Carnegie Mellon

15-826: Multimedia (Databases) and Data Mining

Lecture #16: SVD - part I (definitions)

C. Faloutsos

Carnegie Mellon

Must-read Material

- Numerical Recipes in C ch. 2.6;
- MM Textbook Appendix D

Outline

Goal: ‘Find similar / interesting things’

- Intro to DB
- Indexing - similarity search
- Data Mining

Indexing - Detailed outline

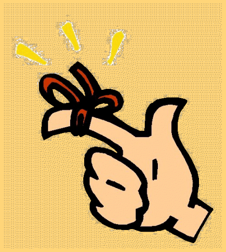
- primary key indexing
- secondary key / multi-key indexing
- spatial access methods
- fractals
- text
- Singular Value Decomposition (SVD)
- multimedia
- ...

SVD - Detailed outline

- Motivation
- Definition - properties
- Interpretation
- Complexity
- Case studies
- Additional properties

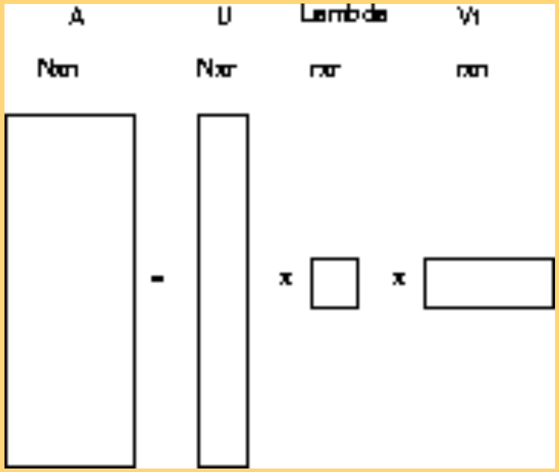
Problem

- How to find ‘concepts’ in a set of doc’s?
  - (~ clusters)



Conclusion

- How to find 'concepts' in a set of doc's?
  - (~ clusters)
- SVD (= LSI) on the document-term matrix



SVD - Motivation

- problem #1: text - LSI: find 'concepts'
- problem #2: compression / dim. reduction

SVD - Motivation

- problem #1: text - LSI: find 'concepts'

term / document	data	information	retrieval	brain	lung
CS-TR1	1	1	1	0	0
CS-TR2	2	2	2	0	0
CS-TR3	1	1	1	0	0
CS-TR4	5	5	5	0	0
MED-TR1	0	0	0	2	2
MED-TR2	0	0	0	3	3
MED-TR3	0	0	0	1	1

SVD - Motivation

- Customer-product, for recommendation system:

$$
\begin{array}{c c}
 & \begin{array}{ccccc} \text{bread} & \text{lettuce} & \text{tomatos} & \text{beef} & \text{chicken} \end{array} \\
\begin{array}{c} \\ \\ \text{vegetarians} \\ \\ \\ \\ \text{meat eaters} \end{array} &
\left[\begin{array}{ccccc}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{array} \right]
\end{array}
$$

SVD - Motivation

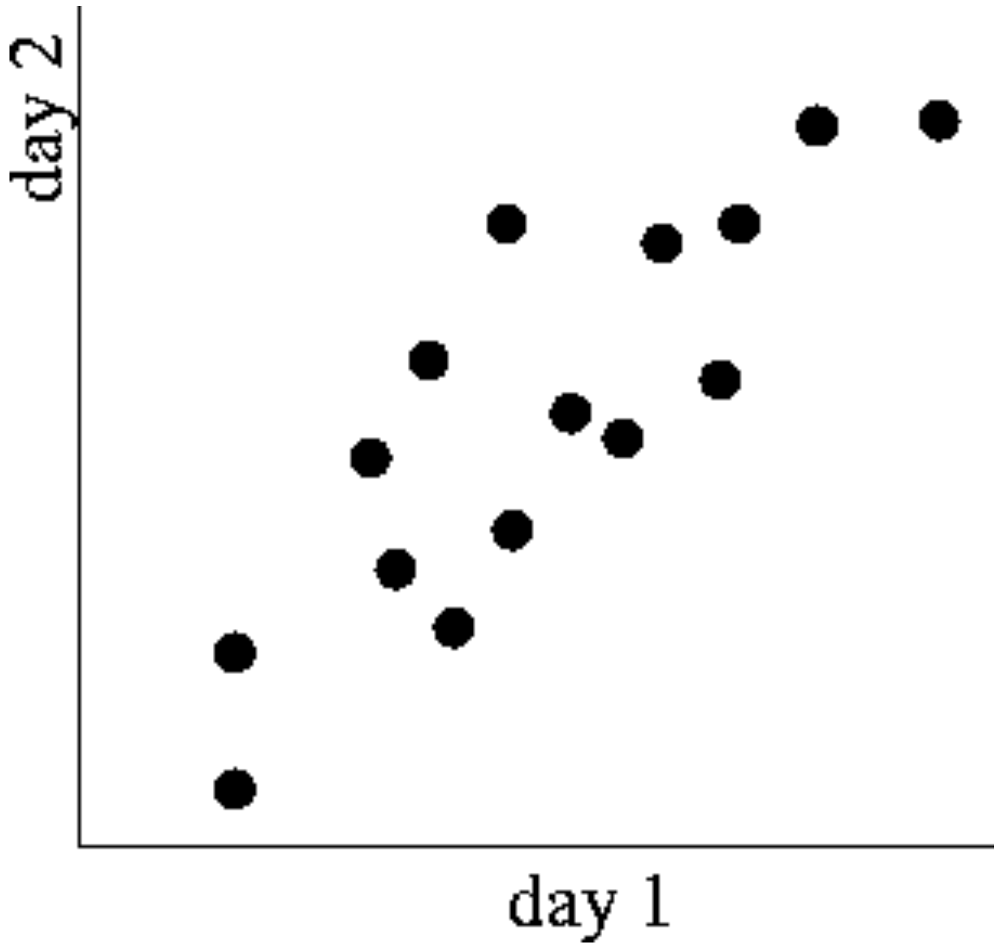
- problem #2: compress / reduce dimensionality

Problem - specs

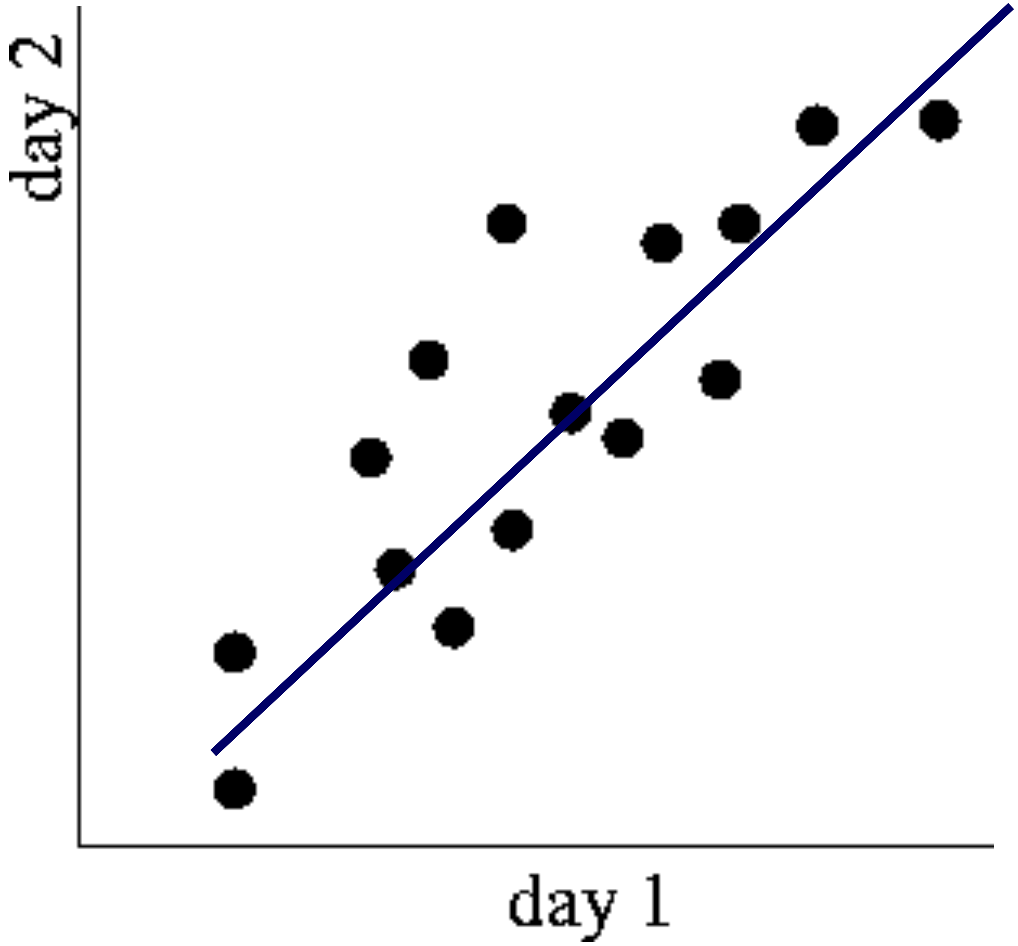
- ~10**6 rows; ~10**3 columns; no updates;
- random access to any cell(s) ; small error: OK

day / customer	We 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

SVD - Motivation



SVD - Motivation



SVD - Detailed outline

- Motivation
- Definition - properties
- Interpretation
- Complexity
- Case studies
- Additional properties

SVD - Definition

(reminder: matrix multiplication

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \times \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \\ \\ \end{bmatrix}$$

3 x *2*   *2* x *1*

SVD - Definition

(reminder: matrix multiplication

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ x } \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} \\ \\ \\ \end{bmatrix}$$

3 x *2* *2* x *1* *3* x *1*

SVD - Definition

(reminder: matrix multiplication

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ x } \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ \\ \\ \end{bmatrix}$$

3 x *2* *2* x *1* *3* x *1*

SVD - Definition

(reminder: matrix multiplication

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{x} \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ \end{bmatrix}$$

3 x *2* *2* x *1* *3* x *1*

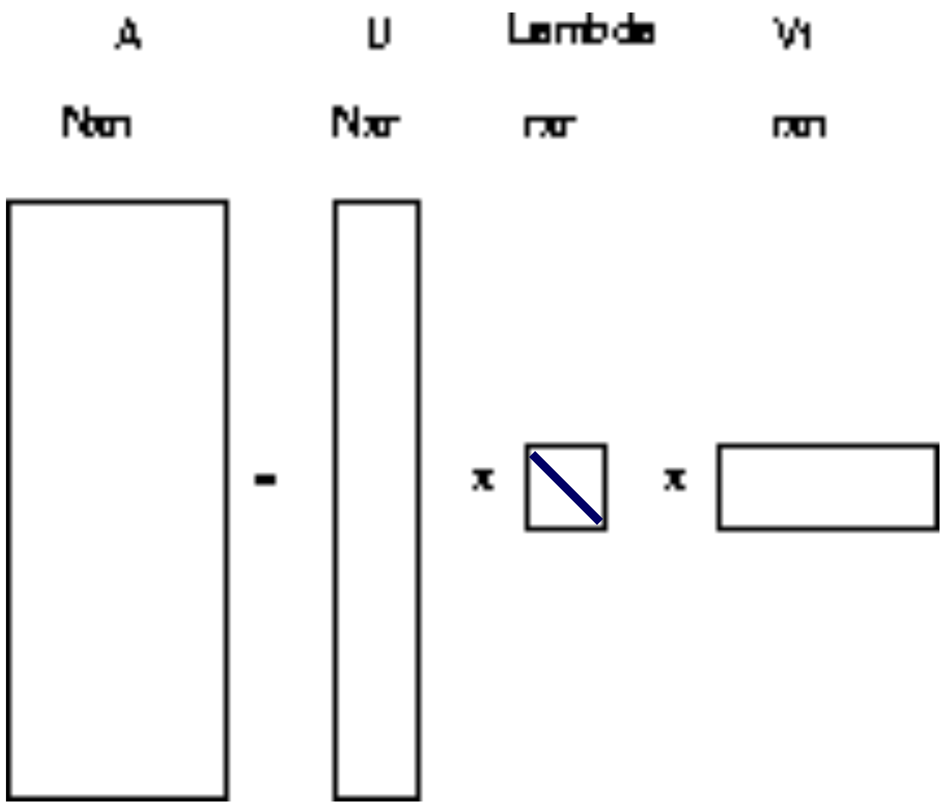
SVD - Definition

(reminder: matrix multiplication

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \text{ x } \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix}$$

SVD - Definition

- $\mathbf{A} = \mathbf{U} \, \mathbf{\Lambda} \, \mathbf{V}^T$ - example:



SVD - Definition

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

- **A**: n x m matrix (eg., n documents, m terms)
- **U**: n x r matrix (n documents, r concepts)
- **Λ**: r x r diagonal matrix (strength of each 'concept') (r : rank of the matrix)
- **V**: m x r matrix (m terms, r concepts)

SVD - Properties

THEOREM [Press+92]: always possible to decompose matrix **A** into **A** = **U** **Λ** **V**^{T} , where

- **U, Λ, V**: unique (*)
- **U**, **V**: column orthonormal (ie., columns are unit vectors, orthogonal to each other)
  - **U**^{T} **U** = **I**; **V**^{T} **V** = **I** (**I:** identity matrix**)**
- **Λ:** singular are positive, and sorted in decreasing order

SVD - Example

- $\mathbf{A} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{V}^T$ - example:

$$
\begin{array}{c}
 \\
\text{CS} \\
 \\
 \\
\text{MD} \\
 \\
\end{array}
\begin{array}{c}
\begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{bmatrix}
\times
\begin{bmatrix}
9.64 & 0 \\
0 & 5.29
\end{bmatrix}
\times
\begin{bmatrix}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{bmatrix}
$$

SVD - Example

- $\mathbf{A} = \mathbf{U}\ \mathbf{\Lambda}\ \mathbf{V}^T$ - example:

$$
\begin{array}{c}
\\
\text{CS} \\
\\
\text{MD}
\end{array}
\begin{array}{c}
\begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix}
\end{array}
=
\begin{array}{c}
\begin{array}{cc} \text{CS-concept} & \text{MD-concept} \end{array} \\
\begin{bmatrix}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{bmatrix}
\end{array}
\text{x}
\begin{bmatrix}
9.64 & 0 \\
0 & 5.29
\end{bmatrix}
\text{x}
\begin{bmatrix}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{bmatrix}
$$

Carnegie Mellon

SVD - Example

- **A = U Λ V^T** - example:

doc-to-concept similarity matrix

CS-concept

MD-concept

$$
\begin{array}{c|c}
 & \begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{array}{c} \text{CS} \\ \\ \\ \\ \text{MD} \\ \\ \\ \end{array} &
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}
\end{array}
=
\begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix}
\times
\begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix}
\times
\begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}
$$

SVD - Example

- $\mathbf{A} = \mathbf{U}\ \mathbf{\Lambda}\ \mathbf{V}^T$ - example:

'strength' of CS-concept

$$
\begin{array}{cc}
 & \begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{array}{c} \text{CS} \\ \\ \\ \\ \text{MD} \\ \\ \\ \end{array} &
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix}
\end{array}
=
\begin{bmatrix}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{bmatrix}
\mathrm{x}
\begin{bmatrix}
9.64 & 0 \\
0 & 5.29
\end{bmatrix}
\mathrm{x}
\begin{bmatrix}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{bmatrix}
$$

SVD - Example

- $\mathbf{A} = \mathbf{U}\, \mathbf{\Lambda}\, \mathbf{V}^T$ - example:

term-to-concept similarity matrix

CS-concept

$$
\begin{array}{cc}
 & \begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{array}{c} \text{CS} \\ \\ \\ \\ \text{MD} \\ \\ \\ \end{array} &
\left[\begin{array}{ccccc}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{array}\right]
\end{array}
=
\left[\begin{array}{cc}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{array}\right]
\times
\left[\begin{array}{cc}
9.64 & 0 \\
0 & 5.29
\end{array}\right]
\times
\left[\begin{array}{ccccc}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{array}\right]
$$

SVD - Example

- $\mathbf{A} = \mathbf{U}\ \mathbf{\Lambda}\ \mathbf{V}^T$ - example:

term-to-concept similarity matrix

CS-concept

$$
\begin{array}{c c}
 & \begin{array}{ccccc} \text{data} & \text{inf.} & \text{retrieval} & \text{brain} & \text{lung} \end{array} \\
\begin{array}{c} \text{CS} \\ \\ \\ \\ \text{MD} \\ \\ \\ \end{array} &
\left[\begin{array}{ccccc}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{array}\right]
\end{array}
=
\left[\begin{array}{cc}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{array}\right]
\times
\left[\begin{array}{cc}
9.64 & 0 \\
0 & 5.29
\end{array}\right]
\times
\left[\begin{array}{ccccc}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{array}\right]
$$

SVD - Detailed outline

- Motivation
- Definition - properties
- Interpretation
- Complexity
- Case studies
- Additional properties

SVD - Interpretation #1

'documents', 'terms' and 'concepts':

- **U**: document-to-concept similarity matrix
- **V**: term-to-concept sim. matrix
- **Λ**: its diagonal elements: 'strength' of each concept

SVD – Interpretation #1

'documents', 'terms' and 'concepts':

Q**:** if **A** is the document-to-term matrix, what is $\mathbf{A}^T \mathbf{A}$?

A:

Q: $\mathbf{A} \mathbf{A}^T$?

A:

SVD – Interpretation #1

'documents', 'terms' and 'concepts':

Q: if **A** is the document-to-term matrix, what is $\mathbf{A}^T \mathbf{A}$?

A: term-to-term ([m x m]) similarity matrix

Q: $\mathbf{A} \mathbf{A}^T$?

A: document-to-document ([n x n]) similarity matrix

Carnegie Mellon

SVD properties

- **V** are the eigenvectors of the *covariance matrix* $\mathbf{A}^T\mathbf{A}$

- **U** are the eigenvectors of the *Gram (inner-product) matrix* $\mathbf{A}\mathbf{A}^T$

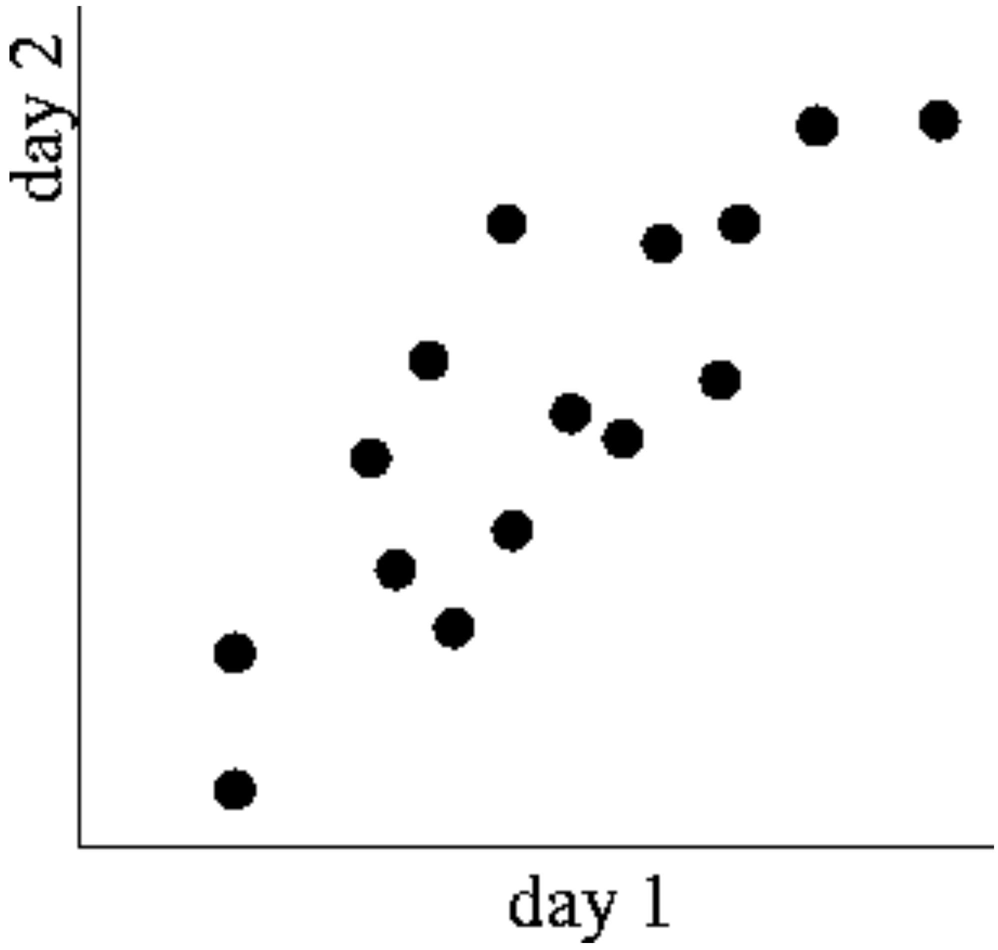
<u>Further reading:</u>

1. Ian T. Jolliffe, *Principal Component Analysis* (2nd ed), Springer, 2002.
2. Gilbert Strang, *Linear Algebra and Its Applications* (4th ed), Brooks Cole, 2005.

SVD - Interpretation #2

- best axis to project on: ('best' = min sum of squares of projection errors)

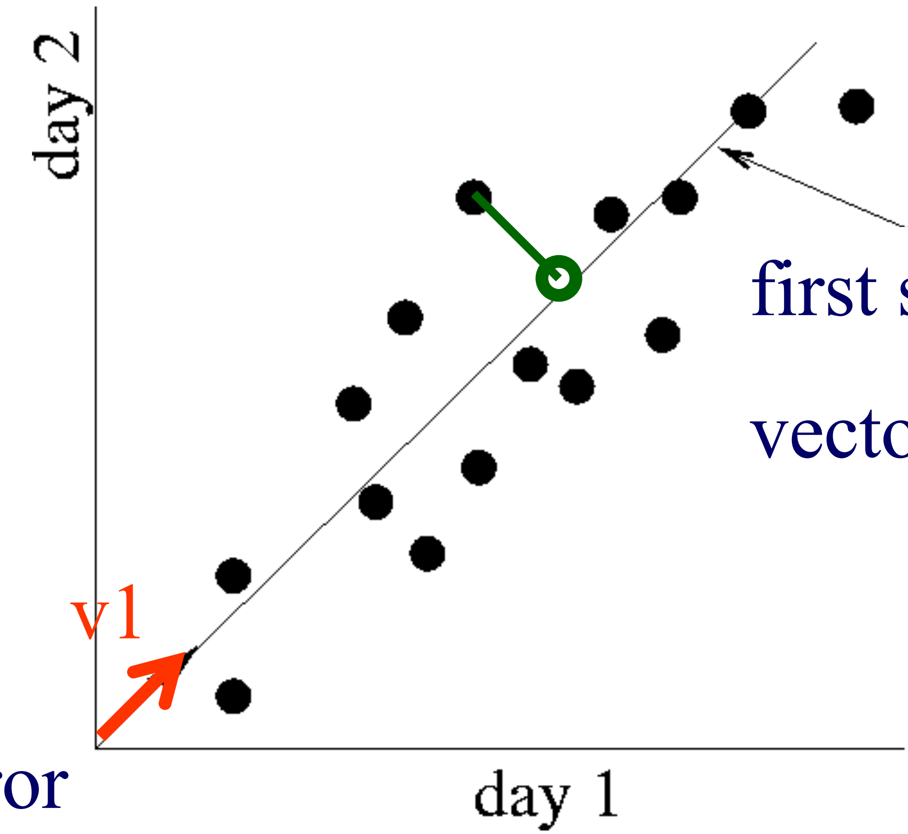
SVD - Motivation



Carnegie Mellon

SVD - interpretation #2

SVD: gives
best axis to project



- minimum RMS error

SVD - Interpretation #2

day customer	We 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

Carnegie Mellon

SVD - Interpretation #2

- $\mathbf{A} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{V}^T$ - example:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \times \begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix} \times \begin{array}{l} \text{v1} \\ \begin{bmatrix} \boxed{0.58 \;\; 0.58 \;\; 0.58 \;\; 0 \;\; 0} \\ 0 \;\; 0 \;\; 0 \;\; 0.71 \;\; 0.71 \end{bmatrix} \end{array}$$

SVD - Interpretation #2

- $\mathbf{A} = \mathbf{U}\ \mathbf{\Lambda}\ \mathbf{V}^T$ - example:

variance ('spread') on the v1 axis

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \mathrm{x} \begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix} \mathrm{x} \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}$$

Carnegie Mellon

SVD - Interpretation #2

- $\mathbf{A} = \mathbf{U}\,\mathbf{\Lambda}\,\mathbf{V}^T$ - example:
  - $\mathbf{U}\,\mathbf{\Lambda}$ gives the coordinates of the points in the projection axis

$$
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \text{x} \begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix} \text{x} \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}
$$

SVD - Interpretation #2

- More details
- Q: how exactly is dim. reduction done?

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \times \begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix} \times \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}$$

SVD - Interpretation #2

- More details
- Q: how exactly is dim. reduction done?
- A: set the smallest singular values to zero:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \times \begin{bmatrix} 9.64 & 0 \\ 0 & \xcancel{5.29} \end{bmatrix} \times \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}$$

Carnegie Mellon

SVD - Interpretation #2

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \mathrm{x} \begin{bmatrix} 9.64 & 0 \\ 0 & \cancel{0} \end{bmatrix} \mathrm{x} \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}$$

SVD - Interpretation #2

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 0.18 & \cancel{0} \\ 0.36 & \cancel{0} \\ 0.18 & \cancel{0} \\ 0.90 & \cancel{0} \\ 0 & \cancel{0.53} \\ 0 & \cancel{0.80} \\ 0 & \cancel{0.27} \end{bmatrix} \text{x} \begin{bmatrix} 9.64 & 0 \\ 0 & \cancel{0} \end{bmatrix} \text{x} \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ \cancel{0} & \cancel{0} & \cancel{0} & \cancel{0.71} & \cancel{0.71} \end{bmatrix}$$

SVD - Interpretation #2

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \sim \begin{bmatrix} 0.18 \\ 0.36 \\ 0.18 \\ 0.90 \\ 0 \\ 0 \\ 0 \end{bmatrix} \times \begin{bmatrix} 9.64 \end{bmatrix} \times \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \end{bmatrix}$$

SVD - Interpretation #2

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

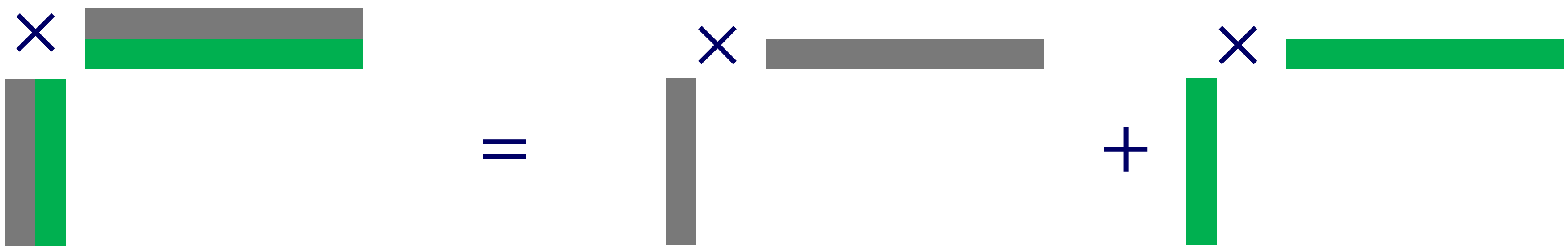
SVD - Interpretation #2

Exactly equivalent:
'spectral decomposition' of the matrix:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 0.18 & 0 \\ 0.36 & 0 \\ 0.18 & 0 \\ 0.90 & 0 \\ 0 & 0.53 \\ 0 & 0.80 \\ 0 & 0.27 \end{bmatrix} \mathrm{x} \begin{bmatrix} 9.64 & 0 \\ 0 & 5.29 \end{bmatrix} \mathrm{x} \begin{bmatrix} 0.58 & 0.58 & 0.58 & 0 & 0 \\ 0 & 0 & 0 & 0.71 & 0.71 \end{bmatrix}$$

Preliminaries: Spectral form

$\mathbf{U} \times \mathbf{V}^{\mathrm{T}}$ $\qquad$ $u_1 \times v_1{}^T$ $\qquad$ $u_2 \times v_2{}^T$



SVD – Spectral form

Exactly equivalent:

'spectral decomposition' of the matrix:

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} | & | \\ u_1 & u_2 \\ | & | \end{bmatrix} \text{ x } \begin{bmatrix} \lambda_1 & \emptyset \\ \emptyset & \lambda_2 \end{bmatrix} \text{ x } \begin{bmatrix} \text{---} & v_1 & \text{---} \\ \text{---} & v_2 & \text{---} \end{bmatrix}$$

SVD – Spectral form

Exactly equivalent:

'spectral decomposition' of the matrix:

$$\begin{array}{cc} & \longleftarrow m \longrightarrow \\ n & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array} = \lambda_1 \quad u_1 \quad {v^T}_1 + \quad \lambda_2 \quad u_2 \quad {v^T}_2 + \ldots$$

SVD – Spectral form

Exactly equivalent:

'spectral decomposition' of the matrix:

$$\begin{array}{cc} & m \\ n & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} \end{array} = \overbrace{\lambda_1 \; u_1 \; v^T{}_1 + \lambda_2 \; u_2 \; v^T{}_2 + \ldots}^{r \text{ terms}}$$

n x 1

1 x m

SVD – Spectral form

approximation / dim. reduction:
by keeping the first few terms (Q: how many?)

$$
\begin{array}{c c}
 & \xleftarrow{\quad m \quad}\!\!\!\rightarrow \\
n & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}
\end{array}
= \lambda_1 \; u_1 \; v^T{}_1 + \lambda_2 \; u_2 \; v^T{}_2 + \ldots
$$

assume: $\lambda_1 >= \lambda_2 >= \ldots$

SVD - Spectral form

A (heuristic - [Fukunaga]): keep 80-90% of 'energy' (= sum of squares of λ_i's)

$$
\begin{array}{cc}
 & \xleftarrow{\quad} m \xrightarrow{\quad} \\
n \updownarrow & \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}
\end{array}
= \lambda_1 \quad u_1 \quad v^T{}_1 + \lambda_2 \quad u_2 \quad v^T{}_2 + \ldots
$$

assume: $\lambda_1 >= \lambda_2 >= \ldots$

SVD - Spectral form

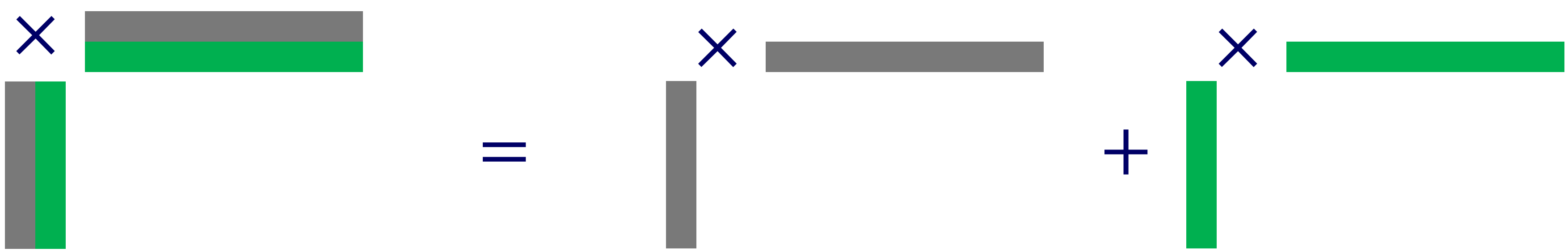
Also, if there are clear blocks, each corresponds to a 'concept'/singular-vector-pair

$$\underbrace{\begin{bmatrix} \boxed{\begin{matrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 1 & 1 & 1 \\ 5 & 5 & 5 \end{matrix}} & \begin{matrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{matrix} \\ \begin{matrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{matrix} & \boxed{\begin{matrix} 2 & 2 \\ 3 & 3 \\ 1 & 1 \end{matrix}} \end{bmatrix}}_{n \times m} = \underline{\lambda_1 \quad u_1 \quad v^T{}_1} + \underline{\lambda_2 \quad u_2 \quad v^T{}_2} + \ldots$$

assume: $\lambda_1 >= \lambda_2 >= \ldots$

SVD – Spectral form

$\mathbf{U} \times \mathbf{V}^{\mathrm{T}}$ $\qquad\qquad$ $u_1 \times v_1^T$ $\qquad\qquad$ $u_2 \times v_2^T$



SVD – Spectral form

Arithmetic example:

$$\begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} \times \begin{bmatrix} 10 & 2 \\ -10 & -2 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} \times \begin{bmatrix} 10 & 2 \end{bmatrix} +$$

$$\begin{bmatrix} 2 \\ 4 \\ 6 \end{bmatrix} \times \begin{bmatrix} -10 & -2 \end{bmatrix}$$

$$= \begin{bmatrix} -10 & -2 \\ -10 & -2 \\ -10 & -2 \end{bmatrix}$$

SVD - Detailed outline

- Motivation
- Definition - properties
- Interpretation
  - #1: documents/terms/concepts
  - #2: dim. reduction
  - #3: picking non-zero, rectangular 'blobs'
- Complexity
- Case studies
- Additional properties

SVD - Interpretation #3

- finds non-zero 'blobs' in a data matrix

$$
\begin{bmatrix}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{bmatrix}
=
\begin{bmatrix}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{bmatrix}
\times
\begin{bmatrix}
9.64 & 0 \\
0 & 5.29
\end{bmatrix}
\times
\begin{bmatrix}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{bmatrix}
$$

SVD - Interpretation #3

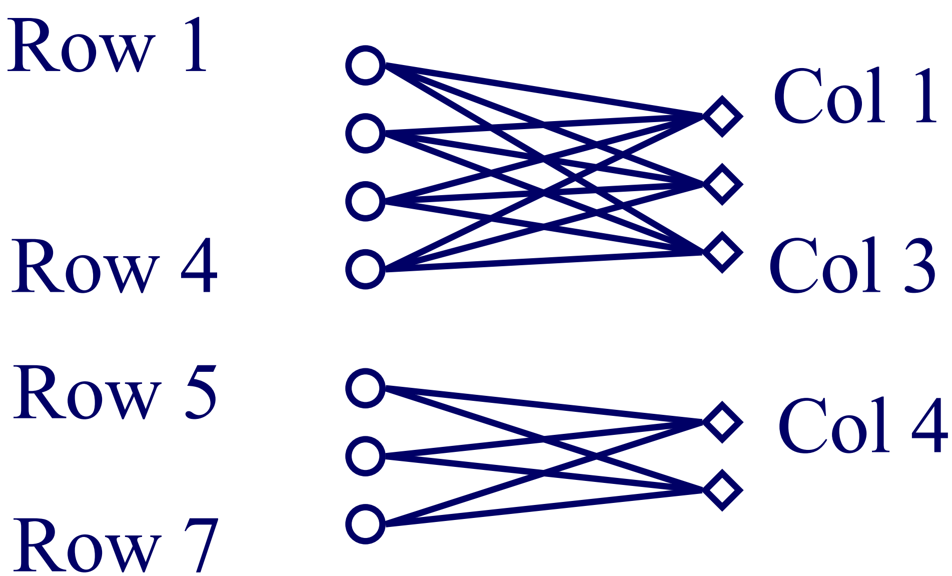
- finds non-zero 'blobs' in a data matrix

$$
\left[\begin{array}{ccc|cc}
1 & 1 & 1 & 0 & 0 \\
2 & 2 & 2 & 0 & 0 \\
1 & 1 & 1 & 0 & 0 \\
5 & 5 & 5 & 0 & 0 \\
\hline
0 & 0 & 0 & 2 & 2 \\
0 & 0 & 0 & 3 & 3 \\
0 & 0 & 0 & 1 & 1
\end{array}\right]
=
\begin{bmatrix}
0.18 & 0 \\
0.36 & 0 \\
0.18 & 0 \\
0.90 & 0 \\
0 & 0.53 \\
0 & 0.80 \\
0 & 0.27
\end{bmatrix}
\times
\begin{bmatrix}
9.64 & 0 \\
0 & 5.29
\end{bmatrix}
\times
\begin{bmatrix}
0.58 & 0.58 & 0.58 & 0 & 0 \\
0 & 0 & 0 & 0.71 & 0.71
\end{bmatrix}
$$

SVD - Interpretation #3

- finds non-zero 'blobs' in a data matrix =
- 'communities' (bi-partite cores, here)

$$\left[\begin{array}{ccc|cc} 1 & 1 & 1 & 0 & 0 \\ 2 & 2 & 2 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 5 & 5 & 5 & 0 & 0 \\ \hline 0 & 0 & 0 & 2 & 2 \\ 0 & 0 & 0 & 3 & 3 \\ 0 & 0 & 0 & 1 & 1 \end{array}\right]$$

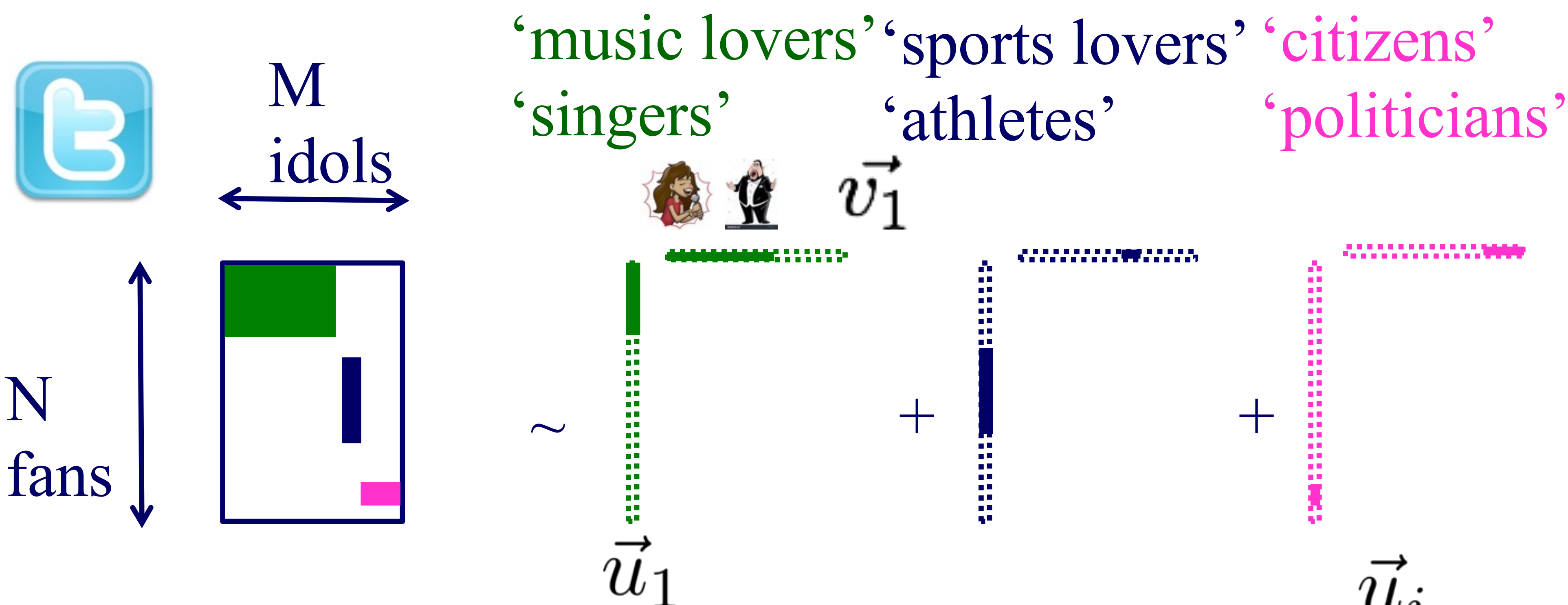


SVD – Interpretation #3

- Another pictorial example
- And its connection to EigenSpokes

SVD – Interpretation #3

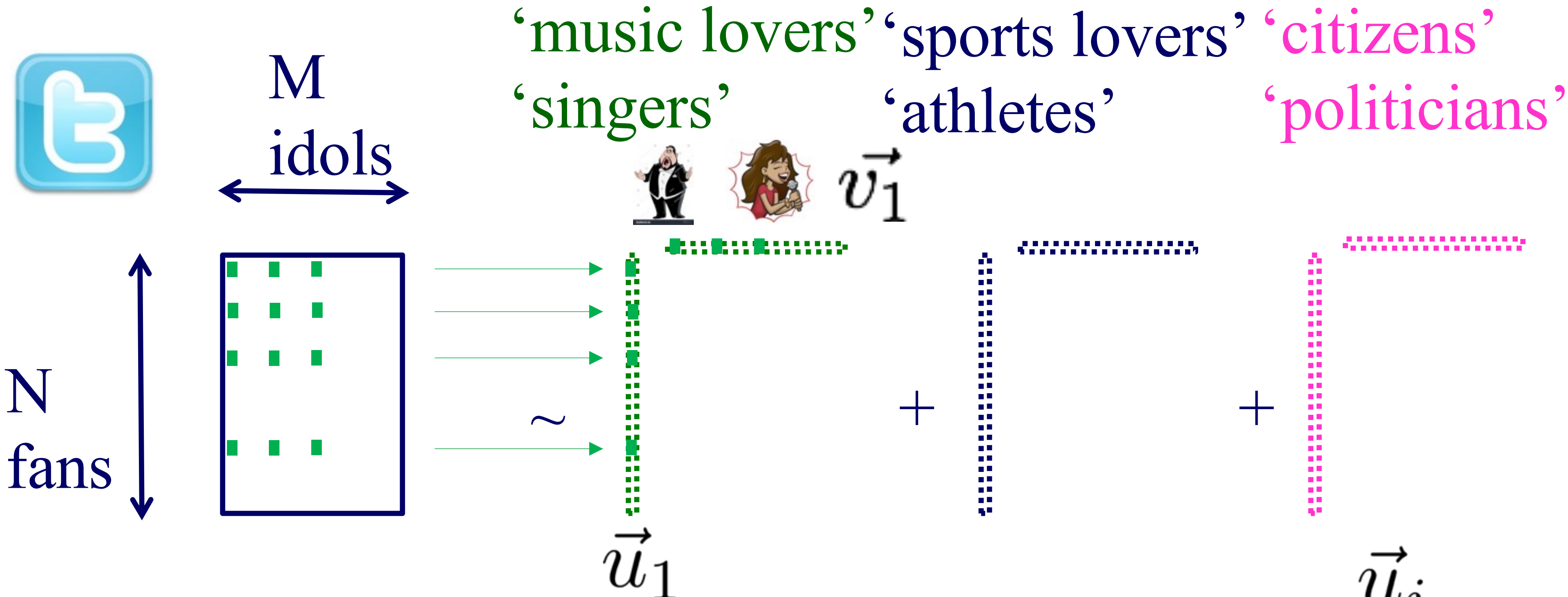
- It finds blocks



SVD – Interpretation #3

- It finds blocks

A) Even if shuffled!

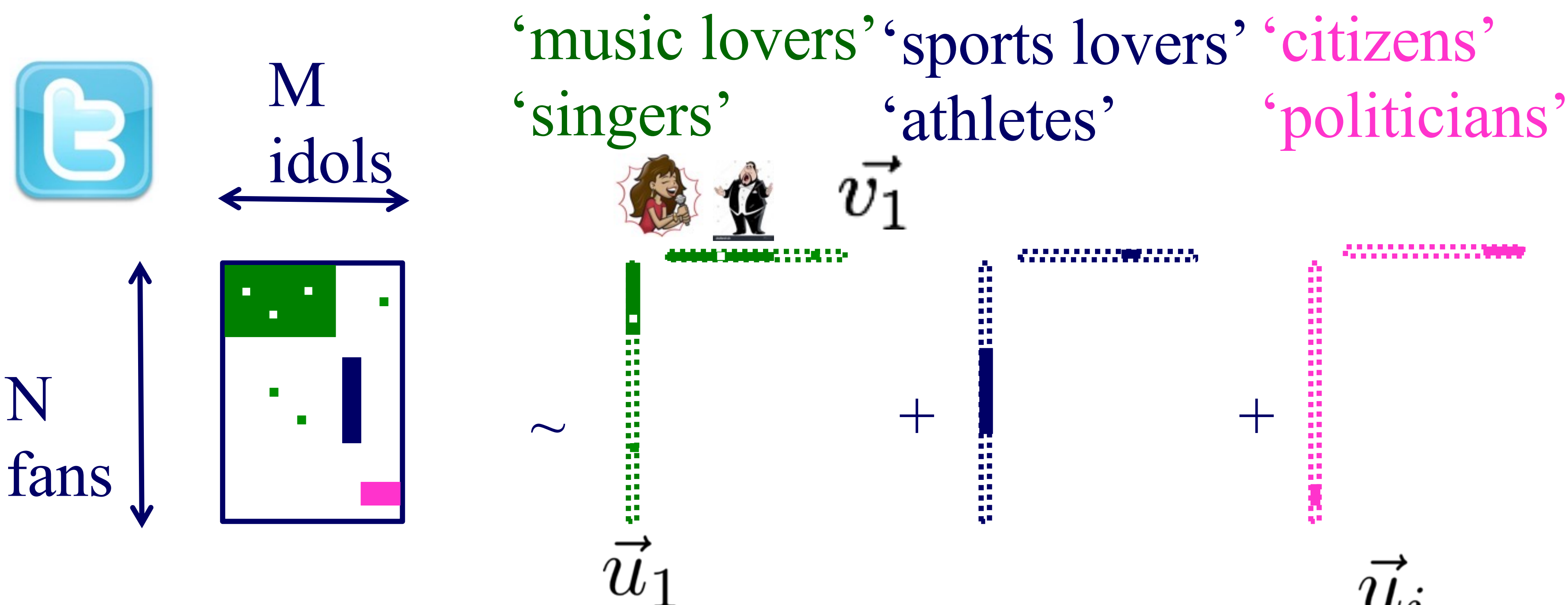


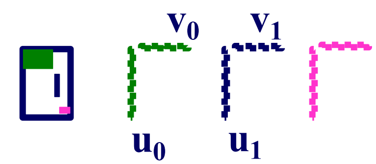
Carnegie Mellon

SVD – Interpretation #3

- It finds blocks

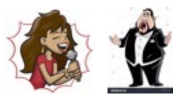
B) Even if ‘salt+pepper’ noise



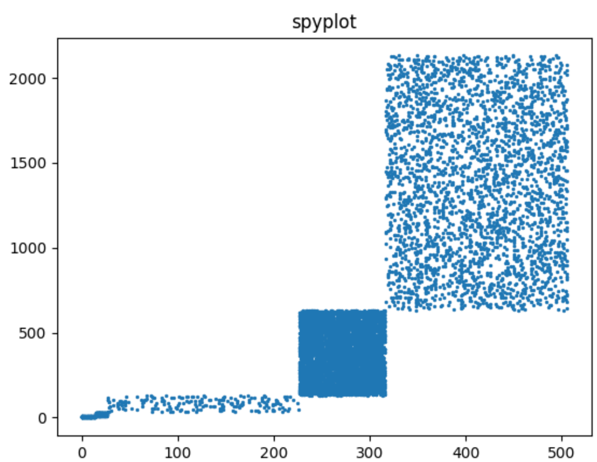


Toy example – EigenSpokes

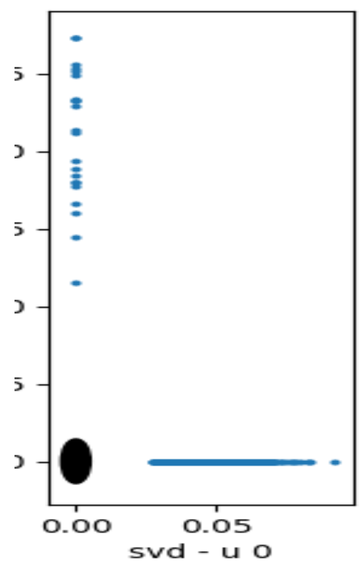
‘idols’



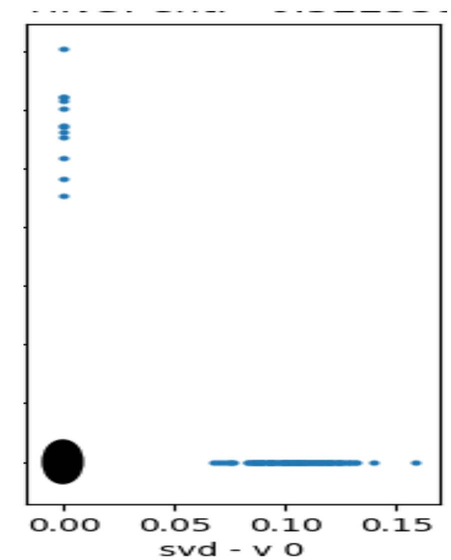
‘fans’



u_1



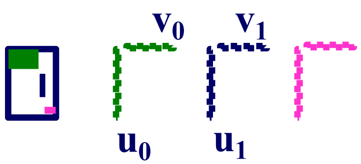
v_1



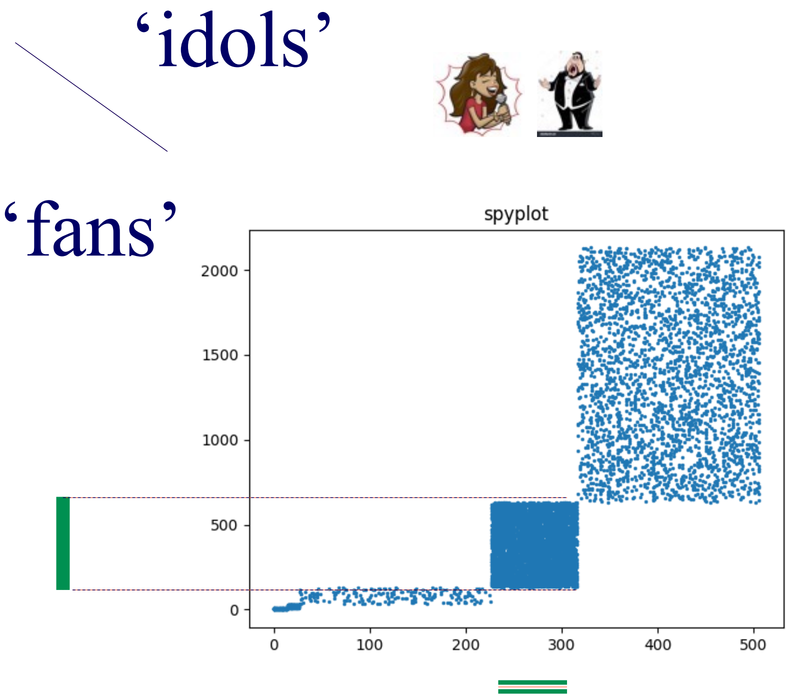
u_0 v_0

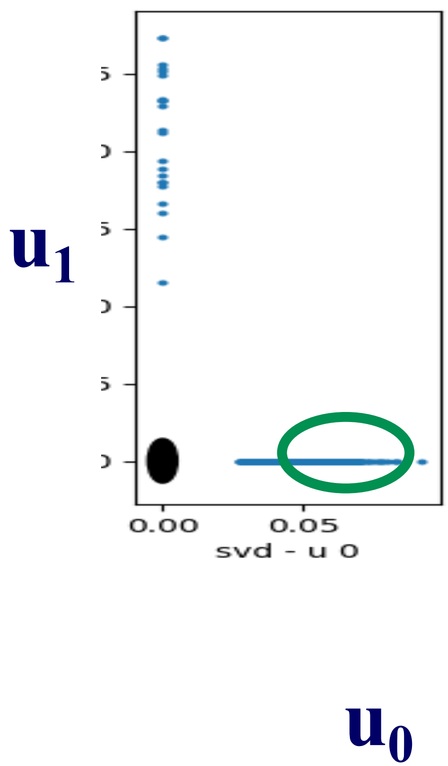
EigenPlots

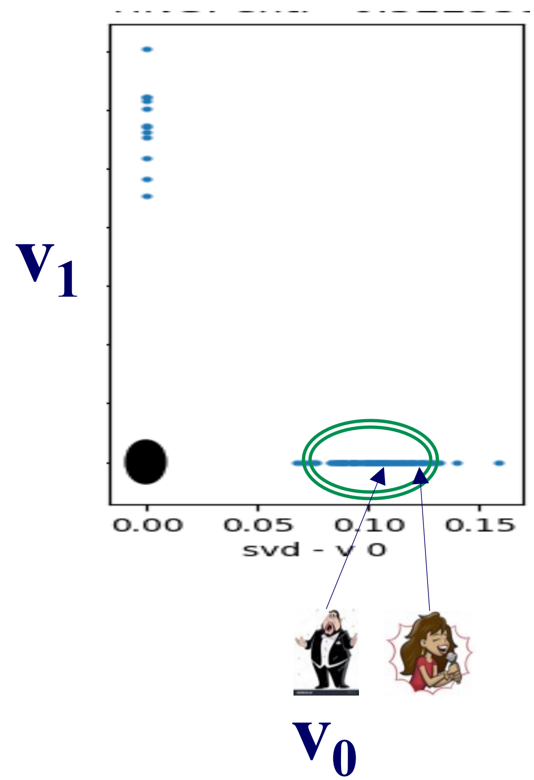
Carnegie Mellon



Toy example – EigenSpokes

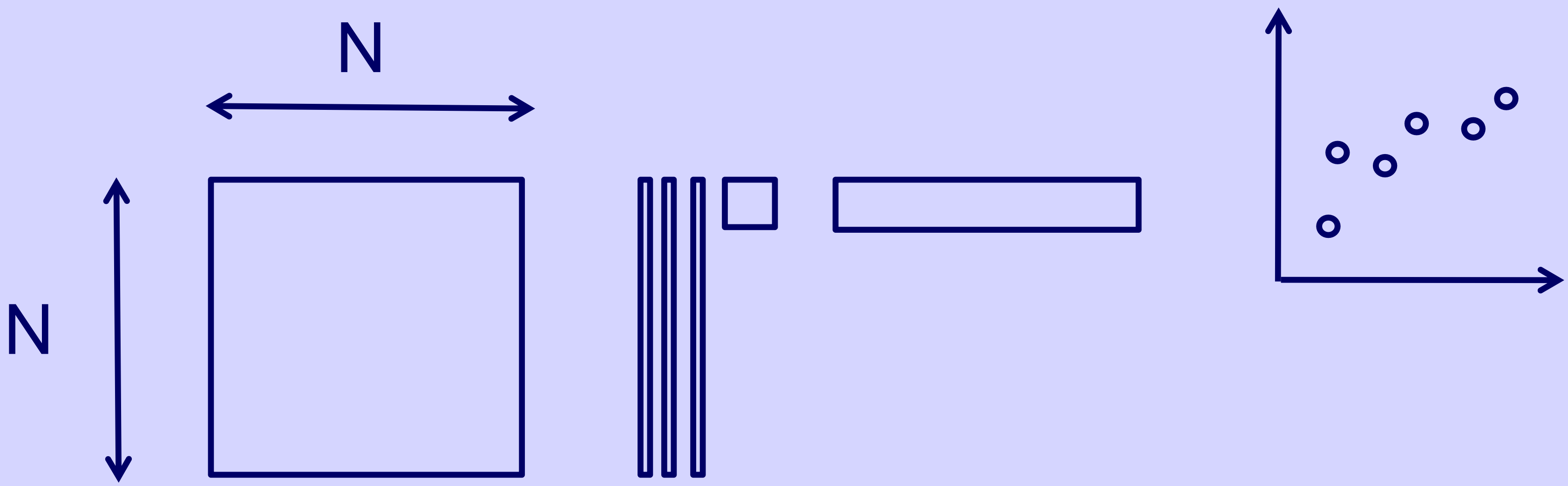






EigenSpokes

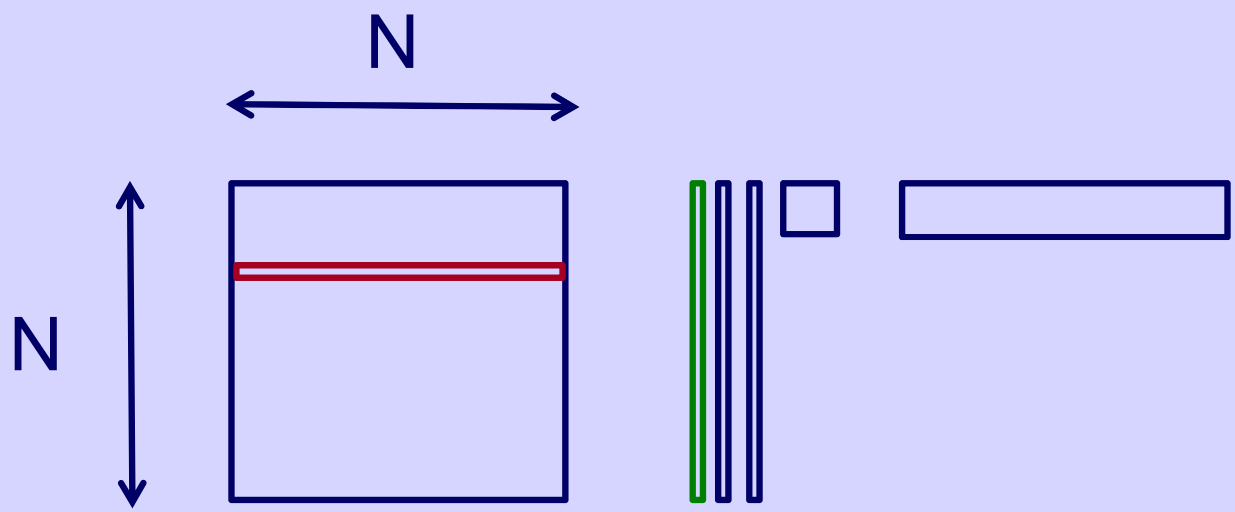
- Eigenvectors of adjacency matrix



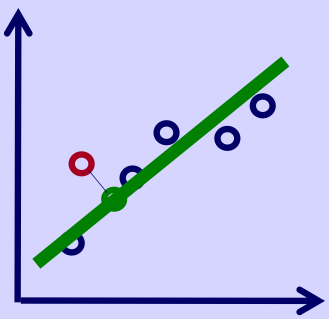
$\vec{u}_1$ $\vec{u}_i$

EigenSpokes

- Eigenvectors of adjacency matrix



$\vec{u}_1$ $\vec{u}_i$



SVD - Drill

- Compute SVD by hand
  - On a block-structured, toy matrix
  - Exploiting the properties of SVD

SVD - Drill

- Drill: find the SVD, 'by inspection' !
- Q: rank = ??

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} ?? \end{bmatrix} \text{ x } \begin{bmatrix} ?? \end{bmatrix} \text{ x } \begin{bmatrix} ?? \end{bmatrix}$$

SVD - Drill

- A: rank = 2 (2 linearly independent rows/cols)

$$\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} | & | \\ ?? & ?? \\ | & | \end{bmatrix} \times \begin{bmatrix} ?? & 0 \\ 0 & ?? \end{bmatrix} \times \begin{bmatrix} \text{—— ?? ——} \\ \text{—— ?? ——} \end{bmatrix}$$

SVD - Drill

- A: rank = 2 (2 linearly independent rows/cols)

$$
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \mathrm{x} \begin{bmatrix} ?? & 0 \\ 0 & ?? \end{bmatrix} \mathrm{x} \begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}
$$

orthogonal??

SVD - Drill

- column vectors: are orthogonal - but not unit vectors:

$$
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/sqrt(3) & 0 \\ 1/sqrt(3) & 0 \\ 1/sqrt(3) & 0 \\ 0 & 1/sqrt(2) \\ 0 & 1/sqrt(2) \end{bmatrix} \times \begin{bmatrix} ?? & 0 \\ 0 & ?? \end{bmatrix} \times \begin{bmatrix} 1/sqrt(3) & 1/sqrt(3) & 1/sqrt(3) & 0 & 0 \\ 0 & 0 & 0 & 1/sqrt(2) & 1/sqrt(2) \end{bmatrix}
$$

SVD - Drill

- and the singular values are:

$$
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\mathrm{sqrt}(3) & 0 \\ 1/\mathrm{sqrt}(3) & 0 \\ 1/\mathrm{sqrt}(3) & 0 \\ 0 & 1/\mathrm{sqrt}(2) \\ 0 & 1/\mathrm{sqrt}(2) \end{bmatrix} \mathrm{x} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \mathrm{x}
$$

$$
\begin{bmatrix} 1/\mathrm{sqrt}(3) & 1/\mathrm{sqrt}(3) & 1/\mathrm{sqrt}(3) & 0 & 0 \\ 0 & 0 & 0 & 1/\mathrm{sqrt}(2) & 1/\mathrm{sqrt}(2) \end{bmatrix}
$$

SVD - Drill

- Q: How to check we are correct?

$$
\begin{bmatrix} 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\mathrm{sqrt}(3) & 0 \\ 1/\mathrm{sqrt}(3) & 0 \\ 1/\mathrm{sqrt}(3) & 0 \\ 0 & 1/\mathrm{sqrt}(2) \\ 0 & 1/\mathrm{sqrt}(2) \end{bmatrix} \text{x} \begin{bmatrix} 3 & 0 \\ 0 & 2 \end{bmatrix} \text{x}
$$

$$
\begin{bmatrix} 1/\mathrm{sqrt}(3) & 1/\mathrm{sqrt}(3) & 1/\mathrm{sqrt}(3) & 0 & 0 \\ 0 & 0 & 0 & 1/\mathrm{sqrt}(2) & 1/\mathrm{sqrt}(2) \end{bmatrix}
$$

SVD - Drill

- A: SVD properties:
  - matrix product should give back matrix **A**
  - matrix **U** should be column-orthonormal, i.e., columns should be unit vectors, orthogonal to each other
  - ditto for matrix **V**
  - matrix **Λ** should be diagonal, with positive values

SVD - Detailed outline

- Motivation
- Definition - properties
- Interpretation
- Complexity
- Case studies
- Additional properties

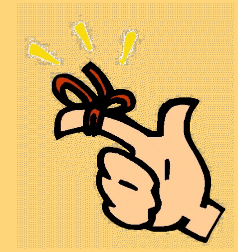
SVD - Complexity

- O(n * m * m) or O(n * n * m) (whichever is less)
- less work, if we just want singular values
- or if we want first *k* singular vectors
- or if the matrix is **sparse** [Berry]
- Implemented: in any linear algebra package (LINPACK, matlab, Splus/R, mathematica ...)

SVD - conclusions so far

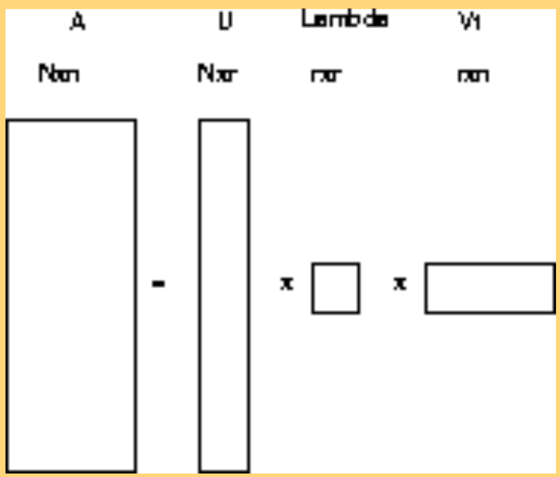
- SVD: **A= U Λ V**^{T} : unique (*)
- **U**: document-to-concept similarities
- **V**: term-to-concept similarities
- **Λ**: strength of each concept
- dim. reduction: keep the first few strongest singular values (80-90% of 'energy')
  - SVD: picks up linear correlations
- SVD: picks up non-zero 'blobs'

Carnegie Mellon



Conclusion

- How to find ‘concepts’ in a set of doc’s?
  – (~ clusters)
- SVD (= LSI) on the document-term matrix



References

- Berry, Michael: http://www.cs.utk.edu/~lsi/
- Fukunaga, K. (1990). Introduction to Statistical Pattern Recognition, Academic Press.
- Press, W. H., S. A. Teukolsky, et al. (1992) (3rd ed.: 2007): *Numerical Recipes in C*, Cambridge University Press.