is to set $s_i = \sigma_i$ (*i=1…k*) and other $s_i = 0$

$$= \min_{s_i} \sum_{i=1}^{k} (\sigma_i - s_i)^2 + \sum_{i=k+1}^{r} \sigma_i^2 = \sum_{i=k+1}^{r} \sigma_i^2$$

# SVD - Interpretation #2

**Equivalent:**

**'spectral decomposition' of the matrix:**

$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
3 & 3 & 3 & 0 & 0 \\
4 & 4 & 4 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 2 & 0 & 4 & 4 \\
0 & 0 & 0 & 5 & 5 \\
0 & 1 & 0 & 2 & 2
\end{bmatrix}
=
\begin{bmatrix}
u_1 & u_2
\end{bmatrix}
\times
\begin{bmatrix}
\sigma_1 & \varnothing \\
\varnothing & \sigma_2
\end{bmatrix}
\times
\begin{bmatrix}
v_1 \\
v_2
\end{bmatrix}
$$

# SVD - Interpretation #2

**Equivalent:**

**'spectral decomposition' of the matrix**

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix} = \sigma_1 \, u_1 \, v^T_1 \; + \; \sigma_2 \, u_2 \, v^T_2 \; + \ldots$$

m

n

k terms

n x 1    1 x m

Assume: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \ldots \geq 0$

**Why is setting small $\sigma_i$ to 0 the right thing to do? Vectors $u_i$ and $v_i$ are unit length, so $\sigma_i$ scales them.**
**So, zeroing small $\sigma_i$ introduces less error.**

# SVD - Interpretation #2

**Q: How many σs to keep?**

**A:** Rule-of-a thumb:

keep 80-90% of 'energy' $= \sum_i \sigma_i^2$

$$\overset{\xleftarrow{\hspace{1em}} m \xrightarrow{\hspace{1em}}}{\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 3 & 3 & 3 & 0 & 0 \\ 4 & 4 & 4 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 2 & 0 & 4 & 4 \\ 0 & 0 & 0 & 5 & 5 \\ 0 & 1 & 0 & 2 & 2 \end{bmatrix}} \; = \; \sigma_1 \quad u_1 \quad v^T_1 \; + \; \sigma_2 \quad u_2 \quad v^T_2 \; + ...$$

n

Assume: $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq ...$

# Relation to Eigen-decomposition

□ **SVD gives us:**

  ■ $A = U \, \Sigma \, V^T$

□ **Eigen-decomposition:**

  ■ $M = X \, L \, X^T$

    □ M is symmetric

    □ U, V, X are orthonormal ($\mathbf{U^T U = I}$),

    □ **L,** $\Sigma$ are diagonal

□ **Now let's calculate:**

  ■ $\mathbf{A A^T} =$

  ■ $\mathbf{A^T A} = \mathbf{V} \, \Sigma^T \, \mathbf{U^T} \, (\mathbf{U} \Sigma \, \mathbf{V^T}) = \mathbf{V} \, \Sigma \Sigma^T \, \mathbf{V^T}$

# Relation to Eigen-decomposition

☐ **SVD gives us:**

■ $A = U \, \Sigma \, V^T$

☐ **Eigen-decomposition:**

■ $M = X \, L \, X^T$

☐ M is symmetric

☐ U, V, X are orthonormal ($\mathbf{U^T U = I}$),

☐ **L,** $\Sigma$ are diagonal

☐ **Now let's calculate:**

■ $\mathbf{AA^T = U\Sigma \, V^T(U\Sigma \, V^T)^T = U\Sigma \, V^T(V\Sigma^T U^T) = U\Sigma\Sigma^T \, U^T}$

$$X \quad L \quad X^T$$

■ $\mathbf{A^T A = V \, \Sigma^T \, U^T \, (U\Sigma \, V^T) = V \, \Sigma\Sigma^T \, V^T}$

$$X \quad L \quad X^T \qquad \textit{where } L = \Sigma^2$$

**Shows how to compute SVD using eigenvalue decomposition!**

$$X \quad L \quad X^T$$

# An Example of SVD Computation

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

**A = U $\Sigma$ V**

**A A$^T$ = X L X$^T$ = U $\Sigma^2$ U$^T$**

**A$^T$ A = Y L Y$^T$ = V $\Sigma^2$ V$^T$**

$$A\,A^T = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

Set AA$^T$= M,  let λ be the eigenvalue of M, then we have

**M X= λ X $\Rightarrow$  (λ I − M) X =0 $\Rightarrow$ | λ I − M| = 0**

**I: identity matrix**

$$\begin{vmatrix} \lambda - 2 & 1 & -1 \\ 1 & \lambda - 1 & 0 \\ -1 & 0 & \lambda - 1 \end{vmatrix} = \lambda(\lambda - 1)(\lambda - 3) = 0$$

$\lambda_1=3, \lambda_2=1, \lambda_3=0$

# An Example of SVD Computation

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad A\,A^T = \begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \qquad \mathbf{A\,A^T = X\,L\,X^T = U\,\Sigma^2\,U^T}$$

$\lambda_1 = 3, \ \lambda_2 = 1, \ \lambda_3 = 0$

When $\lambda_1 = 3$, we have $\mathbf{M\,X = \lambda\,X}$

$$\begin{bmatrix} 2 & -1 & 1 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 3 \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$$

$$x_1{}^2 + x_2{}^2 + x_3{}^2 = 1$$

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \dfrac{2}{\sqrt{6}} \\ \dfrac{-1}{\sqrt{6}} \\ \dfrac{1}{\sqrt{6}} \end{bmatrix} = U_1$$

$$\lambda_1 = 3, \ U_1 = \begin{bmatrix} \dfrac{2}{\sqrt{6}} \\ \dfrac{-1}{\sqrt{6}} \\ \dfrac{1}{\sqrt{6}} \end{bmatrix} \qquad \lambda_2 = 1, \ U_2 = \begin{bmatrix} 0 \\ \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix} \qquad \lambda_3 = 0, \ U_3 = \begin{bmatrix} \dfrac{1}{\sqrt{3}} \\ \dfrac{1}{\sqrt{3}} \\ \dfrac{-1}{\sqrt{3}} \end{bmatrix}$$

# An Example of SVD Computation

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$A = U \, \Sigma \, V$$
$$A \, A^T = X \, L \, X^T = U \, \Sigma^2 \, U^T$$
$$A^T A = Y \, L \, Y^T = V \, \Sigma^2 \, V^T$$

$$A^T A = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \qquad \lambda_1 = 3, \ \lambda_2 = 1$$

$$\lambda_1 = 3, \ V_1 = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{-1}{\sqrt{2}} \end{bmatrix} \qquad \lambda_2 = 1, \ V_2 = \begin{bmatrix} \dfrac{1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

# An Example of SVD Computation

$$A = \begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$$

**A = U $\Sigma$ V**

**A A$^T$ = X L X$^T$ = U $\Sigma^2$ U$^T$**

**A$^T$ A = Y L Y$^T$ = V $\Sigma^2$ V$^T$**

**A = U $\Sigma$ V**

$$\begin{bmatrix} 1 & -1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \dfrac{2}{\sqrt{6}} & 0 & \\ \dfrac{-1}{\sqrt{6}} & \dfrac{1}{\sqrt{2}} & \\ \dfrac{1}{\sqrt{6}} & \dfrac{1}{\sqrt{2}} & \end{bmatrix} \begin{bmatrix} \sqrt{3} & 0 \\ 0 & 1 \\ & \end{bmatrix} \begin{bmatrix} \dfrac{1}{\sqrt{2}} & \dfrac{-1}{\sqrt{2}} \\ \dfrac{1}{\sqrt{2}} & \dfrac{1}{\sqrt{2}} \end{bmatrix}$$

# SVD - Complexity

- ☐ **To compute SVD:**
  - ■ **O(nm²)** or **O(n²m)** (whichever is less)
- ☐ **But:**
  - ■ Less work, if we just want singular values
  - ■ or if we want first *k* singular vectors
  - ■ or if the matrix is sparse

- ☐ **Implemented in** linear algebra packages like
  - ■ LINPACK, Matlab, SPlus, Mathematica ...

# SVD - Conclusions so far

□ **SVD: A= U $\Sigma$ V$^T$: unique**

  ■ **U**: user-to-concept similarities

  ■ **V**: movie-to-concept similarities

  ■ $\Sigma$ : strength of each concept

□ **Dimensionality reduction:**

  ■ keep the few largest singular values (80-90% of 'energy')

  ■ SVD: picks up linear correlations

# Example of SVD & Conclusion

# Case study: How to query?

☐ **Q: Find users that like 'Matrix'**
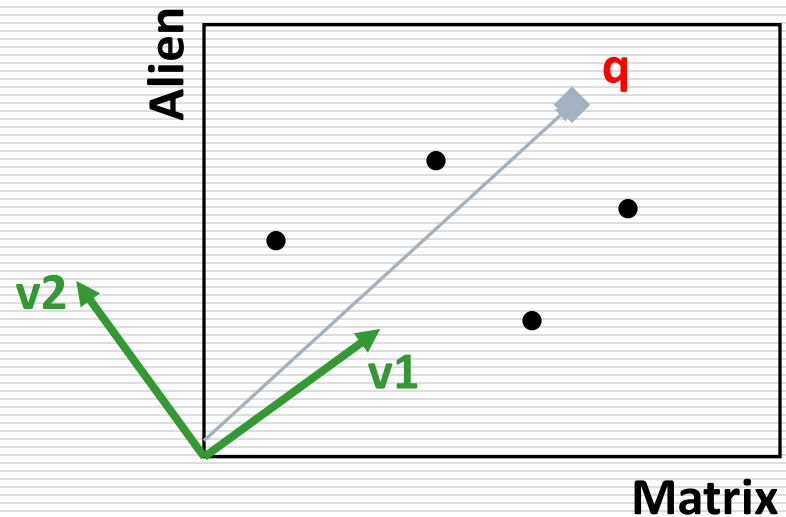
☐ **A: Map query into a 'concept space' – how?**

|    | Matrix | Alien | Serenity | Casablanca | Amelie |
|----|--------|-------|----------|------------|--------|
| u1 | 1      | 1     | 1        | 0          | 0      |
| u2 | 3      | 3     | 3        | 0          | 0      |
| u3 | 4      | 4     | 4        | 0          | 0      |
| u4 | 5      | 5     | 5        | 0          | 0      |
| u5 | 0      | 2     | 0        | 4          | 4      |
| u6 | 0      | 0     | 0        | 5          | 5      |
| u7 | 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 |

**X**

| 12.4 | 0   | 0   |
|------|-----|-----|
| 0    | 9.5 | 0   |
| 0    | 0   | 1.3 |

**X**

| 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 |

# Case study: How to query?

☐ **Q: Find users that like 'Matrix'**

☐ **A: Map query into a 'concept space' – how?**

$$q = \begin{bmatrix} \overset{\text{Matrix}}{5} & \overset{\text{Alien}}{0} & \overset{\text{Serenity}}{0} & \overset{\text{Casablanca}}{0} & \overset{\text{Amelie}}{0} \end{bmatrix}$$

**Project into concept space:**
**Inner product with each**
**'concept' vector $v_i$**

# Case study: How to query?

☐ **Q: Find users that like 'Matrix'**

☐ **A: Map query into a 'concept space' – how?**

$q$ =
$$\begin{bmatrix} 5 & 0 & 0 & 0 & 0 \end{bmatrix}$$

columns: **Matrix**, **Alien**, **Serenity**, **Casablanca**, **Amelie**

**Project into concept space:**
**Inner product with each**
**'concept' vector $v_i$**
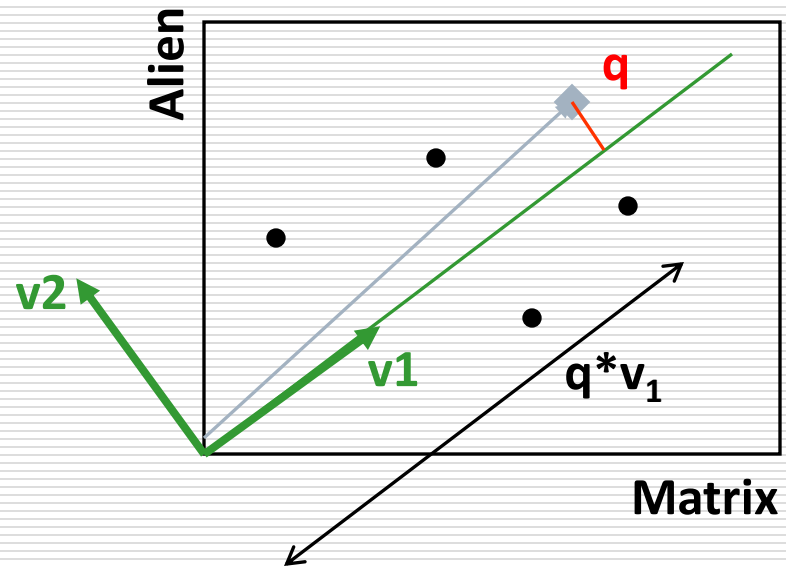
# Case study: How to query?

**Compactly, we have:**

$q_{concept} = q \, V$

**E.g.:**

$q = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \end{bmatrix}$

(columns: Matrix, Alien, Serenity, Casablanca, Amelie)

**x**

$\begin{bmatrix} 0.56 & 0.12 \\ 0.59 & -0.02 \\ 0.56 & 0.12 \\ 0.09 & -0.69 \\ 0.09 & -0.69 \end{bmatrix}$

movie-to-concept
similarities (V)

SciFi-concept

$= \begin{bmatrix} 2.8 & 0.6 \end{bmatrix}$

# Case study: How to query?

☐ **How would the user *d* that rated ('Serenity') be handled?**
**$d_{concept} = d\ V$**

**E.g.:**

$q =$ $\begin{bmatrix} \overset{\text{Matrix}}{0} & \overset{\text{Alien}}{0} & \overset{\text{Serenity}}{4} & \overset{\text{Casablanca}}{0} & \overset{\text{Amelie}}{0} \end{bmatrix}$ **x** $\begin{bmatrix} 0.56 & 0.12 \\ 0.59 & -0.02 \\ 0.56 & 0.12 \\ 0.09 & -0.69 \\ 0.09 & -0.69 \end{bmatrix}$ **=** $\begin{bmatrix} 2.24 & 0.48 \end{bmatrix}$
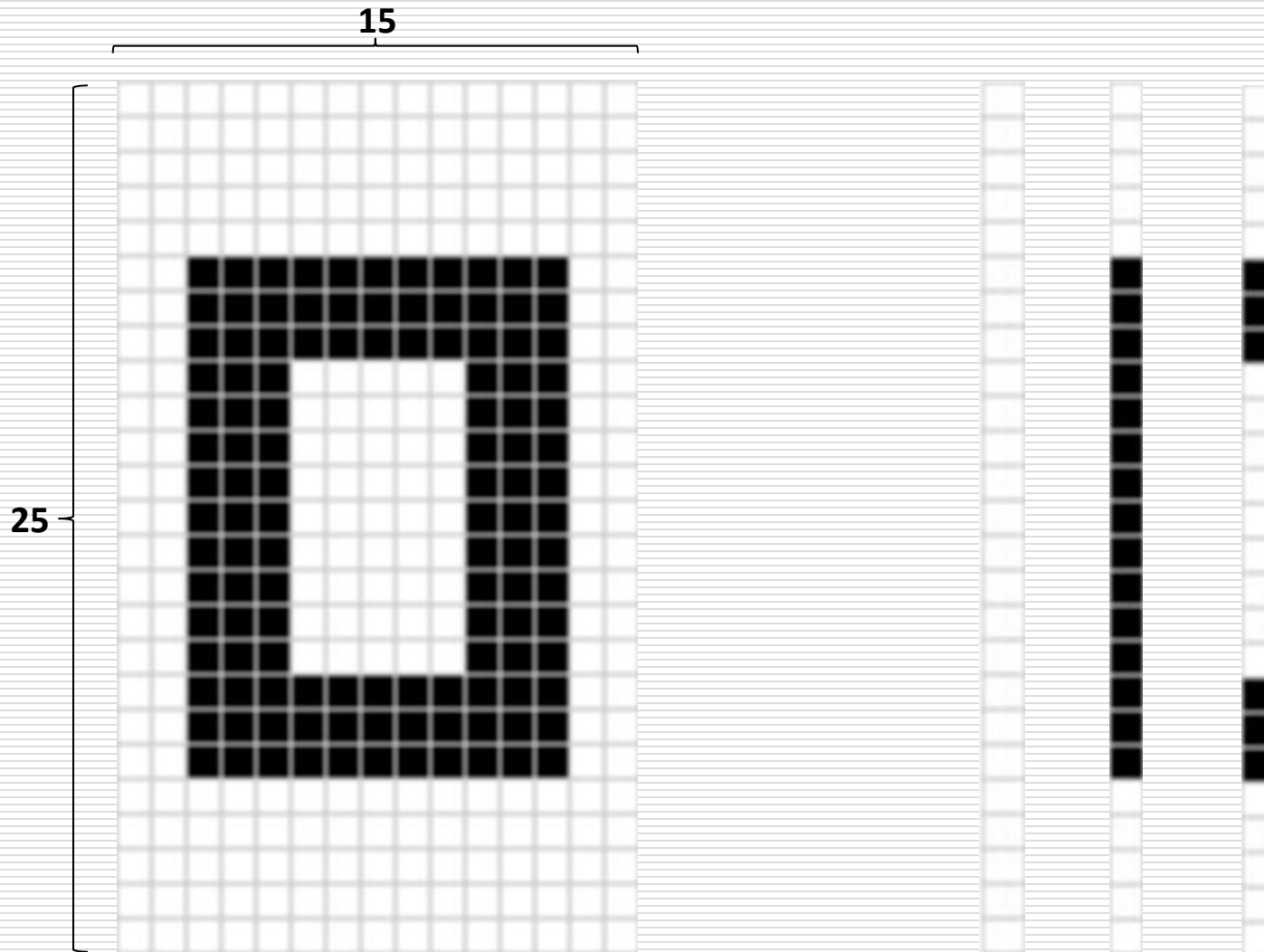
SciFi-concept

**movie-to-concept similarities (V)**

# Case study: How to query?

□ **Observation:** User *d* that rated ('*Serenity*') will be **similar** to user **q** that rated ('*Matrix*'), although *d* and **q** have **zero ratings in common**!
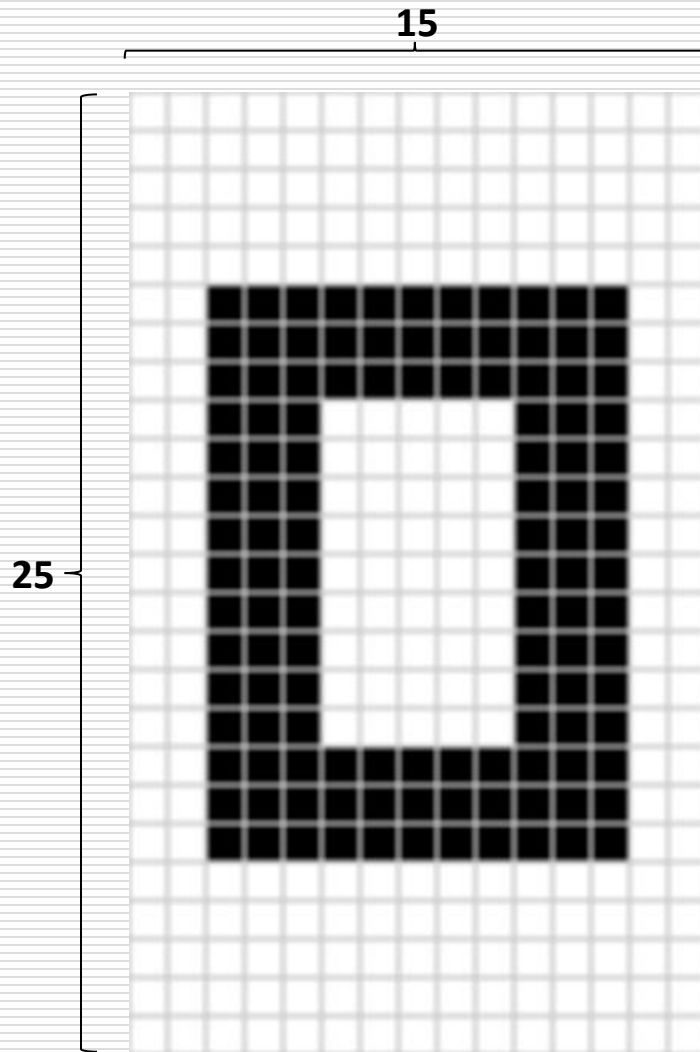
$$d = \begin{bmatrix} 0 & 0 & 4 & 0 & 0 \end{bmatrix} \dashrightarrow \begin{bmatrix} 2.24 & 0.48 \end{bmatrix}$$

$$q = \begin{bmatrix} 5 & 0 & 0 & 0 & 0 \end{bmatrix} \dashrightarrow \begin{bmatrix} 2.8 & 0.6 \end{bmatrix}$$

Columns: **Matrix, Alien, Serenity, Casablanca, Amelie**

**SciFi-concept**

**Zero ratings in common**                    **Similarity ≠ 0**

# Case study: Image compression

**15**

**25**

**Space cost=25*15=375**

# Case study: Image compression



**15**

**25**

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}$$

**Space cost=25*15=375**                    **white color=1, black color=0**

# Case study: Image compression



**15**

**25**

Space cost=25*15=375

SVD:

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

**σ1=14.72, σ2=5.22, σ3=3.31**
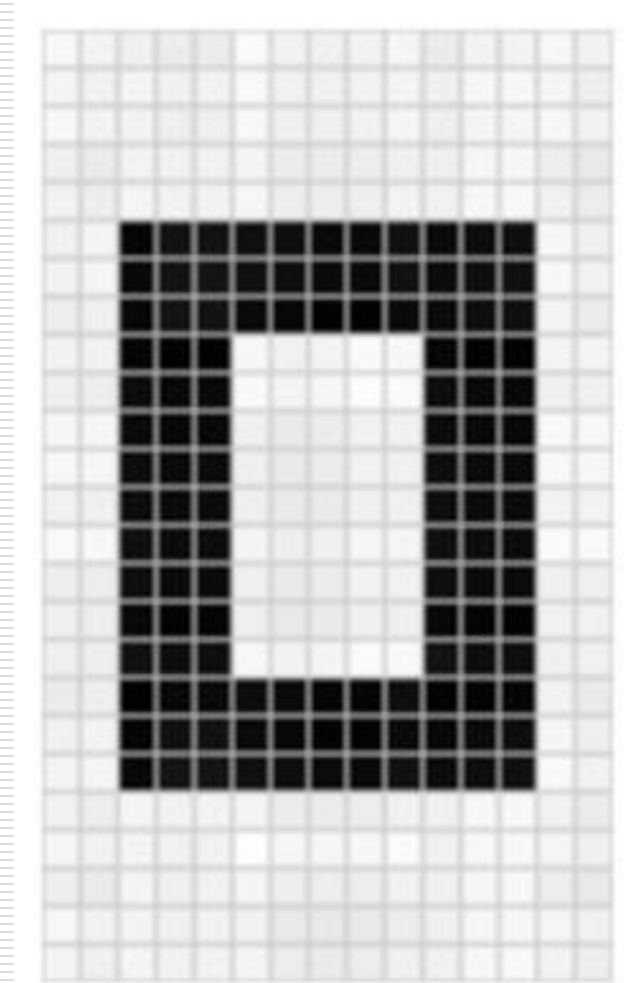
**Space cost=25*3+15*3+3=123**

# Case study: Noise reduction
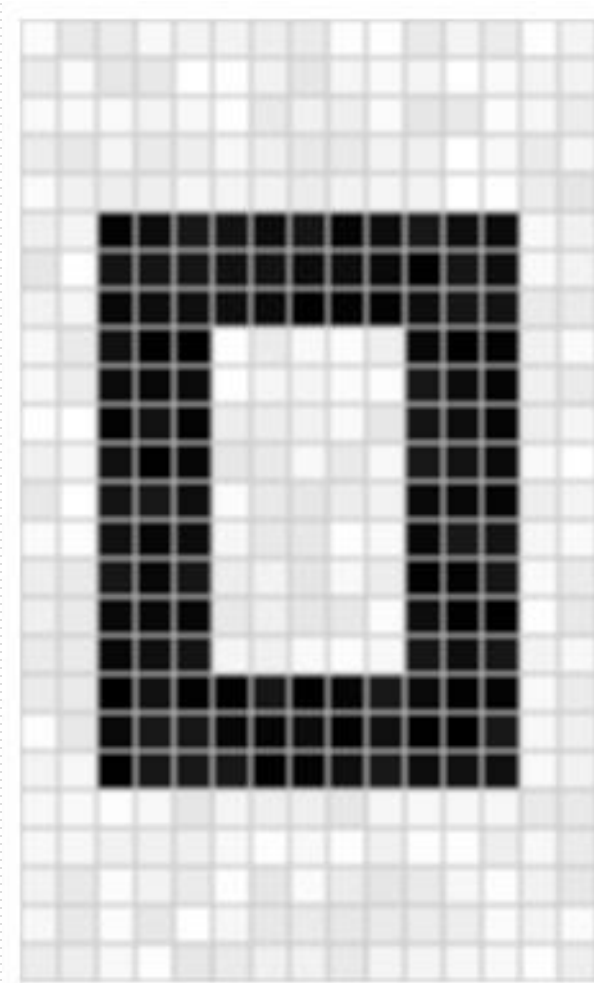


SVD:
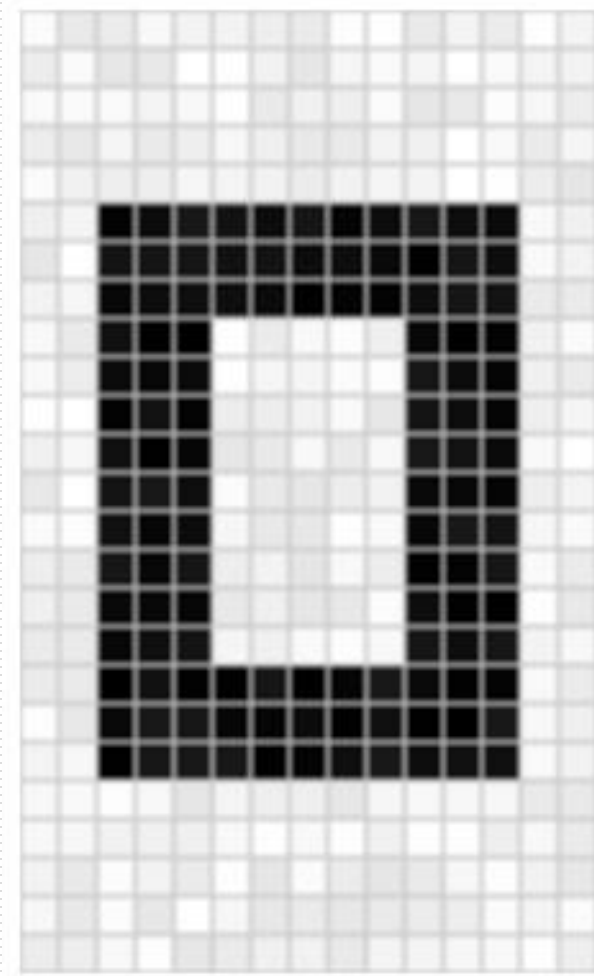
$$\mathbf{A} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T = \sum_i \sigma_i \mathbf{u}_i \circ \mathbf{v}_i$$

**σ1=14.15, σ2=4.67, σ3=3.00, σ4=0.21 , σ5=0.19, …… , σ15=0.05**

$$\mathbf{A} \approx \mathbf{u}_1\sigma_1\,\mathbf{v}_1^{\mathsf{T}} + \mathbf{u}_2\sigma_2\,\mathbf{v}_2^{\mathsf{T}} + \mathbf{u}_3\sigma_3\,\mathbf{v}_3^{\mathsf{T}}$$

# Case study: Noise reduction



**SVD with low rank approximation**

# Acknowledgement

- Slides are adapted from:
  - Prof. Jeffrey D. Ullman
  - Dr. Anand Rajaraman
  - Dr. Jure Leskovec