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 heuristics 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 $x_i = (x_{i,1}, x_{i,2}, x_{i,3})^\top \in \mathbb{R}^3$, $i = 1, 2, \ldots, n$.

Let the distances $d_k > 0$, $k = 1, 2, \ldots, m$ 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]:

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

subject to:

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

\[
a_{ij}^k \in \{0, 1\}, \quad k = 1, 2, \ldots, m, \quad i = 1, 2, \ldots, n-1, \quad j = i+1, \ldots, n, \tag{4}
\]

\[
x_i \in \mathbb{R}^3, \quad i = 1, 2, \ldots, n. \tag{5}
\]

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:

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

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

\[
y_k \geq \alpha_k, \quad y_k \geq -\alpha_k, \quad k = 1, 2, \ldots, m, \tag{8}
\]

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

\[
-(1-a_{ij}^k)M + t_{ij} \leq z_{ijk} \leq t_{ij} + (1-a_{ij}^k)M, \quad i = 1, 2, \ldots, n-1, \quad j = i+1, i+2, \ldots, n, \quad k = 1, 2, \ldots, m, \tag{10}
\]

\[
-a_{ij}^k M \leq z_{ijk} \leq a_{ij}^k M, \quad i = 1, 2, \ldots, n-1, \quad j = i+1, i+2, \ldots, n, \quad k = 1, 2, \ldots, m, \tag{11}
\]

\[
-(1-a_{ij}^k)M + (d_k + \alpha_k) \leq z_{ijk} \leq (d_k + \alpha_k) + (1-a_{ij}^k)M, \quad i = 1, 2, \ldots, n-1, \quad j = i+1, i+2, \ldots, n, \quad k = 1, 2, \ldots, m, \tag{12}
\]

\[
\alpha_k \in \mathbb{R}, \quad y_k \geq 0, \quad k = 1, 2, \ldots, m, \tag{13}
\]

\[
z_{ijk} \geq 0, \quad i = 1, 2, \ldots, n-1, \quad j = i+1, i+2, \ldots, n, \quad k = 1, 2, \ldots, m, \tag{14}
\]

Where:

\[
M > \max_{k=1,2,\ldots,m} \{d_k\}.
\]

In (7), we minimize \(\sum_{k=1}^{m} |\alpha_k|\).
If \( \text{val}(P) = 0 \) we obtain a feasible solution.

( \text{val}(\cdot) \) is the optimum value of the objective function of problem \((\cdot)\).)

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.

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

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

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

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

\[
t_{ij}^2 = \sum_{i=1}^{3} (x_{i,l} - x_{j,l})^2, \quad i = 1, 2, ..., n - 1, \quad j = i + 1, i + 2, ..., n.
\]

### References
