Tutorial 11
Overview

Recommender Systems
  Content-based Filtering
  Collaborative Filtering

Multi-Dimensional Scaling
  Vanilla MDS
  MDS Variants
  ISOMAP Exercise
Recommender Systems

Amazon

Facebook

IMDb

Netflix

Yelp

YouTube
Content-based Filtering

- Supervised learning method (features & labels)
  - Features can describe users (e.g. “average amount of time spent watching Netflix”) and/or items (e.g. “Romantic Comedy”, “Oscar winning”)
  - Labels are ratings
- Fit a model, then at test time recommend the item that would be rated the highest.
Collaborative Filtering

- Unsupervised learning
  - We are given users’ ratings of items but no features of users or items
  - Need to fill in the “user-item matrix”
Collaborative Filtering - Latent Factor Model

- Instead of prespecifying features like content-based filtering, we learn features to represent users and items
  \[ y_{ij} \approx w_j^T z_i \]
  \[ Y \approx Z_{n,k} W_{k,d} \]

- \( z_i \) feature vector for each user
- \( w_j \) feature vector for each movie
- Use a squared loss function with L2 regularization to train model over available ratings, \( R \)
  \[ F(Z, W) = \sum_{(i,j) \in R} (w_j^T z_i - y_{ij})^2 + \frac{\lambda_1}{2} \| Z \|_F^2 + \frac{\lambda_2}{2} \| W \|_F^2 \]
- Can also introduce bias for user, item, or both
  \[ y_{ij} \approx w_j^T z_i + \beta + \beta_i + \beta_j \]
Multi-Dimensional Scaling

- No latent factors, directly optimize the location of the $z_i$ values
- Classic MDS cost function:

$$f(Z) = \sum_{i+1}^{N} \sum_{j=i+1}^{N} (\|z_i - z_j\| - \|x_i - x_j\|)^2$$

- PCA used latent factors, $W$, and represented the data as a linear combination of them
- MDS is non-parametric
MDS Variants

- General MDS cost function:

\[
f(Z) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} d_{3}(d_{1}(z_{i}, z_{j}), d_{2}(x_{i}, x_{j}))
\]

- \(d(\cdot, \cdot)\) can be
  - A norm (classic MDS)
  - Geodesic distance (ISOMAP, \(d_{2}\))
    - Distance along graph formed by k-nearest neighbours

- Sammon’s Mapping

\[
f(Z) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} \frac{(d_{1}(z_{i}, z_{j}) - d_{2}(x_{i}, x_{j}))^{2}}{d_{2}(x_{i}, x_{j})^{2}}
\]
ISOMAP Exercise

Create a distance matrix using geodesic distances, where $k = 2$
ISOMAP Exercise - Answer

<table>
<thead>
<tr>
<th></th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>0</td>
<td>1</td>
<td>(\sqrt{2})</td>
<td>(\sqrt{2} + 1)</td>
<td>(\sqrt{2} + 2)</td>
<td>(\sqrt{2} + 3)</td>
<td>(1.5\sqrt{2} + 3)</td>
<td>(2\sqrt{2} + 3)</td>
</tr>
<tr>
<td>b</td>
<td>0</td>
<td>1</td>
<td>2</td>
<td>(\sqrt{2} + 4)</td>
<td>3</td>
<td>4</td>
<td>(0.5\sqrt{2} + 4)</td>
<td>(\sqrt{2} + 4)</td>
</tr>
<tr>
<td>c</td>
<td>0</td>
<td>1</td>
<td>(\sqrt{2} + 3)</td>
<td>(\sqrt{2} + 3)</td>
<td>2</td>
<td>(0.5\sqrt{2} + 3)</td>
<td>(\sqrt{2} + 3)</td>
<td></td>
</tr>
<tr>
<td>d</td>
<td>0</td>
<td>1</td>
<td>(\sqrt{2} + 2)</td>
<td>(\sqrt{2} + 2)</td>
<td>2</td>
<td>(0.5\sqrt{2} + 2)</td>
<td>(\sqrt{2} + 2)</td>
<td></td>
</tr>
<tr>
<td>e</td>
<td>0</td>
<td>1</td>
<td>(\sqrt{2} + 1)</td>
<td>(\sqrt{2} + 1)</td>
<td>0</td>
<td>(0.5\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td></td>
</tr>
<tr>
<td>f</td>
<td>0</td>
<td>0</td>
<td>(0.5\sqrt{2})</td>
<td>(\sqrt{2})</td>
<td>0</td>
<td>(0.5\sqrt{2})</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>g</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>h</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>