Low-rank Matrix Completion from Noisy Entries

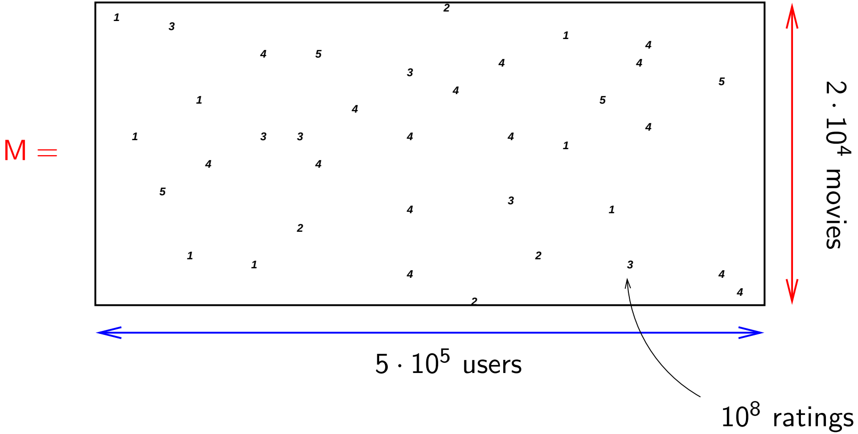
Sewoong Oh

Joint work with Raghunandan Keshavan and Andrea Montanari
Stanford University

Forty-Seventh Allerton Conference
October 1, 2009

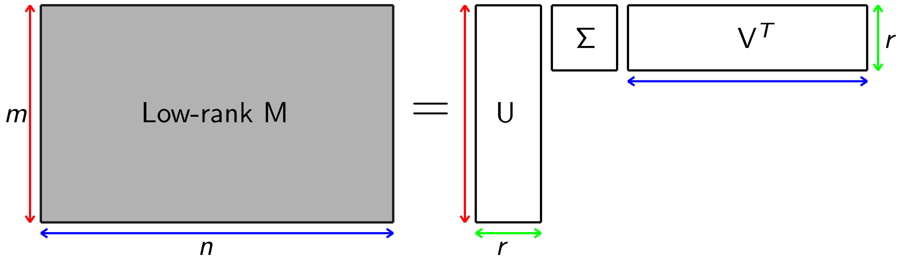
Motivating Example

- Netflix Challenge



The Model

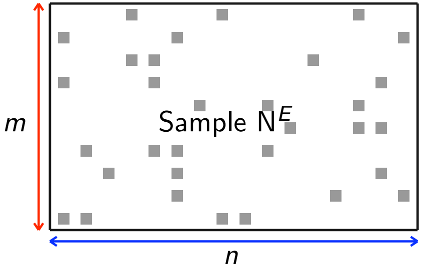
Matrix Completion Problem



1. Low-rank matrix M
2. $\mathrm{N} = \mathrm{M} + \mathrm{Z}$
3. Uniformly random sample E

$$\mathrm{N}^E_{ij} = \begin{cases} \mathrm{M}_{ij} + \mathrm{Z}_{ij} & \text{if } (i,j) \in E\,, \\ 0 & \textit{otherwise}. \end{cases}$$

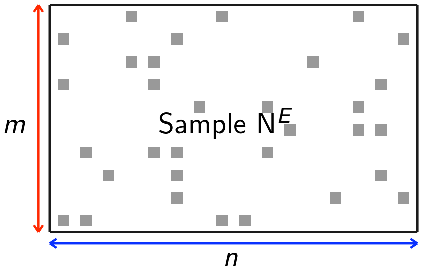
Matrix Completion Problem



1. Low-rank matrix M
2. $\mathrm{N} = \mathrm{M} + \mathrm{Z}$
3. Uniformly random sample E

$$\mathrm{N}^E_{ij} = \begin{cases} \mathrm{M}_{ij} + \mathrm{Z}_{ij} & \text{if } (i,j) \in E \ , \\ 0 & \textit{otherwise}. \end{cases}$$

Matrix Completion Problem



Goal : Estimation $\hat{\mathrm{M}}(E, \mathrm{N}^E)$ that minimizes

$$\mathrm{RMSE} \equiv \left(\frac{1}{mn} \sum_{i,j} (\mathrm{M}_{ij} - \hat{\mathrm{M}}_{ij})^2 \right)^{1/2} .$$

Matrix Completion Problem

- Characterized by 4 parameters:
  - Data size $m \times n$, assume fixed $\alpha \equiv m/n$
  - Rank r
  - Sample size $|E|$
  - Noise Z^E
    $\{$ Running example : $\mathrm{Z}_{ij} \sim$ i.i.d. $\mathrm{N}(0,\sigma_z^2)$ $\}$

Q. Is there an efficient algorithm with performance guarantee?

$$\mathrm{RMSE}_{ALG} \quad \leq \quad F(n, r, |E|, \mathrm{Z}^E)$$

Matrix Completion Problem

- Characterized by 4 parameters:
  - Data size $m \times n$, assume fixed $\alpha \equiv m/n$
  - Rank r
  - Sample size $|E|$
  - Noise Z^E

    { Running example : $\mathrm{Z}_{ij} \sim$ i.i.d. $\mathrm{N}(0,\sigma_z^2)$ }

Q. Is there an efficient algorithm with performance guarantee?

$$\mathrm{RMSE}_{ALG} \quad \leq \quad F(n, r, |E|, \mathrm{Z}^E)$$

Matrix Completion Problem

- Characterized by 4 parameters:
  - Data size $m \times n$, assume fixed $\alpha \equiv m/n$
  - Rank r
  - Sample size $|E|$
  - Noise Z^E

    { Running example : $\mathrm{Z}_{ij} \sim$ i.i.d. $\mathrm{N}(0,\sigma_z^2)$ }

Q. Is there an efficient algorithm with performance guarantee?

$$\mathrm{RMSE}_{OptSpace} \;\leq\; C\,\frac{n\sqrt{r}}{|E|}||\mathrm{Z}^E||_2$$

$$\left(C\,\sigma_z\sqrt{rn\log n/|E|}\;,\;\text{for Gaussian}\right)$$

Matrix Completion Problem

- Characterized by 4 parameters:
  - Data size $m \times n$, assume fixed $\alpha \equiv m/n$
  - Rank r
  - Sample size $|E|$
  - Noise Z^E

    { Running example : $\mathrm{Z}_{ij} \sim$ i.i.d. $\mathrm{N}(0,\sigma_z^2)$ }

Q. Is there an efficient algorithm with performance guarantee?

$$\mathrm{RMSE}_{OptSpace} \leq C\,\frac{n\sqrt{r}}{|E|}||\mathrm{Z}^E||_2$$

$$\left(C\,\sigma_z\,\sqrt{rn\log n/|E|}\,,\ \text{for Gaussian}\right)$$

$$\mathrm{RMSE}_{Oracle} \simeq \frac{1}{\sqrt{|E|}}||\mathrm{Z}^E||_F$$

$$\left(\sigma_z\sqrt{2\,r\,n/|E|}\,,\ \text{for Gaussian}\right)$$

Matrix Completion Problem

- Characterized by 4 parameters:
  - Data size $m \times n$, assume fixed $\alpha \equiv m/n$
  - Rank r
  - Sample size $|E|$
  - Noise Z^E

    { Running example : $Z_{ij} \sim$ i.i.d. $N(0, \sigma_z^2)$ }

Q. Is there an efficient algorithm with performance guarantee?

$$\mathrm{RMSE}_{OptSpace} \leq C \frac{n\sqrt{r}}{|E|} ||Z^E||_2$$

$$\left(C\, \sigma_z \sqrt{rn \log n / |E|}\ , \text{ for Gaussian} \right)$$

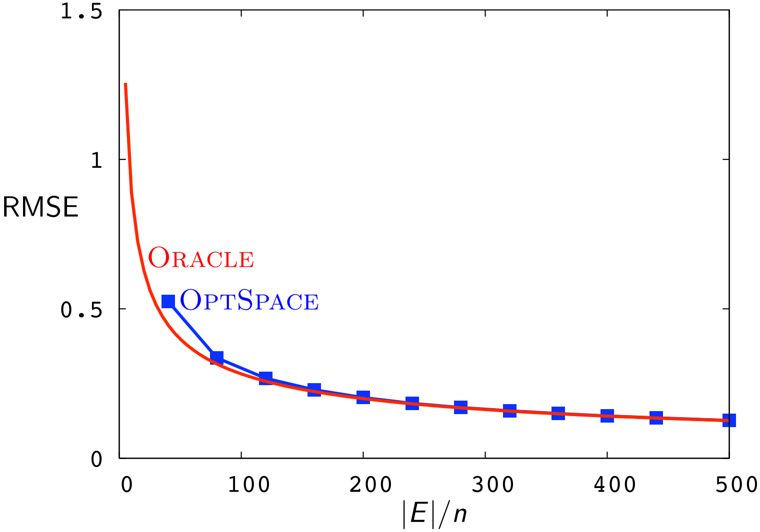
$$\mathrm{RMSE}_{Oracle} \simeq \frac{1}{\sqrt{|E|}} ||Z^E||_F$$

$$\left(\sigma_z \sqrt{2\, r\, n / |E|}\ , \text{ for Gaussian} \right)$$

OptSpace is Near-optimal

Numerical Simulation Results

- Fixed $n = 500, r = 4, \sigma_z = 1$



Main Contribution

OptSpace	
1. Complexity:	Low complexity
2. Theory:	Near-optimal performance guarantee
3. Practice:	Numerical simulations

The Algorithm and Main Theorems

Naïve Approach

$$\mathsf{N}^E = \sum_{k=1}^{n} x_k \sigma_k y_k^T$$

Rank-r projection :

$$\mathcal{P}_r(\mathsf{N}^E) \equiv \frac{mn}{|E|} \sum_{k=1}^{r} x_k \sigma_k y_k^T$$

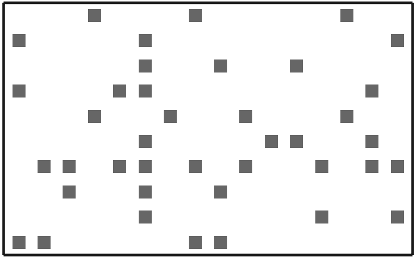
Naïve Approach Fails

- Define : $deg(row_i) \equiv \#$ of samples in row i.
- For $|E| = O(n)$, there exists a row with degree $\Omega(\log n/(\log \log n))$.
- *spurious* singular values of $\Omega(\sqrt{\log n/(\log \log n)})$.

Trimming

- Solution : Trimming

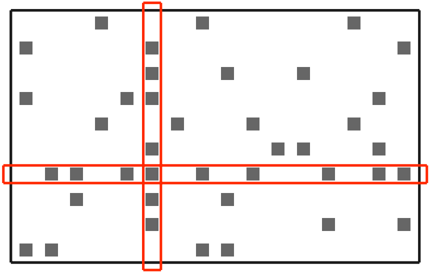
$$\widetilde{\mathsf{N}}^E_{ij} = \begin{cases} 0 & \text{if } deg(row_i) > 2\mathbb{E}[deg(row_i)] \,, \\ 0 & \text{if } deg(col_j) > 2\mathbb{E}[deg(col_i)] \,, \\ \mathsf{N}^E_{ij} & otherwise. \end{cases}$$



Trimming

- Solution : Trimming

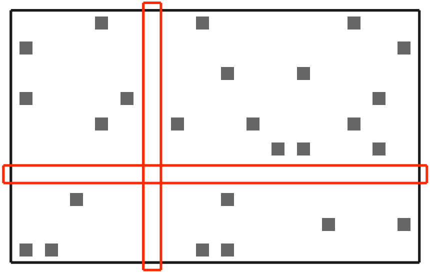
$$\widetilde{\mathrm{N}}_{ij}^{E} = \begin{cases} 0 & \text{if } deg(row_i) > 2\mathbb{E}[deg(row_i)]\ , \\ 0 & \text{if } deg(col_j) > 2\mathbb{E}[deg(col_i)]\ , \\ \mathrm{N}_{ij}^{E} & otherwise. \end{cases}$$



Trimming

- Solution : Trimming

$$\widetilde{\mathsf{N}}_{ij}^{E} = \begin{cases} 0 & \text{if } deg(row_i) > 2\mathbb{E}[deg(row_i)] \ , \\ 0 & \text{if } deg(col_j) > 2\mathbb{E}[deg(col_i)] \ , \\ \mathsf{N}_{ij}^{E} & otherwise. \end{cases}$$



The Algorithm

OptSpace

Input : sample positions E, sample values N^E, rank r

Output : estimation $\hat{\mathrm{M}}$

1: Trim N^E, and let $\widetilde{\mathrm{N}}^E$ be the output;

2: Compute rank-r projection $\mathcal{P}_r(\widetilde{\mathrm{N}}^E) = X_0 S_0 Y_0^T$;

3:

Main Result

Theorem (Keshavan, Montanari, Oh, 2009 Thm. 1.1)

Let M *be an* $n \times n$ *matrix of rank-*r *bounded by* $\mathsf{M}_{\max}$*. Then, w.h.p., rank-*r *projection achieves*

$$\mathrm{RMSE} \;\; \leq \;\; C\mathsf{M}_{\max}\sqrt{\frac{nr}{|E|}} + C'\,\frac{n\sqrt{r}}{|E|}||\mathsf{Z}^E||_2\,.$$

$$\left(\text{Example: } C\mathsf{M}_{\max}\sqrt{\frac{nr}{|E|}} + C'\sigma_z\sqrt{\frac{rn\log n}{|E|}}\right)$$

The Algorithm

OptSpace

Input : sample positions E, sample values N^E, rank r

Output : estimation $\hat{\mathsf{M}}$

1: Trim N^E, and let $\widetilde{\mathsf{N}}^E$ be the output;

2: Compute rank-r projection $\mathcal{P}_r(\widetilde{\mathsf{N}}^E) = X_0 S_0 Y_0^T$;

3: Minimize RMSE by gradient descent starting at (X_0, S_0, Y_0).

Main Result

Theorem (Keshavan, Montanari, Oh, 2009 Thm. 1.1)

Let M *be an* $n \times n$ *matrix of rank-r bounded by* $\mathrm{M}_{\max}$*. Then, w.h.p., rank-r projection achieves*

$$\text{RMSE} \quad \le \quad C\mathrm{M}_{\max}\sqrt{nr/|E|} + C' \, ||\mathrm{Z}^E||_2 \, n\sqrt{r}/|E| \; .$$

Theorem (Keshavan, Montanari, Oh, 2009 Thm. 1.2)

Let M *be an* $n \times n$ *rank-r incoherent matrix with* $\sigma_1(M)/\sigma_r(M) = O(1)$*. If* $|E| \ge C n \, r \max\{r, \log n\}$ *, then, w.h.p.,* OptSpace *achieves*

$$\text{RMSE} \le C'' \frac{n\sqrt{r}}{|E|} ||Z^E||_2 \; ,$$

provided that the RHS is smaller than $\sigma_r(\mathrm{M})$.

$$\left(\text{Example: } C''\sigma_z \sqrt{r \, n \, \log n/|E|}\right)$$

Comparison: Theory

Theorem (Candés, Plan, 2009)

Assume strongly incoherent matrix M. *If* $|E| \geq C\, r\, n\, (\log n)^6$ *then* SEMIDEFINITE PROGRAMMING *achieves, w.h.p.,*

$$\text{RMSE} \leq C' \sqrt{\frac{n}{|E|}}\, ||\mathsf{Z}^E||_F + C'' \frac{1}{n} ||\mathsf{Z}^E||_F \,.$$

$$\left(\text{Example: } C' \sigma_z \sqrt{n} + C'' \sigma_z \frac{\sqrt{|E|}}{n} \right)$$

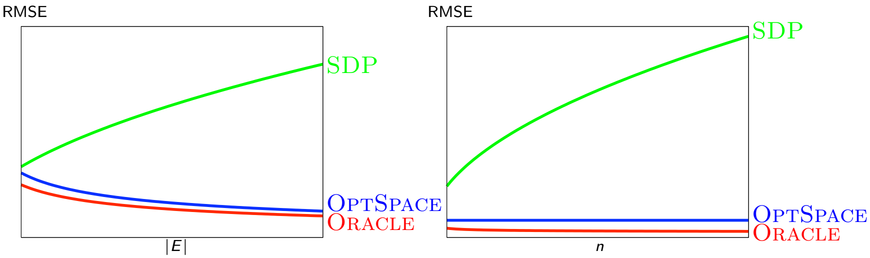
Comparison: Theory

When Z is i.i.d $N(0,\sigma_z^2)$,

$$\textsc{Oracle}: \quad \text{RMSE} \simeq \quad C\sigma_z \sqrt{\frac{r\, n}{|E|}}$$

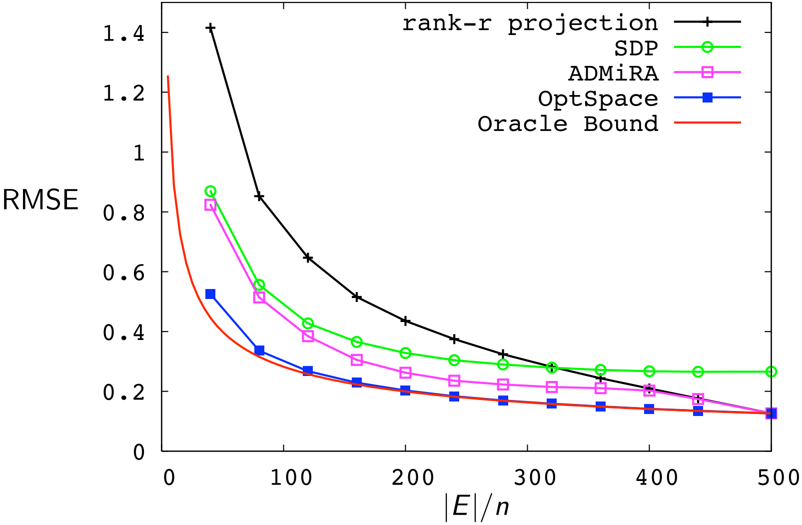
$$\textsc{OptSpace}: \quad \text{RMSE} \leq \quad C'\sigma_z \sqrt{\frac{r\, n \log n}{|E|}}$$

$$\text{SDP}: \quad \text{RMSE} \leq \quad C''\sigma_z \left\{ \sqrt{n} + \frac{\sqrt{|E|}}{n} \right\}$$



Comparison

- Fixed $n = 500, r = 4, \sigma_z = 1$, example from [Candés, Plan, 2009]



Conclusion

	OptSpace
1. Complexity:	Low complexity
2. Theory:	Near-optimal performance guarantee
3. Practice:	Numerical simulations

Thank you!

Conclusion

	OptSpace
1. Complexity:	Low complexity
2. Theory:	Near-optimal performance guarantee
3. Practice:	Numerical simulations

Thank you!