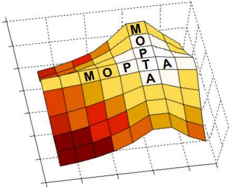
Taking Advantage of Degeneracy in Cone Optimization: with Applications to Sensor Network Localization

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo

at: MOPTA 2012, Lehigh University



Motivation: Loss of Slater CQ/Facial reduction

- optimization algorithms rely on the KKT system;
  and require that some constraint qualification (CQ) holds
  (Slater's CQ/strict feasibility for convex conic optimization)

- However, surprisingly many conic opt, SDP relaxations,
  instances arising from applications (QAP, GP, strengthened MC, SNL, POP, Molecular Conformation)
  do not satisfy Slater's CQ/are degenerate

- lack of Slater's CQ results in: unbounded dual solutions;
  theoretical and numerical difficulties,
  in particular for *primal-dual interior-point methods*.

- solution:
  - theoretical *facial reduction* *(Borwein, W.'81)*
  - preprocess for regularized smaller problem (Cheung, Schurr, W.'11)
  - take advantage of degeneracy (for SNL)
    (Krislock, W.'10; Krislock, Rendl, W.'10)

Outline: Regularization/Facial Reduction

1. Preprocessing/Regularization
   - Abstract convex program
     - LP case
     - CP case
   - Cone optimization/SDP case
2. Applications: QAP, GP, SNL, Molecular conformation ...
   - SNL; highly (implicit) degenerate/low rank solutions

Background/Abstract convex program

$$\text{(ACP)} \quad \inf_x f(x) \text{ s.t. } g(x) \preceq_K 0, x \in \Omega$$

where:

- $f : \mathbb{R}^n \to \mathbb{R}$ convex; $g : \mathbb{R}^n \to \mathbb{R}^m$ is K-convex
  - $K \subset \mathbb{R}^m$ closed convex cone; $\Omega \subseteq \mathbb{R}^n$ convex set
  - $a \preceq_K b \iff b - a \in K$
  - $g(\alpha x + (1-\alpha y)) \preceq_K \alpha g(x) + (1-\alpha)g(y)$,
    $\forall x, y \in \mathbb{R}^n, \forall \alpha \in [0,1]$

Slater's CQ: $\exists \hat{x} \in \Omega$ s.t. $g(\hat{x}) \in -\operatorname{int} K$ $\quad (g(x) \prec_K 0)$

- guarantees strong duality
- essential for efficiency/stability in primal-dual interior-point methods
  ((near) loss of strict feasibility correlates with number of iterations and loss of accuracy)

Case of Linear Programming, LP

Primal-Dual Pair: $A, m \times n$ / $\mathcal{P} = \{1, \ldots, n\}$ constr. matrix/set

$$\text{(LP-P)} \quad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & A^\top y \leq c \end{array} \qquad \text{(LP-D)} \quad \begin{array}{ll} \min & c^\top x \\ \text{s.t.} & Ax = b, \; x \geq 0. \end{array}$$

Slater's CQ for (LP-P) / Theorem of alternative

$\exists \hat{y}$ s.t. $c - A^\top \hat{y} > 0$, $\quad \left((c - A^\top \hat{y})_i > 0, \forall i \in \mathcal{P} =: \mathcal{P}^< \right)$

iff

$$Ad = 0, \; c^\top d = 0, \; d \geq 0 \implies d = 0 \qquad (*)$$

implicit equality constraints: $i \in \mathcal{P}^=$

Finding solution $0 \neq d^*$ to $(*)$ with max number of non-zeros determines ($\mathcal{F}^y$ feasible set)

$$d_i^* > 0 \quad \implies (c - A^\top y)_i = 0, \forall y \in \mathcal{F}^y \quad (i \in \mathcal{P}^=)$$

Rewrite implicit-equalities to equalities/ Regularize LP

Facial Reduction: $A^\top y \leq_f c$; minimal face $f \trianglelefteq \mathbb{R}^n_+$

$$(\text{LP}_{reg}\text{-P}) \quad \begin{array}{ll} \max & b^\top y \\ \text{s.t.} & (A^<)^\top y \leq c^< \\ & (A^=)^\top y = c^= \end{array} \qquad (\text{LP}_{reg}\text{-D}) \quad \begin{array}{ll} \min & (c^<)^\top x^< + (c^=)^\top x^= \\ \text{s.t.} & \begin{bmatrix} A^< & A^= \end{bmatrix} \begin{pmatrix} x^< \\ x^= \end{pmatrix} = b \\ & x^< \geq 0, x^= \text{ free} \end{array}$$

Mangasarian-Fromovitz CQ (MFCQ) holds

(after deleting redundant equality constraints!)

$$\left(\exists \hat{y} : \quad \underset{(A^<)^\top \hat{y} < c^<}{\underline{i \in \mathcal{P}^<}} \quad \underset{(A^=)^\top \hat{y} = c^=}{\underline{i \in \mathcal{P}^=}} \right) \qquad (A^=)^\top \text{ is onto}$$

MFCQ holds iff dual optimal set is compact

Numerical difficulties if MFCQ fails; in particular for interior point methods! Modelling issue?

Facial Reduction/Preprocessing

Linear Programming Example, $x \in \mathbb{R}^2$

$$\begin{array}{ll} \max & \begin{pmatrix} 2 & 6 \end{pmatrix} y \\ \text{s.t.} & \begin{bmatrix} -1 & -1 \\ 1 & 1 \\ 1 & -1 \\ -2 & 2 \end{bmatrix} y \leq \begin{pmatrix} 1 \\ 2 \\ 1 \\ -2 \end{pmatrix} \end{array}$$

$\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ feasible; weighted last two rows $\begin{bmatrix} 1 & -1 & 1 \\ -2 & 2 & -2 \end{bmatrix}$ sum to zero. $\mathcal{P}^< = \{1,2\}, \mathcal{P}^= = \{3,4\}$

Facial reduction to 1 dim; substit. for y

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad -1 \leq t \leq \tfrac{1}{2}, \qquad t^* = \tfrac{1}{2}.$$

Facial Reduction on Dual/Preprocessing

Linear Programming Example, $x \in \mathbb{R}^5$

$$\begin{array}{ll} \min & \begin{pmatrix} 2 & 6 & -1 & -2 & 7 \end{pmatrix} x \\ \text{s.t.} & \begin{bmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & -1 & -1 & 0 & 1 \end{bmatrix} x = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \\ & x \geq 0 \end{array}$$

Sum the two constraints:

$$2x_1 + x_4 + x_5 = 0 \implies x_1 = x_4 = x_5 = 0.$$

yields the equivalent simplified problem in a smaller face

$$\begin{array}{ll} \min & \begin{pmatrix} 6 & -1 \end{pmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} \\ \text{s.t.} & \begin{bmatrix} 1 & 1 \end{bmatrix} \begin{pmatrix} x_2 \\ x_3 \end{pmatrix} = 1 \\ & x_2, x_3 \geq 0, x_1 = x_4 = x_5 = 0 \end{array}$$

Case of ordinary convex programming, CP

$$\text{(CP)} \qquad \sup_y \; b^\top y \text{ s.t. } g(y) \le 0,$$

where

- $b \in \mathbb{R}^m$; $g(y) = \big(g_i(y)\big) \in \mathbb{R}^n$, $g_i : \mathbb{R}^m \to \mathbb{R}$ convex, $\forall i \in \mathbb{P}$
- Slater's CQ: $\exists \hat{y}$ s.t. $g_i(\hat{y}) < 0, \forall i$ (implies MFCQ)
- Slater's CQ fails implies implicit equality constraints exist, i.e.:
  $\mathcal{P}^= := \{i \in \mathcal{P} : g(y) \le 0 \implies g_i(y) = 0\} \ne \emptyset$
  Let $\mathcal{P}^< := \mathcal{P} \backslash \mathcal{P}^=$ and

$$g^< := (g_i)_{i \in \mathcal{P}^<}, g^= := (g_i)_{i \in \mathcal{P}^=}$$

Rewrite implicit equalities to *equalities*/ Regularize CP

(CP) is equivalent to $g(y) \preceq_f 0$, f is minimal face

$$(\mathrm{CP}_{\mathrm{reg}}) \qquad \begin{array}{ll} \sup & b^\top y \\ \text{s.t.} & g^<(y) \preceq 0 \\ & y \in \mathcal{F}^= \end{array} \qquad \text{or } (g^=(y) = 0)$$

where $\mathcal{F}^= := \{y : g^=(y) = 0\}$. Then

$\mathcal{F}^= = \{y : g^=(y) \preceq 0\}$, so is a convex set!

Slater's CQ holds for (CP_{reg}) $\exists \hat{y} \in \mathcal{F}^= : g^<(\hat{y}) \prec 0$

modelling issue again?

Faithfully convex case

Faithfully convex function f (Rockafellar'70)

f affine on a line segment only if affine on complete line containing the segment (e.g. analytic convex functions)

$\mathcal{F}^{=} = \{y : g^{=}(y) = 0\}$ is an affine set

Then:

$\mathcal{F}^{=} = \{y : Vy = V\hat{y}\}$ for some $\hat{y}$ and full-row-rank matrix V.

Then MFCQ holds for

$$(\mathrm{CP}_{\mathrm{reg}}) \qquad \begin{array}{lrcl} \sup & b^\top y & & \\ \text{s.t.} & g^{<}(y) & \leq & 0 \\ & Vy & = & V\hat{y} \end{array}$$

Semidefinite Programming, SDP

$K = \mathcal{S}^n_+ = K^*$ nonpolyhedral cone!

$$\text{(SDP-P)} \quad v_P = \sup_{y \in \mathbb{R}^m} b^\top y \text{ s.t. } g(y) := \mathcal{A}^* y - c \preceq_{\mathcal{S}^n_+} 0$$

$$\text{(SDP-D)} \quad v_D = \inf_{x \in \mathcal{S}^n} \langle c, x \rangle \text{ s.t. } \mathcal{A}x = b, \; x \succeq_{\mathcal{S}^n_+} 0$$

where:

- PSD cone $\mathcal{S}^n_+ \subset \mathcal{S}^n$ symm. matrices
- $c \in \mathcal{S}^n$, $b \in \mathbb{R}^m$
- $\mathcal{A} : \mathcal{S}^n \to \mathbb{R}^m$ is a linear map, with adjoint $\mathcal{A}^*$
  $\mathcal{A}x = (\operatorname{trace} A_i x) \in \mathbb{R}^m$
  $\mathcal{A}^* y = \sum_{i=1}^m A_i y_i \in \mathcal{S}^n$

Slater's CQ/Theorem of Alternative

(Assume feasibility: $\exists \tilde{y}$ s.t. $c - \mathcal{A}^*\tilde{y} \succeq 0$.)

$$\exists \hat{y} \text{ s.t. } s = c - \mathcal{A}^*\hat{y} \succ 0 \qquad (\textit{Slater})$$

iff

$$\mathcal{A}d = 0,\ \langle c, d \rangle = 0,\ d \succeq 0 \implies d = 0 \qquad (*)$$

Faces of Cones - Useful for Charact. of Opt.

Face

A convex cone F is a face of K, denoted $F \unlhd K$, if
$x, y \in K$ and $x + y \in F \implies x, y \in F$
($F \lhd K$ proper face)

Conjugate Face

If $F \unlhd K$, the conjugate face (or complementary face) of F is
$F^c := F^\perp \cap K^* \unlhd K^*$
If $x \in \mathrm{ri}(F)$, then $F^c = \{x\}^\perp \cap K^*$.

Minimal Faces

$f_P :=$ face $\mathcal{F}_P^s \unlhd K$, $\quad \mathcal{F}_P^s$ is primal feasible set
$f_D :=$ face $\mathcal{F}_D^x \unlhd K^*$, $\quad \mathcal{F}_D^x$ is dual feasible set
where: K^* denotes the dual (nonnegative polar) cone;
face S denotes the smallest face containing S.

Regularization Using Minimal Face

Borwein-W.'81 , $f_P = \text{face}\, \mathcal{F}_P^s$

(SDP-P) is equivalent to the regularized

$$(\text{SDP}_{reg}\text{-P}) \qquad v_{RP} := \sup_y \{\langle b, y\rangle \; : \; \mathcal{A}^* y \preceq_{f_P} c\}$$

(slacks: $s = c - \mathcal{A}^* y \in f_p$)

Lagrangian Dual DRP Satisfies Strong Duality:

$$(\text{SDP}_{reg}\text{-D}) \quad v_{DRP} := \inf_x \{\langle c, x\rangle \; : \; \mathcal{A} x = b, \; x \succeq_{f_P^*} 0\}$$
$$= v_P = v_{RP}$$

and v_{DRP} is attained.

(SYMMETRIC) Subspace form

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 P-D Pair in Subspace Form, (e.g. N&N94)

Particular solution + solution of homogeneous equation

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

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

Minimal subspaces

Faces of Recession Directions (feasible case/homog. prob.)

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

Recall: for feasible sets $\mathcal{F}_P^s, \mathcal{F}_D^x$

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

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 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

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 (Slater)
   (from BW: $f_P^* = K^*$)
2. $\mathcal{L}^\perp \cap (f_P - f_P) = \mathcal{L}_{PM}^\perp = \mathcal{L}^\perp \cap (K - K)$ is a CQ (if K is FDC (nice))

SDP Regularization process

Alternative to Slater CQ

$$\mathcal{A}d = 0, \ \langle c, d\rangle = 0, \ 0 \neq d \succeq_{\mathcal{S}^n_+} 0 \qquad (*)$$

Determine a proper face $f \lhd \mathcal{S}^n_+$

Let d solve $(*)$ with $d = P d_+ P^\top$, $d_+ \succ 0$, and $[P \ Q] \in \mathbb{R}^{n\times n}$ orthogonal. Then

$$c - \mathcal{A}^* y \succeq_{\mathcal{S}^n_+} 0 \implies \langle c - \mathcal{A}^* y, d^*\rangle = 0$$

$$\implies \mathcal{F}^s_P \subseteq \mathcal{S}^n_+ \cap \{d^*\}^\perp = Q \mathcal{S}^{\bar{n}}_+ Q^\top \lhd \mathcal{S}^n_+$$

(implicit rank reduction, $\bar{n} < n$)

Regularizing SDP

- at most $n-1$ iterations to satisfy Slater's CQ.
- to check Theorem of Alternative

$$\mathcal{A}d = 0,\ \langle c, d\rangle = 0,\ 0 \neq d \succeq_{\mathcal{S}^n_+} 0, \qquad (*)$$

use stable auxiliary problem

$$(AP) \qquad \min_{\delta, d}\ \delta \ \text{ s.t. } \left\| \begin{bmatrix} \mathcal{A}d \\ \langle c, d\rangle \end{bmatrix} \right\|_2 \leq \delta, \quad \operatorname{trace}(d) = \sqrt{n}, \quad d \succeq 0.$$

- Both (AP) and its dual satisfy Slater's CQ.

Auxiliary Problem

$$(AP) \qquad \min_{\delta,d} \ \delta \ \text{ s.t. } \left\| \begin{bmatrix} \mathcal{A}d \\ \langle c, d \rangle \end{bmatrix} \right\|_2 \leq \delta,$$
$$\operatorname{trace}(d) = \sqrt{n}, d \succeq 0.$$

Both (AP) and its dual satisfy Slater's CQ ... but ...

Cheung-Schurr-W'11, a $k = 1$ step CQ

Strict complementarity holds for (AP)

iff

$k = 1$ steps are needed to regularize (SDP-P).

Regularizing SDP

Minimal face containing $\mathcal{F}_P^s := \{s : s = c - \mathcal{A}^* y \succeq 0\}$

$$f_P = Q\mathcal{S}_+^{\bar{n}} Q^\top$$

for some $n \times n$ orthogonal matrix $U = [P\ Q]$

(SPD-P) is equivalent to

$$\sup_y \; b^\top y \text{ s.t. } g^{\prec}(y) \preceq 0, \; g^{=}(y) = 0,$$

where

$$g^{\prec}(y) := Q^\top(\mathcal{A}^* y - c)Q$$

$$g^{=}(y) := \begin{bmatrix} P^\top(\mathcal{A}^* y - c)P \\ P^\top(\mathcal{A}^* y - c)Q + Q^\top(\mathcal{A}^* y - c)P \end{bmatrix}.$$

(gen.) Slater CQ holds for the reduced program:
$\exists \hat{y}$ s.t. $g^{\prec}(y) \prec 0$ and $g^{=}(y) = 0$.

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)

Strict Complementarity and Nonzero Gaps

Theorem (Tuncel-W.'11): 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 SDP-P and SDP-D with data $(\mathcal{L}, K, \bar{s}, \bar{x})$ has a finite nonzero duality gap.

(Partial Converse)

(2) If
(a) SDP-P and SDP-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 SDP-P and SDP-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); use min. face f_P (and/or implicit rank reduction)
- goal: a backwards stable preprocessing algorithm to handle (feasible) conic problems for which Slater's CQ (almost) fails
- Failure of strict complementarity for associated recession problems is related to existence of finite nonzero duality gap; provides a means of generating instances for testing.

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

Slater CQ $\qquad$ str. compl.

$\updownarrow$

pos. duality gap

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

Part II: Applications of SDP where Slater's CQ fails

Instances of SDP relaxations of NP-hard combinatorial optimization problems with row and column sum and $0, 1$ constraints

- Quadratic Assignment (Zhao-Karish-Rendl-W.'96)
- Graph partitioning (W.-Zhao'99)

Low rank problems

- Sensor network localization (SNL) problem (Krislock-W.'10, Krislock-Rendl-W.'10)
- Molecular conformation (Burkowski-Cheung-W.'11)
- general SDP relaxation of low-rank matrix completion problem

SNL (K-W'10,K-R-W'10)

Highly (implicit) degenerate/low-rank problem

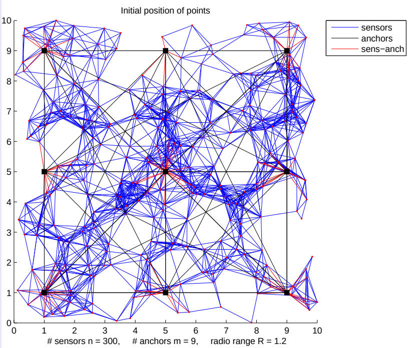
- high (implicit) degeneracy translates to low rank solutions
- fast, high accuracy solutions

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

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

Sensor Localization Problem/Partial EDM

Sensors ○ and Anchors ■



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 $\mathbb{R}^r$: a mapping of nodes $v_i \mapsto p_i \in \mathbb{R}^r$ 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.

Connections to Semidefinite Programming (SDP)

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

$P^\top = \begin{bmatrix} p_1 & p_2 & \dots & p_n \end{bmatrix} \in \mathcal{M}^{r\times n}$;
$B := PP^\top \in \mathcal{S}^n_+$ (Gram matrix of inner products);
$\operatorname{rank} B = r$; let $D \in \mathcal{E}^n$ corresponding EDM; $e = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix}^\top$

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

Euclidean Distance Matrices; Semidefinite Matrices

Moore-Penrose Generalized Inverse $\mathcal{K}^\dagger$

$$B \succeq 0 \quad \Longrightarrow \quad D = \mathcal{K}(B) = \mathrm{diag}\,(B)\, e^\top + e\, \mathrm{diag}\,(B)^\top - 2B \in \mathcal{E}$$

$$D \in \mathcal{E} \quad \Longrightarrow \quad B = \mathcal{K}^\dagger(D) = -\frac{1}{2} J \mathrm{offDiag}\,(D)\, J \succeq 0, Be = 0$$

Theorem (Schoenberg, 1935)

A (hollow) matrix D (with $\mathrm{diag}\,(D) = 0, D \in S_H$) is a
Euclidean distance matrix
if and only if

$$B = \mathcal{K}^\dagger(D) \succeq 0.$$

And

$$\mathrm{embdim}\,(D) = \mathrm{rank}\,\left(\mathcal{K}^\dagger(D)\right), \quad \forall D \in \mathcal{E}^n$$

Popular Techniques; SDP Relax.; Highly Degen.

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

- $\min_{B \succeq 0} \|H \circ (\mathcal{K}(B) - D)\|$; $\operatorname{rank} B = r$;
  typical weights: $H_{ij} = 1/\sqrt{D_{ij}}$, if $ij \in E$, $H_{ij} = 0$ otherwise.
- 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

Basic Single Clique/Facial Reduction

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.

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

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

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

if $\alpha = 1 : k$; embedding dim $\text{embdim}\,(\bar{D}) = t \le r$

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

BASIC THEOREM for Single Clique/Facial Reduction

Let:

- $\bar{D} := D[1:k] \in \mathcal{E}^k$, $k < n$, $\text{embdim}(\bar{D}) = t \leq r$ be given;
- $B := \mathcal{K}^\dagger(\bar{D}) = \bar{U}_B S \bar{U}_B^\top$, $\bar{U}_B \in \mathcal{M}^{k\times t}$, $\bar{U}_B^\top \bar{U}_B = I_t$, $S \in \mathcal{S}_{++}^t$ be full rank orthogonal decomposition of Gram matrix;
- $U_B := \begin{bmatrix} \bar{U}_B & \frac{1}{\sqrt{k}}e \end{bmatrix} \in \mathcal{M}^{k\times(t+1)}$, $U := \begin{bmatrix} U_B & 0 \\ 0 & I_{n-k} \end{bmatrix}$, and $\begin{bmatrix} V & \frac{U^\top e}{\|U^\top e\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal.

Then the minimal face:

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

The minimal face

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

Note that the minimal face is defined by the subspace $\mathcal{L} = \mathcal{R}(UV)$. 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.

Facial Reduction for Disjoint Cliques

Corollary from Basic Theorem

let $\alpha_1, \ldots, \alpha_\ell \subseteq 1{:}n$ pairwise disjoint sets, wlog:
$\alpha_i = (k_{i-1}+1){:}k_i, k_0 = 0$, $\alpha := \bigcup_{i=1}^{\ell} \alpha_i = 1{:}|\alpha|$ let
$\bar{U}_i \in \mathbb{R}^{|\alpha_i| \times (t_i+1)}$ with full column rank satisfy $e \in \mathcal{R}(\bar{U}_i)$ and

$$U_i := \begin{array}{c} \\ k_{i-1} \\ |\alpha_i| \\ n-k_i \end{array} \begin{array}{c} \begin{array}{ccc} k_{i-1} & t_i+1 & n-k_i \end{array} \\ \left[\begin{array}{ccc} I & 0 & 0 \\ 0 & \bar{U}_i & 0 \\ 0 & 0 & I \end{array}\right] \end{array} \in \mathbb{R}^{n \times (n-|\alpha_i|+t_i+1)}$$

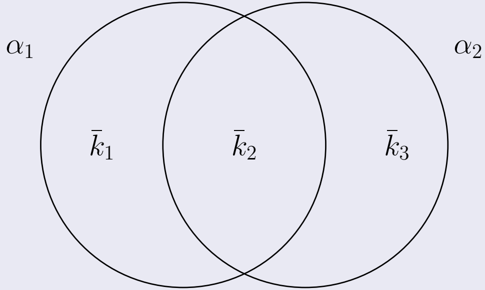
The minimal face is defined by $\mathcal{L} = \mathcal{R}(U)$:

$$U := \begin{array}{c} \\ |\alpha_1| \\ \vdots \\ |\alpha_\ell| \\ n-|\alpha| \end{array} \begin{array}{c} \begin{array}{cccc} t_1+1 & \ldots & t_\ell+1 & n-|\alpha| \end{array} \\ \left[\begin{array}{cccc} \bar{U}_1 & \ldots & 0 & 0 \\ \vdots & \ddots & \vdots & \vdots \\ 0 & \ldots & \bar{U}_\ell & 0 \\ 0 & \ldots & 0 & I \end{array}\right] \end{array} \in \mathbb{R}^{n \times (n-|\alpha|+t+1)},$$

where $t := \sum_{i=1}^{\ell} t_i + \ell - 1$. And $e \in \mathcal{R}(U)$.

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 *complete* the union of two cliques, α_1, α_2, that intersect.

Two (Intersecting) Clique Reduction/Subsp. Repres.

Let:

- $\alpha_1, \alpha_2 \subseteq 1\!:\!n; \quad k := |\alpha_1 \cup \alpha_2|$
- for $i = 1, 2$: $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, embedding dimension t_i;
- $B_i := \mathcal{K}^\dagger(\bar{D}_i) = \bar{U}_i S_i \bar{U}_i^\top$, $\bar{U}_i \in \mathcal{M}^{k_i \times t_i}$, $\bar{U}_i^\top \bar{U}_i = I_{t_i}$, $S_i \in \mathcal{S}^{t_i}_{++}$;
- $U_i := \begin{bmatrix} \bar{U}_i & \frac{1}{\sqrt{k_i}} e \end{bmatrix} \in \mathcal{M}^{k_i \times (t_i+1)}$; and $\bar{U} \in \mathcal{M}^{k \times (t+1)}$

  satisfies $\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right)$, with $\bar{U}^\top \bar{U} = I_{t+1}$
- $U := \begin{bmatrix} \bar{U} & 0 \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{M}^{n \times (n-k+t+1)}$ and $\begin{bmatrix} V & \frac{U^\top e}{\|U^\top e\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal.

Then

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

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:

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

($Q_1 =: (U_1'')^\dagger U_2'', Q_2 = (U_2'')^\dagger U_1''$ orthogonal/rotation)
(Efficiently) satisfies

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

Two (Intersecting) Clique Explicit Delayed Completion

Let:

- Hypotheses of intersecting Theorem (Thm 2) holds
- $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, for $i = 1, 2$, $\beta \subseteq \alpha_1 \cap \alpha_2, \gamma := \alpha_1 \cup \alpha_2$
- $\bar{D} := D[\beta]$ with embedding dimension r
- $B := \mathcal{K}^\dagger(\bar{D}), \quad \bar{U}_\beta := \bar{U}(\beta, :)$, where $\bar{U} \in \mathcal{M}^{k\times(t+1)}$ satisfies intersection equation of Thm 2
- $\begin{bmatrix} \bar{V} & \frac{\bar{U}^\top e}{\|\bar{U}^\top e\|} \end{bmatrix} \in \mathcal{M}^{t+1}$ be orthogonal.
- $\boxed{Z := (J\bar{U}_\beta\bar{V})^\dagger B((J\bar{U}_\beta\bar{V})^\dagger)^\top}$.

<u>THEN</u> $t = r$ in Thm 2, and $Z \in \mathcal{S}^r_+$ is the unique solution of the equation $(J\bar{U}_\beta\bar{V})Z(J\bar{U}_\beta\bar{V})^\top = B$, and the exact completion is

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

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 $D = \mathcal{K}(PP^\top)$
- Solve the orthogonal Procrustes problem:

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

$P_2^\top A = U \Sigma V^\top$ SVD decomposition; set $Q = UV^\top$;
(Golub/Van Loan'79, Algorithm 12.4.1)

- Set $X := P_1 Q$

Summary: Facial Reduction for Cliques

- Using the basic theorem: each clique corresponds to a Gram matrix/corresponding subspace/corresponding face of SDP cone (implicit rank reduction)
- In the case where two cliques intersect, the union of the cliques correspond to the (efficiently computable) intersection of the corresponding faces/subspaces
- Finally, the positions are determined using a Procrustes problem

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:
$M = [3,078,915 \quad 12,315,351 \quad 27,709,309 \quad 76,969,790]$
$N = 10^9 \, [0.2000 \quad 0.8000 \quad 1.8000 \quad 5.0000]$

- J.M. Borwein and H. Wolkowicz, *Characterization of optimality for the abstract convex program with finite-dimensional range*, J. Austral. Math. Soc. Ser. A **30** (1980/81), no. 4, 390–411. MR 83i:90156

- F. Burkowski, Y-L. Cheung, and H. Wolkowicz, *Semidefinite programming and side chain positioning*, Tech. Report CORR 2011, in progress, University of Waterloo, Waterloo, Ontario, 2011.

- Y-L. Cheung, S. Schurr, and H. Wolkowicz, *Preprocessing and reduction for degenerate semidefinite programs*, Tech. Report CORR 2011-02, URL: www.optimization-online.org/DB_HTML/2011/02/2929.html, University of Waterloo, Waterloo, Ontario, 2011, 49 pages.

- G.H. Golub and C.F. Van Loan, *Matrix computations*, 3nd ed., Johns Hopkins University Press, Baltimore, Maryland, 1996.

- N. Krislock, F. Rendl, and H. Wolkowicz, *Noisy sensor network localization using semidefinite representations and facial reduction*, Tech. Report CORR 2010-01, University of Waterloo, Waterloo, Ontario, 2010.

- N. Krislock and H. Wolkowicz, *Explicit sensor network localization using semidefinite representations and facial reductions*, SIAM Journal on Optimization **20** (2010), no. 5, 2679–2708.

- Y.E. Nesterov and A.S. Nemirovski, *Interior point polynomial algorithms in convex programming*, SIAM Publications, SIAM, Philadelphia, USA, 1994.

- G. Pataki, *Bad semidefinite programs: they all look the same*, Tech. report, Department of Operations Research, University of North Carolina, Chapel Hill, 2011.

- R. Tyrrell Rockafellar, *Some convex programs whose duals are linearly constrained*, Nonlinear Programming (Proc. Sympos., Univ. of Wisconsin, Madison, Wis., 1970), Academic Press, New York, 1970, pp. 293–322.
- H. Wolkowicz and Q. Zhao, *Semidefinite programming relaxations for the graph partitioning problem*, Discrete Appl. Math. **96/97** (1999), 461–479, Selected for the special Editors' Choice, Edition 1999. MR 1 724 735
- Q. Zhao, S.E. Karisch, F. Rendl, and H. Wolkowicz, *Semidefinite programming relaxations for the quadratic assignment problem*, J. Comb. Optim. **2** (1998), no. 1, 71–109, Semidefinite programming and interior-point approaches for combinatorial optimization problems (Fields Institute, Toronto, ON, 1996). MR 99f:90103

Thanks for your attention!

Taking Advantage of Degeneracy in Cone Optimization: with Applications to Sensor Network Localization

Henry Wolkowicz

Dept. Combinatorics and Optimization, University of Waterloo

at: MOPTA 2012, Lehigh University