The many faces of degeneracy in conic optimization

Dmitriy Drusvyatskiy
Mathematics, Washington

Joint work with
N. Krislock (NIU), G. Pataki (UNC),
Y.-L. Voronin (Boulder), and H. Wolkowicz (Waterloo)

AFOSR YIP

Review of degeneracy

Primal-dual pair:

$$
\begin{array}{llll}
(P) & \min & \operatorname{tr} CX \\
 & \text{s.t.} & \mathcal{A}(X) = b \\
 & & X \succeq 0
\end{array}
\qquad
\begin{array}{llll}
(D) & \max & b^T y \\
 & \text{s.t.} & \mathcal{A}^* y \preceq C
\end{array}
$$

Review of degeneracy

Primal-dual pair:

$$
\begin{array}{llll}
(P) & \min & \operatorname{tr} CX \\
 & \text{s.t.} & \mathcal{A}(X) = b \\
 & & X \succeq 0
\end{array}
\qquad\qquad
\begin{array}{lll}
(D) & \max & b^T y \\
 & \text{s.t.} & \mathcal{A}^* y \preceq C
\end{array}
$$

where

$$\langle C, X\rangle = \operatorname{tr} CX, \qquad \mathcal{A}(X) = \Big(\langle A_1, X\rangle, \langle A_2, X\rangle, \ldots, \langle A_m, X\rangle\Big)$$

and then

$$\mathcal{A}^* y = \sum_{i=1}^{m} y_i A_i.$$

Review of degeneracy

Primal-dual pair:

(P)	$\min$	$\operatorname{tr} CX$	(D)	$\max$	$b^T y$
	s.t.	$\mathcal{A}(X) = b$		s.t.	$\mathcal{A}^* y \preceq C$
		$X \succeq 0$			

where

$$\langle C, X \rangle = \operatorname{tr} CX, \qquad \mathcal{A}(X) = \Big(\langle A_1, X \rangle, \langle A_2, X \rangle, \ldots, \langle A_m, X \rangle \Big)$$

and then
$$\mathcal{A}^* y = \sum_{i=1}^{m} y_i A_i.$$

Slater's condition for (P): $\exists X \succ 0$ satisfying $\mathcal{A}(X) = b$

Slater's condition for (D): $\exists y$ satisfying $\mathcal{A}^* y \prec C$

Review of degeneracy

Primal-dual pair:

$$
\begin{array}{llll}
(P) & \min & \operatorname{tr} CX \\
 & \text{s.t.} & \mathcal{A}(X) = b \\
 & & X \succeq 0
\end{array}
\qquad\qquad
\begin{array}{llll}
(D) & \max & b^T y \\
 & \text{s.t.} & \mathcal{A}^* y \preceq C
\end{array}
$$

where

$$\langle C, X\rangle = \operatorname{tr} CX, \qquad \mathcal{A}(X) = \Big(\langle A_1, X\rangle, \langle A_2, X\rangle, \ldots, \langle A_m, X\rangle\Big)$$

and then

$$\mathcal{A}^* y = \sum_{i=1}^{m} y_i A_i.$$

Slater's condition for (P): $\exists X \succ 0$ satisfying $\mathcal{A}(X) = b$
Slater's condition for (D): $\exists y$ satisfying $\mathcal{A}^* y \prec C$

Slater in (P) $\Rightarrow$

- Strong duality: $val(P) = val(D)$ and (D) is attained.
- Bounded dual solutions
- Stability relative to b

Review of degeneracy

Primal-dual pair:

(P)	min	$\operatorname{tr} CX$	(D)	max	$b^T y$
	s.t.	$\mathcal{A}(X) = b$		s.t.	$\mathcal{A}^* y \preceq C$
		$X \succeq 0$			

where

$$\langle C, X \rangle = \operatorname{tr} CX, \qquad \mathcal{A}(X) = \Big(\langle A_1, X \rangle, \langle A_2, X \rangle, \ldots, \langle A_m, X \rangle \Big)$$

and then

$$\mathcal{A}^* y = \sum_{i=1}^{m} y_i A_i.$$

Slater's condition for (P): $\exists X \succ 0$ satisfying $\mathcal{A}(X) = b$

Slater's condition for (D): $\exists y$ satisfying $\mathcal{A}^* y \prec C$

Slater in (P) $\Rightarrow$

- Strong duality: $val(P) = val(D)$ and (D) is attained.
- Bounded dual solutions
- Stability relative to b

Slater (D) often holds in applications, but Slater (P) may fail.

Degeneracy

Eg: Structured data

$$\begin{bmatrix} 1 & 1 & ? \\ 1 & 1 & 1 \\ ? & 1 & 1 \end{bmatrix} \succeq 0$$

Degeneracy

Eg: Structured data

$$\begin{bmatrix} 1 & 1 & ? \\ 1 & 1 & 1 \\ ? & 1 & 1 \end{bmatrix} \succeq 0$$

or

$$\begin{bmatrix} x & y & z \\ y & -x & y \\ z & y & 1 \end{bmatrix} \succeq 0$$

Degeneracy

Eg: Structured data

$$\begin{bmatrix} 1 & 1 & ? \\ 1 & 1 & 1 \\ ? & 1 & 1 \end{bmatrix} \succeq 0$$

or

$$\begin{bmatrix} x & y & z \\ y & -x & y \\ z & y & 1 \end{bmatrix} \succeq 0$$

More interesting examples later!

Detecting degeneracy

Exactly one holds (statement of alternative):

- Slater (P)
- The auxiliary system

$$\boxed{0 \neq \mathcal{A}^* y \succeq 0, \quad b^T y \leq 0} \qquad \text{is consistant.}$$

Detecting degeneracy

Exactly one holds (statement of alternative):

- Slater (P)
- The auxiliary system

$$\boxed{0 \neq \mathcal{A}^* y \succeq 0, \quad b^T y \leq 0} \qquad \text{is consistant.}$$

Distance to (P)-infeasibility (Renegar): infimum of

$$\|(\widehat{\mathcal{A}}, \widehat{b})\|$$

such that the system

$$\left\{ \begin{array}{l} (\mathcal{A} + \widehat{\mathcal{A}})(X) = b + \widehat{b} \\ X \succeq 0 \end{array} \right\} \qquad \text{is infeasible.}$$

Detecting degeneracy

Exactly one holds (statement of alternative):

- Slater (P)
- The auxiliary system

$$\boxed{0 \neq \mathcal{A}^* y \succeq 0, \quad b^T y \leq 0} \qquad \text{is consistant.}$$

Distance to (P)-infeasibility (Renegar): infimum of

$$\|(\widehat{\mathcal{A}}, \widehat{b})\| \qquad =: \max(\|\widehat{\mathcal{A}}\|_{op}, \|\widehat{b}\|)$$

such that the system

$$\left\{ \begin{array}{l} (\mathcal{A} + \widehat{\mathcal{A}})(X) = b + \widehat{b} \\ X \succeq 0 \end{array} \right\} \qquad \text{is infeasible.}$$

Detecting degeneracy

Exactly one holds (statement of alternative):

- Slater (P)
- The auxiliary system

$$\boxed{0 \neq \mathcal{A}^* y \succeq 0, \quad b^T y \leq 0} \qquad \text{is consistant.}$$

Distance to (P)-infeasibility (Renegar): infimum of

$$\|(\widehat{\mathcal{A}}, \widehat{b})\| \qquad =: \max(\|\widehat{\mathcal{A}}\|_{op}, \|\widehat{b}\|)$$

such that the system

$$\left\{ \begin{array}{l} (\mathcal{A} + \widehat{\mathcal{A}})(X) = b + \widehat{b} \\ X \succeq 0 \end{array} \right\} \qquad \text{is infeasible.}$$

(Renegar):

$$\text{distance to (P)-infeasibility} = \min_{y : \|y\| = 1} \max\{\|\lambda_-(\mathcal{A}^* y)\|, b^T y\}$$

Facial reduction

(P) feasible and y satisfies the auxiliary system $\Longrightarrow$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^\perp \cap \mathcal{S}_+^n.$$

Facial reduction

(P) feasible and y satisfies the auxiliary system $\implies$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Facial reduction

(P) feasible and y satisfies the auxiliary system $\implies$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Singularity degree (Sturm '98):
d = minimal # of facial reduction iterations required.

Facial reduction

(P) feasible and y satisfies the auxiliary system $\implies$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Singularity degree (Sturm '98):
d = minimal # of facial reduction iterations required.

Connection to error bounds (Sturm '98):

$$\frac{\mathrm{dist}_\Omega(Z)}{\left(\mathrm{dist}_{\mathcal{S}^n_+}(Z) + \mathrm{dist}_{\mathcal{A}^{-1}b}(Z)\right)^{2^d}} \quad \text{is bounded on compact sets}$$

Facial reduction

(P) feasible and y satisfies the auxiliary system $\implies$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^{\perp} \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^{\perp} \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Singularity degree (Sturm '98):
d = minimal # of facial reduction iterations required.

Connection to error bounds (Sturm '98):

$$\frac{\mathrm{dist}_{\Omega}(Z)}{\left(\mathrm{dist}_{\mathcal{S}^n_+}(Z) + \mathrm{dist}_{\mathcal{A}^{-1}b}(Z)\right)^{2^d}} \quad \text{is bounded on compact sets}$$

(D-Pataki-Wolkowicz '14): If (P) is degenerate, then

$$d = 1 \quad \iff \quad \mathrm{face}\Big(b, \mathcal{A}(\mathcal{S}^n_+)\Big) \text{ is exposed.}$$

Facial reduction

(P) feasible and y satisfies the auxiliary system $\implies$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^\perp \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Singularity degree (Sturm '98):
d = minimal # of facial reduction iterations required.

Connection to error bounds (Sturm '98):

$$\frac{\mathrm{dist}_\Omega(Z)}{\left(\mathrm{dist}_{\mathcal{S}^n_+}(Z) + \mathrm{dist}_{\mathcal{A}^{-1}b}(Z)\right)^{2^d}} \qquad \text{is bounded on compact sets}$$

(D-Pataki-Wolkowicz '14): If (P) is degenerate, then

$$d = 1 \quad \iff \quad \mathrm{face}\Big(b, \mathcal{A}(\mathcal{S}^n_+)\Big) \text{ is exposed.}$$

Open question:
When is $\mathcal{A}(\mathcal{S}^n_+)$ facially exposed?

Facial reduction

(P) feasible and y satisfies the auxiliary system $\Longrightarrow$

$$\Omega = \{X \succeq 0 : \mathcal{A}(X) = b\} \quad \subseteq \quad (\mathcal{A}^* y)^{\perp} \cap \mathcal{S}^n_+.$$

Facial reduction (Borwein-Wolkowicz '81):
Replace $\mathcal{S}^n_+$ with $(\mathcal{A}^* y)^{\perp} \cap \mathcal{S}^n_+ \cong \mathcal{S}^r_+$.
Repeat until Slater (P) holds.

Singularity degree (Sturm '98):
d = minimal # of facial reduction iterations required.

Connection to error bounds (Sturm '98):

$$\frac{\mathrm{dist}_{\Omega}(Z)}{\Big(\mathrm{dist}_{\mathcal{S}^n_+}(Z) + \mathrm{dist}_{\mathcal{A}^{-1}b}(Z)\Big)^{2^d}} \qquad \text{is bounded on compact sets}$$

(D-Pataki-Wolkowicz '14): If (P) is degenerate, then

$$d = 1 \quad \Longleftrightarrow \quad \mathrm{face}\Big(b, \mathcal{A}(\mathcal{S}^n_+)\Big) \text{ is exposed.}$$

Open question:
When is $\mathcal{A}(\mathcal{S}^n_+)$ facially exposed? When is $\mathcal{P}_E(\mathcal{S}^n_+)$?

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A} \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A} \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

and its SDP lift

$$\widehat{\mathcal{F}} = \{ X \succeq 0 : \mathcal{A}(X) = 0,\ X_{11} = 1 \} .$$

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A} \begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

and its SDP lift

$$\widehat{\mathcal{F}} = \{ X \succeq 0 : \mathcal{A}(X) = 0,\ X_{11} = 1 \}.$$

(Tunçel '01):

$$\operatorname{aff} \mathcal{F} = \left\{ x : \hat{L} \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \right\} \quad \Longrightarrow \quad \text{can add } \langle \widehat{L}^T \widehat{L}, \cdot \rangle = 0 \text{ to } \mathcal{F}$$

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A}\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

and its SDP lift

$$\widehat{\mathcal{F}} = \{ X \succeq 0 : \mathcal{A}(X) = 0,\ X_{11} = 1 \}.$$

(Tunçel '01):

$$\operatorname{aff} \mathcal{F} = \left\{ x : \hat{L} \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \right\} \quad \Longrightarrow \quad \text{can add } \langle \widehat{L}^T \widehat{L}, \cdot \rangle = 0 \text{ to } \mathcal{F}$$

Then $(\widehat{L}^T \widehat{L})^{\perp} \cap \mathcal{S}^n_+$ regularizes $\widehat{\mathcal{F}}$.

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A}\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

and its SDP lift

$$\widehat{\mathcal{F}} = \{ X \succeq 0 : \mathcal{A}(X) = 0,\ X_{11} = 1 \}.$$

(Tunçel '01):

$$\operatorname{aff} \mathcal{F} = \left\{ x : \hat{L} \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \right\} \quad \Longrightarrow \quad \text{can add } \langle \widehat{L}^T \widehat{L}, \cdot \rangle = 0 \text{ to } \mathcal{F}$$

Then $(\widehat{L}^T \widehat{L})^\perp \cap \mathcal{S}^n_+$ regularizes $\widehat{\mathcal{F}}$.
$\Rightarrow$ $\boxed{d = 1 \text{ and facial reduction is easy}}$.

Example: SDP relaxation

Consider the region

$$\mathcal{F} := \left\{ x \in \mathbf{R}^n : \mathcal{A}\begin{pmatrix} 1 & x^T \\ x & xx^T \end{pmatrix} = 0 \right\}$$

and its SDP lift

$$\widehat{\mathcal{F}} = \{ X \succeq 0 : \mathcal{A}(X) = 0,\ X_{11} = 1 \}.$$

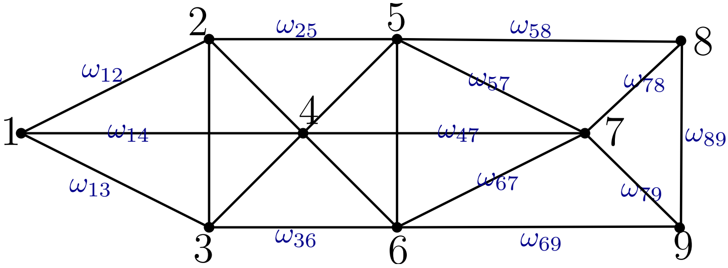
(Tunçel '01):

$$\operatorname{aff} \mathcal{F} = \left\{ x : \hat{L} \begin{bmatrix} 1 \\ x \end{bmatrix} = 0 \right\} \quad \Longrightarrow \quad \text{can add } \langle \widehat{L}^T \widehat{L}, \cdot \rangle = 0 \text{ to } \mathcal{F}$$

Then $(\widehat{L}^T \widehat{L})^\perp \cap \mathcal{S}^n_+$ regularizes $\widehat{\mathcal{F}}$.
$\Rightarrow$ $\boxed{d = 1 \text{ and facial reduction is easy}}$.
Eg. QAP, graph partitioning, second-lift of MAX-CUT.

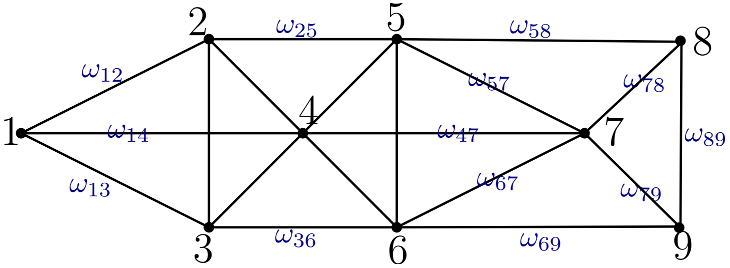
Example: EDM completion

Problem: given a weighed graph $G = (V, E, \omega)$



Example: EDM completion

Problem: given a weighed graph $G = (V, E, \omega)$

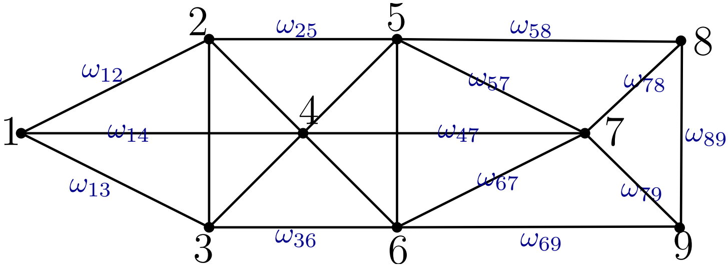


find a **realization**:

$$x_1, \ldots, x_n \in \mathbf{R}^r \quad \text{with} \quad \omega_{ij} = |x_i - x_j|^2.$$

Example: EDM completion

Problem: given a weighed graph $G = (V, E, \omega)$



find a **realization**:

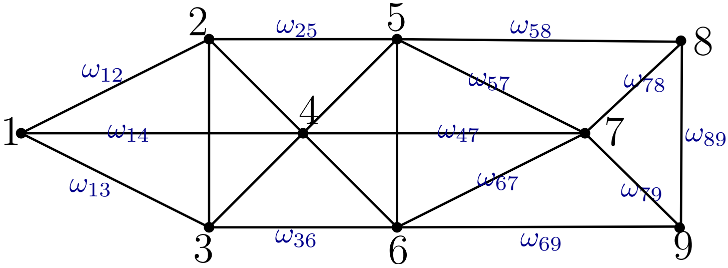
$$x_1, \ldots, x_n \in \mathbf{R}^r \quad \text{with} \quad \omega_{ij} = |x_i - x_j|^2.$$

If possible, then

$$\text{embdim}\, G = \text{ minimal } r.$$

Example: EDM completion

Problem: given a weighed graph $G = (V, E, \omega)$



find a **realization**:

$$x_1, \ldots, x_n \in \mathbf{R}^r \quad \text{with} \quad \omega_{ij} = |x_i - x_j|^2.$$

If possible, then

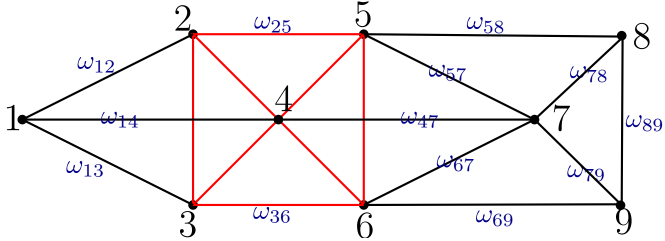
$$\text{embdim}\, G = \text{ minimal } r.$$

Eg: Sensor network localization and molecular conformation

EDM completions

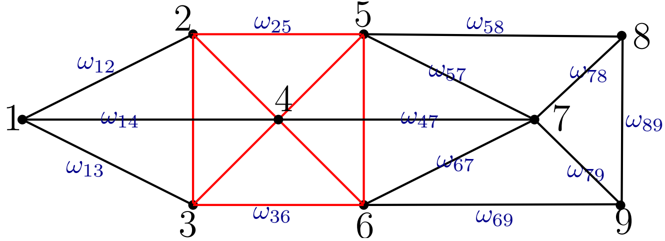
Natural substructures: cliques.

Natural substructures: cliques.



EDM completions

Natural substructures: cliques.

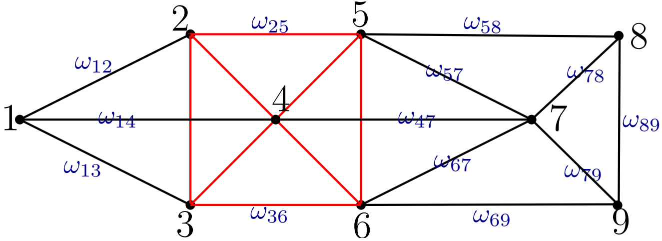


Idea: "Collapse" cliques

- Not clear how to do this

EDM completions

Natural substructures: cliques.



Idea: "Collapse" cliques

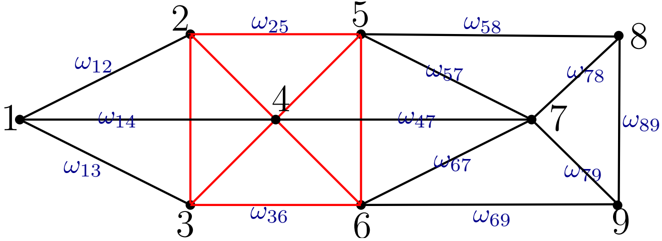
- Not clear how to do this

SDP relaxation

$$\mathcal{F} = \left\{ \begin{array}{ll} X_{ii} + X_{jj} - 2X_{ij} = \omega_{ij} & \text{for all } ij \in E \\ Xe = 0 & \\ X \succeq 0 & \end{array} \right\}$$

EDM completions

Natural substructures: cliques.



Idea: "Collapse" cliques

- Not clear how to do this

SDP relaxation

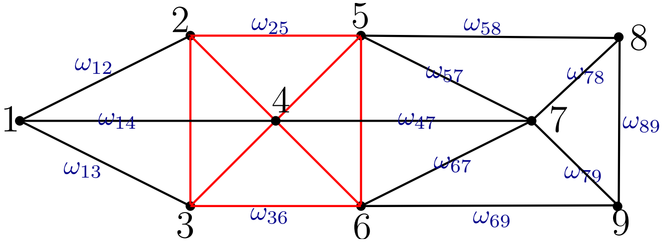
$$\mathcal{F} = \left\{ \begin{array}{ll} X_{ii} + X_{jj} - 2X_{ij} = \omega_{ij} & \text{for all } ij \in E \\ Xe = 0 & \\ X \succeq 0 & \end{array} \right\}$$

Krislock-Wolkowicz '10: For "any" cliques $\chi_1, \dots, \chi_m$ in G

$$\mathcal{F} \quad \subseteq \quad \bigcap_i \left(\mathcal{S}_+^n \cap Y_i^\perp \right)$$

EDM completions

Natural substructures: cliques.



Idea: "Collapse" cliques

- Not clear how to do this

SDP relaxation

$$\mathcal{F} = \left\{ \begin{array}{ll} X_{ii} + X_{jj} - 2X_{ij} = \omega_{ij} & \text{for all } ij \in E \\ Xe = 0 & \\ X \succeq 0 & \end{array} \right\}$$

Krislock-Wolkowicz '10: For "any" cliques $\chi_1, \ldots, \chi_m$ in G

$$\mathcal{F} \quad \subseteq \quad \bigcap_i \left(\mathcal{S}^n_+ \cap Y_i^\perp \right)$$

- Collapse occurs in the SDP!

Rounding for EDM

Real problems have noise in ω!

Rounding for EDM

Real problems have noise in ω!

Key idea:

$$\bigcap_i \left(\mathcal{S}_+^n \cap Y_i^\perp\right) \quad = \quad \mathcal{S}_+^n \cap \Big(Y_1 + \ldots + Y_m\Big)^\perp$$

Rounding for EDM

Real problems have noise in ω!

Key idea:

$$\bigcap_i \left(\mathcal{S}_+^n \cap Y_i^\perp\right) \quad = \quad \mathcal{S}_+^n \cap \Big(Y_1 + \ldots + Y_m\Big)^\perp$$

Algorithmic framework (Cheung-D-Krislock-Wolkowicz '14):

1. Fix a set of cliques χ^i
2. Form "approximate exposing matrices" Y_i from χ_i
3. Form the aggregate

$$Y = Y_1 + \ldots + Y_m.$$

4. Round down Y to a nearest rank $n - r$ matrix $\mathcal{N}$
5. Solve Least Squares on $\mathcal{S}_+^n \cap \mathcal{N}^\perp \cong \mathcal{S}_+^r$

Real problems have noise in ω!

Key idea:

$$\bigcap_i \left(\mathcal{S}_+^n \cap Y_i^\perp\right) \quad = \quad \mathcal{S}_+^n \cap \Big(Y_1 + \ldots + Y_m\Big)^\perp$$

Algorithmic framework (Cheung-D-Krislock-Wolkowicz '14):

1. Fix a set of cliques χ^i
2. Form "approximate exposing matrices" Y_i from χ_i
3. Form the aggregate

$$Y = Y_1 + \ldots + Y_m.$$

4. Round down Y to a nearest rank $n - r$ matrix $\mathcal{N}$
5. Solve Least Squares on $\mathcal{S}_+^n \cap \mathcal{N}^\perp \cong \mathcal{S}_+^r$

Reasonable conditions $\Longrightarrow$

$$\text{output error} \leq \kappa(\text{input noise}).$$

Rounding for EDM

Real problems have noise in ω!

Key idea:

$$\bigcap_i \left(\mathcal{S}^n_+ \cap Y_i^\perp\right) \quad = \quad \mathcal{S}^n_+ \cap \Big(Y_1 + \ldots + Y_m\Big)^\perp$$

Algorithmic framework (Cheung-D-Krislock-Wolkowicz '14):

1. Fix a set of cliques χ^i
2. Form "approximate exposing matrices" Y_i from χ_i
3. Form the aggregate
$$Y = Y_1 + \ldots + Y_m.$$
4. Round down Y to a nearest rank $n-r$ matrix $\mathcal{N}$
5. Solve Least Squares on $\mathcal{S}^n_+ \cap \mathcal{N}^\perp \cong \mathcal{S}^r_+$

Reasonable conditions $\Longrightarrow$

$$\text{output error} \le \kappa(\text{input noise}).$$

Advertisement: see Krislock TD21 for more.

Rounding for EDM

5% noise, 6% density $(n = 1000, r = 2)$:

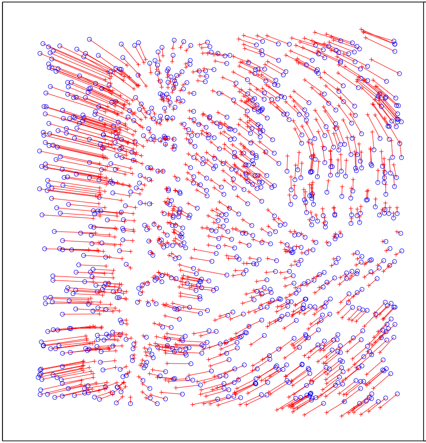


Figure: Before refinement

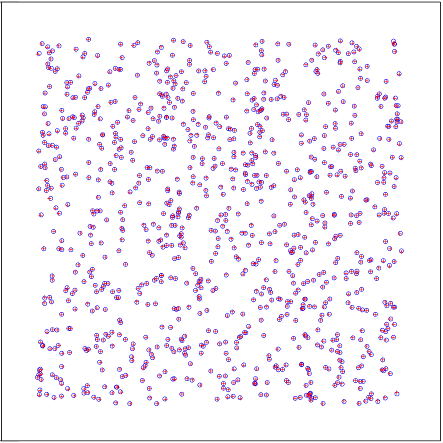


Figure: After refinement

Different idea for noisy EDM

Unfolding heuristic (Weinberger et al. '79):

$$\begin{aligned} \max \quad & \operatorname{tr}(X) \\ \text{s.t.} \quad & \sqrt{\sum_{ij\in E} |X_{ii} - 2X_{ij} + X_{jj} - \omega_{ij}|^2} \leq \sigma \\ & Xe = 0 \\ & X \succeq 0 \end{aligned}$$

Different idea for noisy EDM

Unfolding heuristic (Weinberger et al. '79):

$$\begin{aligned} \max \quad & \operatorname{tr}(X) \\ \text{s.t.} \quad & \sqrt{\sum_{ij \in E} |X_{ii} - 2X_{ij} + X_{jj} - \omega_{ij}|^2} \leq \sigma \\ & Xe = 0 \\ & X \succeq 0 \end{aligned}$$

(Biswas-Liang-Toh-Ye-Wang '06)

Different idea for noisy EDM

Unfolding heuristic (Weinberger et al. '79):

$$\begin{array}{ll} \max & \operatorname{tr}(X) \\ \text{s.t.} & \sqrt{\sum_{ij\in E} |X_{ii} - 2X_{ij} + X_{jj} - \omega_{ij}|^2} \leq \sigma \\ & Xe = 0 \\ & X \succeq 0 \end{array}$$

(Biswas-Liang-Toh-Ye-Wang '06)

Intuition: $$\operatorname{tr}(X) = \frac{1}{2n} \sum_{i,j=1}^{n} \|p_i - p_j\|^2$$

Different idea for noisy EDM

Unfolding heuristic (Weinberger et al. '79):

$$\begin{aligned} \max \quad & \operatorname{tr}(X) \\ \text{s.t.} \quad & \sqrt{\sum_{ij \in E} |X_{ii} - 2X_{ij} + X_{jj} - \omega_{ij}|^2} \leq \sigma \\ & Xe = 0 \\ & X \succeq 0 \end{aligned}$$

(Biswas-Liang-Toh-Ye-Wang '06)

Intuition: $$\operatorname{tr}(X) = \frac{1}{2n} \sum_{i,j=1}^{n} \|p_i - p_j\|^2$$

Flipped problem:

$$\begin{aligned} \psi(\tau) := \min \quad & \sqrt{\sum_{ij \in E} |X_{ii} - 2X_{ij} + X_{jj} - \omega_{ij}|^2} \\ \text{s.t.} \quad & \operatorname{tr} X = \tau \\ & Xe = 0 \\ & X \succeq 0. \end{aligned}$$

Approximate Newton

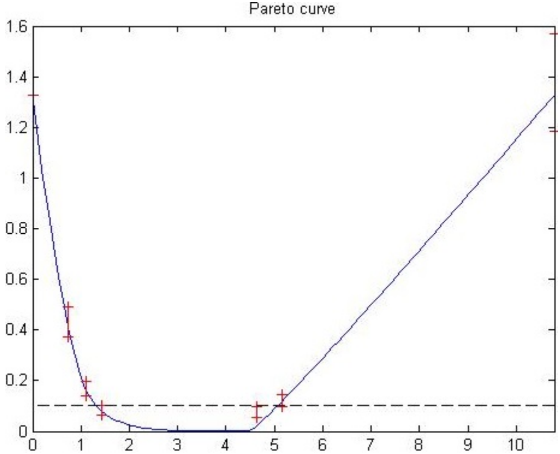


Figure: Graph of ψ

Strategy: approximate Newton method for finding

$$\boxed{\text{maximal } \tau \text{ with } \psi(\tau) \leq \sigma}.$$

Approximate evaluation of ψ with Frank-Wolfe algorithm.

Convergence guarantee: can obtain $X \succeq 0$ with

$$\operatorname{tr} X \geq \text{max-trace} \quad \text{and} \quad \text{residual} \ \leq \sigma + \epsilon$$

using

$$\mathcal{O}\left(\frac{\bar{\tau} \cdot \mathrm{Lip}^2}{\epsilon^2} \ln\left(\frac{(\tau_0 - \bar{\tau}) \cdot \psi_0'}{\epsilon}\right)\right) \quad \text{FW iterations.}$$

Approximate Newton

Convergence guarantee: can obtain $X \succeq 0$ with

$$\operatorname{tr} X \geq \text{max-trace} \quad \text{and} \quad \text{residual} \leq \sigma + \epsilon$$

using

$$\mathcal{O}\left(\frac{\bar{\tau} \cdot \mathrm{Lip}^2}{\epsilon^2} \ln\left(\frac{(\tau_0 - \bar{\tau}) \cdot \psi_0'}{\epsilon}\right)\right) \quad \text{FW iterations.}$$

Related “flippy strategies”: (van den Berg-Friedlander ’08, Harchaoui-Juditsky-Nemirovski ’13)

Max-trace vs Min-trace

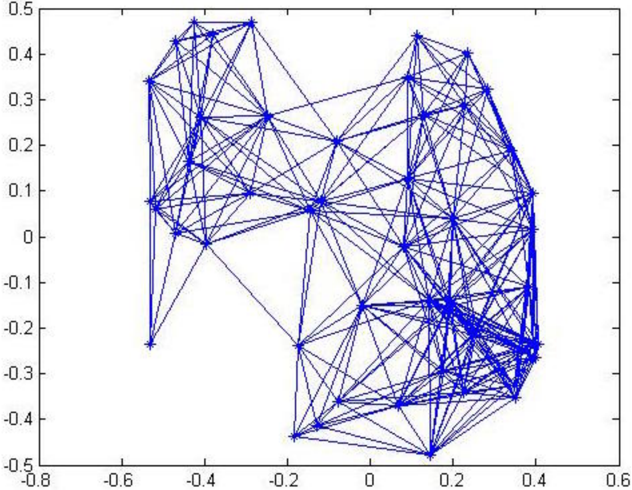
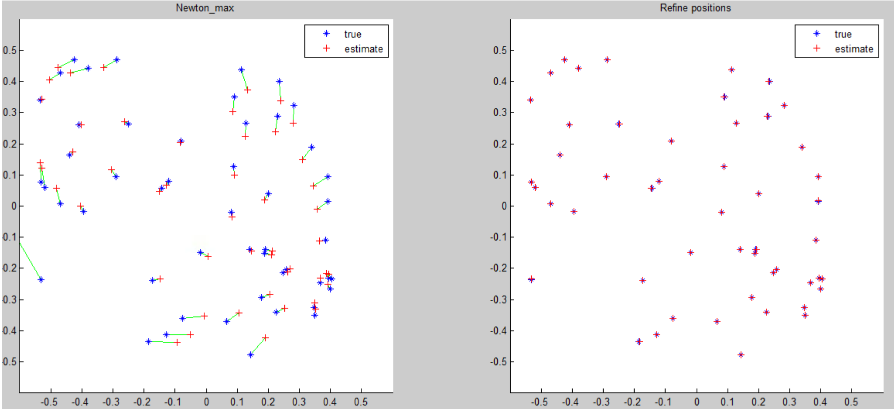


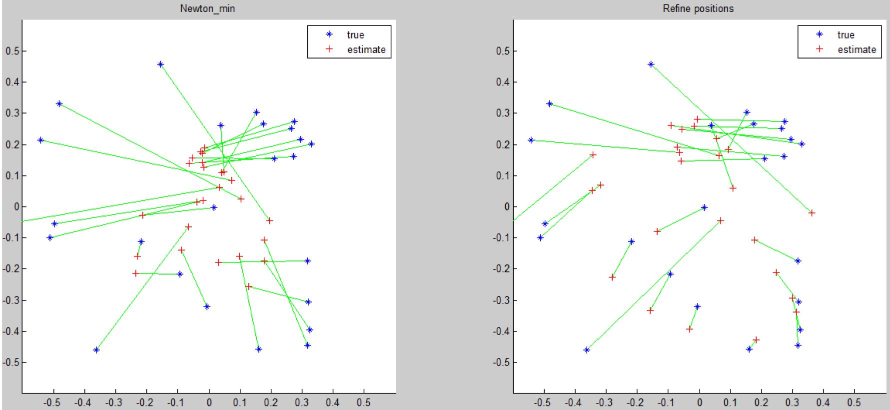
Figure: Sensor network

Max-trace



Newton_max
true
estimate
Refine positions
true
estimate
0.5
0.4
0.3
0.2
0.1
0
-0.1
-0.2
-0.3
-0.4
-0.5

Min-trace



Conclusion & Open questions

- Slater condition: fundamentally important, and can fail in applications.

Conclusion & Open questions

- Slater condition: fundamentally important, and can fail in applications.
- Illustration: noisy, low-rank EDM completions.

Conclusion & Open questions

- Slater condition: fundamentally important, and can fail in applications.
- Illustration: noisy, low-rank EDM completions.
  - randomized rounding
  - Newton with Frank-Wolfe.

Conclusion & Open questions

- Slater condition: fundamentally important, and can fail in applications.
- Illustration: noisy, low-rank EDM completions.
    - randomized rounding
    - Newton with Frank-Wolfe.
- Advertisement:
  Survey paper (with H. Wolkowicz) is forthcoming.

Thank you.