ELSEVIER

Operations Research Letters 22 (1998) 151–157

operations research letters

Accelerating convergence in the Fermat–Weber location problem

Jack Brimberg [a], Reuven Chen [b,*], Doron Chen[c]

[a] *Department of Business Administration, Royal Military College, Kingston, Ont., Canada K7K 5LO*
[b] *School of Physics and Astronomy, Raymond and Beverly Sackler Faculty of Exact Sciences, Tel-Aviv University, Tel-Aviv 69978, Israel*
[c] *Herzliya, Israel*

Received 1 June 1996; revised 1 February 1998

Abstract

This paper presents a simple procedure for accelerating convergence in a generalized Fermat–Weber problem with l_p distances. The main idea is to multiply the predetermined step size of the Weiszfeld algorithm by a factor which is a function of the parameter p. The form of this function is derived from the local convergence properties of the iterative sequence. Computational results are obtained which demonstrate that the total number of iterations to meet a given stopping criterion will be reduced substantially by the new step size, with the most dramatic results being observed for values of p close to 1.

Keywords: Single facility; Minisum location problem; l_p norm; Convergence

1. Introduction

The Fermat–Weber problem concerns the siting of a new facility on the plane in order to minimize the sum of weighted distances to a set of fixed points. For the case where distances are given by the l_p norm, the problem may be formulated as follows:

$$\min W(x) = \sum_{i=1}^{n} w_i l_p(x - a_i), \tag{1}$$

where n is the number of fixed points (also referred to as customers); $a_i = (a_{i1}, a_{i2})^{\mathrm{T}}$ is the known position of customer i; $i = 1, \ldots, n$; $x = (x_1, x_2)^{\mathrm{T}}$ is the unknown position of the new facility; w_i is a positive weighting constant which converts distance travelled between the new facility and customer i into a cost, $i = 1, \ldots, n$; and the distance between any $x, a_i \in \mathscr{R}^2$ is given by

$$l_p(x - a_i) = [|x_1 - a_{i1}|^p + |x_2 - a_{i2}|^p]^{1/p}, \quad p \geqslant 1. \tag{2}$$

The restriction on the parameter p to values greater than or equal to unity ensures that $l_p(\cdot)$ has the properties of a norm. In a practical setting, the factor w_i would be proportional to the demand at customer i, and the term $w_i l_p(x - a_i)$ would be an estimate of the cost of serving the demand at customer i by the facility at x. Thus, the objective in Eq. (1) is to minimize the total service (or distribution) cost, $W(x)$.

* Corresponding author. Fax: 972 9 9561213; e-mail: chenr@ccsg.tau.ac.il.

0167-6377/98/$19.00
PII: S0167-6377(98)00016-9

Problem (1) is a basic model in continuous location theory which has received much attention in the literature. For a detailed review, see for example, [14], Ch. 2 and [19]. The most popular method of solution of the unconstrained problem is given by a one-point iterative procedure which was first proposed for Euclidean distances ($p=2$) by Weiszfeld [18]. The procedure is readily generalized to l_p distances (e.g., [14], Ch. 2).

Despite the long history of the single-facility minisum location problem, there still appears to be much to say about it as evidenced by several recently published articles on this topic. For example, Brimberg and Love [4, 5] discuss local and global convergence properties of the generalized Weiszfeld procedure for l_p distances. These results are extended by Frenk et al. [11] under more general assumptions of quasiconvexity of the objective function. However, these authors use a hyperbolic approximation of the l_p norm in order to avoid the problem of singularities in the iteration functions. The issue of singularities is treated in [1, 6] for Euclidean distances, and Brimberg and Chen [3] for l_p distances.

The notion of accelerating the convergence of descent methods such as the Weiszfeld procedure requires an investigation of alternate step sizes (e.g. [8]). A first attempt at accelerating the Weiszfeld procedure may be attributed to Katz [13], who suggested the use of Steffensen's iteration. This method is not known to be globally convergent, but it may be used to accelerate the local convergence rate of the Weiszfeld procedure from linear to quadratic. However, the gain in convergence rate may be offset by the larger computational time of the Steffensen iterations. Drezner [10] applies a factor (λ) to multiply the step size of the Weiszfeld procedure for the case of Euclidean distances ($p=2$). It is known that convergence will occur for a step size not exceeding double the original ([16]). However, Drezner uses an approximation function of $W(x)$ to calculate a λ at each iteration which may take on values considerably greater than 2. Computational results reported by Drezner appear to indicate, in general, only a modest reduction in the number of iterations using his acceleration method as compared with a fixed λ of 2. The net effect on running time is unclear due to the increased computations associated with the variable step size factor.

Frenk et al. [11] recommend the use of a step size factor between 0 and 2 for the generalization to l_p distances, after showing that the descent property is still maintained for all $p \in [1,2]$. The authors note that the optimal λ must be determined empirically, but do not pursue this direction. Our objective then is to consider the effect of various step sizes on the Weiszfeld procedure applied to problem (1). To the best of our knowledge, this is the first attempt to accelerate the Weiszfeld algorithm for l_p distances.

In the next section, the local convergence properties of the iterative sequence are examined when the original step size is multiplied by a fixed factor λ. This leads to a formula for λ as a simple monotonically decreasing function of p. An interesting outcome of this analysis is that as the parameter p decreases to a value of 1, λ may be increased to $+\infty$ in the limit with local convergence of the sequence maintained. Furthermore, the new step-size factor guarantees local convergence for all $p>1$, a condition not generally true in the original method with $p>2$ (see [4, 5, 11]).

The proposed step size is investigated with several others in an empirical study reported in the subsequent section. The main conclusion drawn from the empirical results is that significant savings in computation time may be obtained using a fixed step-size factor, and further, these savings tend to increase dramatically as p is decreased towards 1.

2. Analysis of local convergence

In this section, we study a revised Weiszfeld procedure where the original step size is multiplied by a constant λ. The effect on local convergence of the new parameter λ is investigated. As noted previously, the analysis is restricted to problems in the plane ($\mathscr{R}^2$), which is typically the case of practical interest. We also assume for the purposes of this analysis that the optimal solution of Eq. (1) occurs at a nonsingular point of the iteration functions. (The singularities of the iteration functions may be avoided alternatively by introducing a smoothing function such as the hyperbolic approximation of the l_p norm. (e.g., see [14]. For a detailed discussion of the singularities, see Brimberg and Love [4, 5].)

The Weiszfeld procedure is derived by setting the first-order partial derivatives of $W(x)$ to zero. For l_p

distances we obtain the following one-point iterative scheme (e.g., see [4, 5]):

$$x_t^{q+1} = \frac{\sum_{i=1}^n w_i |x_t^q - a_{it}|^{p-2} a_{it} / [l_p(x^q - a_i)]^{p-1}}{\sum_{i=1}^n w_i |x_t^q - a_{it}|^{p-2} / [l_p(x^q - a_i)]^{p-1}}, \quad t = 1, 2, \tag{3}$$

where the superscript $q = 0, 1, 2, \ldots$, denotes the iteration number, and the starting point x^0 may be any point in $\mathscr{R}^2$. Introducing the following notation:

$$y_{it}(x) = w_i |x_t - a_{it}|^{p-2} / [l_p(x - a_i)]^{p-1}, \quad \forall i, t, \tag{4}$$

$$S_t(x) = \sum_{i=1}^n y_{it}(x), \quad \forall t \tag{5}$$

and letting $\nabla_t := \partial / \partial x_t$, we can rewrite Eq. (3) in an equivalent form,

$$x_t^{q+1} = x_t^q - \frac{1}{S_t(x^q)} \nabla_t W(x^q), \quad t = 1, 2. \tag{6}$$

Thus, the Weiszfeld procedure for l_p distances is seen to be a modified gradient descent method with predetermined step size.

A more general scheme introduces a step size factor $\lambda > 0$ to obtain

$$x_t^{q+1} = x_t^q - \frac{\lambda}{S_t(x^q)} \nabla_t W(x^q), \quad t = 1, 2. \tag{7}$$

The effect of the new step size on the local convergence properties of the iteration sequence is examined next. The aim of this analysis is to determine a suitable range of values of λ to use in an empirical study. The analysis extends the local convergence results of [4] for the original step size ($\lambda = 1$).

The modified scheme in Eq. (7) may be rewritten as follows,

$$J_t(x) = \Phi_t(x) - \frac{(\lambda - 1)}{S_t(x)} \nabla_t W(x), \quad t = 1, 2, \tag{8}$$

where $\Phi_t(x) = x_t - \nabla_t W(x) / S_t(x)$ is the original iteration function for the x_t ordinate defined in Eq. (6) (or Eq. (3)), and $J_t(x)$ is the new iteration function. Let $J(x) = (J_1(x), J_2(x))$ denote the vector of iteration functions given by Eq. (8). To study the local convergence properties of the new step size, we need to examine the Jacobian matrix of first-order partial derivatives of $J(x)$. Element (t, k) of this matrix is given by

$$\nabla_k J_t(x) = \nabla_k \Phi_t(x) - (\lambda - 1) S_t^{-1}(x) \nabla_{kt}^2 W(x) - (\lambda - 1) \nabla_k (S_t^{-1}(x)) \nabla_t W(x), \quad \forall k, t, \tag{9}$$

where ∇_{kt}^2 denotes the second-order partial derivative with respect to x_k and x_t. However, at the optimal solution x^* we have

$$\nabla_k \Phi_t(x^*) = -S_t^{-1}(x^*) \nabla_{kt}^2 W(x^*), \quad \forall k, t, k \neq t, \tag{10}$$

$$\nabla_t \Phi_t(x^*) = 1 - S_t^{-1}(x^*) \nabla_{tt}^2 W(x^*), \quad \forall t, \tag{11}$$

[4], and furthermore,

$$\nabla_t W(x^*) = 0, \quad \forall t. \tag{12}$$

Combining Eqs. (9)–(12), it follows that the Jacobian matrix evaluated at x^* is given by

$$J'(x^*) = I - \lambda B(x^*), \tag{13}$$

where I is a 2×2 identity matrix, and

$$B(x^*) = \begin{bmatrix} \frac{1}{S_1(x^*)} \frac{\partial^2 W(x^*)}{\partial x_1^2} & \frac{1}{S_1(x^*)} \frac{\partial^2 W(x^*)}{\partial x_2 \partial x_1} \\ \frac{1}{S_2(x^*)} \frac{\partial^2 W(x^*)}{\partial x_1 \partial x_2} & \frac{1}{S_2(x^*)} \frac{\partial^2 W(x^*)}{\partial x_2^2} \end{bmatrix}. \tag{14}$$

The matrix $B(x^*)$ resembles the Hessian matrix of the objective function evaluated at x^*, namely, $H(x^*)$. Since the optimal solution is assumed to occur at a nonsingular point of the iteration functions, it readily follows that $H(x^*)$ and $B(x^*)$ are both well defined. It is well known that if $p > 1$ and the fixed points a_i are noncollinear, then W is a strictly convex function of x. Furthermore, except for the collinear case and a special geometry when $p > 2$, the Hessian matrix can be shown to be positive definite $\forall x$ where it is defined (see [2]). More specifically, the following are assumed to apply in order to ensure the positive definiteness of $H(x^*)$:

Condition 1: $1 < p \leqslant 2$, and the fixed points $a_1, \ldots, a_n$ are noncollinear;

Condition 2: $p > 2$, and at least two indices $r, s \in \{1, \ldots, n\}$ exist such that $x_1^* \notin \{a_{r1}, a_{s1}\}$, $x_2^* \notin \{a_{r2}, a_{s2}\}$, and a_r, a_s, x^* are noncollinear.

The positive definiteness of $H(x^*)$ implies that the eigenvalues of $B(x^*)$ are all positive-valued (e.g., see [4] for a proof). From Eq. (13) it follows that

$$v_j = 1 - \lambda u_j, \quad j = 1, 2, \tag{15}$$

where u_1, u_2 are the eigenvalues of $B(x^*)$ and v_1, v_2 are the eigenvalues of $J'(x^*)$. Also, recalling that $\lambda > 0$, we conclude that

$$v_j < 1, \quad \forall j. \tag{16}$$

Relation (16) provides the following result for location in the plane.

Property 1. *Let* $\mathrm{tr}[J'(x^*)]$ *denote the trace of* $J'(x^*)$. *Then a sufficient condition for local convergence at* x^* *of the modified iteration procedure* $J: x \to J(x)$ *is given by*

$$\mathrm{tr}[J'(x^*)] \geqslant 0. \tag{17}$$

Proof. Since $\mathrm{tr}[J'(x^*)] = v_1 + v_2$, we obtain from Eq. (17), $v_1 \geqslant -v_2$. But $v_1 < 1$ by Eq. (16). It immediately follows that $v_2 > -1$. Likewise we conclude that v_1 is also > -1. Thus, the spectral radius of $J'(x^*)$, $\rho = \max_j |v_j| < 1$, and local convergence is guaranteed (e.g., see [9], or [15]). Furthermore, the local convergence rate will be linear or better, since ρ equals the upper asymptotic convergence bound. □

Using the derivatives calculated in [4], it is readily shown that

$$\mathrm{tr}[J'(x^*)] = 2 - \lambda(p-1)(2-Q^*), \tag{18}$$

where

$$Q^* = \sum_{t=1}^{2} \frac{1}{S_t(x^*)} \sum_{i=1}^{n} \frac{y_{it}(x^*)|x_t^* - a_{it}|^p}{[l_p(x^* - a_i)]^p}. \tag{19}$$

It is clear that $Q^* > 0$. Also, observe that

$$Q^* < \sum_{t=1}^{2} \frac{1}{S_t(x^*)} \sum_{i=1}^{n} y_{it}(x^*) = 2, \tag{20}$$

(where condition 2 is invoked for $p > 2$). Thus, $Q^* \in (0,2)$, $2 - Q^* > 0$, and we conclude from Eq. (17) and (18) that local convergence is guaranteed by the following upper bound on λ:

$$\lambda \leqslant \frac{2}{(p-1)(2-Q^*)}. \tag{21}$$

Relation (21) is of little practical use, since x^* is unknown and hence, Q^* cannot be evaluated beforehand. However, a more conservative bound may be imposed on λ as shown by the following result.

Property 2. *Local convergence at* x^* *is guaranteed if the step size factor* $\lambda \leqslant 1/(p-1)$.

Proof. Since $0 < Q^* < 2$, it follows that $1/(2-Q^*) > \frac{1}{2}$. Thus, $\lambda \leqslant 1/(p-1)$ implies that the upper bound in Eq. (21) is strictly satisfied. We conclude that local convergence will occur for any $\lambda \in (0, 1/(p-1)]$. □

Property 2 provides an interesting insight on local convergence for different values of the parameter p. We see that for p close to 1, the step size may be increased substantially and still maintain local convergence. For example, if $p = 1.1$, the step size may be augmented by a factor of 10. Sluggish convergence rates, approaching sublinear, have been observed with the original step size for p values close to 1 (see [4]). Thus, we may anticipate substantial improvements in the convergence rate when large values of λ are used.

The bound $1/(p-1)$ suggests a factor $\lambda > 1$ be used when p is between 1 and 2, and a $\lambda < 1$ when $p > 2$. Brimberg and Love [4, 5] show that convergence with the original step size ($\lambda = 1$) is not guaranteed for sufficiently large values of p exceeding 2. Property 2 now provides a step-size factor as a simple monotonic function of p which guarantees at least local convergence for all values of p.

As a final observation, we note that for $p = 2$,

$$\begin{aligned} Q^* &= \sum_{t=1}^{2} \left(\sum_{j=1}^{n} \frac{w_j}{l_2(x^* - a_j)} \right)^{-1} \\ &\quad \times \sum_{i=1}^{n} \frac{w_i}{l_2(x^* - a_i)} \frac{(x_t^* - a_{it})^2}{[l_2(x^* - a_i)]^2} \\ &= \left(\sum_{j=1}^{n} \frac{w_j}{l_2(x^* - a_j)} \right)^{-1} \\ &\quad \times \sum_{i=1}^{n} \frac{w_i}{[l_2(x^* - a_i)]^3} \sum_{t=1}^{2} (x_t^* - a_{it})^2 = 1. \end{aligned}$$

Assuming that $Q^* \approx 1$ for other values of p, we obtain as an approximation of the upper bound in (21),

$$\lambda = 2/(p-1). \tag{22}$$

Alternatively, we might consider as a compromise a step-size factor,

$$\lambda = p/(p-1), \tag{23}$$

which falls between the two bounds suggested above for $1 < p \leqslant 2$.

3. Computational results and discussion

In this section we present some empirical results of the numerical solution of the l_p problem. Typically, problems were of 10, 30, 100 and 1000 demand points distributed randomly, using a uniform distribution, over a 100×100 square. The weights (w_i) associated with the demand points were equal in some cases, and randomly chosen between 1 and 100 in the others. The pseudo-random numbers were generated using a built-in standard method. The idea was to repeat the computations for a given set of chosen points with different step sizes, and change at will the set of points and weights by changing the *seed* in the generating program, once we wished to make a different set of computational experiments with varying step size. We also solved problems on a 100×500 rectangle, as well as problems with demand points arranged in clusters which themselves were distributed over a square. No significant differences were observed between the different geometries, except that in the weighted problems, a coincidence between a demand point and the optimal facility location was more common.

This possibility of coincidence deserves a further discussion. In these cases, the algorithm with the original step size converged, though relatively slowly, whereas with larger step sizes, typically, no convergence was reached. A common behavior in these cases was that after a number of steps, there was a hopping between two attraction points around the optimal solution. Hopping among the neighborhoods of two or more attraction points results when the descent property is violated by too large step size. This situation of oscillations between several nonoptimal points rather than convergence to the optimum is common in other iterative procedures such as Newton's method. (For an illustration, see p. 93 in [17]. Also, [5] for an example with the standard Weiszfeld procedure and $p > 2$.) A remedy to the preceding problem is given, however, by Juel and Love [12] who showed that a coincidence of the optimally located facility with a demand point r with coordinates (a_{r1}, a_{r2}) will occur if, and only if,

$$(|R_{r1}|^{p/(p-1)} + |R_{r2}|^{p/(p-1)})^{(p-1)/p} \leqslant w_r, \tag{24}$$

where

$$R_{rk} = \sum_{j=1,\, j\neq r}^{n} \frac{w_j \cdot sign(a_{rk} - a_{jk})|a_{rk} - a_{jk}|^{p-1}}{(l_p(a_r - a_j))^{p-1}}, \quad k = 1, 2. \tag{25}$$

In cases where no convergence is reached, it is recommended to check the demand points for optimality using Eqs. (24) and (25), or simply revert to the original step size.

The sequence of iterations was terminated when two successive iterates were within an ε-distance of each other. This is a standard stopping criterion for such procedures. We specified an ε of 10^{-7}, which was considered sufficiently precise for our purposes. Alternate stopping criteria may use the difference between two successive values of the objective function, or an acceptable deviation from a calculated bound on the optimal value of the objective function.

It is to be noted that in most of the problems where the optimal solution does not coincide with a demand point, the saving in the number of iterations is substantial, whereas a negligible computational effort is required for changing the step size in the direction determined by the original method. Table 1 shows a set of results for a sample problem of 100 demand points distributed over a 100×100 square with equal weights and with different values of p varying between 1.2 and 10.0.

Unfortunately, for $p = 1.1$, nonconvergence occurred for all but the original step size. Once again, an oscillation between two attraction points was observed for the larger step sizes even though the optimal facility location did not coincide with a demand point. This illustrates a second type of ill-conditioned case which takes place when the optimal facility location has one or more of its coordinates nearly coinciding with a coordinate of a demand point. Such a situation normally exists as p approaches a value of 1.0. However, it turns out that if the iteration process starts in these cases from a point sufficiently close to the optimal solution, convergence will occur for the larger step sizes. Moreover, either $\lambda = 1/(p-1)$ or $p/(p-1)$ accelerates the convergence substantially.

Table 1
The number of iterations required for reaching the optimal solution within a given stopping criterion specified in the text for different values of p, and for different step size multipliers λ: (1) $\lambda = 1$; (2) $\lambda = 2$; (3) $\lambda = 1/(p-1)$; (4) $\lambda = p/(p-1)$; (5) $\lambda = Q$

	Time(s)					No. of iterations				
p	(1)	(2)	(3)	(4)	(5)	(1)	(2)	(3)	(4)	(5)
1.2	4.45	2.20	0.71	0.55	0.99	105	52	17	12	9
1.3	3.03	1.42	0.77	0.50	0.93	72	34	18	12	9
1.4	2.25	0.99	0.72	0.44	0.93	53	24	18	11	9
1.5	1.59	0.71	0.72	0.38	0.83	38	17	17	9	9
1.6	1.54	0.65	0.82	0.39	1.09	37	16	20	9	10
1.7	1.43	0.61	0.94	0.44	1.10	35	14	23	10	10
1.8	1.32	0.50	1.05	0.39	1.05	32	12	25	10	10
1.9	1.27	0.44	1.15	0.44	1.15	30	11	27	10	11
2.0	0.33	0.11	0.38	0.11	0.33	29	10	29	10	11
2.5	0.99	0.93	1.64	0.49	1.04	24	23	40	12	12
3.0	0.33	0.55	0.71	0.16	0.39	22	42	50	13	13
3.5	0.88	2.91	2.53	0.55	1.15	21	70	61	13	13
4.0	0.28	1.48	1.04	0.16	0.44	21	106	73	13	13
5.0	0.27	2.52	1.38	0.22	0.44	20	177	97	14	14
7.0	0.33	3.24	2.25	0.22	0.50	22	222	151	17	15
10.0	0.39	4.95	3.79	0.33	0.60	26	336	259	23	17

This behavior should not be too surprising since, after all, the proof given above concerns only *local* convergence. One may even be surprised about the empirical result that in most cases, convergence is reached so quickly even while starting from very far away. Since the preceding case cannot be identified at the outset, the only practical solution is that in the rare cases of no convergence with the larger step size, and no coincidence between a demand point and the optimal solution, one should resort to the original, shorter step size.

Table 1 represents a typical example of the run times and the number of iterations required to reach the given stopping criterion. The different step-size factors used are indicated in the caption. With the exception of column (5), the number of iterations and run times are proportional to each other since the extra computational effort of multiplying the step size by a constant is minimal. As for column (5), even when the number of steps was the smallest, the run times were not, due to the extra computational effort required to evaluate Q (formula (19) with the current iterate x^q replacing x^*). The unusually low run times found with $p = 2$ and other integer values of p can be ascribed to the more efficient manner which is used by the computer for raising numbers to integer powers. It is also to be noted that in all the cases reported in the table, the step size multiplier of $p/(p-1)$ turned out to be more efficient than $1/(p-1)$; this will be discussed below. The $\lambda = 1/(p-1)$ factor was consistently better than the original factor of $\lambda = 1$ for p values up to 2, whereas as compared to the factor of $\lambda = 2$, it was superior only up to $p = 1.5$. Also, note that for the smaller values of p, the number of iterations increases substantially for step size multipliers (1), (2), whereas the number of iterations remains relatively stable for the others.

For values of $p > 2$, the use of $\lambda = 1/(p-1)$ was found in most cases to be inferior to the original step size with $\lambda = 1$. The empirical use of $\lambda = p/(p-1)$ turned out to be better, but of course, when p gets very large, this advantage diminishes as $p/(p-1)$ approaches unity. In a few isolated cases, the original step size ($\lambda = 1$) failed to converge. This is not surprising, since global convergence is only guaranteed for $p \in [1, 2]$ (see [4, 5]). It should, however, be noted that the smaller step size obtained with $\lambda = 1/(p-1)$ managed to converge in all problems tested having $p > 2$. This corroborates the local convergence results of the preceding section.

A conjecture can be made that a step-size multiplier of $\lambda = p/(p-1)$ may be an optimal one. This is based on the fact that it turned out to be the best overall in our numerous test problems, and that for $p = 2$, the optimal step-size multiplier is $\lambda = 2$ as suggested by Ostresh [16] and further discussed in a broader context by Chen [7].

In the numerical examples, two classes of hard-to-solve problems were identified. One is where the optimal solution coincides with a demand point. This difficulty can easily be overcome, for example, by using Eq. (24) to identify the optimal point, in which case one would not have to resort to the iterative procedure. A more troublesome situation has been found in which one of the coordinates of the solution nearly coincides with the corresponding coordinate of one of the demand points. In these cases, only starting the iterations from a point close to the optimal solution results in convergence. It is to be noted that once a close enough point is selected, the local convergence is still substantially accelerated by using $\lambda = 1/(p-1)$ and even more so with $\lambda = p/(p-1)$.

In spite of the difficulties mentioned here, the reduction, which becomes increasingly dramatic for smaller p, in the number of iterations and the run times as compared to the original method, has to be considered seriously. This is especially important in the multi-source location–allocation problem, where the single facility iterative process may be repeated many, sometimes several thousands of, times.

Finally, for future research we can think of two key issues.

1. To prove the conjecture that $\lambda = p/(p-1)$ is optimal and to find its advantages and limitations for local and global convergence.
2. To try to identify the problematic cases of the second kind, namely, cases in which there is a near coincidence of one of the coordinates, or to devise a method to solve these cases efficiently. One option to consider is to use a sliding step-size multiplier, starting, for $1 < p \leqslant 2$, from $\lambda = 1$ and increasing toward $p/(p-1)$ as the iterations approach the optimal solution.

References

[1] J. Brimberg, The Fermat–Weber location problem revisited, Math. Programming 71 (1995) 71–76.

[2] J. Brimberg, R.F. Love, Local convexity results in a generalized Fermat–Weber problem (1998), submitted for publication.

[3] J. Brimberg, R. Chen, A note on convergence in the single facility minisum location problem, Comput. Math. Appl. 35 (1998) 25–31.

[4] J. Brimberg, R.F. Love, Local convergence in a generalized Fermat–Weber problem, Ann. Oper. Res. 40 (1992) 33–66.

[5] J. Brimberg, R.F. Love, Global convergence in a generalized iterative procedure for the minisum location problem with l_p distances, Oper. Res. 41 (1993) 1153–1163.

[6] R. Chandrasekaran, A. Tamir, Open questions concerning Weiszfeld's algorithm for the Fermat–Weber location problem, Math. Programming 44 (1989) 293–295.

[7] R. Chen, Solution of location problems with radial cost functions, Comput. Math. Appl. 10 (1984) 87–94.

[8] A.I. Cohen, Stepsize analysis for descent methods, J. Optim. Theory Appl. 33 (1981) 187–205.

[9] G. Dahlquist, Å. Björck, Numerical Methods (transl. by N. Anderson), Prentice-Hall, Englewood Cliffs, NJ, 1974.

[10] Z. Drezner, A note on the Weber location problem, Ann. Oper. Res. 40 (1992) 153–161.

[11] J.B.G. Frenk, M.T. Melo, S. Zhang, A Weiszfeld method for a generalized l_p distance minisum location model in continuous space, Location Sci. 2 (1994) 111–127.

[12] H. Juel, R.F. Love, Fixed point optimality criteria for the location problem with arbitrary norms, J. Oper. Res. Soc. 32 (1981) 891–897.

[13] I.N. Katz, Local convergence in Fermat's problem, Math. Programming 6 (1974) 89–104.

[14] R.F. Love, J.G. Morris, G.O. Wesolowsky, Facilities Location: Models and Methods, North-Holland, New York, 1988.

[15] J.M. Ortega, W.C. Rheinboldt, Iterative Solutions of Numerical Equations in Several Variables, Academic Press, New York, 1970.

[16] L.M. Ostresh, On the convergence of a class of iterative methods for solving the Weber problem, Oper. Res. 26 (1978) 597–609.

[17] A.L. Peressini, F.E. Sullivan, J.J. Uhl Jr., The Mathematics of Nonlinear Programming, Springer, New York, 1988.

[18] E. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnés est minimum, Tôhoku Math. J. 43 (1937) 355–386.

[19] G.O. Wesolowsky, The Weber problem: history and perspectives, Location Sci. 1 (1993) 5–23.