Theory and Applications of Degeneracy in Cone Optimization

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

IPAM, Oct 15, 2010

Outline

Part I: Degeneracy in Cone Optimization

minimal representations and strong duality

(strict) complementarity and duality gaps

} Numerical difficulties

(With: Y-L Cheung, L. Tuncel, S. Schurr, H. Wei)

Part II: Sensor Network Localization, SNL

- exploiting implicit degeneracy
- solving huge problems
- high accuracy (low rank) solutions

(With: N. Krislock, F. Rendl)

Cone Optimization, (e.g. $K = \mathcal{S}^n_+$, SDP, $K = \mathbb{R}^n_+$, LP)

Primal-Dual Pair of Optimization Problems in Conic Form

$$\text{(assumed finite)} \quad v_P = \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_K c\}, \qquad (\mathbb{P})$$

$$(v_P \leq) \quad v_D = \inf_x \{\langle c, x\rangle \ : \ \mathcal{A}x = b, \ x \succeq_{K^*} 0\}. \qquad (\mathbb{D})$$

where

- $\mathcal{A}$ - an onto linear transformation; adjoint is $\mathcal{A}^*$
- K - a proper convex cone with dual/polar cone $K^* = \{x : \langle s, x\rangle \geq 0, \ \forall s \in K\}$.
- $s' \preceq_K s''(s' \prec_K s'')$ - partial order, $s'' - s' \in K (\in \mathrm{int} K)$

$K = \mathcal{S}^n_+$, Semidefinite Programming (SDP, LMI)

Primal-Dual Pair

$$v_P = \sup_y \{b^T y \ : \ c - \sum_{i=1}^m y_i A_i \succeq 0\}, \qquad (\mathbb{P})$$

$$v_D = \inf_x \{\operatorname{trace} cx \ : \ (\operatorname{trace} A_i x) = b \in \mathbb{R}^m, \ x \succeq 0\}. \qquad (\mathbb{D})$$

$c, A_i \in \mathcal{S}^n, \forall i$

Optimality Conditions

Strong Duality if a Constraint Qualification, CQ, holds

$$v_P = v_D = \langle c, x \rangle, \qquad x \text{ dual optimal}$$

Zero duality gap and dual attainment.

Strict Complementarity

x, z optimal pair;
$\langle x, z \rangle = 0$ complementarity
$x + z \succ 0$ strict complementarity

Part I: Motivation/Outline

In case of nonpolyhedral cones, Strong Duality and/or Strict Complementarity can Fail

- Many Instances: SDP relax. for hard comb. probs. (e.g. QAP, GP, strengthened MC, POP, SNL)
- Fresh look at known Characterizations of Optimality without a CQ using Subspace Formulation
- theme: use MINIMAL REPRESENTATIONS for regularization, efficient solutions
- Connections Complementarity of Homog. Probl. and duality/Numerical implications

Faces of Cones - Useful for Charact. of Opt.

Face

A convex cone F is a face of K, denoted $F \trianglelefteq K$, if

$$x, y \in K \text{ and } x + y \in F \implies x, y \in F.$$

If $F \trianglelefteq K$ and $F \neq K$, write $F \triangleleft K$.

Conjugate Face

If $F \trianglelefteq K$, the conjugate face (or complementary face) of F is

$$F^c := F^\perp \cap K^* \trianglelefteq K^*.$$

If $x \in \mathrm{ri}(F)$, then $F^c = \{x\}^\perp \cap K^*$.

Minimal Face (Minimal Cone)

Feasible sets

$$\begin{array}{lcll} \mathcal{F}_P^y & := & \{y : c - \mathcal{A}^* y \succeq_K 0\} & \text{primal} \\ \mathcal{F}_P^s & := & \{s : s = c - \mathcal{A}^* y \succeq_K 0, \text{ for some } y\} & \text{primal slacks} \\ \mathcal{F}_D^x & := & \{x : \mathcal{A} x = b, x \succeq_K^* 0\} & \text{dual} \end{array}$$

Minimal Faces (Intersection of Faces is a Face)

$$f_P := \operatorname{face} \mathcal{F}_P^s \unlhd K \qquad f_D := \operatorname{face} \mathcal{F}_D^x \unlhd K^*$$

(Modified) SDP Example from Ramana, 1995

Primal SDP, $\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$

$$0 = v_P = \sup_y \left\{ y_2 : \begin{pmatrix} 0 & 0 & y_2 \\ 0 & y_2 & 0 \\ y_2 & 0 & y_1 \end{pmatrix} \preceq \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \right\}$$

2×2 principal submatrix $\preceq 0 \implies y_2 = 0$

$$y^* = \begin{pmatrix} y_1^* & 0 \end{pmatrix}^T, \quad y_1^* \leq 0, \quad s^* = c - \mathcal{A}^* y^* = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -y_1^* \end{pmatrix}$$

$$A_1 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \qquad b = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

Slater's CQ fails for primal and dual

in fact, positive duality gap: $v_D = 1 \quad > \quad v_P = 0$

Dual of SDP Example

Dual Program

$$1 = v_D = \inf_x \{x_{22} \ : \ x_{33} = 0, \ x_{22} + 2x_{13} = 1, \ x \succeq 0\}$$

$$x_{33} = 0, x \succeq 0 \implies x_{13} = 0 \implies x_{22} = 1$$

$$x^* = \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{21} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad x_{11} \geq (x_{12}^2)$$

Slater's CQ for (primal) dual & complementarity fails

positive duality gap: $v_D - v_P = 1 - 0 = 1$,

$$\text{trace}\, x^* s^* = \text{trace} \begin{pmatrix} x_{11} & x_{12} & 0 \\ x_{21} & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -y_1^* \end{pmatrix} = 1 > 0$$

Minimal Face for Ramana Example

Feasible Set/Minimal Face

$\mathcal{F}_P^y = \{y \in \mathbb{R}^2 : y_1 \leq 0,\ y_2 = 0\}$

$$\begin{aligned} f_P &= \bigcap\{F \trianglelefteq K : \mathcal{F}_P^s = c - \mathcal{A}^*(\mathcal{F}_P^y) \subset F\} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & \mathbb{S}^2_+ \end{pmatrix} \lhd \mathbb{S}^3_+ \end{aligned}$$

Rotate/project to get Smaller Problem with Slater's CQ

$y \in \mathcal{F}_P^y$ iff $\begin{bmatrix} 0 & I \end{bmatrix} (c - y_1 A_1) \begin{bmatrix} 0 & I \end{bmatrix}^T \in \mathcal{S}^2_+$, A_2 disappears

Slater CQ and Minimal Face

If $(\mathbb{P})$ is feasible, then

$$c - \mathcal{A}^* y \not\succ_K 0\ \forall y \text{ (Slater's CQ fails for } (\mathbb{P}) \text{) } \iff f_P \lhd K$$

Regularization of $(\mathbb{P})$ Using Minimal Face

Borwein-W (1981), $f_P = \text{face}\,\mathcal{F}_P^s$

$(\mathbb{P})$ is equivalent to regularized $(\mathbb{P})$

$$v_{RP} := \sup_y \{\langle b, y\rangle \ : \ \mathcal{A}^* y \preceq_{f_P} c\}. \qquad (\mathbb{RP})$$

Lagrangian Dual DRP Satisfies Strong Duality:

$$v_P = v_{RP} = v_{DRP} := \inf_x \{\langle c, x\rangle \ : \ \mathcal{A} x = b, \ x \succeq_{f_P^*} 0\} \qquad (\mathbb{DRP})$$

and v_{DRP} is attained

smaller cone in primal $f_P \subset K$; larger cone in dual $K^* \subset f_P^*$

(SYMMETRIC) Subspace Form for ($\mathbb{P}$) and ($\mathbb{D}$)

Assume Linear Feasibility for $\tilde{s}, \tilde{y}$, $\tilde{x}$; with data A, b, c, K

$$\mathcal{A}^*\tilde{y} + \tilde{s} = c \qquad \mathcal{A}\tilde{x} = b$$

$$\mathcal{L}^\perp = \mathcal{R}(\mathcal{A}^*) \text{ (range)} \qquad \mathcal{L} = \mathcal{N}(\mathcal{A}) \text{ (nullspace)}$$

Equivalent Primal-Dual Pair in Subspace Form, (e.g. N&N '94)

Particular solution + solution of homogeneous equation

$$v_P = c\tilde{x} - \inf_s \left\{ s\tilde{x} : s \in (\tilde{s} + \mathcal{L}^\perp) \cap K \right\}. \qquad (\mathbb{P})$$

$$v_D = \tilde{y}b + \inf_x \{\tilde{s}x : x \in (\tilde{x} + \mathcal{L}) \cap K^*\}. \qquad (\mathbb{D})$$

For $(\mathbb{P})$ and $(\mathbb{D})$

Faces of Recession Directions (feasible case)

$f_P^0 := \mathrm{face}\,\left(\mathcal{L}^\perp \cap K\right) (\subset f_P)$, $\quad f_D^0 := \mathrm{face}\,\left(\mathcal{L} \cap K^*\right) (\subset f_D)$

Recall

minimal faces: $f_P = \mathrm{face}\,\mathcal{F}_P^s$, $\quad f_D = \mathrm{face}\,\mathcal{F}_D^x$

Minimal Subspaces/Linear Transformations

min. subsp.: $\mathcal{L}_{PM}^\perp := \mathcal{L}^\perp \cap (f_P - f_P)$, $\quad \mathcal{L}_{DM} := \mathcal{L} \cap (f_D - f_D)$

min. Lin. Tr.: $\mathcal{A}_{PM}^*$, $\quad \mathcal{A}_{DM}$

Regularization of $(\mathbb{P})$ Using Minimal Subspace

Assume K Facially Dual Complete, FDC (Pataki/07, 'nice')

i.e. $F \lhd K \implies K^* + F^\perp$ is closed. (e.g. $\mathcal{S}^n_+, \mathbb{R}^n_+$,SOC).

$\mathcal{L}^\perp_{PM} = \mathcal{L}^\perp \cap (f_P - f_P)$

$$v_{RP} = c\tilde{x} - \inf_s \left\{ s\tilde{x} : s \in (\tilde{s} + \mathcal{L}^\perp_{MP}) \cap K \right\} \qquad (\mathbb{RP})$$

Lagrangian Dual DRP Satisfies Strong Duality:

$$v_P = v_{RP} = v_{DRP} = \tilde{y}b + \inf_x \{\tilde{s}x : x \in (\tilde{x} + \mathcal{L}_{MP}) \cap K^*\} \quad (\mathbb{DRP})$$

and v_{DRP} is attained

Nice and Devious Cones

Lemma for SDP Case (Ramana,Tuncel,W./97)

Let $0 \neq F \lhd \mathcal{S}_+^n$. Then
$\mathcal{S}_+^n + F^\perp$ is closed (nice)
$\mathcal{S}_+^n + \mathrm{span} F^c$ is not closed (devious)
$\mathcal{S}_+^n + F^\perp = \overline{\mathcal{S}_+^n + \mathrm{span} F^c}$

Infinite Duality Gap for Devious cones

Let $\mathcal{L} = \mathrm{span} F^c$; choose $c = \tilde{s} = 0$ and
$\tilde{x} \in (\mathcal{S}_+^n + F^\perp) \backslash (\mathcal{S}_+^n + \mathrm{span} F^c)$; (subspace repr. (P),(D): (13)).
then $0 = v_P < v_D = \infty$.

Strong Duality for (P) ($v_P = v_D$ and v_D is attained)

Minimal Face and Minimal Subspace CQs for (P)

1. $f_P = K$ is a CQ
   (from BW: $f_P^* = K^*$)
2. $\mathcal{L}^\perp \cap (f_P - f_P) = \mathcal{L}_{PM}^\perp = \mathcal{L}^\perp$ is a CQ (if K is FDC (nice))
   ($\tilde{s} \in f_P - f_P : x^* = x_K^* + x_f^* \in f_P^* = K^* + f_P^\perp \implies x^*(\tilde{s} + \mathcal{L}^\perp) = x_K^*(\tilde{s} + \mathcal{L}^\perp)$)

Universal CQ, UCQ for (P) (i.e. independent of feasible data c, b)

$\mathcal{L}^\perp \subset f_P^0 - f_P^0$ is a UCQ (if K is FDC)
(wlog choose $\tilde{s} \in K, \tilde{x} \in K^*$; shows that $f_P^0 \subset f_P, f_D^0 \subset f_D$)

Backward Stable Reglarization (in progress)

Goals: Detect (near) Loss of Slater CQ/Regularize

- Solve a backward stable auxiliary problem
- Alternate projection onto smaller face/subspace to finally obtain a regularized problem

i.e. $f_P = K$ and $\mathcal{L}^\perp \cap (f_P - f_P) = \mathcal{L}^\perp_{PM} = \mathcal{L}^\perp$

(Near) Loss of Slater Condition/Strict Feasibility

Theoretical/Numerical Difficulties

- Primal Slater condition implies strong duality, i.e. zero duality gap AND dual attainment.
- (Near) loss of strict feasibility is used as a measure in complexity theory. (e.g. Renegar/95, Freund/01, Lara and Tuncel/02)
- (Near) loss of strict feasibility correlates with number of iterations and loss of accuracy in interior-point methods (e.g. Freund/Ordonez/Toh 2006)

Loss of Strict Complementarity, (SC)

Strict Complementary Optimal Primal-Dual Pair

- There exists an optimal primal-dual pair x, s such that
  $x + s \succ 0 \qquad (\in \operatorname{int}(K + K^*))$

Theoretical Difficulties/Convergence

- Convergence proofs for asymptotic quadratic superlinear convergence require SC.
- Proofs of convergence to the analytic center require SC

Numerical Difficulties/Relation to Duality Gaps???

increased number of iterations? loss of accuracy?

Generating Hard SDP Instances (Wei and W. 2006)

Maximal Complementary Solution Pair:

- A p-d pair of optimal solutions $(\bar{s}, \bar{x})$ is a *maximal complementary solution pair* if the pair maximizes the sum $\mathrm{rank}\,(s) + \mathrm{rank}\,(x)$ over all p-d optimal (s, x).

Strict Complementarity Nullity, g:

- $g = n - \mathrm{rank}\,(\bar{s}) - \mathrm{rank}\,(\bar{x})$, where $(\bar{s}, \bar{x})$ is a maximal complementary solution pair

Hard SDP Instances:

- problems where nullity is nonzero

Empirical Observations

Numerical Difficulties Correlate with Large Nullity

- There is a strong correlation between the iteration number to achieve the desired stopping tolerance and the size of the complementarity nullity, when the accuracy requirement is high.
- Large nullity instances cause problems for SDP solvers.
- Local asymptotic convergence rate is slower when nullity is larger.

Theoretical Connections Complementarity/Duality?

Numerical Difficulties

(Both) loss of Slater CQ (strict feasibility) and loss of strict complementarity independently result in theoretical difficulties and numerical difficulties for interior-point methods.

Theoretical Connection?

Is there a theoretical connection between loss of duality (from loss of a CQ) and loss of strict complementarity?

Complementarity Partition

Recall Faces of Recession Directions

$$f_P^0 := \text{face}\left(\mathcal{L}^\perp \cap K\right), \qquad f_D^0 := \text{face}\left(\mathcal{L} \cap K^*\right)$$

The pair f_P^0, f_D^0 define a Complementarity Partition

- $\text{face}(f_P^0) \subset \text{face}(f_D^0)^c$ and $\text{face}(f_D^0) \subset \text{face}(f_P^0)^c$.
- it is a strict complementarity partition if both $[\text{face}(f_P^0)]^c = \text{face}(f_D^0)$ and $[\text{face}(f_D^0)]^c = \text{face}(f_P^0)$;
- it is proper if f_P^0 and f_D^0 are both nonempty.

SDP Picture

For SDP (after a rotation)

$$\begin{bmatrix} f_D^0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & f_P^0 \end{bmatrix}$$

Form Primal-Dual Pair

$$\tilde{x} = \tilde{s} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & v \succ 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \implies \langle s, x \rangle \geq \|v\|_F^2,$$

for all feasible pairs s, x. (gap is dimension of v)

Theorem: K is a proper cone

(1) If f_P^0, f_D^0 define a proper complementarity partition with a gap of dimension 1, so, the partition is not a strict complementarity partition, then there exists $\bar{s}$ and $\bar{x}$ such that $(\mathbb{P})$–$(\mathbb{D})$ with data $(\mathcal{L}, K, \bar{s}, \bar{x})$ has a finite nonzero duality gap.

(Partial Converse)

(2) If
(a) $(\mathbb{P})$–$(\mathbb{D})$ with data $(\mathcal{L}, K, \bar{s}, \bar{x})$ has a finite nonzero duality gap with both optimal values attained, and
(b) the objective functions are constant along all recession directions of $(\mathbb{P})$ and $(\mathbb{D})$,
then f_P^0, f_D^0 has a proper complementarity partition but not a strict complementarity partition.

Conclusion Part I

- Minimal Representations of the data regularize (P)
  min. face f_P and/or the min. L.T. $\mathcal{A}_{PM}$ or $\mathcal{L}^*_{PM}$
- goal: a stable algorithm to solve (feasible) conic problems for which Slater's CQ fails
- Failure of strict complementarity for the associated recession problems is closely related to the existence of instances having a finite nonzero duality gap; provides a means of generating instances for testing.

$\uparrow \quad \rightarrow$ Numer. Difficulties $\leftarrow \quad \uparrow$

Slater CQ   str. compl.

$\updownarrow$

pos. duality gap

$\uparrow \quad \leftarrow$ str. compl. partit. homog. prob. $\rightarrow \quad \uparrow$

Part II: Sensor Network Localization, SNL, Problem (Exploiting (Implicit) Degeneracy)

SNL - a Fundamental Problem of Distance Geometry; easy to describe - dates back to Grasssmann 1886

- n ad hoc wireless sensors (nodes) to locate in $\mathbb{R}^r$, (r is embedding dimension; sensors $p_i \in \mathbb{R}^r, i \in V := 1, \ldots, n$)
- m of the sensors are anchors, $p_i, i = n - m + 1, \ldots, n$) (positions known, using e.g. GPS)
- pairwise distances $D_{ij} = \|p_i - p_j\|^2, ij \in E$, are known within radio range $R > 0$
- $$P^T = \begin{bmatrix} p_1 & \ldots & p_n \end{bmatrix} = \begin{bmatrix} X^T & A^T \end{bmatrix} \in \mathbb{R}^{r \times n}$$

Applications

Horst Stormer (Nobel Prize, Physics, 1998), "21 Ideas for the 21st Century", Business Week. 8/23-30, 1999

Untethered micro sensors will go anywhere and measure anything - traffic flow, water level, number of people walking by, temperature. This is developing into something like a nervous system for the earth, a skin for the earth. The world will evolve this way.

Tracking Humans/Animals/Equipment/Weather (smart dust)

- geographic routing; data aggregation; topological control; soil humidity; earthquakes and volcanos; weather and ocean currents.
- military; tracking of goods; vehicle positions; surveillance; random deployment in inaccessible terrains.
- brain scans

Underlying Graph Realization/Partial EDM NP-Hard

Graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$

- node set $\mathcal{V} = \{1, \ldots, n\}$
- edge set $(i,j) \in \mathcal{E}$; $\omega_{ij} = \|p_i - p_j\|^2$ known approximately
- The anchors form a clique (complete subgraph)
- Realization of $\mathcal{G}$ in $\Re^r$: a mapping of node $v_i \to p_i \in \Re^r$ with squared distances given by ω.

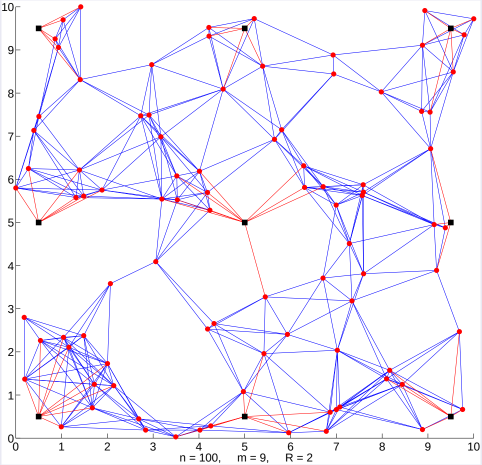
Corresponding Partial Euclidean Distance Matrix, EDM

$$D_{ij} = \begin{cases} d_{ij}^2 & \text{if } (i,j) \in \mathcal{E} \\ 0 & \text{otherwise (unknown distance),} \end{cases}$$

$d_{ij}^2 = \omega_{ij}$ are known squared Euclidean distances between sensors p_i, p_j; anchors correspond to a clique.

Sensor Localization Problem/Partial EDM

Sensors ○ and Anchors ■



Connections to Semidefinite Programming (SDP)

$\mathcal{S}^n_+$, Cone of (symmetric) SDP matrices in $\mathcal{S}^n$; $x^TAx \geq 0$

inner product $\langle A, B\rangle = \operatorname{trace} AB$
Löwner (psd) partial order $A \succeq B, A \succ B$

$D = \mathcal{K}(B) \in \mathcal{E}^n$, $B = \mathcal{K}^\dagger(D) \in \mathcal{S}^n \cap \mathcal{S}_C$ (centered $Be = 0$)

$P^T = \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} \in \mathcal{M}^{r\times n}$; $B := PP^T \in \mathcal{S}^n_+$;
$\operatorname{rank} B = r$; $D \in \mathcal{E}^n$ be corresponding EDM.

$$\begin{array}{rcl} \text{(to } D \in \mathcal{E}^n) \quad D & = & \left(\|p_i - p_j\|_2^2\right)_{i,j=1}^n \\ & = & \left(p_i^Tp_i + p_j^Tp_j - 2p_i^Tp_j\right)_{i,j=1}^n \\ & = & \boxed{\operatorname{diag}(B)e^T + e\operatorname{diag}(B)^T - 2B} \\ & =: & \mathcal{D}_e(B) - 2B \\ & =: & \mathcal{K}(B) \quad (\text{from } B \in \mathcal{S}^n_+). \end{array}$$

Current Techniques; SDP Relax.; Highly Degen.

Nearest, Weighted, SDP Approx. (relax $\operatorname{rank} B$)

- $\min_{B \succeq 0, B \in \Omega} \|H \circ (\mathcal{K}(B) - D)\|$; $\operatorname{rank} B = r$;
  typical weights: $H_{ij} = 1/\sqrt{D_{ij}}$, if $ij \in E$.
- with rank constraint: a non-convex, NP-hard program
- SDP relaxation is convex, BUT: expensive/low accuracy/implicitly highly degenerate (cliques restrict ranks of feasible Bs)

Instead: (Shall) Take Advantage of Degeneracy!

clique $\alpha, |\alpha| = k$ (corresp. $D[\alpha]$) with embed. dim. $= t \leq r < k$
$\implies \operatorname{rank} \mathcal{K}^\dagger(D[\alpha]) = t \leq r \implies \operatorname{rank} B[\alpha] \leq \operatorname{rank} \mathcal{K}^\dagger(D[\alpha]) + 1$
$\implies \operatorname{rank} B = \operatorname{rank} \mathcal{K}^\dagger(D) \leq n - \boxed{(k - t - 1)} \implies$
Slater's CQ (strict feasibility) fails

Linear Transformations: $\mathcal{D}_v(B), \mathcal{K}(B), \mathcal{T}(D)$

- allow: $\mathcal{D}_v(B) := \text{diag}(B)\, v^T + v\, \text{diag}(B)^T$;
  $\mathcal{D}_v(y) := yv^T + vy^T$
- adjoint $\mathcal{K}^*(D) = 2(\mathrm{Diag}\,(De) - D)$.
- $\mathcal{K}$ is $1-1$, onto between centered & hollow subspaces:
  $\mathcal{S}_C := \{B \in \mathcal{S}^n : Be = 0\}$;
  $\mathcal{S}_H := \{D \in \mathcal{S}^n : \mathrm{diag}\,(D) = 0\} = \mathcal{R}\,(\mathrm{offDiag}\,)$
- $J := I - \frac{1}{n}ee^T$ (orthogonal projection onto $M := \{e\}^\perp$);
- $\mathcal{T}(D) := -\frac{1}{2} J \mathrm{offDiag}\,(D) J \qquad (= \mathcal{K}^\dagger(D))$

Semidefinite Cone, Faces

Faces of cone K

- $F \subseteq K$ is a face of K, denoted $\boxed{F \trianglelefteq K}$, if
  $\left(x, y \in K,\ \frac{1}{2}(x+y) \in F\right) \implies \left(\operatorname{cone}\{x, y\} \subseteq F\right)$.
- $\boxed{F \triangleleft K}$, if $F \trianglelefteq K$, $F \neq K$; F is $\boxed{\text{proper face}}$ if $\{0\} \neq F \triangleleft K$.
- $F \trianglelefteq K$ is $\boxed{\text{exposed}}$ if: intersection of K with a hyperplane.
- $\boxed{\operatorname{face}(S)}$ denotes smallest face of K that contains set S.

$\mathcal{S}^n_+$ is a Facially Exposed Cone

All faces are exposed.

Facial Structure of SDP Cone; Equivalent SUBSPACES

Face $F \unlhd \mathcal{S}^n_+$ Equivalence to $\mathcal{R}(U)$ Subspace of $\mathbb{R}^n$

$F \unlhd \mathcal{S}^n_+$ determined by range of any $S \in \operatorname{relint} F$,
i.e. let $S = U\Gamma U^T$ be compact spectral decomposition; $\Gamma \in \mathcal{S}^t_{++}$ is diagonal matrix of pos. eigenvalues; $\boxed{F = U\mathcal{S}^t_+ U^T}$
(F associated with $\mathcal{R}(U)$)

$$\dim F = t(t+1)/2.$$

face F representation by subspace $\mathcal{L}$

(subspace) $\mathcal{L} = \mathcal{R}(T)$, T is $n \times t$ full column, then:

$$F := T\mathcal{S}^t_+ T^T \unlhd \mathcal{S}^n_+$$

Further Notation

Matrix with Fixed Principal Submatrix

For $Y \in \mathcal{S}^n$, $\alpha \subseteq \{1, \ldots, n\}$: $Y[\alpha]$ denotes principal submatrix formed from rows & cols with indices α.

Sets with Fixed Principal Submatrices

If $|\alpha| = k$ and $\bar{Y} \in \mathcal{S}^k$, then:

- $\mathcal{S}^n(\alpha, \bar{Y}) := \left\{ Y \in \mathcal{S}^n : Y[\alpha] = \bar{Y} \right\}$,
- $\mathcal{S}^n_+(\alpha, \bar{Y}) := \left\{ Y \in \mathcal{S}^n_+ : Y[\alpha] = \bar{Y} \right\}$
  i.e. the subset of matrices $Y \in \mathcal{S}^n$ ($Y \in \mathcal{S}^n_+$) with principal submatrix $Y[\alpha]$ fixed to $\bar{Y}$.

Basic Single Clique/Facial Reduction

$\bar{D} \in \mathcal{E}^k$, $\alpha \subseteq 1\!:\!n$, $|\alpha| = k$

Define $\mathcal{E}^n(\alpha, \bar{D}) := \left\{D \in \mathcal{E}^n : D[\alpha] = \bar{D}\right\}$.

Given $\bar{D}$; find a corresponding $B \succeq 0$; find the corresponding face; find the corresponding subspace.

if $\alpha = 1:k$; embed. dim of $\bar{D}$ is $t \leq r$

$$D = \begin{bmatrix} \bar{D} & \cdot \\ \cdot & \cdot \end{bmatrix},$$

BASIC THEOREM for Single Clique/Facial Reduction

THEOREM 1: Single Clique/Facial Reduction

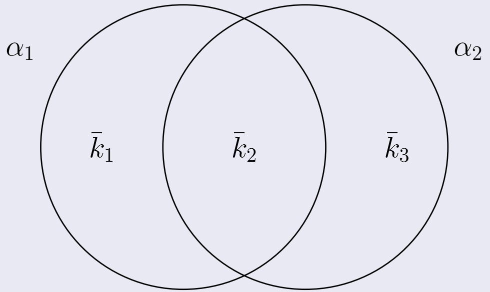
Let: $\bar{D} := D[1\!:\!k] \in \mathcal{E}^k$, $k < n$, with embedding dimension $t \leq r$; $B := \mathcal{K}^\dagger(\bar{D}) = \bar{U}_B S \bar{U}_B^T$, $\bar{U}_B \in \mathcal{M}^{k \times t}$, $\bar{U}_B^T \bar{U}_B = I_t$, $S \in \mathcal{S}^t_{++}$.
Furthermore, let $U_B := \begin{bmatrix} \bar{U}_B & \frac{1}{\sqrt{k}} e \end{bmatrix} \in \mathcal{M}^{k \times (t+1)}$,
$U := \begin{bmatrix} U_B & 0 \\ 0 & I_{n-k} \end{bmatrix}$, and let $\begin{bmatrix} V & \frac{U^T e}{\|U^T e\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal. Then:

$$\begin{aligned} \operatorname{face} \mathcal{K}^\dagger\left(\mathcal{E}^n(1\!:\!k, \bar{D})\right) &= \left(U \mathcal{S}_+^{n-k+t+1} U^T\right) \cap \mathcal{S}_C \\ &= (UV) \mathcal{S}_+^{n-k+t} (UV)^T \end{aligned}$$

Note that we add $\frac{1}{\sqrt{k}} e$ to represent $\mathcal{N}(\mathcal{K})$; then we use V to eliminate e to recover a centered face.

Sets for Intersecting Cliques/Faces

$\alpha_1 := 1{:}(\bar{k}_1 + \bar{k}_2); \quad \alpha_2 := (\bar{k}_1 + 1){:}(\bar{k}_1 + \bar{k}_2 + \bar{k}_3)$



For each clique $|\alpha| = k$, we get a corresponding face/subspace ($k \times r$ matrix) representation. We now see how to handle two cliques, α_1, α_2, that intersect.

THEOREM 2: Clique/Facial Intersection Using Subspace Intersection

$\{\ \alpha_1, \alpha_2 \subseteq 1\!:\!n; \quad k := |\alpha_1 \cup \alpha_2|$

For $i = 1, 2$: $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, embedding dimension t_i;

$B_i := \mathcal{K}^{\dagger}(\bar{D}_i) = \bar{U}_i S_i \bar{U}_i^T$, $\bar{U}_i \in \mathcal{M}^{k_i \times t_i}$, $\bar{U}_i^T \bar{U}_i = I_{t_i}$, $S_i \in \mathcal{S}^{t_i}_{++}$;

$U_i := \begin{bmatrix} \bar{U}_i & \frac{1}{\sqrt{k_i}} e \end{bmatrix} \in \mathcal{M}^{k_i \times (t_i+1)}$; and $\bar{U} \in \mathcal{M}^{k \times (t+1)}$ satisfies

$$\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right), \text{ with } \bar{U}^T \bar{U} = I_{t+1}$$

cont. . .

Two (Intersecting) Clique Reduction, cont...

THEOREM 2 Nonsing. Clique/Facial Inters. cont...

cont... with

$$\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right), \text{with } \bar{U}^T\bar{U} = I_{t+1};$$

let: $U := \begin{bmatrix} \bar{U} & 0 \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{M}^{n\times(n-k+t+1)}$ and

$\begin{bmatrix} V & \frac{U^Te}{\|U^Te\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal. Then

$$\begin{aligned} \bigcap_{i=1}^2 \text{face}\, \mathcal{K}^\dagger\left(\mathcal{E}^n(\alpha_i, \bar{D}_i)\right) &= \left(U\mathcal{S}_+^{n-k+t+1}U^T\right) \cap \mathcal{S}_C \\ &= (UV)\mathcal{S}_+^{n-k+t}(UV)^T \end{aligned}$$

Expense/Work of (Two) Clique/Facial Reductions

Subspace Intersection for Two Intersecting Cliques/Faces

Suppose:

$$U_1 = \begin{bmatrix} U_1' & 0 \\ U_1'' & 0 \\ 0 & I \end{bmatrix} \quad \text{and} \quad U_2 = \begin{bmatrix} I & 0 \\ 0 & U_2'' \\ 0 & U_2' \end{bmatrix}$$

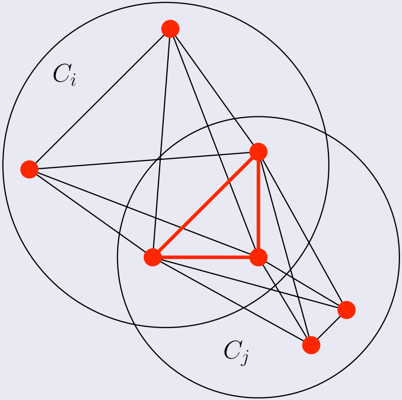
Then:

$$U := \begin{bmatrix} U_1' \\ U_1'' \\ U_2'(U_2'')^\dagger U_1'' \end{bmatrix} \quad \text{or} \quad U := \begin{bmatrix} U_1'(U_1'')^\dagger U_2'' \\ U_2'' \\ U_2' \end{bmatrix}$$

(Efficiently) satisfies:

$$\mathcal{R}(U) = \mathcal{R}(U_1) \cap \mathcal{R}(U_2)$$

Two (Intersecting) Clique Reduction Figure



Completion: missing distances can be recovered if desired.

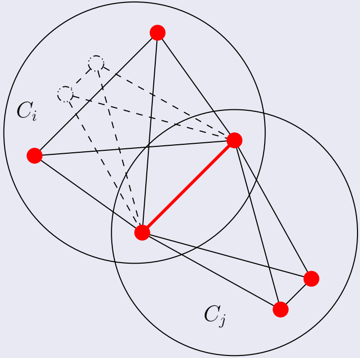
Two (Intersecting) Clique Explicit Delayed Completion

COR. Intersection with Embedding Dim. r/Completion

Hypotheses of Theorem 2 holds. Let $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, for $i = 1, 2$, $\beta \subseteq \alpha_1 \cap \alpha_2, \gamma := \alpha_1 \cup \alpha_2, \bar{D} := D[\beta], B := \mathcal{K}^\dagger(\bar{D})$, $\quad \bar{U}_\beta := \bar{U}(\beta, :)$, where $\bar{U} \in \mathcal{M}^{k \times (t+1)}$ satisfies intersection equation of Theorem 2. Let $\begin{bmatrix} \bar{V} & \frac{\bar{U}^T e}{\|\bar{U}^T e\|} \end{bmatrix} \in \mathcal{M}^{t+1}$ be orthogonal. Let $\boxed{Z := (J\bar{U}_\beta \bar{V})^\dagger B((J\bar{U}_\beta \bar{V})^\dagger)^T}$. If the embedding dimension for $\bar{D}$ is r, THEN $t = r$ in Theorem 2, and $Z \in \mathcal{S}^r_+$ is the unique solution of the equation $(J\bar{U}_\beta \bar{V})Z(J\bar{U}_\beta \bar{V})^T = B$, and the exact completion is

$$\boxed{D[\gamma] = \mathcal{K}\left(PP^T\right)} \quad \text{where} \quad \boxed{P := UVZ^{\frac{1}{2}} \in \mathbb{R}^{|\gamma| \times r}}$$

Two (Intersecting) Clique Reduction Figure/Singular Case



Use R as lower bound in singular/nonrigid case.

COR. Clique-Sing.; Intersect. Embedding Dim. $r-1$

Hypotheses of previous COR holds. For $i=1,2$, let $\beta \subset \delta_i \subseteq \alpha_i$, $A_i := J\bar{U}_{\delta_i}\bar{V}$, where $\bar{U}_{\delta_i} := \bar{U}(\delta_i,:)$, and $B_i := \mathcal{K}^\dagger(D[\delta_i])$. Let $\bar{Z} \in \mathcal{S}^t$ be a particular solution of the linear systems

$$\begin{array}{rcl} A_1 Z A_1^T & = & B_1 \\ A_2 Z A_2^T & = & B_2. \end{array}.$$

If the embedding dimension of $D[\delta_i]$ is r, for $i=1,2$, but the embedding dimension of $\bar{D} := D[\beta]$ is $r-1$, then the following holds. cont. . .

COR. Clique-Degen. cont. . .

The following holds:

1. $\dim \mathcal{N}(A_i) = 1$, for $i = 1, 2$.
2. For $i = 1, 2$, let $n_i \in \mathcal{N}(A_i)$, $\|n_i\|_2 = 1$, and $\Delta Z := n_1 n_2^T + n_2 n_1^T$. Then, Z is a solution of the linear systems if and only if
   $Z = \bar{Z} + \tau \Delta Z$, for some $\tau \in \mathcal{R}$
3. There are at most two nonzero solutions, τ_1 and τ_2, for the generalized eigenvalue problem $-\Delta Z v = \tau \bar{Z} v$, $v \neq 0$. Set $Z_i := \bar{Z} + \frac{1}{\tau_i} \Delta Z$, for $i = 1, 2$. Then the exact completion is one of $D[\gamma] \in \{\mathcal{K}(\bar{U}\bar{V}Z_i\bar{V}^T\bar{U}^T) : i = 1, 2\}$

Completing SNL (Delayed use of Anchor Locations)

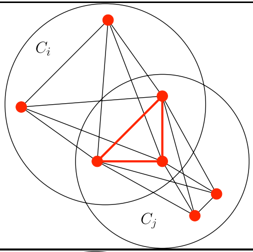
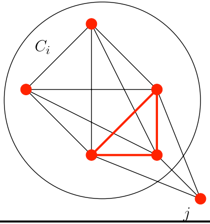
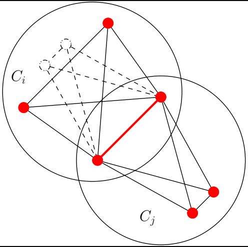
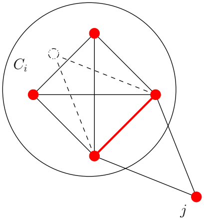
Rotate to Align the Anchor Positions

- Given $P = \begin{bmatrix} P_1 \\ P_2 \end{bmatrix} \in \mathbb{R}^{n \times r}$ such that $D = \mathcal{K}(PP^T)$
- Solve the orthogonal Procrustes problem:

$$\begin{array}{ll} \min & \|A - P_2 Q\| \\ \text{s.t.} & Q^T Q = I \end{array}$$

  $P_2^T A = U \Sigma V^T$ SVD decomposition; set $Q = UV^T$;
  (Golub/Van Loan, Algorithm 12.4.1)
- Set $X := P_1 Q$

Algorithm: Four Cases

	Clique Union	Node Absorption
Rigid		
Non-rigid		

ALGOR: clique union; facial reduct.; delay compl.

Initialize: Find initial set of cliques.

$C_i := \left\{ j : (D_p)_{ij} < (R/2)^2 \right\}$, for $i = 1, \ldots, n$

Iterate

- For $|C_i \cap C_j| \geq r + 1$, do Rigid Clique Union
- For $|C_i \cap \mathcal{N}(j)| \geq r + 1$, do Rigid Node Absorption
- For $|C_i \cap C_j| = r$, do Non-Rigid Clique Union (lower bnds)
- For $|C_i \cap \mathcal{N}(j)| = r$, do Non-Rigid Node Absorp. (lower bnds)

Finalize

When $\exists$ a clique containing all anchors, use computed facial representation and positions of anchors to solve for X

Results - Data for Random Noisless Problems

- 2.16 GHz Intel Core 2 Duo, 2 GB of RAM
- Dimension $r = 2$
- Square region: $[0,1] \times [0,1]$
- $m = 9$ anchors
- Using only Rigid Clique Union and Rigid Node Absorption
- Error measure: Root Mean Square Deviation

$$\text{RMSD} = \left(\frac{1}{n} \sum_{i=1}^{n} \| p_i - p_i^{\text{true}} \|^2 \right)^{1/2}$$

Results - Large n (SDP size $O(n^2)$)

n # of Sensors Located

n # sensors \ R	0.07	0.06	0.05	0.04
2000	2000	2000	1956	1374
6000	6000	6000	6000	6000
10000	10000	10000	10000	10000

CPU Seconds

# sensors \ R	0.07	0.06	0.05	0.04
2000	1	1	1	3
6000	5	5	4	4
10000	10	10	9	8

RMSD (over located sensors)

n # sensors \ R	0.07	0.06	0.05	0.04
2000	$4e{-16}$	$5e{-16}$	$6e{-16}$	$3e{-16}$
6000	$4e{-16}$	$4e{-16}$	$3e{-16}$	$3e{-16}$
10000	$3e{-16}$	$5e{-16}$	$4e{-16}$	$4e{-16}$

Results - N Huge SDPs Solved

Large-Scale Problems

# sensors	# anchors	radio range	RMSD	Time
20000	9	.025	$5e-16$	25s
40000	9	.02	$8e-16$	1m 23s
60000	9	.015	$5e-16$	3m 13s
100000	9	.01	$6e-16$	9m 8s

Size of SDPs Solved: $N = \binom{n}{2}$ (# vrbls)

$\mathcal{E}_n(\text{density of } \mathcal{G}) = \pi R^2$; $M = \mathcal{E}_n(|E|) = \pi R^2 N$ (# constraints)

Size of SDP Problems:

$M = [3,078,915 \quad 12,315,351 \quad 27,709,309 \quad 76,969,790]$

$N = 10^9\,[0.2000 \quad 0.8000 \quad 1.8000 \quad 5.0000]$

Locally Recover Exact EDMs

Nearest EDM

- Given clique α; corresp. $\mathrm{EDM}\, D_\epsilon = D + N_\epsilon$, N_ϵ noise
- we need to find the smallest face containing $\mathcal{E}^n(\alpha, D)$.
- $$\begin{cases} \min & \|\mathcal{K}(X) - D_\epsilon\| \\ \text{s.t.} & \operatorname{rank}(X) = r, Xe = 0, X \succeq 0 \\ & X \succeq 0. \end{cases}$$
- Eliminate the constraints: $Ve = 0, V^T V = I$, $\mathcal{K}_V(X) := \mathcal{K}(VXV^T)$:

$$\begin{array}{rcll} U_r^* & \in & \operatorname{argmin} & \frac{1}{2}\|\mathcal{K}_V(UU^T) - D_\epsilon\|_F^2 \\ & & \text{s.t.} & U \in M^{(n-1)r}. \end{array}$$

The nearest EDM is $D^* = \mathcal{K}_V(U_r^*(U_r^*)^T)$.

Solve Overdetermined Nonlin. Least Squares Prob.

Newton (expensive) or Gauss-Newton (less accurate)

$$F(U) := \text{us2vec}\left(\mathcal{K}_V(UU^T) - D_\epsilon\right), \quad \min_U f(U) := \frac{1}{2}\|F(U)\|^2$$

Derivatives: gradient and Hessian

$$\nabla f(U)(\Delta U) = \langle 2\left(\mathcal{K}_V^*\left[\mathcal{K}_V(UU^T) - D_\epsilon\right]\right)U, \Delta U\rangle$$

$$\nabla^2 f(U) = 2\text{vec}\left(\mathcal{L}_U^*\mathcal{K}_V^*\mathcal{K}_V\mathcal{S}_\Sigma\mathcal{L}_U + \mathcal{K}_V^*\left(\mathcal{K}_V(UU^T) - D_\epsilon\right)\right)\text{Mat}$$

where $\mathcal{L}_U(\cdot) = \cdot U^T$; $\mathcal{S}_\Sigma(U) = \frac{1}{2}(U + U^T)$

random noisy probs; $r = 2, m = 9, nf = 1e-6$

- Using only Rigid Clique Union, preliminary results:

remaining cliques

n / R	1.0	0.9	0.8	0.7	0.6
1000	1.00	5.00	11.00	40.00	124.00
2000	1.00	1.00	1.00	1.00	7.00
3000	1.00	1.00	1.00	1.00	1.00
4000	1.00	1.00	1.00	1.00	1.00
5000	1.00	1.00	1.00	1.00	1.00

cpu seconds

n / R	1.0	0.9	0.8	0.7	0.6
1000	9.43	6.98	5.57	5.04	4.05
2000	12.46	12.18	12.43	11.18	9.89
3000	18.08	18.50	19.07	18.33	16.33
4000	25.18	24.01	24.02	23.80	22.12
5000	38.13	31.66	30.26	30.32	29.88

max-log-error

n / R	1.0	0.9	0.8	0.7	0.6
1000	-3.28	-4.19	-2.92	*Inf*	*Inf*
2000	-3.63	-3.81	-3.82	-2.39	-3.73
3000	-3.51	-3.98	-3.25	-3.90	-3.28
4000	-4.15	-4.05	-3.52	-3.04	-3.33
5000	-4.80	-4.38	-3.89	-4.13	-3.40

Summary

- SDP relaxation of SNL is highly (implicitly) degenerate: The feasible set of this SDP is restricted to a low dim. face of the SDP cone, causing the Slater constraint qualification (strict feasibility) to fail
- We take advantage of this degeneracy by finding explicit representations of intersections of faces of the SDP cone corresponding to unions of intersecting cliques
- Without using an SDP-solver (eg. SeDuMi or SDPT3), we quickly compute the exact solution to the SDP relaxation

Thanks for your attention!

Theory and Applications of Degeneracy in Cone Optimization

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

IPAM, Oct 15, 2010