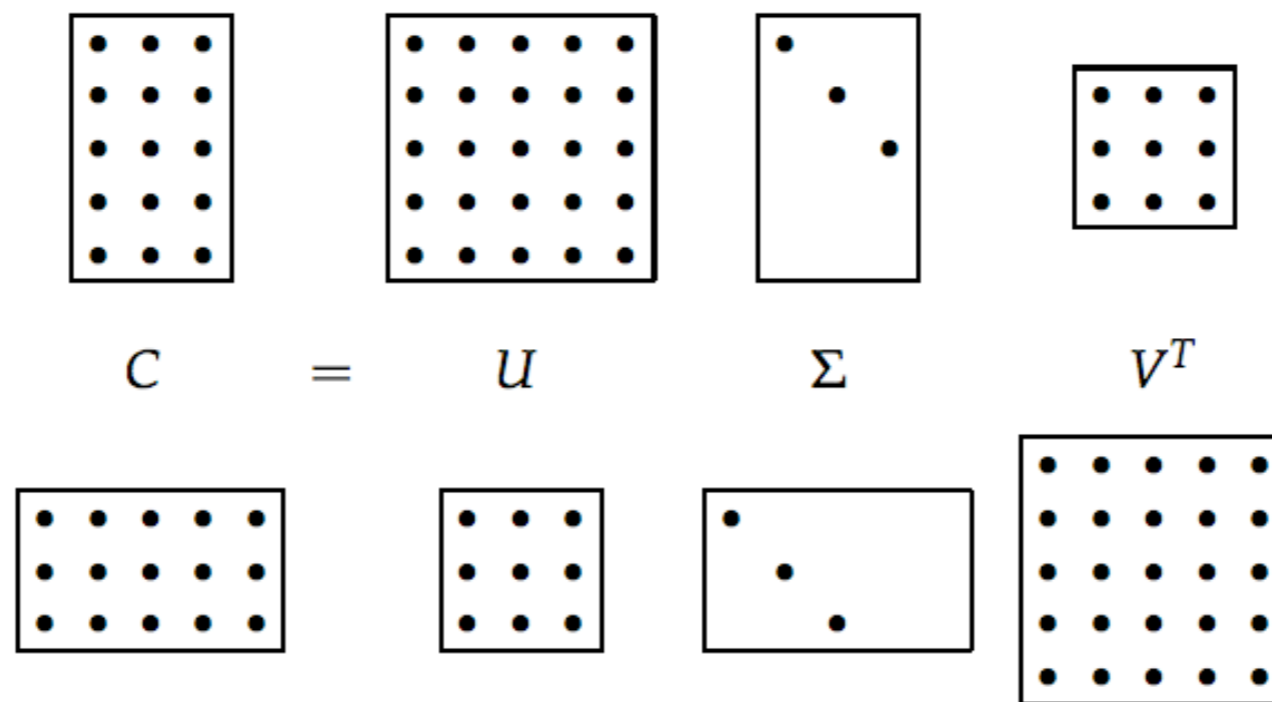


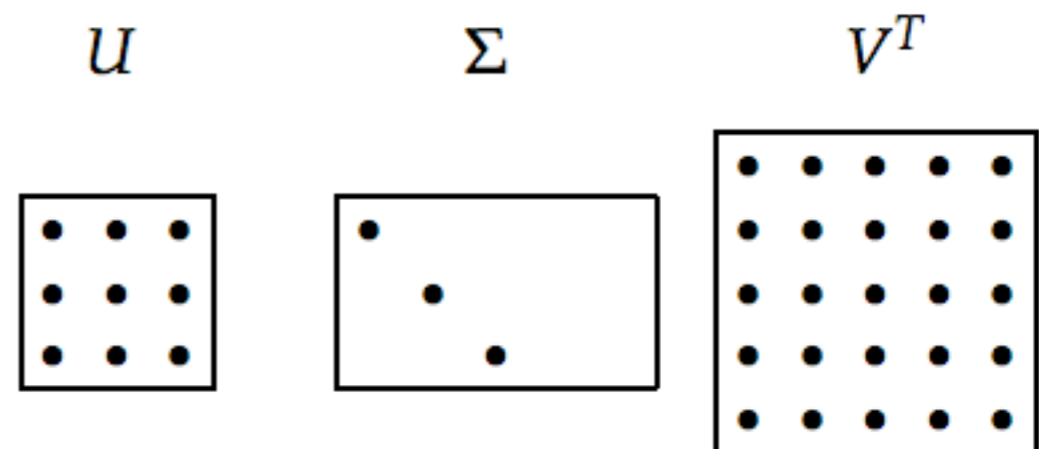
## Matrix Decomposition

- Singular Value Decomposition
  - Is a technique that splits a matrix into three components with these properties.



## Matrix Decomposition

- Singular Value Decomposition
  - SVD reduces the size of your term-document matrix
    - reduces the **dimensionality** or the **rank**
    - you can arbitrarily reduce the dimensionality by putting zeros in the bottom right of sigma
    - this is the mathematically optimal way of reducing dimensions



## Matrix Decomposition

- Singular Value Decomposition
- If the old dimensions were based on **terms**
  - after reducing the rank of the matrix the dimensionality is based on **concepts** or **semantics**
- a concept is a **linear combination** of terms

$$SVD_{dimension_1} = a * td_{dim_1} + b * td_{dim_2} + c * td_{dim_3} + d * td_{dim_4}$$

$$SVD_{dimension_2} = a' * td_{dim_1} + b' * td_{dim_2} + c' * td_{dim_3} + d' * td_{dim_4}$$

$$SVD_{dimension_3} = a'' * td_{dim_1} + b'' * td_{dim_2} + c'' * td_{dim_3} + d'' * td_{dim_4}$$



## Matrix Decomposition

- Singular Value Decomposition

$$SVD_{dimension_1} = a * td_{dim_1} + b * td_{dim_2} + c * td_{dim_3} + d * td_{dim_4}$$

$$SVD_{dimension_2} = a' * td_{dim_1} + b' * td_{dim_2} + c' * td_{dim_3} + d' * td_{dim_4}$$

$$SVD_{dimension_3} = a'' * td_{dim_1} + b'' * td_{dim_2} + c'' * td_{dim_3} + d'' * td_{dim_4}$$

$$\begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix}$$



## Matrix Decomposition

- Singular Value Decomposition

$$SVD_{dimension_1} = a * td_{dim_1} + b * td_{dim_2} + c * td_{dim_3} + d * td_{dim_4}$$

$$SVD_{dimension_2} = a' * td_{dim_1} + b' * td_{dim_2} + c' * td_{dim_3} + d' * td_{dim_4}$$

$$SVD_{dimension_3} = a'' * td_{dim_1} + b'' * td_{dim_2} + c'' * td_{dim_3} + d'' * td_{dim_4}$$

$$\begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix} * \begin{vmatrix} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{vmatrix}$$



## Matrix Decomposition

- Singular Value Decomposition

$$SVD_{dimension_1} = a * td_{dim_1} + b * td_{dim_2} + c * td_{dim_3} + d * td_{dim_4}$$

$$SVD_{dimension_2} = a' * td_{dim_1} + b' * td_{dim_2} + c' * td_{dim_3} + d' * td_{dim_4}$$

$$SVD_{dimension_3} = a'' * td_{dim_1} + b'' * td_{dim_2} + c'' * td_{dim_3} + d'' * td_{dim_4}$$

$$\begin{vmatrix} SVD_{dim_1} \\ SVD_{dim_2} \\ SVD_{dim_3} \end{vmatrix} = \begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix} * \begin{vmatrix} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{vmatrix}$$



## Matrix Decomposition

- Singular Value Decomposition

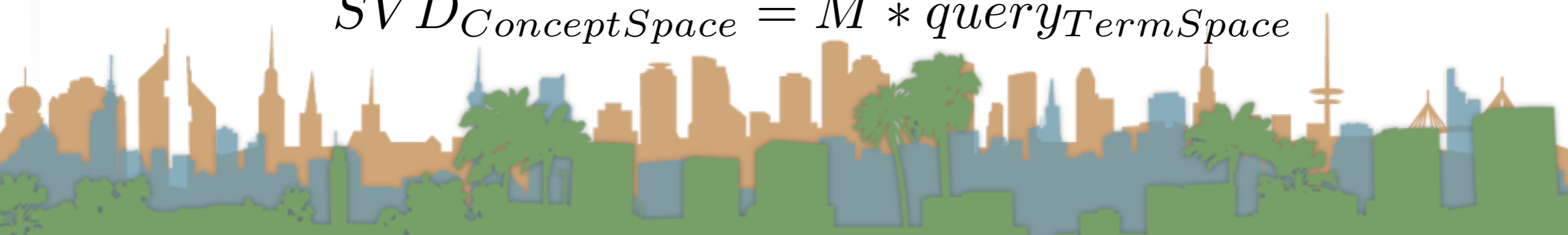
$$SVD_{dimension_1} = a * td_{dim_1} + b * td_{dim_2} + c * td_{dim_3} + d * td_{dim_4}$$

$$SVD_{dimension_2} = a' * td_{dim_1} + b' * td_{dim_2} + c' * td_{dim_3} + d' * td_{dim_4}$$

$$SVD_{dimension_3} = a'' * td_{dim_1} + b'' * td_{dim_2} + c'' * td_{dim_3} + d'' * td_{dim_4}$$

$$\begin{vmatrix} SVD_{dim_1} \\ SVD_{dim_2} \\ SVD_{dim_3} \end{vmatrix} = \begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix} * \begin{vmatrix} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{vmatrix}$$

$$SVD_{ConceptSpace} = M * queryTermSpace$$



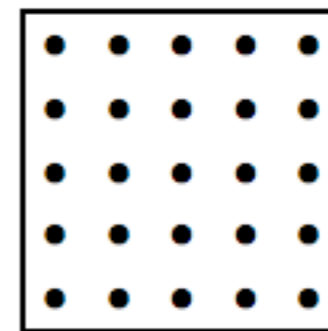
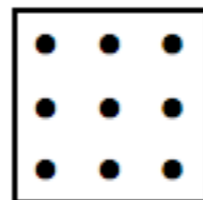
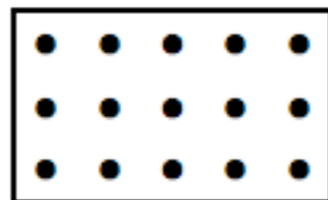
## Matrix Decomposition

- Singular Value Decomposition

$$\begin{array}{|l} SV D_{dim_1} \\ SV D_{dim_2} \\ SV D_{dim_3} \end{array} = \begin{array}{|l} a \quad b \quad c \quad d \\ a' \quad b' \quad c' \quad d' \\ a'' \quad b'' \quad c'' \quad d'' \end{array} * \begin{array}{|l} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{array}$$

$$SV D_{ConceptSpace} = M * query_{TermSpace}$$

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





## Matrix Decomposition

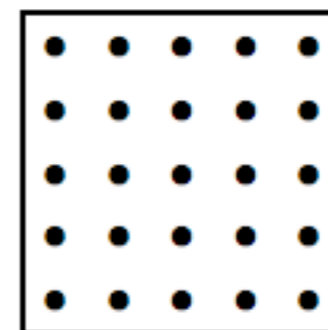
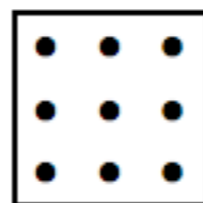
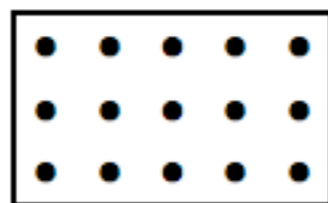
- Singular Value Decomposition

$$\begin{array}{|l} SV D_{dim_1} \\ SV D_{dim_2} \\ SV D_{dim_3} \end{array} = \begin{array}{|l} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{array} * \begin{array}{|l} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{array}$$

$$SV D_{ConceptSpace} = M * query_{TermSpace}$$

$$M = \Sigma_k^{-1} U_k^T$$

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



## Matrix Decomposition

- Singular Value Decomposition

$$\begin{vmatrix} SVD_{dim_1} \\ SVD_{dim_2} \\ SVD_{dim_3} \end{vmatrix} = \begin{vmatrix} a & b & c & d \\ a' & b' & c' & d' \\ a'' & b'' & c'' & d'' \end{vmatrix} * \begin{vmatrix} td_{dim_1} \\ td_{dim_2} \\ td_{dim_3} \\ td_{dim_4} \end{vmatrix}$$

$$SVD_{ConceptSpace} = M * query_{TermSpace}$$

$$M = \Sigma_k^{-1} U_k^T$$

$$query_{ConceptSpace} = \Sigma_k^{-1} U_k^T query_{TermSpace}$$



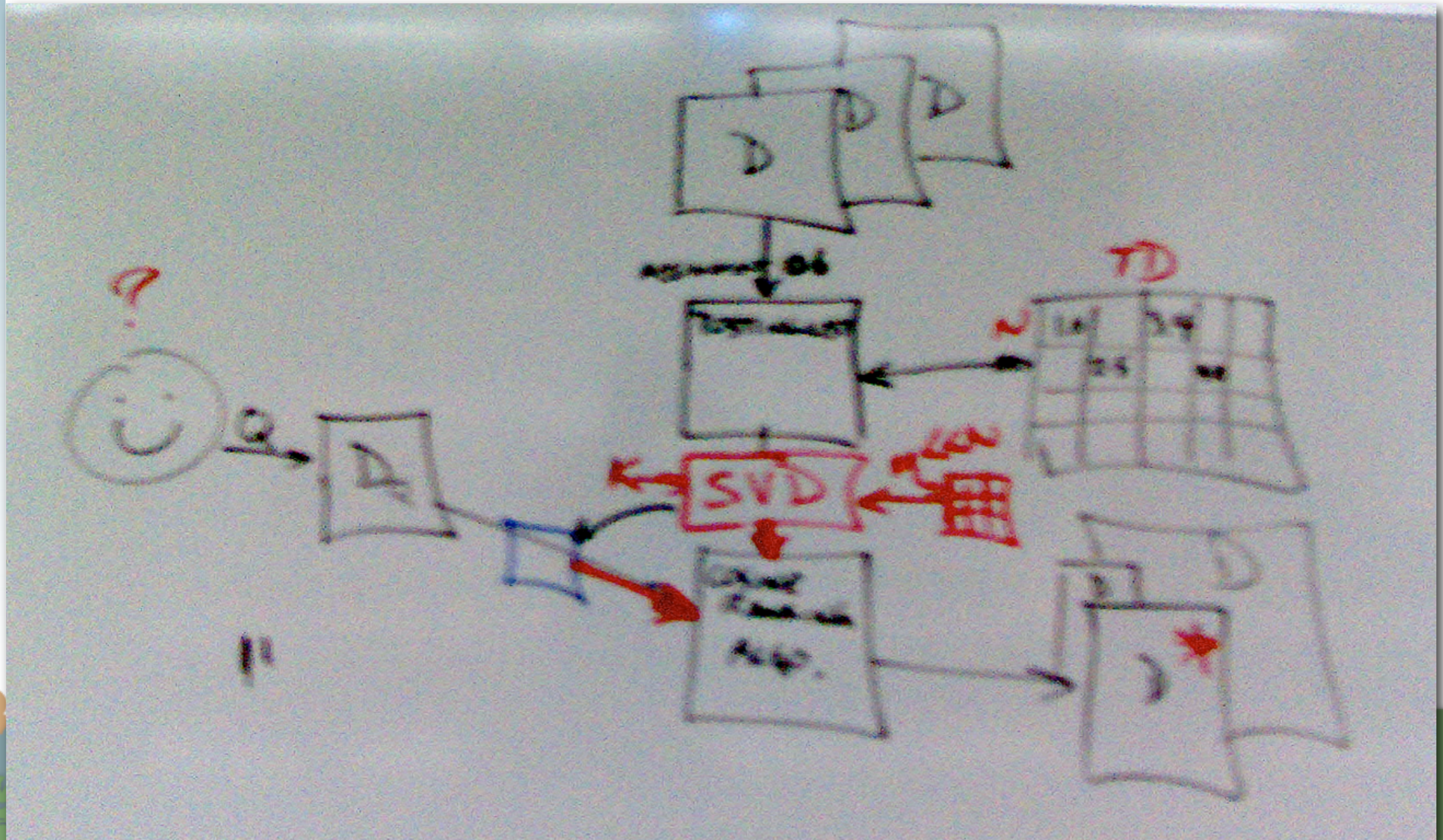
## Matrix Decomposition

- Singular Value Decomposition
  - SVD is an algorithm that gives us  $\Sigma U V^T$
  - With these quantities we can reduce dimensionality
  - With reduced dimensionality
    - **synonyms are mapped onto the same location**
      - “bat” “chiroptera”
    - **polysemies are mapped onto different locations**
      - “bat” (baseball) vs. “bat” (small furry mammal)



# Matrix Decomposition

- Big Picture



## Matrix Decomposition

- “I am not crazy”
- Netflix

**NETFLIX**

# Netflix Prize

Home Rules Leaderboard Register Update Submit Download

**NETFLIX**

Browse Recommendations Friends Queue Buy DVDs

Home Genres New Releases Previews Netflix Top 100 Crit

## Movies For You

Randy, the following movies were chosen based on your interest in:  
[Bowling for Columbine](#)  
[Carnivale: Season 1](#)  
[Fahrenheit 9/11](#)

**The Big One**  
★★★★☆  
A rivetingly creative series continues with a documentary about the lives of a motley crew of filmmakers who've made the Oscar shortlist.

**Carnivale: Season 2**  
★★★★★  
Disc Series

**Red Eye**  
★★★★☆  
In this biographical satire, a woman is stalked by a man who is obsessed with her.

**Rear Window**  
★★★★★  
A classic film noir about a man who becomes convinced his neighbor is having an affair.

**You really liked it...**  
Now only for just \$5.99  
Shop as low as low titles

**Original artwork**

**Guides:**  
Member Favorites  
Easter Eggs  
By Decade  
By Studio  
Movies You've Seen

**Welcome!**

The Netflix Prize seeks to substantially improve the accuracy of predictions about how much someone is going to love a movie based on their movie preferences. Improve it enough and you win one (or more) Prizes. Winning the Netflix Prize improves our ability to connect people to the movies they love.

Read the [Rules](#) to see what is required to win the Prizes. If you are interested in joining the quest, you should [register a team](#).

You should also read the [frequently-asked questions](#) about the Prize. And check out how various teams are doing on the [Leaderboard](#).

Good luck and thanks for helping!

Give a friend

## Matrix Decomposition

- “I am not crazy”
- Netflix
- Machine translations
  - Just like “bat” and “chiroptera” map the same
  - “bat” and “murciélago” can map to the same thing

The math is hard but it's beautiful and powerful

La matemáticas es dura pero es hermosa y de gran  
alcance

Jene mathematisch ist hart, aber ist und an  
langer Reichweite schön

That one mathematically is hard, but is beautiful and at long  
range

