Comput. & Ops. Res. Vol. 11, No. 3, pp. 285–294, 1984
Printed in the U.S.A.

0305-0548/84 $3.00 + .00

LOCATION PROBLEMS WITH COSTS BEING SUMS OF POWERS OF EUCLIDEAN DISTANCES

REUVEN CHEN†
Department of Physics and Astronomy and Faculty of Management, Tel-Aviv University, Tel-Aviv, Israel

Scope and purpose—A crucial problem while considering the location of a new facility with respect to existing facilities is to determine the kind of distance dependent function which is to be minimized in order to get an optimum in a desired sense[1]. For certain problems, one wishes to minimize the weighted sums of Euclidean (straight line) distances, e.g. to minimize transportation costs between the given facilities and the new one, under the assumption that costs are strictly proportional to the distances (minisum).

In many cases, however, this nicety of a strict linear relation between distance and cost of transportation seems to be an oversimplification (e.g. [2], [6], [12], [14–16], [21–22]). In some cases, one can expect the cost to be a sublinear function of the distance[16] which is an example of the well known situation of "economies of scale". On the other hand, a strongly superlinear cost function should be considered in problems such as the location of emergency facilities, in which service to the farthest customer is the main concern (minimax). Cases of moderately superlinear cost functions are also of importance when one wishes to solve the hybrid problem which is intermediate between the minimax and the minisum ones[22]. The most obvious choice for such sublinear and superlinear functions are the power function r_i^n, with $n < 1$ in the former case and $n > 1$ in the latter[12, 14].

The present work presents an amendment to a method previously suggested by Cooper[12] to solve this problem of minimizing the sum of weighted powers of the Euclidean distances. This amendment enables the solution of the problem even for rather large powers of the distances, which could not be done by Cooper's original method. It also expedites substantially the iterative procedure for smaller values of the power n.

Abstract—The Cooper problem of minimizing the sum of weighted powers of the Euclidean distances is further studied. The practical significance of this problem is that instead of the Weber problem where the weighted sum of the Euclidean distances is minimized, here economies and diseconomies of scale are allowed for, which results in a problem of minimizing the sum of weighted powers of the Euclidean distances. The original Cooper solution has been a modification of the Weiszfeld method for the Weber problem; it can be shown to be a steepest descent method with a step size determined by an expression depending on the weights and distances, which is not optimal in any sense. In the present work, the matter of the best step size to be taken is investigated. It is shown that multiplying the step size given by Cooper by $2/n$ where n is the power of the Euclidean distances, improves the iterative process substantially. For the cases with $n \geq 1$ considered by Cooper, namely $1 \leq n \leq 3$, the number of steps needed to reach a certain termination criterion is reduced. For higher values of n, the original Cooper step size is usually too large and there is no convergence. With the present amended method, the problem can be solved, using an appropriate normalization, even for powers of 100 and more. A semi-intuitive proof for this step size is given as well as numerical examples of solutions with $n = 1$, 10 and 100. The method can easily be extended to location problems in three (and more) dimensions.

INTRODUCTION

The Weber location problem, also known as the Fermat or Steiner problem, is that of locating a "service facility" while a set of m facilities located at (a_i, b_i), $i = 1 \ldots m$, is given, so that the sum of weighted distances is minimized. The "metrics" by which the (generalized) distance between two points is considered may be different in various cases. (For a recent discussion on the impact of "distance" on location problems see Krarup and Pruzan[2].) Here we concentrate on Euclidean distances or powers of Euclidean distances. The weighted problem with Euclidean distances is

$$\min_{x,y} f(x, y) = \sum_{i=1}^{m} w_i r_i(x, y) \tag{1}$$

†Dr. Reuven Chen is affiliated with the Physics Department, and the Faculty of Management, Tel-Aviv University, Israel. He received his M.Sc. and Ph.D. degrees in Physics from the Hebrew University, Jerusalem. He has held visiting posts in the Polytechnic Institute of Brooklyn, N.Y., Queens University, and Simon Fraser Univ., Canada, Oxford and Birmingham University, England.

He published 50 papers in solid state physics and co-authored a book *Analysis of Thermally Stimulated Process* (Pergamon Press, 1981). Recently, he became interested in operations research and in particular, in location problems.

where

$$r_i(x, y) = [(x - a_i)^2 + (y - b_i)^2]^{\frac{1}{2}} \tag{2}$$

and where w_i are the weights associated with the given points (a_i, b_i) (w_i representing, say, the amounts to be shipped between the two points in a given period of time). (x, y) is any point in E^2; the extension to E^K for $K > 2$ will also be briefly mentioned below. An iterative solution of this problem has been suggested by Weiszfeld[3] for the equiweighted problem ($w_i = 1$), which has later been studied and extended to include the non-equiweighted cases[4–11]. The conditions for getting a minimum in (1) are reached by partial differentiation of $f(x, y)$ with respect to x and y and equating the derivatives to zero. This yields

$$x = \sum_{i=1}^{m} (w_i a_i / r_i) \Big/ \sum_{i=1}^{m} (w_i / r_i); \quad y = \sum_{i=1}^{m} (w_i b_i / r_i) \Big/ \sum_{i=1}^{m} (w_i / r_i) \tag{3}$$

where, of course, the r_i's are functions of x and y. Weiszfeld's iterative algorithm was based on this, namely, (x^{N+1}, y^{N+1}) is calculated in the $(N+1)$th iterative step from (x^N, y^N), using

$$x^{N+1} = \sum_{i=1}^{m} [w_i a_i / r_i(x^N, y^N)] \Big/ \sum_{i=1}^{m} [w_i / r_i(x^N, y^N)] \tag{4a}$$

$$y^{N+1} = \sum_{i=1}^{m} [w_i b_i / r_i(x^N, y^N)] \Big/ \sum_{i=1}^{m} [w_i / r_i(x^N, y^N)]. \tag{4b}$$

As a first approximation one can use the "center of gravity" solution, namely,

$$x^0 = \sum_{i=1}^{m} w_i a_i \Big/ \sum_{i=1}^{m} w_i; \quad y^0 = \sum_{i=1}^{m} w_i b_i \Big/ \sum_{i=1}^{m} w_i. \tag{5}$$

Cooper[12] suggested the extension of problem (1) so that the minimization will be

$$\min_{x, y} f(x, y) = \sum_{i=1}^{m} w_i r_i^n \tag{6}$$

where r_i is given by (2). The case of $n < 1$ is that of economies of scale whereas $n > 1$ means diseconomies of scale. In other words, $n < 1$ represents the sublinear case in which the cost element $f_i(r)$ is such that $d^2 f_i / dr^2 < 0$. $n > 1$ represents the superlinear case in which $d^2 f_i / dr^2 > 0$. Cooper showed that the objective function is convex when $n \geq 1$ and therefore every local minimum is a global minimum. When $n < 1$, he showed that the objective function is neither convex nor concave, therefore any iterative process we choose will lead to a local minimum which by no means is necessarily global. It is to be noted that the function (6) is not differentiable at the given points (a_i, b_i) for $n \leq 1$. This difficulty can be by-passed in the original Weber problem[8] ($n = 1$), anyway we shall concentrate on the cases where $n > 1$.

Cooper extended the Weisfeld method by following the same steps of equating the derivatives to zero and writing the index $N + 1$ on the l.h.s. and N on the r.h.s.,

$$x^{N+1} = \sum_{i=1}^{m} w_i a_i [r_i(x^N, y^N)]^{n-2} \Big/ \sum_{i=1}^{m} w_i [r_i(x^N, y^N)]^{n-2} \tag{7a}$$

$$y^{N+1} = \sum_{i=1}^{m} w_i b_i [r_i(x^N, y^N)]^{n-2} \Big/ \sum_{i=1}^{m} w_i (x^N, y^N)]^{n-2}. \tag{7b}$$

Cooper solved problems of this kind with values of n up to 3.

Only few of the researchers dealing with problem (1) stated that the Weiszfeld method is nothing but the steepest descent method with a step size determined by the denominator in eqns (4a) and (4b) (e.g. see Ostresh[9] and Avriel[13]).

We shall show in this work that the Cooper method for solving problem (6) is a steepest descent method as well. It will be shown that the step size determined by the denominator in (7a) and (7b) is far from being optimal. Empirical results as well as a semi-intuitive proof show that the optimal step size is $2/n$ times that given by the Cooper method. Whereas the original Cooper method does not usually converge for $n > 3$, the amended method suggested here converges very efficiently (say, starting from the initial point of the "center of gravity") even for values of n as high as 100 and more. The implication of such a solution for very large values of n as an approximation to the minimax problem will be mentioned.

MOTIVATION

A number of authors discussed the Cooper problem with $n < 1$ and $n > 1$ (e.g. see a review paper by Scott[14]). It seems indeed that non-linear elements of the cost functions are to be considered (e.g. see Dearing[15]). However, it seems that when such non-linearity occurs, the dependence is more often than not sublinear rather than superlinear (e.g.[16]). The method proposed here can expedite the iterative procedure for $n < 1$ as well as for $n \geq 1$, however, in the former case, the problem is not convex and therefore, a solution reached by any iterative procedure may not be global. Also, as discussed below, at the demand points (a_i, b_i) the objective function is not differentiable when $n < 1$ which may impair the process.

An important application of solving the Cooper problem for large power n is to get a good approximation for the minimax problem which is often used for the location of emergency facilities (for a review on the problem and other possible ways to solve it see Ref.[17]). As described by Drezner and Wesolowsky[18], for any set of positive numbers $C_1, \ldots, C_m$

$$\max\{C_1, \ldots, C_m\} = \lim_{n \to \infty} (C_1^n + \cdots + C_m^n)^{1/n}. \tag{8}$$

Thus, a good approximation for the minimax problem

$$\min_{x,y} \max_i w_i r_i \tag{9}$$

will be given for large value of n by

$$\min_{x,y} \left\{ \sum_{i=1}^{m} (w_i r_i)^n \right\}^{1/n}. \tag{10}$$

Drezner and Wesolowsky continued their solution from this point by resorting to the solution of differential equations. The solution will remain the same if we raise the expression to the nth power, namely solving

$$\min_{x,y} \sum_{i=1}^{m} (w_i r_i)^n \tag{11}$$

which is the Cooper problem with the weights being w_i^n. The solution of the minimax problem by using this approximation will be described elsewhere. At this point we only mention that, as compared to other minimax algorithms, the one suggested here is attractive due to its very simple extension to multi-dimensional minimax problems. In the context of approximating a minimax problem by a minisum one, a recent paper by Krarup and Pruzan[19] should be mentioned. Although conceptually they are dealing with the

same kind of transformation minimax→minisum, their treatment is different since they solve 0–1 programming problems.

The main motivation for solving Cooper's problem with $n > 1$ is thus the utilisation of this solution as a good approximation for minimax problems. In fact, the extension to even more general minimax problems, namely

$$\min_{x,y} \max_{i} f_i(r_i) \tag{12}$$

where f_i are any increasing cost functions[15] seems rather straightforward. A point should be made, however, for studying problem (6) with $n > 1$, n not being very large. As claimed by Burkard *et al.*[20], "neither the minisum nor the minimax criterion alone capture most elements of a location problem where it is important to consider both the total costs of serving clients as well as the service provided to those clients who are located far away from the facility. The minisum criterion alone will not take the worst case solutions into account and may result in solutions which are unacceptable from a service point of view. On the other hand, the minimax criterion if used alone may result in very costly systems". A similar point had been made previously by Halpern[21] and Handler and Mirchandani[22] with respect to location on networks. Halpern[21] and Burkard *et al.*[20] attacked this hybrid location problem by taking a convex combination of the minimax and minisum terms with a parameter α that varied from 0 to 1 while the problem changed continuously from minimax to minisum. Handler and Mirchandani[22] preferred to deal with this "medi-center" problem by suggesting a "penalty approach". Starting with the median framework for minimizing average distance, they want to penalize greater distances between facility and demand locations at increasing rates by generalizing the linear distance function $d(x, y)$ in the minisum formulation to a nondecreasing convex (super-linear) function of the distance, $f(d(x, y))$. The most obvious choice for such a function is the power r_i^n with $n > 1$ which in our case of a continuous space leads to problem (6). Of course, one has to determine beforehand the desirable value of n to be chosen; a larger n will be chosen when the "service to farthest customer" element is to be given a higher weight whereas a small (closer to 1) value will be taken when one wants to emphasize the "cost" aspect.

AMENDMENT OF THE COOPER METHOD

Let us show first that the direction determined by eqns (7a)–(7b) including the special case for $n = 1$, eqns (4a)–(4b), is the steepest descent direction. The partial derivatives of (6) at a point (x, y) are

$$\partial f/\partial x = n \sum_{i=1}^{m} w_i r_i^{n-1} (\partial r_i/\partial x) = n \sum_{i=1}^{m} w_i r_i^{n-1} (x - a_i)/r_i \tag{13a}$$

$$\partial f/\partial y = n \sum_{i=1}^{m} w_i r_i^{n-1} (y - b_i)/r_i. \tag{13b}$$

These expressions include the terms $r_i^{n-2}(x - a_i)$ and $r_i^{n-2}(y - b_i)$ which, at first sight may appear not to be defined for $(x, y) = (a_i, b_i)$ for any i (which, of course implies $r_i = 0$), when $n \leq 2$. A closer look at eqns (13a) and (13b) show that the problem occurs only for $n \leq 1$. This is so since the terms $(x - a_i)/r_i$ and $(y - b_i)/r_i$, though not uniquely defined in the vicinity of the given points (a_i, b_i) have finite values between zero and unity. For $n > 1$ these are multiplied by r_i^{n-1} which goes to zero for $(x, y) \to (a_i, b_i)$ and thus the contributions to $\partial f/\partial x$ and $\partial f/\partial y$ in (13a) and (13b) are nill. For $n \leq 1$, the gradient is indeed not defined at (a_i, b_i). As mentioned above, this has to be separately considered in the original Weber case ($n = 1$) and the "economies of scale" cases ($n < 1$).

In order to see the direction given by the Cooper method (eqns 7a–7b), let us consider the step determined by two subsequent points in the iteration,

$$x^{N+1} - x^N = \sum_{i=1}^{m} w_i a_i r_i^{n-2} \Big/ \sum_{i=1}^{m} w_i r_i^{n-2} - x^N = \sum_{i=1}^{m} w_i (a_i - x^N)$$

$$\times [r_i(x^N, y^N)]^{n-2} \Big/ \sum_{i=1}^{m} w_i [r_i(x^N, y^N)]^{n-2} \tag{14a}$$

and in the same way

$$y^{N+1} - y^N = \sum_{i=1}^{m} w_i(b_i - y^N) r_i^{n-2} \Big/ \sum_{i=1}^{m} w_i r_i^{n-2}. \tag{14b}$$

Comparing eqns (14a)–(14b) and (13a)–(13b) readily shows that the Cooper step (including the Weiszfeld step for $n = 1$) is in the direction of ∇f, i.e. the steepest descent. The step size is determined by the factor $\Sigma_{i=1}^{m} w_i r_i^{n-2}$, the denominator in eqns (7) and (14). Two questions should be asked at this point.

(1) Can we consider alternatives to using the steepest descent method?

(2) If we do decide to use the steepest descent method, what is the best step size to choose.

The alternative of using the Newton method has been considered by Harris[23] and El-Shaieb[24] for the original Weber problem ($n = 1$). In the Newton method one has to calculate the inverse Hessian at each step, which is, of course, time consuming. This is usually compensated by the reduction in the number of steps needed to reach a certain termination criterion. The location problem as discussed so far is a two dimensional problem, therefore the additional complication of calculating the inverse Hessian does not seem too costly. However, as pointed out by Harris, two additional problems sometimes occur while using Newton's method. His experience is based on solving the Weber problem ($n = 1$); similar difficulties were encountered in course of the present work while solving Cooper's problem with $n > 1$.

(1) The Newton method has a risk of getting outside of the convex hull of the given points, in which case the method may break down.

(2) Sometimes there may be a tendency for the solution process to oscillate so that the termination criterion may never be fulfilled. Harris[23] points out that the oscillation may be reduced and some of the advantages of the Newton method retained if only the principal diagonal of the Hessian is used. This point will be further discussed below.

COMPUTATIONAL EXPERIENCE

We shall concentrate here on choosing a better step size, and show that very satisfactory results can be reached following this route. The first thing that comes to mind is to perform a line search along the gradient direction which would, of course, increase the time needed for each step, but probably reduce the number of steps needed. This has been done by Cooper and Katz[25] for the Weber problem ($n = 1$); they found that for some choices of line search, the method was rather inefficient. Only one search method, namely Armijo's[26] inexact line search has made a considerable improvement of saving about one third of the CPU time. The price to be paid is, of course, the complication in the programming, i.e. an increase in the programmer's time. In the method suggested in this work, a similar saving of CPU time has been achieved with the problem tested by Cooper and Katz with practically no complication added to the program. As mentioned already, the main feature of the present method is that it can be used for $n > 1$ and even for very large values of n with only a slight change in the original program.

While aiming at getting a good approximation to minimax problems, an attempt has been made in the present work to solve problem (6) with n being, say, 10 or more, by using eqns (7a)–(7b). In all these attempts, it has been readily seen that no convergence was achieved with such values of n. Empirically, smaller step sizes were taken so that instead of eqns (7) or (14), the following have been utilised

$$x^{N+1} = x^N + C \sum_{i=1}^{m} w_i(a_i - x^N) r_i^{n-2} \Big/ \sum_{i=1}^{m} w_i r_i^{n-2} \tag{15a}$$

$$y^{N+1} = y^N + C \sum_{i=1}^{m} w_i(b - y^N) r_i^{n-2} \Big/ \sum_{i=1}^{m} w_i r_i^{n-2}. \tag{15b}$$

By trial and error, the constant C was chosen so as to get convergence of the iterative

process. Table 1 demonstrates some of the results for $n = 1$, 10, 100 and for different choices of C. (a_i, b_i) were 100 points, the coordinates of which were "random" numbers (generated by the computer) between 0 and 100. Two sets of calculations were taken, one for the equi-weighted case (all $w_i = 1$) and in the other, the weights w_i were also random numbers between 0 and 100.

A practical point concerning the magnitudes of the numbers involved should be mentioned. For large n, say $n = 100$, and values of r_i of up to ~ 100, r_i^n may be too large to be handled by the computer one uses. This problem can be remedied quite easily. The solution of problem (6) will remain unchanged if all the terms in the sum are divided (or multiplied) by the same constant. One should choose a constant r_0 such that $\Sigma_{i=1}^m w_i(r_i/r_0)^n$ will not exceed the capacity of the computer to handle large numbers. In this "normalization", each term in the sum in problem (6) has thus been divided by the same constant r_0^n and therefore, the solution is not changed. The choice of r_0 depends on the actual computer one is using. For the small values of r_i, an underflow may occur while $(r_i/r_0)^n$ is computed. The term is replaced by zero and does not contribute to the sum, as it should be. This shows, however, that one should not choose a too large value of r_0 since too many terms may "disappear". The best policy to be used is to take r_0 such that the largest $w_i(r_i/r_0)^n$ is, say, two orders of magnitude below the maximum number handled by the computer, to allow for the contribution of the other terms.

The results in Table 1 give the number of steps needed for reaching a termination criterion as a function of the chosen constant C in eqns (15) which determines the step size. For $n = 1$, values of C from 0.2 to 3.2 are taken in steps of 0.2. Both in the equi-weighted and weighted problems, the optimum is seen to be at $C = 2.0$. It has already been mentioned by Ostresh[27] that "there are theoretical reasons for suspecting that $C = 2$ gives the fastest rate of convergence" if (x, y) does not coincide with any (a_i, b_i). He also says that $C = 1$ is the best when the solution is one of the given points, however, as mentioned by Katz[8], it is preferable to identify this situation directly and not to resort to the iterative process when such coincidence occurs. Anyway, it seems that this problem will not be of significance for $n > 1$ since, as mentioned above, there is no discontinuity in the derivative when $n > 1$. For $n = 10$ the optimum is readily seen to be at $C = 0.2$ and for $n = 100$ it is $C = 0.02$. The empirical conclusion that one should always expect to have $C = 2/n$ seems rather straightforward. As shown in the table, this is so for both the equi-weighted and weighted cases. The termination of the table is in each case at the value of C where no convergence is reached. It is rather obvious why Cooper's choice of $C = 1$

Table 1. Number of steps in iterations for various step sizes

$n = 1; \frac{2}{n} = 2.0$			$n = 10; \frac{2}{n} = 0.2$			$n = 100; \frac{2}{n} = 0.02$		
c	$w_i = 1$	w_i random	c	$w_i = 1$	w_i random	c	$w_i = 1$	w_i random
0.2	57	51	0.02	54	31	0.002	135	116
0.4	31	28	0.04	29	19	0.004	76	67
0.6	22	20	0.06	20	13	0.006	54	48
0.8	16	15	0.08	15	10	0.008	42	38
1.0	13	12	0.10	12	8	0.010	34	31
1.2	11	10	0.12	10	7	0.012	29	26
1.4	9	8	0.14	8	6	0.014	25	23
1.6	8	7	0.16	7	5	0.016	22	20
1.8	6	6	0.18	6	4	0.018	20	18
2.0	6	5	0.20	5	3	0.020	18	16
2.2	7	8	0.22	7	4	0.022	24	26
2.4	10	10	0.24	9	5	0.024	209	231
2.6	13	14	0.26	11	6			
2.8	18	25	0.28	15	8			
3.0	26	29	0.30	21	10			
3.2	50		0.32	33	14			
			0.34	63	19			
			0.36		34			
			0.38		119			

does not bring about convergence for rather large values of n. It is to be noted that very similar results were found for other problems, namely, for other given data.

As shown below in the semi-intuitive proof to this result, the condition for getting the result that $C = 2/n$ is that the number of given points m is rather large and that they are distributed more or less uniformly in a given range of x and y. In fact it has been found that the choice of $C = 2/n$ is a very good one even when m is small, say $m = 5$ and when the given points are not distributed over a square but say, on a rectangle of 10×100. The rate of convergence does reduce substantially if the points are distributed nearly linearly, e.g. on a rectangle of 1×100. At large values of n, divergence may occur for a C which is only slightly larger than the optimal. To be on the safe side, one may decide to choose a value for C which is slightly below $2/n$ (say, $1.8/n$). The number of steps may slightly increase, but the risk of not having convergence is substantially reduced.

A SEMI-INTUITIVE PROOF

The main point about the semi-intuitive proof which is presented here is based on a comparison between the steepest descent and the Newton methods. Whereas in the steepest descent method the iteration is along $-\nabla f$, in the Newton case the step is $-H^{-1}\nabla f$ where H is the Hessian matrix at the current point. The effect of the multiplication of the gradient by H^{-1} is twofold, it changes the direction of search and determines the step size. In spite of the difficulties sometimes occurring with the Newton method as mentioned above, it is usually expected to be quite efficient. In the present account, we would like to simplify the Newton step while retaining some of its desirable properties; in this sense we follow Harris[23] who empirically suggested to use (for $n = 1$ only) the diagonal elements of the Hessian. We shall show that in a favorable case in which the existing facilities are quite numerous and are distributed uniformly over, say, a square, the H^{-1} matrix will be approximately diagonal (i.e. the off-diagonals will be much smaller than the diagonals) and the diagonal terms will be approximately equal. The magnitude of the diagonal terms will determine the step size to be $2/n$ times the Cooper stepsize.

The elements of the Hessian will be found by differentiating eqns (13a)–(13b) (see also [12]),

$$\partial^2 f/\partial x^2 = n \sum_{i=1}^{m} w_i r_i^{n-2} + n(n-2) \sum_{i=1}^{m} w_i r_i^{n-4}(x - a_i)^2, \tag{16}$$

$$\partial^2 f/\partial y^2 = n \sum_{i=1}^{m} w_i r_i^{n-2} + n(n-2) \sum_{i=1}^{m} w_i r_i^{n-4}(y - b_i)^2, \tag{17}$$

$$\partial^2 f/\partial x \partial y = n(n-2) \sum_{i=1}^{m} w_i r_i^{n-4}(x - a_i)(y - b_i). \tag{18}$$

Let us consider now the contribution of a certain (a, b_i) to the sums in eqns (16)–(18). Figure 1 shows the location of (a_i, b_i) relative to the current (x, y). θ_i is the angle at which (a_i, b_i) is seen from (x, y) or, more explicitly, the angle between the horizontal direction

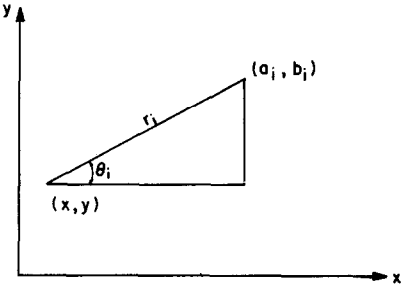

Fig. 1. The ith given facility (a_i, b_i) as seen from a point (x, y). r_i is the Euclidean distance and θ_i is the angle between the horizontal and the line connecting the two points.

and the direction connecting (x, y) and (a_i, b_i). The following relations are rather obvious,

$$(x-a_i)^2/r_i^2 = \cos^2\theta_i;\ (y-b_i)^2/r_i^2 = \sin^2\theta_i;\ (x-a_i)(y-b_i)/r_i^2 = \sin\theta_i\cos\theta_i. \tag{19}$$

Using these, eqns (16)–(18) can be written as

$$\partial^2 f/\partial x^2 = n\sum_{i=1}^{m} w_i r_i^{n-2} + n(n-2)\sum_{i=1}^{m} w_i r_i^{n-2}\cos^2\theta_i, \tag{16'}$$

$$\partial^2 f/\partial y^2 = n\sum_{i=1}^{m} w_i r_i^{n-2} + n(n-2)\sum_{i=1}^{m} w_i r_i^{n-2}\sin^2\theta_i, \tag{17'}$$

$$\partial^2 f/\partial x\partial y = n(n-2)\sum_{i=1}^{m} w_i r_i^{n-2}\sin\theta_i\cos\theta_i. \tag{18'}$$

If the given points are distributed uniformly and if the current point in the iteration is not too far from the final result, the angle θ_i can be considered to be uniformly distributed, $0 \le \theta_i < 2\pi$. Denoting the average value of a function (over, say, 2π) *by* $\langle\cdot\rangle$, we can write

$$\langle\cos^2\theta_i\rangle = \langle\sin^2\theta_i\rangle = \frac{1}{2};\langle\sin\theta_i\cos\theta_i\rangle = 0. \tag{20}$$

The following point is the weak part of the proof, due to which it has been termed "semi-intuitive". We replace $\cos^2\theta_i$, $\sin^2\theta_i$ and $\sin\theta_i\cos\theta_i$ in eqns (16')–(18') by their average values and get

$$\partial^2 f/\partial x^2 \approx n\sum_{i=1}^{m} w_i r_i^{n-2} + \frac{1}{2}n(n-2)\sum_{i=1}^{m} w_i r_i^{n-2} = \frac{1}{2}n^2\sum_{i=1}^{m} w_i r_i^{n-2} \tag{16''}$$

and, in the same way

$$\partial^2 f/\partial y^2 \approx \frac{1}{2}n^2\sum_{i=1}^{m} w_i r_i^{n-2} \tag{17''}$$

and

$$\partial^2 f/\partial x\partial y \approx 0. \tag{18''}$$

The validity of these results has been checked in actual cases and it has been found that typically, the off-diagonal elements were an order of magnitude smaller than the diagonals. The diagonal terms differed from one another by up to a factor of 1.5 which, again, shows the approximate nature of these considerations. Combining the relation (16")–(18"), we get the approximate Hessian

$$H \approx \frac{1}{2}n^2\sum_{i=1}^{m} w_i r_i^{n-2} \times I \tag{21}$$

where I is the two dimensional unity matrix. The inverse Hessian will therefore be approximately

$$H^{-1} \approx \left[2\Big/\left(n^2\sum_{i=1}^{m} w_i r_i^{n-2}\right)\right] \times I. \tag{22}$$

Combining eqns (13a, b) and (22) we readily have

$$H^{-1}\nabla f \approx (2/n)\begin{pmatrix} \sum_{i=1}^{m} w_i r_i^{n-2}(x-a_i)\Big/\sum_{i=1}^{m} w_i r_i^{n-2} \\ \sum_{i=1}^{m} w_i r_i^{n-2}(y-b_i)\Big/\sum_{i=1}^{m} w_i r_i^{n-2} \end{pmatrix} \tag{23}$$

hence we found the Cooper iteration step with step-size multiplied by $2/n$.

DISCUSSION

A substantial improvement has been suggested in this work to the method originally proposed by Cooper[12] to the problem (6) of minimizing the weighted sums of nth powers of the Euclidean distances. Empirically, it has been found that the iterative procedure is expedited by multiplying the Cooper step size by $2/n$. This method enabled the solution of the problem for very large values of n, as high as 100 and more whereas the original Cooper method did not converge usually for $n > 3$. The semi-intuitive proof that is given as a partial support to the empirical results is based on the particular structure of the problem, namely, that the elements of the sum to be minimized are "radial", i.e. depend on the Euclidean distances (rather than depending in any other way on (x, y)). In the proof, it is shown that the steepest descent direction should not differ by much from the Newton direction (under the appropriate assumptions), and that the Newton step-size can be approximated by $2/n$ times the Cooper step-size. It is somewhat surprising, however, that the performance of the present method, which as shown here can be regarded as an approximation to the Newton method, is superior to the latter in the sense that the two possible ailments mentioned by Harris[23], i.e. divergence and oscillations never occurred in the problems solved in this work. It is possible that another proof, not based on the Newton step can be devised in order to justify the multiplication of the Cooper step-size by $2/n$. Attempts in this direction have not succeeded so far.

An extension of the method to three dimensional (as well as multidimensional) problems is straightforward. Three dimensional problems have been solved in course of the present work with results almost as good as described above for the two dimensional problems. The given points were here (a_i, b_i, c_i) for $i = 1, \ldots m$; the initial values in eqn (5) should include $z^0 = \Sigma_{i=1}^m w_i c_i / \Sigma_{i=1}^m w_i$ and we added to the iteration equations (15)

$$z^{N+1} = z^N + C \sum_{i=1}^{m} w_i (c_i - z^N) r_i^{n-2} \Big/ \sum_{i=1}^{m} w_i r_i^{n-2} \tag{15c}$$

where still $C = 2/n$ was found to be the best choice in (15a–c). r_i in eqns (15a–c), is now

$$r_i(x, y, z) = [(x - a_i)^2 + (y - b_i)^2 + (z - c_i)^2]^{\frac{1}{2}}. \tag{2'}$$

The term "best choice" should be understood here, in the following "average" meaning. In some cases, in particular in problems where $n \approx 1$, the smallest number of iterative steps was achieved for values of the factor $C \approx 1.5/n$. For large values of n, it was closer to $2.5/n$. Anyway, the choice of $C = 2/n$ always resulted in quite fast convergence.

Although all the considerations in this work were related to the location of a single facility, the same principles can be applied in the solution of multifacility location and location–allocation problems. These will be elaborated elsewhere.

REFERENCES

1. R. L. Francis and J. A. White, *Facility Layout and Location: An Analytical approach*. Prentice-Hall, Englewood Cliffs, New Jersey (1974).
2. J. Krarup and P. M. Pruzan, The impact of distance on location problems. *Eur. J. Op. Res.* **4**, 256 (1980).
3. E. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Math. J.* **43**, 355 (1937).
4. H. W. Kuhn and R. E. Kuenne, An efficient algorithm for the numerical solution of the generalized Weber problem in spatial economics. *J. Reg. Sci.* **4**, 21 (1962).
5. H. W. Kuhn, On a pair of dual nonlinear programs. *Nonlinear Programming* (Edited by J. Abadie), p. 37. North Holland, Amsterdam (1967).
6. I. N. Katz, On the convergence of a numerical scheme for solving some locational equilibrium problems. *SIAM J. Appl. Math.* **17**, 1224 (1969).
7. H. W. Kuhn, A note on Fermat's problem. *Math. Prog.* **4**, 98 (1973).
8. I. N. Katz, Local convergence in Fermat's problem, *Math. Prog.* **6**, 89 (1974).
9. L. M. Ostresh Jr., On the convergence of a class of iterative methods for solving the Weber location problem, *Op. Res.* **26**, 597 (1978).
10. Z. Drezner, Bounds on the optimal location to the Weber problem under conditions of uncertainty. *J. Opl Res. Soc.* **30**, 923 (1979).
11. U. Eckhardt, Weber's problem and Weiszfeld's algorithm in general space. *Math Prog.* **18**, 186 (1980).
12. L. Cooper, An extension to the generalized Weber problem, *J. Reg. Sci.* **8**, 181 (1968).

13. M. Avriel, A geometric programming approach to the solution of location problems. *J. Reg. Sci.* **20**, 239 (1980).
14. A. J. Scott, Location-allocation systems: A review. *Georg. Anal.* **II**, 95 (1970).
15. P. M. Dearing, Minimax location problems with nonlinear costs. *J. Res. Nat. Bur. Stand.* **82**, 65 (1977).
16. P. Kolesar, W. Walker and J. Hausner, Determining the relation between fire engine travel times and travel distances in New-York City. *Ops. Res.* **23**, 614 (1975).
17. D. W. Hearn and J. Jesunathadas, Analysis and extensions of algorithms for Sylvester's matrix location problem. Res. Report. 80–1, Ind. Syst. Engng Dept., University of Florida (1980).
18. Z. Drezner and G. O. Wesolowsky, A new method for the multifacility minimax location problem. *Op. Res. Soc.* **29**, 1095 (1978).
19. J. Krarup and P. M. Pruzan, Reducibility of minimax to minisum 0–1 programming problems, *Eur. J. Op. Res.* **6**, 125 (1981).
20. R. E. Burkard, J. Krarup and P. M. Pruzan, Efficiency and optimum in minisum, minimax 0–1 programming problems, *J. Opl Res. Soc.* **33**, 137 (1982).
21. J. Halpern, The location of a center-median convex combination on an undirected tree. *J. Reg. Sci.* **16**, 237 (1976).
22. G. Y. Handler and P. B. Mirchandani, *Location on Networks, Theory and Algorithms*, Chap. 4. The MIT Press, Cambridge, Mass. London, England, (1979).
23. B. Harris, Speeding up iterative algorithms—the generalized Weber problem. *J. Reg. Sci.* **16**, 411 (1978).
24. A. M. El-Shaieb, The single source Weber problem—survey and extensions, *J. Opl Res. Soc.* **29**, 467 (1978).
25. L. Cooper and I. N. Katz, The Weber problem revisited. *Comput. Math. Applics* **7**, 225 (1981).
26. L. Armijo, Minimization of functions having continuous derivatives. *Pacific J. Math.* **16**, 1 (1966).
27. L. M. Ostresh, Jr., The multifacility location problem. Applications and descent theorems. *J. Reg. Sci.* **17**, 409 (1977).