Taking Advantage of Degeneracy in Cone Optimization with Applications to Low Rank Matrix Approximation

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

at: Dept. Math. & Stat, June 8, 2012

Outline

Part I/Theory: Degeneracy in Cone Optimization

$$\text{minimal representations } \underline{\text{and}} \text{ strong duality } \left.\right\} \begin{array}{c}\text{theoretical,}\\ \text{numerical}\\ \text{difficulties}\end{array}$$

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

Part II: Low Rank Matrix Approximations (SNL)

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

(With: N. Krislock, F. Rendl)

Primal-Dual Pair of Optimization Problems in Conic Form

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

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

where

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

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

Primal-Dual Pair

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

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

$c, A_i \in \mathcal{S}^n, \forall i$ (real, symmetric matrices)

Optimality Conditions

Strong Duality holds if a Constraint Qualification, CQ, holds

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

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, modelling issue

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 $F \triangleleft K$.

Minimal Faces (Intersection of Faces is a Face)

$$f_P := \mathrm{face}\, \mathcal{F}_P^s \trianglelefteq K \qquad f_D := \mathrm{face}\, \mathcal{F}_D^x \trianglelefteq K^*$$

Examples of Faces (self replicating)

LP: $K = \mathbb{R}^n_+, \mathcal{I} \subseteq \{1, \ldots, n\}$

$$f := \{x \in \mathbb{R}^n : x_i = 0, \forall i \in \mathcal{I}\} \trianglelefteq \mathbb{R}^n_+$$

SDP: $K = \mathcal{S}^n_+, P$ is $n \times t$, $\mathrm{Range}(P) \subseteq \mathcal{S}^n_+$

$$f := P\mathcal{S}^t_+ P^T \trianglelefteq \mathcal{S}^n_+$$

(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: $\quad v_D = 1 \quad > \quad v_P = 0$

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: $v_D - v_P = 1 - 0 = 1$,

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

Minimal Face for Ramana Example

Feasible Set/Minimal Face

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

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

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), $f_P = \operatorname{face} \mathcal{F}_P^s$

($\mathbb{P}$) is equivalent to regularized ($\mathbb{P}$);
replace cone K by (smaller cone) minimal face $f_P \subseteq K$.

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

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 (by $x^* \in f_P^* \supseteq K$, larger dual cone)

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

Conclusion Part I

- Minimal Representations of the data regularize (P)
  min. face f_P (and/or the min. L.T. $\mathcal{A}_{PM}$ or $\mathcal{L}^*_{PM}$)
- 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; radiation levels.
- 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 $v_i \rightarrow p_i \in \Re^r$ with squared distances given by ω.

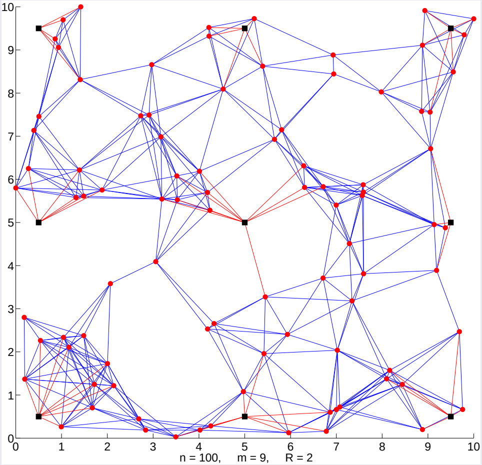
Corresponding Partial Euclidean Distance Matrix, EDM

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

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

Sensor Localization Problem/Partial EDM

Sensors ○ and Anchors ■



Molecular Conformation: $r = 3$, no anchors

Distance Geometry Description

From Experimental data, e.g. NMR spectroscopy

1. a list of distances (lower and upper bounds on the distances between pairs of atoms)
2. chirality constraints (chirality of its rigid quadruples of atoms)

Connections to Semidefinite Programming (SDP)

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

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 $Be = 0$)

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

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

Euclidean Distance, EDM, and Semidefinite, SDP, Matrices

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

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

where $J := I - \frac{1}{n} ee^T$ (orthogonal projection onto $M := \{e\}^\perp$)

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

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

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

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

Properties of Linear Transformations

$\mathcal{K}$, $\mathcal{T}$, Diag, $\mathcal{D}_e$

$$\mathcal{R}(\mathcal{K}) = \mathcal{S}_H; \qquad \underline{\mathcal{N}(\mathcal{K}) = \mathcal{R}(\mathcal{D}_e)};$$

$$\mathcal{R}(\mathcal{K}^*) = \mathcal{R}(\mathcal{T}) = \mathcal{S}_C; \qquad \mathcal{N}(\mathcal{K}^*) = \mathcal{N}(\mathcal{T}) = \mathcal{R}(\mathrm{Diag});$$

$$\mathcal{S}^n = \mathcal{S}_H \oplus \mathcal{R}(\mathrm{Diag}) = \mathcal{S}_C \oplus \mathcal{R}(\mathcal{D}_e).$$

$$\mathcal{T}(\mathcal{E}^n) = \mathcal{S}_+^n \cap \mathcal{S}_C \quad \text{and} \quad \mathcal{K}(\mathcal{S}_+^n \cap \mathcal{S}_C) = \mathcal{E}^n.$$

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

- $\min_{B \succeq 0, B \in \Omega} \|H \circ (\mathcal{K}(B) - D)\|$; $\operatorname{rank} B = r$;
  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

Facial Reduction

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

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

Facial Structure of SDP Cone; Equivalent SUBSPACES

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

$F \trianglelefteq \mathcal{S}^n_+$ 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 $\mathcal{L} = \mathcal{R}(T)$

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

$$F := T\mathcal{S}^t_+ T^T \trianglelefteq \mathcal{S}^n_+, \qquad \operatorname{relint}(F) = T\mathcal{S}^t_{++} T^T$$

Facial Reduction for SDP

Minimal Face

Suppose that the minimal face is:

$$\text{face}\left(\{X \in \mathcal{S}^n_+ : \text{trace}\, A_i X = b_i, i = 1, \ldots, m\}\right) = U\mathcal{S}^t_+ U^T$$

Facially Reduced Program

Then (note that $\text{trace}\, GH = \text{trace}\, HG$):

$$\begin{aligned} p^* &= \min\left\{\text{trace}\, CX : \text{trace}\, A_i X = b_i, i = 1, \ldots m, X \in \mathcal{S}^n_+\right\} \\ &= \min\Big\{\text{trace}\,(U^T C U) Z : \text{trace}\,(U^T A_i U) Z = b_i, i = 1, \ldots m, \\ &\qquad\qquad Z \in \mathcal{S}^t_+\Big\} \end{aligned}$$

and
$X^* = UZ^*U^T$

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 $\alpha = 1:k$; 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

THEOREM 1: Single Clique/Facial Reduction

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

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

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

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

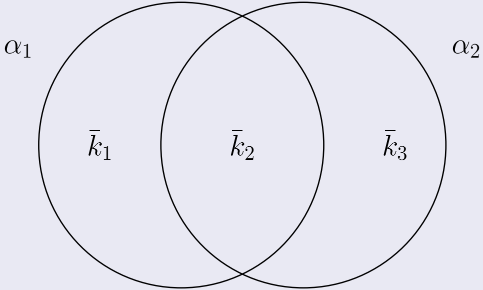
The minimal face is defined by $\mathcal{L} = \mathcal{R}(U)$:

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

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 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

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

cont. . .

Two (Intersecting) Clique Reduction, cont...

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

cont...with

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

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

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

Expense/Work of (Two) Clique/Facial Reductions

Subspace Intersection for Two Intersecting Cliques/Faces

Suppose:

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

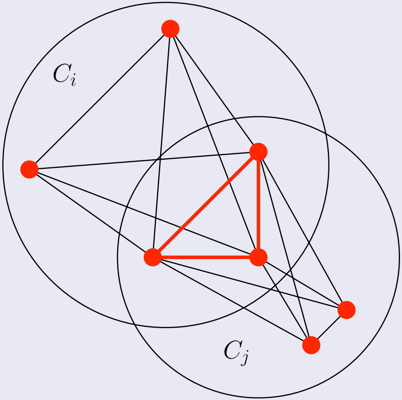
Then:

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

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

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

Two (Intersecting) Clique Reduction Figure



Completion: missing distances can be recovered if desired.

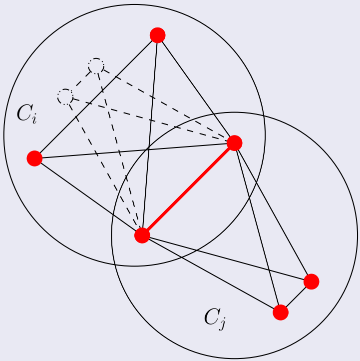
Two (Intersecting) Clique Explicit Delayed Completion

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, <u>THEN</u> $t = r$ 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.

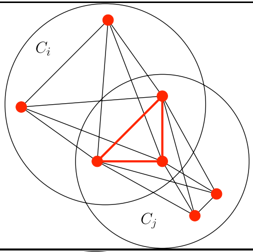
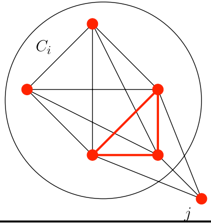
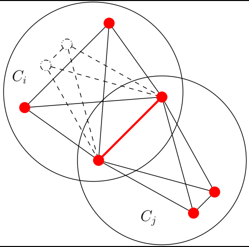
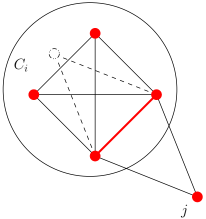
Rotate to Align the Anchor Positions

- 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:

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

  $P_2^T A = U \Sigma V^T$ 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.

$C_i := \{j : (D_p)_{ij} < (R/2)^2\}, \quad$ 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:

$M = \begin{bmatrix} 3,078,915 & 12,315,351 & 27,709,309 & 76,969,790 \end{bmatrix}$

$N = 10^9 \begin{bmatrix} 0.2000 & 0.8000 & 1.8000 & 5.0000 \end{bmatrix}$

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: $Ve = 0, V^T V = I$, $\mathcal{K}_V(X) := \mathcal{K}(VXV^T)$:

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

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

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

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

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

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 Part II

- 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!

Taking Advantage of Degeneracy in Cone Optimization with Applications to Low Rank Matrix Approximation

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

at: Dept. Math. & Stat, June 8, 2012