Semidefinite Programming and Applications to Computationally Hard Optimization Problems

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

CO769/2010

Sensor Network Localization, SNL

SNL, Special Case of EDMC?

- 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, $p_i, i = n - m + 1, \ldots, n$) (e.g. using GPS)
- pairwise distances known within radio range R, $D_{ij} = \|p_i - p_j\|^2, ij \in E$

Current Techniques: Nearest, Weighted, SDP Approx.

$\min_{Y \succeq 0, Y \in \Omega} \|H \circ (\mathcal{K}(Y) - D)\|$
SDP program is: Expensive/low accuracy/implicitly highly degenerate (cliques restrict ranks of feasible Ys)

Applications

Tracking Humans/Animals/Equipment

- monitoring: natural habitat; earthquakes and volcanos; weather and ocean currents.
- military, tracking of goods, vehicle positions, surveillance, (where open-air positioning is not feasible), random deployment in inaccessible terrains or disaster relief operations

sensors everywhere - by 2084!

Lobster migration

Underlying Graph Realization/Partial EDM

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 node $v_i \to p_i \in \mathbb{R}^r$ with squared distances given by ω.

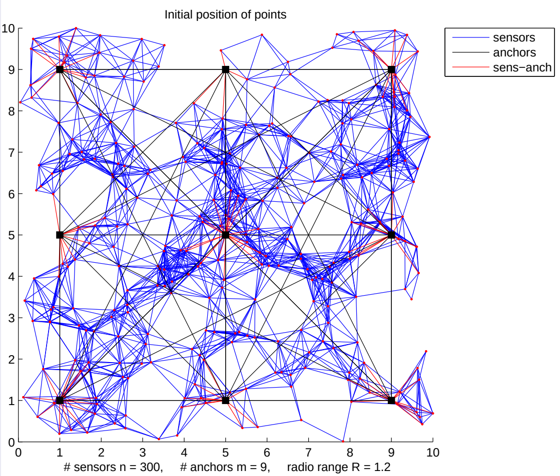
Corresponding Partial Euclidean Distance Matrix, EDM

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

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

Sensor Localization Problem/Partial EDM

Sensors ○ and Anchors ■



SDP: Cone and Faces

$\mathcal{S}^k_+$, Cone of SDP matrices in $\mathcal{S}^k$

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

Faces of cone K

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

Facial Structure of SDP Cone

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

All faces are exposed.

Faces of $\mathcal{S}^k_+$ Equivalence to Subspaces of $\mathbb{R}^k$

- face $F \trianglelefteq \mathcal{S}^k_+$ is determined by range of any, $S \in \operatorname{relint} F$,
  i.e. let $S = U \Gamma U^T$ be compact spectral decomposition, with diagonal matrix of eigenvalues $\Gamma \in \mathcal{S}^t_{++}$, then:
  $$\boxed{F = U \mathcal{S}^t_+ U^T} \qquad (F \text{ associated with } \mathcal{R}(U))$$
- $\dim F = t(t+1)/2.$

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

SNL Connection to SDP (Lin. Trans., Adjoints)

$v = \operatorname{diag}(S) \in \mathbb{R}^n$, $S = \operatorname{Diag}(v) \in \mathcal{S}^n$

$\operatorname{diag} = \operatorname{Diag}^*$ (adjoint)
for $B \in \mathcal{S}^n$, $\operatorname{offDiag}(B) := B - \operatorname{Diag}(\operatorname{diag}(B))$

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

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

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

$\mathcal{K} : \mathcal{S}^n_+ \cap \mathcal{S}_C \to \mathcal{E}^n$

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

Properties of Linear Transformations

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

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

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

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

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

Rank minimization heuristics for the EDM completion problem

heuristic to encourage low rank solution of semidefinite relaxation, Weinberger, 2004

- try to flatten the graph associated with partial EDM ;
- push the nodes of the graph away from each other as much as possible;
- analogy: a loose string on the table can occupy two dimensions, but the same string pulled taut occupies just one dimension.

Low Rank

maximize the objective function

$$\sum_{i,j=1}^{n} \|p_i - p_j\|^2 = e^T \mathcal{K}(PP^T)e, \quad P := \begin{bmatrix} p_1^T \\ \vdots \\ p_n^T \end{bmatrix},$$

s.t. distance constraints holding.

Include centering $P^T e = 0$

$e^T \mathcal{K}(PP^T)e = \langle ee^T, \mathcal{K}(PP^T)\rangle = \langle \mathcal{K}^*(ee^T), PP^T\rangle =$
$\langle 2(\mathrm{Diag}(ee^T e) - ee^T), PP^T\rangle = \langle 2(nI - ee^T), PP^T\rangle =$
$\langle 2nI, PP^T\rangle - \langle ee^T, PP^T\rangle = 2n \cdot \mathrm{trace}(PP^T)$
and $\mathrm{trace}(PP^T) = \sum_{i=1}^{n} \|p_i\|^2$
so, pushing nodes away from each other is equivalent to pushing nodes away from origin

Regularization

regularized semidefinite relaxation of the low-dimensional Euclidean distance matrix completion problem

assume points centered at origin; normalize by dividing by $2n$; substitute $Y = PP^T$:

$$\begin{array}{ll} \text{maximize} & \text{trace}(Y) \\ \text{subject to} & H \circ \mathcal{K}(Y) = H \circ D \\ & Y \in \mathcal{S}^n_+ \cap \mathcal{S}^n_C. \end{array}$$

Nuclear Norm Heuristic for Rank Minimization

nuclear norm of a matrix $X \in \mathbb{R}^{m\times n}$

$\|X\|_* := \sum_{i=1}^{k} \sigma_i(X)$,
$\sigma_i(X)$ is i^{th} largest singular value of X, $\operatorname{rank}(X) = k$.

nuclear norm of symmetric matrix $Y \in \mathcal{S}^n$ is

$\|Y\|_* = \sum_{i=1}^{n} |\lambda_i(Y)|$.

trace min vs max

numerical tests show min (max) yields high (low) rank
alternatively, try max log-det heuristic

Single Clique/Facial Reduction for SNL

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

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.

Basic Theorem for Single Clique/Facial Reduction

THEOREM 1: Single Clique Reduction

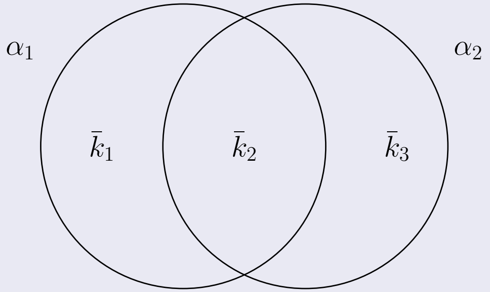
Let $\bar{D} := D[1:k] \in \mathcal{E}^k$, $k < n$, with embedding dimension $t \leq r$, and $B := \mathcal{K}^\dagger(\bar{D}) = \bar{U}_B S \bar{U}_B^T$, where $\bar{U}_B \in \mathcal{M}^{k \times t}$, $\bar{U}_B^T \bar{U}_B = I_t$, and $S \in \mathcal{S}^t_{++}$. Furthermore, let $U_B := \begin{bmatrix}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

Positive Integers for Intersecting Cliques

Two Intersection Sets, α_1, α_2.



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.

Two (Intersecting) Clique Reduction/Subsp. Repres.

THEOREM 2: Clique Intersection Reduction/Subsp. Repres.

$$\alpha_1 := 1:(\bar{k}_1+\bar{k}_2), \quad \alpha_2 := (\bar{k}_1+1):(\bar{k}_1+\bar{k}_2+\bar{k}_3) \quad \subseteq \quad 1:n,$$
$$k_1 := |\alpha_1| = \bar{k}_1+\bar{k}_2, \quad k_2 := |\alpha_2| = \bar{k}_2+\bar{k}_3,$$
$$k := \bar{k}_1+\bar{k}_2+\bar{k}_3.$$

For $i=1,2$, let $\bar{D}_i := D[\alpha_i] \in \mathcal{E}^{k_i}$, with embedding dimension t_i, and $B_i := \mathcal{K}^\dagger(\bar{D}_i) = \bar{U}_i S_i \bar{U}_i^T$, where $\bar{U}_i \in \mathcal{M}^{k_i \times t_i}$, $\bar{U}_i^T \bar{U}_i = I_{t_i}$, and $S_i \in \mathcal{S}^{t_i}_{++}$. Furthermore, for $i=1,2$, let $U_i := \begin{bmatrix} \bar{U}_i & \frac{1}{\sqrt{k_i}} e \end{bmatrix} \in \mathcal{M}^{k_i \times (t_i+1)}$, and let $\bar{U} \in \mathcal{M}^{k \times (t+1)}$ satisfy

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

cont. . .

Two (Intersecting) Clique Reduction, cont...

THEOREM 2 Nonsing. Clique Inters. cont...

cont...with

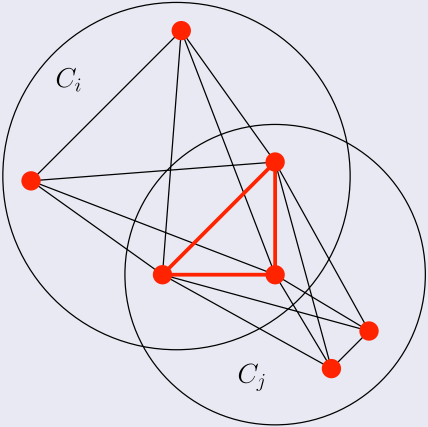
$$\mathcal{R}(\bar{U}) = \mathcal{R}\left(\begin{bmatrix} U_1 & 0 \\ 0 & I_{\bar{k}_3} \end{bmatrix}\right) \cap \mathcal{R}\left(\begin{bmatrix} I_{\bar{k}_1} & 0 \\ 0 & U_2 \end{bmatrix}\right), \text{ with } \bar{U}^T \bar{U} = I_{t+1},$$

let $U := \begin{bmatrix} \bar{U} & 0 \\ 0 & I_{n-k} \end{bmatrix} \in \mathcal{M}^{n \times (n-k+t+1)}$ and let $\begin{bmatrix} V & \frac{U^T e}{\|U^T e\|} \end{bmatrix} \in \mathcal{M}^{n-k+t+1}$ be orthogonal. Then

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

Basic work: find $\bar{U}$ to repres. inters. of 2 subspaces.

Two (Intersecting) Clique Reduction Figure



Completion: missing distances can be recovered if desired.

Two (Intersecting) Clique Explicit Completion

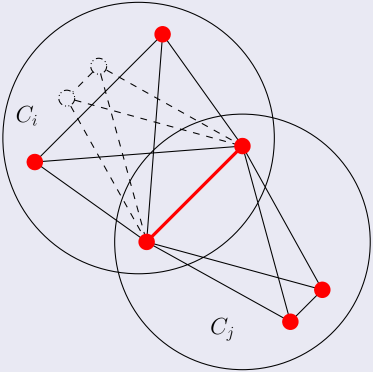
COR. Intersection with Embedding Dim. r/Completion

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

$$\boxed{D[\gamma] = \mathcal{K}\left((\bar{U}\bar{V}) Z (\bar{U}\bar{V})^T\right)}.$$

2 (Inters.) Clique Red. Figure/Singular Case

Two (Intersecting) Clique Reduction Figure/Singular Case



Use *R* as lower bound in singular/nonrigid case.

Two (Inters.) Clique Explicit Compl.; Sing. 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...

2 (Inters.) Clique Expl. Compl.; Degen. 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 $D[\gamma] \in \{\mathcal{K}(\bar{U}\bar{V} Z_i \bar{V}^T \bar{U}^T) : i = 1, 2\}$

Clique Initialization

LEMMA Clique Initialization (by triangle inequality)

For each $i \in \{1, \ldots, n\}$, use half the radio range and define the set

$$C_i := \left\{ j \in \{1, \ldots, n\} : D_{ij} \leq (R/2)^2 \right\}.$$

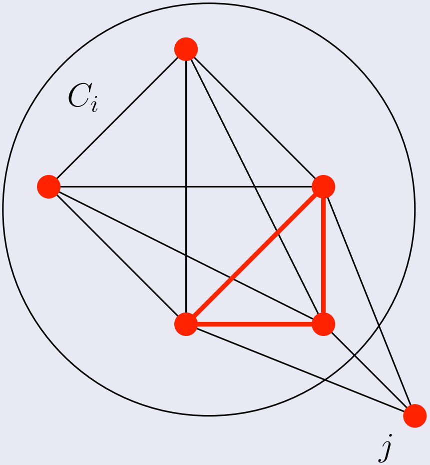
Then each C_i corresponds to a clique of sensors that are within radio range of each other.

Current Set of Cliques, $\{C_i\}_{i \in \mathcal{C}}$

indices $\mathcal{C} \subseteq 1 : n$ corresponding to at most n cliques $\{C_i\}_{i \in \mathcal{C}}$.

Node Absorbtion

Clique C_i Absorbs Node j



Clique Absorption

COROLLARY Clique Absorption

Let C_k, for $k \in \mathcal{C}$, be a given clique with node $l \notin C_k$, $\beta := \{j_1, j_2, j_3\} \subseteq C_k$, and $\max_{i=1,2,3} D_{lj_i} \leq R^2$. If

$$\text{rank } \mathcal{K}^\dagger(D[\beta]) = r,$$

then l can be absorbed by the clique C_k; and, we can complete the missing elements in column (row) l of $D[C_k \cup \{l\}]$.

COROLLARY Clique Absorption-Degenerate

Similar to the degenerate case for the intersection, we can use the lower bound and obtain clique absorption for the case that the embedding dimension is only $r - 1$.

Algorithm

Initialize

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

Algorithm

Initialize

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

Algorithm

Initialize

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

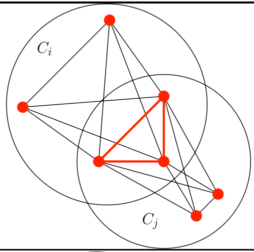
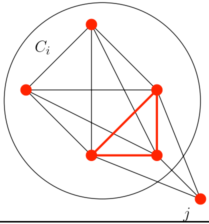
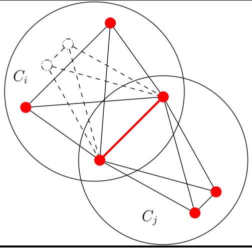
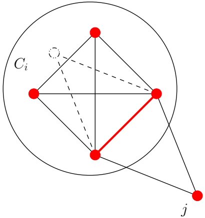
Iterate

- 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

Algorithm

	Clique Union	Node Absorption
Rigid		
Non-rigid		

Results

Rigid Clique Union

n / R	0.7	0.6	0.5	0.4
2000	1	7	91	362
4000	1	1	1	16
6000	1	1	1	1
8000	1	1	1	1
10000	1	1	1	1

Remaining Cliques

n / R	0.7	0.6	0.5	0.4
2000	4.8	4.6	4.2	4.1
4000	9.2	9.4	9.1	9.2
6000	16.0	14.7	15.3	14.9
8000	22.9	22.5	20.9	21.0
10000	38.3	32.7	29.1	30.7

CPU Seconds

n / R	0.7	0.6	0.5	0.4
2000	−10.1	−10.8	—	—
4000	−10.9	−11.0	−10.5	−9.6
6000	−11.6	−10.7	−10.6	−10.0
8000	−11.1	−11.0	−10.7	−9.2
10000	−11.0	−11.0	−10.2	−10.4

Max log(Error)

Rigid Clique Union and Node Absorption

n / R	0.7	0.6	0.5	0.4
2000	1	1	2	78
4000	1	1	1	1
6000	1	1	1	1
8000	1	1	1	1
10000	1	1	1	1

Remaining Cliques

n / R	0.7	0.6	0.5	0.4
2000	4.9	4.9	6.1	13.2
4000	9.2	9.5	9.1	9.8
6000	16.1	15.1	15.1	14.8
8000	22.7	22.4	21.0	21.3
10000	32.5	32.4	28.8	30.6

CPU Seconds

n / R	0.7	0.6	0.5	0.4
2000	−10.1	−10.8	−9.8	−8.8
4000	−10.9	−11.0	−10.5	−9.6
6000	−11.6	−10.7	−10.6	−10.0
8000	−11.1	−11.0	−10.7	−9.2
10000	−11.0	−11.0	−10.2	−10.4

Max log(Error)

EDM Completion Problem, EDMC

- given certain fixed elements of a EDM matrix A
- the other elements are unknown (free)
- complete this matrix to an EDM

$$\mathcal{S} = \{(i,j) : A_{i,j} = \frac{1}{\sqrt{2}} b_k \text{ is known, fixed, } i < j\}, \ |S| = m,$$

$$(\text{EDMC}) \qquad \begin{array}{rcl} \mu^* := & \min & f(X) := \frac{1}{2}\|X\|_F^2 \\ & \text{subject to} & \mathcal{A}(X) = b \\ & & X \succeq 0, \end{array}$$

constraint $\mathcal{A} = \mathcal{I} \cdot \mathcal{L} : \mathcal{S}^{n-1} \to \mathbb{R}^{|S|}$ yields interpolation conditions

$$\mathcal{A}(X)_{ij} = \operatorname{trace} E_{ij} \mathcal{L}(X) = b_k, \quad \forall k \cong (ij) \in S,$$

Duality/Optimality for EDMC

- strict convexity, coercivity implies compact level sets
- EDMC attained and no duality gap (actually primal and dual attainment)

Lagrangian dual

$$\mu^* = \nu^* := \max_{\Lambda \succeq 0, y \in \mathbb{R}^{|S|}} \min_X \frac{1}{2}\|X\|_F^2 + y^T(b - \mathcal{A}(X)) - \operatorname{trace} \Lambda X$$

characterization of optimality

THEOREM Suppose that the feasible set of EDMC is not the empty set. Then the optimal solution of EDMC is $D = \mathcal{L}([\mathcal{A}^*(y)]_+)$, where y is the unique solution of the single equation

$$\mathcal{A}([\mathcal{A}^*(y)]_+) = b,$$

and B_+ denotes the projection of the symmetric matrix $B \in \mathcal{S}^{n-1}$ onto the cone $\mathcal{P}_{n-1}$.

Proof

optimality conditions after differentiation

$$
\begin{array}{rlc}
X = \mathcal{A}^*(y) + \Lambda \succeq 0, \quad \Lambda \succeq 0, & & \text{dual feasibility} \\
\mathcal{A}(X) = b & & \text{primal feasibility} \\
\Lambda X = 0 & & \text{complementary slackness}
\end{array}
$$

This means that $\mathcal{A}^*(y) = X - \Lambda$, where both $X \succeq 0, \Lambda \succeq 0$, and $\Lambda X = 0$. Therefore the three symmetric matrices $W = \mathcal{A}^*(y), X, \Lambda$ are mutually diagonalizable. We write $X = PD_X P^T, \ \Lambda = PD_\Lambda P^T$, i.e. we conclude that $W = \mathcal{A}^*(y) = P(D_X - D_\Lambda)P^T$, $D_X D_\Lambda = 0$. Therefore $[\mathcal{A}^*(y)]_+ = PD_X P^T = X$.
QED

Efficient/Explicit Solution if $y \geq 0$

large class (generic?) can be solved in polytime.

COROLLARY The linear operator $\mathcal{A}$ is onto and $\mathcal{A}\mathcal{A}^*$ is nonsingular. Suppose that $y = (A\mathcal{A}^*)^{-1}b \in \mathbb{R}^m_+$. Then

$$D = \mathcal{L}\left(\mathcal{A}^*(y)\right)$$

is the unique solution of EDMC.
Proof: That $\mathcal{A}$ is onto follows from the definitions.
If $y \geq 0$, then the matrix $\mathcal{I}(y) \geq 0$ with 0 diagonal. Therefore, $X = \mathcal{L}^*(\mathcal{I}(y))$ is diagonally dominant with nonnegative diagonal, i.e. $X \succeq 0$ by Gersgorin's disk theorem. This implies that D is a distance matrix and it satisfies the interpolation conditions, i.e. it satisfies the optimality conditions in the Theorem.
QED

Numerics: dim vs dens with $\#$ of failures in 100 tests

$y = \mathcal{A}^{\dagger} b \geq 0$ may not hold in general, we still get a distance matrix D, i.e. $\mathcal{A}^*(y) \succeq 0$, $n = 10 : 10 : 100$; density $.1 : .1 : .8$.

$n \backslash$density	.1	.2	.3	.4	.5	.6	.7	.8
10	19	27	29	25	32	27	20	38
20	6	20	23	22	27	21	28	28
30	8	8	9	9	11	16	17	24
40	2	2	6	5	14	17	20	17
50	2	0	2	8	7	8	15	12
60	1	1	1	1	3	8	15	11
70	2	0	3	1	5	7	6	15
80	1	0	0	4	2	4	9	9
90	1	0	0	1	3	2	5	6
100	0	0	0	0	1	6	5	5

Summary

- Many hard (combinatorial) problems can be modelled using quadratic objectives and constraints, QQPs.
- QQPs are generally NP-hard problems. But, the Lagrangian relaxation can be solved efficiently using the equivalent SDP (relaxation).
- The special structure of the SDP relaxations can be exploited in order to get efficient solutions for large scale problems.
- Many SDP relaxations of combinatorial problems are degenerate. But, this degeneracy can be exploited. In particular, the SDP relaxation of SNL is highly (implicitly) degenerate. This degeneracy allows for a fast, accurate solution technique.

Thanks for your attention!

Semidefinite Programming
and
Applications to
Computationally Hard Optimization Problems

Henry Wolkowicz

Dept. of Combinatorics and Optimization
University of Waterloo

CO769/2010