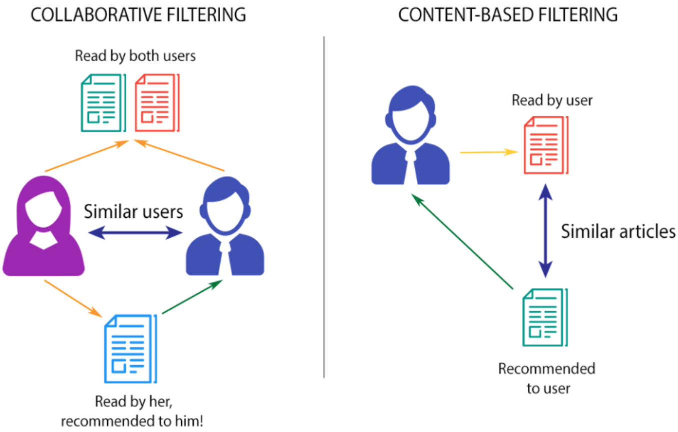
Recommendation Systems Based on Matrix Completion

Joy Song

December 9, 2019

Contents

- **Popular Algorithms:**



Model: Matrix Completion

Collaborative Filtering $\Rightarrow$ Matrix Completion Problem

	Inside Out	Good Will Hunting	Mean Girls	Terminator	Titanic	Warrior
Tina Fey	3	1	5	1	?	1
Helen Mirren	2	?	?	2	5	1
Sylvester Stallone	1	3	1	4	2	5
Tom Hanks	?	3	1	?	4	3
George Clooney	2	2	1	3	1	4

Approach: Matrix Factorization & Rank Minimization

Matrix Factorization

Decompose the sparse data matrix into a user-feature matrix and an item-feature matrix, which capture the latent relationships between users and items.

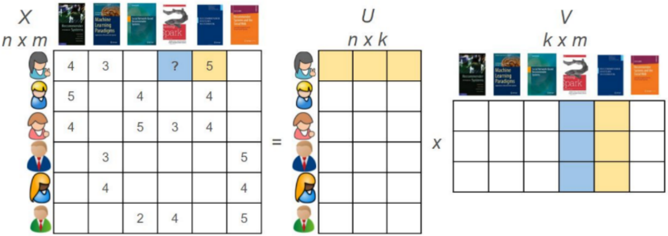
Rank Minimization

Based on the assumption that the complete data matrix $\mathbf{X}$ is of low rank

$$\begin{aligned} \min \quad & \text{rank}(\mathbf{X}) \\ \text{s.t.} \quad & \mathbf{X}_{ij} = \mathbf{A}_{ij}, \quad \forall (i,j) \in \Omega \end{aligned}$$

Contents

Matrix Factorization: SVD



- Fill in the missing entries of $\mathbf{A} \rightarrow \tilde{\mathbf{A}}$
- Singular value decomposition: $\tilde{\mathbf{A}} = \mathbf{U}\boldsymbol{\Sigma}\mathbf{V}^T$
- Dimensionality reduction: $\hat{\mathbf{U}} = \mathbf{U}(:, 1:k), \quad \hat{\mathbf{V}} = \mathbf{V}(:, 1:k)$
- User feature matrix: $\hat{\mathbf{U}}\boldsymbol{\Sigma}^{1/2}$; Item feature matrix: $\hat{\mathbf{V}}\boldsymbol{\Sigma}^{1/2}$
- Estimation: $\hat{\mathbf{A}}_{ij} = \mathbf{U}_F(i,:) \cdot \mathbf{V}_F(j,:)$

Matrix Factorization: Optimization

Loss function

$$E = \frac{1}{2}\sum_{\substack{i=1\\(i,j)\in\Omega}}^{m}\sum_{j=1}^{n}(A_{ij} - \mathbf{u}_i\mathbf{v}_j^T)^2 = \frac{1}{2}\sum_{\substack{i=1\\(i,j)\in\Omega}}^{m}\sum_{j=1}^{n}e_{ij}^2$$

Gradient descent

$$\frac{\partial E}{\partial U_{pq}} = -\sum_{\substack{j=1\\(p,j)\in\Omega}}^{n} e_{pj}v_{jq}; \quad \mathbf{U} \leftarrow \mathbf{U} - \alpha\nabla_u$$

$$\frac{\partial E}{\partial V_{pq}} = -\sum_{\substack{j=1\\(i,p)\in\Omega}}^{n} e_{ip}u_{iq}; \quad \mathbf{V} \leftarrow \mathbf{V} - \alpha\nabla_v$$

Contents

Rank Minimization: Nuclear Norm Minimization

Nuclear Norm

$$\|\mathbf{X}\|_* = \sum_{i=1}^{r} \sigma_i(\mathbf{X})$$

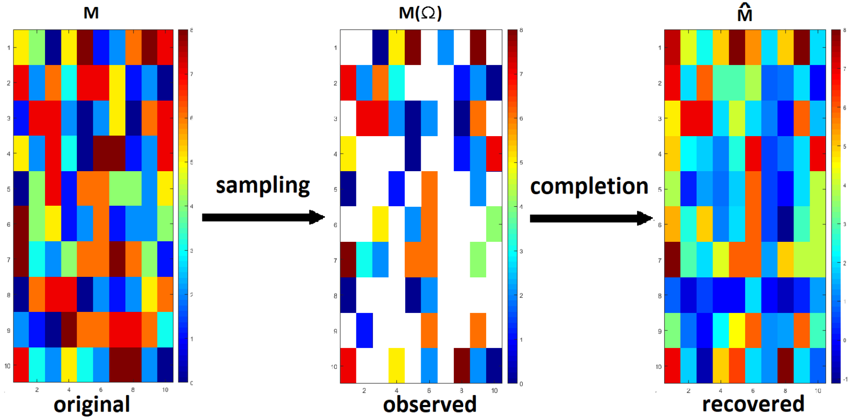
where $\sigma_i(\mathbf{X})$ are the singular values of $\mathbf{X}$.

Optimization Problem

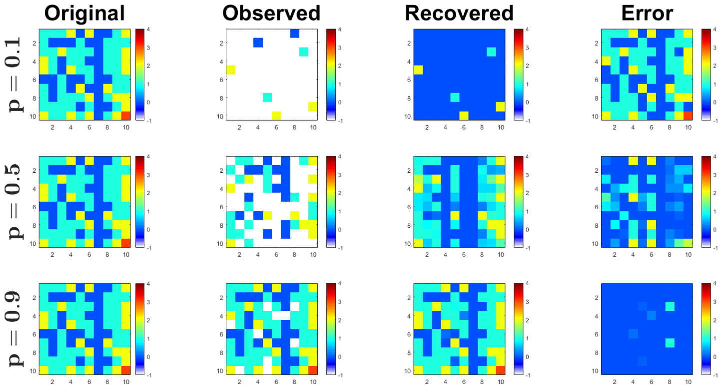
$$\begin{aligned} \min \quad & \text{rank}(\mathbf{X}) \quad \text{s.t.} \quad \mathbf{X}_{ij} = \mathbf{A}_{ij}, \quad \forall (i,j) \in \Omega \\ \min \quad & \|\mathbf{X}\|_* \quad \text{s.t.} \quad \mathcal{P}_\Omega(\mathbf{X}) = \mathcal{P}_\Omega(\mathbf{A}) \end{aligned}$$

Contents

Methodology: Pipeline

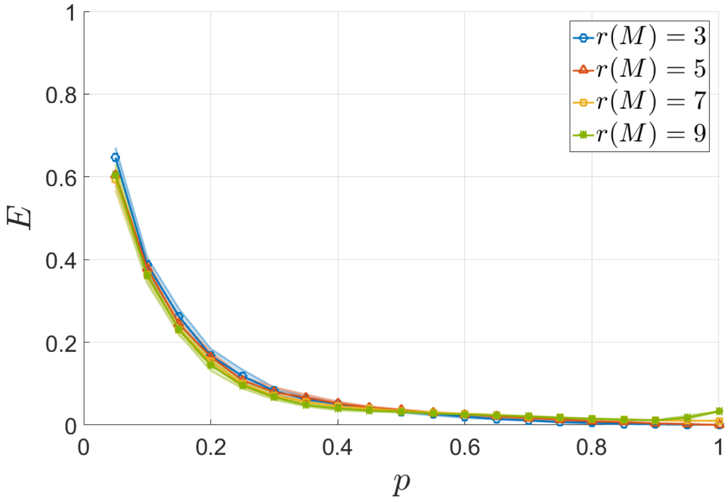


Methodology: Sampling Probabilities

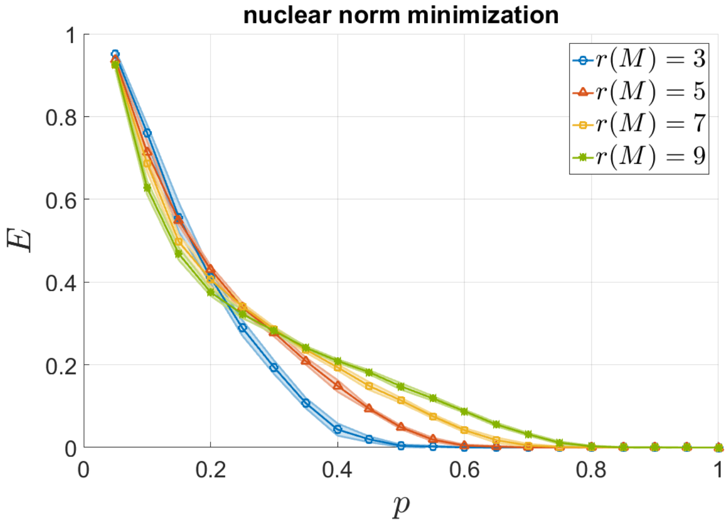


Contents

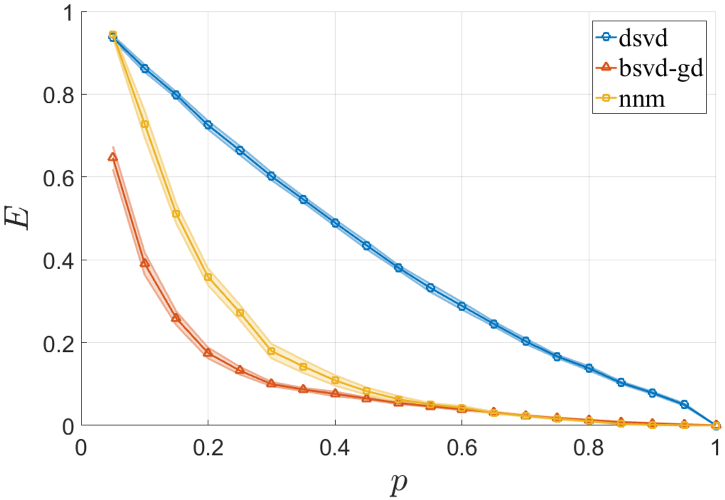
Results: Error vs Pr for MF



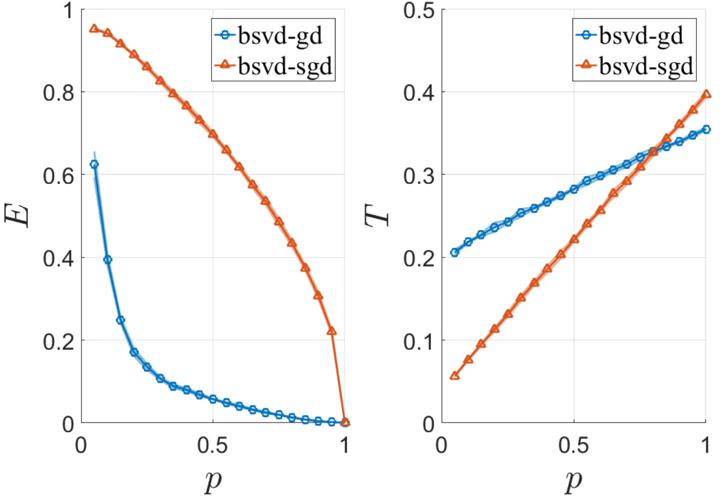
Results: Error vs Pr for NNM



Results: SVD vs MF vs NNM



Results: GD vs SGD for MF



Contents

Conclusions

- As more entries are observed (sampling probabilities $\rightarrow 1$), the errors of different methods are all decreasing to zero.
- When the truncated size k is larger than the rank of the complete matrices, the rank seems to have no influence on the errors, while when the truncated size is smaller than the rank, the error cannot converge to zero. The error of rank minimization method will decrease as the rank of the complete matrices decreases.
- On synthetic data, matrix factorization method via optimization has better performance over the other two methods, direct singular value decomposition produces the largest errors.
- On synthetic data for matrix factorization methods, SGD produces much larger errors than GD, while SGD is much faster than GD.

References

A. Ramlatchan, M. Yang, Q. Liu, M. Li, J.Wang and Y. Li (2018)
A Survey of Matrix Completion Methods for Recommendation Systems
Big Data Mining and Analytics 1, 308 – 323.

Charu C. Aggarwal (2016)
Recommendation Systems
Chapter 3.6 Latent Factor Models 90 – 118.

Badrul M. Sarwar, George Karypis, Joseph A. Konstan, John T. Riedl (2000)
Application of Dimensionality Reduction in Recommender System – A Case Study

E.J. Candès and B. Recht (2009)
Exact Matrix Completion via Convex Optimization
Foundations of Computational Mathematics 9, 771 – 772.

D. Molitor, D. Needell (2018)
Matrix Completion for Structured Observations
arXiv preprint arXiv:1801.09657