Introduction to Data Science

Center for Data Science
Iddo Drori, Spring 2019

Content-Based Recommendation

Problem Formulation

- Items $i=1..n$
- Users $j=1..p$
- Ratings Y_{ij}
- Binary mask if rating is available M_{ij}

- Problem: matrix completion

Naive Solution

- If k dimensional feature vectors $x^{(i)}$ are known for each item $i=1..n$
- Learn parameter vectors $\theta^{(j)}$ for each user $j=1..p$
- Predict user j rating item i by $\theta^{(j)^T} x^{(i)}$

$$\underset{\theta^{(j)}}{\text{minimize}} \frac{1}{2} \sum_{i:M_{ij}=1} \left(\theta^{(j)^T} x^{(i)} - Y_{ij} \right)^2 + \frac{\lambda}{2} \sum_k \theta^{(j)^2}$$

Naive Solution

- If k dimensional feature vectors $x^{(i)}$ are known for each item $i=1..n$
- Learn parameter vectors $\theta^{(j)}$ for **all users** $j=1..p$ using gradient descent
- Predict user j rating item i by $\theta^{(j)^T} x^{(i)}$

$$\underset{\theta^{(1)}\ldots\theta^{(p)}}{\text{minimize}} \frac{1}{2} \sum_{i,j:M_{ij}=1} \left(\theta^{(j)^T} x^{(i)} - Y_{ij}\right)^2 + \frac{\lambda}{2} \sum_j \sum_k \theta^{(j)^2}$$

Collaborative Filtering

Problem

- If k dimensional feature vectors $x^{(i)}$ are unknown
- Given parameter vectors $\theta^{(j)}$ for all users $j=1..p$ learn feature vectors $x^{(i)}$

$$\underset{x^{(1)}\ldots x^{(n)}}{\text{minimize}}\ \frac{1}{2}\sum_{i,j:M_{ij}=1}\left({\theta^{(j)}}^T x^{(i)} - Y_{ij}\right)^2 + \frac{\lambda}{2}\sum_{i}\sum_{k} {x_k^{(i)}}^2$$

Iterative Solution

- Given $\{x^{(1)}, \dots, x^{(n)}\}$ learn $\{\theta^{(1)}, \dots, \theta^{(p)}\}$

- Given $\{\theta^{(1)}, \dots, \theta^{(p)}\}$ learn $\{x^{(1)}, \dots, x^{(n)}\}$

Collaborative Filtering Solution

- Learn $\{x^{(1)}, \ldots, x^{(n)}\}$ and $\{\theta^{(1)}, \ldots, \theta^{(p)}\}$ together

$$\underset{x^{(1)} \ldots x^{(n)}, \theta^{(1)} \ldots \theta^{(p)}}{\text{minimize}} \frac{1}{2} \sum_{i,j:M_{ij}=1} \left({\theta^{(j)}}^T x^{(i)} - Y_{ij} \right)^2 + \frac{\lambda}{2} \sum_i \sum_k {x_k^{(i)}}^2 + \frac{\lambda}{2} \sum_j \sum_k {\theta_k^{(j)}}^2$$

- Predictions ${\theta^{(j)}}^T x^{(i)}$

Collaborative Filtering Solution

- Low rank factorization of rating into product of feature matrix and parameter matrix

$$\begin{bmatrix} {\theta^{(1)}}^T x^{(1)} & \cdots & {\theta^{(p)}}^T x^{(1)} \\ \vdots & \ddots & \vdots \\ {\theta^{(1)}}^T x^{(n)} & \cdots & {\theta^{(p)}}^T x^{(n)} \end{bmatrix}$$