7 APPLYING NELDER MEAD'S OPTIMIZATION ALGORITHM FOR MULTIPLE GLOBAL MINIMA

Ştefan ŞTEFĂNESCU[*]

Abstract

The iterative deterministic optimization method could not more find multiple global minima of a given objective function ([6]).

Generally, the probabilistic optimization algorithms have not this restrictive behaviour, to determine only a single global minimum point. In this context we'll prove experimentally that Nelder-Mead's heuristic procedure can detect successfully multiple global extremal points.

Key words : global optimization, Nelder-Mead algorithm, multiple minima

JEL Classification code : C61, C02.

1. Introduction

For an arbitrary function $h: D \to R$ with $D \subset R^m$ we intend to find those points $x^* \in D$, $x^* = (x_1^*, x_2^*, x_3^*, ..., x_m^*)$ such that

$$x^* = \arg\min_{w \in D} h(w) \tag{1}$$

Therefore

$$h(x^*) = \min_{x \in D} h(x) \tag{2}$$

where $x = (x_1, x_2, x_3, ..., x_m)$.

In fact $h(x^*)$ is the minimum global value for the function $h(x)$, $x \in D$.

In the literature ([6]) are very present the classical *derivative optimization methods*, based on the gradient direction for finding the minimum global value $h(x^*)$.

[*] *Associate Professor, PhD., Faculty of Mathematics and Informatics, University of Bucharest, e-mail : stefanst@fmi.unibuc.ro*

But always in practice the exact expression of the gradient function could be extremely hard for computing. For this reason the gradient expression is often approximated by finite differences.

The *non-derivative methods* use directly only some selected values $h(x)$ ([1]-[5], [7]-[10]). In this context we remark *the model-based variant* and *the geometry-based method* too.

More precisely, the model-based procedures work with an interpolation or also with a least-squares approximation of the objective function $h(x)$ to compute the next iteration in searching process of x^*.

Contrary, the geometry-based algorithms do not necessary involve an explicit auxiliary form of the function $h(x)$ and essentially produce samples from $x \in D$ which have imposed properties.

The Nelder-Mead (*NM*) method is oriented for solving a continuous unconstrained optimization problem of type (2).

A *NP* type algorithm is clearly an authentic geometry-based procedure whose flexibility is given by its four parameters $\alpha, \beta, \gamma, \delta$ which adjust the search process for the minimum function values.

In general, the geometry-based procedures and particularly the *NP* algorithm are easily to be programmed. Their major advantage is imposed by a relative non frequently evaluation of the function $h(x)$. Usually, in practice, the computation of a complex objective function $h(x)$ is very time-consuming. Often the evaluation of $h(x)$ demands before an auxiliary data collected activity.

2. An implementation of the *NM* algorithm

The iterative optimization procedures generally use only a starting point $x_1 \in D$, chosen by specific rules.

Contrary, the *NP* algorithm consider a nondegenerate simplex inside the domain $D \subset R^m$ as starting figure. At every iteration step the *NP* algorithm modifies a single vertex of the current simplex by applying a λ-transform. In this way it results another nondegenerate simplex.

More precisely, for any two points $y \in R^m$ and $z \in R^m$ we can produce a new point $w \in R^m$ by using a λ-rule, that is

$$w = z + \lambda(y - z) \quad , \quad \lambda \in R \tag{3}$$

So, if $y = (y_1, y_2, y_3, ..., y_m)$, $z = (z_1, z_2, z_3, ..., z_m)$, $w = (w_1, w_2, w_3, ..., w_m)$

we get
$$w_j = z_j + \lambda(y_j - z_j) \quad , \quad 1 \le j \le m \tag{4}$$

Depending on the value of the coefficient λ, $\lambda \in \{\alpha, \beta, \gamma, \delta\}$, and also on the individual significance of the points y and z, we can simulate more geometric type operations as a α-reflection, a β-expansion, a γ-contraction or a δ-shrinkage ([1}, [2], [5], [9]).

The classical Nelder-Mead algorithm [5] has a lot of little modified forms (compare, for example, the *NP* procedures presented in [1]-[3], [7], [9]). For the present study it was implemented in MatLab the variant given in [3]. This variant operates with the following λ-parameters :

$$\alpha = 1 \qquad \beta = 2 \qquad \gamma = 0.5 \qquad \delta = 0.5 \tag{5}$$

3. Multiple global minima

In the subsequent we intend to test the *NM* algorithm when the function $h(x)$ has multiple minima. We are interested to see if the *NM* procedure could find all the global extremal values $x*$.

The following example will give us the right answer.

Example 1. For $m = 2$ we will consider the function $h_1 : D \to R$ with

$$D = [0\,,6]\ x\ [-3\,,12] \qquad D \subset R^2 \tag{6}$$

$$h_1(w) = h_1((w_1, w_2)) = 4 + |(w_1 - 1)(w_1 - 5)| + |w_2 - (w_1 - 2)(w_1 - 3)|$$

Obviously

$$h_1(s) = h_1(t) = \inf_{w \in D} h_1(w) = 4 \tag{7}$$

where

$$s = (1, 2) \qquad t = (5\,, 6) \tag{8}$$

and more $s \in D$, $t \in D$.

From a straightforward reasoning we deduce that the function $h_1(w)$ has, on the domain D, only two global extremal points. These special points are just the vectors s and t defined by the formulas (8).

Graphic 1 gives us an imagine about how the function $h_1(w)$ fluctuates.

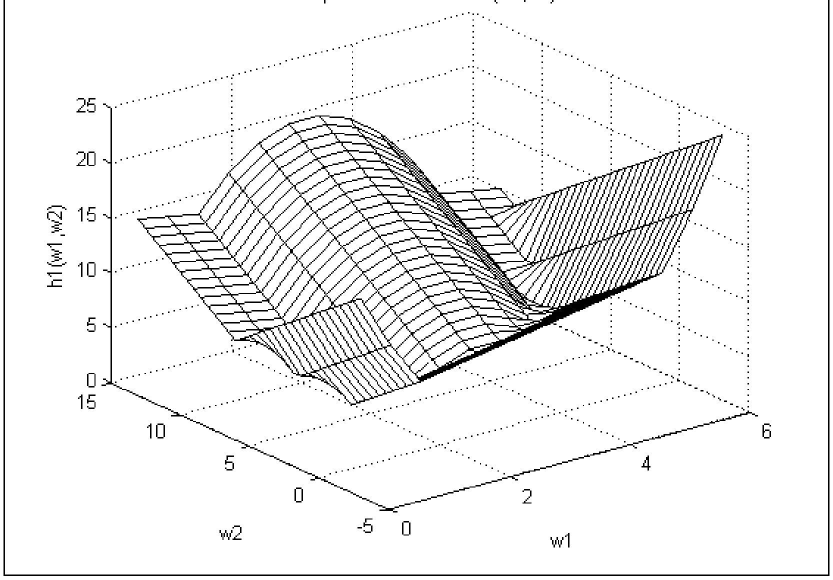
We intend to verify if the *NP* procedure could find both minimizer points s and t. The Graphic 1 does not suggest us clearly the exact places where we have the two global extremal points s, t .

For this reason we can study the variability of the function $h_2(w)$,

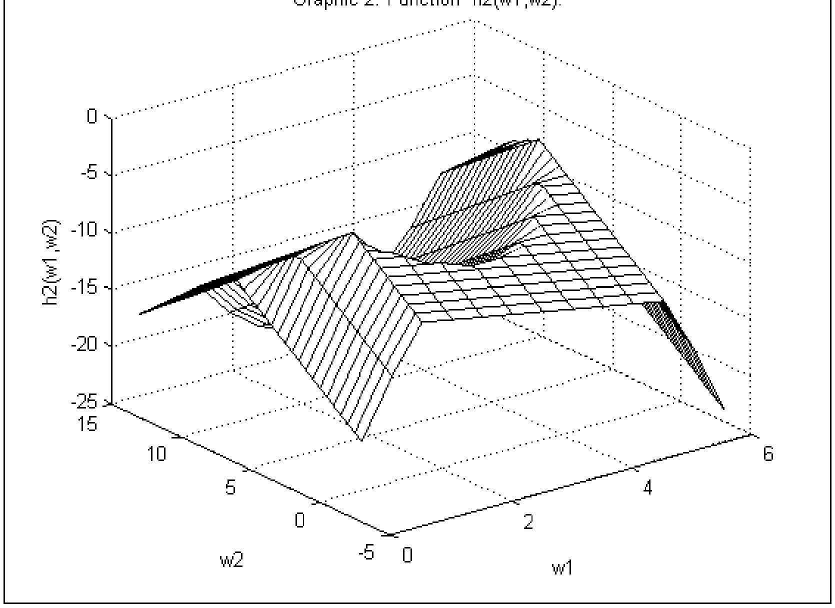
$$h_2(w) = h_2((w_1, w_2)) = -h_1((w_1, w_2)) \tag{9}$$

The minimum values of the function $h_1(w)$ became the maximum values for the application $h_2(w)$. The Graphic 2 suggests at least two global maximization points for the function $h_2(w)$. So, $h_1(w)$ has multiple global minimizer points.

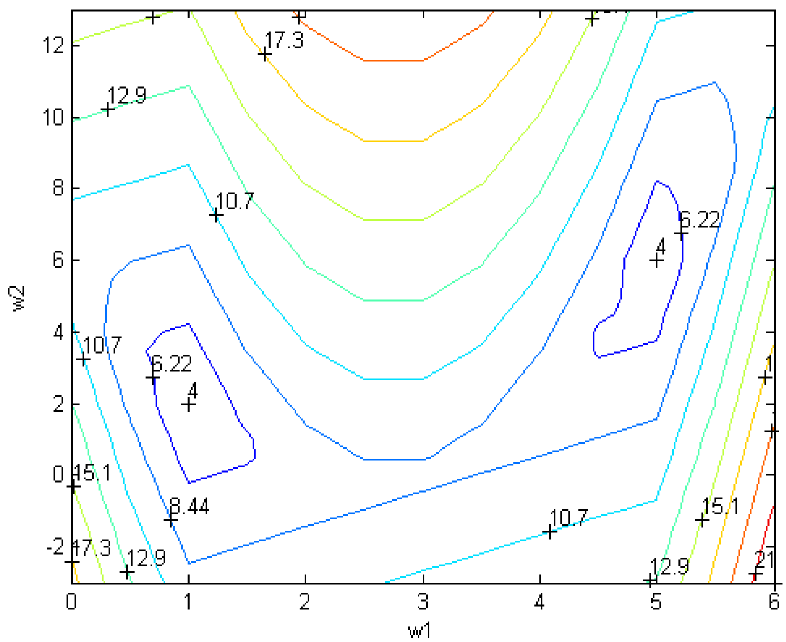
Graphic 1. Function h1(w1,w2).



Graphic 2. Function h2(w1,w2).



Graphic 3. The contour lines for the function h1(w1,w2).



But the correct answer regarding the number of the global extremal points of $h_1(w)$ is obtain after an interpretation of the contour lines structure. So, we conclude that the function $h_1(w)$ has only two minimizer points (see Graphic 3).

Running 100 times the *NM* algorithm we get always only the minimizer vectors s or t but after a different number n of iterations. More, the variants s and t appeared randomly and around the same proportion (see *Table 1*).

Table 1

The minimization value x^* obtained after n iterations

(*NM* algorithm, $x^* \in \{s, t\}$, function $h_1(w)$).

x*	n	x*	n	x*	n	x*	n	x*	n
t	63	s	58	S	60	t	56	s	56
t	61	s	56	S	56	t	59	t	63
s	66	t	61	S	59	s	60	t	55
s	52	s	59	S	54	t	60	t	72
s	59	s	58	T	61	t	66	s	56
s	56	t	67	T	54	s	56	s	53
s	53	s	48	S	54	t	59	t	62
s	56	s	59	T	55	t	56	s	55

s	57	t	54	T	66	s	55	s	55
t	61	t	91	S	55	s	62	t	58
t	70	t	81	S	62	s	55	t	68
s	62	s	65	T	69	t	60	s	57
s	57	t	66	S	55	t	59	s	60
s	123	t	54	T	54	t	58	s	56
t	65	t	59	T	62	s	55	t	52
t	53	s	59	S	68	t	57	t	57
s	58	s	73	S	56	t	56	t	116
s	55	s	66	S	61	s	57	t	62
s	55	t	57	S	72	s	63	t	77
t	61	t	53	T	62	t	60	t	60

4. Concluding remarks

It is very known from the literature that the iterative deterministic optimization methods could not usualy find more multiple minima of a given objective function $h(w)$ (details in [6]).

But this behavioural restriction isn't generally true for the probabilistic optimization algorithms.

In the present paper we proved experimentaly that the Nelder-Mead heuristic procedure can detect successfully multiple extremal global points.

More, in example 1, the *NP* procedure identified approximately in the same proportion the both global minimizer points (see *Table 1*).

References

R, Barton, J. S. Ivey, "Nelder Mead simplex modifications for simulation optimization", *Management Science* 42, 7(1996), 954–973.

L. Han, M. Newmann, "Effect of dimensionality on the Nelder-Mead simplex method", *Optimization Methods and Software*, 21, 1(2006), 1-16.

J.C. Lagarias, J.A. Reeds, M.H. Wright, P.E. Wright, "Convergence properties of the Nelder-Mead simplex method in low dimensions", *SIAM Journal on Optimization*, 9, 1(1998), 112-147.

K.I.M. McKinnon, "Convergence of the Nelder-Mead simplex method to a nonstationary point", *SIAM Journal on Optimization*, 9, (1998), 148-158.

J.A. Nelder, R. Mead, "A simplex method for function minimization", *Computer Journal 7*, (1965), 308-313.

W.H. Press, B.P. Flannery, S.A. Teukolsky, W.T. Vettering, *Numerical Recipes in C. Cambridge University Press*, Cambridge, UK, 1988.

C.J. Price, I.D. Coope, D. Byatt, "A convergent variant of the Nelder-Mead algorithm", *Journal of Optimization Theory and Applications*, 113, (2002), 5–19.

A.S. Rykov, 'Simplex algorithms for unconstrained optimization", *Problems of Control and Information Theory*, 12, (1983), 195-208.

P. Tseng, "Fortified-descent simplicial search method: a general approach" *SIAM Journal on Optimization 10*, (2000), 269–288.

W.C. Yu, "The convergence property of the simplex evolutionary techniques", *Scientia Sinica,Special issue of Mathematics 1*, (1979), 68–77.