On the complexity of discretization vertex orders for distance geometry problems.

PGMO's Days 2013

A. Cassioli[1]

a joint work with

O.Gunluck[4] - C. Lavor[2] - L. Liberti[1,4]

[1]LIX – Ecole Polytechnique - Palaiseau (FR)
[2]IMECC-UNICAMP, Barao Geraldo, Campinas, (BR)
[4]IBM "T.J. Watson" Research Laboratory, Yorktown Heights (USA)

04/10/2013

Outline

1. Distance geometry problems are typically cast as continuous GO problems.

Outline

1. Distance geometry problems are typically cast as continuous GO problems.
2. DGP instances can be represented as graphs.

Outline

1. Distance geometry problems are typically cast as continuous GO problems.
2. DGP instances can be represented as graphs.
3. Under suitable assumptions the feasible set becomes finite and discrete.

Outline

1. Distance geometry problems are typically cast as continuous GO problems.
2. DGP instances can be represented as graphs.
3. Under suitable assumptions the feasible set becomes finite and discrete.
4. A key point of the *discretization* is the availability of a suitable vertex order.

Distance Geometry Problem (DGP)

Given

- $K > 0$;
- a simple, undirected, nonnegatively weighted graph $G = (V, E)$;
- a function $d : E \to \mathbb{R}^+$;

find a realization $x : V \to \mathbb{R}^K$ such that:

$$\forall \{u, v\} \in E \quad \|x_u - x_v\|^2 = d_{uv}.$$

Vertex Order (VO)

Given a graph $G = (V, E)$, a total vertex order can be seen as

- a vertex permutation;
- a function (labeling) $\rho : V \to \mathbb{N}$ that induces a total order on V;

Applications: scheduling, labeling-based algorithm,...

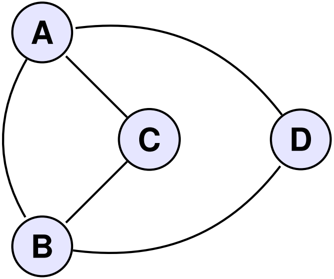
Assumptions and Notation

- we focus on **simple, connected and undirected** graphs;
- **adjacent vertices**: $N_v = \{w \in V : (v, w) \in E\}$
- **predecessor set**: $U_v = \{w \in V : w < v\}$
- we investigate the problem of determining the **existence of discretization vertex orders**

^{K}DDGP Order Definition

A ^{K}DDGP order for $G, K > 0$ is defined as follows:

- $\{v \in V | v \leq K\}$ is a $K-$clique;
- $\forall v \in V$ with $v > K, |U_v \cap N_v| \geq K$.

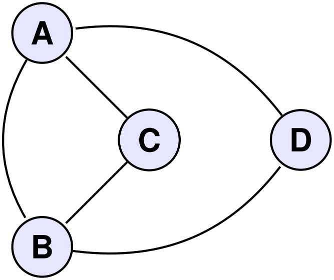


^{K}DDGP Orders for $K = 2$: ABCD, ABDC, ACBD,...

^{K}DMDGP Order Definition

A ^{K}DMDGP order for $G, K > 0$ is defined as follows:

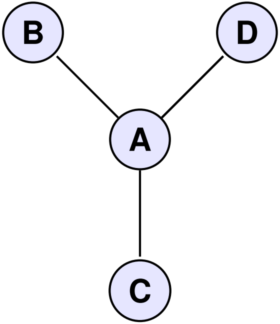
- $\{v \in V | v \leq K\}$ is a $K-$clique;
- $\forall v \in V$ with $v > K, v - K \leq w < v$ then $w \in N_v$.



^{K}DMDGP Orders for $K = 2$: DBAC, CBAD,...

^{K}DDGPO$\equiv{}^{K}$DMDGPO ?

Some graphs are YES for ^{K}DDGPO but NO for ^{K}DMDGPO:



The $S(K, n)$ family

We define for $K > 0$ and $n \leq K + 3$ a family of graphs $S(K, n)$ with n vertices $\{1, 2, \ldots, n\}$ and edges

$$\{\{u, v\} | u < v \leq K\} \cup \{\{u, v\} | u \leq K \vee v > K\}$$

The $S(K, n)$ family

We define for $K > 0$ and $n \leq K + 3$ a family of graphs $S(K, n)$ with n vertices $\{1, 2, \ldots, n\}$ and edges

$$\{\{u, v\} | u < v \leq K\} \cup \{\{u, v\} | u \leq K \vee v > K\}$$

(in the previous slide $S(1, 4)$ is depicted.)

The $S(K, n)$ family

We define for $K > 0$ and $n \leq K + 3$ a family of graphs $S(K, n)$ with n vertices $\{1, 2, \ldots, n\}$ and edges

$$\{\{u, v\} | u < v \leq K\} \cup \{\{u, v\} | u \leq K \vee v > K\}$$

(in the previous slide $S(1, 4)$ is depicted.)

Lemma

For a positive integer K and $n \geq K + 3$, no graph $S(K, n)$ can have a ^{K}DMDGP order.

^{K}DDGPO Properties

- DDGPO is trivially in *NP*;
- for fixed K a polynomial time algorithm exists [1].

[1] Lavor, C., Lee, J., Lee-St. John, A., Liberti, L., Mucherino, A., Sviridenko, M.: Discretization orders for distance geometry problems. Optimization Letters (2012)

^{K}DDGPO Properties

- DDGPO is trivially in *NP*;
- for fixed *K* a polynomial time algorithm exists [1].

Proposition

[1] *DDGPO is equivalent to Spanning-Tree on an undirected simple graph.*

---

[1] Lavor, C., Lee, J., Lee-St. John, A., Liberti, L., Mucherino, A., Sviridenko, M.: Discretization orders for distance geometry problems. Optimization Letters (2012)

^{K}DDGPO Problem

Proposition

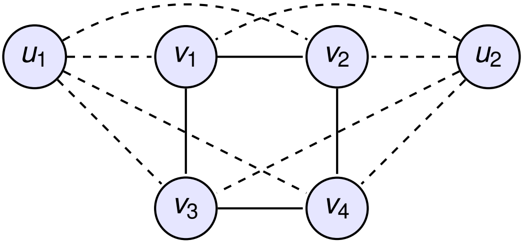
The ^{K}DDGPO problem is **NP***-complete.*

The proof uses a reduction from $K-$Clique problem.

^{K}DDGPO Problem – Reduction

Given $G = (V, E)$ we construct a ^{K}DDGPO instance G' using the following mapping:

- $V' = V \cup U$ with $U = \{u_1, u_2, \ldots, u_K\}, U \cap V = \emptyset$;
- $E' = E \cup \{(u, v) : v \in V \vee u \in U\}$

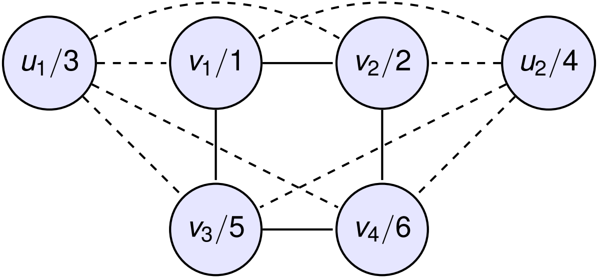


G' for $K = 2$

^{K}DDGPO Problem – Reduction

Let $\{v_1, \ldots, v_K\}$ be a K-clique in G. Then we can construct a ^{K}DDGPO:

1. postponing any permutation of U;
2. adding greedly vertices in $V \setminus \{v_1, \ldots, v_K\}$.



^{K}DDGPO Problem – Reduction

Let assume G is NO for K-clique but YES for ^{K}DDGPO.

1. the initial $K-$clique must contain one vertex w from U and $K-1$ from V.

2. the $K+1th$ vertex z cannot be in U because it is connected to w.
3. z is connected to all the other vertices of the initial K-clique.

Therefore G contains a $K-$clique and we got a contraddiction. Thus the reduction maps YES to YES.

^{K}DMDGPO Problem

- It is trivially in **NP**.

^{K}DMDGPO Problem

- It is trivially in **NP**.
- ^{K}DMDGPO is **NP**-complete.

^{K}DMDGPO Problem

- It is trivially in **NP**.
- ^{K}DMDGPO is **NP**-complete.
- A ^{K}DMDGPO order is symmetric

^{K}DMDGPO Problem

- It is trivially in **NP**.
- ^{K}DMDGPO is **NP**-complete.
- A ^{K}DMDGPO order is symmetric
- A graph G is YES for ^{K}DMDGPO only if it is for ^{K}DDGPO.

^{K}DMDGPO Problem

Proposition

The ^{1}DMDGPO is **NP***-complete.*

Proof.

The definition of ^{1}DMDGPO is equivalent of that HP.

□

^{K}DMDGPO Problem

Theorem

The ^{K}DMDGPO is **NP***-complete for any* $K > 1$

Proof idea:

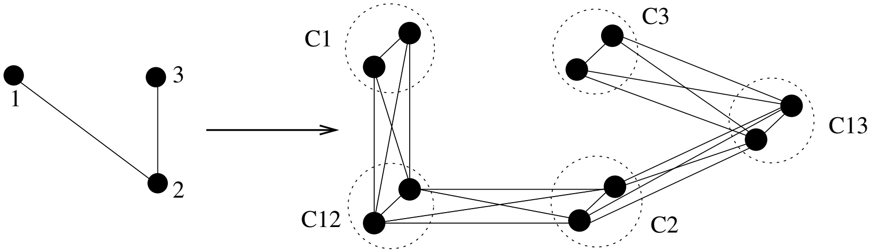
- consider a restriction with given initial $K-$clique (rDMDGPO);
- reduce from Hamiltonian Path with a given starting vertex (rHP);
- repeat for all possible $K-$cliques.

rDMDGPO Problem – Reduction

Given $G = (V, E)$ we construct a rDMDGPO instance G' using the following mapping τ_K:

- a $K-$clique C_v for all $v \in V$
- a $K-$clique C_{ij} for each $(i, j) \in E$
- edges to have bicliques C_{ij}, C_i and C_{ij}, C_j for each $(i, j) \in E$

rDMDGPO Problem – Reduction



Transformation τ_2

rDMDGPO Problem – Preliminary Results (1)

Lemma

Let $(G = (V, E)\,, s)$ be a YES instance of rHP, and $K > 1$. Then $G' = \tau_K(G)$ has a consecutive ^{K}DMDGP order starting with C_s.

Proof.

Let $\beta = \{v_1, \dots, v_{|V|}\}$ a HP in G with $s \equiv v_1$, which corresponds to the clique order

$$\gamma = \{C_1, C_{12}, C_2, \dots\}.$$

Taking $C_i = \{u_{i1}, \dots, u_{iK}\}$ and expanding the cliques we get

$$\alpha = \{u_{11}, \dots, u_{1K}, u_{121}, \dots, u_{12K}, u_{21}, \dots, u_{2K}, \dots\},$$

which is a ^{K}*DMDGP* order in G'.

rDMDGPO Problem – Preliminary Results (2)

Lemma

Let $(G = (V, E)\ , s)$ *be an instance of rHP,* $K > 1$*, and* α *be a* ^{K}*DMDGP order in* $G' = \tau_K(G)$*. Then G is a YES instance for rHP.*

Proof.

If α spans the cliques of G' sequentially then $\tau^{-1}(\alpha)$ is a HP in G. By Lemma .. □

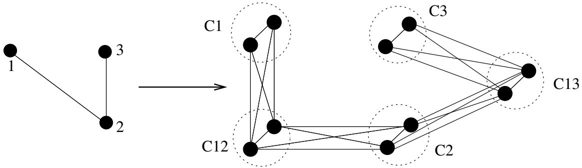
rDMDGPO Problem – Preliminary Results (3)

Lemma

Let C_1, C_2, C_3 be three $K-$cliques in $G' = \tau_K(G)$. There is no vertex triplet $v_1 \in C_1, v_2 \in C_2, v_3 \in C_3$ inducing a triangle in G'.

Proof.

Idea: only cliques of different types are connected on G', i.e. a triangle cannot be *closed*.



^{K}DMDGPO Problem

Theorem

The ^{K}DMDGPO is **NP***-complete for any* $K > 1$

Proof.

By the first two lemmata τ_K maps YES/NO instances to YES/NO instances between rHP and rDMDGPO. Then we can reduce rDMDGPO to ^{K}DMDGPO dropping the initial $K-$clique. □

^{K}DMDGPO ILP Formulation

$$\text{find } x \in \{0,1\}^{|V|\times|V|}$$

$$\sum_{i=1}^{|V|} x_{vi} = 1 \qquad \forall v \in V$$

$$\sum_{v \in V} x_{vi} = 1 \qquad \forall i \in \{1,\ldots,|V|\}$$

$$\sum_{u \in N_v} \sum_{j<i} x_{uj} \geq i x_{vi} \qquad \forall i \in \{1,\ldots,K\}$$

$$\sum_{u \in N_v} \sum_{i-K \leq j < i} x_{uj} \geq K x_{vi} \qquad \forall i > K, \forall v \in V$$

Computational Results (1)

5 random instances for $K \in \{1, 2, 3\}$ with 1 hour cpu time limit with CPLEX 12.5.

$\|V\|$	10%	20%
10	5/0.2/1.676/4.06	5/0.1/1.186/3.22
20	5/20.43/104.578/207.72	5/7.13/116.242/232.02
30	2/7.46/767.415/1527.37	1/2294.5/2294.5/2294.5
40	0/-/-/-	0/-/-/-
50	0/-/-/-	0/-/-/-
60	0/-/-/-	0/-/-/-
70	0/-/-/-	0/-/-/-
80	0/-/-/-	0/-/-/-
90	0/-/-/-	0/-/-/-
100	0/-/-/-	0/-/-/-

Computational Results (2)

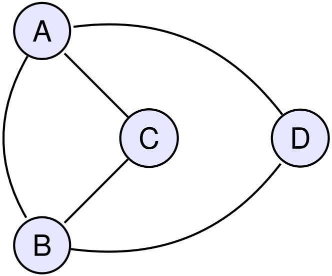
5 random instances for $K \in \{1, 2, 3\}$ with 1 hour cpu time limit with CPLEX 12.5.

$\lvert V \rvert$	50%	90%
10	5/0.02/0.094/0.14	5/0.02/0.022/0.03
20	5 / 4.43/15.998/38.77	5/0.1/0.194/0.3
30	5/109.25/351.612/829.44	5/0.29/0.606/1.03
40	1/1960.15/1960.15/1960.15	5/0.84/1.21/1.96
50	2/2244.54/2403.365/2562.19	5/1.62/4.098/6.54
60	0/-/-/-	5/3.52/8.586/17.87
70	0/-/-/-	5/7.4/17.708/28.75
80	0/-/-/-	5/20.94/31.368/52.05
90	0/-/-/-	5/38.96/57.918/78.25
100	0/-/-/-	5/69.85/132.258/165.43

^{K}DMDGPO ILP Formulation – Extensions (1)

Let π be any total order on V. Then, by the symmetry of any ^{K}DMDGP order it holds that

$$x_{\pi(w)0} \leq x_{\pi(v)|V|} \quad \forall v, w : \pi(v) < \pi(w)$$



$CBAD \equiv DABC$ (π as alphabetical order...)

^{K}DMDGPO Problem

Necessary conditions for G:

- $N_v \geq K$ for all $v \in V$;

^{K}DMDGPO Problem

Necessary conditions for G:

- $N_v \geq K$ for all $v \in V$;
- for $i \in \{K, \ldots, 2K-1\}$ it must hold that

$$|\{v \in V : |N_v| = i\}| \leq 2(i - K + 1);$$

^{K}DMDGPO Problem

Necessary conditions for G:

- $N_v \geq K$ for all $v \in V$;
- for $i \in \{K, \ldots, 2K-1\}$ it must hold that

$$|\{v \in V : |N_v| = i\}| \leq 2(i - K + 1);$$

- a least K $(K+1)$-cliques in G;

^{K}DMDGPO Problem

Necessary conditions for G:

- $N_v \geq K$ for all $v \in V$;
- for $i \in \{K, \ldots, 2K - 1\}$ it must hold that

$$|\{v \in V : |N_v| = i\}| \leq 2(i - K + 1);$$

- a least K $(K + 1)$-cliques in G;
- G YES for the ^{K}DDGPO;

^{K}DMDGPO ILP Formulation – Extensions (2)

Let $\bar{C}_K$ the set of $K-$cliques for which a ^{K}DDGPO can be found. Let $y_c, w_c \in \{0,1\}$ indicates whether the clique c is used to start or end the order, respectively.

$$\sum_{i=1,\dots,K} \sum_{v \in c} x_{vi} \geq K y_c \qquad \forall c \in \bar{C}_K$$

$$\sum_{i=|V|-K,\dots,|V|} \sum_{v \in c} x_{vi} \geq K w_c \qquad \forall c \in \bar{C}_K$$

$$\sum_{c \in \bar{C}_K} y_c = \sum_{c \in \bar{C}_K} w_c = 1$$

$$y_c + w_c \leq 1 \qquad \forall c \in \bar{C}_K$$

^{K}DMDGPO ILP Formulation – Extensions (3)

Assuming $|V| \geq 2K$, we can ask that the first and last cliques to be disjoint. Let

$$\bar{C}_K(l) = \{w \in \bar{C}_k : l \cap w = \emptyset\}.$$

Then the following constraints hold:

$$\sum_{t \in C_K(l)} w_t \geq y_c \qquad \forall c \in \bar{C}_K$$

$$\sum_{t \in C_K(l)} y_t \geq w_c \qquad \forall c \in \bar{C}_K$$

Conclusions

1. we proof ^{K}DDGPO to be **NP**-complete.

Conclusions

1. we proof ^{K}DDGPO to be **NP**-complete.
2. we proof ^{K}DMDGPO to be **NP**-complete for any fixed $K > 0$.

Conclusions

1. we proof ^{K}DDGPO to be **NP**-complete.
2. we proof ^{K}DMDGPO to be **NP**-complete for any fixed $K > 0$.
3. we provide an infinity family of instances that are YES for ^{K}DDGPO but NO for ^{K}DMDGPO.

Future Directions

1. Tests the extended formulation.

Future Directions

1. Tests the extended formulation.
2. Some other orders in between ^{K}DDGPO and ^{K}DMDGPO?

Future Directions

1. Tests the extended formulation.
2. Some other orders in between ^{K}DDGPO and ^{K}DMDGPO?
3. Find good heuristics, if any.

Future Directions

1. Tests the extended formulation.
2. Some other orders in between ^{K}DDGPO and ^{K}DMDGPO?
3. Find good heuristics, if any.

Future Directions

1. Tests the extended formulation.
2. Some other orders in between ^{K}DDGPO and ^{K}DMDGPO?
3. Find good heuristics, if any.

Thank You! Questions??