

Item-Item with Implicit Feedback

Introduction

- We've talked about item-item over rating data
- Also works well on unary data (implicit feedback)
 - clicks
 - plays
 - purchases
- But some tweaks are needed

Data Representation

- Rating values: user-item rating matrix
- Need some matrix to represent data
 - Logical (1/0) user-item 'purchase' matrix
 - Purchase count matrix
- Problem: what is a 0?
 - We just ignore that for item-item

Data Normalization

- Standard mean-centering not meaningful
- But we can normalize user vectors to unit vectors
 - Intuition: users who like many items provide less information about any particular pair
- Could also consider: logging counts

Computing Similarities

- Cosine similarity still works $w_{ij} = \cos(\hat{r}_i, \hat{r}_j)$
- Can also use conditional probability $w_{ij} = P(r_i | r_j)$
 - See Deshpande and Karypis paper
 - Becomes quite similar to association rules $P(r_i | r_j) \neq P(r_j | r_i)$

Aggregating Scores

- Weighted average works for non-binary
 - counts
- For binary (0/1), just sum neighbor similarities
 - fixed neighborhood size means this isn't unbounded

$$s(i; u) = \sum_{j \in N} w_{ij}$$

- Neighborhood selection unchanged (most similar)

$$\frac{\sum_j w_{ij} r_{ui}}{\sum_j |w_{ij}|} \in \{0, 1\} = 1$$

$$S(u, i) = \sum_{j \in \mathcal{I}_u} w_{ij}$$

└ can truncate
to top k

Conclusion

- Item-item basically works for unary data
- A few tweaks to algorithm components needed to make it well-behaved
- Test variants with your data/context
 - Evaluation tools in next course will help with this

Item-Item with Implicit Feedback