Artificial Intelligence

Instructor: Jie Shen

Department of Computer Science

Last time: PCA

$X = (x_1 \; \cdots \; x_n)$ $\quad d \times n$

Find M that best fits X

rank(M) is small

$$\Rightarrow \begin{cases} \min\limits_{M} \; \| M - X \|_F^2 \\ \text{s.t.} \quad \text{rank}(M) \le k. \end{cases}$$

Can be solved by SVD

$[U, S, V] = \text{SVD}(X)$

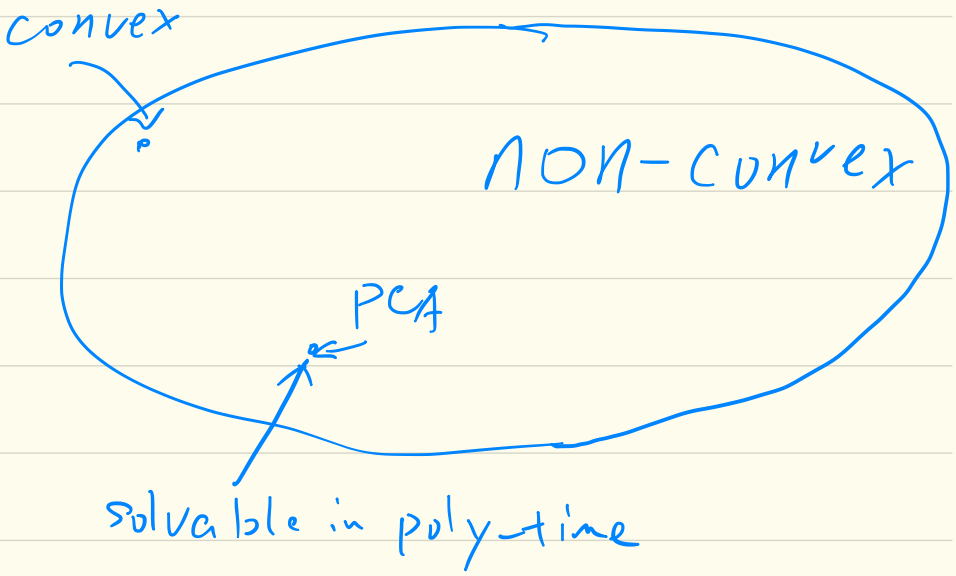
Assume: diagonal elements in S in descending order (default in python/matlab)

U_k = first k columns in U

$X \rightarrow U_k^T \cdot X$

$d \times n$ $\quad$ $k \times n$

↑ new data



$$N = \{0, 1, 2, 3, \cdots \infty\}$$

$$|N| = \infty$$

$\underline{R}$

$$|R| = \infty$$

$$|N| < |R|$$

Q: How to compare size of a set ?

A B

finite < infinite

$\{0, 1\}$ N

finite finite

→ count

Count:

$A = \{1, 3, 10, 7\}$

$1 \quad 2 \quad 3 \quad 4$

any finite set $\Leftrightarrow \{1, 2, \cdots k\}$

k

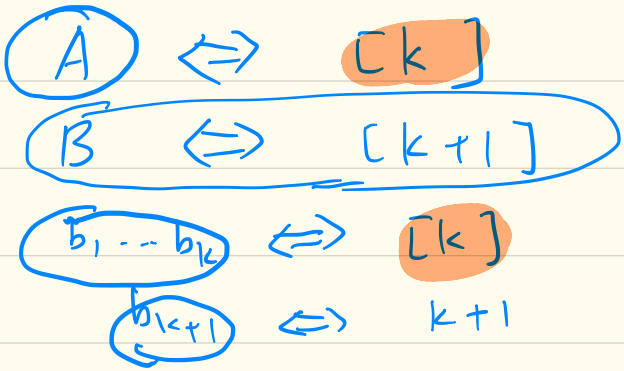
$A \overset{\text{1-to-1}}{\Longleftrightarrow} [k]$

$B \overset{\text{1-to-1}}{\Longleftrightarrow} [k]$

$\Rightarrow \quad A \overset{\text{1-to-1}}{\Longleftrightarrow} B$

A, B finite. $|A| = k$

$|B| = k+1$



$A \Leftrightarrow \{b_1 \cdots b_k\}$

but no ele in $A \Leftrightarrow b_{k+1}$

$|A| = \infty$, $|B| = \infty$

if ∃ 1-to-1 mapping

$A \Leftrightarrow B$.

$|A| = |B|$

$|\mathbb{N}| = |2\mathbb{N}|$

if no mapping f,

such that $\exists$ some $b \in B$

cannot be mapped from $a \in A$

then $|A| < |B|$

$A = \{0, 1\}$

$B = \{0, 1, 2\}$

Cantor Theorem. Power set of A is always $>$ A

$A = \{0, 1\}$

$P(A) = \{\phi, \{0\}, \{1\}, \{0, 1\}\}$

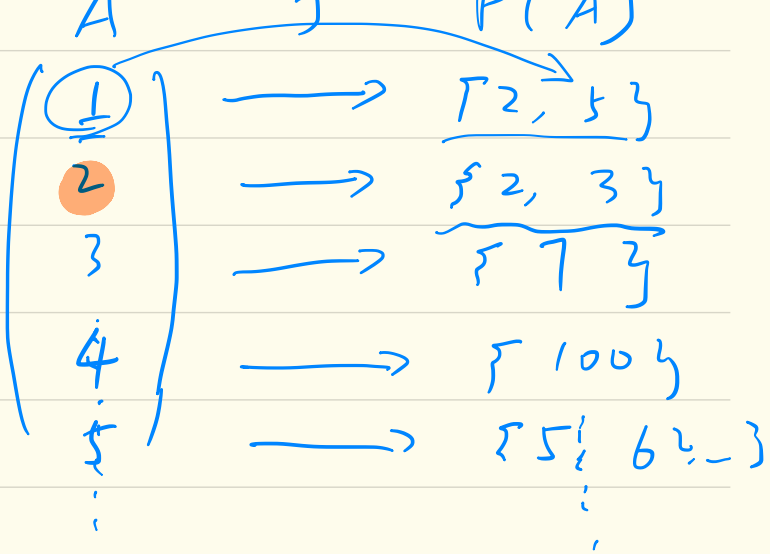
Assume for contradiction: $\exists$ 1-to-1 $f: A \to P(A)$

$A = \{a_1, \dots a_n\}$

$P(A) = \{\{a_1\}, \{a_2\}, \dots \{a_n\}, \phi\}$

$|A| = \infty$, $|P(A)| > |A|$

Proof:



Selfish number: $a \in A$.

$a \in f(a)$ $\in P(A)$

"2" is selfish "5" is

"1" not .

$A = A_1 \cup A_2$

$A_1 = \{$ all selfish numbers $\}$

$A_2 = \{$ all non-selfish $\}$

$A_1 \subset A$, $A_2 \subset A$

$\Rightarrow A_2 \in P(A)$

$A \xrightarrow{f} P(A)$

$a \longrightarrow A_2$

Case 1: $a \in A_2$ contradiction

$\Rightarrow a$ is selfish

$\Rightarrow a \in A_1$

$\Rightarrow a \notin A_2$

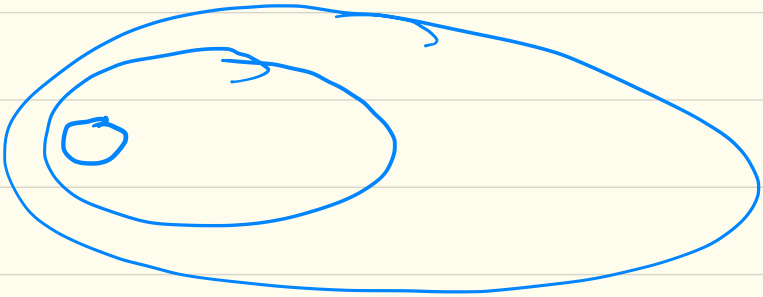
Case 2: $a \notin A_2$

$\Rightarrow a$ is non-selfish

$\Rightarrow a \in A_2$

$\Rightarrow$ No 1-to-1 mapping from A to $P(A)$.

$\Rightarrow$ $|P(A)| > |A|$ □

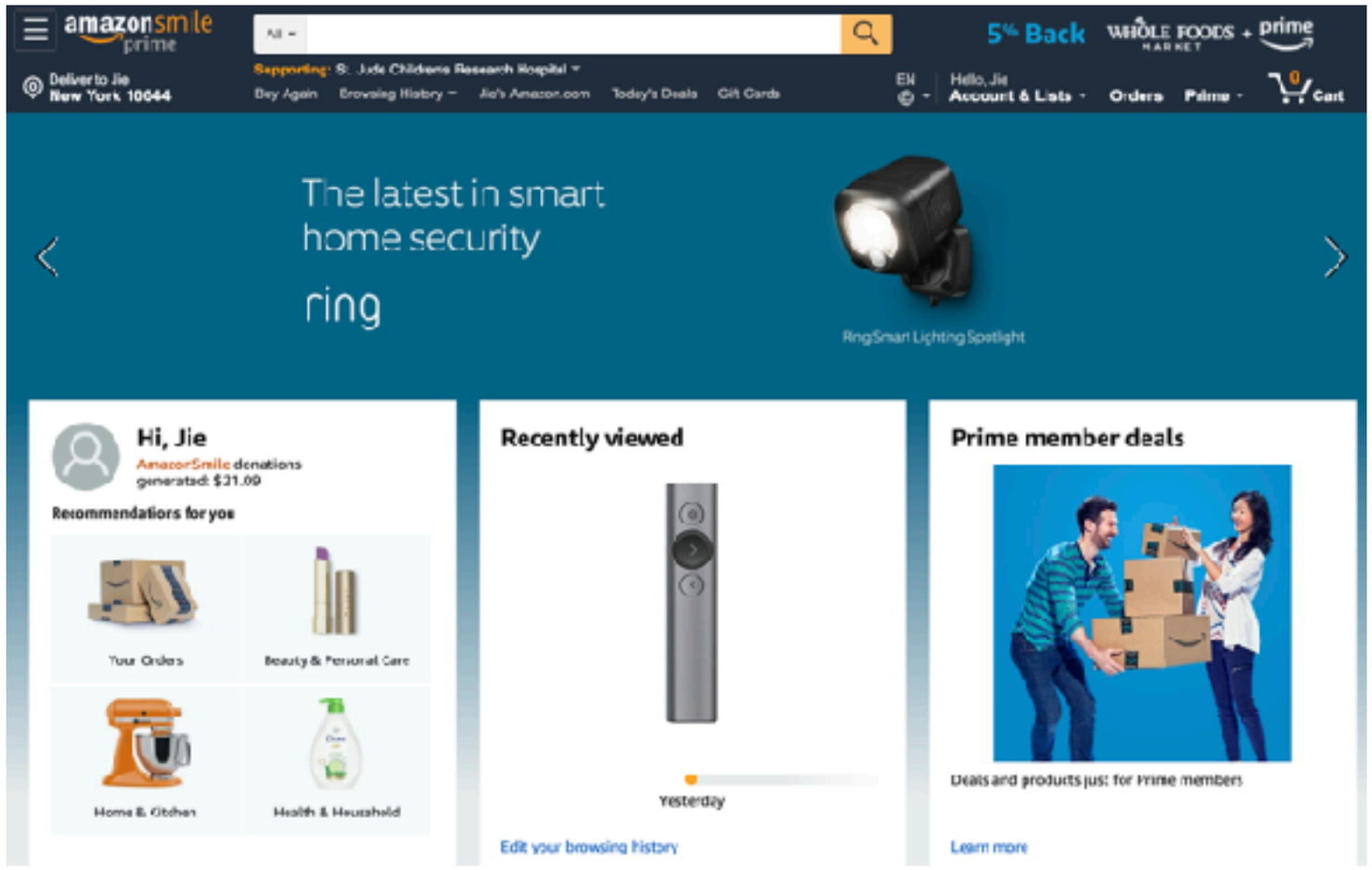


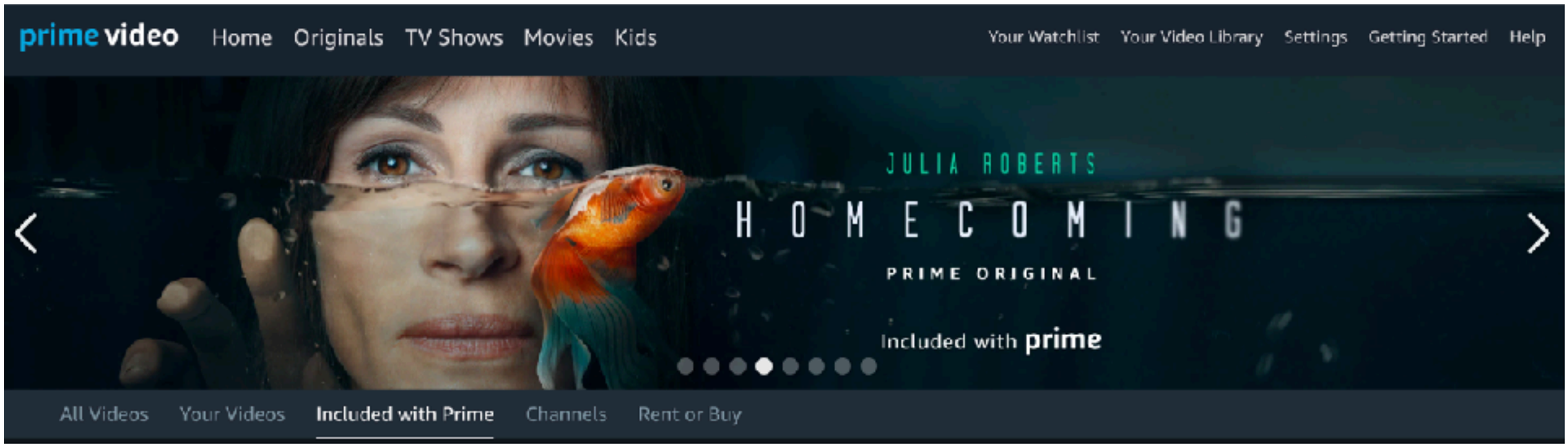
No largest set

$|\mathbb{R}| = |P(\mathbb{N})|$

$\rightsquigarrow$ Cantor.

Background


amazonsmile
prime
Deliver to Jie
New York 10044
Supporting St. Jude Children's Research Hospital
Buy Again
Browsing History
Jie's Amazon.com
Today's Deals
Gift Cards
5% Back
WHOLE FOODS MARKET + prime
EN
Hello, Jie
Account & Lists
Orders
Prime
0
Cart
The latest in smart home security
ring
RingSmart Lighting Spotlight
Hi, Jie
AmazonSmile donations generated: $21.09
Recommendations for you
Your Orders
Beauty & Personal Care
Home & Kitchen
Health & Household
Recently viewed
Yesterday
Edit your browsing history
Prime member deals
Deals and products just for Prime members
Learn more


prime video
Home
Originals
TV Shows
Movies
Kids
Your Watchlist
Your Video Library
Settings
Getting Started
Help
JULIA ROBERTS
HOMECOMING
PRIME ORIGINAL
Included with prime
All Videos
Your Videos
Included with Prime
Channels
Rent or Buy

Background

amazonsmile

Delivered Jan 12, 2019

Oral-B White Pro 1000 Power Rechargeable Electric Toothbrush, Powered by Braun

Sold by: Amazon.com Services, Inc

Return window closed on Feb 11, 2019

$39.94

Buy it again

Get product support

Order details

Write a product review

Archive order

10 Braun Oral-B Cross Action Replacement Toothbrush Heads by Oral-B

by Oral B

106

$36.99

prime | FREE One-Day.

Oral-B FlossAction Electric Toothbrush Replacement Brush Heads Refill (2 count)

by Oral-B

2,886

$14.88 $17.15

Save 5% more with Subscribe & Save

prime | FREE One-Day.

Toothbrush Replacement Heads Refill for Oral-B Electric Toothbrush, 4 Floss Action, 4 cross action, 8 Count

by MelMel

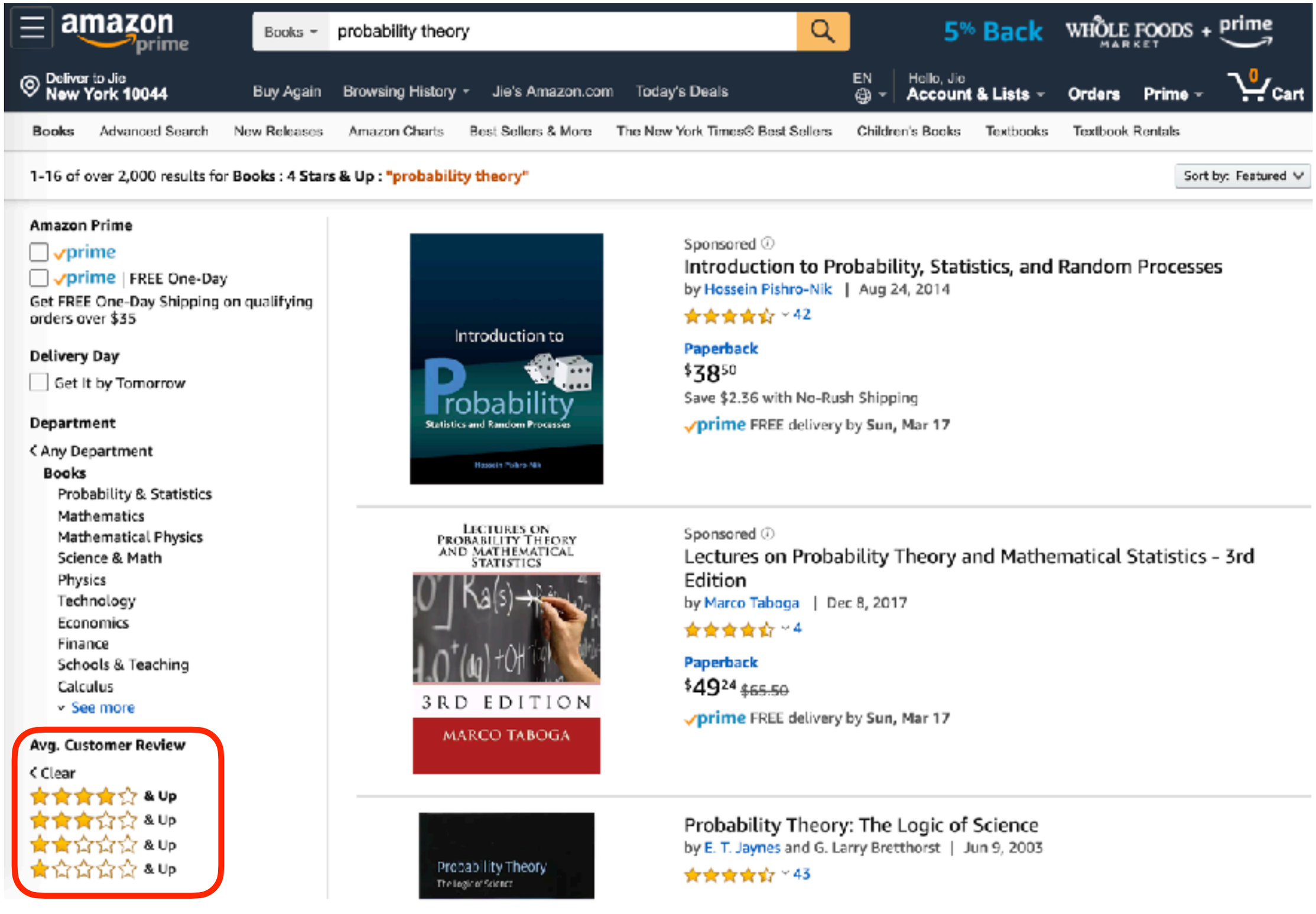
95

$10.89

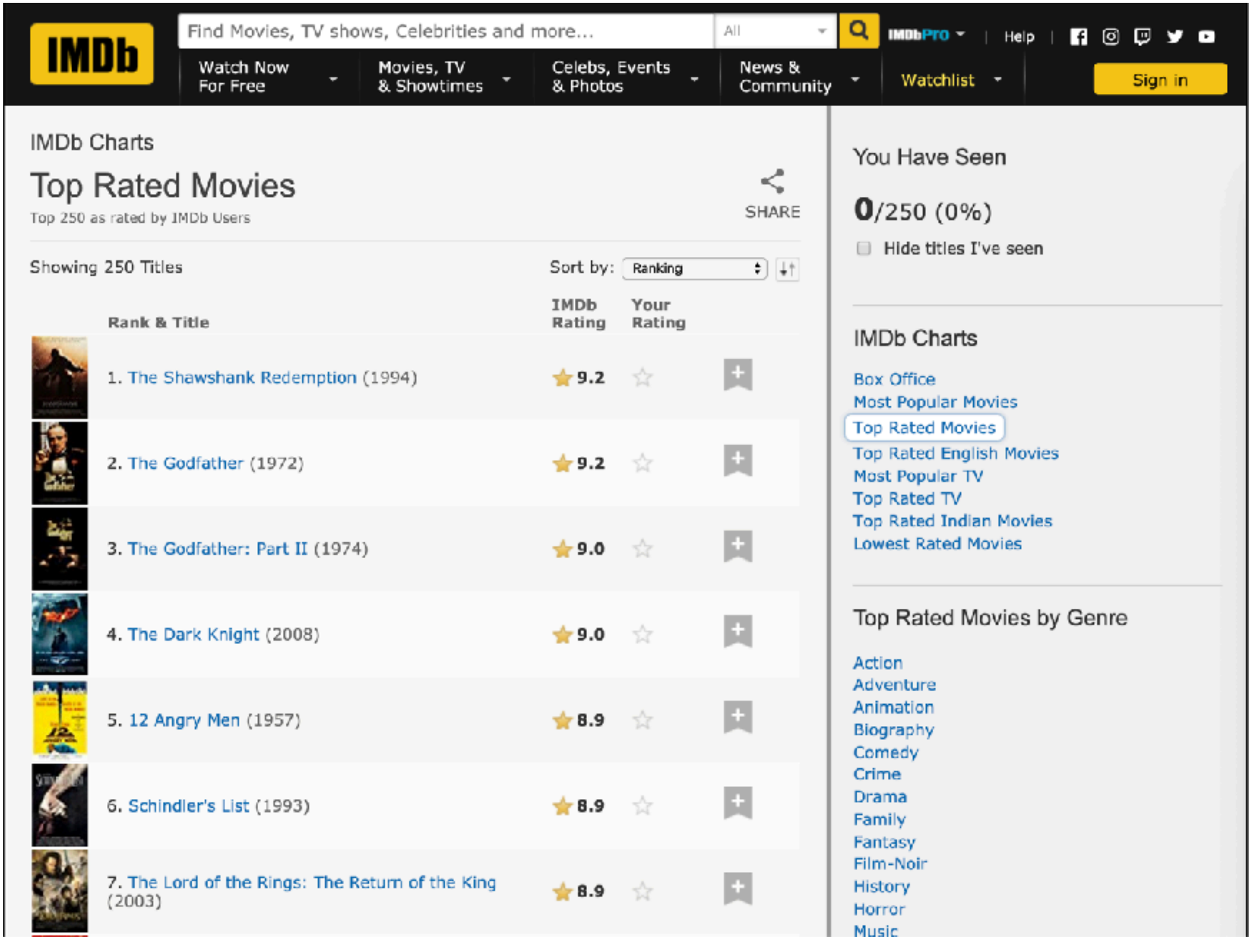
Save 5% with coupon

prime | FREE One-Day.

Rating System: 1 Star - 5 Stars



IMDb: 1 - 10 stars



"Real-Valued" Data

rating matrix / data matrix

User 1 rates "5" to item p

$\Rightarrow Z_{1,p} = 5$

User 2 rate '1' to item 1

$\Rightarrow Z_{2,1} = 1$

Z is very sparse.

1B

@ 1k → 20

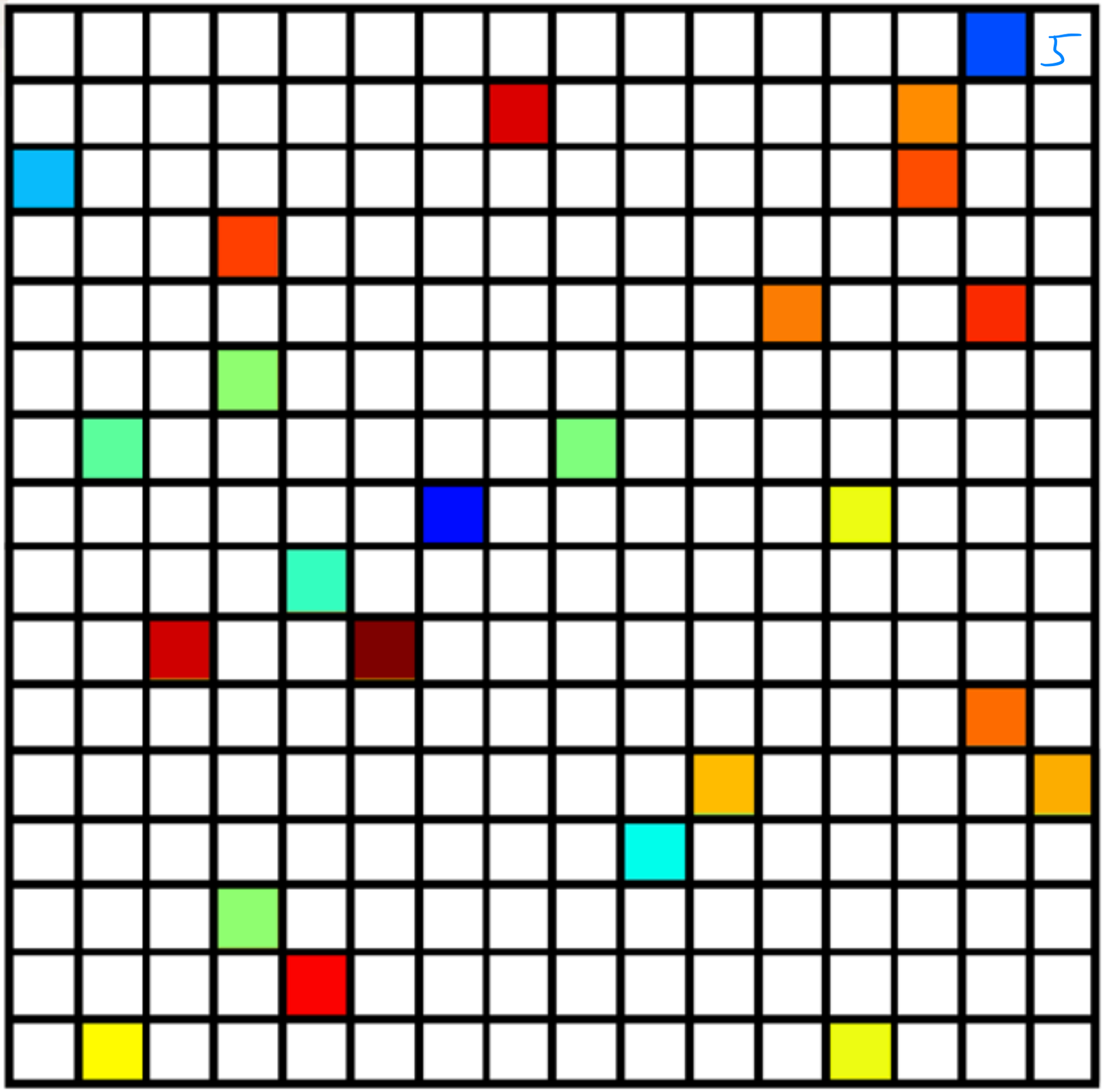
20/1B

Item 1 . . . . . Item p

User 1

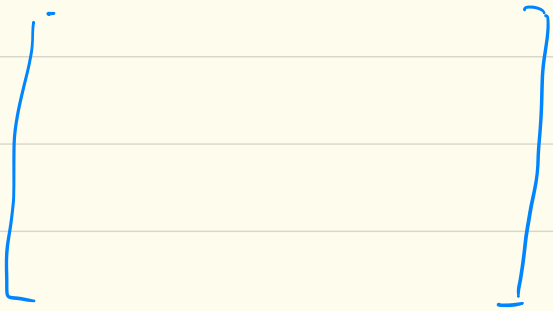
User 2

User n



Z

In 2010. "Netflix Prize"



$\underline{f_0}$ → observed.

estimate $g \approx g_0$.

↓

2013.

Collaborative Filtering for 1 - 5 Rating

- Traditional collaborative filtering (CF):

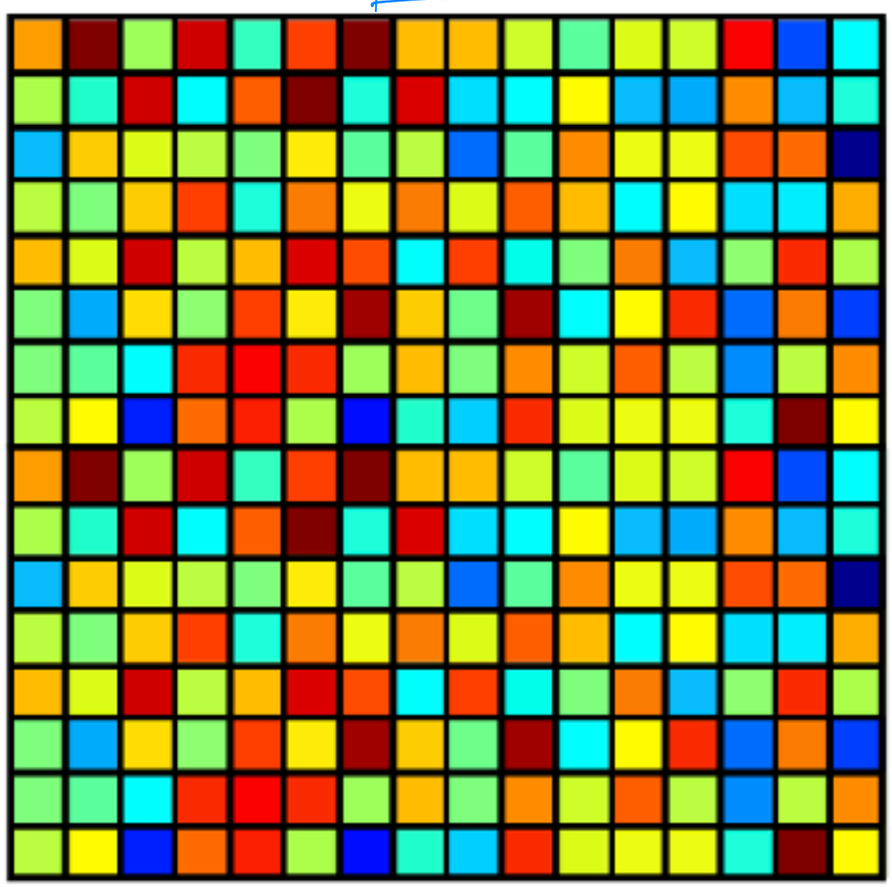
NP-hard.

$$\min_X \|(Z - X)_\Omega\|_F^2, \quad \text{s.t. } \text{rank}(X) \le r$$

Z is data matrix

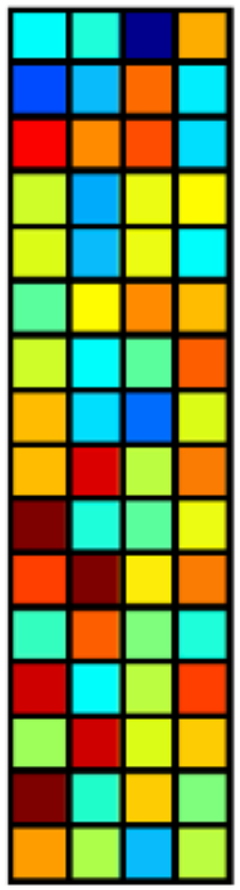
1M x 1M

r = 20.

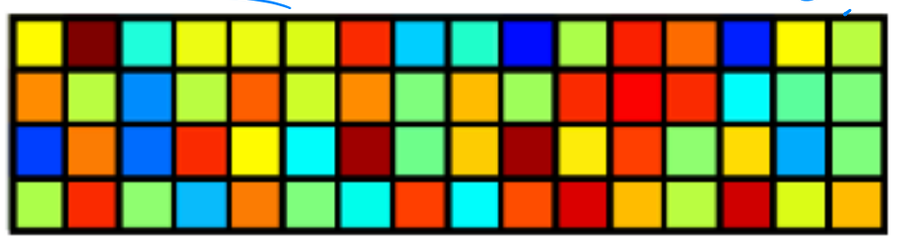


X

$=$



U



V

$\Omega = \{(i,j), Z_{ij}$ is observed $\}$.

$Z =$

$\begin{pmatrix} ? & 1 & 2 \\ 3 & ? & ? \\ ? & 5 & ? \end{pmatrix}$

$\Omega = \{(1,2), (1,3), (2,1), (3,2)\}$

known

- CF: $\min_{U,V} \|(Z - UV^\top)_\Omega\|_F^2 + \lambda(\|U\|_F^2 + \|V\|_F^2)$

$$Z = \begin{pmatrix} 3 & 1 & 2 \\ & & 5 \end{pmatrix}.$$

$$\min_X \; (X_{12}-1)^2 + (X_{13}-2)^2 + (X_{21}-3)^2 + (X_{32}-5)^2 \rightarrow \|(X-Z)_{\Omega}\|_F^2$$

$$\text{s.t. } \operatorname{rank}(X) \le r$$

4:45

X is full, real-preference to all items.

low-rank

Nintendo		
Mario -1	-2	-3
5	5	5
1	1	1

$$\begin{cases} \min_{X} \ \|(Z - X)_{\Omega}\|_F^2 \\ \text{s.t.} \quad \operatorname{rank}(X) \le r \end{cases}$$

Step 1. $X = U \cdot V^T$ $\quad n \times p$

U: $n \times r$, V^T: $r \times p$

Step 2:

$$\min_{U,V} \|(Z - UV^T)_{\Omega}\|_F^2$$

Step 3: run gradient descent on U, V.

2014 Stoc

~~$X \sim N(0,1)$~~

$X = U \cdot V^T$

$\rightarrow$ $U, V \sim N(0, I)$

$\min (1 - UV)^2$

↓

non-convex

$|\Omega| \ge r^2 \cdot \text{poly} \log n$

What's the Problem?

Customer reviews

8

5.0 out of 5 stars

5 star	100%
4 star	0%
3 star	0%
2 star	0%
1 star	0%

Linear Operator Theory in Engineering and Science (Applied Mathematical Sci...

by Arch W. Naylor

Format: Paperback | Change

Price: $117.56 ✓prime

Write a review

Look inside ↓

Char

i would give 6 stars!

February 26, 2012

Format: Paperback | **Verified Purchase**

I'm doing a PhD in econometrics and I need to apply operator theories in constructing a linear or nonlinear operator to help explain individual economic behaviour. This book contains numerous useful ideas and applications with exercises thoroughly designed; one of the questions in the exercise gave me an idea of creating a matrix for describing a nonlinear operator. That question asks for a matrix that describes a second order differential operator and that gave me an idea that taylor series approximation can be used to linearise a nonlinear operator and hence a nonlinear operator may also be described by a matrix.

Binary Feedback

Lorenzo Martin

★★★★★ **Actually impressed.**

October 17, 2018

Color: Heather Gray | Configuration: Echo Dot | **Verified Purchase**

I was debating between this and the regular echo. I've had both and we both know what the echo does so all I'll talk about is the sound quality.

I'm going to get to the chase. This is the BEST speaker I've ever heard for under $50. The functions of the echo along with the music playing functionality and the sound quality makes it a steal at this price point.

The sound is great for its size. It can get pretty loud but the bass is definitely lacking. The bass isn't god awful but it just isn't there.

Now I will say this... if your primary motivation for a speaker is to listen to music, go with the echo. If it is but your budget is under $50 then go with the dot. It won't blow anyone away during a get together in your living room but it'll get the job done. The dot will only flourish is an enclosed space. It needs walls nearby to bounce off to cover up the weakness of the bass.

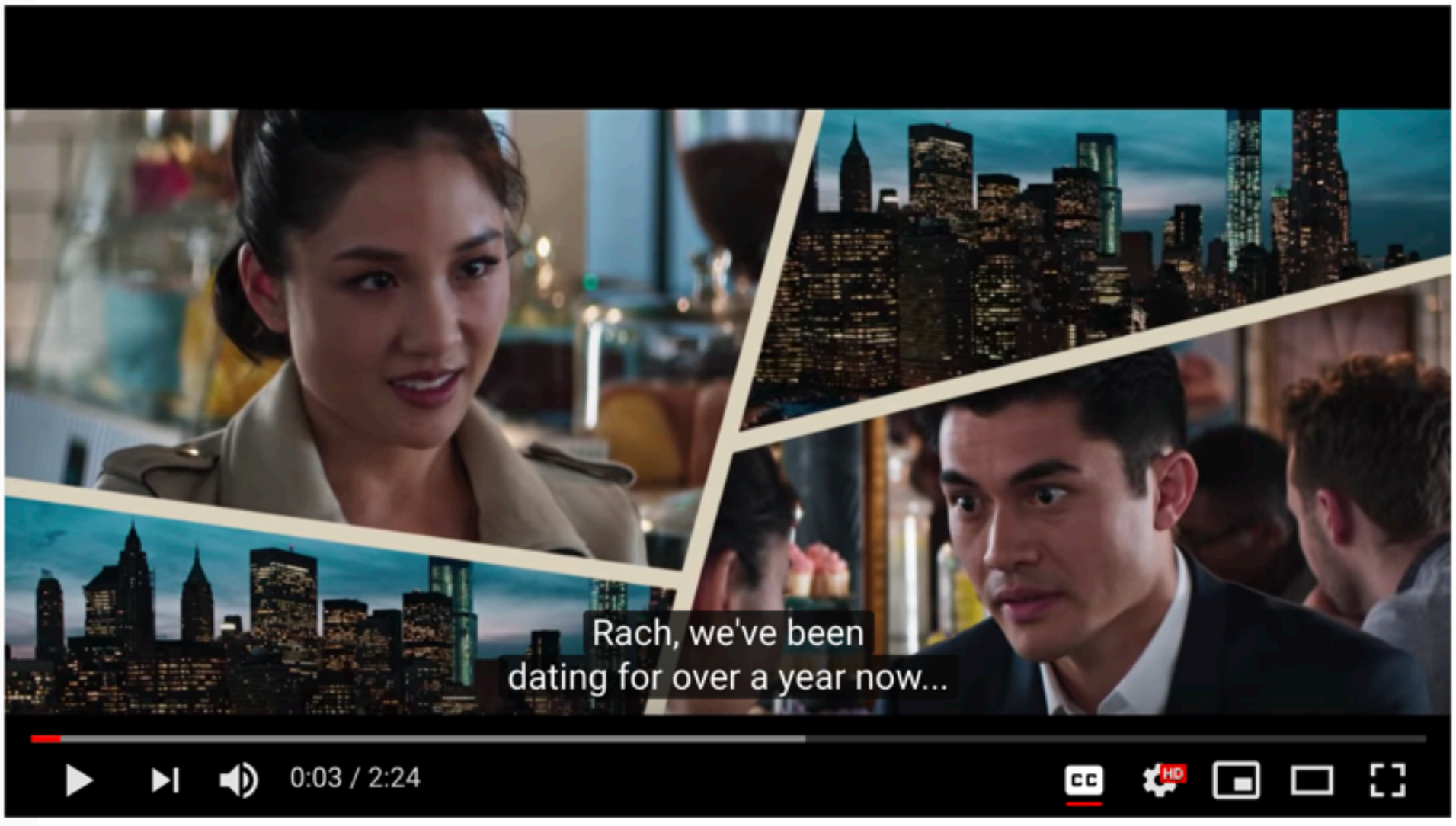
If I had to compare the sound to another speaker, I'd say the UE Wonderboom or the older Beats Pill, all of which suffer from weak bass. For the size, I was impressed by the sound but with the Echo at its refurbished price of $69 I'll have to aim towards that direction.

I'd recommend this for someone who needs an echo for their kitchen, bedroom or bathroom. It's just not going to hold down a living room unfortunately but overall, I like the sound and if I need a second echo, I'll definitely get a dot.

234 people found this helpful

Helpful | Comment | Report abuse

Binary Feedback

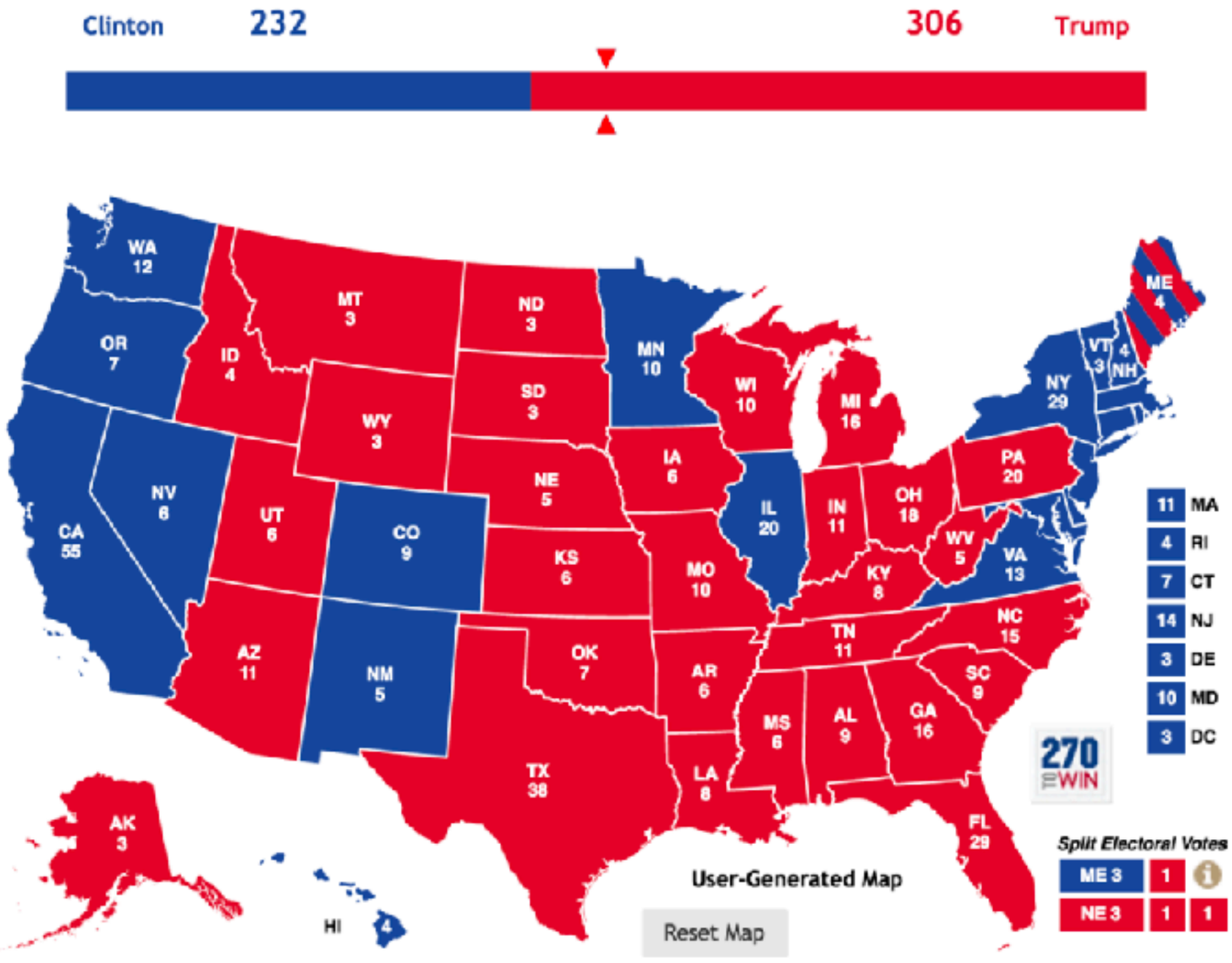


#CrazyRichAsians

CRAZY RICH ASIANS - Official Trailer

23,910,027 views 170K 9K SHARE SAVE

Binary Feedback



Binary Feedback

- Market survey
- Student response data
- And more...

10/8: guest lecture
(contrastive learning!)

10/15 no class.

10/22 mid-term

Prob. RP. PCA. RecSys.

Binary Feedback: Upside

- Ease the process of data acquisition

- Save the storage

But...

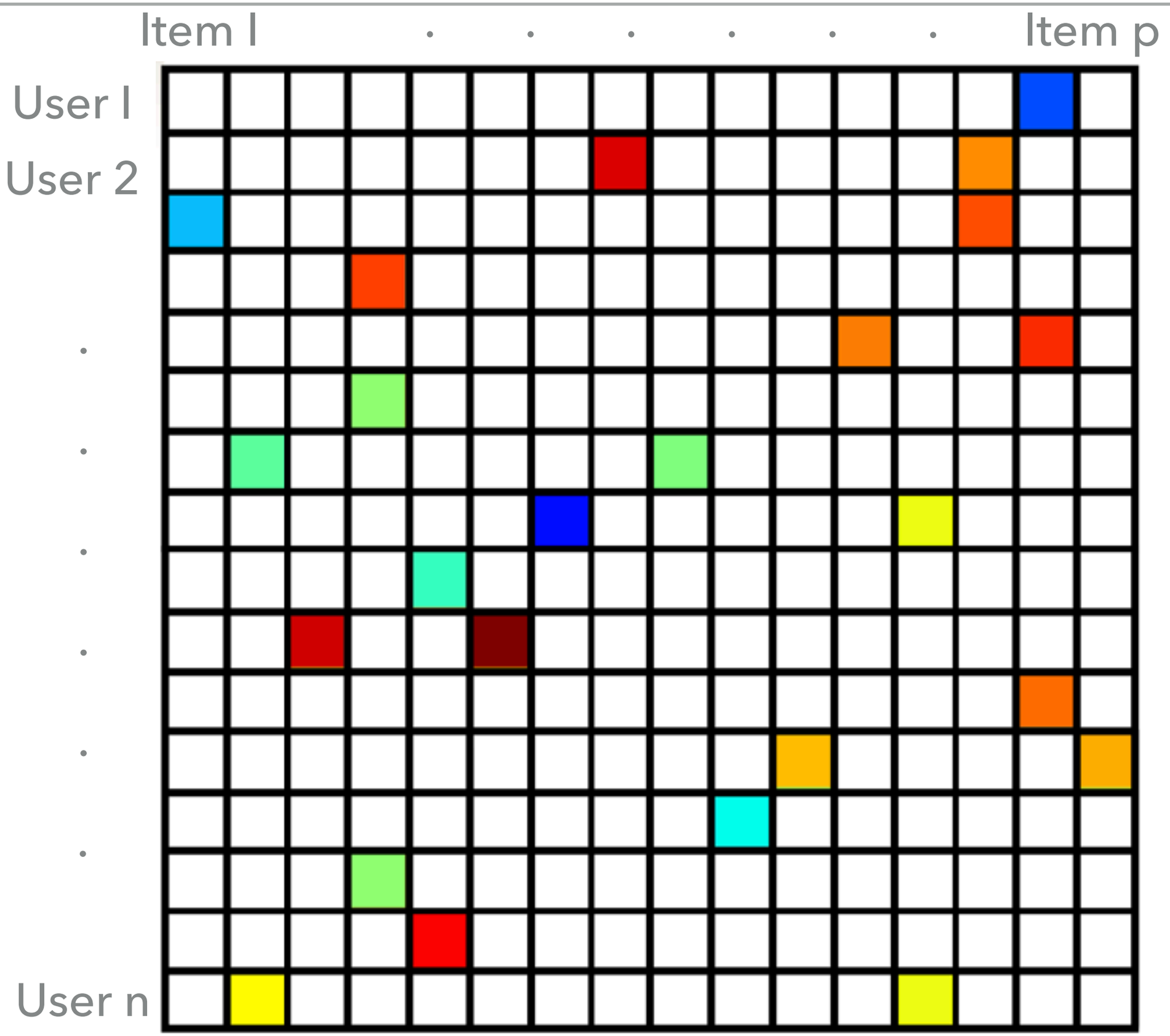
- Binary feedback *destroys low-rank structure*

$$\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix} \longrightarrow \begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$$

Rank 1

Rank 2 - full rank !

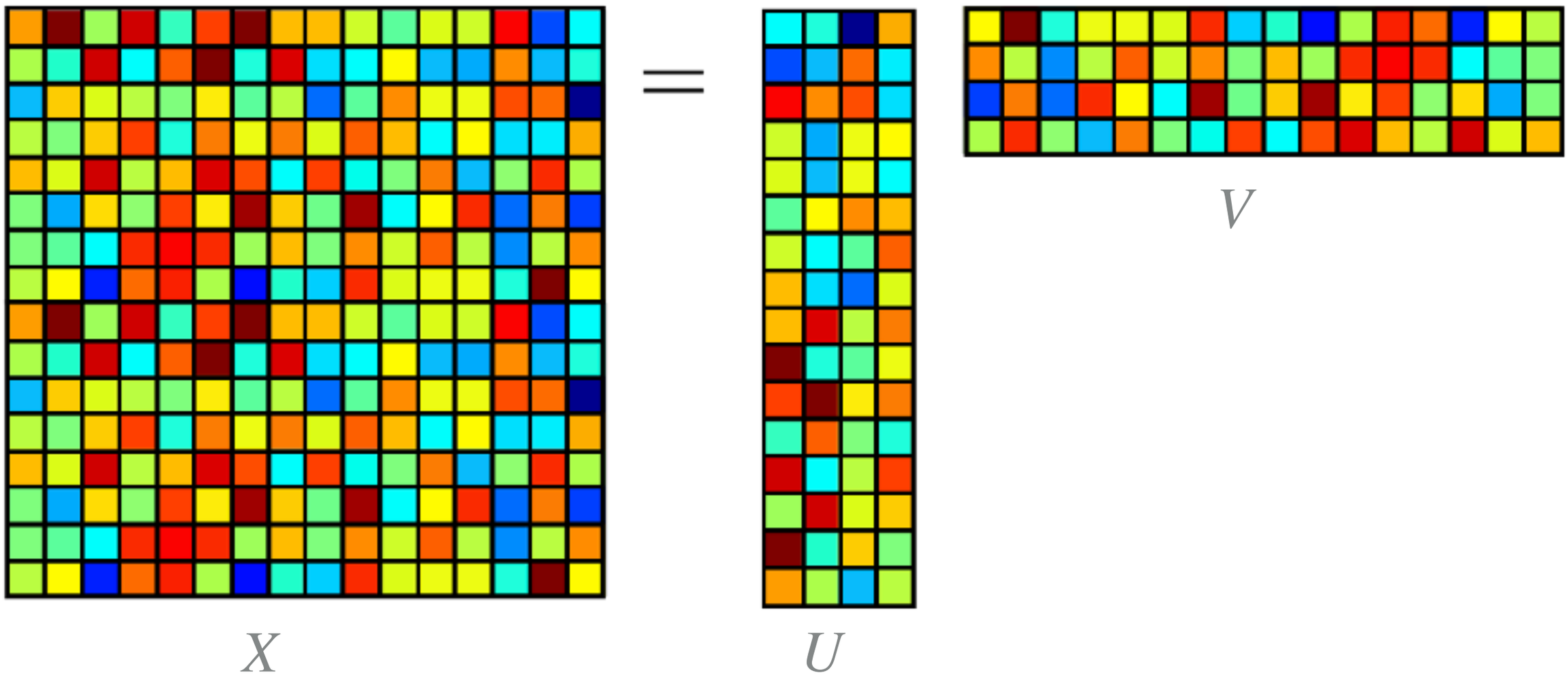
1 - 5 Stars Data



Collaborative Filtering for 1 - 5 Rating

- Traditional collaborative filtering (CF):

$$\min_{X} \|(Z - X)_{\Omega}\|_F^2, \quad \text{s.t. } \text{rank}(X) \leq r$$



- CF: $\min_{U,V} \|(Z - UV^{\top})_{\Omega}\|_F^2 + \lambda(\|U\|_F^2 + \|V\|_F^2)$

Another Technical Challenge

- Predict true preference of user based on the binary observation

- $\begin{bmatrix} 2 & 4 \\ 1 & 2 \end{bmatrix}$? $\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix}$

Natural Approach

- CF for real-valued data:

$$\min_{X} \|(Z - X)_{\Omega}\|_F^2, \quad \text{s.t. } \text{rank}(X) \leq r$$

- CF for binary data:

$$\min_{X} \|(Z - \text{sign}(X))_{\Omega}\|_F^2, \quad \text{s.t. } \text{rank}(X) \leq r$$

Why Natural Approach Fails?

$$Z \qquad X \qquad \tilde{X}$$

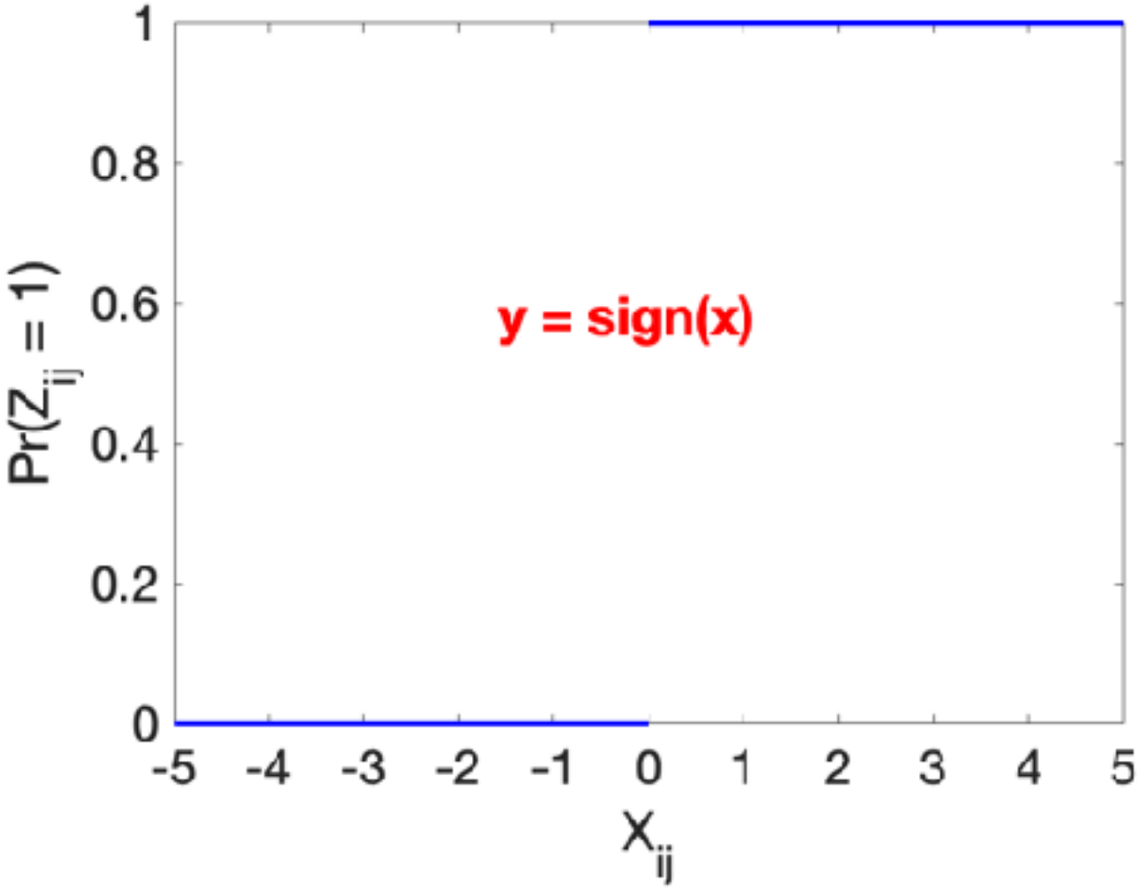
$$\begin{bmatrix} -1 & 1 \\ -1 & -1 \end{bmatrix} \quad \begin{bmatrix} -1 & 3 \\ -2 & -4 \end{bmatrix} \quad \begin{bmatrix} -1 & 5 \\ -5 & -1 \end{bmatrix}$$

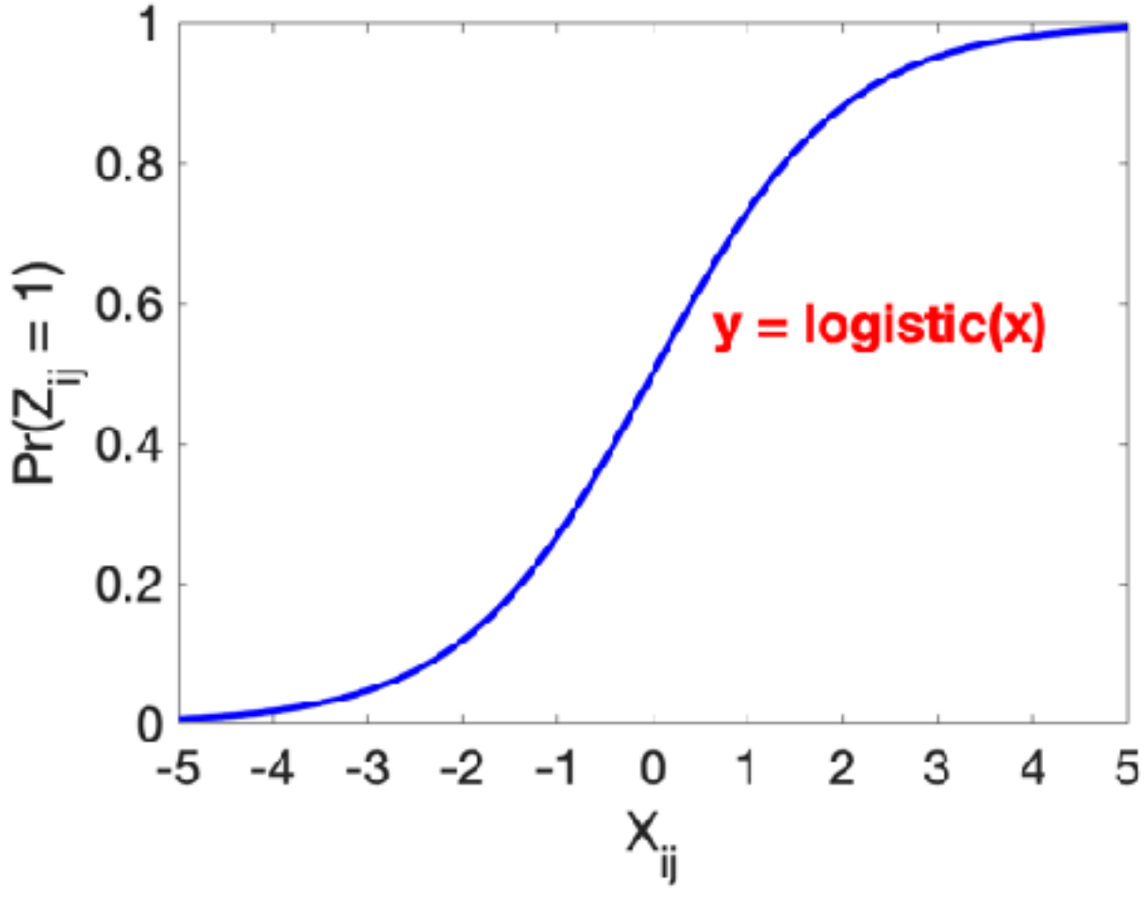
- Sign function absorbs magnitude

Probabilistic Model

- For $(i,j) \in \Omega \quad Z_{i,j} = \begin{cases} +1, & \text{with probability } f(X_{i,j}) \ , \\ -1, & \text{with probability } 1 - f(X_{i,j}) \ . \end{cases}$

- $f(X_{i,j}) = \dfrac{1}{1 + e^{-X_{i,j}}}$ is logistic function





Maximum Likelihood Estimation

- Negative log-likelihood function:

$$L(X; Z, \Omega) = -\sum_{(i,j)\in\Omega} \left(\mathbf{1}_{[Z_{i,j}=1]} \log(f(X_{i,j})) + \mathbf{1}_{[Z_{i,j}=-1]} \log(1 - f(X_{i,j})) \right)$$

$$= \sum_{(i,j)\in\Omega} \left(\log(1 + \exp(-Z_{i,j} X_{i,j})) \right)$$

- Maximum Likelihood Estimation:

$$\min L(X; Z, \Omega), \quad \text{s.t.} \quad \text{rank}(X) \leq r$$

Algorithm: stochastic gradient descent

- The gradients with respect to u_i and v_j are given by

$$\nabla_u L(U, V; Z, \Omega) = \frac{-Z_{i,j}}{1 + \exp(Z_{i,j} u_i v_j^\top)} v_j,$$

$$\nabla_v L(U, V; Z, \Omega) = \frac{-Z_{i,j}}{1 + \exp(Z_{i,j} u_i v_j^\top)} u_i .$$

Stochastic algorithm

- For iteration $t = 1, \cdots, T$
  - Randomly choose an entry $Z_{i,j}$ with $(i, j) \in \Omega$
  - Update the new iterate

$$u_i \longleftarrow \Pi(u_i - \alpha_t \nabla_u)$$
$$v_j \longleftarrow \Pi(v_i - \alpha_t \nabla_v)$$