Facial Reduction for Cone Optimization

Henry Wolkowicz
Dept. Combinatorics and Optimization
University of Waterloo

Tuesday Apr. 21, 2015

(with: Drusvyatskiy, Krislock, (Cheung) Voronin)

Motivation: Loss of Slater CQ/Facial reduction

- optimization algorithms key is KKT system;
  (Slater's CQ/strict feasibility for convex conic optimization)
- Slater CQ holds generically
  Surprisingly, for many conic opt, SDP relaxations, arising from applications Slater's fails,
  e.g., SNL, POP, Molecular Conformation, QAP, GP, strengthened MC
- Slater fails $\Longrightarrow$: -unbounded dual solutions;
  -theoretical and numerical difficulties
  (in particular for *primal-dual interior-point methods*).
- solutions?
  - theoretical *facial reduction* *(Borwein, W. '81)*
  - preprocess for regularized smaller problem (Cheung, Schurr, W.'11)
  - take advantage of degeneracy (e.g. recent: Krislock, W.'10; Cheung, Drusvyatskiy, Krislock, W.'14; Reid, Wang, W. Wu'15)

Outline: Regularization/Facial Reduction

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

Facial Reduction on (dual) LP, $Ax = b, x \geq 0$

Theorem of alternative, A full row rank

$\exists \hat{x}$ s.t. $A\hat{x} = b, \hat{x} > 0$

iff

$A^\top y \geq 0,\ b^\top y = 0,\ \implies\ y = 0 \qquad (**)$

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 (use $y^T = (1\ 1)$ in (**)):

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

yields equivalent simplified problem:

$$\boxed{\min \quad 6x_2 - x_3 \quad \underline{\text{s.t.}} \quad x_2 + x_3 = 1, x_2, x_3 \geq 0}$$

Facial Reduction on Primal, $A^T y \leq c$

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

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

active set $\{2,3,4\}$

$\begin{pmatrix} 3/2 \\ 1/2 \end{pmatrix}$ is optimal, $p^* = 6$

weighted last two rows $\begin{bmatrix} 1 & -1 & 1 \\ -2 & 2 & -2 \end{bmatrix}$ sum to zero:

set of implicit equalities: $\mathcal{P}^e := \{3,4\}$

Facial reduction to 1 dim. after substit. for y

$$\begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \max\left\{2 + 8t : -1 \leq t \leq \tfrac{1}{2}\right\}, \quad t^* = \tfrac{1}{2}.$$

General Case?

- Can we do facial reduction in general?
- Is it efficient/worthwhile?
- applications?

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
  - $a \preceq_K b \iff b - a \in K$, $a \prec_K b \iff b - a \in \operatorname{int} K$
  - $g(\alpha x + (1-\alpha y)) \preceq_K \alpha g(x) + (1-\alpha)g(y)$,
    $\forall x, y \in \mathbb{R}^n, \forall \alpha \in [0,1]$

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

- guarantees strong duality
- (near) loss of strict feasibility, nearness to infeasibility, correlates with number of iterations & loss of accuracy

Back to: Case of Linear Programming, LP

Primal-Dual Pair: A onto, $m \times n$, $\mathcal{P} = \{1, \dots, n\}$

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

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

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

iff

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

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

Find $0 \neq d^*$ to $(*)$ with max number of non-zeros
(exposes minimal face containing feasible slacks)
$d_i^* > 0 \implies (c - A^\top y)_i = 0, \forall y \in \mathcal{F}^y \quad (i \in \mathcal{P}^e)$
(where $\mathcal{F}^y$ is primal feasible set)

Make implicit-equalities explicit/ Regularizes LP

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

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

Mangasarian-Fromovitz CQ (MFCQ) holds

(after deleting redundant equality constraints!)

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

MFCQ holds iff dual optimal set is compact

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

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

$\mathcal{P}^e := \{i \in \mathcal{P} : g(y) \le 0 \implies g_i(y) = 0\} \neq \emptyset$

Let $\mathcal{P}^{lt} := \mathcal{P} \backslash \mathcal{P}^e$ and

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

Rewrite implicit equalities to *equalities*/ Regularize CP

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

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

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

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

Slater's CQ holds for (CP$_{reg}$) $\boxed{\exists \hat{y} \in \mathcal{F}^e : g^{lt}(\hat{y}) \prec 0}$

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)

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

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

Then MFCQ holds for regularized

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

Faces of Cones - Useful for Charact. of Opt.

Face

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

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

Polar (Dual) Cone

$$K^* := \{\phi : \langle \phi, k \rangle \geq 0, \ \forall k \in K\}$$

Conjugate Face

If $F \trianglelefteq K$, the conjugate face of F is

$$\boxed{F^c := F^\perp \cap K^* \trianglelefteq K^*}$$

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

Recall: (ACP) $\inf_x f(x)$ s.t. $g(x) \preceq_K 0, x \in \Omega$

- polar cone: $K^* = \{\phi : \langle \phi, y \rangle \geq 0, \forall y \in K\}$.
- $K^f :=$ face(F) minimal face containing feasible set F.

Lemma (Facial Reduction; find exposing vector ϕ)

Suppose $\bar{x}$ is feasible. Then the LHS system

$$\left\{ \begin{array}{ll} (\Omega - \bar{x})^+ \cap \partial \langle \phi, g(\bar{x}) \rangle \neq \emptyset \\ \phi \in K^+, & \langle \phi, g(\bar{x}) \rangle = 0 \end{array} \right\} \quad \textit{implies} \quad K^f \subseteq \phi^\perp \cap K.$$

Proof

line 1 of system implies $\bar{x}$ global min for convex function $\langle \phi, g(\cdot) \rangle$ on Ω; i.e., $0 = \langle \phi, g(\bar{x}) \rangle \leq \langle \phi, g(x) \rangle \leq 0, \forall x \in F$; implies $-g(F) \subseteq \phi^\perp \cap K$. □

Semidefinite Programming, SDP, $\mathcal{S}_+^n$

$K = \mathcal{S}_+^n = K^*$: nonpolyhedral, self-polar, facially exposed

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

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

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 an onto linear map, with adjoint $\mathcal{A}^*$

$$\mathcal{A}x = (\operatorname{trace} A_i x) = (\langle A_i, x \rangle) \in \mathbb{R}^m, \quad A_i \in \mathcal{S}^n$$
$$\mathcal{A}^* y = \sum_{i=1}^m A_i y_i \in \mathcal{S}^n$$

Slater's CQ/Theorem of Alternative

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

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

<u>iff</u>

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

Regularization Using Minimal Face

Borwein-W.'81 , $f_P = \text{face}\, \mathcal{F}^s_P$; min. face of feasible slacks

(SDP-P) is equivalent to the regularized

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

f_p is mininimal face of primal feasible slacks

$$\{ s \succeq 0 : s = c - \mathcal{A}^* y \} \subseteq f_p \trianglelefteq \mathcal{S}^n_+$$

Lagrangian Dual DRP Satisfies Strong Duality:

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

and v_{DRP} is attained.

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 $f_p \trianglelefteq f = Q\mathcal{S}^{\bar{n}}_+ Q^T \lhd \mathcal{S}^n_+$

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

$$\begin{aligned} c - \mathcal{A}^*y \succeq_{\mathcal{S}^n_+} 0 &\implies \langle c - \mathcal{A}^*y, d^*\rangle = 0 \\ &\implies \mathcal{F}^s_P \subseteq \mathcal{S}^n_+ \cap \{d^*\}^\perp = Q\mathcal{S}^{\bar{n}}_+ Q^\top \lhd \mathcal{S}^n_+ \end{aligned}$$

- (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

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

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

Auxiliary Problem

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

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).

$k = 1$ always holds in LP case.

Conclusion Part I

- Minimal representations of the data regularize (P); use min. face f_P (and/or implicit rank reduction)
- ideal goal: a backwards stable preprocessing algorithm to handle (feasible) conic problems for which Slater's CQ (almost) fails

Puzzle?

• Minimal representations are needed for stability in cone optimization.
• But adding redundant constraints to quadratic models before lifting often strengthens SDP relaxation.

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

Instances SDP relaxations of NP-hard comb. opt.

- Quadratic Assignment (Zhao-Karish-Rendl-W.'96)
- Graph partitioning (W.-Zhao'99)
- Strengthened Max-Cut (Anjos-W'02)

Low rank problems

- Systems of polynomial equations (Reid-Wang-W.-Wu'15)
- 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, D-K-V-W'14)

Highly (implicit) degenerate/low-rank problem

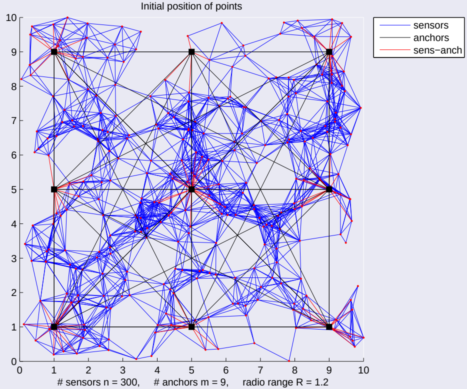
- high (implicit) degeneracy translates to low rank solutions
- take advantage of degeneracy; fast, high accuracy solutions

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; $e = \begin{pmatrix} 1 & \dots & 1 \end{pmatrix}^\top$

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

Euclidean Distance Matrices; Semidefinite Matrices

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

$$\begin{array}{lcl} B \succeq 0 & \Longrightarrow & D = \mathcal{K}(B) = \operatorname{diag}(B)\, e^\top + e \operatorname{diag}(B)^\top - 2B \in \mathcal{E} \\ D \in \mathcal{E} & \Longrightarrow & B = \mathcal{K}^\dagger(D) = -\frac{1}{2} J \operatorname{offDiag}(D)\, J \succeq 0,\ Be = 0 \end{array}$$

Theorem (Schoenberg, 1935)

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

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

And !!!!

$$\boxed{\operatorname{embdim}(D) = \operatorname{rank}\left(\mathcal{K}^\dagger(D)\right), \quad \forall D \in \mathcal{E}^n} \tag{1}$$

Popular Techniques; SDP Relax.; Highly Degen.

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

- $\min_{B\succeq 0} \|H \circ (\mathcal{K}(B) - D)\|$; $\mathrm{rank}\, B = r$;
  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. $= t \leq r < k$

$$\implies \mathrm{rank}\, \mathcal{K}^\dagger(D[\alpha]) = t \leq r \implies \mathrm{rank}\, B[\alpha] \leq \mathrm{rank}\, \mathcal{K}^\dagger(D[\alpha]) + 1$$

$$\implies \mathrm{rank}\, B = \mathrm{rank}\, \mathcal{K}^\dagger(D) \leq n - \boxed{(k-t-1)} \implies$$

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 $\operatorname{embdim}(\bar{D}) = t \leq r$

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

BASIC THEOREM for Single Clique/Facial Reduction

Let:

- $\bar{D} := D[1:k] \in \mathcal{E}^k$, $k < n$, $\operatorname{embdim}(\bar{D}) = t \leq r$ 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:

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

The minimal face

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

Note that the minimal face is defined by the subspace $\mathcal{L}=\mathcal{R}\left(UV\right)$. We add $\frac{1}{\sqrt{k}}e$ to represent $\mathcal{N}\left(\mathcal{K}\right)$; 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

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

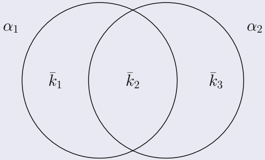
The minimal face is defined by $\mathcal{L} = \mathcal{R}(U)$:

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

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

Sets for Intersecting Cliques/Faces

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



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

  be orthogonal.

Then

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

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:

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

($Q_1 =: (U_1'')^\dagger U_2''$, $Q_2 = (U_2'')^\dagger U_1''$ 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.
- $\boxed{Z := (J\bar{U}_\beta \bar{V})^\dagger B((J\bar{U}_\beta \bar{V})^\dagger)^\top}$.

THEN $t = r$ 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 (from 2010) - 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:

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

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

Noisy SNL Case

200 Sensors; [-0.5,0.5] box; noise 0.05; radio range 0.1

- use sum of exposing vectors rather than intersection of faces obtained from cliques to do facial reduction
- use motivation: roundoff error cancels

show video

Thanks for your attention!

Facial Reduction for Cone Optimization

Henry Wolkowicz
Dept. Combinatorics and Optimization
University of Waterloo

Tuesday Apr. 21, 2015

(with: Drusvyatskiy, Krislock, (Cheung) Voronin)