Message Passing for Matrix Factorization

Raghunandan Keshavan   Andrea Montanari

Model

- Let $N \in \mathbb{R}^{m \times n}$ be 'approximately' low rank

$$N = UV^T + W$$

- A small subset E of entries revealed
- U and V are typically low rank
- Compute $\widehat{UV^T}$ from the subset of entries revealed

Motivation

Recommendation Systems

- N is the matrix of user ratings
- N_{ij} is rating given by user i to product j
- We want to learn unrevealed N_{ij}
- Recommend product j to customer i if $\widehat{N_{ij}}$ is large

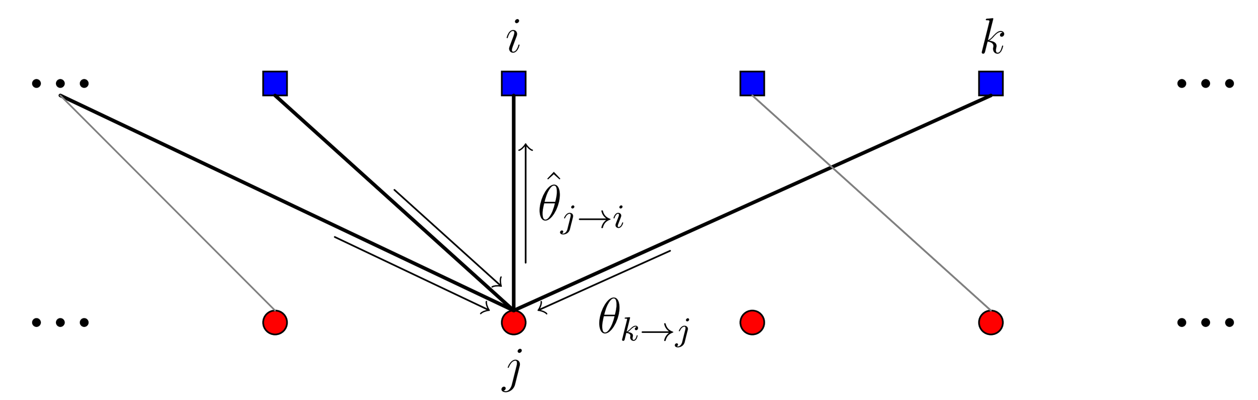
Sensor Localization

- $N_{ij} = D_{ij}^2$: D_{ij}, distance between sensors i and j
- Sensors can only detect nearby sensors
  - Only small N_{ij} are revealed
- Matrix $\{D_{ij}^2\}$ has rank 5
- Want to learn absolute positions of all the sensors
- Applications in indoor positioning

- Image and Video Processing
- Ultrasound Tomography
- ...

Algorithm

$$\mathcal{F}_E(X,Y) \equiv \|\mathcal{P}_E(XY^T - N)\|_F^2 + \lambda \|X\|_F^2 + \lambda \|Y\|_F^2$$



$$\hat{\theta}_{j\to i} = \left(\lambda + \sum_{k \in \partial j \setminus i} \theta_{k\to j}\theta_{k\to j}^T\right)^{-1} \left(\sum_{k \in \partial j \setminus i} N_{jk}\theta_{k\to j}\right)$$

- Initialize $\{\theta_{k\to j}\}$ and $\{\hat{\theta}_{j\to k}\}$ randomly
- Iteratively update the messages $\{\theta_{k\to j}\}$ and $\{\hat{\theta}_{j\to k}\}$
- The variable updates are :

$$x_j = \left(\lambda + \sum_{k \in \partial j} \theta_{k\to j}\theta_{k\to j}^T\right)^{-1} \left(\sum_{k \in \partial j} N_{kj}\theta_{k\to j}\right)$$

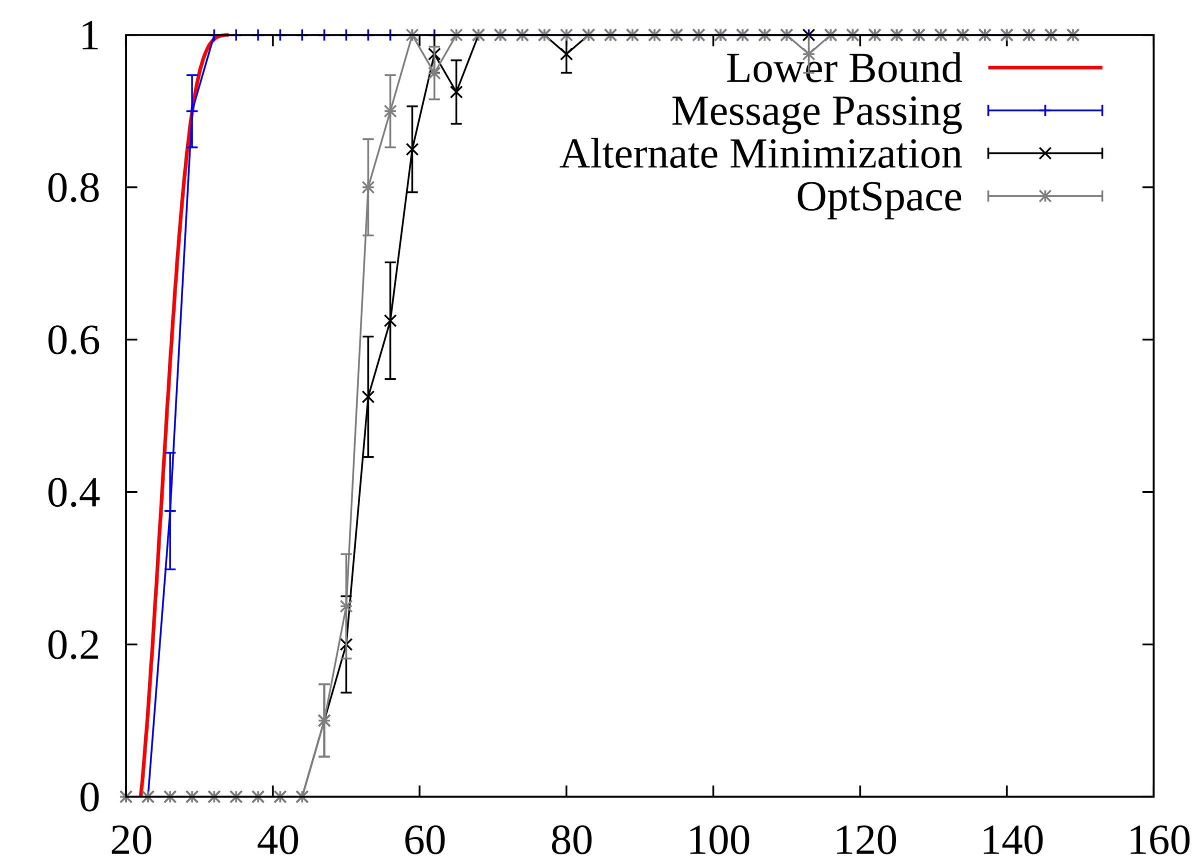


FIGURE 1: Recon. rate vs. $|E|$

Results

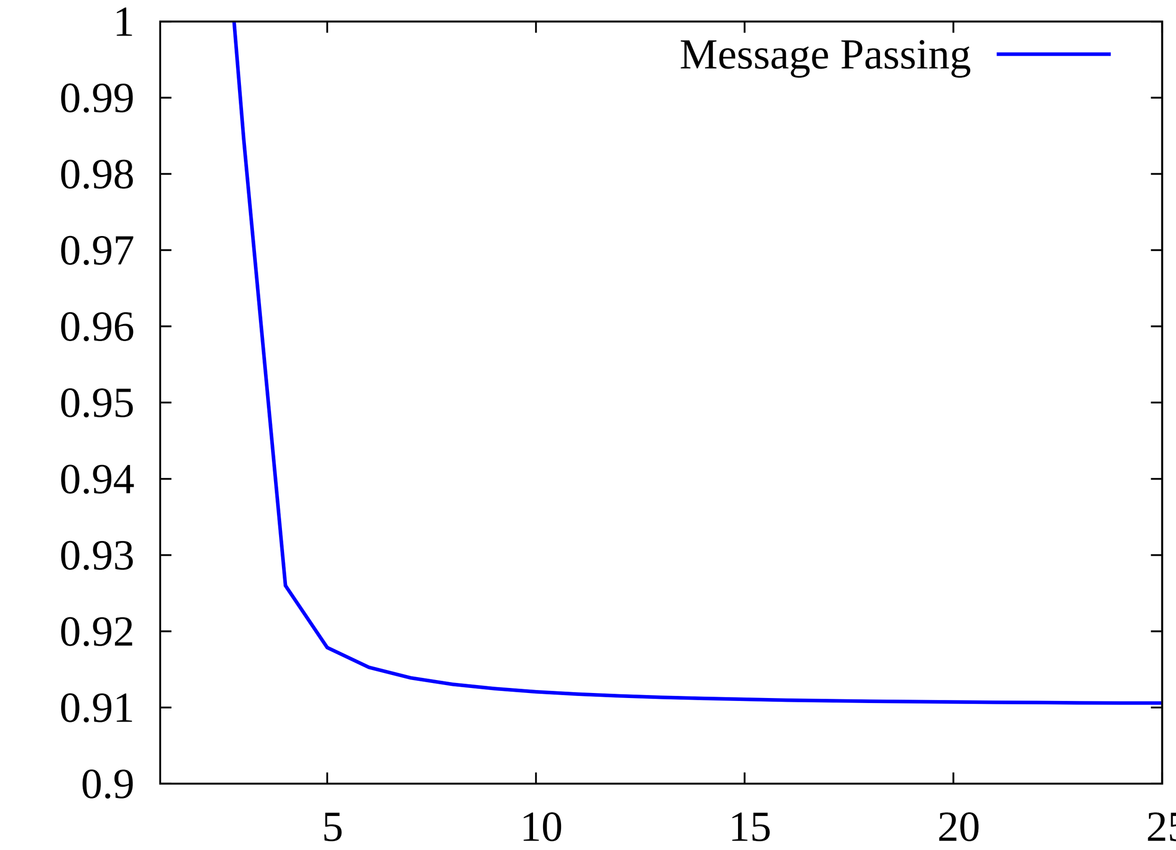


FIGURE 2: RMSE vs. iteration for the Netflix Dataset.

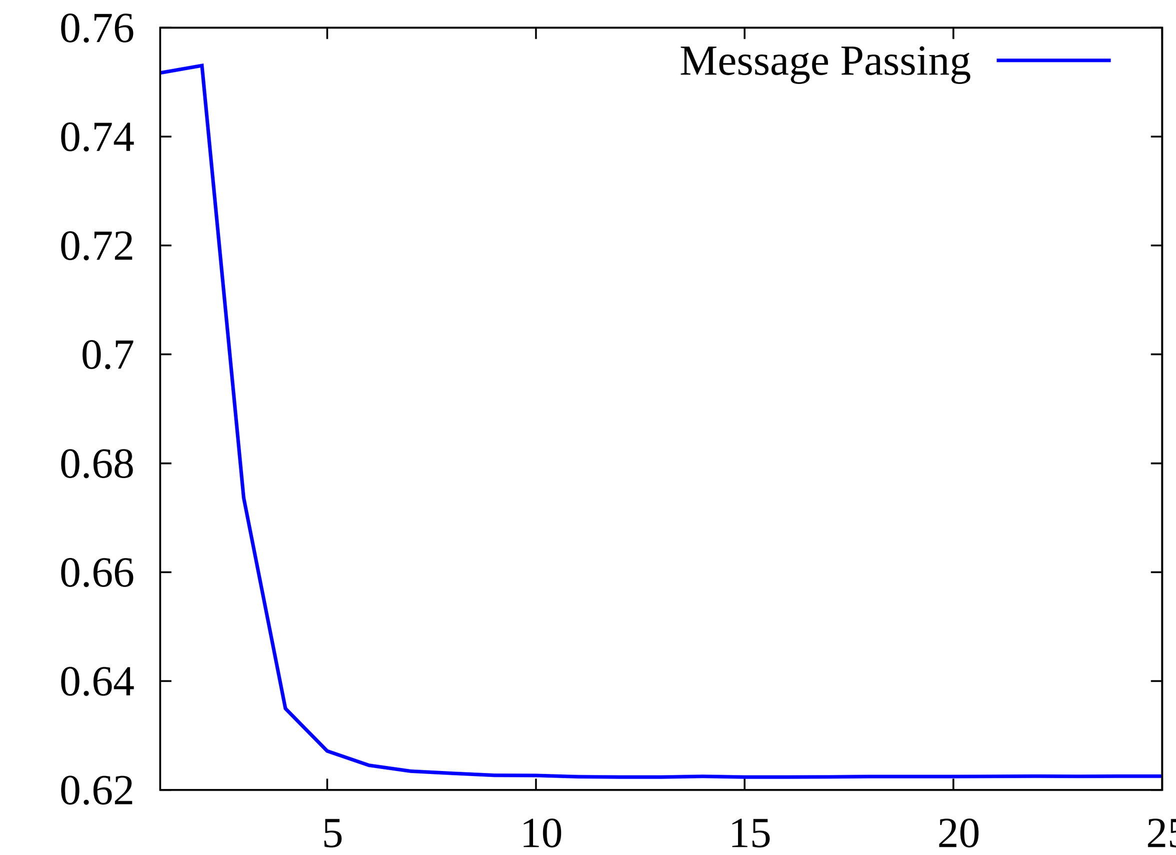


FIGURE 3: MAE vs. iteration for the Movielens Dataset. Best reported MAE : 0.6382 [2]

References

[1] R. Keshavan, A. Montanari, *Message Passing for Matrix Factorization*, in preparation.

[2] L. Candillier, F. Meyer, M. Boullè, *Comparing state-of-the-art collaborative filtering systems*, Lecture Notes in Computer Science 4571 (2007), 548562.