Introduction to Recommender Systems

Course: Database Theory & Applications
University of Glasgow

Dr. Sham Puthiya Parambath

April 1, 2021

Guest Lecture

Introduction

What is a Recommender System ?

- Recommender systems are software agents to find the best matching between users and items.

Introduction

What is a Recommender System ?

- Recommender systems are software agents to find the best matching between users and items.
- Recommender systems are at the heart of the internet world

Introduction

Where is it used?

- Ecommerce - Amazon, Ebay etc

Introduction

Where is it used?

- Ecommerce - Amazon, Ebay etc
- Video Streaming - Youtube, Netflix etc

Introduction

Where is it used?

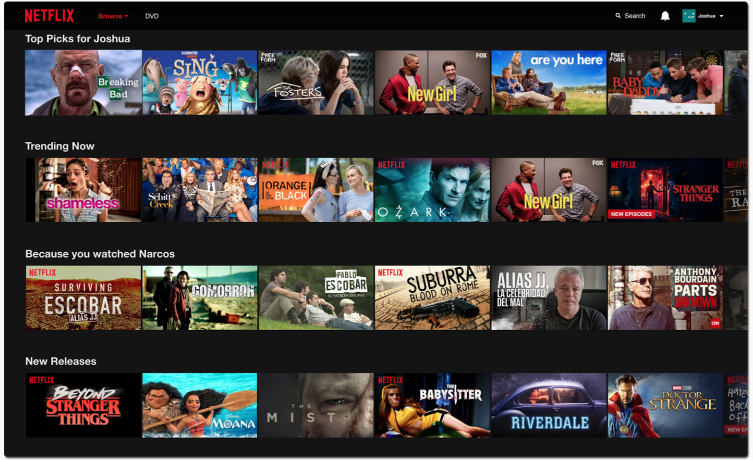
- Ecommerce - Amazon, Ebay etc
- Video Streaming - Youtube, Netflix etc
- Social Networks - Facebook, Twitter etc

Introduction

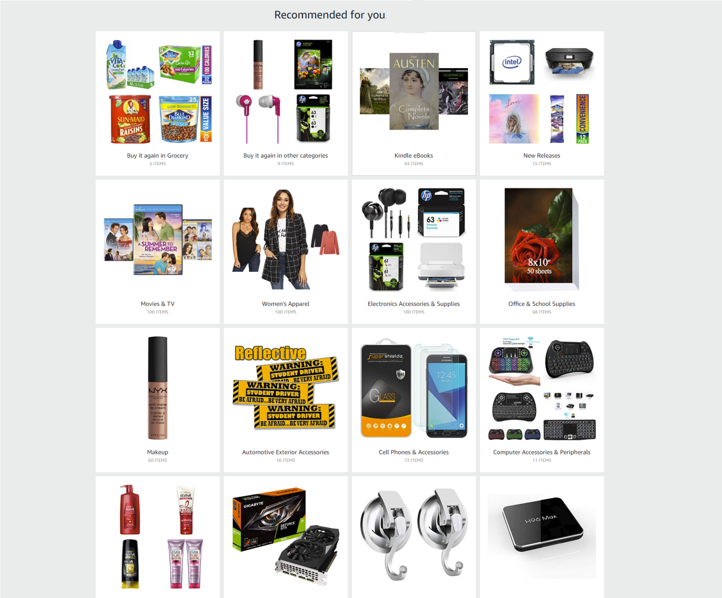
Where is it used?

- Ecommerce - Amazon, Ebay etc
- Video Streaming - Youtube, Netflix etc
- Social Networks - Facebook, Twitter etc
- Public Decision Making, Elections

Introduction



Introduction



Introduction

Why is it important?

- If users can't be matched against proper content, they will lose interest in the service and drop out gradually.

Introduction

Why is it important?

- If users can't be matched against proper content, they will lose interest in the service and drop out gradually.
- Over 75% of netflix viewing, over 70% of Youtube viewing, over 38% of Google News clicks and over 35% of Amazon sales are solely from recommendations

Recommendation Problem

Science of Recommendations

- Recommendation comes in different flavours: *(i)* personalized *(ii)* consensus and *(iii)* group recommendations

Recommendation Problem

Science of Recommendations

- Recommendation comes in different flavours: *(i)* personalized *(ii)* consensus and *(iii)* group recommendations
  - Personalized: given a large collection of items, we have to choose a fixed number of items that match user's intrinsic preference

Recommendation Problem

Science of Recommendations

- Recommendation comes in different flavours: *(i)* personalized *(ii)* consensus and *(iii)* group recommendations
  - Personalized: given a large collection of items, we have to choose a fixed number of items that match user's intrinsic preference
  - Consensus: select a single consensus item from the item set which appeals to the entire population

Recommendation Problem

Science of Recommendations

- Recommendation comes in different flavours: *(i)* personalized *(ii)* consensus and *(iii)* group recommendations
    - Personalized: given a large collection of items, we have to choose a fixed number of items that match user's intrinsic preference
    - Consensus: select a single consensus item from the item set which appeals to the entire population
    - Group: middle ground between fully personalized and consensus recommendation

Personalized Recommendations

Science of Recommendations

- In typical recommendation settings, what data is available? Typical data includes content data and collaborative data

Personalized Recommendations

Science of Recommendations

- In typical recommendation settings, what data is available? Typical data includes content data and collaborative data
  - Content data refers to the metadata associated with the users and items
  - Collaborative data refers to the interaction data between the entities that are collaborating

Personalized Recommendations

Science of Recommendations

- In typical recommendation settings, what data is available? Typical data includes content data and collaborative data
  - Content data refers to the metadata associated with the users and items
  - Collaborative data refers to the interaction data between the entities that are collaborating
- We represent the set of m users as $\mathcal{U}$ and n items as $\mathcal{I}$

Personalized Recommendations

Science of Recommendations

- In typical recommendation settings, what data is available? Typical data includes content data and collaborative data
  - Content data refers to the metadata associated with the users and items
  - Collaborative data refers to the interaction data between the entities that are collaborating
- We represent the set of m users as $\mathcal{U}$ and n items as $\mathcal{I}$
- We represent user content as $\boldsymbol{C}$ and content data associated with items as $\boldsymbol{D}$

Personalized Recommendations

C and D are matrices with rows representing content features.

Personalized Recommendations

C and D are matrices with rows representing content features.

$$C_{m\times p} = \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & \cdots & x_p \end{array} \\ \begin{bmatrix} 1 & 0 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \\ 0 & 0 & \cdots & 1 \end{bmatrix} \end{array} \quad D_{n\times q} = \begin{array}{c} \begin{array}{cccc} y_1 & y_2 & \cdots & y_q \end{array} \\ \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ 1 & 1 & \cdots & 0 \end{bmatrix} \end{array}$$

Personalized Recommendations

$$\boldsymbol{C}[1,:] = \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{bmatrix} \mathit{Female} & 30 & \mathit{Glasgow} & \mathit{Thriller} \end{bmatrix} \end{array}$$

Personalized Recommendations

$$\boldsymbol{C}[1,:] = \begin{array}{c} \begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \end{array} \\ \begin{bmatrix} \mathit{Female} & 30 & \mathit{Glasgow} & \mathit{Thriller} \end{bmatrix} \end{array}$$

$$\boldsymbol{D}[1,:] = \begin{array}{c} \begin{array}{ccccc} y_1 & y_2 & y_3 & y_4 & y_5 \end{array} \\ \begin{bmatrix} \mathit{Thriller} & C.\mathit{Nolan} & J.\mathit{Washington} & \mathit{UK} & 2020 \end{bmatrix} \end{array}$$

Personalized Recommendations

$$\boldsymbol{C}[1,:] = \begin{array}{c} \\ \end{array}\begin{array}{cccc} x_1 & x_2 & x_3 & x_4 \\ \end{array}$$

$$\boldsymbol{C}[1,:] = \begin{array}{c} x_1 \quad\quad x_2 \quad\quad x_3 \quad\quad x_4 \\ \begin{bmatrix} Female & 30 & Glasgow & Thriller \end{bmatrix} \end{array}$$

$$\boldsymbol{D}[1,:] = \begin{array}{c} y_1 \quad\quad y_2 \quad\quad y_3 \quad\quad y_4 \quad\quad y_5 \\ \begin{bmatrix} Thriller & C.Nolan & J.Washington & UK & 2020 \end{bmatrix} \end{array}$$

In practice these matrices will have 1000s of columns and instead of raw values we use encoding methods to represent the values. In real world applications, these are stored in in-memory databases.

Personalized Recommendations

- We represent the collaborative data using the matrix $\boldsymbol{R}$

Personalized Recommendations

- We represent the collaborative data using the matrix $\boldsymbol{R}$
- Rows correspond to users and columns for movies. What is the dimension of this matrix ?

Personalized Recommendations

- We represent the collaborative data using the matrix $\boldsymbol{R}$
- Rows correspond to users and columns for movies. What is the dimension of this matrix ?

$$\boldsymbol{R} = \begin{array}{cc} & \begin{array}{cccccc} i^1 & i^2 & i^3 & \cdots & i^{n-1} & i^n \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \\ \\ \\ u^m \end{array} & \begin{bmatrix} \star & 2 & 5 & \cdots & \star & 4 \\ 2 & \star & \star & \cdots & 4 & 4 \\ \star & 5 & 2 & \cdots & 3 & \star \\ \cdots & \cdots & & & \cdots & \\ \cdots & \cdots & & & \cdots & \\ 1 & 3 & \star & \cdots & \star & 1 \end{bmatrix} \end{array}$$

Personalized Recommendations

.

$$
\begin{array}{cc}
 & \begin{array}{cccccc} i^1 & i^2 & i^3 & \cdots & i^{n-1} & i^n \end{array} \\
R[1,:] = & \begin{bmatrix} \star & 2 & 5 & \cdots & \star & 4 \end{bmatrix}
\end{array}
$$

Personalized Recommendations

.

$$
\boldsymbol{R}[1,:] = \begin{array}{c} \begin{array}{cccccc} i^1 & i^2 & i^3 & \cdots & i^{n-1} & i^n \end{array} \\ \begin{bmatrix} \star & 2 & 5 & \cdots & \star & 4 \end{bmatrix} \end{array}
$$

- The ratings are assumed to be in ordinal scale like 1 star, 2 star, 5 star etc (like in Amazon), like $(+1)$ or dislike (-1) as in Youtube etc.

Recommendation Problem

- We redefine personalized recommendation task as a data driven approach.

Recommendation Problem

- We redefine personalized recommendation task as a data driven approach.
- What items to be recommended ? How many items to be recommended ?

Recommendation Problem

- We redefine personalized recommendation task as a data driven approach.
- What items to be recommended ? How many items to be recommended ?
- Can we cast the recommendation problem as a subset selection problem ?

Recommendation Problem

- We redefine personalized recommendation task as a data driven approach.
- What items to be recommended ? How many items to be recommended ?
- Can we cast the recommendation problem as a subset selection problem ?
- How do we quantify usefulness ? Let f measure the usefulness of item i to user u i.e. $f : \mathcal{U} \times \mathcal{I} \rightarrow \Re$

Recommendation Problem

- We redefine personalized recommendation task as a data driven approach.
- What items to be recommended ? How many items to be recommended ?
- Can we cast the recommendation problem as a subset selection problem ?
- How do we quantify usefulness ? Let f measure the usefulness of item i to user u i.e. $f : \mathcal{U} \times \mathcal{I} \rightarrow \Re$
- The problem can be cast as a subset selection problem that maximizes a utility function

Recommendation As An Optimization Problem

Input : Content or/and Collaborative Data & f

Output : k items with highest values for f from $\mathcal{A}$

Problem Statement: $\operatorname{argmax}_{\substack{\mathcal{S} \subseteq \mathcal{I} \setminus \mathcal{O} \\ |\mathcal{S}| \leq \mathsf{K}}} f(u, \mathcal{S})$

Recommendation Algorithms

Content Based Recommendations

Users

Items

$$U = \begin{array}{c c} & \begin{array}{cccc} genre_1 & genre_2 & genre_3 & genre_4 \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \end{array} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \end{array} \mathrm{M} = \begin{array}{c c} & \begin{array}{cccc} genre_1 & genre_2 & genre_3 & genre_4 \end{array} \\ \begin{array}{c} i^1 \\ i^2 \\ i^3 \end{array} & \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \end{array}$$

$$f(u, i) = \qquad \sum_j u_j i_j$$

Content Based Recommendations

Users	Items	Utility ($f(u, i)$)
u^1	i^1	1
u^1	i^2	2
u^1	i^3	1
u^2	i^1	1
u^2	i^2	2
u^2	i^3	1
u^3	i^1	1
u^3	i^2	2
u^3	i^3	1

Content Based Recommendations

Advantages & Disadvantages

- Data is easy to obtain

Content Based Recommendations

Advantages & Disadvantages

- Data is easy to obtain
- But we might not have enough user metadata to make a good prediction

Content Based Recommendations

Advantages & Disadvantages

- Data is easy to obtain
- But we might not have enough user metadata to make a good prediction
- Performs very poorly in real world applications

Recommendation Algorithms

Collaborative Recommendations

- State-Of-The-Art algorithms used in industrial level recommender systems

Recommendation Algorithms

Collaborative Recommendations

- State-Of-The-Art algorithms used in industrial level recommender systems
- Can be subdivided into two
  - Neighbourhood models
  - Latent Factor Models

Neighbourhood models

- Based on the similarity relationship between the collaborating entities.

Neighbourhood models

- Based on the similarity relationship between the collaborating entities.
- User based neighbourhood models

Neighbourhood models

- Based on the similarity relationship between the collaborating entities.
- User based neighbourhood models
- Item based neighbourhood models

User-User Collaborative Filtering

- Based on the principle that a user can be identified using her neighbors i.e. ‘similar users have similar tastes‘

User-User Collaborative Filtering

- Based on the principle that a user can be identified using her neighbors i.e. ‘similar users have similar tastes‘
- Given a target user, how can we identify her neighbors ? Similarity function ?

User-User Collaborative Filtering

- Based on the principle that a user can be identified using her neighbors i.e. 'similar users have similar tastes'
- Given a target user, how can we identify her neighbors ? Similarity function ?

$$
\boldsymbol{R} = \begin{array}{c c}
 & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\
\begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \star & 5 & 4 & 5 & \star \\ 2 & 2 & \star & \star & 4 & 5 \\ \star & 4 & 3 & \star & 4 & 2 \\ 3 & 5 & \star & 2 & \star & 1 \end{pmatrix}
\end{array}
$$

User-User Collaborative Filtering

$$\cos(a,b) = \begin{array}{c} \\ u^1 \\ u^2 \\ u^3 \\ u^4 \end{array}\begin{array}{c} \begin{array}{cccc} u^1 & u^2 & u^3 & u^4 \end{array} \\ \begin{pmatrix} 1 & 0.38 & 0.64 & 0.22 \\ 0.38 & 1 & 0.72 & 0.48 \\ 0.64 & 0.72 & 1 & 0.53 \\ 0.22 & 0.48 & 0.53 & 1 \end{pmatrix} \end{array}$$

$$u^1 = \begin{pmatrix} 1 & \star & 5 & 4 & 5 & \star \end{pmatrix}$$

$$u^2 = \begin{pmatrix} 2 & 2 & \star & \star & 4 & 5 \end{pmatrix}$$

User-User Collaborative Filtering

Algorithm

- Identify set of users most similar to the target user according to a similarity function

User-User Collaborative Filtering

Algorithm

- Identify set of users most similar to the target user according to a similarity function
- Identify the products these similar users liked

User-User Collaborative Filtering

Algorithm

- Identify set of users most similar to the target user according to a similarity function
- Identify the products these similar users liked
- Generate predictions from the liked items and recommend the items with highest utility value

User-User Collaborative Filtering

Illustration

$$
\begin{aligned}
f(u_1, i_2) &= 0.38 \times 2 + 0.64 \times 4 + 0.22 \times 5 = 4.42 \\
f(u_1, i_6) &= 0.38 \times 5 + 0.64 \times 2 + 0.22 \times 1 = 3.4 \\
f(u_2, i_3) &= 0.38 \times 5 + 0.72 \times 3 = 4.06 \\
f(u_2, i_4) &= 0.38 \times 4 + 0.48 \times 2 = 2.48 \\
f(u_3, i_1) &= 0.64 \times 1 + 0.72 \times 2 + 0.53 \times 3 = 3.67 \\
f(u_3, i_4) &= 0.64 \times 4 + 0.53 \times 2 = 3.62 \\
f(u_4, i_3) &= 0.22 \times 5 + 0.53 \times 3 = 2.69 \\
f(u_4, i_5) &= 0.22 \times 5 + 0.48 \times 4 + 0.53 \times 4 = 5.14
\end{aligned}
$$

User-User Collaborative Filtering

$$R = \begin{array}{c} \\ u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} \begin{array}{c} \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{pmatrix} 1 & \mathbf{4.42} & 5 & 4 & 5 & \mathbf{3.4} \\ 2 & 2 & \mathbf{4.06} & \mathbf{2.48} & 4 & 5 \\ \mathbf{3.67} & 4 & 3 & \mathbf{3.62} & 4 & 2 \\ 3 & 5 & \mathbf{2.69} & 2 & \mathbf{5.14} & 1 \end{pmatrix} \end{array}$$

User CF will recommend i^2, i^3, i^1, i^5 for u^1, u^2, u^3 and u^4 respectively.

Item-Item Collaborative Filtering

- Based on the principle that an item can be identified using its neighbors i.e ‘similar items have similar preferences‘

Item-Item Collaborative Filtering

- Based on the principle that an item can be identified using its neighbors i.e ‘similar items have similar preferences‘

$$
\boldsymbol{R} = \begin{array}{c c} & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \star & 5 & 4 & 5 & \star \\ 2 & 2 & \star & \star & 4 & 5 \\ \star & 4 & 3 & \star & 4 & 2 \\ 3 & 5 & \star & 2 & \star & 1 \end{pmatrix} \end{array}
$$

Item-Item Collaborative Filtering

$$\cos(a, b) = \begin{array}{cc} & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{array}{c} i^1 \\ i^2 \\ i^3 \\ i^4 \\ i^5 \\ i^6 \end{array} & \begin{pmatrix} 1 & 0.76 & 0.23 & 0.60 & 0.46 & 0.63 \\ 0.76 & 1 & 0.31 & 0.33 & 0.47 & 0.63 \\ 0.23 & 0.31 & 1 & 0.77 & 0.84 & 0.19 \\ 0.60 & 0.33 & 0.77 & 1 & 0.59 & 0.08 \\ 0.46 & 0.47 & 0.84 & 0.59 & 1 & 0.68 \\ 0.63 & 0.63 & 0.19 & 0.08 & 0.68 & 1 \end{pmatrix} \end{array}$$

Item-Item Collaborative Filtering

Illustration

$$
\begin{aligned}
f(u_1, i_2) &= 0.76 \times 1 + 0.31 \times 5 + 0.33 \times 4 + 0.47 \times 5 = 6 \\
f(u_1, i_6) &= 0.63 \times 1 + 0.19 \times 5 + 0.08 \times 4 + 0.68 \times 5 = 5.3 \\
f(u_2, i_3) &= 0.23 \times 2 + 0.31 \times 2 + 0.84 \times 4 + 0.19 \times 5 = 5.4 \\
f(u_2, i_4) &= 0.60 \times 2 + 0.33 \times 2 + 0.59 \times 4 + 0.08 \times 5 = 4.6 \\
f(u_3, i_1) &= 0.76 \times 4 + 0.23 \times 3 + 0.46 \times 4 + 0.63 \times 2 = 5.4 \\
f(u_3, i_3) &= 0.31 \times 4 + 0.77 \times 2 + 0.84 \times 4 + 0.19 \times 5 = 7.1 \\
f(u_4, i_3) &= 0.23 \times 3 + 0.31 \times 5 + 0.77 \times 2 + 0.19 \times 1 = 4.0 \\
f(u_4, i_5) &= 0.46 \times 3 + 0.47 \times 5 + 0.59 \times 2 + 0.68 \times 1 = 5.6
\end{aligned}
$$

Item-Item Collaborative Filtering

$$R = \begin{array}{c c} & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \mathbf{6} & 5 & 4 & 5 & \mathbf{5.3} \\ 2 & 2 & \mathbf{5.4} & \mathbf{4.6} & 4 & 5 \\ \mathbf{5.4} & 4 & 3 & \mathbf{7.1} & 4 & 2 \\ 3 & 5 & \mathbf{4} & 2 & \mathbf{5.6} & 1 \end{pmatrix} \end{array}$$

Item CF will recommend recommend i^2, i^3, i^4, i^5 for u^1, u^2, u^3, u^4 respectively.

Latent Factor Models

- Based on the principle of ‘there exists fixed number of latent factors for users and items‘

Latent Factor Models

- Based on the principle of ‘there exists fixed number of latent factors for users and items‘
- What is a latent factor ?

Latent Factor Models

- Based on the principle of ‘there exists fixed number of latent factors for users and items‘
- What is a latent factor ?

Latent Factor Models

- Based on the principle of ‘there exists fixed number of latent factors for users and items‘
- What is a latent factor ?

$$R = \begin{array}{c c} & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \star & 5 & 4 & 5 & \star \\ 2 & 2 & \star & \star & 4 & 5 \\ \star & 4 & 3 & \star & 4 & 2 \\ 3 & 5 & \star & 2 & \star & 1 \end{pmatrix} \end{array}$$

Latent Factor Models

$$
\boldsymbol{R} = \begin{array}{c c}
 & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\
\begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \star & 5 & 4 & 5 & \star \\ 2 & 2 & \star & \star & 4 & 5 \\ \star & 4 & 3 & \star & 4 & 2 \\ 3 & 5 & \star & 2 & \star & 1 \end{pmatrix}
\end{array}
$$

We assume that the observed rating matrix is a linear function of user latent factor and item latent factor

$$\boldsymbol{R} = \boldsymbol{U}\boldsymbol{V}$$

where $\boldsymbol{U}$ is the $n \times k$ user latent matrix and $\boldsymbol{V}$ is the $m \times k$ item latent factor matrix

Latent Factor Models

- Given $\boldsymbol{R}$, how can we estimate the latent matrices $\boldsymbol{U}$ and $\boldsymbol{V}$?

Latent Factor Models

- Given $\boldsymbol{R}$, how can we estimate the latent matrices $\boldsymbol{U}$ and $\boldsymbol{V}$?
- We can use optimization techniques to estimate $\boldsymbol{U}$ and $\boldsymbol{V}$

Latent Factor Models

- Given $\boldsymbol{R}$, how can we estimate the latent matrices $\boldsymbol{U}$ and $\boldsymbol{V}$?
- We can use optimization techniques to estimate $\boldsymbol{U}$ and $\boldsymbol{V}$

$$\underset{\boldsymbol{U},\boldsymbol{V}}{\operatorname{argmin}} \|\boldsymbol{U}\boldsymbol{V} - \boldsymbol{R}\|^2 \tag{1}$$

Latent Factor Models

$$R = \begin{array}{c c} & \begin{array}{cccccc} i^1 & i^2 & i_3 & i_4 & i_5 & i_6 \end{array} \\ \begin{array}{c} u^1 \\ u^2 \\ u^3 \\ u^4 \end{array} & \begin{pmatrix} 1 & \mathbf{2.55} & 5 & 4 & 5 & \mathbf{4.45} \\ 2 & 2 & \mathbf{3.85} & \mathbf{3.05} & 4 & 5 \\ \mathbf{2.07} & 4 & 3 & \mathbf{2.29} & 4 & 2 \\ 3 & 5 & \mathbf{2.29} & 2 & \mathbf{3.15} & 1 \end{pmatrix} \end{array}$$

Latent factor model will recommend i^6, i^3, i^4, i^5 for u^1, u^2, u^3, u^4 respectively.

References i

- Paolo Cremonesi, Yehuda Koren, and Roberto Turrin.
  Performance of recommender algorithms on top-n recommendation tasks.
  In *RecSys*, 2010.
- Mohamed Sarwat, James Avery, and Mohamed F Mokbel.
  Recdb in action: recommendation made easy in relational databases.
  VLDB, 2013.

References ii

- Mohamed Sarwat.
  Database management system support for collaborative filtering recommender systems.
  2014.
- Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl.
  Item-based collaborative filtering recommendation algorithms.
  In *WWW*, 2001.