7-1: Introduction to Dimensionality Reduction Recommenders

Learning Objectives

- To understand the motivation, history, and intuition behind dimensionality reduction recommendation algorithms
- To gain a basic understanding of the algorithm idea, preparing you to master the details later this module
- To understand some of the practical strengths and weaknesses dimensionality reduction approaches

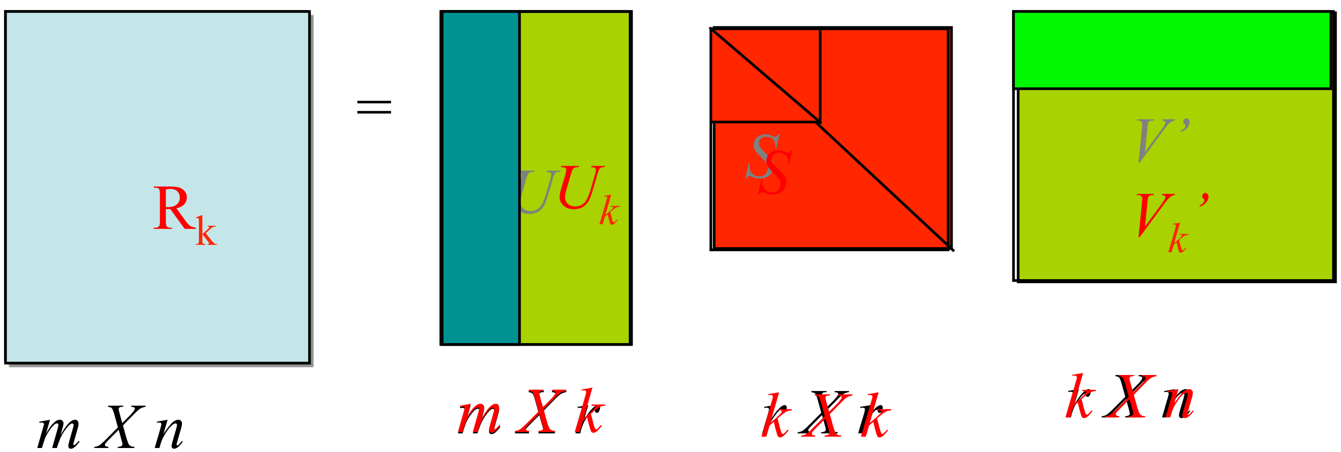
Motivation and Intuition

- Ratings matrix is an overfit representation of user tastes and item descriptions
  - Leads to problems of synonymy – what happens if I like *Hamlet* and *King Lear* and you like *Shakespeare: The Histories and Tragedies*
  - Also leads to computational complexity, potentially poorer results
  - Ideal would be to have a more compact representation of user tastes and item descriptions – but how?

History: Latent Semantic Indexing

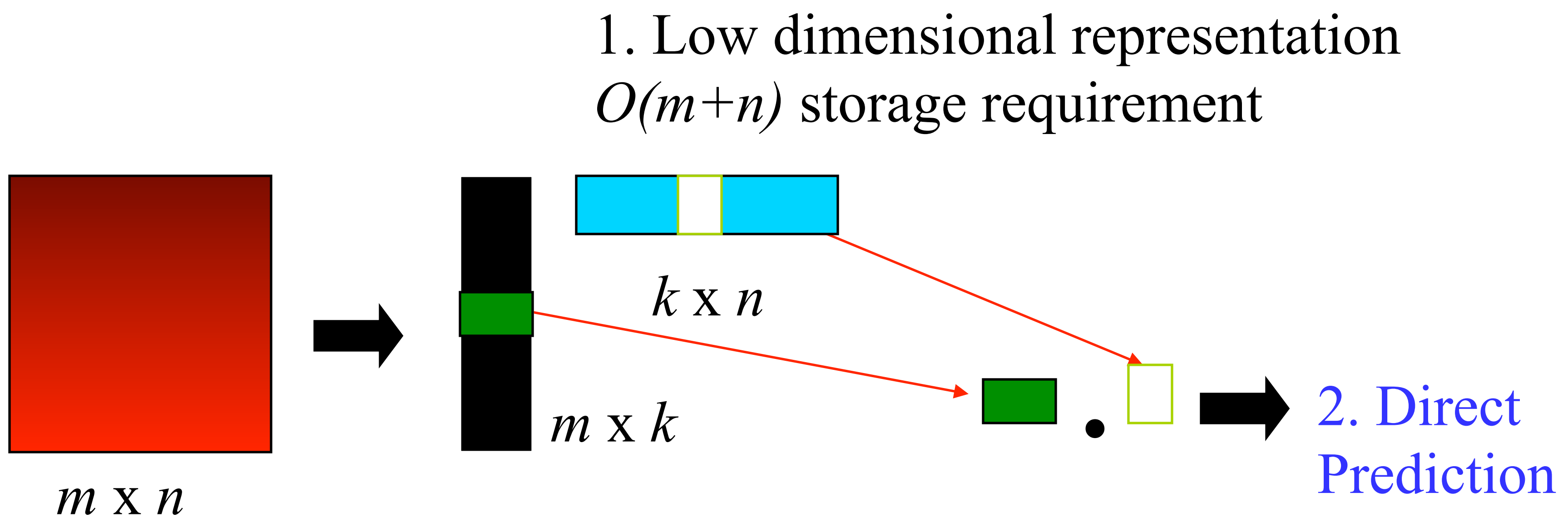
- The information retrieval community addressed this problem earlier (1988)
  - They faced the same issue – keyword vectors had the problem that queries and documents were poorly represented. They wanted to recognize concepts, not words.
- Singular Value Decomposition was used to create a solution
  - Intuitive description: reduce space to a smaller taste space that is compact and robust

SVD: Mathematical Background



The reconstructed matrix $R_k = U_k.S_k.V_k'$ is the closest *rank-k* matrix to the original matrix *R*.

SVD for Collaborative Filtering



Singular Value Decomposition

- Reduce dimensionality of problem
  - Results in small, fast model
  - Richer neighbor network
  - Need to experiment to find appropriate value of k for a domain (for movies, roughly 13-20)
- Challenge #1: missing values
  - Need some way to fill them
  - Several alternatives, including clever averages and predictions

Singular Value Decomposition

- Challenge #2: computational complexity
  - SVD computation is $O(m^2n + n^3)$
  - Some practical approaches
    - Folding in (keep factorization, add new users, item, data) – factorization slowly worsens
    - Probabalistic and incremental approaches
- Challenge #3: lack of transparency / explainability
  - Optimal dimensions do not correspond to user-comprehensible concepts

SVD: Take-Aways

- Clever and useful approach
  - Reduces problems of synonymy and overfitting
  - Computational advantages at run-time
- Significant challenges in model building
  - Particularly for large models
  - One key compromise can be sacrificing model optimality for performance
- SVD is growing in use, but still not dominant in the field

Moving Forward

- Next Lectures
  - Looking at the details of SVD
    - Preparing the matrix (handling missing values)
    - Linear algebra
  - Optimizations and practical variants
    - Gradient descent approaches
    - Simon Funk’s approach from the Netflix Prize

7-1: Introduction to Dimensionality Reduction Recommenders