

SOLUTION OF LOCATION PROBLEMS WITH RADIAL COST FUNCTIONS

REUVEN CHEN

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

(Received April 1983)

Communicated by I. Norman Katz

Abstract—The solution of the Cooper location problem $\min_{x,y} \sum_{i=1}^m w_i r_i^n$ where r_i is the radial (Euclidean) distance between the i th given location (a_i, b_i) and the center (x, y) to be located is further investigated. The iterative method given by Cooper (which includes the well known Weiszfeld procedure for $n = 1$) was previously amended using semi-intuitive arguments. In the present work a better proof is offered for the results given before. Furthermore, using the same line of argument, a broader group of problems previously mentioned by Katz and others can be efficiently solved. These are the problems $\min_{x,y} \sum_{i=1}^m \phi_i(r_i)$ where ϕ_i are non-decreasing functions of the Euclidean distances. The method is also extended to solve similar problems in E^K with $K > 2$. Apart from the theoretical account, computational experience is reported for the three dimensional Cooper problem with different values of n . Computational results of the $\min_{x,y} \sum_{i=1}^m \exp(\alpha w_i r_i - C)$ which is a different member of the Katz class of problems, are also presented.

1. INTRODUCTION

The Weber problem of optimally locating a service facility with respect to the location of a number of demand points in E^2 , was given its first iterative solution by Weiszfeld [1]. This problem, in which costs are assumed to be proportional to the Euclidean distances, was further studied by a number of authors [2-5]. Cooper [6] extended the Weber problem by posing the problem of minimizing the weighted sums of powers of the Euclidean distances

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

where $r_i = [(x - a_i)^2 + (y - b_i)^2]^{1/2}$ and where (a_i, b_i) , $i = 1, \dots, m$ are the given fixed points having weights of w_i , and (x, y) the variable point in E^2 to be located optimally. Cooper pointed out that $n > 1$ represents the case of "diseconomies of scale" whereas $n < 1$ is that of "economies of scale". $n = 1$ is, of course, the original Weber problem. Cooper also showed that for $n \geq 1$, the problem is convex and therefore, a local minimum reached by any iterative process must be global. When $n < 1$, the objective function is not convex and therefore several local minima may exist. Cooper [6] suggested an iterative process to solve problem (1) which was an extension to Weiszfeld's method. Starting from any point in the convex hull of the given points (a_i, b_i) , usually from the "center of gravity"

$$x^0 = \frac{\sum_{i=1}^m w_i a_i}{\sum_{i=1}^m w_i}; \quad y^0 = \frac{\sum_{i=1}^m w_i b_i}{\sum_{i=1}^m w_i} \quad (2)$$

Cooper's N th step is given by

$$x^{N+1} = \frac{\sum_{i=1}^m w_i a_i [r_i(x^N, y^N)]^{n-2}}{\sum_{i=1}^m w_i [r_i(x^N, y^N)]^{n-2}} \quad (3a)$$

$$y^{N+1} = \frac{\sum_{i=1}^m w_i b_i [r_i(x^N, y^N)]^{n-2}}{\sum_{i=1}^m w_i [r_i(x^N, y^N)]^{n-2}}, \quad (3b)$$

whereas Weiszfeld's step is the special case in which $n = 1$. Other authors (e.g. [4, 7])

showed that Weiszfeld's iterative process is just the steepest descent method with a step-size determined by the denominator $\sum_{i=1}^m w_i r_i^{-1}$. Chen[8] showed that in the more general case of n not being necessarily unity, Cooper's iterative step is also along the steepest descent direction, the N th step is given by

$$x^{N+1} - x^N = \sum_{i=1}^m w_i (a_i - x^N) [r_i(x^N, y^N)]^{n-2} \bigg/ \sum_{i=1}^m w_i [r_i(x^N, y^N)]^{n-2} \quad (4a)$$

$$y^{N+1} - y^N = \sum_{i=1}^m w_i (b_i - y^N) r_i^{n-2} \bigg/ \sum_{i=1}^m w_i r_i^{n-2}. \quad (4b)$$

Again, the step-size is determined by the denominator $\sum_{i=1}^m w_i r_i^{n-2}$. Cooper[6] solved his problem with values of n of up to 3. In an attempt to approximate the minimax solution, Chen[8] tried to solve problem (1) with large values of n and found that for $n > 3$, the process usually did not converge at all. Empirical results as well as a "semi-intuitive" proof showed that Cooper's step-size is too large for high values of n and Chen[8] showed that convergence occurs even for very high values of n if equations [4] are replaced by

$$x^{N+1} = x^N + (2/n) \sum_{i=1}^m w_i (a_i - x^N) r_i^{n-2} \bigg/ \sum_{i=1}^m w_i r_i^{n-2} \quad (5a)$$

$$y^{N+1} = y^N + (2/n) \sum_{i=1}^m w_i (b_i - y^N) r_i^{n-2} \bigg/ \sum_{i=1}^m w_i r_i^{n-2}. \quad (5b)$$

For large values of n this reduces substantially (by a factor of $2/n$) the Cooper step-size whereas for the original Weber problem ($n = 1$) this doubles the step-size. This particular result (for $n = 1$) is in accord with a statement by Ostresh[9] to the same effect. The proof given by Chen[8] was based on the properties of the Newton step (without search). It was shown that the Hessian matrix is nearly diagonal and that the diagonal elements are approximately equal to each other under the assumption that the given points are distributed over a given area. One of the purposes of the present work is to provide a more accurate proof for the result expressed by equation (5) which removes many of the restrictions in the previous work. Katz[10] and Cordellier and Fiorot[11] extended the Cooper problem so that more general cost functions are included. The objective function is now written as

$$\min_{x, y} f(x, y) = \sum_{i=1}^m \phi_i(r_i) \quad (6)$$

where the ϕ_i 's are "radial" cost functions, i.e. they depend on the Euclidean distances r_i between (x, y) which is to be optimally located and the given points (a_i, b_i) , rather than in any other way on (x, y) . These authors showed that $f(x, y)$ is convex if ϕ_i are convex, differentiable and nondecreasing, which ensures that a local minimum is necessarily global. This group of functions includes the Cooper case in which $\phi_i(r_i) = w_i r_i^n$ with $n > 1$, as well as many other convex increasing functions. In analogy with the Weiszfeld[1] and Cooper[6] iterative methods, Katz suggested the following iterative procedure for problem (6)

$$x^{N+1} = x^N + \sum_{i=1}^m [\phi_i'(r_i^N)(a_i - x^N)/r_i^N] \bigg/ \sum_{i=1}^m [\phi_i'(r_i^N)/r_i^N] \quad (7a)$$

$$y^{N+1} = y^N + \sum_{i=1}^m [\phi_i'(r_i^N)(b_i - y^N)/r_i^N] \bigg/ \sum_{i=1}^m [\phi_i'(r_i^N)/r_i^N]. \quad (7b)$$

This also is a steepest descent iteration with the step-size determined by the denominator in equation (7). A better step-size will be suggested here which includes the second derivative $d^2\phi_i/dr_i^2$. Finally, the solution of three dimensional and multi-dimensional location problems with cost functions similar to those mentioned above will also be briefly mentioned.

2. PROOF OF THE AMENDED COOPER STEP-SIZE

As mentioned above, Chen[8] found the amended Cooper process expressed by equations (5) empirically, and supported it by a "semi-intuitive" proof based on the properties of the Hessian. The proof that equations (5) constitute the best step-size to be taken is changed here, based on the diagonal terms of the Hessian only. Following some early works[12-14], Cohen[15] has recently summed up the matter of rate of convergence of descent methods. For the case of steepest descent he showed that if we write the i th iterative step as

$$x^{N+1} = x^N - \theta^N \nabla f^N \quad (8)$$

where $\theta^N \in [\delta, 2/\lambda_{\max} - \delta]$, $\delta \in (0, 1/\lambda_{\max})$ and λ_{\max} is the largest eigenvalue of the Hessian at x^N , then the descent method converges linearly. The best estimate of the convergence ratio occurs when

$$\theta^N = 2/(\lambda_{\max} + \lambda_{\min}) \quad (9)$$

where λ_{\min} is the smallest eigenvalue of the Hessian, the resultant ratio being

$$q = (\lambda_{\max} - \lambda_{\min})/(\lambda_{\max} + \lambda_{\min}). \quad (10)$$

In the present location problem which is two dimensional, the Hessian has, of course only two eigenvalues, namely, λ_{\min} and λ_{\max} . Since the trace of a matrix is invariant, we can write for these problems

$$\lambda_{\min} + \lambda_{\max} = \partial^2 f / \partial x^2 + \partial^2 f / \partial y^2. \quad (11)$$

The first and second derivatives of (1) as given by Cooper[6] are

$$\partial f / \partial x = n \sum_{i=1}^m w_i r_i^{n-2} (x - a_i) \quad (12a)$$

$$\partial f / \partial y = n \sum_{i=1}^m w_i r_i^{n-2} (y - b_i) \quad (12b)$$

and

$$\partial^2 f / \partial x^2 = n \sum_{i=1}^m w_i r_i^{n-2} + n(n-2) \sum_{i=1}^m w_i r_i^{n-4} (x - a_i)^2 \quad (13a)$$

$$\partial^2 f / \partial y^2 = n \sum_{i=1}^m w_i r_i^{n-2} + n(n-2) \sum_{i=1}^m w_i r_i^{n-4} (y - b_i)^2. \quad (13b)$$

Adding (13a) and (13b) yields

$$\partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 = 2n \sum_{i=1}^m w_i r_i^{n-2} + n(n-2) \sum_{i=1}^m w_i r_i^{n-4} [(x - a_i)^2 + (y - b_i)^2]. \quad (14)$$

However, since $r_i^2 = (x - a_i)^2 + (y - b_i)^2$ we have

$$\partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 = [2n + n(n-2)] \sum_{i=1}^m w_i r_i^{n-2} = n^2 \sum_{i=1}^m w_i r_i^{n-2}. \quad (15)$$

By using equations (8), (9), (11), (12) and (15) we get the step-size given by equations (5) as the one giving the best convergence ratio according to Cohen[15]. Thus, the step-size found empirically by Chen[8] to be the best, is now shown to be optimal in this sense. It is to be mentioned, that the semi-intuitive proof given by Chen[8], showing that the Hessian is "nearly", diagonal with diagonal terms nearly equal to one another indicate that the two eigenvalues are rather close to one another. This is so since the eigenvalues of a matrix are continuous functions of the elements of the matrix[16] and since in the extreme case of a diagonal matrix, the eigenvalues are, of course, the diagonal elements themselves. This would result in the ratio of convergence q in equation (10) being a small number which guarantees a fast linear convergence.

3. EXTENSION TO GENERAL RADIAL COST FUNCTIONS

We would like to extend the expression for the step-size so that it includes the general cost functions given in (6). By differentiating (6) we get

$$\nabla f = \begin{pmatrix} \sum_{i=1}^m (d\phi_i/dr_i)(x - a_i)/r_i \\ \sum_{i=1}^m (d\phi_i/dr_i)(y - b_i)/r_i \end{pmatrix}. \quad (16)$$

Second differentiation yields

$$\partial^2 f / \partial x^2 = \sum_{i=1}^m [(d^2\phi_i/dr_i^2)(x - a_i)^2/r_i^2 + (d\phi_i/dr_i)(y - b_i)^2/r_i^3] \quad (17a)$$

$$\partial^2 f / \partial y^2 = \sum_{i=1}^m [(d^2\phi_i/dr_i^2)(y - b_i)^2/r_i^2 + (d\phi_i/dr_i)(x - a_i)^2/r_i^3]. \quad (17b)$$

By adding up equations (17) we have

$$\lambda_{\min} + \lambda_{\max} = \partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 = \sum_{i=1}^m [(d^2\phi_i/dr_i^2) + (d\phi_i/dr_i)/r_i]. \quad (18)$$

The best step-size, which is the extension of equation (5) is given now by

$$[2/(\lambda_{\min} + \lambda_{\max})]\nabla f = \left\{ 2 \left/ \sum_{i=1}^m [d^2\phi_i/dr_i^2 + (d\phi_i/dr_i)/r_i] \right. \right\} \begin{pmatrix} \sum_{i=1}^m (d\phi_i/dr_i)(x - a_i)/r_i \\ \sum_{i=1}^m (d\phi_i/dr_i)(y - b_i)/r_i \end{pmatrix}. \quad (19)$$

Equations (19) differ from those given by Katz[10] (equations 7 above) by the factor 2 as well as by the terms $d^2\phi_i/dr_i^2$. This difference will be of importance in particular in cases where the functions $\phi_i(r_i)$ have large second derivatives. The above example of $\phi_i(r_i) = w_i r_i^n$ with large value of n is such a case. On the other hand, if $n = 1$, the second derivative is certainly nill and therefore the step-size is double that given by Weiszfeld as pointed out by Ostresh[4]. It is to be noted that the arguments mentioned above concerning the fast (linear) convergence due to the closeness of the eigenvalues to one another are valid in the general case (of any $\phi_i(r_i)$) as well.

Another example that comes to mind is $\phi_i(r_i) = \exp(\alpha w_i r_i)$ for large values of α . Computer experiments with the problem

$$\min_{x,y} f(x,y) = \sum_{i=1}^m \exp(\alpha w_i r_i - C) \quad (20)$$

where α and C are constants, have been made since it can rather easily be shown that for

a large enough value of α , (20) is a good approximation to the minimax problem

$$\min_{x,y} \max_i w_i r_i \quad (21)$$

The step to be used according to (19) is

$$\frac{2}{\lambda_{\min} + \lambda_{\max}} \nabla f = 2 \left(\begin{array}{c} \sum_{i=1}^m [w_i \exp(w_i r_i - C)(x - a_i)/r_i] / \sum_{i=1}^m (w_i^2 + w_i/r_i) \exp(w_i r_i - C) \\ \sum_{i=1}^m [w_i \exp(w_i r_i - C)(y - b_i)/r_i] / \sum_{i=1}^m (w_i^2 + w_i/r_i) \exp(w_i r_i - C) \end{array} \right) \quad (22)$$

where the α has been included in the w_i 's for the sake of simplicity. In order to check this result numerically, a given problem has been solved using equation (22) with different constants replacing the 2. Similar to the results reported before [8], for $\phi_i(r_i) = w_i r_i^n$, the smallest number of steps to reach a given termination criterion was found when the value of the constant was indeed ~ 2 . When the constant taken was too large (e.g. 3 or more), the process did not converge at all.

4. THREE AND MULTIDIMENSIONAL PROBLEMS

We try now to extend this method to multidimensional problems, the three dimensional one being the one which is most likely to be of practical use. The problem we would like to solve now is the same as (6), minimization being over x, y, z, \dots and where

$$r_i^2 = (x - a_i)^2 - (y - b_i)^2 + (z - c_i)^2 + \dots \quad (23)$$

Equation (16) should be extended to include a third component (and possibly more) of the gradient,

$$\partial f / \partial z = \sum_{i=1}^m (d\phi_i / dr_i)(z - c_i) / r_i \quad (16c)$$

and similarly more terms if the number of dimensions K is larger than 3. In the second differentiation, equations (17a, b) should be slightly modified and one or more equations should be added as follows:

$$\partial^2 f / \partial x^2 = \sum_{i=1}^m \{ (d^2 \phi_i / dr_i^2)(x - a_i)^2 / r_i + (d\phi_i / dr_i)[r_i^2 - (x - a_i)^2] / r_i^3 \} \quad (24a)$$

$$\partial^2 f / \partial y^2 = \sum_{i=1}^m \{ (d^2 \phi_i / dr_i^2)(y - b_i)^2 / r_i + (d\phi_i / dr_i)[r_i^2 - (y - b_i)^2] / r_i^3 \} \quad (24b)$$

$$\partial^2 f / \partial z^2 = \sum_{i=1}^m \{ (d^2 \phi_i / dr_i^2)(z - c_i)^2 / r_i + (d\phi_i / dr_i)[r_i^2 - (z - c_i)^2] / r_i^3 \} \quad (24c)$$

and similar equations if $K > 3$ dimensions are involved.

According to the discussion of the previous section the best step size in the steepest descent method can be determined by $2/(\lambda_{\min} + \lambda_{\max})$. However, in the multidimensional case it is impossible to find exactly $\lambda_{\min} + \lambda_{\max}$ as was done for the two dimensional problems. We shall postulate here that the average eigenvalue of the Hessian is approximately the same as the average of λ_{\min} and λ_{\max} , namely,

$$(\lambda_1 + \dots + \lambda_K) / K \cong (\lambda_{\min} + \lambda_{\max}) / 2. \quad (25)$$

This cannot be proved, of course, precisely, but the arguments given in [8] and above, namely, that the Hessian is rather close to be diagonal with nearly equal diagonal elements,

strongly support this assertion. An empirical test of the resulting step-size for a concrete problem is given below. Again, since the trace of a matrix is an invariant we can write

$$\lambda_1 + \dots + \lambda_K = \partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 + \partial f / \partial z^2 + \dots \quad (26)$$

Combining (25) and (26) we have

$$2/(\lambda_{\min} + \lambda_{\max}) \cong K/(\lambda_1 + \dots + \lambda_K) = K/(\partial^2 f / \partial x^2 + \partial^2 f / \partial y^2 + \dots), \quad (27)$$

and summing up equations (24a, b, c . . .) we get

$$2/(\lambda_{\min} + \lambda_{\max}) \cong K / \sum_{i=1}^m [d^2 \phi_i / dr_i^2 + (K-1) (d\phi_i / dr_i) / r_i]. \quad (28)$$

This term should now replace the one preceding the gradient vector on the r.h.s. of equation (19). Of course, the gradient vector itself is now K dimensional ($K \geq 3$) rather than the two dimensional one in (19).

Let us consider now the special K dimensional problem in which $\phi_i(r_i) = w_i r_i^n$. The steepest descent step to be taken here is

$$[2/(\lambda_{\min} + \lambda_{\max})] \nabla f \cong \left\{ [K/(n+K-2)] / \sum_{i=1}^m w_i r_i^{n-2} \right\} \nabla f, \quad (29)$$

namely, the original Cooper step-size should be multiplied by $C_T = K/(n+K-2)$. Of course, this reduces back to $2/n$ in the two dimensional case.

In order to check the applicability of this method, including the validity of the postulate (25), three dimensional Cooper problems (equation 1) have been solved. The problem consisted of 100 given points distributed in random over a cube of $100 \times 100 \times 100$. The problem has been solved twice for the same points, once with $w_i = 1$ for all i and once when the weights were also random numbers between 0 and 100. Three values of n have been taken, namely, $n = 1$, $n = 10$ and $n = 100$. The coefficient in equation (29) is expected to be here, for $K = 3$, $C_T = 3/(n+1)$. These problems have been solved, with a given termination criterion and with different values of C replacing the "optimal" C_T . In each case, the iteration started from the center of gravity point (equation (2) extended to three dimensions). The results are shown in Table 1. Each column in the table terminates when such a value of C is taken which results in a divergence of the iterative process. It can be seen in the table that for "moderate" values of n , choose the value of $C = C_T$ reached theoretically above, yields the best results. In the case of very steep functions, exemplified here by $\phi_i(r_i) = w_i r_i^n$ with $n = 100$, divergence occurs at relatively small values of C , and the optimum found here is seen to be $C = 0.8C_T$. The reason for this seems to be that the postulate in equation (25), emanating from the special properties of the Hessian when the cost functions are radial[8], is less accurate when steep functions are concerned. The empirical conclusion is to take a step-size somewhat smaller than C_T , say, $C = 0.8C_T$. Similar results have been found in the solution of an exponential three dimensional problem of the form (20) with rather large value of α ; the details are not given here. As can be seen in the table, taking $C \cong 0.8C_T$ may be a good choice whenever one is in doubt whether the functions in hand are to be considered "steep" or not. The number of iterative steps may increase slightly, but the risk of having divergence decreases substantially.

5. CONCLUSION

Continuous location problems with costs which are non-linear functions of the Euclidean distances have further been investigated in this work. A result pertaining to the best step-size in the solution of Cooper's problem (1), which was previously found empirically and supported in a semi-intuitive way[8], has now been given a more rigorous proof. The same proof has been extended to the problem (6) first by Katz[10] in which

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

C given in units of $C_T = 3/(n+1)$	$n = 1$; w_i random	$C_T = 1.5$ $w_i = 1$	$n = 10$; w_i random	$C_T = 0.273$ $w_i = 1$	$n = 100$; w_i random	$C_T = 0.0297$ $w_i = 1$
0.1	61	61	59	51	228	169
0.2	33	33	32	27	124	93
0.3	22	23	22	19	86	65
0.4	17	17	16	14	67	51
0.5	13	13	13	11	54	41
0.6	10	11	10	9	46	35
0.7	9	9	9	7	40	30
0.8	7	7	7	6	35	26
0.9	6	6	6	5	37	31
1.0	5	5	6	6		431
1.1	6	6	8	8		
1.2	8	8	12	11		
1.3	10	10	18	14		
1.4	13	13	31	20		
1.5	17	18	96	32		
1.6	23	26		69		
1.7	36	44				
1.8	68	119				

the cost elements are $\phi_i(r_i)$, general functions of the Euclidean distances. As mentioned already, the objective function $f(x, y)$ is convex if ϕ_i are convex, differentiable and nondecreasing (e.g. the Cooper problem in which $\phi_i(r_i) = w_i r_i^n$ with $n \geq 1$), and therefore, a minimum reached by any converging iterative process is necessarily a global minimum. However, the iterative procedure developed here is expected to be very efficient in the cases where $f(x, y)$ is not convex as well. The usefulness of this point is, of course, rather limited since in these cases (e.g. $\phi_i(r_i) = w_i r_i^n$ with $n < 1$), the resulting solution is a local minimum and not necessarily global.

The method has also been extended to multidimensional problems of the same kind, of which at least the three dimensional ones seem to be of potential practical importance. The proof for the step-size here is less rigorous since it is based on the postulate that the eigenvalues of the Hessian are not very much different from one another. From empirical results, it appears that this assertion and the resulting step-size are correct for "moderate" functions. For "steeper" functions, the empirical conclusion is that a step-size somewhat smaller than that found theoretically should be taken for best convergence. It is to be mentioned that such a problem with steep cost functions (e.g. $\phi_i(r_i) = (w_i r_i)^n$ with $n \approx 100$) or problem like (20) with large α can be utilised for the solution of minimax location problems (21) in Euclidean space[8]. It appears that at this time, this is the only feasible method for solving three dimensional minimax location problems.

REFERENCES

1. E. Weiszfeld, Sur le point pour lequel la somme des distances de n points donnés est minimum. *Tôhoku Math. J.* **43**, 335-386 (1937).
2. 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**, 21-33 (1962).
3. H. W. Kuhn, A note on Fermat's problem. *Math. Prog.* **4**, 98-107 (1973).
4. L. M. Ostresh Jr., On the convergence of a class of iterative methods for solving the Weber location problem. *Ops. Res.* **26**, 597-609 (1978).
5. I. N. Katz, Local convergence in Fermat's problem. *Math. Prog.* **6**, 89-104 (1974).
6. L. Cooper, An extension to the generalized Weber problem. *J. Reg. Sci.* **8**, 181-197 (1968).
7. M. Avriel, A geometric programming approach to the solution of location problems. *J. Reg. Sci.* **20**, 239-247 (1980).
8. R. Chen, Location problems with costs being sums of powers of Euclidean distance. *Comput. Ops Res.* In Press.
9. L. M. Ostresh, Jr., The multifacility location problem. Applications and descent theorems. *J. Reg. Sci.* **17**, 409-419 (1977).
10. I. N. Katz, On the convergence of a numerical scheme for solving some locational equilibrium problems, *SIAM J. Appl. Math.* **17**, 1224-1231 (1969).

11. F. Cordellier and J. C. Fiorot, On the Fermat–Weber problems with convex cost functions. *Math. Prog.* **14**, 295–311 (1978).
12. J. B. Crockett and H. Chernoff, Gradient methods of maximization. *Pacific J. Math.* **5**, 33–50 (1955).
13. A. Goldstein, Cauchy's method of minimization. *Numer. Math.* **4**, 146–150 (1962).
14. B. T. Polyak, Gradient methods for the minimization of functionals. *USSR Comput. Math. and Math. Phys.* **3**, 864–878 (1963).
15. A. I. Cohen, Stepsize analysis for descent methods. *J. Opt. Theory Applics* **33**, 187–205 (1981).
16. J. M. Ortega. *Numerical Analysis*, p. 45. Academic Press, New York (1972).