Universal Journal of Applied Mathematics & Computation **3 (2015), 63-69**
www.papersciences.com

Determination of Stationary Points of Multivariable Functions using Quadratic Interpolation

G.B. Muravskii

Technion – Israel Institute of Technology
12 Havatselet Str.
Nesher, 36730, Israel
gmuravsk@tx.technion.ac.il

Abstract
In a vicinity of a stationary point of a multivariable function, the quadratic interpolating function containing along with unknown coefficients also coordinates of the stationary point itself is constructed. The coefficients are determined (independently of the stationary point) exactly and rather simply due to the specific choice of interpolation points in the considered n-dimensional space. Knowing the coefficients allows us to obtain a system of n linear equations for an approximation for n coordinates of the stationary point. This process is repeated iteratively with placing the new set of the interpolation points into immediate proximity of the stationary point found approximately at the previous iteration (with decreasing distances between points of the set). The method is free from determination of derivatives and requires only a single iteration when the given function is quadratic itself. A connection between the method and the Newton method is considered. The important new result could be formulated as follows: if the stationary point of a quadratic function is searched, the finite-difference analog of Newton equations can be used one time, and the location of the initial point and values of steps along coordinate axes can be taken in an arbitrary way. This can serve as a good base for practical applications in engineering and mathematical physics.

2010 Mathematical Subject Classification: 41A05, 49M05.
Key Word and Phrases
Stationary point, Interpolation, Iterations, Newton Method.

1. Introduction

Among numerous methods of optimization, parabolic interpolation represents a widely used method for the one dimensional cases. The attempts to spread this quadratic interpolation method over the domain of multivariable functions apparently have not been made whilst the problem becomes rather simple when choosing a specific set of interpolation points. As result we obtain a method which is free from determination of derivatives and is exact and not requiring iterations when the objective function is quadratic itself. In the general case, the essence of the method suggests that the interpolation points become more and more close to the sought-for stationary point, i.e. an iteration process should be applied in which distances between the interpolation points are decreased and the whole set is placed more and more nearly to the previously found approach to the stationary point. In this paper, an analysis of the existing methods of optimization is not carried out, one could it found in books (see, e.g., [1]-[6]).

2. Description of the Method

The suggested method is characterized by including coordinates of the stationary point into an iteration process which leads to more and more exact values of the coordinates. In a vicinity of a stationary point S, we represent an enough smooth function $F(x_1,\ldots,x_n)$ of n variables $x_1,\ldots,x_n$ in the form of the quadratic function:

$$F(x_1,...,x_n) \approx G(x_1,...,x_n) = A_0 + \sum_{i=1}^{n} A_{ii}(x_i - S_i)^2 + \sum_{i=1}^{n-1}\sum_{j=i+1}^{n} A_{ij}(x_i - S_i)(x_j - S_j) \qquad (2.1)$$

In the last group of terms, only the coefficients with $j > i$ are considered; the terms of the first power are not present due to the presumed stationarity of the point S. The number of the coefficients A_{ii}, A_{ij} in (2.1) is $n+n(n-1)/2$; it can be said that they form a symmetrical matrix whose diagonal elements and elements of the upper triangular matrix enter into (2.1). The representation (2.1) contains coefficients A_0, A_{ii}, A_{ij} which are to be found and also unknown coordinates of the assumed stationary point, $S_1,S_2,\ldots,S_n$; thus the total number of unknowns is $1+2n+n(n-1)/2$. For determination of the unknowns we require that the function represented by (2.1) gives the values of the function $F(x_1,\ldots,x_n)$ at $1 + 2n + n(n-1)/2$ points located in a vicinity of the point S. The more close the points to the point S are the more exact the equation (2.1) is and more exactly the stationary point will be found. Note that in paper by Powell [7] the possibility of finding the above mentioned function $G(x_1,\ldots,x_n)$ using the required number of function values of $F(x_1,\ldots,x_n)$ was pointed out and estimated as "likely to be a bad method". However a specific choice of interpolation points allows obtaining very effective solution for the corresponding system of equations which leads simultaneously to an approach to the coordinates of the stationary point.

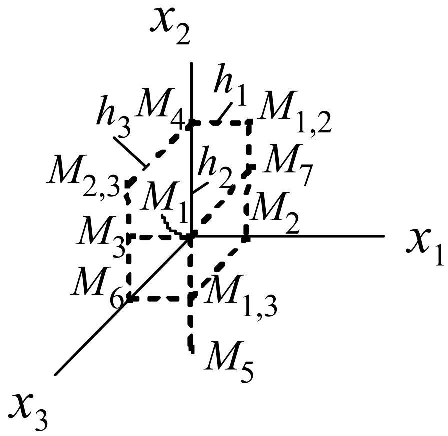

Fig. 1 . Set of interpolation points (3-dimensional case). Lines have been shown which go from point M_1 parallel to coordinate axes.

The following set of the points is advisable (see Fig. 1 which illustrates the case $n = 3$): the points $M_1,\ldots,M_{1+2n}$ are taken at lines "parallel" to coordinate axes (M_1 and two additional points with a step h_i for each such line), and the points $M_{i,j}$ ($i = 1,2,\ldots,n-1$; $j = i+1,\ldots,n$) lie in the planes parallel to the planes (x_i,x_j) and have x_i-coordinate equal x_i-coordinate of M_1 plus h_i and x_j-coordinate equal x_j-coordinate of M_1 plus h_j; the parameters h_i (steps along axes) should be decreased during the process of iterations in which a new set of points for the next iteration is located near the point S found at the previous iteration; it is recommended to place the new point M_1 into the found point S. We see that the number of the points corresponds to the number of the unknowns. Denoting the i-th coordinate of the points M_k, $M_{r,s}$ by $x_{i,k}$, $x_{i;r,s}$, respectively, the system of equations can be written in the form:

$$A_0+\sum_{i=1}^{n}A_{ii}(x_{i,k}-S_i)^2+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}(x_{i,k}-S_i)(x_{j,k}-S_j)=F_k$$
$$(k=1,2,...,2n+1)$$
$$A_0+\sum_{i=1}^{n}A_{ii}(x_{i;r,s}-S_i)^2+\sum_{i=1}^{n-1}\sum_{j=i+1}^{n}A_{ij}(x_{i;r,s}-S_i)(x_{j;r,s}-S_j)=F_{r,s}$$
$$(r=1,2,...,n-1;s=r+1,...,n) \tag{2.2}$$

Here $F_k = F(x_{1,k},\ldots,x_{n,k})$, $F_{r,s} = F(x_{1;r,s},\ldots,x_{n;r,s})$ are values of the objective function in the considered points; the equations for the points M_k, lying at coordinate "lines", and the points $M_{r,s}$ lying in the corresponding coordinate "planes" are conveniently written separately. Owing to the choice of the points the system (2.2) allows rather simple solution. The coefficients A_{ii} (i = 1,2, …n) are determined considering n groups of three points lying at coordinate "lines" (see Fig. 1): M_1,M_2,M_3; M_1,M_4,M_5;… M_1,M_{2n},M_{2n+1}. In each group, the points have identical coordinates other than x_1 (for the first group), x_2 (for the second group) and so on. This allows us to eliminate all unknowns except the considered A_{ii}. We subtract from the equation for the point M_{2i} the doubled equation for the first point of i-th group (i.e. M_1) and add the equation for the third point of the group (i.e. M_{2i+1}). Accounting that the terms not containing x_i cancel out, we write below only the terms in which the coordinate $x_{i,1}$, $x_{i,2i}$, $x_{i,2i+1}$, are present:

$$A_{ii}\left(-2(x_{i,1}-S_i)^2+(x_{i,1}+h_i-S_i)^2+(x_{i,1}-h_i-S_i)^2\right)$$
$$+\sum_{j=i+1}^{n}A_{ij}\left(-2(x_{i,1}-S_i)(x_{j,1}-S_j)+(x_{i,1}+h_i-S_i)(x_{j,1}-S_j)+(x_{i,1}-h_i-S_i)(x_{j,1}-S_j)\right)$$
$$+\sum_{j=1}^{i-1}A_{ji}\left(-2(x_{j,1}-S_j)(x_{i,1}-S_i)+(x_{j,1}-S_j)(x_{i,1}+h_i-S_i)+(x_{j,1}-S_j)(x_{i,1}-h_i-S_i)\right)$$
$$=F_{2i}-2F_1+F_{2i+1} \tag{2.3}$$

All terms in (3), except those containing A_{ii}, disappear which results in:

$$\tilde{A}_{ii}=h_i^2A_{ii}=\frac{F_{2i}-2F_1+F_{2i+1}}{2}$$
$$(i=1,2,...,n) \tag{2.4}$$

The coefficients A_{ii} are equal to halved finite difference approximations of second partial derivatives of the given function. We emphasize that in our treatment this result corresponds to an exact solution of the system of equation (2.2) regardless of accuracy of derivatives' approximations. For determining the coefficients A_{ij} ($j > i$) we consider $n(n–1)/2$ groups of four points lying in coordinate "planes" (x_i,x_j), i.e. the points $M_1,M_{2i},M_{2j},M_{i,j}$. The following combinations of equations for each group are considered: the second equation (i.e. for the second point of the group) is subtracted from the first one plus the fourth equation minus the third one. In this combination finally only the terms with A_{ij} ($j > i$) remain and we obtain:

$$\tilde{A}_{ij}=A_{ij}h_ih_j=F_1-F_{2i}-F_{2j}+F_{i,j}$$
$$(i=1,2,...,n-1;j=i+1,...,n) \tag{2.5}$$

The coefficients A_{ij} represent finite difference approximations for the mixed second partial derivatives; emphasize again that we have obtained the exact solution of the considered system of equations. The elements of the low triangular matrix $\tilde{A}_{ij}$ are adopted using the symmetry. Note that all coefficients have been found independently of unknown values S_i.

After determining the coefficients A_{ii}, A_{ij}, the approximate coordinates S_i of the stationary point can be found. It is advisable to represent the coordinates in the form:

$$S_i = x_{i,1} - h_i\delta_i \tag{2.6}$$

using the new unknowns δ_i (normalized coordinate deviations of the point M_1 from the point S). The corresponding n equations can be written by subtraction the equation for the point M_1 from equations from the point M_{2i} (i = 1, 2,…,n). It is clear that the terms with coefficients A_{ps} where $p \neq i$, $s \neq i$ can be omitted from the outset and we keep only the terms with A_{ji} (j = 1,...,i–1), A_{ii}, A_{ij} (j = i +1,…,n), i.e. the terms in which the coordinate x_i is present (being different for the two considered points). The system of equations has the form:

$$\sum_{j=1}^{i-1} A_{ji}(x_{j,1}-S_j)\left((x_{i,1}+h_i-S_i)-(x_{i,1}-S_i)\right)+A_{ii}\left((x_{i,1}+h_i-S_i)^2-(x_{i,1}-S_i)^2\right)$$
$$+\sum_{j=i+1}^{n} A_{ij}\left((x_{i,1}+h_i-S_i)-(x_{i,1}-S_i)\right)(x_{j,1}-S_j)=F_{2i}-F_1$$

(2.7)

$$(i = 1,\ldots,n)$$

Or accounting for (2.6)

$$\sum_{j=1}^{i-1}\tilde{A}_{ji}\delta_j+2\tilde{A}_{ii}\delta_i+\sum_{j=i+1}^{n}\tilde{A}_{ij}\delta_j=-\left(F_{2i}-F_1-\tilde{A}_{ii}\right)=\frac{F_{2i}-F_{2i+1}}{2} \tag{2.8}$$

$$(i=1,\ldots..,n)$$

The coefficients of the equations are determined by equations (2.4), (2.5); along with the solution of (2.8) (assuming that the corresponding matrix is not singular) the coefficients represent an exact solution of the system (2.2). Therefore in the case when the given function is quadratic itself, the single application of equations (2.8) (single iteration) leads to the determination of the stationary point independently of the location of the point M_1 and the steps h_j. Note that the right-hand sides in equations (2.8) correspond to the derivative (multiplied by h_i) at the point M_1 defined by the centered difference approximation.

1. In the case n =1, above formulas give:

$$\delta_1=\frac{F_2-F_3}{4\tilde{A}_{11}}, S_1=x_{1,1}-h_1\delta_1 \tag{2.9}$$

This result corresponds to the well known method of inverse parabolic interpolation.

2. The corresponding solution for n = 2 will be:

$$D=4\tilde{A}_{11}\tilde{A}_{22}-\tilde{A}_{12}^2,\ b_1=\frac{F_2-F_3}{2},\ b_2=\frac{F_4-F_5}{2},$$
$$\delta_1=\frac{2\tilde{A}_{22}b_1-\tilde{A}_{12}b_2}{D},\delta_2=\frac{2\tilde{A}_{11}b_2-\tilde{A}_{12}b_1}{D},S_1=x_{1,1}-h_1\delta_1,S_2=x_{2,1}-h_2\delta_2 \tag{2.10}$$

3. For n =3 we present an expression for the third right-hand part of system (2.8) in addition to those given already in (2.10):

$$b_3 = \frac{F_6 - F_7}{2}. \tag{2.11}$$

The solution can be represented using determinants:

$$\begin{aligned}
&D = 2\left(4\tilde{A}_{11}\tilde{A}_{22}\tilde{A}_{33} + \tilde{A}_{12}\tilde{A}_{13}\tilde{A}_{23} - \tilde{A}_{11}\tilde{A}_{23}^2 - \tilde{A}_{22}\tilde{A}_{13}^2 - \tilde{A}_{33}\tilde{A}_{12}^2\right),\\
&\delta_1 = \frac{b_1\left(4\tilde{A}_{22}\tilde{A}_{33} - \tilde{A}_{23}^2\right) - b_2\left(2\tilde{A}_{33}\tilde{A}_{12} - \tilde{A}_{13}\tilde{A}_{23}\right) + b_3\left(\tilde{A}_{12}\tilde{A}_{23} - 2\tilde{A}_{22}\tilde{A}_{13}\right)}{D},\\
&\delta_2 = \frac{b_2\left(4\tilde{A}_{33}\tilde{A}_{11} - \tilde{A}_{13}^2\right) - b_3\left(2\tilde{A}_{11}\tilde{A}_{23} - \tilde{A}_{12}\tilde{A}_{13}\right) + b_1\left(\tilde{A}_{13}\tilde{A}_{23} - 2\tilde{A}_{33}\tilde{A}_{12}\right)}{D},\\
&\delta_3 = \frac{b_3\left(4\tilde{A}_{11}\tilde{A}_{22} - \tilde{A}_{12}^2\right) - b_1\left(2\tilde{A}_{22}\tilde{A}_{13} - \tilde{A}_{23}\tilde{A}_{12}\right) + b_2\left(\tilde{A}_{12}\tilde{A}_{13} - 2\tilde{A}_{11}\tilde{A}_{23}\right)}{D},\\
&S_i = x_{i,1} - h_i\delta_i \quad (i = 1,2,3)
\end{aligned} \tag{2.12}$$

3. Relation to Newton Method

The system (2.8), the above formulas for the coefficients A_{ii}, A_{ij} and right-hand parts of the system (2.8) show that the quadratic interpolation applied corresponds to Newton method with replacements of the Hessian terms with the corresponding finite difference approach (see (2.4), (2.5)), and instead of the first partial derivatives the centered difference approximations are used. Note that the values of δ_i correspond to the deviations from the supposed stationary points (see (2.6)), whereas in the Newton method we deal with the values opposite in sign to δ_i, so in right hand sides also the signs are opposite. The important new result could be formulated as follows: if the stationary point of a quadratic function is searched, the Newton equations with indicated replacements can be used one time, and the location of the initial point M_1 and values of steps h_j can be taken in an arbitrary way.

In the general case, for sufficient exact determination of the stationary point of a twice continuously differentiated function, the corresponding iteration process should be performed in which the steps h_j along coordinate lines become more and more smaller (e.g. divided by a value greater than 1 after each iteration) whereas the point M_1 is placed nearly the already found point S (current approach to the stationary point); it is advisable to use for the coordinate of the new point M_1 the relationship $x_{i,1}= S_i$. These characteristics of the suggested algorithm follow from the fact that the considered quadratic approach is searched as the interpolation function becoming in the iteration process more and more close to the quadratic part of the corresponding Taylor series; the method during the progress of the iterations is more and more close to the classic Newton method.

4. Some Examples

Three method: method of Davies, Swann and Campey; Powell's method and Smith's method considered in the paper by Fletcher [8], are compared with the suggested interpolation method for the following functions.

1. Rosenbrock's function [9]

$$F(x_1, x_2) = 100\left(x_2 - x_1^2\right)^2 + (1 - x_1)^2 \tag{4.13}$$

Starting at the set with M_1 having coordinates (–1.2,1) and the steps h_1=h_2=0.03 with subsequent dividing them by 3 after each iteration we achieve after 8 iteration the point S(0.999995,0.999990) with the function value 7.3×10^{-11}; it should be noted that the behavior of the function is not monotonic in the course of iterations. The number of function determinations is 8×6=48, whereas using other method (see [4]) requires significantly more function determinations for reaching the same smallness of the function value.

2. A function of 4 variables [7]:

$$F(x_1, x_2, x_3, x_4) = (x_1 + 10x_2)^2 + 5(x_3 - x_4)^2 + (x_2 - 2x_3)^4 + 10(x_1 - x_4)^4 \tag{4.14}$$

with starting point M_1(3,–1,0,1) and initial equal steps h_i=2.1 which are divided by 2 after each iteration. A point with the value of the function 2.22×10^{-9} has been achieved after 7 iteration, so the number of function determinations equals 7×15=105 which again is smaller than for other methods [8]. It should be noted rather large values of steps in this example which are far from those ensuring an acceptable finite-difference approximation for derivatives; of course, this approximation is not needed in the considered method of interpolation.

3. A helical valley [10]:

$$\begin{gathered} F(x_1, x_2, x_3) = 100\left((x_3 - 10\theta)^2 + (r - 1)^2\right) + x_3^2 \\ x_1 = r\cos(2\pi\theta),\ x_2 = r\sin(2\pi\theta) \end{gathered} \tag{4.15}$$

with the initial point M_1(–1,0,0), equal initial steps h_j=0.5 which are divided by 3 after each iteration. After 11 iterations we come to the function value 2.92×10^{-12}. Thus the number of function determination is 11×10=110 which is smaller than in the case of the three methods considered in [8].

It should be noted that the indicated good convergence of the suggested method takes place only for a suitable choice of the initial steps and the rate of steps decrease.

5. Conclusions

Although applications of finite-difference approximations for derivatives in the context of Newton method represent a well known treatment of the optimization problem, apparently it was not recognized before that this treatment (with the above indicated kinds of finite-difference approximations) leads to an exact method of quadratic interpolation and therefore allows obtaining the instantaneous solution when the function under consideration is quadratic itself. A quick convergence should take place for functions close to quadratic ones. It should be emphasized the possibility of a method improvement by including the corresponding linear optimization after each realization of the quadratic interpolation.

References

1. *Dennis, J. E. and Schnabel, R. B., 'NumericalMethods for Unconstrained Optimization',* Prentice-Hall, Englewood Cliffs, New Jersey*, 1983.*
2. *Edgar, T. E., Himmelblau, D. M., Lasdon, L. S., 'Optimization of chemical processes',* McGraw-Hill, New York*, 2001.*
3. *Fletcher, R.,' Practical methods of optimization, volume 1, Unconstrained optimization',* John Wiley & Sons, New York, *1980*
4. *Gill, P.E., Murray, W., Wright, M.,'Practical Optimization',* Academic Press, London*, 1982.*

5. *Nocedal, J. and Wright,S. J., 'Numerical Optimization',* Springer-Verlag, New York*, 1999.*
6. *Singiresu, S. Rao., 'Engineering Optimization: Theory and Practice',* Fourth Edition, John Wiley & Sons, New York, *2009.*
7. *Powell, M. J. D.,* 'An iterative method for finding stationary values of a function of several variables'*, The Computer Journal,* ***5*** *(1962), 147-151.*
8. *Fletcher, R., '*Function minimization without evaluating derivatives—a review'*, The Computer Journal,* ***8*** *(1965), 33-41.*
9. *Rosenbrock, H. H.,* 'An automatic method for finding the greatest or the least value of a function'*, The Computer Journal,* ***3*** *(1960), 175-184.*
10. *Fletcher, R. and Powell, M. J. D.,* 'A rapidly convergent descent method for minimization'*, The Computer Journal,* ***6*** *(1963), 163-168.*