18

Combinatorial and Global Optimization, pp. 283–296
P.M. Pardalos, A. Migdalas and R. Burkard, Editors
© 2002 World Scientific Publishing Co.

A Derivative Free Minimization Method For Noisy Functions

V.P. Plagianakos (`vpp@math.upatras.gr`)
Department of Mathematics, University of Patras,
University of Patras Artificial Intelligence Research Center,
GR–26110 Patras, Greece.

M.N. Vrahatis (`vrahatis@math.upatras.gr`)
Department of Mathematics, University of Patras,
University of Patras Artificial Intelligence Research Center,
GR–26110 Patras, Greece.

Abstract

An unconstrained minimization method which is based on Powell's derivative free method is presented. The proposed method retains the termination properties of Powell's method and it can be successfully applied to problems with imprecise function values. The ability of this method to cope with imprecise or noisy problems is due to the fact that it proceeds solely by comparing the relative size of the function values. The method has been implemented and tested, and performance information is given.

Keywords: Powell's method, Rosenbrock's method, bisection method, unconstrained optimization, imprecise problems, noisy functions.

1   Introduction

Several methods for finding the extrema of a function $f: \mathcal{D} \subset \mathrm{R}^n \to \mathrm{R}$, where $\mathcal{D}$ is open and bounded, have been proposed, with many applications in different scientific fields (mathematics, physics, engineering, computer science etc.). Most of them

require precise function and gradient values. In many applications though, precise values are either impossible or time consuming to obtain. For example, when the function and gradient values depend on the results of numerical simulations, then it may be difficult or impossible to get very precise values. Or, in other cases, it may be necessary to integrate numerically a system of differential equations in order to obtain a function value, so that the precision of the computed value is limited. Furthermore, in many problems the values of the function to be minimized are computationally expensive. Such problems are common in real life applications as in the optimization of parameters in chemical experiments or finite element calculations, where a single measurement (function evaluation) takes several hours or even days. With such applications in mind, robust methods are needed which make good progress with the fewest possible number of function evaluations.

In this contribution a new method is presented for the computation of a minimum x^* of an n–dimensional real valued function f. The proposed algorithm is based on Powell's method (see [14, 28]), which minimizes a function without calculating derivatives, and proceeds solely by comparing the relative size of the function values. Thus, although this method retains the termination properties of Powell's method, it can be successfully applied to problems with imprecise function values.

In Section 2 we give a discussion of optimization of noisy functions as well as a simulation of the influence of noise (proportional to a Gaussian distributed random number with zero mean and various variances). In Section 3 a brief overview of Powell's method and a detailed description of the new method are presented, while in Section 4 numerical results are presented. Finally, in Section 5, we give some concluding remarks.

2   Optimization of noisy functions

The problem of optimization of noisy or imprecise (not exactly known) functions occurs in various applications, as for instance, in the task of experimental optimization. Also, the problem of locating local maxima and minima of a function from approximate measurement results is vital for many physical applications. In spectral analysis, chemical species are identified by locating local maxima of the spectra. In radioastronomy, sources of celestial radio emission and their subcomponents are identified by locating local maxima of the measured brightness of the radio sky. Elementary particles are identified by locating local maxima of the experimental curves.

The theory of local optimization provides a large variety of efficient and effective methods for the computation of an optimizer of a smooth function f. For example, Newton–type and quasi–Newton methods show superlinear convergence in the vicinity of a nondegenerate optimizer. However, these methods require the gradient or the

Hessian, respectively, in contrast to other optimization procedures, like the simplex method [12], the method of Hook and Jeeves [9], the direction set method of Powell (see [7, pp.87–92]), or some other recently proposed methods [5, 6, 23].

In some applications, however, the function to be minimized is only known within some (often unknown and low) precision. This might be due to the fact that evaluation of the function means measuring some physical or chemical quantity or performing a finite element calculation in order to solve partial differential equations. The function values obtained are corrupted by noise, namely stochastic measurement errors or discretization errors. This means that, although the underlying function is smooth, the function values available show a discontinuous behavior. Moreover, no gradient information is available. For small variations in a neighborhood of a point the corresponding function values reflect the local behavior of the noise rather than that of the function. Thus, a finite–difference procedure to estimate the gradient fails [5].

The traditional method for optimizing noisy functions is the simplex or polytope method proposed by Nelder and Mead [12] (cf. [7, p.18], [13, p.202], [15, p.230]). This method surpasses other well–known optimization methods when dealing with the large noise case. However, this is not valid in the noiseless case. The ability of this method to cope with noise, is due to the fact that it proceeds solely by comparing the relative size of the function values, as the proposed method does. The simplex method does not use a local model of the function f and works without the assumption of continuity. Although this method has poor convergence properties (for a convergence proof of a modified version see [18]), yet it has been proved to be a useful method in many applications. The simplex method converges whenever the standard deviation of the function at the vertices of the current simplex is smaller than some prescribed small quantity. This method can be deficient when the current simplex is very "flat". This can be avoided by suitable variants (see for example [18]). More sophisticated methods in this direction are discussed by Powell [16].

To study the influence of the imprecise information (regarding the values of the objective function and the gradient), we simulate imprecisions with the following approach. Information about $f(x)$ is obtained in the form of $f^\sigma(x)$, where $f^\sigma(x)$ is an approximation to the true function value $f(x)$, contaminated by a small amount of noise η. Specifically, the function values are obtained as [5, 6]:

$$f^\sigma(x) = f(x)(1 + \eta), \qquad \eta \sim N(0, \sigma^2), \qquad (1)$$

where $N(0, \sigma^2)$ denotes a Gaussian distributed random number with zero mean and variance σ^2, i.e., relative stochastic errors are used for the test problems. To obtain η we apply the method of Box and Muller [1] using various variances σ.

3 A derivative–free minimization method for imprecise problems and its convergence

In this section we briefly describe Powell's algorithm for solving the nonlinear unconstrained minimization problem without calculating derivatives. Also, we propose a derivative free minimization method which is based on Powell's method and we study the termination properties of the new method.

Powell's method [14] is based on the use of conjugate directions and the main idea of his approach is that the minimum of a positive–definite quadratic form can be found by performing at most n successive line searches along mutually conjugate directions, where n is the number of variables. Also, this procedure can be applied to non–quadratic functions by adding a new composite direction at the end of each cycle of n line searches. In this case finite termination is no longer expected. One iteration of Powell's basic procedure consists of the following steps, where x^0 is the initial approximation to the minimum, and u_i, $i = 1, 2, \ldots, n$ determine the initial set of directions which are equal to the basis vectors e_i, $i = 1, 2, \ldots, n$:

Step 1. For $i = 1, 2, \ldots, n$ compute λ_i to minimize $f(x^{i-1} + \lambda_i u_i)$ and define $x^i = x^{i-1} + \lambda_i u_i$.

Step 2. For $i = 1, 2, \ldots, n - 1$, replace u_i by u_{i+1}.

Step 3. Replace u_n by $x^n - x^0$.

Step 4. Compute λ to minimize $f(x^n + \lambda u_n)$, and set $x^0 = x^n + \lambda u_n$.

For a general (non–quadratic) function, the iteration is repeated until some stopping criterion is satisfied. If f is quadratic we minimize along n conjugate directions $u_1, u_2, \ldots, u_n$, and the minimum is reached if the u_i are all nonzero. This is true if $\lambda_1 \neq 0$ at each iteration, for then the directions $u_1, u_2, \ldots, u_n$ cannot become linearly dependent. The problem is that Powell's method has a tendency to choose search directions that are linearly dependent on each other, especially in ill–conditioned problems. There are various procedures to cope with the linear dependence in Powell's algorithm. The simpler way to avoid linear dependence of the search directions is to reset the set of directions u_i, to the basis vectors e_i after n or $(n + 1)$ iterations of the basic procedure. This is the procedure we will follow in our approach and it retains the quadratic termination property if $\lambda_1 \neq 0$.

From the above discussion, it is evident, that the only computable information required by Powell's method is the values of f when the one–dimensional subminimization procedures used to minimize $f(x + \lambda u)$ are performed without utilizing the derivatives of f. One such subminimization procedure has been proposed by Powell [14].

Next, we propose a modification of Powell's method for the numerical computation of a minimizer of f utilizing only the relative size of the function values. To do this we have to minimize $f(x + \lambda u)$ by comparing the relative size of the values of f. In our case, we have to minimize $f(x^0 + \lambda u)$ along the direction u. One way to do this, by applying one–dimensional rootfinding methods, is to compute the value of $\lambda \neq 0$ such that:

$$f(x^0 + \lambda u) - f(x^0) = 0. \tag{2}$$

Now, if $\hat{\lambda}$ is the solution of the above equation, then, of course, the point $\hat{x}^0 = x^0 + \hat{\lambda} u$ possesses the same function value as the point x^0, so it belongs to the contour line of x^0. Then, we can choose a point which belongs to the line segment with endpoints x^0 and $\hat{x}^0$ possessing a smaller function value than these endpoints. With this fact in mind we can now choose such a point, say for example:

$$x^1 = x^0 + \gamma \left(\hat{x}^0 - x^0 \right), \quad \gamma \in (0,1).$$

To solve the one–dimensional Eq. (2), we use our modification of the bisection method which is briefly described below. A solution of the equation $f(x) = 0$, where the function $f \colon [a, b] \subset \mathbb{R} \to \mathbb{R}$ is continuous, is guaranteed to exist in the interval (a, b) if the following criterion is fulfilled:

$$f(a)\, f(b) < 0, \quad \text{or} \quad \operatorname{sgn} f(a) \operatorname{sgn} f(b) = -1,$$

where sgn is the well known three valued sign function. This criterion is known as Bolzano's existence criterion (for a generalization of this criterion to higher dimensions see [21]). Based on this criterion various rootfinding methods, as for example the bisection method, were created. Here we shall use the bisection method which has been modified to the following simplified version described in [19, 20, 25]. There it is reported that, in order to compute a root of $f(x) = 0$ where $f \colon [a, b] \subset \mathbb{R} \to \mathbb{R}$ is continuous, a simplified version of the bisection method leads to the following iterative formula:

$$r^{p+1} = r^p + c \cdot \operatorname{sgn} f(r^p)/2^{p+1}, \quad p = 0, 1, \ldots, \lceil \log_2(b - a)\, \varepsilon^{-1} \rceil, \tag{3}$$

with $c = \operatorname{sgn} f(r^0)(b - a)$, $r^0 = a$ and where ε is the required accuracy. Of course the iterations (3) converge to a root $r^* \in (a, b)$ such that $|r^{p+1} - r^*| \leq \varepsilon$ if for some r^p, $p = 1, 2, \ldots$, the following holds:

$$\operatorname{sgn} f(r^0) \operatorname{sgn} f(r^p) = -1.$$

Instead of the iterative formula (3) we can also use the following

$$r^{p+1} = r^p - c \cdot \operatorname{sgn} f(r^p)/2^{p+1}, \quad p = 0, 1, \ldots, \lceil \log_2(b - a)\, \varepsilon^{-1} \rceil, \tag{4}$$

with $c = \operatorname{sgn} f(r^0)(b - a)$, $r^0 = b$.

The one–dimensional rootfinding portion of our method employs the above modified bisection method. We use this method since it is a globally convergent method, it always converges within the given interval and it is optimal, i.e. it possesses asymptotically the best possible rate of convergence [17]. In addition, it has a known behavior concerning the number of iterations required, when we seek a root with a predetermined accuracy ε, and last, but not least, it is the only method that can be applied to problems with imprecise function values, as it is evident from (3) and (4), where the only computable information required by the bisection method is the algebraic signs of the function f. As a consequence, for problems where the function value is given as a result of an infinite series (e.g. Bessel or Airy functions), it can be shown [24, 26] that the sign stabilizes after a relatively small number of terms of the series and the calculations can be speed up considerably.

In our case to solve Eq. (2) for λ, along the direction u, the above modified bisection method assumes the form:

$$\lambda^{p+1} = \lambda^p + c \cdot \text{sgn} \left[f\left(a + \lambda^p u\right) - f\left(x^0\right) \right] \Big/ 2^{p+1}, \quad p = 0, 1, \ldots, \lceil \log_2(b-a)\,\varepsilon^{-1} \rceil, \quad (5)$$

with $c = \text{sgn}\left[f(a) - f(x^0)\right](b - a)$, $\lambda^0 = 0$, $h = b - a$ and where h indicates the stepsize. Of course, utilizing our process we are able to obtain also local maxima along the line u. But if we choose the endpoints a and b in a proper way, then this method deals with a minimum. This can be easily handled by applying the iterative scheme (5) and taking the endpoints a and b from the following relations:

$$a = x^0 - s\beta - \frac{1}{2}\left(1 + s\right)h \qquad \text{and} \qquad b = a + h, \tag{6}$$

where $s = \text{sgn}\left[f(x^0 + \beta u) - f(x^0)\right]$ and β is a small positive real number (e.g. $\beta \propto \sqrt{epsmch}$ where $epsmch$ denotes the relative machine precision).

To study the termination properties of our approach we give the following result that states that any search method involving minimization along a set of n linearly independent conjugate directions has the quadratic termination property.

Theorem 3.1 [14, 28]. *If a quadratic function $f(x)$ of dimension n is minimized sequentially, once along each direction of a set of n linearly independent, conjugate directions, the global minimum of f will be located in n or less cycles independent of the starting point as well as the order in which the minimization directions are used.*

Remark 3.1 *A method that minimizes f according to the requirements of the above theorem has the property known as* quadratic termination. *Also, the order in which the directions are used is immaterial.*

Theorem 3.2 [14, 28]. *The directions generated in Powell's method are conjugate.*

Theorem 3.3 *The proposed method locates the minimum of an n–dimensional quadratic function f(x), in n or less iterations, utilizing only the relative size of the function values of f, independent of the starting point as well as the order in which the minimization directions are used.*

Proof: Since the proposed method uses the direction set u_i of Powell's method, by Theorem 3.2 these directions are conjugate. Also, the starting search directions are coordinate directions and hence they are linearly independent. Furthermore, the proposed method avoids linear dependence of the search directions by resetting the set of directions u_i, to the coordinate directions after n or $(n + 1)$ iterations of the basic procedure. Thus the assumptions of Theorem 3.1 are fulfilled and the result follows.

As it is well known the quadratic termination property is very powerful because most of the general (non–quadratic) functions can be approximated very closely by a quadratic one near their minima. Thus, this property is expected to speed up the convergence even for general functions. On the other hand Powell's method as well as the proposed method require generally more than, the theoretically estimated number, of n cycles for quadratic functions. The proof of the quadratic termination property has been established with the assumption that the exact minimum is found in each of the one–dimensional subminimizations. In practice, the actual subminimizer is approximated and hence the subsequent directions will not be conjugate. Thus, the methods require more iterations for achieving the overall convergence. For a proof of convergence, when the directions of the basis vectors are used and inexact one–dimensional subminimization procedures are applied, see [22, p.377].

4 Numerical applications

The above procedures were implemented using a new portable Fortran program named SIGNOPT, which has been applied to several test functions and has been compared with the well known Powell's and Rosenbrock's minimization methods.

Rosenbrock's method is a simple iterative procedure that can be utilized to locate the minimum of a function of one variable. An initial point x^0 and an initial step length h are chosen and the function $f(x)$ evaluated at x^0 and $(x^0 + h)$. If $f(x^0 + h) < f(x^0)$, the step is called a success, otherwise it is called a failure. The successive steps are taken according to the following simple algorithm:

Step 1.   Choose x^0, h.

Step 2.   **Repeat**

Step 3.   **If** $f(x^0 + h) < f(x^0)$.

Then $x^0 = x^0 + h$ and $h = 2h$.

Else $h = -h/2$.

Step 4.   **Until** converged.

The method described above can also be used to find the minimum of a multi-variable function, if the function $f(x)$ is minimized sequentially, once along each direction of a set of n linearly independent, conjugate directions. We have used as such directions the same directions we used for the SIGNOPT method, namely the Powell's directions.

SIGNOPT was tested on various test problems and our experience is that it behaves predictably and reliably. Some typical computational results, for some classical test problems, are given below. For each of the following examples, the reported parameters are:

— n dimension,

— σ^2 the value of the variance of the simulated noise,

— $x^0 = (x_1, x_2, \ldots, x_n)$ starting point,

— h the stepsize for the bisection and the Rosenbrock's method.

In the following tables the number of function evaluations needed for the methods to converge is given. If the field for the number of function evaluations contains a hyphen, the method failed to converge. Here we exhibit results from the following test problems.

EXAMPLE 1. *Himmelblau function,* [8]. In this case f is given by:

$$f(x) = \left(x_1^2 + x_2 - 11\right)^2 + \left(x_1 + x_2^2 - 7\right)^2.$$

The starting point was $x^0 = (1,1)$ and the steplength for the bisection and the Rosenbrock's method were $h = 9$ and $h = 1$ respectively. In Table 1 we exhibit the corresponding results obtained by the three methods. Powell's method was faster in the noiseless case, but when we added noise to the function values, SIGNOPT exhibited better performance. The Rosenbrock's method needed much more function evaluations in order to converge.

EXAMPLE 2. *Kearfott function,* [10]. The objective function f is given by:

$$f(x) = \left(x_1^2 + x_2^2 - 2\right)^2 + \left(x_1^2 - x_2^2 - 1\right)^2.$$

Table 1: Function Evaluations for solving Example 1.

σ	Powell	Rosenbrock	SIGNOPT
0	321	727	432
0.01	726	10283	645
0.10	845	14963	645
0.20	927	18809	645
0.30	1900	21649	751

Table 2: Function Evaluations for solving Example 2.

σ	Powell	Rosenbrock	SIGNOPT
0	144	292	325
0.01	598	5763	324
0.10	-	14534	324
0.20	-	13286	324
0.30	-	31260	324

In this problem we started the methods from the point $x^0 = (1,1)$ and the steplength for the bisection was $h = 9$ and for the Rosenbrock's method was $h = 1$. SIGNOPT required the same function evaluations, even when we have used noisy function values, while Powell's method failed to converge for $\sigma > 0.01$ and Rosenbrock's method needed many function evaluations. In Table 2 we exhibit the obtained results.

EXAMPLE 3. *Weber–Werner singular function*, [27]. In this example the function is given by:

$$f(x) = \left(x_1^2 - 2x_1 + \frac{x_2^3}{3} + \frac{2}{3}\right)^2 + \left(x_1^3 - x_1 x_2 - 2x_1 + \frac{x_2^2}{2} + \frac{3}{2}\right)^2. \qquad (7)$$

As our starting values we utilized, $x^0 = (10, 10)$. The steplength of the bisection method was $h = 9$ and the steplength for Rosenbrock's method was $h = 1$. The results are shown in Table 3. Once again Powell's method failed to converge when we increased the variance of the noise. Rosenbrock's method converged even when we applied noisy function values, but once again it needed many function evaluations. On the other hand, SIGNOPT converged to the minimum within at most 1404 function evaluations, in all cases.

EXAMPLE 4. *Broyden banded function*, [11]. In this example f is given by:

$$f(x) = \sum_{i=1}^{n} f_i^{\,2}(x), \qquad (8)$$

Table 3: Function Evaluations for solving Example 3.

σ	Powell	Rosenbrock	SIGNOPT
0	2352	733	511
0.01	4433	10397	898
0.10	4036	14002	1404
0.20	-	31233	1404
0.30	-	47460	1404

with:

$$f_i(x) = x_i(2 + 5x_1^2) + 1 - \sum_{j \in J_i} x_j(1 + x_j), \qquad (9)$$

where:

$$J_i = \{j : j \neq i, \max(1, i - m_l) \leq j \leq \min(n, i + m_u)\} \quad \text{and} \quad m_l = 5, \ m_u = 1. \quad (10)$$

In this problem, for $n = 2$, we used as starting values $x^0 = (1, 1)$, $h = 9$ for the bisection and $h = 1$ for Rosenbrock's method. For $n = 3$ we started the methods from the point $x^0 = (1, 1, 1)$, using the same steplengths. In Table 4 we exhibit the

Table 4: Function Evaluations for solving Example 4.

n	σ	Powell	Rosenbrock	SIGNOPT
2	0	246	599	536
	0.01	1549	12437	964
	0.10	9369	20524	1289
	0.20	-	15819	1180
	0.30	-	30756	1824
3	0	2182	1426	1002
	0.01	-	43654	3429
	0.10	-	42231	2574
	0.20	-	90159	4000
	0.30	-	-	9289

obtained results. SIGNOPT converged in all cases and had predictable performance, while Powell's method diverged as we increased σ and Rosenbrock's method once again exhibited slow convergence.

EXAMPLE 5. *Hilbert function*, [2]. In this example f is given by:

$$f(x) = x^\top A x, \qquad (11)$$

where A is an n by n Hilbert matrix, i.e.:

$$a_{ij} = \frac{1}{i+j-1}, \quad 1 \leq i, j \leq n. \tag{12}$$

This is a positive definite quadratic function, but its condition number increases rapidly with n. The starting point was $x^0 = (1,1)$ for $n = 2$, $x^0 = (1,1,1)$ for $n = 3$ and $x^0 = (1,1,1,1)$ for $n = 4$. The bisection and Rosenbrock's steps were $h = 9$ and $h = 1$ respectively. In Table 5 we exhibit the corresponding results obtained by the methods. Powell's method for $n = 4$ has shown a quite random behavior and Rosenbrock's method had a slow convergence. Once again, SIGNOPT outperformed the other two methods; was faster and more predictable.

Table 5: Function Evaluations for solving Example 5.

n	σ	Powell	Rosenbrock	SIGNOPT
2	0	106	205	322
	0.01	1151	1213	432
	0.10	-	2411	538
	0.20	-	38882	324
	0.30	-	65861	1825
3	0	439	507	428
	0.01	2523	3553	716
	0.02	3370	3814	716
	0.03	11370	3514	716
	0.04	15221	3623	858
	0.05	15598	4343	858
	0.10	24533	5040	4225
4	0	882	717	533
	0.01	6160	3810	359
	0.02	7174	4426	359
	0.03	4297	7136	359
	0.04	9757	6659	359
	0.05	1624	9373	538
	0.10	4505	11332	4824

5 Concluding remarks

A method for the computation of a minimum of an n–dimensional noisy or inexact objective function f is presented. The proposed method is based on Powell's derivative–

free minimization method and utilizes a modification of the one–dimensional rootfinding bisection procedure. The method proceeds solely by comparing the relative size of the function values. Thus, although it retains the termination properties of Powell's method, it can be successfully applied to problems with imprecise function values.

Our experimental results clearly show that the proposed method outperforms Powell's and Rosenbrock's methods when noisy function values were provided. Powell's method utilizing exact one–dimensional line search was faster when accurate function values were used. In the presence of noise, Powell's method either needed many function evaluations or failed to converge and it behaved rather randomly. On the other hand, the performance of SIGNOPT, when noisy function values were provided, was satisfactory and predictable, and its number of function evaluations seemed to increase with n in a smooth and predictable way, which makes it a reliable method for minimizing imprecise or noisy functions.

References

[1] Box G.E.P. and Muller M.E., A note on the generation of random normal deviates, *Ann. Math. Statistics*, **29**, 610–611, (1958).

[2] Brent R.P., *Algorithms for minimization without derivatives*, Prentice–Hall, Englewood Cliffs, NJ, (1973).

[3] Dennis J.E. and Schnabel R.B., *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice–Hall, Englewood Cliffs, NJ, (1983).

[4] Dennis J.E. and Torczon V., Direct search methods on parallel machines, *SIAM J. Optimization*, **1**, No. 4, 448-474, (1991).

[5] Elster C. and Neumaier A., A grid algorithm for bound constrained optimization of noisy functions, *IMA J. Numer. Anal.*, **15**, 585–608, (1995).

[6] Elster C. and Neumaier A., A method of trust region type for minimizing noisy functions, *Computing*, **58**, 31–46, (1997).

[7] Fletcher R., Practical Methods of Optimization, Wiley, New York, (1987).

[8] Himmelblau D.M., *Applied nonlinear programming*, McGraw–Hill, New York, (1972).

[9] Hook R. and Jeeves T., Direct search solution of numerical and statistical problems, *J. ACM*, **7**, 212–229, (1969).

[10] Kearfott R.B., An efficient degree–computation method for a generalized method of bisection, *Numer. Math.*, **32**, 109–127, (1979).

[11] Moré B.J., Garbow B.S. and Hillstrom K.E., Testing unconstrained optimization, *ACM Trans. Math. Software* **7**, 17–41, (1981).

[12] Nelder J.A. and Mead R., A simplex method for function minimization. *Computer J.*, **7**, 308–313, (1965).

[13] Nocedal J., *Theory of algorithms for unconstrained optimization*, In: Acta Numerica 1992, Iserles A., ed., Cambridge University Press, Cambridge, 199–242, (1992).

[14] Powell M.J.D., An efficient method for finding the minimum of a function of several variables without calculating derivatives, *Computer J.*, **7**, 155–162, (1964).

[15] Powell M.J.D., *A review of algorithms for nonlinear equations and unconstrained optimization*, In: ICIAM 1987 Proceedings, McKenna J. and Temam R., eds., SIAM, Philadelphia, 220–232, (1988).

[16] Powell M.J.D., A direct search optimization method that models the objective and constraint functions by linear interpolation, *Numerical Analysis Reports*, DAMTP 1992/NA5, Department of Applied Mathematics and Theoretical Physics, University of Cambridge, England, April (1992).

[17] Sikorski K., Bisection is optimal, *Numer. Math.*, **40**, 111–117, (1982).

[18] Torczon V., On the convergence of the multidimensional search algorithm, *SIAM J. Optimization*, **1**, No. 1, 123–145, (1991).

[19] Vrahatis M.N., Solving systems of nonlinear equations using the nonzero value of the topological degree, *ACM Trans. Math. Software*, **14**, 312–329, (1988).

[20] Vrahatis M.N., CHABIS: A mathematical software package for locating and evaluating roots of systems of nonlinear equations, *ACM Trans. Math. Software*, **14**, 330–336, (1988).

[21] Vrahatis M.N., A short proof and a generalization of Miranda's existence theorem, *Proc. Amer. Math. Soc.*, **107**, 701–703, (1989).

[22] Vrahatis M.N., Androulakis G.S., Lambrinos J.N. and Magoulas G.D., A class of gradient unconstrained minimization algorithms with adaptive stepsize, *J. Comput. Appl. Math.*, **114**, 367–386, (2000).

[23] Vrahatis M.N., Androulakis G.S. and Manoussakis M.E., A new unconstrained optimization method for imprecise function and gradient values, *J. Math. Anal. Appl.*, **197**, 586–607, (1996).

[24] Vrahatis M.N, Grapsa T.N., Ragos O. and Zafiropoulos F.A., On the localization and computation of zeros of Bessel functions, *Z. Angew. Math. Mech.*, **77**, 467–475, (1997).

[25] Vrahatis M.N. and Iordanidis K.I., A rapid generalized method of bisection for solving systems of nonlinear equations, *Numer. Math.*, **49**, 123–138, (1986).

[26] Vrahatis M.N, Ragos O., Zafiropoulos F.A. and Grapsa T.N., Locating and computing zeros of Airy functions, *Z. Angew. Math. Mech.*, **76**, 419–422, (1996).

[27] Weber H. and Werner W., On the accurate determination of nonisolated solutions of nonlinear equations, *Computing*, **26**, 315–326, (1981).

[28] Zangwill W.I., Minimizing a function without calculating derivatives, *Computer J.*, **10,** 293–296, (1967).