Journal of Sciences, Islamic Republic of Iran 21(2): 181-186 (2010) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

A Continuous Optimization Model for Partial Digest Problem

H. Salehi Fathabadi* and R. Nadimi

Department of Applied Mathematics, School of Mathematics, Statistics, and Computer Science, Faculty of Sciences, University of Tehran, Tehran, Islamic Republic of Iran

Received: 26 January2009 / Revised: 16 February 2010 / Accepted: 23 February 2010

Abstract

The pupose of this paper is modeling of Partial Digest Problem (PDP) as a mathematical programming problem. In this paper we present a new viewpoint of PDP. We formulate the PDP as a continuous optimization problem and develope a method to solve this problem. Finally we constract a linear programming model for the problem with an additional constraint. This later model can be solved by the simplex method in which a restricted basis-entry-rule is defined.

Keywords: Molecular biology; Continuous optimization; Simplex method; DNA; PDP

Introduction

One of the interesting tasks in computational biology is Restriction Site Mapping. When a particular restriction enzyme is added to a DNA, the DNA is cut at particular restriction sites. The goal of restriction site mapping is to determine the location of every site for a given enzyme. Using gel electrophoresis, one can find the distance between each pair of restriction sites. In the Partial Digest Problem (PDP), the distances arising from digestion experiments with one enzyme are given, and the locations of all restriction sites must be computed.

Let $X = \{x_0, x_1, ..., x_n\}$ be the set of restriction site locations on a DNA strand. We denote the "multiset" of all $N = \binom{n+1}{2}$ pairwise distances between these site locations by $\Delta X = \{x_j - x_i \mid x_j > x_i, i, j = 0,1,...,n\}$.

Suppose that, in PDP, a multiset $B = \{b_1, b_2, ..., b_N\}$ of distances is given. Our goal is to find a set $Y = \{y_0, y_1, ..., y_n\}$ of points on a line such that B is the pairwise distance multiset of Y. We denote the minimum and maximum of B by b_m and b_M, respectively.

This problem was presented in the 1930's in the area of X-ray crystallography [1]. In 1988 P. Lemke and M. Werman, solved it by a pseudo-polynomial time algorith[2]. The running time of the presented algorithm depends on b_M. Skiena et al[3] suggested a backtracking algorithm to solve the problem where its running time was depended only on n. In 1994, Z. Zhang, by an example, showed that the running time of backtracking algorithm in worst case is exponential [4]. Then, in 2000, T. Dakice in his Ph.D. thesis presented a $0-1$ quadratic programming model for PDP and solved it by a heuristic successive semidefinite programming algorithm [5]. Finally, in 2005, M. Cieliebak et al. proved that Partial Digest Problem is hard to solve for erroneous input data [6].

Modeling this problem using mathematical programming techniqiues, connects it to known and powerful algorithms in the area of mathematical optimization. In this paper we present a continuous

* Corresponding author, Tel.: +98-21-66412178, Fax: +98-21-66412178, E-mail: salehi@khayam.ut.ac.ir

optimization model that can be solved by the well-known simplex method with restricted entry rule for non-basic variables.

In Section 1, we introduce the continuous optimization model for PDP. In section 2 the model is converted to a linear programming problem with an additional set of constraints and dvelop an extended version of the simplex algorithm to solve the problem.

Continuous Optimization Model

Suppose that there are $N = n(n+1)/2$ line segments with lengths of $b_1, b_2, ..., b_N$. We want to place them in a line interval $[0, b_M]$ such that the multiset of endpoints of these line segments equals to $B = \{b_1, b_2, ..., b_N\}$. In other word, we are to provide a solution of PDP with endpoints of line segments. Let a line segment with length b_j be denoted as "b_j-*segment*". It is obvious that the beginning point of b_M-segment is zero and its endpoint is b_M. Let the variables x_j and x_{j+N}, $j = 1, 2, ..., N$ show the beginning and end points of b_j-segment in the interval $[0, b_M]$, i.e. $x_{j+N} - x_j = b_j$ for all $j = 1, 2, ..., N$.

We design an optimization model with $X = \{x_1, x_2, ..., x_{2N}\}$ as the decision variables such that, at optimality, X has exactly $(n+1)$ different values and the multiset of these values is equal to B. We define the new set $\bar{X}$ by eliminating the replicated members of X. (The number of different values in X is equal to the cardinality of $\bar{X}$, $|\bar{X}|$).

Each set of values of x_j's that are between zero and b_M, and satisfy the constraints $x_{j+N} - x_j = b_j$, $j = 1, 2, ..., N$, is defined as a "***placement***" of line segments $b_1, b_2, ..., b_N$ in interval $[0, b_M]$. It is clear that a placement in which the number of its endpoints is not equal to $(n+1)$ is not desirable. Moreover, in the following example, we show that it is possible to place the line segments in interval $[0, b_M]$ with $(n+1)$ endpoints such that the multiset of the endpoints is not equal to B.

Example 1. let $B = \{2,2,2,4,4,4,6,6,8,10\}$ be the input data of PDP. Then we have $N = 10$, $n = 4$ and the target interval is $[0,10]$. Also we have:

$$b_1 = 2, b_2 = 2, b_3 = 2, b_4 = 4, b_5 = 4, b_6 = 4, b_7 = 6, b_8 = 6, b_9 = 8, b_{10} = 10.$$

Now we present two different placements of these line segments in interval [0,10] with $n+1 = 5$ endpoints, such that one of them is a solution of PDP but the other one is not. In the presented placement in the Table 1 (placement(1)), $\bar{X}$ is equal to $\{0,4,6,8,10\}$ and $\Delta\bar{X}$ is equal to B. Therefore, $\bar{X}$ is a solution of PDP. In the placement(2), Table 2, $\bar{X}$ is equal to $\{0,2,4,8,10\}$ but $\Delta\bar{X} = \{2,2,2,4,4,6,6,8,8,10\}$ is not equal to B and ,hence, $\bar{X}$ is not a solution of PDP. □

Review of differences between placement(1) and placement(2) is useful to provide the rule of correct placing. In placement(1) for each $\bar{x}_k$ from $\bar{X}$ there are exactly n=4 members of X equal to $\bar{x}_k$ (see Figure 1), but in placement(2) only x_{13} is equal to 2 and there are more than 4 members of X equal to 4 (see Figure 2).

In placement(2) some $b_j - segments$ coincide with each other, i.e. they have the same beginning and end points. For example b_4, b_5 and b_6 are coincided together. But in placement(1) there is no coincidence.

In the following lemma and theorem it is proved that each placement of B with $(n+1)$ different endpoints that has no coincidence, is a solution of PDP.

Lemma 1. If a placement has no coincidence, then its set of endpoints consists of at least $(n+1)$ different values.

Proof. If a placement X does not have any coincidence, then each pair of members of $\bar{X}$ corresponds at most to one member of B. Suppose that $\bar{X}$ has r members. The number of distinct pairs of the members is greater than or equal to N ($N = |B|$). This means:

Table 1. x_i's values of placement(1)

i	1	2	3	4	5	6	7	8	9	10
b_i	2	2	2	4	4	4	6	6	8	10
x_i	4	6	8	0	4	6	0	4	0	0
x_{i+10}	6	8	10	4	8	10	6	10	8	10

Table 2. x_i's values of placement(2)

i	1	2	3	4	5	6	7	8	9	10
b_i	2	2	2	4	4	4	6	6	8	10
x_i	8	8	0	0	0	0	4	4	0	0
x_{i+10}	10	10	2	4	4	4	10	10	8	10

$$\binom{r}{2} \geq \frac{n(n+1)}{2} = \binom{n+1}{2}$$

Therefore: $r \geq n+1$ □

Theorem 1. Let X be a placement with no coincidence and corresponding $\bar{X}$ has exactly $(n+1)$ members, then $\bar{X}$ is a solution of PDP.

Proof. In a placement with no coincidence each b_j has a unique corresponding pair $(\bar{x}_k, \bar{x}_l)$, so that $\bar{x}_k = x_j$ and $\bar{x}_l = x_{j+N}$. Therefore there are N distinct pair of members of $\bar{X}$ corresponding to members of B. On the other hand, there is only N members in $\Delta\bar{X}$, hence there is no member of $\Delta\bar{X}$ that is not in B. That is $\Delta\bar{X} = B$. □

Now, in oredr to avoid coincidence, we define a set of constraints in placing b_j's. A coincidence occurs when two line segments with the same length have equal beginning and end points. If we consider the following constraint in the placing process, then we have no coincidence in a placement:

$$x_j - x_i \geq b_m \quad if \quad b_j = b_i \quad and \quad j > i \qquad (1)$$

Suppose that X is a placement for $B = \{b_1, b_2, ..., b_N\}$. We use the term "$\varepsilon$-colony" to present a subset of X such that distance between each pair of it's members is less than ε. In other words, the subset $S = \{x_{l_1}, x_{l_2}, ..., x_{l_r}\}$ is an ε-colony if $max\{x_{l_j} - x_{l_i} \mid i, j = 1, 2, ..., r\} < \varepsilon$. It is clear that if $\varepsilon < b_m$, then there is no line segment placed in an ε-colony. Moreover with respect to the constraints 1, if $\varepsilon < \min\{\frac{b_j - b_i}{3} \mid b_j > b_i, i, j = 1, 2, ..., N\}$, then for each pair of ε-colonies S_1 and S_2, there is at most one b_i such that $x_i \in S_1$ and $x_{i+N} \in S_2$. Therefore, immediately, we have the next theorem:

Theorem 2. In a placement with respect to constraints 1, if $\varepsilon < min\{b_m, min\{\frac{b_j - b_i}{3} \mid b_j > b_i, i, j = 1, 2, ..., N\}\}$, then there are at least $(n+1)$ ε-colonies such that the distance between each pair of them is greater than ε. □

Now we constract an optimization model with X as the set of decision variables so that, at optimality, the corresponding $\bar{X}$ is a solution of PDP. We denote the distance between x_i and x_j by z_{ij}, i.e. $z_{ij} = |x_i - x_j|$. To avoid duplication we define z_{ij} only for $j \geq i$.

In order to complete the model we define an objective function and a set of constraints.These constraints and the objective function ensure that there are exactly $(n+1)$ different values for x_j in the optimal solution. Let's define new variables $w_{ij}, i, j = 1, 2, ..., 2N., j \geq i$ which indicate the share of z_{ij} 's in the objective function.

$$w_{ij} = \begin{cases} (\frac{1-\varepsilon}{\varepsilon}) z_{ij} & if \quad z_{ij} \leq \varepsilon, \\ \frac{\varepsilon}{b_M}(z_{ij} - \varepsilon) + (1-\varepsilon) & if \quad z_{ij} > \varepsilon. \end{cases}$$

Figure 3 shows the curve of this function.

In the definition of w_{ij}, ε is a positive real value less than 1. When ε is much smaller than b_M, the gradient of the first segment of the curve is much steeper than the gradient of the second segment.

Using the following constraints, the next lemma shows that adding constraints (2) to the placing process restricts the number of ε-colonies to $(n+1)$.

$$\sum_{j>i} w_{ij} + \sum_{j<i} w_{ji} \leq 2N - n \quad for \quad i = 1, 2, ..., 2N. \qquad (2)$$

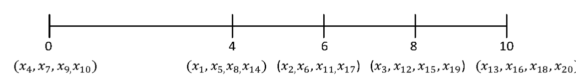

Figure 1. Placement(1).

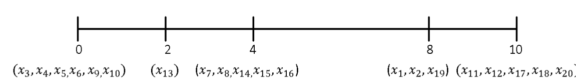

Figure 2. Placement(2).

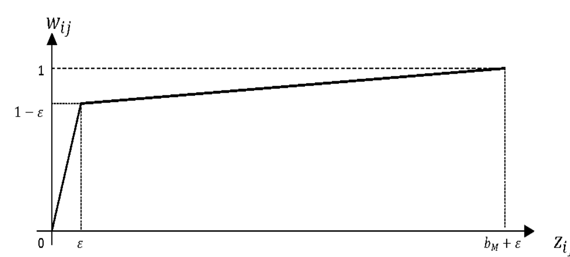

Figure 3. Curve of w_{ij}.

Lemma 2. If we choose $\varepsilon < \frac{1}{2N-n+1}$ in any placement that satisfies constraint (2), then each neighbourhood $N_\varepsilon(x_i)$ has at least n members for any i, where $N_\varepsilon(x_i) = \{x \mid |x - x_i| < \varepsilon\}$.

Proof. Suppose that there exist an index i such that $N_\varepsilon(x_i)$ has less than n members. Then, there are at least $(2N-n+1)$ variables x_j, such that $z_{ij} = |x_i - x_j| > \varepsilon$, and:

$$\sum_{j>i} w_{ij} + \sum_{j<i} w_{ji} > (2N-n+1)(1-\varepsilon)$$

or

$$\sum_{j>i} w_{ij} + \sum_{j<i} w_{ji} > (2N-n)$$

which violate constraints 2. □

With respect to Theorem 2 and Lemma 2, in any placement satisfying constraints (1) and (2) for

$$\varepsilon < min\{\frac{1}{2N-n+1}, min\{\frac{b_j - b_i}{3} \mid b_j > b_i, i,j = 1,2,\ldots,N, b_m, 1\}$$

we have exactly $(n+1)$ ε-colonies such that each of them has exactly n members. (The number of members of X is $2N = n(n+1)$).

Now we can state the complete model as follow:

$$(P) \quad minimize: \; f = \sum_{i<j} w_{ij}$$

subject to:

$$w_{ij} = \begin{cases} (\frac{1-\varepsilon}{\varepsilon}) z_{ij} & if \quad z_{ij} \le \varepsilon, \\ \frac{\varepsilon}{b_M}(z_{ij} - \varepsilon) + (1-\varepsilon) & if \quad z_{ij} > \varepsilon. \end{cases}$$

$z_{ij} = |x_i - x_j|$ for $j > i$, $i,j = 1,2,\ldots,2N$,

$x_{j+N} - x_j = b_j$ for $j = 1,2,\ldots,N$,

$x_j - x_i \ge b_m$ for $b_j = b_i$, $j > i$, $i,j = 1,2,\ldots,2N$,

$\sum_{j>i} w_{ij} + \sum_{j<i} w_{ji} \le 2N - n$ for $i = 1,2,\ldots,2N$,

$0 \le x_j \le b_M$ for $j = 1,2,\ldots,2N$,

$x_N = 0$.

In problem(P), ε is a real number less than:

$$min\{\frac{1}{2N-n+1}, min\{\frac{b_j - b_i}{3} \mid b_j > b_i, i,j = 1,2,\ldots,N\}, \frac{b_M}{b_M + n^3}, b_m\}$$

Remember that the reasons for constraints, $\varepsilon < \frac{1}{2N-n+1}$, $\varepsilon < b_m$ and $\varepsilon < min\{\frac{b_j - b_i}{3} \mid b_j > b_i, i,j = 1,2,\ldots,N\}$ were discussed earlier. In the next Theorem we show that the constraint $\varepsilon < \frac{b_M}{b_M + n^3}$, (which yields $\varepsilon < 1$) implies that an optimal solution of P, is a solution to the corresponding PDP.

Theorem 3. If (X^1, Z^1, W^1) and (X^2, Z^2, W^2) are two feasible solutions of problem P such that $\bar{X}^1$ is a solution to the corresponding PDP and $\bar{X}^2$ is not, then we have:

$$f(X^1, Z^1, W^1) < f(X^2, Z^2, W^2)$$

Proof. Let (X, Z, W) be an arbitrary feasible solution of P. Any feasible solution of P satisfies constraints 1 and 2. Therefore there are exactly $(n+1)$ ε-colonies such that each of them has exactly n members. We denote these ε-colonies by $[x]_i, i = 1,2\ldots,(n+1)$ and their length by $\delta_i, i = 1,2\ldots,(n+1)$.

Each pair $([x]_r, [x]_l)$ $(l > r)$ of these ε-colonies defines a unique b_j in B, so that $x_j \in [x]_r$ and $x_{j+N} \in [x]_l$. Let's $\underline{I}_j(x)$, $\bar{I}_j(x)$ and $I_j(X)$ are defined as follow:

$$\underline{I}_j(x) = \{i \mid |x_i - x_j| < \varepsilon\}, \bar{I}_j(x) = \{i \mid |x_i - x_{j+N}| < \varepsilon\}$$

$$I_j(x) = \{(m,n) \mid m,n \in \underline{I}_j(x) \cup \bar{I}_j(x), z_{mn} > \varepsilon, n > m\}$$

It is clear that $|I_j(x)| = n^2$. Let the length of colonies that include the beginning and end of b_j be denoted by $\underline{\delta}_j$ and $\bar{\delta}_j$ respectively, and define $\hat{\delta}_j$ as follow:

$$\hat{\delta}_j = \underline{\delta}_j + \bar{\delta}_j$$

We split the objective function into f_1 and f_2 with the following properties:

$$f = f_1 + f_2, f_1(X, Z, W) = \sum_{\substack{i<j \\ z_{ij} \geq \varepsilon}} w_{ij}, f_2(X, Z, W) = \sum_{\substack{i<j \\ z_{ij} < \varepsilon}} w_{ij}$$

Now compare the values of the objective function in (X^1, Z^1, W^1) and (X^2, Z^2, W^2). It is clear that for each $(m,n) \in I_j(x^2)$ ($j = 1, 2, \cdots, N$), we have: $z^2_{mn} \geq b_j - \hat{\delta}^2_j$ and for each $(m,n) \in I_j(x^1)$, $z^1_{mn} = b_j$. Therefore:

$$\sum_{(i,j) \in I_j(x^2)} w^2_{ij} \geq \sum_{(i,j) \in I_j(x^1)} w^1_{ij} - n^2 (\frac{\varepsilon}{b_M} \hat{\delta}_j^{\,2})$$

which implies

$$\sum_{j=1}^{N} \sum_{(i,j) \in I_j(x^2)} w^2_{ij} \geq \sum_{j=1}^{N} \sum_{(i,j) \in I_j(x^1)} w^1_{ij} - n^2 (\frac{\varepsilon}{b_M} \sum_{j=1}^{N} \hat{\delta}_j^{\,2})$$

Now we have

$$f_1(X^2, Z^2, W^2) \geq f_1(X^1, Z^1, W^1) - n^2 (\frac{\varepsilon}{b_M} \sum_{j=1}^{N} \hat{\delta}_j^{\,2})$$

On the other hand

$$\sum_{j=1}^{N} \hat{\delta}_j^{\,2} = n \sum_{i=1}^{n+1} \delta_i^2$$

Hence

$$f_1(X^2, Z^2, W^2) \geq f_1(X^1, Z^1, W^1) - \frac{n^3 \varepsilon}{b_M} \sum_{i=1}^{n+1} \delta_i^2 \quad (3)$$

Moreover

$$f_2(X^1, Z^1, W^1) = 0$$

and

$$f_2(X^2, Z^2, W^2) = \frac{1-\varepsilon}{\varepsilon} \sum_{\substack{i<j \\ z^2_{ij} < \varepsilon}} z^2_{ij} \geq \frac{1-\varepsilon}{\varepsilon} \sum_{i=1}^{n+1} \delta_i^2$$

Therefore

$$f_2(X^2, Z^2, W^2) \geq f_2(X^1, Z^1, W^1) + \frac{1-\varepsilon}{\varepsilon} \sum_{i=1}^{n+1} \delta_i^2 \quad (4)$$

With respect to equations 3 and 4 we have

$$f(X^2, Z^2, W^2) \geq f(X^1, Z^1, W^1) - \frac{n^3 \varepsilon}{b_M} \sum_{i=1}^{n+1} \delta_i^2 + \frac{1-\varepsilon}{\varepsilon} \sum_{i=1}^{n+1} \delta_i^2$$

In order to complete the proof, it is sufficient to show that

$$\frac{n^3 \varepsilon}{b_M} < \frac{1-\varepsilon}{\varepsilon}$$

In problem P we assumed that $\varepsilon < \frac{b_M}{b_M + n^3}$, therefore

$$\frac{b_M + n^3}{b_M} < \frac{1}{\varepsilon}$$

or :

$$\frac{n^3}{b_M} < \frac{1-\varepsilon}{\varepsilon}$$

According to $\varepsilon < \frac{1}{2N - n + 1}$, ε is less than 1 and

$$\frac{n^3 \varepsilon}{b_M} < \frac{1-\varepsilon}{\varepsilon} \quad \square$$

Result: If a PDP has a solution, the corresponding model P has a feasible solution and the optimal solution of P is a solution of PDP.

It should be mentioned that if PDP has any solution, it has 2^k (for some integer k) different solutions ,[6], for which the value of the objective function in corresponding P is the same for all of them.

Presenting a Ppiecewise Linear Programming Model

In the previous section we presented a continuous optimization model for PDP. In this model w_{ij} 's are nonlinear(piecewise linear) function of z_{ij} 's. In this section we convert this model to a linear programming problem with an extra constraint. We write w_{ij} 's and z_{ij} 's as convex combinations of the break points in curve of w_{ij} 's.

If $0 \leq z_{ij} \leq \varepsilon$, then there are $\lambda^0_{ij} \geq 0$ and $\lambda^1_{ij} \geq 0$ such that:

$$\begin{cases} \lambda^0_{ij} + \lambda^1_{ij} = 1 \\ z_{ij} = 0\lambda^0_{ij} + \varepsilon\lambda^1_{ij} = \varepsilon\lambda^1_{ij} \\ w_{ij} = 0\lambda^0_{ij} + (1-\varepsilon)\lambda^1_{ij} = (1-\varepsilon)\lambda^1_{ij} \end{cases}$$

and if $\varepsilon \leq z_{ij} \leq \varepsilon + b_M$, then there are $\lambda^1_{ij} \geq 0$ and $\lambda^2_{ij} \geq 0$ so that:

$$\begin{cases} \lambda_{ij}^1 + \lambda_{ij}^2 = 1 \\ z_{ij} = \varepsilon\lambda_{ij}^1 + (b_M + \varepsilon)\lambda_{ij}^2 \\ w_{ij} = (1-\varepsilon)\lambda_{ij}^1 + 1\lambda_{ij}^2 \end{cases}$$

These equations yield:

$$\begin{cases} z_{ij} = 0\lambda_{ij}^0 + \varepsilon\lambda_{ij}^1 + (b_M + \varepsilon)\lambda_{ij}^2 = \varepsilon\lambda_{ij}^1 + (b_M + \varepsilon)\lambda_{ij}^2 \\ w_{ij} = 0\lambda_{ij}^0 + (1-\varepsilon)\lambda_{ij}^1 + 1\lambda_{ij}^2 = (1-\varepsilon)\lambda_{ij}^1 + 1\lambda_{ij}^2 \\ \lambda_{ij}^0 + \lambda_{ij}^1 + \lambda_{ij}^2 = 1 \\ \lambda_{ij}^0\lambda_{ij}^2 = 0 \end{cases}$$

We substitute these results in problem P. Before substitution, consider that with respect to the objective function, the constraint $z_{ij} = |x_i - x_j|$ is equivalent to the following constraints:

$$\begin{cases} z_{ij} + x_j - x_i \geq 0 \\ z_{ij} - x_j + x_i \geq 0 \end{cases}$$

After substitutions, we have a linear programming problem with an additional set of constraints: $\lambda_{ij}^0\lambda_{ij}^2 = 0$. We denote the new problem by P'.

$$(P')\ minimize:\ f = \sum_{i<j}((1-\varepsilon)\lambda_{ij}^1 + \lambda_{ij}^2)$$

subject to:

$$\varepsilon\lambda_{ij}^1 + (b_M + \varepsilon)\lambda_{ij}^2 + x_j - x_i \geq 0 \qquad for \quad j > i,\quad i, j = 1,2,...,2N,$$

$$\varepsilon\lambda_{ij}^1 + (b_M + \varepsilon)\lambda_{ij}^2 - x_j + x_i \geq 0 \qquad for \quad j > i,\quad i, j = 1,2,...,2N,$$

$$x_{j+N} - x_j = b_j \qquad for \quad j = 1,2,...,N,$$

$$x_j - x_i \geq b_m \qquad for \quad b_j = b_i, j > i,\quad i, j = 1,2,...,N,$$

$$\sum_{j>i}((1-\varepsilon)\lambda_{ij}^1 + \lambda_{ij}^2) + \sum_{j<i}((1-\varepsilon)\lambda_{ji}^1 + \lambda_{ji}^2) \leq 2N - n \qquad for \quad i = 1,2,...,2N,$$

$$\lambda_{ij}^0 + \lambda_{ij}^1 + \lambda_{ij}^2 = 1 \qquad for \quad j > i,\quad i, j = 1,2,...,2N,$$

$$\lambda_{ij}^0 . \lambda_{ij}^2 = 0 \qquad for \quad j > i,\quad i, j = 1,2,...,2N,$$

$$\lambda_{ij}^0, \lambda_{ij}^1, \lambda_{ij}^2 \geq 0 \qquad for \quad j > i,\quad i, j = 1,2,...,2N,$$

$$0 \leq x_j \leq b_M \qquad for \quad j = 1,2,...,2N,$$

$$x_N = 0.$$

As in the problem (P), here ε is a real number less than

$$min\{\frac{1}{2N - n + 1}, \frac{b_j - b_i}{3} | b_j > b_i, i, j = 1,2,...,N\}, \frac{b_M}{b_M + n^3}, b_m\}.$$

Problem(P') could be solved by simplex method with *"restricted basis-entry-rule"*. In this rule, λ_{ij}^0 or λ_{ij}^2 is introduced into the basis only if it improves the objective function and if the new basis has only one of λ_{ij}^0 or λ_{ij}^2 [1]. In [1], this method has been used for obtaining an approximating solution to the separable programming.

Results

In this paper our goul was to construct and analyse an exact mathematical model of Partial Digest Problem and propose a proper algorithm to solve it. Hence, we developed a continuous optimization model, which is solvable by the simplex method with restricted basis-entry-rule. Theoretically the running time of the simplex method is exponential but, in practice, the simplex method works surprisingly well and exhibits linear complexity; proportional to $n + m$ where n and m are the number of variables and constraints respectively in the problem. The computational complexity of Partial Digest Problem is an open problem. Neither a proof of NP-hardness nor a polynomial time algorithm is known for this problem. Mathematical programming models have a long history in optimization theory and there are many powerful methods to solve them. We hope this model of PDP, open a new view to this problem.

References

1. Bazaraa M. S., Sherali H. D., and Shetty C. M. *Nonlinear programming theory and algorithms, Second edition*, p:512, John Wiley and Sons, Inc, (1993).
2. Cieliebak M., Eidenbenz S., and Penna P. *Partial Digest is hard to solve for erroneous input data.* Theor. Comput. Sci., **349**(3): 361-381, (2005).
3. Dakic T., *On the Turnpike Problem*. PhD thesis, Simon Fraser University, (2000).
4. Lemke P., Werman M. *On the complexity of inverting the autocorrelation function of a finite integer sequence, and the problem of locating n points on a line, given the* $\binom{n}{2}$ *unlabelled distances between them*. Preprint **453**, Institute for Mathematics and its Application IMA, (1988)
5. Patterson A. L. A direct method for the determination of the components of interatomic distances in crystals, Zeitschr. Krist. **90**: 517-542, (1935).
6. Skiena S. S., Smith W., and Lemke P. *Reconstructing sets from interpoint distances*. In Proc. of the 6th ACM Symposium on Computational Geometry (SoCG 1990), pages 332-339, (1990).
7. Zhang Z., *An exponential example for a partial digest mapping algorithm.* Journal of Computational Biology, **1**(3): 235-239, (1994).