NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

CONCEPTUAL NEWTON METHOD FOR SOLVING MULTIVALUED INCLUSIONS: SCALAR CASE

E.A. Nurminski

June 1979
WP-79-50

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

ABSTRACT

Due to different reasons, the actual state of economic, environmental and even mechanical systems is often known only as a set of possible values of the systems indexes. Another source of uncertainty is an unspecified reaction of the system to the changes in control or unpredicted changes in systems inputs. The theory of set-valued mapping provides the mathematical tools for analysis and construction of such systems and is of great importance to system analysis methodology.

This paper is concerned with one of the basic problems of application of set valued mapping - solving multivalued inclusions. It uses the original definition of a set valued derivative and develops the Newton-like method for solving this problem. The remarkable feature of the proposed method is a quadratical rate of convergency.

CONCEPTUAL NEWTON METHOD FOR SOLVING MULTIVALUED INCLUSIONS: SCALAR CASE

E.A. Nurminski

INTRODUCTION

Many problems of economic theory, optimal control, mathematical programming etc. can be formulated as a problem of finding the roots of specific multivalued mappings. This problem can be stated as follows: for given multivalued mapping (multifunction, set-valued map) $Y(x)$ which for every $x \in X$ define a convex subset $Y(x) \subset Y$ find such x^* that

$$0 \in Y(x^*) \ . \tag{1}$$

Many methods were proposed for solving the problem in specific formulations connected with applications mentioned above. In general formulation this problem has been studied less thoroughly. The method of simple iteration has been considered within the framework of the theory of monotone mappings.

This method has a form:

$$x^{s+1} = x^s - \rho_s y^s \ , \ s = 0,1,\ldots \tag{2}$$

$$y^s \in Y(x^s)$$

where ρ_s - step-size multipliers, and it is generally known that the convergency of this method is rather slow. In nondifferentiable convex optimization for instance where $Y(x) = \partial f(x)$ - subgradient mapping, one can consult J.L. Goffin (1977) on that matter.

The problem of speeding up the methods for solving (1) is of great theoretical interest and practical importance. Here we undertook the development of the Newton-like method which for single-valued cases has a reputation of extremely fast convergency. This development requires some new notions to be introduced and the first part of the paper deals with the notion of the derivative of the multivalued mappings and its properties. It is mainly based on the author's paper (1978).

The second part is devoted to the method itself. Here we present the conceptual variant of the method which is not straightforward implementable. The problem of developing the implementable variant of this method is an independent one and will be solved in future.

The main result of this part is a quadratical rate of convergency of the proposed method. It makes the method very attractive from the computational point of view and justifies the research efforts in this direction.

DERIVATIVES OF THE MULTIFUNCTIONS

The different definitions of the multifunction derivatives were proposed by Tjurin (1965), Hukuhara (1967), Hermes (1968), Bridgland (1970), Banks, Jacobs (1970), Pschenichiy (1977,1976), Makarov, Rubinov (1973), Martelli, Vignoli (1974), De Blasi (1976), Bradley, Datko (1977) and possibly by many others. As a survey of these definitions is not an aim of this paper, we restrict ourselves to the mention of these works and a detailed description of the author's approach which was originally published in 1978.

We consider here finite dimensional space X and Y with standard euclidean norm

$$\|a\| = (aa)^{1/2} \quad , \quad a \in X \text{ or } Y$$

and correspondently the inner product of two vectors a and b will be written as ab. Also, let denote for a set A which can be either a subset of X or Y

$$\|A\| = \sup_{a \in A} \|a\|$$

We will also use specific notation for some relation between the two sets A and B. It is written as

$$A = B|\theta$$

and it means that a set C exists such that

$$A \subset B + C$$

$$B \subset A + C$$

and

$$\|C\| \leq \theta$$

Certainly it is equivalent to the fact that Hausdorf distance between A and B less θ, but for our purposes the proposed notation is more obvious.

Following the definition given in Nurminski (1978) multi-valued mapping Y(x) is called a differentiable at the point x^0 if a set $Y'(x^0)$ exists such that whether

(A1) $$Y(x) = Y(x^0) + Y'(x^0)(x-x^0) \mid o(\|x-x^0\|)$$

for $x > x^0$ or

(B1) $$Y(x^0) = Y(x) - Y'(x^0)(x-x^0) \mid o(\|x-x^0\|)$$

for $x > x^0$ and whether

(A2) $$Y(x) = Y(x^0) - Y'(x^0)(x^0-x) \mid o(\|x-x^0\|)$$

or

(B2) $$Y(x^0) = Y(x) + Y'(x^0)(x^0 - x) \mid o(\|x - x^0\|)$$

for $x < x^0$.

It was supposed that we deal with scalar x. The generalization of the given definition on a multidimensional case and the consequences will be discussed in a separate paper.

For (A1,2) we will call $Y'(x^0)$ right-oriented and for (B1,2)- left-oriented derivative.

If we consider the support function

$$V_p(x) = \sup_{y \in Y(x)} py$$

then the following criteria of differentiability can be proposed.

Theorem 1 (Nurminski 1978). A multivalued mapping Y(x) is differentiable if, and only if, the support function is differentiable, and the derivative of the support function at the point x^0 is a convex or concave function of p.

The proof of this statement presents no special difficulty and so is omitted.

NEWTON METHOD

With the definition of the derivative of multivalued mapping given above, we consider the multivalued analogues of the Newton method for solving the problem (1).

Let us define the following procedure:

Step 1. For given x^s define set X^{s+1} in a manner which is described in the table below.

$Y'(x^s)$	$Y(x^s) > 0$	$Y(x^s) < 0$
right-oriented	$Y'(x^s)(x - X^{s+1}) = Y(x^s)$	$X^{s+1} = x^s - (Y'(x^s))^{-1} Y(x^s)$
left-oriented	$X^{s+1} = x^s - (Y'(x^s))^{-1} Y(x^s)$	$Y'(x^s)(x^s - X^{s+1}) = Y(x^s)$

Step 2. Solve the problem

$$\inf_{x \in X^{s+1}} \inf_{y \in Y(x)} \|y\| = \inf_{y \in Y(x^{s+1})} \|y\|$$

and go to Step 1.

This procedure generates a sequence of points $\{x^s, s=0,1,\dots\}$ with a given initial point x^0. We will prove further that if the initial point is close enough to the solution of (1) then the sequence generated by the method has a quadratic rate of convergence.

Theorem 2. Let $Y(x)$- be differentiable and the derivative $Y'(x)$ satisfies Lipschitz condition with constant L. Let us assume that the inverse of the derivative is uniformly bounded from above

$$\|Y^{-1}(x)\| \le C^{-1}$$

and the initial point x^0 is such that

$$\Delta_0 = \inf_{y \in Y(x^0)} \|y\| \le \left(\frac{L}{C^2}\right)^{-1}$$

Then the sequence x^s converges to the solution of (1) in a sense that

$$\Delta_s = \inf_{y \in Y(x^s)} \|y\| \to 0$$

and

$$\Delta_{s+1} \le \frac{L}{C^2} \Delta_s^2 \quad .$$

Proof. First of all it is easy to see that if x^s is not a solution of the problem (1) then the set X^{s+1} does not contain point x^s. It means that the method cannot stop at a point which is not a solution.

Further we assume that, without loss of generality, $Y'(x) > 0$. Then there are two other possibilities to consider: right and left-oriented derivatives.

(A3) Right-oriented derivative

In this case if $Y(x^s) > 0$ then the following inequalities are valid:

$$\Delta_{s+1} \triangleq \inf_{x \in X^{s+1}} \inf_{y \in Y(x)} \| y \| =$$

$$= \inf_{x \in X^{s+1}} \inf_{y \in Y(x^s)} \sup_{p \in S} py =$$

$$= \inf_{x \in X^{s+1}} \sup_{p \in S} \inf_{y \in Y(x^s)} py =$$

$$= \inf_{x \in X^{s+1}} \sup_{p \in S} \{W_p(x)\} =$$

$$= \inf_{x \in X^{s+1}} \sup_{p \in S} \{W_p(x^s) + W'_p(x^s)(x - x^s) + R_p(x - x^s)\} =$$

$$= \inf_{x \in X^{s+1}} \sup_{p \in S} \{W_p(x^s) + V'_{-p}(x^s)(x^s - x) + R_p(x - x^s)\}$$

Consider the second term in the brackets

$$V'_{-p}(x^s)(x^s - x) = (x^s - x) \sup_{y' \in Y'(x^s)} - py' =$$

$$= \sup_{y' \in Y'(x^s)} - py'(x^s - x) = \sup_{y \in Y'(x^s)(x^s - x)} - py \le$$

$$\le \sup_{y \in -Y(x^s)} py = V_{-p}(x^s) = -W_p(x^s)$$

as by definition of X^{s+1}

$$Y(x^s) = Y'(x^s)(x^s - X^{s+1})$$

Consequently

$$\Delta_{s+1} \le \inf_{x \in x^{s+1}} \sup_{p \in S} \{W_p(x^s) - W_p(x^s) + R_p(x - x^s)\} =$$

$$= \inf_{x \in X^{s+1}} \sup_{p \in S} R_p(x - x^s)$$

For the remainder $R_p(x - x^s)$ it is possible to make the following estimates: so far as derivative $Y'(x)$ is Lipschitz continuous with respect to x, it is easy to show that $W'_p(x)$ is also Lipschitz continuous. Then uniformly with respect to $p \in S$

$$|R_p(x - x^s)| \leq L\|x - x^s\|^2$$

and hence

$$\Delta_{s+1} \leq L \inf_{x \in X^{s+1}} \|x - x^s\|^2$$

$$= L(\inf_{x \in X^{s+1}} \|x - x^s\|)^2$$

Furthermore

$$\Delta_s = \inf_{y \in Y(x^s)} \|y\| = \inf_{\substack{y' \in Y'(x^s) \\ x \in X^{s+1}}} \|y'(x - x^s)\| \geq$$

$$\geq \inf_{x \in X^{s+1}} \|x - x^s\| \inf_{y' \in Y'(x^s)} \|y'\| \geq$$

$$\geq C \inf_{x \in X^{s+1}} \|x - x^s\|$$

Finally we obtain

$$\Delta_{s+1} \leq \frac{L}{C^2} \Delta_s^2$$

which proves our statements for the case under consideration.

(B3) Left-oriented derivative

Here we consider the case when $Y(x^s) \geq 0$ because otherwise it is equivalent to (A3). This case is somewhat simpler than (A3).

Here

$$\Delta_{s+1} \triangleq \inf_{x \in X^{s+1}} \inf_{y \in Y(x)} \|y\|$$

and the estimate for the right-hand side we obtain in the following manner:

$$Y(x) + R(x - x^s) \supset Y(x^s) + Y'(x^s)(x - x^s)$$

where

$$\|R(x - x^s)\| \leq r(x - x^s) \leq L\|x - x^s\|^2$$

then

$$\inf_{\substack{y \in Y(x) \\ z \in R(x - x^s)}} \|y+z\| \leq \inf_{y \in Y(x^s) + Y'(x^s)(x - x^s)} \|y\| = 0$$

as

$$0 \in Y(x^s) + Y'(x^s)(x - x^s)$$

for any $x \in X^{s+1}$.

On the other hand

$$\inf_{\substack{y \in Y(x) \\ x \in R(x - x^s)}} \|y+z\| \geq \inf_{y \in Y(x)} \|y\| - \sup_{z \in R(x - x^s)} \|z\| \geq$$

$$\geq \inf_{y \in Y(x)} \|y\| - L\|x - x^s\|^2$$

Eventually

$$\inf_{y \in Y(x)} \|y\| \leq L\|x - x^s\|^2$$

and consequently

$$\Delta_{s+1} \leq \inf_{x \in X^{s+1}} L\|x - x^s\|^2 \leq \frac{L}{c^2} \Delta_s^2$$

where the last estimate can be obtained as in the previous case.

All other possibilities are equivalent to either A3 or B3, so the theorem is proved.

REFERENCES

Banks, H.T. and M.Q. Jacobs (1970) A differential calculus for multifunctions. Journal of Mathematical Analysis and Applications 29(2):246-272.

Bradley, M. and R. Datko (1977) Some analytic and measure theoretic properties of set valued mappings. SIAM Journal Control and Optimization, 15(4):625-635.

Bridgland,T.F. (1970) Trajectory integrals of set valued functions. Pacific Journal of Mathematics. 33:43-68.

De Blasi, F.S. (1976) On the differentiability of multifunctions. Pacific Journal of Mathematics 66(1):67-82.

Goffin, J.L. (1977) On convergence rates of subgradient optimization methods. Mathematical Programming 13:329-347.

Hermes, H. (1968) Calculus of set valued functions and control. Journal of Mathematics and Mechanics 18:47-60.

Hukuhara,M. (1967) Integration des applications measurables dont la valuer est un compact convexe. Funkcialaj Ekvacioj 10:203-223.

Makarov, V.L. and A.M. Rubinov (1977) Mathematical theory of economic growth. Berlin:Springer-Verlag.

Martelli, M. and A. Vignoli (1974) On differentiability of multivalued maps. Unione Matematica Italiana Bolletino (Bologne) 4(10):701-712.

Nurminski, E.A. (1978) On differentiability of multivalued mappings. Cybernetics 5:46-48 ((in Russian).

Pshenichnij, B.N. (1972) Convex multifunctions and their conjugates. Cybernetics 3:94-102 (in Russian).

Pshenichnij, B.N. (1976) Necessary optimality conditions for differential inclusions, Cybernetics 6:60-73 (in Russian).

Tjurin, Yu.N. (1965) Mathematical formulation of the simplified model of industrial planning. Ekonomika i Mat. Metody 1(3):391-410 (in Russian).