Recent progress in Matrix Rank Minimization

Henry Wolkowicz

Parts of this talk represent work with B. Ames, N. Krislock and S. Vavasis.
Dept. of Combinatorics and Optimization
University of Waterloo

AFOSR Optimization & Discrete Math. Program Review
April 19-21, 2010, Arlington, VA

Part I: (Brendan Ames and Stephen Vavasis)

Convex relaxation/Compressive sensing: for the planted clique, biclique and cluster problems

Part II: (Nathan Krislock and Henry Wolkowicz)

Explicit Sensor Network Localization using Semidefinite Representations and Facial Reductions

Outline: Part I

1. Examples of compressive sensing (find sparse solutions using convex relaxations)
2. Algorithm for Maximum clique and biclique (NP hard problems)
3. combinatorial clustering
4. tools: matrix rank minimization using nuclear norm relaxation

Convex relaxation example: compressive sensing

sparsest vector (NP-hard) problem

find $\mathbf{x}$ with fewest number of nonzero entries satisfying $A\mathbf{x} = \mathbf{b}$

Spherical section property:

$\nexists \mathbf{v} \in \mathrm{Null}(A)$ with $\|\mathbf{v}\|_1 / \|\mathbf{v}\|_2$ unusually small value

[Cf. Donoho; Candès, Romberg and Tao; Zhang; others.]

- If A has spherical section property, and $\exists \mathbf{x}^*$ a suff. sparse soluton: then convex relax. $\min\{\|x\|_1 : A\mathbf{x} = \mathbf{b}\}$ yields $\mathbf{x}^*$.
- AND: Only exponentially small subset of $\mathbb{R}^{m \times n}$ fails to have spherical section property; so can choose $A \in \mathbb{R}^{m \times n}$ randomly.

Maximum clique and biclique, NP hard, problems

Clique:

Given an undirected graph (V, E), find k vertices mutually interconnected such that k is maximized

Biclique:

Given a bipartite graph (U, V, E), find a subgraph (U^*, V^*, E^*) containing all possible $|U^*| \cdot |V^*|$ edges such that $|U^*| \cdot |V^*|$ is maximized

They are simple models of information retrieval problems

Results for clique and biclique

convex relaxation of clique or biclique finds exact solution provided input graph is constructed as:

- Start with a single reasonably large clique or biclique.
- Insert additional ‘noise’ edges either chosen by an adversary (arbitrary) or at random.
- The algorithm tolerates a modest number (but the largest possible number, up to constants) adversary-chosen noise edges.
- The algorithm tolerates a much larger number of random noise edges.

Combinatorial clustering problem

NP-hard ($s = 1$ case is classical max clique problem)

Can be posed very generally as follows. Given a graph on n data points, where edges indicate compatibility, find a set of s disjoint cliques that cover as many nodes as possible.

Results

- Suppose the graph consists of s cliques plus some inter-clique noise edges plus additional noise nodes not part of any cliques. Then our convex relaxation will recover the original cliques exactly.
- Many more noise edges and nodes tolerated in the randomized (than in arbitrary) case.

Biclique reformulation as rank minimization

Existence of an mn biclique as rank minimization:

$$\begin{array}{rll} \min & \operatorname{rank}(X) & \\ \text{s.t.} & X(i,j) \in [0,1] & \forall (i,j) \in U \times V \\ & X(i,j) = 0 & \forall (i,j) \in (U \times V) - E \\ & \sum_{(i,j)} X(i,j) \geq mn & \end{array}$$

(Similar formulation exists for clique)

Matrix rank minimization

- Matrix rank minimization is an optimization problem: $\min \operatorname{rank}(X)$ s.t. $X \in C$, $C \subseteq \mathbb{R}^{m \times n}$ convex
- Certain well-known cases solvable efficiently: $C = B(X_0, \delta, \|.\|_2)$ or $C = B(X_0, \delta, \|.\|_F)$,
- In general, the problem is NP-hard. Nonetheless, many very interesting problems are expressed as matrix rank minimization.
  - Sensor network localization
  - Matrix completion problem

Matrix rank min., nuclear norm, compr. sensing

- *Nuclear norm* of X, $\|X\|_*$, is sum of singular vals of X's
- refs:, e.g., Fazel thesis (2002), suggested nuclear norm (convex function) as relaxation of rank.
- Recht, Fazel, Parrilo (2007) showed nuclear norm relaxation is exact for an interesting class of matrix rank minimization problems
- RFP extended compressive sensing properties to rank minimization: If $A \in \mathbb{R}^{m \times n \times p}$ satisfies a certain property, $\hat{X}$ is sufficiently low rank, and $\mathbf{b} = A\hat{X}$, then $\hat{X}$ can be recovered by minimizing $\|X\|_*$ subject to $AX = \mathbf{b}$.
- Nuclear norm minimization can be rewritten as SDP.

Nuclear norm relaxation

Nuclear norm (convex) relaxation of biclique:

$$(NNR) \quad \begin{array}{rll} \min & \|X\|_* & \\ \text{s.t.} & X(i,j) \geq 0 & \forall (i,j) \in U \times V, \\ & X(i,j) = 0 & \forall (i,j) \in (U \times V) - E, \\ & \sum_{(i,j)} X(i,j) \geq mn. & \end{array}$$

Conclusions and open questions, Part I

- Convex relaxation can find a clique or biclique in a graph that contains the clique and biclique plus many diversionary edges.
- If the diversionary edges are placed at random, then the algorithm can tolerate many more of them.
- Would be interesting to extend the technique to other information retrieval problems, e.g., nonnegative matrix factorization.
- Efficient and accurate solvers needed.

Outline Part II

Sensor Network Localization (SNL)/ Facial Reduction

Algorithm based on Euclidean Matrix Completions, EDM and exploiting implicit degeneracy in SDP relaxation.
No SDP solver is used.
anchors ignored

Outline

- Problem Description
- clique union and facial reduction algorithm
- delayed Euclidean Matrix Completion

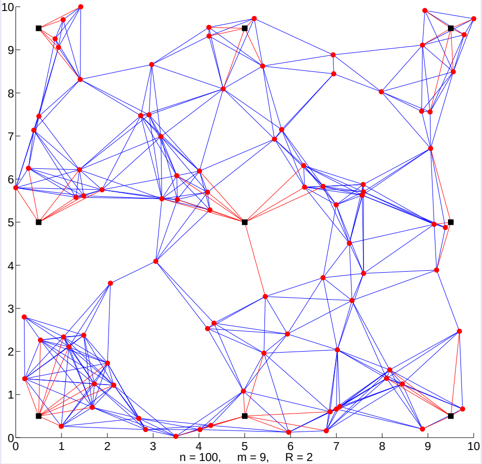
Sensor Network Localization (SNL)/ Facial Reduction

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 $R > 0$
- $$P = \begin{bmatrix} p_1^T \\ \vdots \\ p_n^T \end{bmatrix} = \begin{bmatrix} X \\ A \end{bmatrix} \in \mathbb{R}^{n \times r}$$

Sensor Localization Problem/Partial EDM

Sensors ○ and Anchors ■



Applications

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

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

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

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

Underlying Graph Realization/Partial EDM NP-Hard

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

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

Corresponding Partial Euclidean Distance Matrix, EDM

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

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

Underlying Graph Realization/Partial EDM NP-Hard

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

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

Corresponding Partial Euclidean Distance Matrix, EDM

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

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

Connections to Semidefinite Programming (SDP)

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

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

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

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

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

Connections to Semidefinite Programming (SDP)

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

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

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

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

(to $D \in \mathcal{E}^n$)

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

Current Techniques; SDP Relax.; Highly Degen.

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

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

Instead: (Shall) Take Advantage of Degeneracy!

clique $\alpha, |\alpha| = k$ (corresp. $D[\alpha]$) with embed. dim. $= t \le r < k$

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

$\implies \operatorname{rank} B = \operatorname{rank} \mathcal{K}^\dagger(D) \le n - (k - t - 1) \implies$

Slater's CQ (strict feasibility) fails

Current Techniques; SDP Relax.; Highly Degen.

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

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

Instead: (Shall) Take Advantage of Degeneracy!

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

Facial Structure of SDP Cone; Equivalent SUBSPACES

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

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

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

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

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

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

Facial Structure of SDP Cone; Equivalent SUBSPACES

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

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

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

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

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

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

BASIC THEOREM for Single Clique/Facial Reduction

THEOREM 1: Single Clique/Facial Reduction

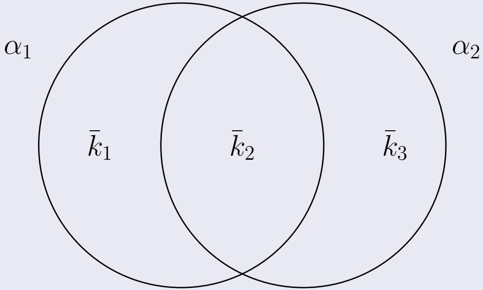
Let: $\bar{D} := D[1{:}k] \in \mathcal{E}^k$, $k < n$, with embedding dimension $t \leq r$; $B := \mathcal{K}^\dagger(\bar{D}) = \bar{U}_B S \bar{U}_B^T$, $\bar{U}_B \in \mathcal{M}^{k\times t}$, $\bar{U}_B^T \bar{U}_B = I_t$, $S \in \mathcal{S}^t_{++}$.
Furthermore, let $U_B := \begin{bmatrix} \bar{U}_B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

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

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

Sets for Intersecting Cliques/Faces

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



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

Two (Intersecting) Clique Reduction/Subsp. Repres.

THEOREM 2: Clique/Facial Intersection Using Subspace Intersection

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

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

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

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

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

cont. . .

Two (Intersecting) Clique Reduction, cont...

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

cont...with

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

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

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

Expense/Work of (Two) Clique/Facial Reductions

Subspace Intersection for Two Intersecting Cliques/Faces

Suppose:

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

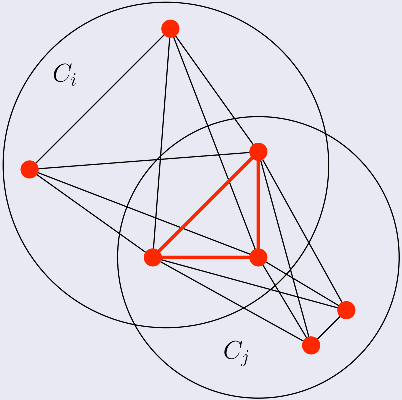
Then:

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

(Efficiently/stably) satisfies:

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

Two (Intersecting) Clique Reduction Figure



Completion: missing distances can be recovered if desired.

Two (Intersecting) Clique Explicit Delayed Completion

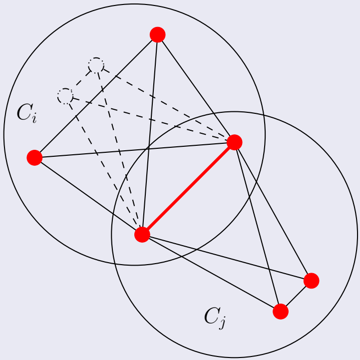
COR. Intersection with Embedding Dim. r/Completion

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

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

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

Two (Intersecting) Clique Reduction Figure/Singular Case



Use R as lower bound in singular/nonrigid case.

Completing SNL (Delayed use of Anchor Locations)

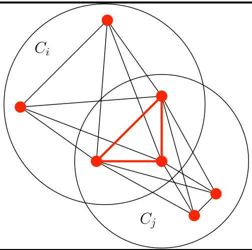
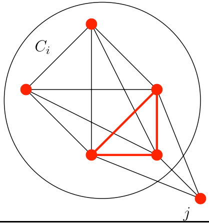
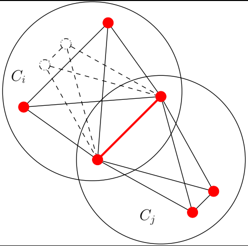
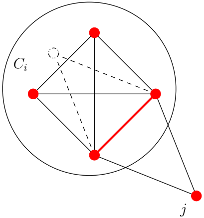
Rotate to Align the Anchor Positions

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

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

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

Algorithm: Four Cases

	Clique Union	Node Absorption
Rigid		
Non-rigid		

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:

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

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

Noisy Case: Locally Recover Exact EDMs

Nearest EDM

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

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

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

Solve Overdetermined Nonlin. Least Squares Prob.

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

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

Derivatives: gradient and Hessian

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

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

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

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

Conclusions, Part II

- exploit: fact that SDP relaxation of SNL is highly (implicitly) degenerate (feasible set of SDP is restricted to low dim. face of SDP cone (low rank matrices))
- take advantage of degeneracy by finding explicit representations of intersections of faces of the SDP cone corresponding to unions of intersecting cliques
- Without any SDP-solver, quickly compute exact solution to SDP relaxation (order of magnitude improvement in: cputime, accuracy, quality)
- anchors are red herring; ignore anchors; delay completion; rotate at end to recover anchor positions.

- P. BISWAS, T.-C. LIAN, T.-C. WANG, and Y. YE, *Semidefinite programming based algorithms for sensor network localization*, ACM Trans. Sen. Netw. **2** (2006), no. 2, 188–220.

- P. BISWAS, T.-C. LIANG, Y. YE, K-C. TOH, and T.-C. WANG, *Semidefinite programming approaches for sensor network localization with noisy distance measurements*, IEEE Transactions onAutomation Science and Engineering **3** (2006), no. 4, 360–371.

- P. BISWAS and Y. YE, *Semidefinite programming for ad hoc wireless sensor network localization*, IPSN ’04: Proceedings of the 3rd international symposium on Information processing in sensor networks (New York, NY, USA), ACM, 2004, pp. 46–54.

- P. BISWAS and Y. YE, *A distributed method for solving semidefinite programs arising from ad hoc wireless sensor network localization*, Multiscale optimization methods and applications, Nonconvex Optim. Appl., vol. 82, Springer, New York, 2006, pp. 69–84. MR MR2191577

- M.W. CARTER, H.H. JIN, M.A. SAUNDERS, and Y. YE, *SpaseLoc: an adaptive subproblem algorithm for scalable wireless sensor network localization*, SIAM J. Optim. **17** (2006), no. 4, 1102–1128. MR MR2274505 (2007j:90005)

- A. CASSIOLI, *Global optimization of highly multimodal problems*, Ph.D. thesis, Universita di Firenze, Dipartimento di sistemi e informatica, Via di S.Marta 3, 50139 Firenze, Italy, 2008.

- K. CHAKRABARTY and S.S. IYENGAR, Springer, London, 2005.

- J. DATTORRO, *Convex optimization & Euclidean distance geometry*, Meboo Publishing, USA, 2005.
- Y. DING, N. KRISLOCK, J. QIAN, and H. WOLKOWICZ, *Sensor network localization, Euclidean distance matrix completions, and graph realization*, Optimization and Engineering **to appear** (2006), no. CORR 2006-23, to appear.
- B. HENDRICKSON, *The molecule problem: Determining conformation from pairwise distances*, Ph.D. thesis, Cornell University, 1990.
- H. JIN, *Scalable sensor localization algorithms for wireless sensor networks*, Ph.D. thesis, Toronto University, Toronto, Ontario, Canada, 2005.

D.S. KIM, *Sensor network localization based on natural phenomena*, Ph.D. thesis, Dept, Electr. Eng. and Comp. Sc., MIT, 2006.

N. KRISLOCK and H. WOLKOWICZ, *Explicit sensor network localization using semidefinite representations and clique reductions*, Tech. Report CORR 2009-04, University of Waterloo, Waterloo, Ontario, 2009, Available at URL: `www.optimization-online.org/DB_HTML/2009/05/2297.html`

S. NAWAZ, *Anchor free localization for ad-hoc wireless sensor networks*, Ph.D. thesis, University of New South Wales, 2008.

T.K. PONG and P. TSENG, *(Robust) edge-based semidefinite programming relaxation of sensor network localization*, Tech. Report Jan-09, University of Washington, Seattle, WA, 2009.

- K. ROMER, *Time synchronization and localization in sensor networks*, Ph.D. thesis, ETH Zurich, 2005.
- S. URABL, *Cooperative localization in wireless sensor networks*, Master's thesis, University of Klagenfurt, Klagenfurt, Austria, 2009.
- Z. WANG, S. ZHENG, S. BOYD, and Y. YE, *Further relaxations of the semidefinite programming approach to sensor network localization*, SIAM J. Optim. **19** (2008), no. 2, 655–673. MR MR2425034

Thanks for your attention!

Recent progress in Matrix Rank Minimization

Henry Wolkowicz

Parts of this talk represent work with B. Ames, N. Krislock and S. Vavasis.
Dept. of Combinatorics and Optimization
University of Waterloo

AFOSR Optimization & Discrete Math. Program Review
April 19-21, 2010, Arlington, VA