

Introduction to Recommender System

Sunghee Yun

Recommender Systems : Example

Customers who bought this item also bought

Page 1 of 19



Nikon WU-1a Wireless Mobile Adapter 27081 for Nikon Df, Nikon 1 S2, COOLPIX P530, D3300, ...
★★★★☆ 1,999
\$39.00 ✓prime



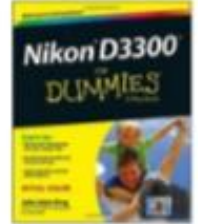
SanDisk 32GB Ultra Class 10 SDHC UHS-I Memory Card Up to 80MB/s, Grey/Black (SDSDUNC...)
★★★★☆ 6,446
#1 Best Seller in SecureDigital Memory Cards
\$12.99 ✓prime



Medium Soft Padded Digital SLR Camera Travel Case/Bag with Clip-on Detachable and...
★★★★☆ 24
\$10.99 ✓prime



NIKON AF-S DX NIKKOR 55-200mm f/4-5.6G ED VR II Lens - White Box (New)
★★★★☆ 258
\$146.95 ✓prime



Nikon D3300 For Dummies (For Dummies Series) by Julie Adair King
★★★★☆ 194
Paperback
\$17.86 ✓prime



Customers who viewed this item also viewed

Page 1 of 11



Nikon D3400 w/ AF-P DX NIKKOR 18-55mm f/3.5-5.6G VR (Black)
★★★★☆ 474
\$496.95 ✓prime



Canon EOS Rebel T6 Digital SLR Camera Kit with EF-S 18-55mm f/3.5-5.6 IS II Lens, Built...
★★★★☆ 106
\$348.90 ✓prime



Nikon D3400 DSLR Camera with AF-P DX NIKKOR 18-55mm f/3.5-5.6G VR and AF-P DX NIKKOR 70...
★★★★☆ 474
\$596.95 ✓prime



Nikon D3400 Digital SLR Camera & 18-55mm VR DX AF-P Zoom Lens (Black) - (Certified Refurbished)
★★★★☆ 48
\$413.00



Nikon D3400 DSLR Camera w/ AF-P DX NIKKOR 18-55mm f/3.5-5.6G VR Lens, Black (Certified...
★★★★☆ 8
1 offer from \$503.98



I. Recommender Systems : Example

amazon prime


Deliver to Manish
New York 10033

Buy Again Browsing History Manish's Amazon.com Today's Deals Gift Cards Whole Foods Registry Sell

EN Hello, Manish
Account & Lists Orders Prime


Your Amazon.com Your Browsing History Improve Your Recommendations Your Profile Learn More

Recommended for you, Manish




Buy It Again in Health and Personal Care

22 ITEMS




Buy It Again in Grocery

3 ITEMS



Buy It Again in Other Categories

6 ITEMS



Candy & Chocolate

35 ITEMS

Tranquil series 1000 pcs Garden of Delight

Recommender Systems : Example

Frequently bought together



Total price: **\$256.78**

Add all three to Cart

Add all three to List

i These items are shipped from and sold by different sellers. [Show details](#)

- ✓ **This item:** Nikon 7540 MONARCH 3 8x42 Binocular (Black) **\$228.55**
- ✓ Nikon 6121 PROSTAFF Bino Harness **\$18.24**
- ✓ Nikon LensPen Cleaning System **\$9.99**

Recommendations: More Amazon Examples

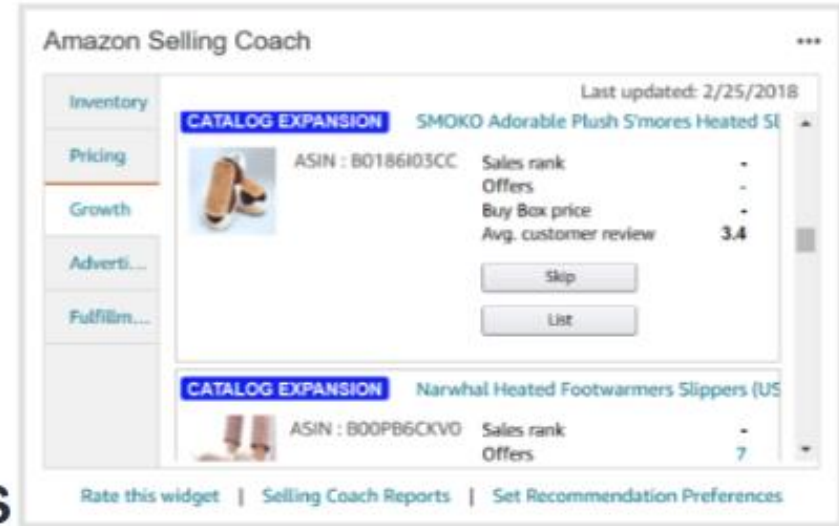
- Prime Video : movie recommendations



- Recommendations for Sellers

- Other examples:

- Amazon Music : song recommendations
- Prime Pantry recommendations based on non-pantry items
- Audio books (Audible)
- Personalized Search (searched for XXX also searched for YYY)



Why is this important to Amazon ?

- Better user experience and increased loyalty
- Opportunity to meet a variety of special needs and tastes
- Increased earnings

[McKinsey Report \(2013\)](#)

35% of the purchases on Amazon are the result of their recommender system, according to **McKinsey**.

[Jeff Bezos - 2016 Letter to Shareholders](#)

But much of what we do with machine learning happens beneath the surface. Machine learning drives our algorithms for demand forecasting, product search ranking, product and deals recommendations, merchandising placements, fraud detection, translations, and much more. Though less visible, much of the impact of machine learning will be of this type – quietly but meaningfully improving core operations.

Other Commercial Recommender Systems

- Alibaba
 - Personalized landing pages -> 20% increase in conversion rate
- YouTube
 - 70% videos are from recommendations
- NetFlix
 - 75% of movies watched are from recommendations (McKinsey)
 - Recommendations save NetFlix \$1B per year (Netflix Execs)
- Best Buy
 - 23.7% increase in sales in 2016 due to recommender system
- Many others: e.g. Spotify, Twitter followers, online dating apps.

Recommender Systems & MFDR

What does MF or DR have to do with Recommender Systems ?

As the Netflix Prize competition has demonstrated, matrix factorization models are superior to classic nearest-neighbor techniques for producing product recommendations, allowing the incorporation of additional information such as implicit feedback, temporal effects, and confidence levels.

Yehuda Koren, *Yahoo Research*

Robert Bell and Chris Volinsky, *AT&T Labs—Research*

User/Ratings Matrix for Toy Example

User Preferences for Movies

	Romance vs. Action	Light vs. Serious
Betty	1	0
Dave	0	1
Russ	0.4	0.6

X

Movie Content

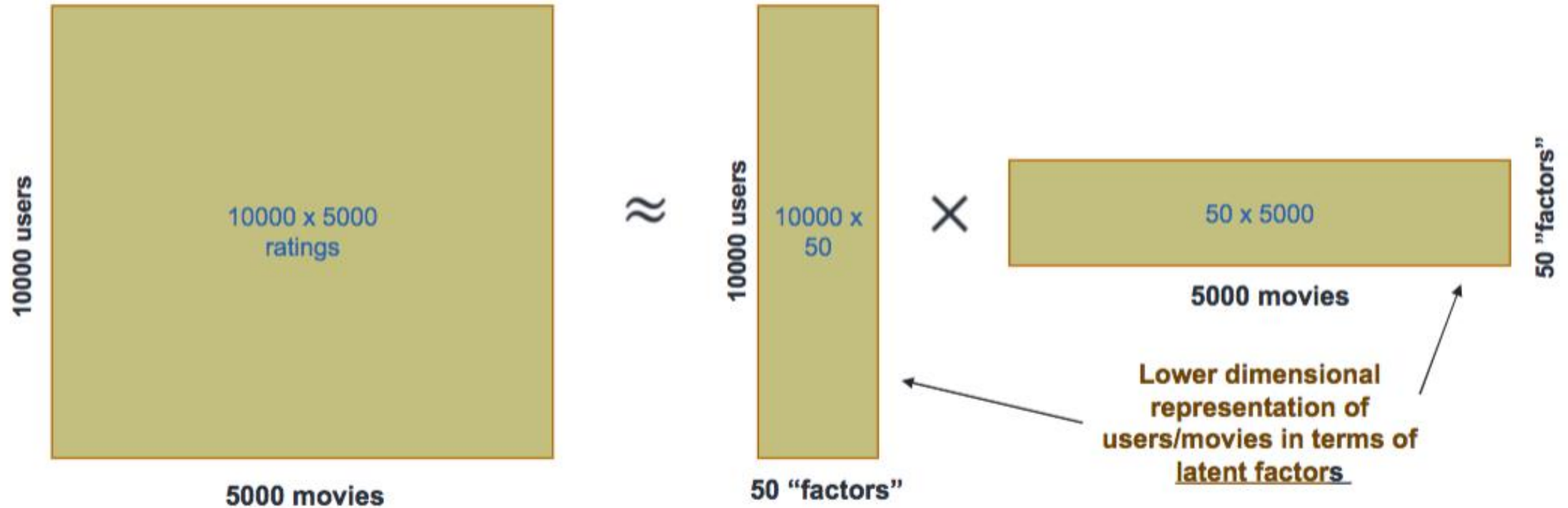
	Color Purple	Princes s Diary	Lion King
Romance vs. Action	8	7	4
Light vs. Serious	2	3	6

=

	Color Purple	Princess Diary	Lion King
Betty	8	7	4
Dave	2	3	6
Russ	4.2	3.9	5.2

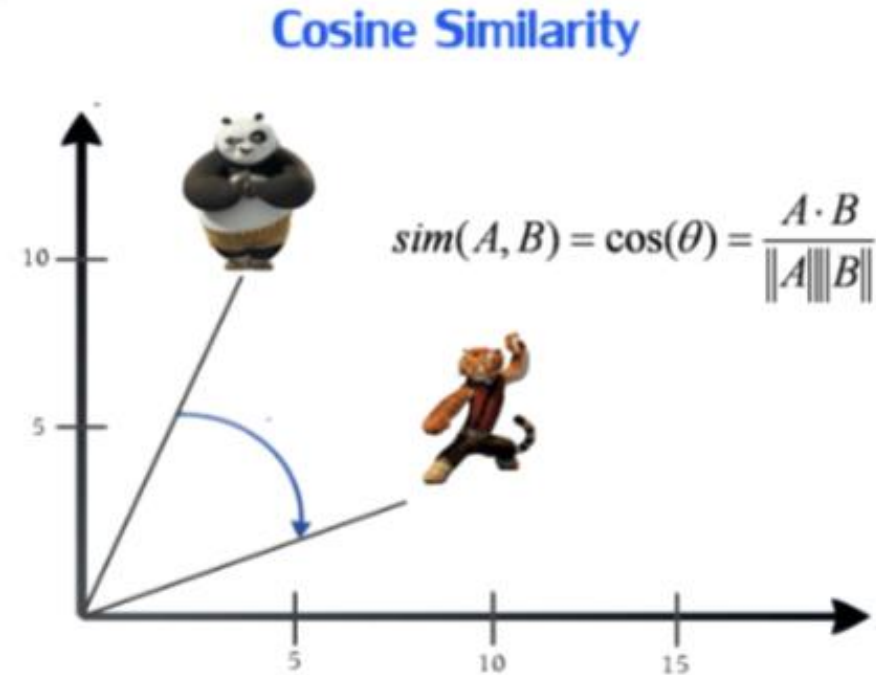
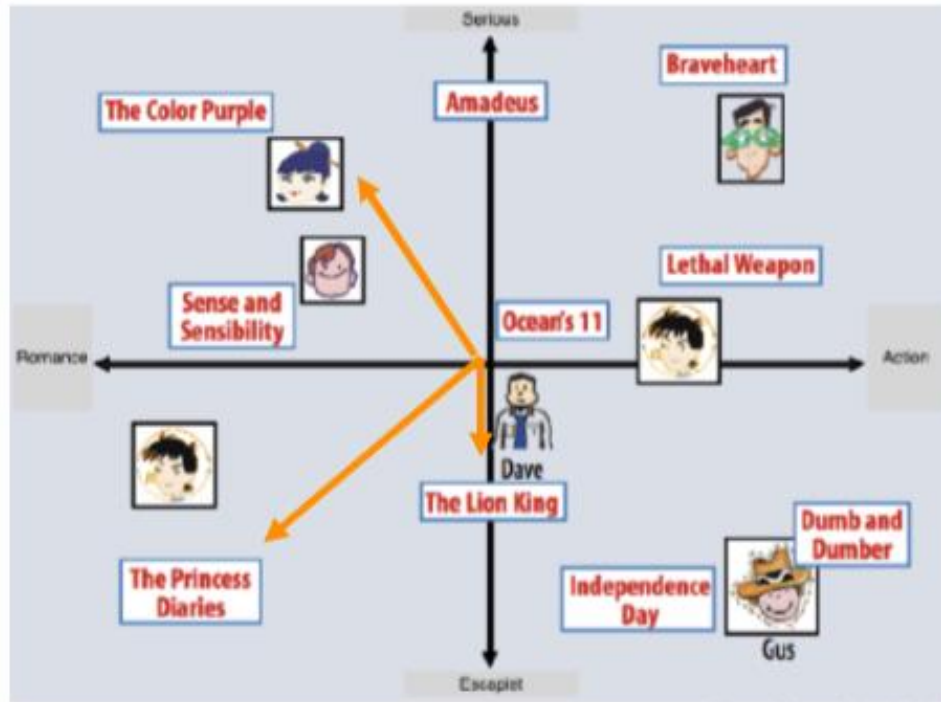
Invert the Problem Now

*Suppose we are given a rating matrix instead, we do not know genres and user preferences
But if we could “factorize” the matrix as*

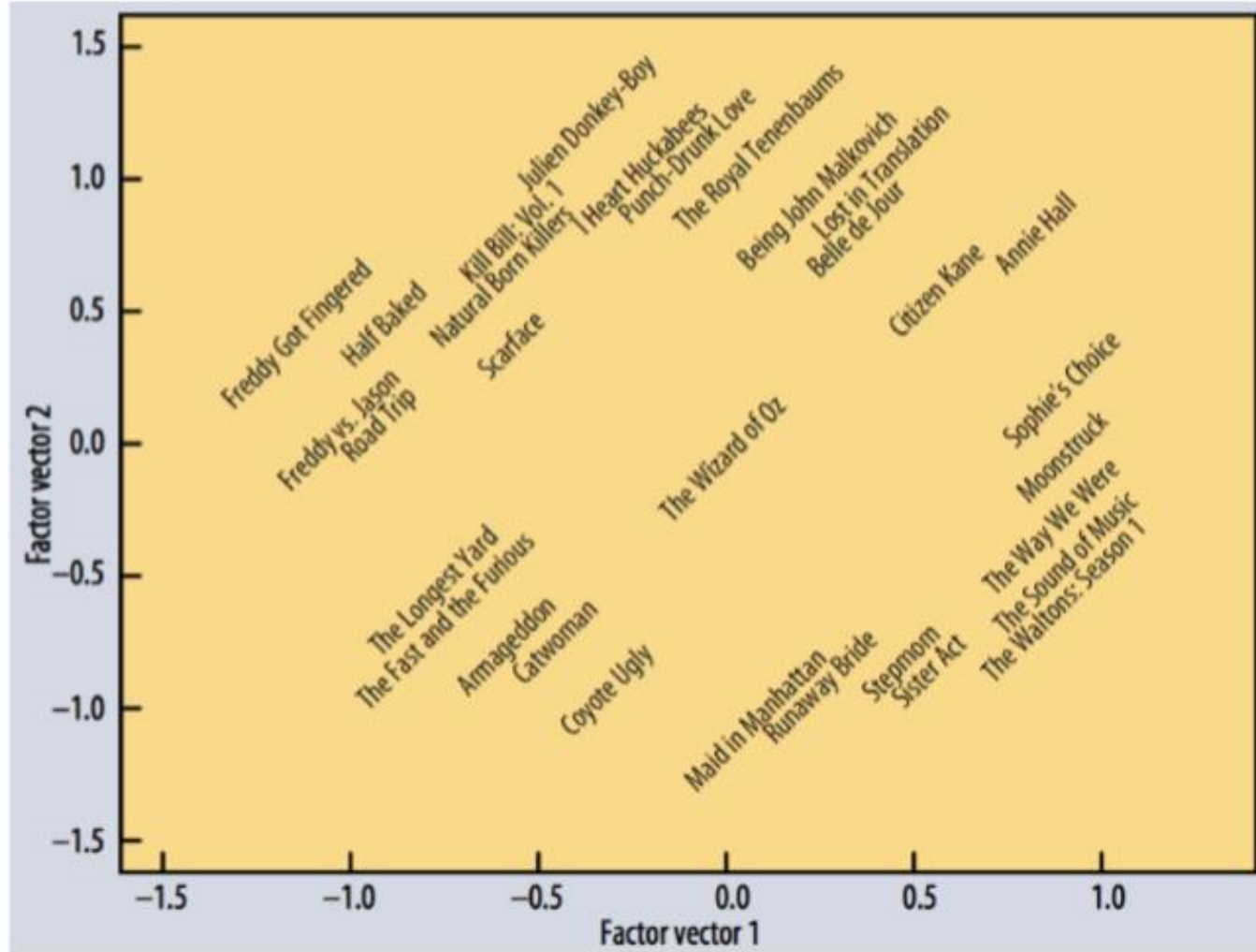


Then we can take the dot product of the user and movie representation in terms of latent factors to predict rating for a movie user has not yet rated. That is, we can make recommendations !!!

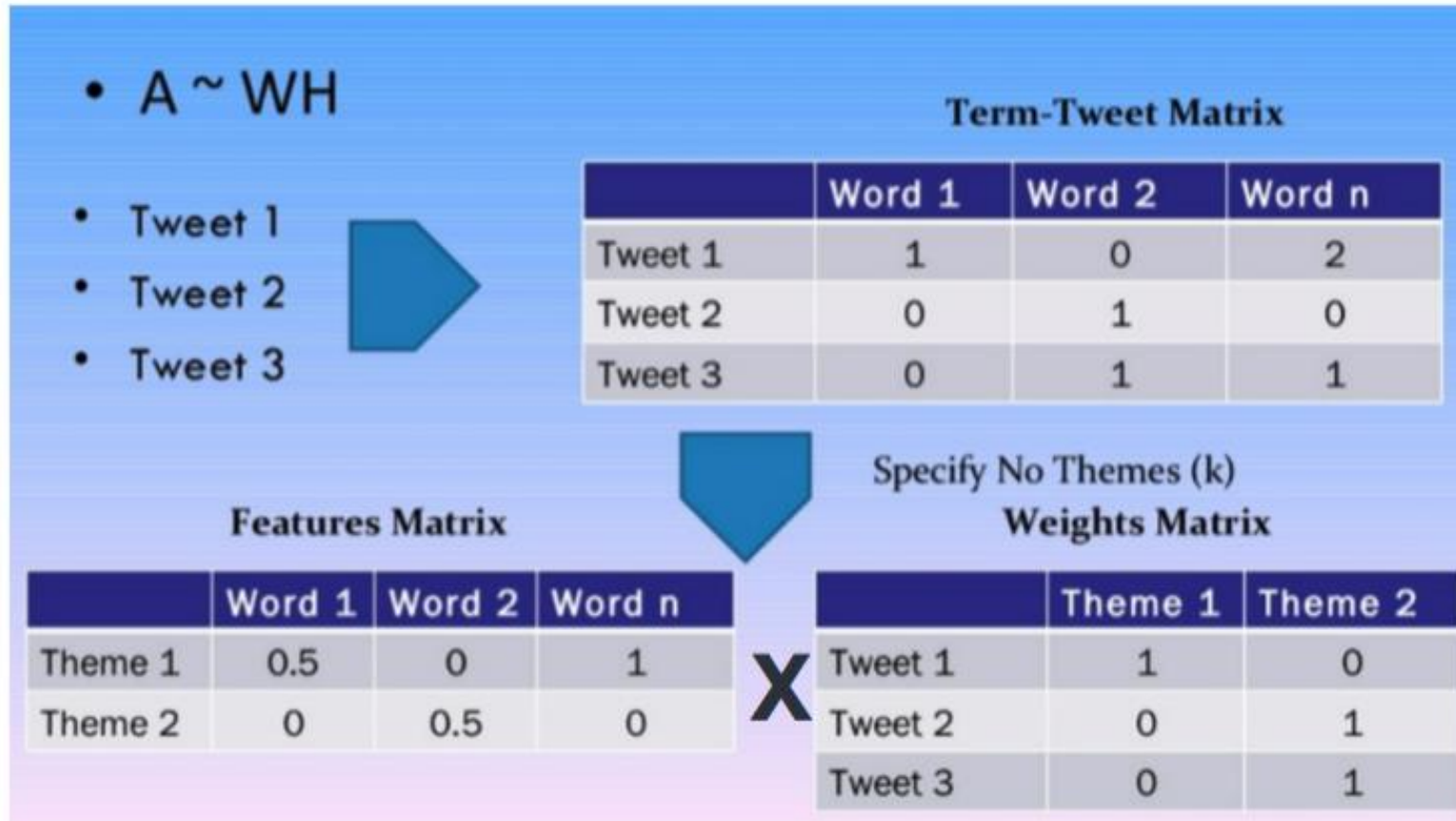
Representation in terms of Latent Factors



Example of Latent Factors: Netflix



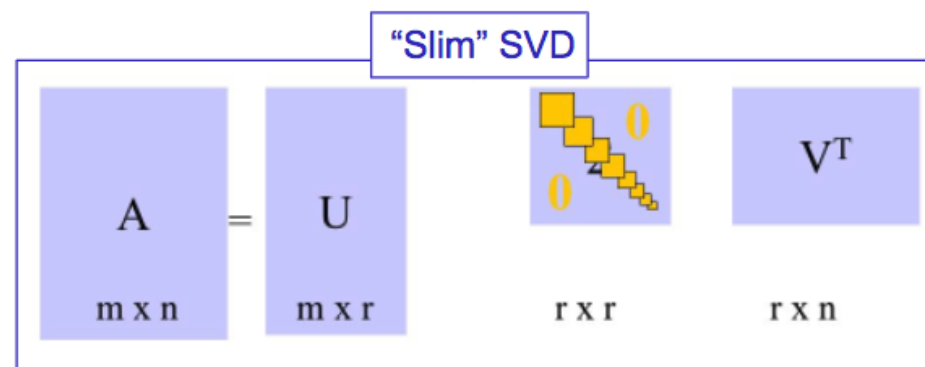
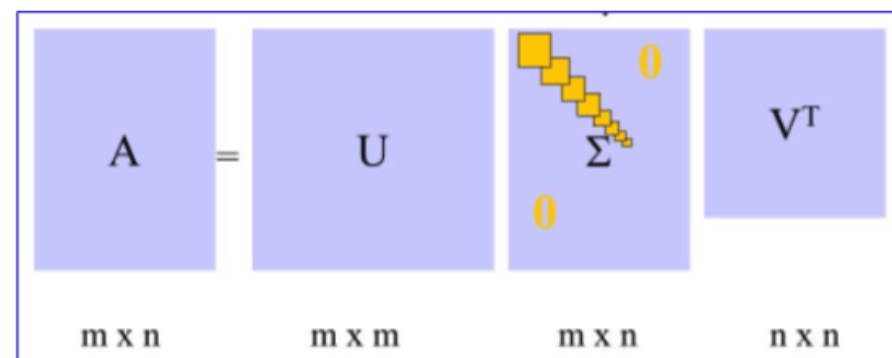
Other Applications: Topic Modeling



Singular Value Decomposition (SVD)

Any $m \times n$ matrix \mathbf{X} admits an SVD given by $\mathbf{X} = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^T$, where

- \mathbf{U} is $m \times m$ and orthogonal (orthonormal), i.e. $\mathbf{U}^T\mathbf{U} = \mathbf{I}$
- \mathbf{V} is $n \times n$ and orthogonal (orthonormal), i.e. $\mathbf{V}^T\mathbf{V} = \mathbf{I}$
- $\mathbf{\Sigma}$ is a diagonal $m \times n$ with non-negative real numbers on the diagonal
- \mathbf{U} form a basis for the columns of \mathbf{X} .
- \mathbf{V} form a basis for the rows of \mathbf{X} .
- Unique if diagonal entries of $\mathbf{\Sigma}$ are in decreasing order, and \mathbf{U}, \mathbf{V} are normalized.
- Diagonal entries in $\mathbf{\Sigma}$ are called *singular values* of \mathbf{X}
- Columns of \mathbf{U} are *left singular vectors* of \mathbf{X}
- Columns of \mathbf{V} are *right singular vectors* of \mathbf{X}
- Rank of \mathbf{X} is the number of non-zero diagonal entries of $\mathbf{\Sigma}$



Application: User/Movie Ratings

	Matrix	Alien	Star Wars	Casablanca	Titanic
Joe	1	1	1	0	0
Jim	3	3	3	0	0
John	4	4	4	0	0
Jack	5	5	5	0	0
Jill	0	2	0	4	4
Jenny	0	0	0	5	5
Jane	0	1	0	2	2

=

$$\begin{bmatrix} .13 & .02 & -.01 \\ .41 & .07 & -.03 \\ .55 & .09 & -.04 \\ .68 & .11 & -.05 \\ .15 & -.59 & .65 \\ .07 & -.73 & -.67 \\ .07 & -.29 & .32 \end{bmatrix} \begin{bmatrix} 12.4 & 0 & 0 \\ 0 & 9.5 & 0 \\ 0 & 0 & 1.3 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ .12 & -.02 & .12 & -.69 & -.69 \\ .40 & .80 & .40 & .09 & .09 \end{bmatrix}$$

U
 Σ
 V^T

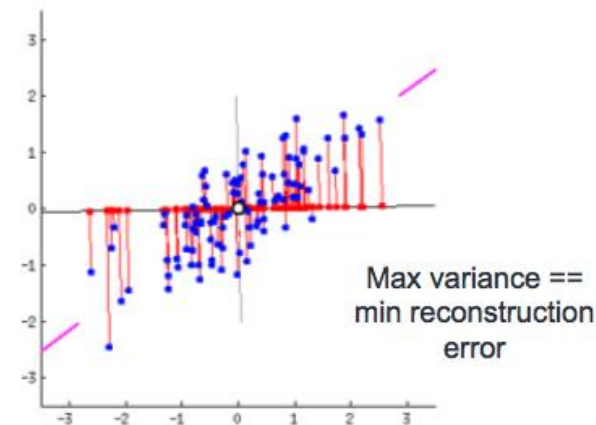
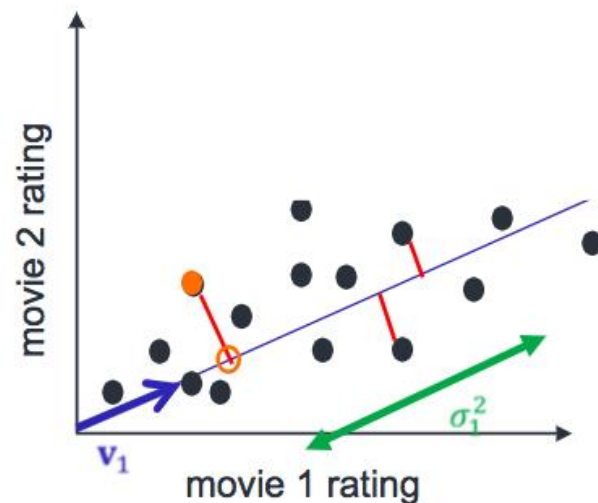
≈

$$\begin{bmatrix} 0.93 & 0.95 & 0.93 & .014 & .014 \\ 2.93 & 2.99 & 2.93 & .000 & .000 \\ 3.92 & 4.01 & 3.92 & .026 & .026 \\ 4.84 & 4.96 & 4.84 & .040 & .040 \\ 0.37 & 1.21 & 0.37 & 4.04 & 4.04 \\ 0.35 & 0.65 & 0.35 & 4.87 & 4.87 \\ 0.16 & 0.57 & 0.16 & 1.98 & 1.98 \end{bmatrix}$$

Complete Example in
Eider Notebook (Link
TBD)

DR with SVD: Intuition

	Matrix	Alien	Star Wars	Casablanca	Titanic
Joe	1	1	1	0	0
Jim	3	3	3	0	0
John	4	4	4	0	0
Jack	5	5	5	0	0
Jill	0	2	0	4	4
Jenny	0	0	0	5	5
Jane	0	1	0	2	2



New coordinates of users in the basis defined by \mathbf{V} will be given by

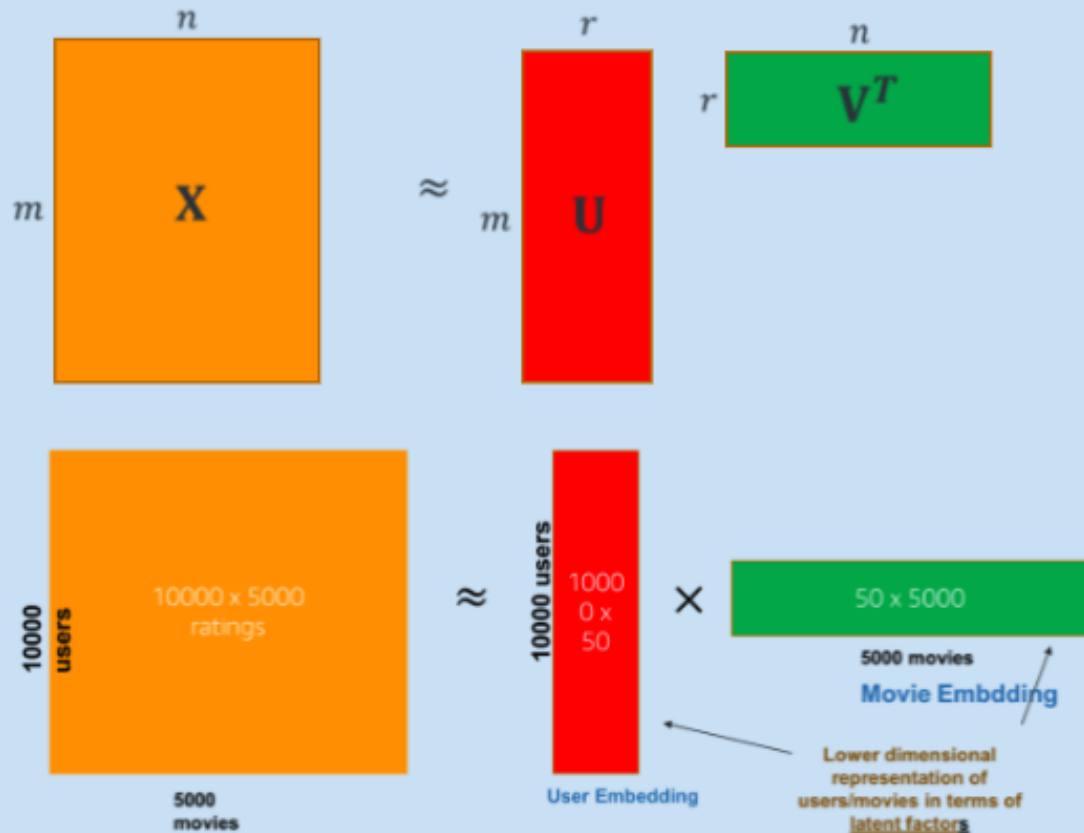
$$\mathbf{XV} = \mathbf{U}\Sigma$$

$$\mathbf{X} = \begin{bmatrix} .13 & .02 \\ .41 & .07 \\ .55 & .09 \\ .68 & .11 \\ .15 & -.59 \\ .07 & -.73 \\ .07 & -.29 \end{bmatrix} \begin{bmatrix} 12.4 & 0 \\ 0 & 9.5 \end{bmatrix} \begin{bmatrix} .56 & .59 & .56 & .09 & .09 \\ .12 & -.02 & .12 & -.69 & -.69 \end{bmatrix}$$

$\mathbf{U} \quad \quad \quad \Sigma \quad \quad \quad \mathbf{V}^T$

Generic Matrix Factorization

Matrix Factorization



MF

$$\min_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{n \times r}} \|X - UV^T\|_F^2$$

Regularized MF

$$\min_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{n \times r}} \|X - UV^T\|_F^2 + \lambda_1 \|U\|_F^2 + \lambda_2 \|V\|_F^2$$

Non-negative MF (NMF)

$$\min_{U \in \mathbb{R}^{m \times r}, V \in \mathbb{R}^{n \times r} \geq 0} \|X - UV^T\|_F^2$$

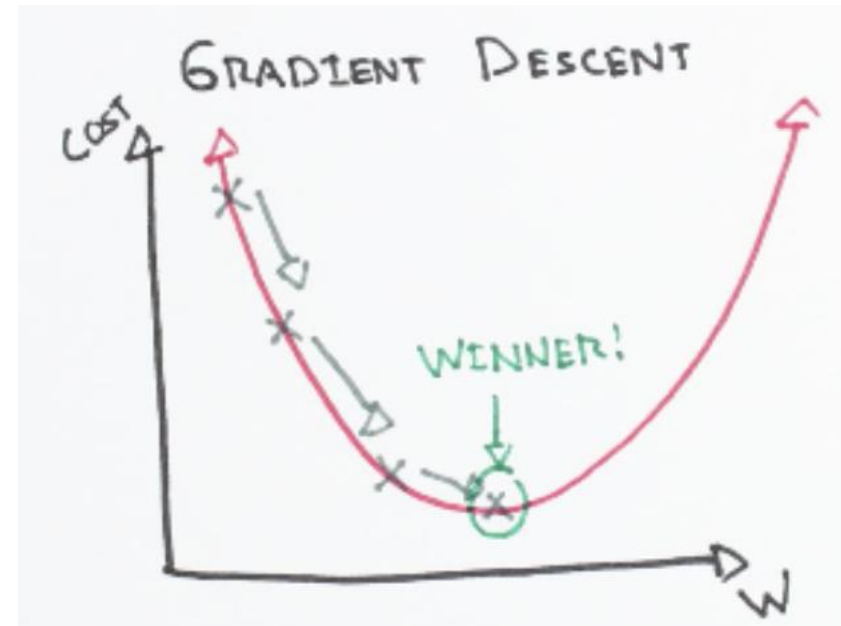
Gradient method

Suppose that $f : \mathbf{R}^n \rightarrow \mathbf{R}$. An unconstrained optimization problem is

$$\text{minimize } f(x) \quad (10.1)$$

The gradient method is

$$x^{k+1} = x^k - \alpha^k \nabla f(x^k)$$



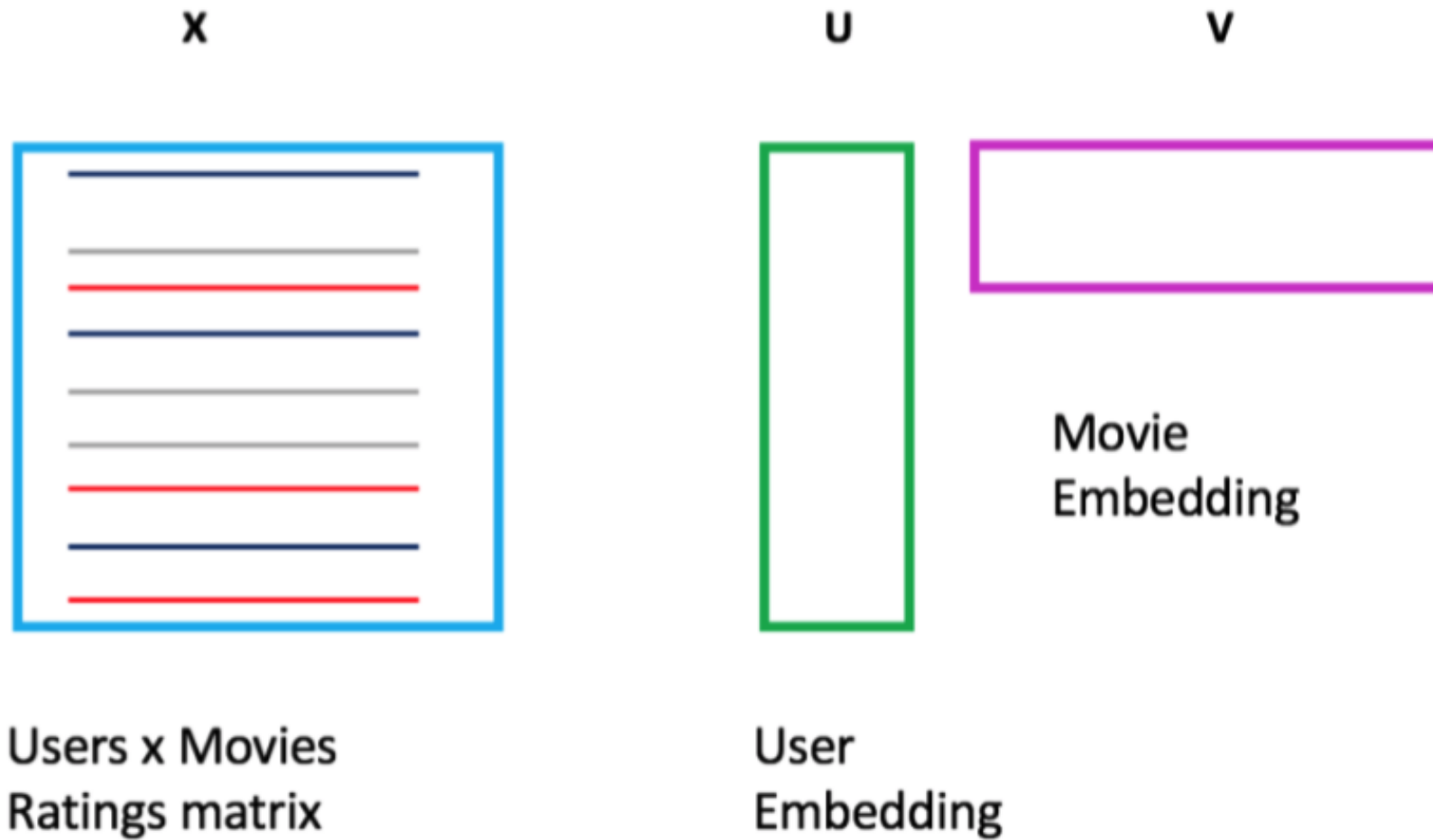
Gradient Descent

1. Given $m \times n$ matrix \mathbf{X} and loss function $L(\mathbf{U}, \mathbf{V}; \mathbf{X})$
2. Initialize \mathbf{U}, \mathbf{V} randomly.
3. At each iteration step:
 - Set $(\mathbf{U}, \mathbf{V}) = (\mathbf{U} - \lambda \nabla_{\mathbf{U}} L(\mathbf{U}, \mathbf{V}; \mathbf{X}), \mathbf{V} - \lambda \nabla_{\mathbf{V}} L(\mathbf{U}, \mathbf{V}; \mathbf{X}))$

Stochastic Gradient Descent - MiniBatch

1. Given $m \times n$ matrix \mathbf{X} and loss function $L(\mathbf{U}, \mathbf{V}; \mathbf{X})$
2. Initialize \mathbf{U}, \mathbf{V} randomly.
3. For each epoch,
 - For each batch,
 - Extract b data points from \mathbf{X} at random to form \mathbf{X}_b
 - Set $\mathbf{U} = \mathbf{U} - \lambda \nabla_{\mathbf{U}} L(\mathbf{U}, \mathbf{V}; \mathbf{X}_b)$
 - Set $\mathbf{V} = \mathbf{V} - \lambda \nabla_{\mathbf{V}} L(\mathbf{U}, \mathbf{V}; \mathbf{X}_b)$

SGD Mini-batch



Incomplete Matrix Factorization

Let R be the set of all $(user, movie)$ pairs that have ratings.

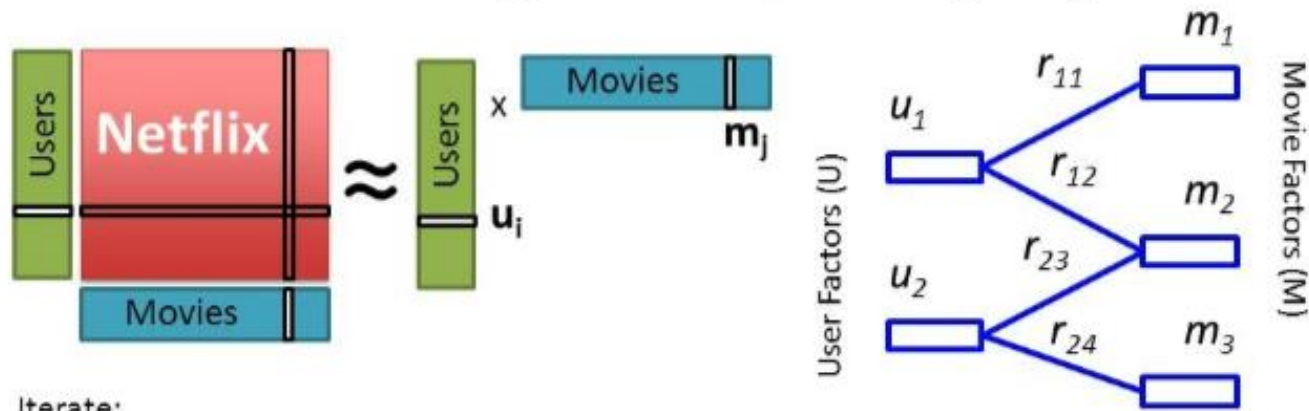
$$\min_{U,V} \sum_{(i,j) \in R} (X_{ij} - U_i^T V_j)^2$$

- Use only known ratings to learn U, V
- Required assumption: every user has rated at least one movie and every movie has at least one user rating (no zeros/columns)
- All algorithms work exactly the same except only use available data points.

ALS with Incomplete Matrix

sense
learn
act

Matrix Factorization Alternating Least Squares (ALS)



Iterate:

$$u_i = \arg \min_w \sum_{j \in N[i]} (r_{ij} - m_j \cdot w)^2$$

$$m_j = \arg \min_w \sum_{i \in N[j]} (r_{ij} - u_i \cdot w)^2$$

Topic Modeling: Examples

- What's trending on twitter?
 - 200 billion tweets a year., 0.5 billion tweets daily
- What's being discussed by congress?
 - 100s of congress bills/year?
- What research topics are hot?
 - 10k active NIH grants

What is Topic Modeling?

- Given a large corpus of documents:
 - Find groups of words that are semantically related (**topics**)
 - Find topics present in each document
- A Bi-clustering problem (**words** and **documents**)
 - "soft" clustering (allow multiple assignments)
- A form of feature reduction
 - Vector representation of a document using topics as features

Topic Model - Example

gene 0.04
dna 0.02
genetic 0.01
...

life 0.02
evolve 0.01
organism 0.01
...

brain 0.04
neuron 0.02
nerve 0.01
...

data 0.02
number 0.02
computer 0.01
...

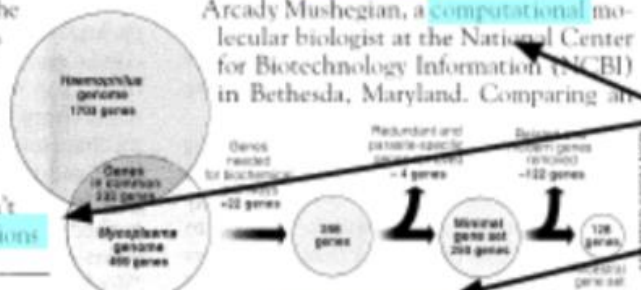
Topics

Seeking Life's Bare (Genetic) Necessities

COLD SPRING HARBOR, NEW YORK—How many genes does an organism need to survive? Last week at the genome meeting here,* two genome researchers with radically different approaches presented complementary views of the basic genes needed for life. One research team, using computer analyses to compare known genomes, concluded that today's organisms can be sustained with just 250 genes, and that the earliest life forms required a mere 128 genes. The other researcher mapped genes in a simple parasite and estimated that for this organism, 800 genes are plenty to do the job—but that anything short of 100 wouldn't be enough.

Although the numbers don't match precisely, those predictions

"are not all that far apart," especially in comparison to the 75,000 genes in the human genome, notes Siv Andersson, Uppsala University in Sweden, who arrived at the 800 number. But coming up with a consensus answer may be more than just a genetic numbers game, particularly as more and more genomes are completely mapped and sequenced. "It may be a way of organizing any newly sequenced genome," explains Arcady Mushegian, a computational molecular biologist at the National Center for Biotechnology Information (NCBI) in Bethesda, Maryland. Comparing an



* Genome Mapping and Sequencing, Cold Spring Harbor, New York, May 8 to 12.

Stripping down. Computer analysis yields an estimate of the minimum modern and ancient genomes.

SCIENCE • VOL. 272 • 24 MAY 1996

Documents

Topic proportions and assignments

THANK YOU!