A new formulation for the unassigned distance geometry problem

Gabriel Braun (UFRJ), Nelson Maculan (UFRJ), Renan Vicente Pinto (UFRRJ),
Carlile Lavor (UNICAMP), Felipe França (UFRJ)

August 09, 2023

Abstract

A new formulation is proposed to solve the unassigned distance geometry problem.

1 Introduction

The unassigned distance geometry problem $(UDGP)$ was defined in [1] and mathematical optimization formulations and heuristis for solving $(UDGP)$ are presented in [2]. Changing the objective function of the proposed formulations in [2], we obtained two new formulations, which will be presented in this work.

2 A formulation presented in [2]

With each vertex v_i it is associated $x_i = (x_{i,1}\ x_{i,2}\ x_{i,3})^\top \in R^3,\ i = 1, 2, ..., n.$

Let the distances $d_k > 0,\ k = 1, 2, ..., m$ be given.

Each d_k is associated with two different vertices $v_i, v_j,\ i < j,$ but we don't know which ones.

Let $||x_i - x_j||_2 = \sqrt{\sum_{l=1}^{3}(x_{i,l} - x_{j,l})^2},\ i \neq j.$

The formulations presented in [2]:

$$(P_0): \quad \min \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (\sum_{k=1}^{m} a_{ij}^k(\ ||x_i - x_j||_2^2\ - d_k^2\ )^2), \tag{1}$$

subject to:

$$\sum_{i=1}^{n-1} \sum_{j=i+1}^{n} a_{ij}^k = 1,\ k = 1, 2, ...m, \tag{2}$$

$$\sum_{k=1}^{m} a_{ij}^{k} \leq 1, \quad i = 1, 2, ..., n-1, \quad j = i+1, i+2, ..., n, \tag{3}$$

$$a_{ij}^{k} \in \{0, 1\}, \; k = 1, 2, ..., m, \; i = 1, 2, .., n-1, \; j = i+1, ..., n, \tag{4}$$

$$x_i \in R^3, \; i = 1, 2, ..., n. \tag{5}$$

Where the binary variable $a_{ij}^{k} = 1$ if the distance d_k is assigned to the pair (v_i, v_j), and $a_{ij}^{k} = 0$ otherwise.

3 A new formulation

We propose the following formulation:

$$(P_1) : \min \; \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} (\sum_{k=1}^{m} a_{ij}^{k} | \; ||x_i - x_j||_2 \; - d_k \; |), \tag{6}$$

subject to: (2 - 5).

We will write (P_1) in another form.

$$(P) : \min \; \sum_{k=1}^{m} y_k, \tag{7}$$

subject to (2 - 5), and

$$y_k \geq \alpha_k, \quad y_k \geq -\alpha_k, \quad k = 1, 2, ..., m, \tag{8}$$

$$t_{ij} \geq 0, \quad t_{ij}^2 = \sum_{l=1}^{3} (x_{i,l} - x_{j,l})^2, \; i = 1, 2, ..., n-1, \; j = i+1, i+2, ..., n, \tag{9}$$

$$-(1-a_{ij}^{k})M + t_{ij} \leq z_{ijk} \leq t_{ij} + (1-a_{ij}^{k})M. \; i = 1, 2, ..., n-1, \; j = i+1, i+2, ..., n, \; k = 1, 2, ..., m, \tag{10}$$

$$-a_{ij}^{k} M \leq z_{ijk} \leq a_{ij}^{k} M, \; i = 1, 2, ..., n-1, \; j = i+1, i+2, ..., n, \; k = 1, 2, ..., m, \tag{11}$$

$$-(1-a_{ij}^{k})M + (d_k + \alpha_k) \leq z_{ijk} \leq (d_k + \alpha_k) + (1-a_{ij}^{k})M. \; i = 1, 2, ..., n-1, \; j = i+1, i+2, ..., n, \; k = 1, 2, ..., m, \tag{12}$$

$$\alpha_k \in R, \; y_k \geq 0, \quad k = 1, 2, ..., m, \tag{13}$$

$$z_{ijk} \geq 0, \; i = 1, 2, ..., n-1, \; j = i+1, i+2, ..., n, \; k = 1, 2, ..., m, \tag{14}$$

Where:

$$M > \max_{k=1,2,...,m} \{d_k\}.$$

In (7), we minimize $\sum_{k=1}^{m} |\alpha_k|$.

If $val(P) = 0$ we obtain a feasible solution.

($val(.)$ is the optimum value of the objective function of problem (.)).

We can consider $x_1 = (0\ 0\ 0)^\top$.

3.1 New ideas

Unfortunately the continuous relaxation of (P) is not convex. Therefore, we will propose a modified model.

$$(PP) : \min \sum_{k=1}^{m} y_k + \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} [t_{ij}^2 - \sum_{l=1}^{3} (x_{i,l} - x_{j,l})^2] \ , \tag{15}$$

subject to (2 - 8) , (10 - 14), and we replace (9) by :

$$t_{ij} \geq 0, \ \ t_{ij}^2 \geq \sum_{l=1}^{3} (x_{i,l} - x_{j,l})^2, \ i = 1, 2, ..., n-1, \ j = i+1, i+2, ..., n. \tag{16}$$

The continuous relaxation of (PP) has a non-convex objective function, and its set of constraints convex.

From (16) we can say that any local optimum of (PP) will imply

$$t_{ij}^2 \ = \ \sum_{l=1}^{3} (x_{i,l} - x_{j,l})^2, \ i = 1, 2, ..., n-1, \ j = i+1, i+2, ..., n. \tag{17}$$

References

[1] P. Duxbury, L. Granlund, S. R. Gujarathi, P. Juhas, and S. J. L. Billinge. The unassigned distance geometry problem. *Discrete Applied Mathematics*, 204:117–132, 2016.

[2] P. Duxbury, C. Lavor, L. Liberti, and L.L. de Salles-Neto. Unassigned distance geometry and molecular conformation problems. *Journal of Global Optimization*, 83:73–82, 2022.