Available online at www.sciencedirect.com

SCIENCE DIRECT®

European Journal of Operational Research 174 (2006) 11–22

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/ejor

Discrete Optimization

A note on duality gap in the simple plant location problem

N. Mladenović [a,*], J. Brimberg [b], P. Hansen [a]

[a] *GERAD and HEC, University of Montreal, 3000, chemin de la Cote-Sainte-Catherine, Montréal, Canada H3T 2A7*
[b] *Royal Military College of Canada, Kingston, Canada K7K 7B4*

Received 5 January 2004; accepted 9 December 2004
Available online 2 April 2005

Abstract

This paper studies the duality gap in the simple plant location problem, and presents general formulas for the gap when certain complementary slackness conditions are satisfied. We show that the duality gap derived by Erlenkotter [A dual-based procedure for uncapacitated facility location, Operations Research 26 (1978) 992–1009], and which has been widely used in the literature, is a special case of the formulas presented here. A counterexample demonstrates that an underlying assumption in Erlenkotter may be violated. The results may be used to obtain improved lower bounds for branch-and-bound algorithms.

Keywords: Simple plant location problem; Duality gap

---

1. Introduction

The simple plant location problem (SPLP), also known as the uncapacitated facility location problem, was first introduced by Balinski [2] in 1965. This model may be classified as one of the most important models in location theory. Numerous articles have been published in the literature that address this problem, e.g., see surveys [6,7,10]. Multi-objective models that embed the SPLP have also been investigated, e.g., see [12] and [4]. One of the most popular solution methods for SPLP was proposed by Erlenkotter in 1978 [8].

Erlenkotter's method is the first dual ascent algorithm proposed in the literature. The method heuristically solves the dual of the LP relaxation of an integer programming formulation of SPLP by exploring

---

[*] Corresponding author. Address: School of Mathematics, University of Birmingham, Edgbaston, Birmingham B15 2TT, United Kingdom. Tel.: +44 121 414 6198; fax: +44 121 414 3389.
E-mail address: n.mladenovic@bham.ac.uk (N. Mladenović).

0377-2217/$ - see front matter
doi:10.1016/j.ejor.2004.12.022

some of the complementary slackness conditions to generate dual solutions and integer solutions to the primal problem. The dual bound provided is not necessarily as good as that obtained by the linear programming method, but the algorithm is fast and the quality of the bound is in general quite high. As a consequence the method can be used in a branch-and-bound algorithm.

In this paper we show that the duality gap in Erlenkotter [8] is derived from certain assumptions on the complementary slackness conditions. Surprisingly, these assumptions have not been further investigated in the literature. Therefore, the primary purpose here is to study the duality gap more generally, and thus, to develop alternate formulas for the gap. There are two advantages of this approach: (i) the alternate dual solutions may provide better estimates of the duality gap, and (ii) some alternate assumptions are able to reduce the size of the dual considerably, making the problem more compatible for solution by general LP solvers. Indeed, Erlenkotter's formula is only one special case of the formulas presented here.

In the next section, we briefly review the model formulation for the primal and dual problems. This is followed by an analysis of the duality gap (Section 3) where various formulas are developed related to feasibility and complementary slackness conditions. This leads to a set of necessary and sufficient conditions for the gap to be zero. In Section 4, a numerical example is given to illustrate the advantages of our alternate approach, also that Erlenkotter's method may lead to poor estimates of the duality gap. The last section presents a brief summary, and possible directions for future research.

2. Mathematical model

Let us denote by I a set of potential facilities ($i = 1, \ldots, m$) and J a set of users or customers ($j = 1, \ldots, n$). The simple plant location problem (SPLP for short) may be written as

$$\min_{x,y} \; z_P = \sum_{i=1}^{m} f_i y_i + \sum_{i=1}^{m} \sum_{j=1}^{n} c_{ij} x_{ij}, \tag{1}$$

s.t.

$$\sum_{i=1}^{m} x_{ij} = 1, \quad \forall j \in J, \tag{2}$$

$$x_{ij} - y_i \leqslant 0, \quad \forall i \in I, \; \forall j \in J, \tag{3}$$

$$y_i \in \{0, 1\}, \quad \forall i \in I, \tag{4}$$

$$x_{ij} \geqslant 0, \quad \forall i \in I, \; \forall j \in J, \tag{5}$$

where f_i denotes the fixed cost for opening facility i, c_{ij} is the distribution cost for satisfying the demand of user j from facility i, y_i is a Boolean variable equal to 1 if facility i is opened, and 0 otherwise, x_{ij} is the fraction of demand of user j satisfied from facility i.

This problem has $mn + m$ variables and $mn + n$ constraints. The objective function expresses that the sum of fixed costs for open facilities and distribution costs from (one of) these facilities to each user must be minimized. The f_i and c_{ij} are assumed to be positive. Once the y_i are fixed, the values of the x_{ij} are easily determined; for each j, x_{ij} is chosen as equal to 1 for some i with $y_i = 1$ and minimum distribution cost c_{ij} (ties may be broken arbitrarily).

If constraints (4) are relaxed to

$$0 \leqslant y_i \leqslant 1, \quad \forall i \in I, \tag{6}$$

one gets the *strong linear programming relaxation*. Another relaxation, called the *weak linear programming relaxation*, may be obtained by replacing constraints (3) by

$$\sum_{j=1}^{n} x_{ij} \leqslant n y_i, \quad \forall i \in I, \tag{7}$$

so that the number of constraints is only $m + n$ instead of $mn + n$. But this latter relaxation is less tight than the strong one, and does not appear to be of interest. For many types of practical instances, it is known (e.g., see [13,4]) that the strong LP reduction is *integer-friendly*, i.e., most variables y_i take an integer value in the optimal vector; the duality gap is also small.

A particular class of test instances has been studied in depth [3] and often used in empirical studies: potential locations for facilities are the same as the given locations of users, and are points taken from a uniform distribution on the unit square. Distribution costs c_{ij} are equal to Euclidean distances between i and j; fixed costs are equal for all facilities, and set at $\sqrt{n}/10$, $\sqrt{n}/100$ and $\sqrt{n}/1000$. For Euclidean instances the relative duality gap is known to be small, i.e., 0.00189 almost surely (see [11,1]). It is proved that any branch-and-bound algorithm using the strong relaxation (without further cutting planes) will require a number of branches which increases exponentially with m (or n). Nevertheless near optimal solutions may be readily obtained for fairly large instances. For example, Barahona and Chudak [3] recently solved with an error of at most 1% instances with $m = n$ up to 3000.

The dual of the *strong linear programming relaxation* is

$$\max_{v,w,t} \left(\sum_{j=1}^{n} v_j - \sum_{i=1}^{m} t_i \right), \tag{8}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} w_{ij} - t_i \leqslant f_i, \quad \forall i \in I, \tag{9}$$

$$v_j - w_{ij} \leqslant c_{ij}, \quad \forall i \in I, \ \forall j \in J, \tag{10}$$

$$t_i, \quad w_{ij} \geqslant 0, \quad \forall i \in I, \ \forall j \in J. \tag{11}$$

Note that the variables v_j are not restricted in sign. However, since the equality sign in (2) may be replaced by $\geqslant$ without affecting the optimal solution, the v_j must be nonnegative. This problem has $mn + m + n$ variables and $mn + m$ constraints. As in the primal, it is large in both dimensions. Fortunately, the dual may be simplified in various ways. First one may observe that the variables t_i each appear only in the objective function, with negative sign, and in a single constraint in (9). Using (9) and (11), we have

$$t_i = \max\left\{ \sum_{j=1}^{n} w_{ij} - f_i, 0 \right\} = \left(\sum_{j=1}^{n} w_{ij} - f_i \right)^{+}, \tag{12}$$

where $a^{+} = \max\{a, 0\}$. Reducing one t_i by an amount Δt_i allows decreasing $\sum_{j \in J} w_{ij}$ and hence $\sum_{j \in J} v_j$ by the same amount. Hence, the t_i can be reduced one at a time without reducing the objective function's value. From here it is clear that all variables, t_i, can be fixed at zero, and we get the simpler LP formulation of the dual that usually appears in the literature:

$$\max_{v,w} \quad z_D = \sum_{j=1}^{n} v_j, \tag{13}$$

s.t.

$$\sum_{j=1}^{n} w_{ij} \leqslant f_i, \quad \forall i \in I, \tag{14}$$

$$v_j - w_{ij} \leqslant c_{ij}, \quad \forall i \in I, \ \forall j \in J, \tag{15}$$

$$w_{ij} \geqslant 0, \quad \forall i \in I, \ \forall j \in J. \tag{16}$$

Erlenkotter [8] has observed that for any fixed v_j's, the w_{ij} may be reduced without affecting feasibility; that is the w_{ij} may be made as small as possible. Constraints (14)–(16) imply that we can set [8]

$$w_{ij} = \max\{v_j - c_{ij}, 0\} = (v_j - c_{ij})^+, \quad \forall i, j. \tag{17}$$

Now substitute (17) into (14) to get a nonlinear programming formulation in n variables with m constraints, known as the *restricted dual*:

$$\max_v \sum_{j=1}^{n} v_j, \tag{18}$$

s.t.

$$\sum_{j=1}^{n} (v_j - c_{ij})^+ \leqslant f_i, \quad \forall i \in I. \tag{19}$$

3. Duality gap

Throughout the paper, we always assume that some feasible primal integer solution (y, x) is given. The dual solution may or may not be feasible as noted. The following notation is used:

$$I^+ = \{i | y_i = 1, i \in I\}, \tag{20}$$

$$J_i = \{j | x_{ij} = 1\}, \quad i \in I^+, \tag{21}$$

$$c_j^{(1)} = \min_{i \in I^+} c_{ij}, \quad j \in J. \tag{22}$$

Therefore, I^+ is the index set of open facilities, J_i is the index set of users served by facility i ($i \in I^+$), and $c_j^{(1)}$ denotes the distribution cost from the closest open facility for the user j. It follows that without loss of generality we may restrict the set of feasible primal solutions (y, x) to only those where customers are served by the closest open facility with ties broken arbitrarily. Thus, the user subsets J_i are uniquely defined for any given I^+. Assuming that two or more facility sites are opened ($|I^+| \geqslant 2$), let us further denote as i_j^+ the index of the closest open facility to user j i.e., $i_j^+ = i$, for all $j \in J_i$, $i \in I^+$ ($c_j^{(1)} = c_{i_j^+, j}$, for all $j \in J$), and let $c_j^{(2)}$ be the distribution cost from the second closest open facility of the user j, i.e.,

$$c_j^{(2)} = \min_{i \in I^+,\ i \neq i_j^+} c_{ij}; \quad \left(c_j^{(2)} = c_{i_j^{++}, j}\right). \tag{23}$$

We now derive some properties that hold for any primal–dual pair (y, x) and (v, w), after which we shall assume that some complementary slackness conditions (related to the strong LP relaxation and its dual) are satisfied. This is motivated by the fact that the dual problem may be very large, and hence, the duality gap may be difficult to determine accurately. Determining the exact duality gap is in fact NP-complete [9]. The assumptions allow us to reduce the size of the dual (by setting a number of variables to zero), and to estimate the duality gap in different ways.

In the remainder of the text we shall use the following relations:

$$\sum_{j=1}^{n} v_j = \sum_{i \in I^+} \sum_{j \in J_i} v_j, \tag{24}$$

$$(a - b) = (a - b)^+ - (b - a)^+, \tag{25}$$

$$(i \in I^+, j \in J_i) \Rightarrow (v_j - c_{ij})^+ = (v_j - c_j^{(1)})^+. \tag{26}$$

Equalities (24) and (25) are obvious. The last one holds because customers are assigned to the closest open facility, i.e., $i_j^+ = i$, $\forall j \in J_i$.

Proposition 1. *For any (feasible or infeasible) solution of the dual,*

$$z_P - z_D = \sum_{i \in I^+} \left(f_i - \sum_{j \in J_i} \left(v_j - c_j^{(1)} \right)^+ + \sum_{j \in J_i} \left(c_j^{(1)} - v_j \right)^+ \right).$$

Proof. Using the definitions of the primal (1) and dual (13), and relation (24), the difference between any integer feasible primal and any dual solution is

$$z_P - z_D = \sum_{i \in I^+} f_i + \sum_{i \in I^+} \sum_{j \in J_i} c_{ij} - \sum_{i \in I^+} \sum_{j \in J_i} v_j.$$

Combining with (26), it follows that

$$z_P - z_D = \sum_{i \in I^+} \left(f_i - \sum_{j \in J_i} \left(v_j - c_j^{(1)} \right) \right). \tag{27}$$

Substitution of $(v_j - c_j^{(1)})$ with (25) then gives the result. □

Alternately, Proposition 1 may be presented as

$$\text{gap} = \sum_{i \in I^+} \text{gap}_i,$$

where

$$\text{gap}_i = f_i - \sum_{j \in J_i} \left(v_j - c_j^{(1)} \right)^+ + \sum_{j \in J_i} \left(c_j^{(1)} - v_j \right)^+, \quad \forall i \in I^+. \tag{28}$$

Hence, gap is an additive (or separable) function, where each component is related to an open facility.

Proposition 2. *For any feasible dual solution (i.e., satisfying (14) and (17)),*

$$\sum_{j \in J_i} \left(c_j^{(1)} - v_j \right)^+ + \sum_{j \notin J_i} (v_j - c_{ij})^+ \leqslant \text{gap}_i, \quad \forall i \in I^+.$$

Proof. From the feasibility of the dual solution (14) and the definition of w_{ij} (17), we have

$$\sum_{j=1}^{n} (v_j - c_{ij})^+ - f_i = \sum_{j \in J_i} (v_j - c_{ij})^+ + \sum_{j \notin J_i} (v_j - c_{ij})^+ - f_i \leqslant 0.$$

If $i \in I^+$ the last relation may be rewritten as

$$\sum_{j \in J_i} \left(v_j - c_j^{(1)} \right)^+ - f_i - \sum_{j \in J_i} \left(c_j^{(1)} - v_j \right)^+ + \sum_{j \in J_i} \left(c_j^{(1)} - v_j \right)^+ + \sum_{j \notin J_i} (v_j - c_{ij})^+ \leqslant 0.$$

Since the first three terms add to gap_i (28), the result follows. □

By Proposition 2 a lower bound of each component of the duality gap is given, if the dual solution is feasible. This result will be used later when deriving different duality gap formulas.

We now derive properties that hold for specific primal–dual pairs that satisfy certain (one or more) complementary slackness conditions of the strong LP relaxation. It is well-known from the *strong duality theorem* that $z_P^* = z_D^*$ if and only if all complementary slackness conditions are met:

$$v_j\left(\sum_{i=1}^{m} x_{ij} - 1\right) = 0, \quad \forall j \in J, \tag{29}$$

$$w_{ij}(y_i - x_{ij}) = 0, \quad \forall i \in I, \ \forall j \in J, \tag{30}$$

$$\left(f_i - \sum_{j=1}^{n} w_{ij}\right) y_i = 0, \quad \forall i \in I, \tag{31}$$

$$(c_{ij} - v_j + w_{ij})x_{ij} = 0, \quad \forall i \in I, \ \forall j \in J, \tag{32}$$

where (y, x) and (v, w) denote respectively primal and dual solutions.

Note that (29) is automatically satisfied by the feasible primal solution (see constraint (2)). However, (30)–(32) are not necessarily true if we add integrality constraints on the primal variables y_i, and that is the source of the duality gap $(z_P - z_D)$. In the next result we assume that the dual is restricted by complementary slackness condition (31); i.e., since $y_i = 1$, $\forall i \in I^+$,

$$\sum_{j=1}^{n} w'_{ij} = f_i, \quad \forall i \in I^+, \tag{33}$$

for some dual solution $z'_D = z_D(v', w')$.

Proposition 3. *For any feasible dual (v', w') satisfying (31), the duality gap is given by*

$$z_P - z'_D = \sum_{i \in I^+}\left(\sum_{j \in J_i}\left(c_j^{(1)} - v'_j\right)^+ + \sum_{j \notin J_i}(v'_j - c_{ij})^+\right).$$

Proof. Follows immediately from the proof in Proposition 2, where the "$\leqslant$" is replaced by "$=$" for $i \in I^+$ as a result of (33). In other words, the proof of this claim is the same as the proof of the previous one. The only difference is that the condition (14) in the proof of the previous proposition is replaced by condition (33). □

In the above proposition a duality gap formula is given for any feasible primal dual pair that satisfies (31). The gap is seen to be positive if there is a dual variable v_j either smaller than $c_j^{(1)}$ or larger than $c_j^{(2)}$.

The next two propositions give equivalent but simpler conditions to (32) and (30). We turn first to complementary slackness condition (32).

Proposition 4. *Any dual solution (v'', w''), where w'' is given by (17), satisfies complementary slackness condition (32) if, and only if,*

$$v''_j \geqslant c_j^{(1)}, \quad \forall j \in J.$$

Proof. Since $x_{ij} = 1$, $\forall i \in I^+, j \in J_i$, it follows that if (32) is satisfied

$$c_j^{(1)} - v''_j + w''_{ij} = 0 \Rightarrow v''_j - c_j^{(1)} = w''_{ij} \geqslant 0 \Rightarrow v''_j \geqslant c_j^{(1)}, \quad \forall j \in J.$$

In the reverse direction, if $v''_j \geqslant c_j^{(1)}$, $\forall j \in J$, and $w''_{ij} = (v''_j - c_{ij})^+$, $\forall i, j$, then $w''_{ij} = (v''_j - c_j^{(1)})^+ = v''_j - c_j^{(1)}$, $\forall i \in I^+, j \in J_i$, and thus, (32) is satisfied. □

Let us now denote by $(\bar{v}, \bar{w})$ and $\bar{z}_D$ a *feasible* set of dual variables and objective function value satisfying complementary slackness condition (30). Several nice properties follow from this assumption.

Proposition 5. *Any dual solution* $(\bar{v}, \bar{w})$, *where* $\bar{w}$ *is given by* (17), *satisfies complementary slackness condition* (30) *if, and only if,*

$$\bar{v}_j \leqslant c_j^{(2)}, \quad \forall j \in J.$$

Proof. If $y_i = 1$, $x_{ij} = 0$, and (30) is true, then

$$\bar{w}_{ij} = (\bar{v}_j - c_{ij})^+ = 0, \quad i \in I^+, \quad \forall j \notin J_i \Rightarrow \bar{v}_j \leqslant c_{ij}, \quad \forall j \in J, \quad i \in I^+ \setminus \{i_j^+\} \Rightarrow \bar{v}_j \leqslant c_j^{(2)}, \quad \forall j \in J.$$

In the reverse direction, if $\bar{v}_j \leqslant c_j^{(2)}$, $\forall j \in J$, and $\bar{w}_{ij} = (\bar{v}_j - c_{ij})^+$, it clearly follows that $\bar{w}_{ij} = 0$, $\forall i \in I^+$, $j \notin J^i$, and hence, (30) must be satisfied. □

Proposition 6. *If* $(\bar{v}, \bar{w})$ *is feasible and* $\bar{v}_j \leqslant c_j^{(2)}$, $\forall j \in J$, *then*

$$\sum_{j \in J_i} (\bar{v}_j - c_j^{(1)})^+ \leqslant f_i, \quad \forall i \in I^+.$$

Proof. Noting that (30) is satisfied, and employing (14) we may infer that

$$\sum_{j=1}^{n} \bar{w}_{ij} \leqslant f_i \Rightarrow \sum_{j \in J_i} \bar{w}_{ij} + \sum_{j \notin J_i} \bar{w}_{ij} = \sum_{j \in J_i} \bar{w}_{ij} \leqslant f_i, \quad \forall i \in I^+,$$

furthermore, from (17) and (26), we have $\bar{w}_{ij} = (\bar{v}_j - c_j^{(1)})^+$, $\forall i \in I^+, j \in J_i$. □

Note that strict equality holds in the previous proposition if additionally the $\bar{w}_{ij}$ satisfy (31).

We now comment on the duality gap formula derived by Erlenkotter [8], where an approximate dual solution v^+ is first found (the w_{ij}^+ are then determined by (17)), and a corresponding primal solution y^+ is defined such that $y_i^+ = 1$ if $i \in I^+$ and 0 otherwise, where

$$I^+ = \left\{ i | \sum_{j=1}^{n} (v_j^+ - c_{ij})^+ = f_i \right\}.$$

The idea behind this connection between dual and primal is obtained from the complementary slackness condition (31). Erlenkotter also assumes that $v_j^+ \geqslant c_j^{(1)}$ (or $(c_j^{(1)} - v_j^+)^+ = 0$), $\forall j$, i.e., complementary slackness condition (32) is satisfied. However, for such defined dual and primal solutions, the duality gap is a special case of Proposition 3.

Proposition 7 (Erlenkotter gap). *For any feasible dual* (v^+, w^+) *satisfying complementary slackness conditions* (31) *and* (32),

$$z_P - z_D^+ = \sum_{i \in I^+} \sum_{j \notin J_i} (v_j^+ - c_{ij})^+.$$

The other special case of Proposition 3 is given by the following result.

Proposition 8. *For any feasible dual* $(\tilde{v}, \tilde{w})$ *satisfying* (30) *and* (31),

$$z_P - \tilde{z}_D = \sum_{j=1}^{n} \left(c_j^{(1)} - \tilde{v}_j \right)^+.$$

Suppose (y^*, x^*) is a local optimal solution in a one-interchange neighborhood obtained by opening any one facility in $I^- = I \backslash I^+$, or closing any one facility in I^+, or doing both simultaneously. This implies that the following three conditions are satisfied:

(i) opening any facility $i \in I^-$ cannot improve the current solution; i.e.,

$$\sum_{j\in J}\left(c_j^{(1)} - c_{ij}\right)^+ \leqslant f_i, \quad \forall i \in I^-, \tag{34}$$

(ii) closing any facility $i \in I^+$ cannot improve the current solution; i.e.,

$$\sum_{j\in J_i}\left(c_j^{(2)} - c_j^{(1)}\right) \geqslant f_i, \quad \forall i \in I^+, \tag{35}$$

(iii) opening any $i_1 \in I^-$ and closing any $i_2 \in I^+$ cannot improve the current solution; i.e.,

$$\sum_{j\in J_{i_2}}\left(\min\{c_j^{(2)}, c_{i_1 j}\} - c_j^{(1)}\right) - \sum_{j\notin J_{i_2}}\left(c_j^{(1)} - c_{i_1 j}\right)^+ \geqslant f_{i_2} - f_{i_1}, \quad \forall i_1 \in I^-, \quad i_2 \in I^+. \tag{36}$$

Proposition 9. *If*

$$\sum_{j\in J}\left(c_j^{(2)} - c_{ij}\right)^+ \leqslant f_i, \quad \forall i \in I^-, \tag{37}$$

then (y^*, x^*) *is an optimal solution to the strong LP relaxation of SPLP.*

Proof. To prove this result we need to find a feasible dual solution that satisfies complementary slackness conditions (30)–(32). (Recall that (29) is automatically satisfied by the feasibility of the primal.) If the w_{ij} are defined according to (17), and the shadow prices v_j^* are specified such that $v_j^* \leqslant c_j^{(2)}$, $\forall j \in J$, it follows from Proposition 5 that (30) holds. Using (35), we may further restrict the v_j^* such that

$$c_j^{(1)} \leqslant v_j^* \leqslant c_j^{(2)}, \quad \forall j \in J, \quad \text{and} \quad \sum_{j\in J_i}\left(v_j^* - c_j^{(1)}\right) = f_i, \quad \forall i \in I^+,$$

i.e., (31) and (32) also hold. All that remains to show is that (14) is satisfied for feasibility of the dual. This is readily seen since

$$\sum_{j\in J}(v_j^* - c_{ij})^+ \leqslant \sum_{j\in J}\left(c_j^{(2)} - c_{ij}\right)^+ \leqslant f_i, \quad \forall i \in I^-,$$

by (37). Thus, a feasible dual is specified such that complementary slackness conditions (30)–(32) are satisfied; hence, the duality gap is zero and (y^*, x^*) is an optimal solution of the LP relaxation. It is also clear that if I^- is the null set, (y^*, x^*) is an optimal solution of the LP relaxation. □

Given a local optimal solution (y^*, x^*), Proposition 9 says that if we impose that each customer be served by its *second-closest* facility in I^+ (instead of first closest), it is sufficient that the saving in transportation cost by opening any site $i \in I^-$ be not greater than the opening cost of that facility for (y^*, x^*) to solve the LP relaxation. Alternatively, we may define for any solution (y, x) a "coefficient of attractiveness" for each facility site i, as

$$\alpha_i = \sum_{j\in J}\left(c_j^{(2)} - c_{ij}\right)^+ - f_i. \tag{38}$$

It suffices to have $\alpha_i \leqslant 0$, $\forall i \in I^-$, in order for (y, x) to solve the LP relaxation. Note also from (35) that $\alpha_i \geqslant 0$, $\forall i \in I^+$ is a necessary condition of optimality.

It follows that facility sites that have relatively large opening costs or are relatively remote from the customer set will tend to have negative coefficients of attractiveness, and remain closed. We may strengthen Proposition 9 with the following main result.

Theorem 1. *Given a primal solution* (y^*, x^*), *the duality gap will be zero*, *if*, *and only if*, *a vector* v^* *may be found such that*

$$c_j^{(1)} \leqslant v_j^* \leqslant c_j^{(2)}, \quad \forall j \in J, \tag{39}$$

$$\sum_{j \in J_i} \left(v_j^* - c_j^{(1)} \right) = f_i, \quad \forall i \in I^+, \tag{40}$$

and

$$\sum_{j \in J} \left(v_j^* - c_{ij} \right)^+ \leqslant f_i, \quad \forall i \in I^-. \tag{41}$$

Proof. Given that the gap is zero, it follows that complementary slackness conditions (30)–(32) must be satisfied by the optimal dual solution (v^*, w^*). This implies that (39) and (40) are true (see Propositions 4 and 5, and relation (33)). Feasibility of (v^*, w^*) implies that (41) is true.

That the given conditions are sufficient immediately follows from feasibility of the dual, and Proposition 1 after noting that $(c_j^{(1)} - v_j)^+ = 0$, $\forall j \in J$ by (39), and

$$f_i - \sum_{j \in J_i} \left(v_j - c_j^{(1)} \right)^+ = 0, \quad \forall i \in I^+,$$

by (40). (Alternatively, we may note that the given conditions (39)–(41) are sufficient, since a feasible dual solution is provided such that all the complementary slackness conditions are satisfied.) □

Note that the necessary condition for the inequality (39) has already been shown (see Proposition 4.5 in chapter 4 of [5]).

4. An illustrative example

We illustrate our theoretical results on a small test instance where the set of users J is identical to the set of potential facilities I. In this example $m = n = 9$ points are chosen in the rectangle $[0, 1] \times [0, 0.5]$ as shown in Fig. 1. Their coordinates are:

$$a_1 = (0.868910, 0.207614), \quad a_2 = (0.995115, 0.044832), \quad a_3 = (0.944689, 0.401347),$$
$$a_4 = (0.725728, 0.240687), \quad a_5 = (0.635106, 0.070752), \quad a_6 = (0.752709, 0.436559),$$
$$a_7 = (0.524239, 0.282562), \quad a_8 = (0.724439, 0.393242), \quad a_9 = (0.396442, 0.447605).$$

As in Barahona and Chudak [3], Euclidean distances between pairs of points are calculated and multiplied by 10,000 and rounded to get an integer variable cost matrix $C = (c_{ij})$ (see Table 1). It is assumed that fixed costs (f_i) are all equal to 4000.

The optimal solution can easily be found by inspection or by using any integer programming software. It is given by: $y_1 = y_8 = 1$, $y_2 = y_3 = y_4 = y_5 = y_6 = y_7 = y_9 = 0$. Now we can define sets I^+, J_1 and J_8 as follows: $I^+ = \{1, 8\}$; $J_1 = \{1, 2, 3, 4, 5\}$; $J_8 = \{6, 7, 8, 9\}$ (see Fig. 1). The primal objective function value is

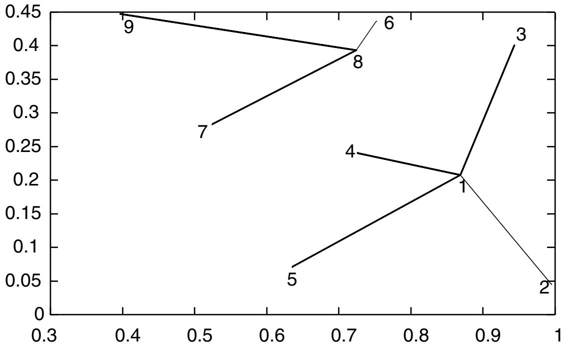

Fig. 1. An Euclidean example with optimal solution $I^+ = \{1, 8\}$.

Table 1
Variable cost matrix (c_{ij}) for numerical example ($n = m = 9$)

	1	2	3	4	5	6	7	8	9
1	0	2059	2080	1469	2709	2567	3527	2352	5299
2	2059	0	3600	3330	3609	4606	5274	4411	7215
3	2080	3600	0	2715	4529	1951	4369	2203	5501
4	1469	3330	2715	0	1925	1977	2057	1525	3889
5	2709	3609	4529	1925	0	3842	2390	3346	4460
6	2567	4606	1951	1977	3842	0	2755	517	3564
7	3527	5274	4369	2057	2390	2755	0	2287	2087
8	2352	4411	2203	1525	3346	517	2287	0	3324
9	5299	7215	5501	3889	4460	3564	2087	3324	0

$$\begin{aligned} z_P^* &= (f_1 + c_{1,2} + c_{1,3} + c_{1,4} + c_{1,5}) + (f_8 + c_{8,6} + c_{8,7} + c_{8,9}) \\ &= (4000 + 2059 + 2080 + 1469 + 2709) + (4000 + 517 + 2287 + 3324) = 22,445. \end{aligned}$$

4.1. Dual solutions

Four different dual solutions, labelled (a)–(d), are given below. Their v_j values and the corresponding duality gap appear in the respective v_j columns in Table 2 in the same order as the solutions are presented. The lower bound on the gap in the last row of the table is obtained by Proposition 2.

(a) *Unrestricted dual solution* (v^*). In the *general LP dual* (13)–(16), there are $m \cdot n + n = 90$ variables and $m \cdot n + m = 90$ constraints. The optimal solution (v^*, w^*) was obtained by Cplex 7.1 in 44 iterations (“no LP presolve or aggregation reductions”). We see that a duality gap exists:

$$z_P^* - z_D^* = 22,445 - 22,200 = 245.$$

We now check our theoretical results. First calculate the following sums (only non-zero terms are shown):

$$\sum_{j=1}^{9} (v_j - c_{1j})^+ = (2352 - 0) + (3516 - 2059) + (2203 - 2080) + (1537 - 1469) = 4000 = f_1,$$

$$\sum_{j=1}^{9} (v_j - c_{8j})^+ = (1537 - 1525) + (1977 - 517) + (2138 - 0) + (3714 - 3324) = 4000 = f_8.$$

Table 2
Four different dual solutions

j	j^+	$c_j^{(1)}$	v_j^*	v_j^+	$v_j^{+'}$	$\bar{v}_j$	$c_j^{(2)}$	j^{++}
1	1	0	2352	1823	2059	2352	2352	8
2	1	2059	3516	3600	3600	3528	4411	8
3	1	2080	2203	**2460**	**2394**	2203	2203	8
4	1	1469	**1537**	**1725**	**1554**	1525	1525	8
5	1	2709	**2676**	2709	2709	**2676**	3346	8
6	8	517	1977	1998	1977	1977	2567	1
7	8	2287	**2087**	2287	2287	**2087**	3527	1
8	8	0	2138	1905	2096	2138	2352	1
9	8	3324	3714	3481	3481	3714	5299	1
$z_D = \sum v_j$			22,200	21,988	22,158.5	22,200		
$z_P - z_D$			245	457	286.5	245		
Lower bound			245	457	220	233		

Boldfaced values violate the condition $c_j^{(1)} \leqslant v_j \leqslant c_j^{(2)}$.

Therefore the complementary slackness condition (31) is satisfied, and Proposition 3 should hold. Indeed, this dual solution violates condition $c_j^{(1)} \leqslant v_j^* \leqslant c_j^{(2)}$ three times; thus,

$$z_P^* - z_D^* = (1537 - 1525) + (2287 - 2087) + (2709 - 2676) = 245.$$

(b) *Dual ascent solution* (v^+). By *dual ascent* [8], that uses ranked distances given in Table 3, we get $z_D = 21{,}967$ and

$$v = (2080, 3600, 2203, 1725, 2709, 1977, 2287, 1905, 3481)^{\mathrm{T}},$$

$$s = (0, 379, 1771, 0, 1291, 383, 0, 278, 319)^{\mathrm{T}},$$

where vector s denotes the slacks for the dual constraints (14). Since $s_1 = s_4 = s_7 = 0$, according to the dual ascent procedure, I^+ is defined as $I^+ = \{1, 4, 7\}$, and the corresponding primal solution has $z_P = 23{,}653$. This gives $z_P - z_D = 23{,}653 - 21{,}976 = 1686$. Then *dual adjustment* is applied and in the third iteration the following dual solution is obtained:

$$v = (1823, 3600, 2460, 1725, 2709, 1998, 2287, 1905, 3481)^{\mathrm{T}},$$

$$s = (0, 400, 1493, 236, 1291, 105, 0, 0, 319)^{\mathrm{T}},$$

with $z_D = 21{,}988$, $I^+ = \{1, 8\}$, $z_P = 22{,}445$ and $z_P - z_D = 457$. The gap may also be calculated using Proposition 7: $z_P - z_D = (2460 - 2203) + (1725 - 1525) = 457$.

Table 3
Ranked distances with corresponding indices in brackets

	1	2	3	4	5	6	7	8	9
1	1469 (4)	2059 (1)	1951 (6)	1469 (1)	1925 (4)	517 (8)	2057 (4)	517 (6)	2087 (7)
2	2059 (2)	3330 (4)	2080 (1)	1525 (8)	2390 (7)	1951 (3)	2087 (9)	1525 (4)	3324 (8)
3	2080 (3)	3600 (3)	2203 (8)	1925 (5)	2709 (1)	1977 (4)	2287 (8)	2203 (3)	3564 (6)
4	2352 (8)	3609 (5)	2715 (4)	1977 (6)	3346 (8)	2567 (1)	2390 (5)	2287 (7)	3889 (4)
5	2567 (6)	4411 (8)	3600 (2)	2057 (7)	3609 (2)	2755 (7)	2755 (6)	2352 (1)	4460 (5)
6	2709 (5)	4606 (6)	4369 (7)	2715 (3)	3842 (6)	3564 (9)	3527 (1)	3324 (9)	5299 (1)
7	3527 (7)	5274 (7)	4529 (5)	3330 (2)	4460 (9)	3842 (5)	4369 (3)	3346 (5)	5501 (3)
8	5299 (9)	7215 (9)	5501 (9)	3889 (9)	4529 (3)	4606 (2)	5274 (2)	4411 (2)	7215 (2)

Therefore, the optimal primal solution is found after applying *dual adjustment*, but the dual solution obtained is smaller than the optimal dual value (21,988 instead of 22,200). This confirms that the duality gap obtained by the Erlenkotter [8] method may be larger than the real gap (in this case, 457 instead of 245). One reason for this is the assumption that $v_j \geqslant c_j^{(1)}$, $\forall j$, which we see is not necessarily true in the optimal solution.

(c) *Restricted dual* ($v^{+'}$). Here we add lower bound constraints $v_j \geqslant c_j^{(1)}$, $\forall j \in J$, to the original dual, and obtain the optimal solution for the restricted problem. There is no practical advantage to this approach since the number of variables in the model remains the same while the number of constraints increases by $|J| = 9$. However, it is interesting to observe that this solution is considerably better than the dual ascent solution in (b).

(d) *Restricted dual* ($\bar{v}$). Here we add upper bound constraints, $v_j \leqslant c_j^{(2)}$, $\forall j \in J$, to the original dual. In general, this will allow a significant reduction in the size of the problem, since w_{ij} may be set to zero whenever $c_{ij} \geqslant c_j^{(2)}$. Examining each column in turn in Table 2, we see that $(5+4+6+7+5+5+3+4+3) = 42$ of the w_{ij} are set to zero in this way for the given example. Thus, the restricted dual has $90 - 42 = 48$ variables and $90 - 42 + 9 = 57$ constraints. Also observe in this case that the result is superior to (c); in fact, the solution is also optimal in the original dual.

5. Conclusions

In this paper several properties are derived for the duality gap in the fundamental problem in location theory known as the simple plant location problem (SPLP). A formula for the duality gap of any primal–dual pair is given, and shown to be separable, with each term being related to an open facility. The duality gap in Erlenkotter [8] is seen to be a special case of the general formula. Other special cases are derived by making different assumptions on the complementary slackness conditions. Necessary and sufficient conditions are also provided for the duality gap to be zero. We show that upper bound constraints on the shadow prices (v_j) based on the distance (cost) to the second closest open facility allow a considerable reduction in the size of the dual. Furthermore, a numerical example demonstrates that this reduced dual may provide better estimates of the gap than the dual ascent and dual adjustment solution of Erlenkotter [8].

The results presented here may be useful in the future in algorithm design, and to obtain improved lower bounds for branch-and-bound procedures. The authors are currently exploring this issue.

References

[1] S. Ahn, G. Cooper, A.M. Frieze, Probabilistic analysis of a relaxation for the k-median problem, Mathematics of Operations Research 13 (1998) 1–31.

[2] M. Balinski, Integer programming: Methods, uses, computation, Management Science 12 (1965) 254–313.

[3] F. Barahona, F. Chudak, Solving large scale uncapacitated location problems, IBM Research Report RC 21515, 1999.

[4] J. Brimberg, C. ReVelle, The maximum return-on-investment plant location problem, Journal of the Operational Research Society 51 (2000) 729–735.

[5] P. Chardaire, Location optimization and games, Ph.D. thesis, School of Computing Sciences, University of East Anglia, Norwich, United Kingdom, 1998.

[6] M. Daskin, Network and Discrete Location: Models, Algorithms and Applications, Wiley, New York, 1995.

[7] Z. Drezner, Facility Location: a Survey of Applications and Methods, Springer, New York, 1995.

[8] D. Erlenkotter, A dual-based procedure for uncapacitated facility location, Operations Research 26 (1978) 992–1009.

[9] M. Goemans, M. Skutella, Cooperative facility location games, Journal of Algorithms 50 (2004) 194–214.

[10] P.B. Mirchandani, R.L. Francis, Discrete Location Theory, Wiley-Interscience, New York, 1990.

[11] C.H. Papadimitrou, Worst-case probabilistic analysis of a geometric location problem, SIAM Journal on Computing 10 (1981) 542–557.

[12] C. ReVelle, G. Laporte, The plant location problem: New models and research prospects, Operations Research 44 (1996) 864–874.

[13] C. ReVelle, Facility siting and integer friendly programming, European Journal of Operational Research 65 (1993) 147–158.