Low-Rank Matrix Completion with Facial Reduction

Shimeng Huang[1]   Henry Wolkowicz[2]

[1]Department of Statistics and Actuarial Science
University of Waterloo

[2]Department of Combinatorics and Optimization
University of Waterloo

Tues. May 23, 9:30AM:

Low-Rank Matrix Completion

Example (Partial Matrix with Noise —— BUT Low Rank)

1.01	2	?
1	?	2.99

Low-Rank Matrix Completion

Example (Partial Matrix with Noise ——- BUT Low Rank)

1.01	2	?
1	?	2.99

1	2	3
1	2	3

Low-Rank Matrix Completion

Example (Partial Matrix with Noise —— BUT Low Rank)

1.01	2	?
1	?	2.99

1	2	3
1	2	3

Problem Statement (non-convex & intractable)

Given a real partial matrix $z \in \mathbb{R}^{\hat{E}}$ with some level of noise,

$$(\text{LRMC}) \qquad \begin{array}{ll} \min & \operatorname{rank}(M) \\ s.t. & \|\mathcal{P}_{\hat{E}}(M) - z\| \leq \delta, \quad M \in \mathbb{R}^{m\times n} \end{array}$$

- $\hat{E}$ indices for known entries (sampled data) in $Z \in \mathbb{R}^{m\times n}$;
  with coordinate projection/partial matrix $z = \mathcal{P}_{\hat{E}}(Z) \in \mathbb{R}^{\hat{E}}$
- $\delta > 0$ is a tuning parameter

Applications Include:

- data science
- model reduction
- collaborative filtering (Netflix problem)
- sensor network localization
- pattern recognition
- various machine learning scenarios

Low-Rank Matrix Completion

Minimizing rank is a hard nonconvex problem

Rank is a lower semi-continuous function.

Nuclear Norm Minimization (convex relaxation)

The problem (LRMC) can be approximated by

$$\text{(NN-LRMC)} \quad \begin{array}{ll} \min & \|M\|_* \\ s.t. & \|\mathcal{P}_{\hat{E}}(M) - z\| \leq \delta \end{array}$$

- $\|M\|_* = \sum_i \sigma_i(M)$, sum of singular values, nuclear norm (Schatten 1-norm, Ky-Fan r-norm, trace norm)
- $\|UXV^T\|_* = \|X\|_*$ unitarily invariant

Nuclear Norm Minimization, Fazel-'02 thesis

Theorem (Fazel,Hindi,Boyd '01)

$\|X\|_*$ *is the convex envelope of* rank X *on* $\{X \in \mathbb{R}^{m \times n} : \|X\| \leq 1\}$.

Properties of nuclear norm:

- *"best"* convex lower approximation of rank function
- The nuclear ball is the convex hull of the intersection of rank-1 matrices with the unit ball:
  $\text{conv}\{uv^T : u \in \mathbb{R}^n, v \in \mathbb{R}^m, \|u\| = 1, \|v\| = 1\}$
- SDP-representable
- Related references by: Candes,Fazel,Parrilo,Recht

SDP Representable

SDP Embedding Lemma

Let $M \in \mathbb{R}^{m\times n}$ and $t \in \mathbb{R}$. Then:

$\|M\|_* \leq t$

if, and only if,

there exist (symmetric) W_1 and W_2 such that

$$\begin{bmatrix} W_1 & M \\ M^T & W_2 \end{bmatrix} \succeq 0, \quad \text{trace}(W_1) + \text{trace}(W_2) \leq 2t.$$

- compact SVD: $M = U\Sigma V^T, \qquad \|M\|_* = \text{trace}\,\Sigma \leq t$
- $\begin{bmatrix} U\Sigma^{1/2} \\ V\Sigma^{1/2} \end{bmatrix} \begin{bmatrix} U\Sigma^{1/2} \\ V\Sigma^{1/2} \end{bmatrix}^T = \begin{bmatrix} U\Sigma U^T & U\Sigma V^T \\ V\Sigma U^T & V\Sigma V^T \end{bmatrix} \succeq 0$
- For necessity, set $W_1 = U\Sigma U^T$, $W_2 = V\Sigma V^T$; for sufficiency, exploit $\boxed{\text{range}\, M \subseteq \text{range}\, W_1, \text{range}\, M^T \subseteq \text{range}\, W_2}$

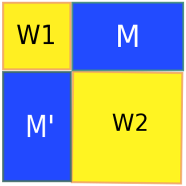
Nuclear Norm Low Rank Problem, (NN-LRMC)

Semidefinite Embedding: Trace Minimization

Problem (NN-LRMC) can be formulated as:

$$\text{(SDP-LRMC)} \qquad \begin{array}{ll} \min & \frac{1}{2}\operatorname{trace}(Y) \\ s.t. & \|\mathcal{P}_{\bar{E}}(Y) - z\| \leq \delta \\ & Y \succeq 0 \end{array}$$

where $Q = \begin{bmatrix} 0 & Z \\ Z^T & 0 \end{bmatrix}$, $z = \mathcal{P}_{\hat{E}}(Z) = \mathcal{P}_{\bar{E}}(Q)$;
$\bar{E}$ is set of indices in Q corresponding to known entries of Z.

Y = 

First, an Example of Facial Reduction, FR

Example (Facial Reduction in Linear Programming)

$$
\begin{array}{ll}
\min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\
\text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\
 & x \geq 0,\ x \in \mathbb{R}^5
\end{array}
$$

First, an Example of Facial Reduction, FR

Example (Facial Reduction in Linear Programming)

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0,\ x \in \mathbb{R}^5 \end{array}$$

If we sum the two constraints we get a *facial* constraint

$$2x_2 + x_3 + 5x_4 = 0 \implies x \in \mathcal{F} = \left\{ x \in \mathbb{R}^5_+ : x_2 = x_3 = x_4 = 0 \right\}$$

First, an Example of Facial Reduction, FR

Example (Facial Reduction in Linear Programming)

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0, \; x \in \mathbb{R}^5 \end{array}$$

If we sum the two constraints we get a *facial* constraint

$$2x_2 + x_3 + 5x_4 = 0 \quad \Longrightarrow \quad x \in \mathcal{F} = \left\{ x \in \mathbb{R}^5_+ : x_2 = x_3 = x_4 = 0 \right\}$$

strict feasibility fails; problem can be reduced

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 7 \end{pmatrix} v \\ \text{s.t.} & \begin{pmatrix} 1 & 1 \end{pmatrix} v = 1 \\ & v \geq 0 \end{array}$$

First Example of Facial Reduction, cont...

Example (Facial Reduction in Linear Programming)

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0 \end{array}$$

First Example of Facial Reduction, cont...

Example (Facial Reduction in Linear Programming)

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Find y with $y^T b = 0, 0 \neq w = A^T y \geq 0$ to get:

$$y = (1\ 1)^T,\ 0 \neq w^T = (A^T y)^T = (0 \quad 2 \quad 1 \quad 5 \quad 0) \geq 0.$$

Then w is an exposing vector of the feasible set:

First Example of Facial Reduction, cont...

Example (Facial Reduction in Linear Programming)

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 5 & -1 & 4 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & -1 & 3 & 1 \\ -1 & 1 & 2 & 2 & -1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Find y with $y^T b = 0, 0 \neq w = A^T y \geq 0$ to get:

$$y = (1\ 1)^T, 0 \neq w^T = (A^T y)^T = \begin{pmatrix} 0 & 2 & 1 & 5 & 0 \end{pmatrix} \geq 0.$$

Then w is an exposing vector of the feasible set:

$$w^T x = 0, \forall \text{ feasible } x \implies x = \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_5 \end{bmatrix}; x_2 = x_3 = x_4 = 0;$$

(simplified) FR problem is

$$\min \left\{ \begin{pmatrix} 2 & 7 \end{pmatrix} v \ : \ \begin{pmatrix} 1 & 1 \end{pmatrix} v = 1, \ v \geq 0 \right\}$$

Faces of a Closed Convex Cone, ccc

Face of a ccc $\mathcal{K}$, $\quad \mathcal{K}+\mathcal{K} \subseteq \mathcal{K}$, $\mathbb{R}\mathcal{K} \subseteq \mathcal{K}$

Let $\mathcal{K}$ be a ccc. A cone $F \subseteq \mathcal{K}$ is a face of $\mathcal{K}$, $F \trianglelefteq \mathcal{K}$, if

$$x, y \in \mathcal{K}, \quad x + y \in F \quad \Rightarrow \quad x, y \in F,$$

If $\emptyset \neq F \subsetneq \mathcal{K}$, then it is called a proper face.

Faces of a Closed Convex Cone, ccc

Face of a ccc $\mathcal{K}$, $\quad \mathcal{K}+\mathcal{K}\subseteq\mathcal{K}, \mathbb{R}\mathcal{K}\subseteq\mathcal{K}$

Let $\mathcal{K}$ be a ccc. A cone $F\subseteq\mathcal{K}$ is a face of $\mathcal{K}$, $F\trianglelefteq\mathcal{K}$, if

$$x,y\in\mathcal{K},\quad x+y\in F \quad\Rightarrow\quad x,y\in F,$$

If $\emptyset\neq F\subsetneq\mathcal{K}$, then it is called a proper face.

Characterization of Faces of PSD Cone $\mathbb{S}^n_+$

Let $X\in\operatorname{relint}(F)$, $F\trianglelefteq\mathbb{S}^n_+$;

let $X=\begin{bmatrix}U & V\end{bmatrix}\begin{bmatrix}D & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}U & V\end{bmatrix}, D\in\mathbb{S}^k_{++}$

be the spectral decomposition.

two views are: $\qquad F=U\mathbb{S}^k_+U^T=\mathbb{S}^n_+\cap(VV^T)^\perp$

Properties of Faces

Some Useful Facts about Faces

- a face of a face is a face;
- an intersection of two faces is a face
- $F_i \trianglelefteq K, F_i = K \cap \phi_i^\perp$, $i = 1, \ldots, k$, implies

$$\cap_i F_i = K \cap (\sum_i \phi_i)^\perp$$

i.e., intersection exposed faces - exposed by sum of exposing vectors

Properties of Faces

Some Useful Facts about Faces

- a face of a face is a face;
- an intersection of two faces is a face
- $F_i \trianglelefteq K, F_i = K \cap \phi_i^\perp$, $i = 1, \ldots, k$, implies

$$\cap_i F_i = K \cap (\sum_i \phi_i)^\perp$$

  i.e., intersection exposed faces - exposed by sum of exposing vectors

For PSD cone

- Self-replicating: a face of a PSD cone is *still* a PSD cone;
- Facially exposed: every face of PSD cone has exposing vector
- Self-dual: $\mathcal{K} = \mathcal{K}^* = \{x : \langle x, y\rangle \geq 0, \forall y \in \mathcal{K}\}$

Back to the Low-Rank Matrix Completion Problem

Recall (SDP-LRMC) Problem: Given $z \in \mathbb{R}^{\hat{E}}$ a partial matrix, find the matrix Z of minimum rank to complete z,
i.e., $\mathcal{P}_{\hat{E}}(Z) = \mathcal{P}_{\bar{E}}(Q) = z$,

Minimize nuclear norm using SDP

$$\text{(SDP-LRMC)} \qquad \begin{array}{ll} \min & \|Y\|_* = \frac{1}{2}\,\text{trace}(Y) \\ \text{s.t.} & \mathcal{P}_{\bar{E}}(Y) = z \\ & Y \succeq 0, \end{array}$$

where $\bar{E}$ is the set of indices in Y that correspond to $\hat{E}$, the known entries of the upper right block of $\begin{bmatrix} 0 & Z \\ Z^T & 0 \end{bmatrix} \in \mathbb{S}_+^{m+n}$.

Back to the Low-Rank Matrix Completion Problem

Recall (SDP-LRMC) Problem: Given $z \in \mathbb{R}^{\hat{E}}$ a partial matrix, find the matrix Z of minimum rank to complete z,
i.e., $\mathcal{P}_{\hat{E}}(Z) = \mathcal{P}_{\bar{E}}(Q) = z$,

Minimize nuclear norm using SDP

$$\text{(SDP-LRMC)} \qquad \begin{array}{rl} \min & \|Y\|_* = \frac{1}{2}\operatorname{trace}(Y) \\ \text{s.t.} & \mathcal{P}_{\bar{E}}(Y) = z \\ & Y \succeq 0, \end{array}$$

where $\bar{E}$ is the set of indices in Y that correspond to $\hat{E}$, the known entries of the upper right block of $\begin{bmatrix} 0 & Z \\ Z^T & 0 \end{bmatrix} \in \mathbb{S}_+^{m+n}$.

- Since the diagonal is free, note that the Slater condition (strict feasibility) does hold for (SDP-LRMC). (And it holds for its dual.)

Facial Reduction of (SDP-LRMC) for Optimal Face

Bipartite Graph, $G_Z = (U_m, V_n, \hat{E})$

With Z and the sampled elements we get a bipartite graph G_Z.

Find Fully Known Submatrix X – a biclique α, $X \cong z[\alpha] \in \mathbb{R}^{p \times q}$

After permutation of rows and columns, WLOG

$$Z = \begin{bmatrix} Z_1 & Z_2 \\ X & Z_3 \end{bmatrix}, \quad z = Z[\hat{E}], \quad \alpha \subseteq \hat{E}, \quad X \cong z[\alpha].$$

Facial Reduction of (SDP-LRMC) for Optimal Face

Bipartite Graph, $G_Z = (U_m, V_n, \hat{E})$

With Z and the sampled elements we get a bipartite graph G_Z.

Find Fully Known Submatrix X – a biclique α, $X \cong z[\alpha] \in \mathbb{R}^{p \times q}$

After permutation of rows and columns, WLOG

$$Z = \begin{bmatrix} Z_1 & Z_2 \\ X & Z_3 \end{bmatrix}, \quad z = Z[\hat{E}], \quad \alpha \subseteq \hat{E}, \quad X \cong z[\alpha].$$

Our algorithm is based on finding bicliques in G_Z; we do this by finding (nontrivial/nondiagonal-block) cliques within symmetric matrix Y.

$$Y = \begin{bmatrix} W_1 & Z \\ Z^T & W_2 \end{bmatrix}$$

Bipartite Graph and Biclique

Partial matrix

$$z \cong \begin{bmatrix} -5 & NA & 10 & -20 & NA & -6 \\ 4 & 0 & 4 & 4 & 6 & 6 \\ -3 & NA & NA & 32 & 27 & NA \\ 5 & NA & 0 & 10 & 12 & NA \\ NA & -30 & NA & NA & 27 & NA \\ 3 & -5 & -2 & 8 & NA & 4 \\ 5 & 5 & NA & 0 & 3 & NA \end{bmatrix}, \quad \hat{E} = \{11, 13, 14, 16, 21, \ldots, 74, 75\}$$

biclique indices: $\bar{U}_m = \{6, 1, 2\}$, $\bar{V}_n = \{1, 4, 3, 6\}$, $\alpha = \{61, 64, 63, 66, 11, \ldots, 26\}$

$$z[\alpha] \equiv X = \begin{bmatrix} 3 & 8 & -2 & 4 \\ -5 & -20 & 10 & -6 \\ 4 & 4 & 4 & 6 \end{bmatrix}.$$

$$Y[\alpha] = \begin{bmatrix} \text{FREE} & \begin{matrix} 3 & 8 & -2 & 4 \\ -5 & -20 & 10 & -6 \\ 4 & 4 & 4 & 6 \end{matrix} \\ \begin{matrix} 3 & -5 & 4 \\ 8 & -20 & 4 \\ -2 & 10 & 4 \\ 4 & -6 & 6 \end{matrix} & \text{FREE} \end{bmatrix}$$

Our View of Facial Reduction and Exposed Faces

Theorem (Drusvyatskiy, Pataki, W. '15)

Linear transformation $\mathcal{M}: \mathbb{S}^n \to \mathbb{R}^m$, *adjoint* $\mathcal{M}^*$; *feasible set* $\mathcal{F} := \{X \in \mathbb{S}^n_+ : \mathcal{M}(X) = b\} \neq \emptyset$, $b \in \mathbb{R}^m$. *Then a vector* v *exposes a proper face of* $\mathcal{M}(\mathbb{S}^n_+)$ *containing* $b \iff v$ *satisfies the auxiliary system*

$$0 \neq \mathcal{M}^* v \in \mathbb{S}^n_+ \quad \text{and} \quad \langle v, b \rangle = 0.$$

Let N *denote smallest face of* $\mathcal{M}(\mathbb{S}^n_+)$ *containing* b*. Then:*

1. $\mathbb{S}^n_+ \cap \mathcal{M}^{-1} N = \text{face}(\mathcal{F})$, *the smallest face containing* $\mathcal{F}$.
2. *For any vector* $v \in \mathbb{R}^m$ *the following equivalence holds:*

$$v \text{ exposes } N \iff \mathcal{M}^* v \text{ exposes } \text{face}(\mathcal{F})$$

Noisy sensor network localization: robust facial reduction and the Pareto frontier

D. Drusvyatskiy, N. Krislock, Y-L. Cheung Voronin, and H. W. '16

Facial Reduction for (SDP-LRMC), r is target rank for Z

Biclique $\alpha \cong$ of G_Z, $z[\alpha] \equiv X \in \mathbb{R}^{p\times q}$

target rank $r \le \min\{p,q\} < \max\{p,q\}$;

WLOG

$$Z = \begin{bmatrix} Z_1 & Z_2 \\ X & Z_3 \end{bmatrix},$$

SVD: $$X = \begin{bmatrix} U_1 & U_X \end{bmatrix} \begin{bmatrix} \Sigma \in \mathbb{S}^r_{++} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} V_1 & V_X \end{bmatrix}^T$$

We get full rank factorization

$$X = \bar{P}\bar{Q}^T = U_1 \Sigma V_1^T, \quad \bar{P} = U_1 \Sigma^{1/2}, \ \bar{Q} = V_1 \Sigma^{1/2}.$$

Since *rank* is lower semi-continuous: $\operatorname{rank} X = \operatorname{rank} Z$ generically.
In fact our tests form: $Z = \bar{P}\bar{Q}^T$
with $\bar{P}, \bar{Q}$ random, i.i.d. and full column rank r.

FR using Optimal Y

Rewrite Optimal Y

Assuming we have obtained the desired target rank $Y = r$

$$0 \preceq Y = \begin{bmatrix} U \\ P \\ Q \\ V \end{bmatrix} D \begin{bmatrix} U \\ P \\ Q \\ V \end{bmatrix}^T = \left[\begin{array}{c|cc|c} UDU^T & UDP^T & UDQ^T & UDV^T \\ \hline PDU^T & PDP^T & \boxed{PDQ^T} & PDV^T \\ QDU^T & QDP^T & QDQ^T & QDV^T \\ \hline VDU^T & VDP^T & VDQ^T & VDV^T \end{array}\right]$$

And assume rank $X = r$

$$X = PDQ^T = \bar{P}\bar{Q}^T.$$

implies the ranges satisfy
$U_1^T U_X = P^T U_X = 0, V_1^T V_X = Q^T V_X = 0$

$$\begin{array}{rclclcl} \operatorname{range}(X) & = & \operatorname{range}(P) & = & \operatorname{range}(\bar{P}) & = & \operatorname{range}(U_1), \\ \operatorname{range}(X^T) & = & \operatorname{range}(Q) & = & \operatorname{range}(\bar{Q}) & = & \operatorname{range}(V_1). \end{array}$$

Constructing Exposing Vectors

Key for facial reduction

We can use an exposing vector formed as $U_X U_X^T$ for block PDP^T as well as $V_X V_X^T$ for block QDQ^T and add appropriate blocks of zeros:

$$
\begin{array}{rcl}
W_X & = & \left[\begin{array}{c|cc|c} 0 & 0 & 0 & 0 \\ \hline 0 & U_X U_X^T & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right] + \left[\begin{array}{c|cc|c} 0 & 0 & 0 & 0 \\ \hline 0 & 0 & 0 & 0 \\ 0 & 0 & V_X V_X^T & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right] \\
 & = & \left[\begin{array}{c|cc|c} 0 & 0 & 0 & 0 \\ \hline 0 & U_X U_X^T & 0 & 0 \\ 0 & 0 & V_X V_X^T & 0 \\ \hline 0 & 0 & 0 & 0 \end{array}\right].
\end{array}
$$

All three matrices provide exposing vectors.

Facial reduction from exposing vector

$$
F^* \trianglelefteq T \mathbb{S}_+^{((n+m)-(p+q-2r))} T^T, \quad \text{range } T = \text{null } W_X.
$$

Measuring Noise of Biclique $\alpha \in \Theta$

Biclique: $\alpha \subseteq \hat{E}$, $z[\alpha] \cong X \in \mathbb{R}^{p\times q}$, target rank r

singular values of X: $\sigma_1 \geq ... \geq \sigma_{\min\{p,q\}}$

biclique noise: $$u_X^P := \frac{\sum_{i=r+1}^{\min\{p,q\}} \sigma_i^2}{0.5p(p-1)} \quad u_X^Q := \frac{\sum_{i=r+1}^{\min\{p,q\}} \sigma_i^2}{0.5q(q-1)}$$

Assign biclique weight

Total noise of all bicliques: $S = \sum_{X\in\Theta}(u_X^P + u_X^Q)$

for each $\alpha \in \Theta$: $$w_X^P = 1 - \frac{u_X^P}{S}, \quad w_X^Q = 1 - \frac{u_X^Q}{S}$$

Facial Reduction Process

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes
- Find corresponding exposing vectors $\{Y_\alpha^{expo}\}_{\alpha\in\Theta}$ calculate their weights $\{\omega_\alpha\}_{\alpha\in\Theta}$

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes
- Find corresponding exposing vectors $\{Y_\alpha^{expo}\}_{\alpha\in\Theta}$ calculate their weights $\{\omega_\alpha\}_{\alpha\in\Theta}$
- Calculate the weighted sum of all the exposing vectors

$$Y_{Final}^{expo} = \sum_{\alpha\in\Theta} \omega_\alpha Y_\alpha^{expo}$$

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes
- Find corresponding exposing vectors $\{Y_\alpha^{expo}\}_{\alpha\in\Theta}$ calculate their weights $\{\omega_\alpha\}_{\alpha\in\Theta}$
- Calculate the weighted sum of <u>all</u> the exposing vectors

$$Y_{Final}^{expo} = \sum_{\alpha\in\Theta} \omega_\alpha Y_\alpha^{expo}$$

- Find full column rank V such that range V = null Y_{Final}^{expo}.

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes
- Find corresponding exposing vectors $\{Y_\alpha^{expo}\}_{\alpha\in\Theta}$ calculate their weights $\{\omega_\alpha\}_{\alpha\in\Theta}$
- Calculate the weighted sum of <u>all</u> the exposing vectors

$$Y_{Final}^{expo} = \sum_{\alpha\in\Theta} \omega_\alpha Y_\alpha^{expo}$$

- Find full column rank V such that range $V =$ null Y_{Final}^{expo}.
- Solve equivalent smaller problem based on smaller dimensional matrix R, where

$$Y = VRV^T$$

Facial Reduction Process

- Find set of bicliques Θ, of appropriate sizes
- Find corresponding exposing vectors $\{Y_\alpha^{expo}\}_{\alpha\in\Theta}$ calculate their weights $\{\omega_\alpha\}_{\alpha\in\Theta}$
- Calculate the weighted sum of all the exposing vectors

$$Y_{Final}^{expo} = \sum_{\alpha\in\Theta} \omega_\alpha Y_\alpha^{expo}$$

- Find full column rank V such that range $V = \text{null } Y_{Final}^{expo}$.
- Solve equivalent smaller problem based on smaller dimensional matrix R, where

$$Y = VRV^T$$

- (Follows the framework in Drusvyatskiy/Krislock/Cheung-Voronin/W.)

Exploit block structure

Y_{Final}^{expo} has block structure so V has a block structure too:

$$Y_{Final}^{expo} = \begin{bmatrix} \sum_{X \in \mathcal{C}} w_X^P W_X^P & 0 \\ 0 & \sum_{X \in \mathcal{C}} w_X^Q W_X^Q \end{bmatrix}, \quad V = \begin{bmatrix} V_P & 0 \\ 0 & V_Q \end{bmatrix}$$

Exploit block structure

Y_{Final}^{expo} has block structure so V has a block structure too:

$$Y_{Final}^{expo} = \begin{bmatrix} \sum_{X \in \mathcal{C}} w_X^P W_X^P & 0 \\ 0 & \sum_{X \in \mathcal{C}} w_X^Q W_X^Q \end{bmatrix}, \quad V = \begin{bmatrix} V_P & 0 \\ 0 & V_Q \end{bmatrix}$$

allows a computational speed up for eigenvalue subproblems.

Noiseless Case

FR dramatically reduces dimension of now overdetermined problem:

$$\begin{array}{ll} \min & \operatorname{trace}(R) \qquad\qquad (= \operatorname{trace}(VRV^T)) \\ \text{s.t.} & \mathcal{P}_{\bar{E}}(V_P R_{pq} V_Q^T) = z \\ & R = \begin{bmatrix} R_p & R_{pq} \\ R_{pq}^T & R_q \end{bmatrix} \succeq 0. \end{array}$$

Noiseless Case

FR dramatically reduces dimension of now overdetermined problem:

$$\begin{array}{ll} \min & \operatorname{trace}(R) \qquad\qquad (= \operatorname{trace}(VRV^T)) \\ \text{s.t.} & \mathcal{P}_{\bar{E}}(V_P R_{pq} V_Q^T) = z \\ & R = \begin{bmatrix} R_p & R_{pq} \\ R_{pq}^T & R_q \end{bmatrix} \succeq 0. \end{array}$$

remove the redundant constraints

Use a compact QR to find well-conditioned full rank matrix representation. A simple semidefinite constrained least squares solution may be enough!

$$\min_{R \in \mathbb{S}_+^{r_v}} \|\mathcal{P}_{\tilde{E}}(V_P R_{pq} V_Q^T) - \tilde{z})\|.$$

(Here $\tilde{E}, \tilde{z}$ denote the corresponding entries after removing redundant constraints. Often R found explicitly.)

Noisy Case

Cannot simply remove redundant constraints;
use random sketch matrix A to reduce the number of constraints;
first solve:

$$\delta_0 = \min_{R \in \mathbb{S}_+^{r_v}} \left\| A \left(\mathcal{P}_{\hat{E}}(V_P R_{pq} V_Q^T) - z \right) \right\|.$$

and hopefully obtain the target rank!

Noisy Case

Cannot simply remove redundant constraints;
use random sketch matrix A to reduce the number of constraints;
first solve:

$$\delta_0 = \min_{R \in \mathbb{S}_+^{r_v}} \left\| A \left(\mathcal{P}_{\hat{E}}(V_P R_{pq} V_Q^T) - z \right) \right\|.$$

and hopefully obtain the target rank!
Otherwise, we use a refinement step.

Refinement Step in the Noisy Case

We would like to reduce the rank after the previous step using a parametric approach:

$$\begin{array}{lcrl} \min & \operatorname{trace}(R) & & \\ \text{s.t.} & \left\| A\left(\mathcal{P}_{\hat{E}}(V_P R_{pq} V_Q^T) - b\right)\right\| & \leq & \delta_0 \\ & R & \succeq & 0. \end{array}$$

Refinement Step in the Noisy Case

We would like to reduce the rank after the previous step using a parametric approach:

$$\begin{array}{lrcl} \min & \text{trace}(R) & & \\ \text{s.t.} & \left\|A\left(\mathcal{P}_{\hat{E}}(V_P R_{pq} V_Q^T) - b\right)\right\| & \leq & \delta_0 \\ & R & \succeq & 0. \end{array}$$

To ensure the rank can be reduce, we flip the problem:

$$\begin{array}{rl} \varphi(\tau) := \min & \left\|A\left(\mathcal{P}_{\hat{E}}(V_P R_{pq} V_Q^T) - b\right)\right\| + \gamma\|R\|_F \\ \text{s.t.} & \text{trace}(R) \leq \tau \\ & R \succeq 0. \end{array}$$

where γ is a regularization parameter, since the least squares problem can be underdetermined.

Sample Results ($\approx 3x10^6$ variables)

Table: noiseless: $r = 8$; $m \times n$ size; density p; mean 20 instances.

Specifications			r_v	Rcvrd (%Z)	Time (s)	Rank	Residual (%Z)
m	n	mean(p)					
1000	3000	0.53	16.10	96.39	37.29	8.0	1.1072e-10
1000	3000	0.50	17.65	88.99	36.50	8.0	4.6569e-10
1000	3000	0.48	32.15	71.66	72.14	8.5	2.0413e-07

Table: noisy: $r = 2$; $m \times n$ size; density p; mean 20 instances.

Specifications				Rcvd (%Z)	Time (s)		Rank		Residual (%Z)	
m	n	% noise	p		initial	refine	initial	refine	initial	refine
1100	3000	0.50	0.33	100.00	33.72	48.53	2.00	2.00	8.53e-03	8.53e-03
1100	3000	1.00	0.33	100.00	33.67	49.09	2.00	2.00	2.70e-02	2.70e-02
1100	3000	2.00	0.33	100.00	34.13	48.84	2.00	2.00	9.75e-02	9.75e-02
1100	3000	3.00	0.33	100.00	36.34	92.73	5.00	5.00	5.48e-01	1.40e-01
1100	3000	4.00	0.33	100.00	51.45	186.28	11.00	8.00	1.25e+00	1.28e-01

Conclusion

Preprocessing

- Though strict feasibility holds generically, failure appears in many applications. Loss of strict feasibility is directly related to ill-posedness and difficulty in numerical methods.
- Preprocessing based on structure can both *regularize* and simplify the problem. In many cases one gets an optimal solution without the need of any SDP solver.
  (New Survey FR: Drusvyatskiy and W. '17)

Exploit structure at optimum

For low-rank matrix completion the structure at the optimum can be exploited to apply FR even though strict feasibility holds.

To do: reduce density/more refinement; real life applications

Thanks for your attention! Questions?

Low-Rank Matrix Completion with Facial Reduction

Shimeng Huang[1] Henry Wolkowicz[2]

[1]Department of Statistics and Actuarial Science
University of Waterloo

[2]Department of Combinatorics and Optimization
University of Waterloo

References

E.J. Candès and B. Recht, *Exact matrix completion via convex optimization*, Found. Comput. Math. 9 (2009), no. 6, 717–772. MR 2565240

D. Drusvyatskiy, N. Krislock, Y-L. Cheung Voronin, and H. Wolkowicz, *Noisy sensor network localization: robust facial reduction and the Pareto frontier*, SIAM Journal on Optimization (2017), 1–30, arXiv:1410.6852, 20 pages, to appear.

D. Drusvyatskiy, G. Pataki, and H. Wolkowicz, *Coordinate shadows of semidefinite and Euclidean distance matrices*, SIAM J. Optim. 25 (2015), no. 2, 1160–1178. MR 3357643

D. Drusvyatskiy, S. Sremac, and H. Wolkowicz, *Three views of facial reduction in cone optimization*, Tech. report, University of Waterloo, Waterloo, Ontario, 2017, survey in progress.

M. Fazel, *Matrix rank minimization with applications*, Ph.D. thesis, Stanford University, Stanford, CA, 2001.

M. Fazel, H. Hindi, and S.P. Boyd, *A rank minimization heuristic with application to minimum order system approximation*, Proceedings American Control Conference, 2001, pp. 4734–4739.

B. Recht, M. Fazel, and P. Parrilo, *Guaranteed minimum-rank solutions of linear matrix equations via nuclear norm minimization*, SIAM Rev. 52 (2010), no. 3, 471–501. MR 2680543