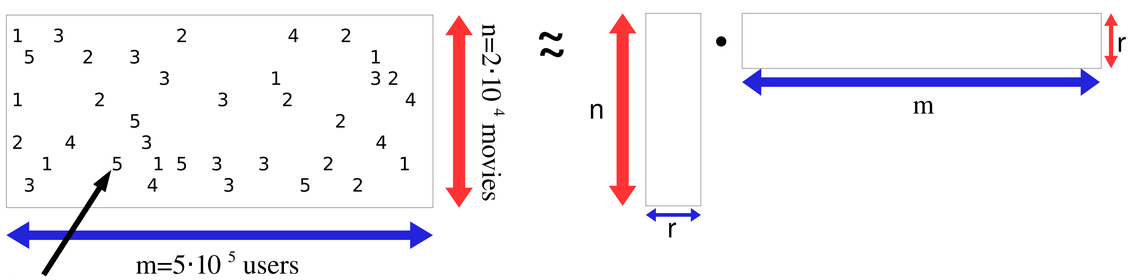
Matrix Completion from a Few Entries

Department of Electrical Engineering
Raghunandan H. Keshavan, Sewoong Oh, Andrea Montanari

What is Matrix Completion?

- Netflix Challenge : A million dollar prize for RMSE < 0.8563

$$M \cong U \Sigma V^T$$

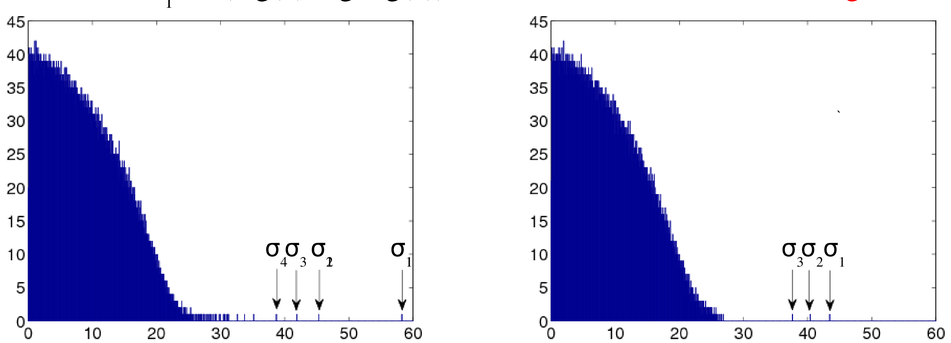


- Problem : How many revealed entries |E| do we need to get $\frac{1}{mn}\|M-\widehat{M}\|_F^2 \le \delta$?

Naïve Approach

- Degrees of freedom = (n+m)r - r^2
- SVD : $Tr(M^E) = \frac{mn}{|E|}\sum_{i=1}^{r} x_i \sigma_i y_i^T$

Problem : $\sigma_1 = \Omega(\log(n)/\log\log(n))$ Solution : Trimming



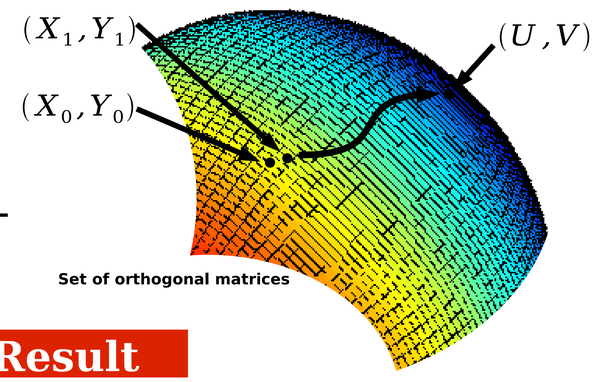
Histogram of singular values of a partially revealed random rank 3 matrix before(left) and after(right) trimming

Algorithm

Algorithm [*Spectral Matrix Completion*]

- Trim : Trim M^E to $\tilde{M}^E$;
- Project : Project $\tilde{M}^E$ onto $\mathrm{Tr}(\tilde{M}^E)$;
- Clean : Minimize Cost F(X,Y), s.t. X,Y orthogonal.

$$F(X,Y) = \min_{S} \sum_{i,j \in E} \left((XSY^T)_{ij} - M_{ij}\right)^2$$

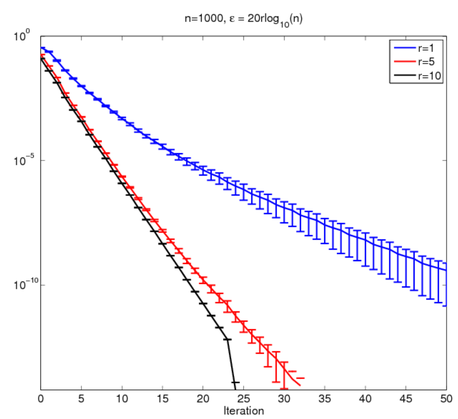


Main Result

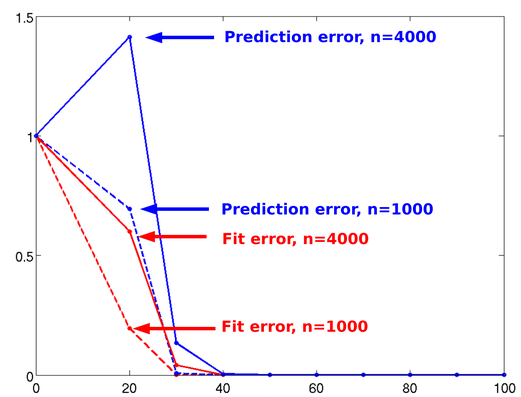
Theorem (Keshavan, Montanari, Oh, 2009)

Assume $r = O(1)$ plus technical conditions (incoherence). If $\epsilon \ge C \log n$ then SPECTRAL MATRIX COMPLETION *returns, whp, the matrix* M.

Implementation



RMSE decreases exponentially w.r.t. the number of iterations. Each point is averaged over 100 instances.



Fit error and prediction error of n x n random rank 5 matrices. Each point is averaged over 50 instances.