NOT FOR QUOTATION
WITHOUT PERMISSION
OF THE AUTHOR

SOFTWARE FOR REGIONAL STUDIES: ON THE DIFFERENCE-APPROXIMATION APPROACH TO SOLVING SYSTEMS OF NONLINEAR EQUATIONS

Alexander G. Birjukov*

December 1982
PP-82-8

*Research Scholar,
Chair of Mathematical Bases of Management,
Moscow Physico-Technical Institute (M.Ph.T.I.)
Moscow District,
Dolgoprudnyi, 141700 USSR

Professional Papers do not report on work of the International Institute for Applied Systems Analysis, but are produced and distributed by the Institute as an aid to staff members in furthering their professional activities. Views or opinions expressed are those of the author(s) and should not be interpreted as representing the view of either the Institute or its National Member Organizations.

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

PREFACE

The success of regional institutions frequently depends on the quality of the software they use. Thus, software development problems rank as key issues in the field of regional studies. For this reason much of the research effort of the Regional Development Group has been devoted to examining such problems and many software elements have been developed, tested, and implemented with positive results.

This article by Alexander Birjukov describes some schemes of unconstrained optimization and methods for solving nonlinear equations that have been found to be among the most effective.

August, 1982

Boris Issaev†
Leader
Regional Development Group

SUMMARY

A scheme of generating efficient methods for solving nonlinear equations and optimization problems which is based on a combined application of the computation methods of linear algebra and the finite-difference approximations of derivatives is proposed. Examples of the new methods constructed with the help of the approach proposed as well as the examples of its possible applications are given below.

SOFTWARE FOR REGIONAL STUDIES: ON THE DIFFERENCE-APPROXIMATION APPROACH TO SOLVING SYSTEMS OF NONLINEAR EQUATIONS

Alexander G. Birjukov

1. INTRODUCTION

The difference-approximation approach (DAA) to solving nonlinear systems and optimization problems usually is understood as embracing various forms of application of difference-approximations of derivatives. In the present work we interpret the term DAA like that: it is the use of difference-approximations only in the frames of applications of computation methods of linear algebra for solving the above mentioned problems. The examples are: discrete Newton method (see [1]), generalized Gaussian elimination method [2], different variants of conjugate directions method [3-6], unconstrained minimization problem solution method which exploits approximations of eigenvectors and eigennumbers of Hessian matrix [7] and other methods. From the above list of methods it is possible to single out two principally different forms of the use of approximations, or forms of DAA. The first form consists of two steps: first a certain linearized problem is being generated with the help of difference-approximations, and then this problem is being solved by a method of linear algebra [1,7]. The second form does not tailor an explicitly formulated linearized problem, but in the process of solving the problem by one of the linear algebra methods difference-approximations of the elements of the method are applied [2-6]. In the present work on the basis of the analysis of the available methods we gave the definition of the DAA with respect to solutions of nonlinear systems of equations and the problem of unconstrained minimization and we proposed a scheme of generating the methods of the second form of the DAA. By way of illustrating the realization of the proposed scheme, new methods are put forward and investigated. These methods are shown to possess the quadratic convergency rate of the discrete Newton method, and at the same time they require considerably less memory (7n instead of $\frac{n \cdot (n+5)}{2}$, where n - is the dimension of the problem). Some other cases of possible uses of the DAA are also given in the work.

1. THE DISCUSSION OF THE DAA SCHEME PROPOSED CAN BE EASILY SEEN ON THE FOLLOWING PROBLEMS

Problem 1. Find $x \in E^n$ such that $g(x^*) = 0$, where $g: E^n \to E^n$.

Problem 2. Find $x^* = \operatorname{argmin} f(x)$, where $f: E^n \to E^1$.

Problem 3. (Auxiliary). Find the solution $x^* \in E^n$ of the linear system of equations $Az + b = 0$, where $z, b \in E^n$, $A \in E^{n \times n}$.

It is supposed that g and f are sufficiently smooth and the solutions of the problems do exist. Some additional constraints depend upon the chosen method of solution.

Note that problem 2 is reducible to problem 1 because the point x^* in problem 2 satisfies equation $f'(x^*) = 0$. But at the same time problem 2 has its own specific traits and that is why it is discussed parallel to problem 1. Note as well that linear approximations of problem 1 and equation $f'(x) = 0$ result in problem 3.

It is known (see [8]) that nonlinear programming and optimal control problems can be reduced to problem 1 and to problem 2. Thus, the DAA is applicable to such problems of optimization as well.

We shall search the solution of problems 1 and 2, as usual, as a limit of the sequence of the form:

$$x_k = x_{k-1} + \lambda_k p_k,\ p_k \in E^n,\ \lambda_k \in E^1,\ k = 1,2,\ldots, \quad . \qquad (1)$$

The values λ_k for problems 1 and 2 can be chosen, for example, the following way:

$$\text{a)} \quad \lambda_k = \underset{\lambda}{\text{argmin}} \, \|g(x_{k-1} + \lambda p_k)\|^2 \quad ,$$

$$\text{b)} \quad \lambda_k = \underset{\lambda}{\text{argmin}} \, f(x_{k-1} + \lambda p_k) \quad . \qquad (2)$$

The purpose of the approach under consideration is, first of all, the generation of methods to choose vector p_k. We proceed from the following considerations: if to assume problem 1 we have $A = g'(x)$, and for problem 2 we have $A = f''(x)$ and $b = f'(x)$, then the values

$$As,\ r^T As,\ s^T b \quad , \quad \text{where} \quad r,\ s \in E^n \quad , \qquad (3)$$

can be approximated, for example, by the following well known formulae:

$$\text{a)} \quad As \cong [g(x+hs) - g(x)]/h \quad ;$$

$$\text{b)} \quad As \cong [f'(x+hs) - f'(x)]/h \quad ;$$

$$\text{c)} \quad r^T As \cong [f(x+hs+hr) - f(x+hs) - f(x+hr) + f(x)]/h \quad ; \qquad (4)$$

$$\text{d)} \quad s^T f'(x) \cong [f(x+hs) - f(x)]/h \quad ;$$

where $|h| \neq 0$ is a small number.

Application of expression (4) instead of (3) in the first and the second forms of approximations implementation is the essence of the difference-approximation approach to the solution of problems 1 and 2. If it is desirable one can apply formulae of higher level of accuracy in h, of course, if g(x) and f(x) are smooth enough. It is easily seen, that f'(x) in (4b) can also be approximated, with the accuracy required, with the help of corresponding formulae, which opens the possibility of using only the values of f(x) for solving problem 2.

The methods of solving problems 1 and 2 on the basis of the first form of DAA are well investigated and are not discussed here. Below under the term DAA its second form is ment.

The following scheme of generating methods for solving problems 1 and 2 can be formulated for this form: 1) take <u>any method</u> for solving problem 3 in which the values of the form (3) are used; 2) in the process of computing vector p_k to substitute values (2) by their approximate values (4); 3) to apply some method of choosing λ_k.

The possibilities of the scheme above can be extended if instead of the methods of solution of problem 3 to use in this scheme other computational methods of linear algebra, for example, the methods of finding eigenvalues and eigenvectors of matrices g'(x) or f"(x), note that eigenvectors or their combinations are taken for vector p_k.

It is natural that for each of the methods generated thorough analysis of its stability against the difference-approximation errors is required.

2. EXAMPLES OF THE NEW METHODS GENERATED ON THE BASIS OF DAA

For the methods 1, 2, and 3 described below, for solving problem 1 we assume that g(x) is differentiable and matrix g'(x) is symmetrical. The principal merit of the methods, in addition to the high rate of convergence, is a relatively small (of the order of 7n numbers) volume of the memory required for the realization. In the discussion of the methods proposed we compare the generated vectors $p_k = p(x_{k-1})$ with the Newton vector $p_k^H = p^H(x_{k-1}) = -g'(x_{k-1})^{-1} g(x_{k-1})$ on the sequence of points (1), however, the evaluations obtained for $\| p_k - p_k^H \|$ are true for an arbitrary point $x \in E^n$ because the initial point $x_0 \in E^n$ in (1) is chosen arbitrarily.

<u>Method 1</u>. Vector p_k^I in (1) is calculated using the relation

$$p_k^I = \sum_{j=1}^{m} a_j s_j \quad ,$$

where

$$s_1 = r_o \equiv -g(x_{k-1}) \quad , \quad z_j = [g(x_{k-1} + hs_j) - g(x_{k-1})]/h \quad ,$$

$$a_j = r_{j-1}^T r_{j-1}/s_j^T z_j \quad , \quad r_j = r_{j-1} - a_j z_j \quad ,$$

$$b_j = r_j^T r_j / r_{j-1}^T r_{j-1} \quad , \quad s_{j+1} = r_j + b_j s_j \quad ; \; j = 1, \ldots, m \quad ,$$

$$m \leq n$$

with m being the minimal member for which $\| r_m \| \leq \varepsilon_k \| g(x_{k-1}) \|$, where $0 < \varepsilon_k < 1$, and $|h| \neq 0$, is a sufficiently small number.

Method 1 is the result of application of DAA to the conjugate gradient method [9]. It was proposed in [5] and in a slightly different form [6]. Here and below the index k with the values a_j, s_j, r_j, z_j, b_j, and so on, which are in the formulae for calculating p_k is omitted to shorten the notation.

Theorem 1. Let $g'(x)$ and ε_k answer the conditions:

$$\gamma_1 \|y\|^2 \leq y^T g'(x) y \leq \gamma_2 \|y\|^2, \; \gamma_2 \geq \gamma_1 > 0, \; \forall x, y \in E^n \quad (5)$$

$$\|g'(x) - g'(y)\| \leq L \| x-y \| \quad , \quad L \geq 0 \quad \forall x, j \in E^n \quad (6)$$

$$\varepsilon_k = \min(\varepsilon, M \|g(x_{k-1})\|), \quad 0 < \varepsilon < 1, \; M > 0 \quad . \quad (7)$$

Then with a sufficiently small $|h|$ for method 1, the evaluation

$$\|p_k^I - p_k^H\| \leq (\varepsilon_k + C_1 |h| \, \|g(x_{k-1})\|) \, \|g(x_{k-1})\| / \gamma_1$$

where $C_1 > 0$ depends only on γ_1, γ_2, L is admissible.

If (5) is not satisfied and $g'(x)$ has negative eigenvalues, method 1 might turn to be unstable because of upzeroing product $s_j^T z_j$. This deficiency is absent in the following method.

Method 2. Vector p_k^{II} in (1) is calculated using the relation

$$p_k^{II} = \sum_{j=i}^{m} a_j s_j \quad ,$$

where

$$s_1 = r_0 \equiv - [g(x_{k-1} + h\, g(x_{k-1})) - g(x_{k-1})]/h \quad ,$$

$$w_j = [g(x_{k-1} + h\, s_j) - g(x_{k-1})]/h \quad ,$$

$$z_j = [g(x_{k-1} + h\, w_j) - g(x_{k-1})]/h \quad ,$$

$$a_j = r_{j-1}^T r_{j-1} / s_j^T z_j \quad , \quad r_j = r_{j-1} - a_j z_j \quad ,$$

$$b_j = r_j^T r_j / r_{j-1}^T r_{j-1} \quad , \quad s_{j+1} = r_j + b_j s_j \quad ; \quad j = 1, \ldots, m, \ m \leq n$$

where m is the minimal number for which $\|r_m\| \leq \varepsilon_k \|g(x_{k-1})\|$, with $0 < \varepsilon_k < 1$, and $|h| \neq o$ - is a sufficiently small number.

This method is proposed in [10] and it is a result of application of DAA to the conjugate gradient method for solving system of the form: AAp + Ab = 0 [9]. Note that though the field of application of method 2 is wider than that of method 1, the former is double labor-consuming (the number of operations to calculate p_k when m = n is meant here).

Theorem 2. Let conditions (6) and (7) were satisfied, and the following unequality be true

$$\delta_1 \|y\|^2 \leq \|g'(x)y\|^2 \leq \delta_2 \|y\|^2, \quad \delta_2 \geq \delta_1 > 0 \qquad (8)$$
$$\forall\, x, y \in E^n$$

Then with a sufficiently small $|h|$ for method 2 we have the evaluation

$$\|p_k^{II} - p_k^H\| \leq (\varepsilon_k + C_2 |h| \; \|g(x_{k-1})\|) \, \|g(x_{k-1})\| / \delta_1$$

where $C_2 > 0$ depends only on δ_1, δ_2, L.

Method 3. Vector p_k^{III} in (1) to be calculated with the help of the following expression

$$p_k^{III} = \sum_{j=1}^{m} a_j s_j, \quad \text{where } s_0 = r_0 = -g(x_{k-1}) \quad ,$$

$$s_1 = [g(x_{k-1} + hr_0) - g(x_{k-1})]/h \quad ,$$

$$z_j = [g(x_{k-1} + hs_j) - g(x_{k-1})]/h \quad ,$$

$$\alpha_j = z_j^T s_j / s_j^T s_j \quad , \quad \beta_1 = 0 \quad ,$$

$$\beta_j = s_j^T s_j / s_{j-1}^T s_{j-1} \quad , \quad j \geq 2 \quad ,$$

$$a_j = r_{j-1}^T s_{j-1} / s_j^T s_j \quad , \quad r_j = r_{j-1} - a_j z_j \quad ,$$

$$s_{j+1} = z_j - \alpha_j s_j - \beta_j s_{j-1} \quad , \quad j = 1, \ldots, m \quad , \quad m \leq n \quad ,$$

here m is the minimal number for which $\|r_m\| \leq \varepsilon_k \|g(x_{k-1})\|$, where $0 < \varepsilon_k = < 1$, and $|h| \neq 0$ is a sufficiently small number.

The described method was proposed in [11] and is a result of application of DAA to a modification of the minimal iterations methods [12], (see also [9]). The method has the same merits as method 2, and its labor-consumption is almost equal to that of method 1.

Theorem 3. Let condition (6), (7), (8) be given. Then, if in method 3 $||r^0_{j-1}|| \leq D\,||s^0_j||$, $j = 1,2,\ldots,m-1$, where r^0_{j-1} and s^0_j are obtained $s^0_1 = g'(x_{k-1})\cdot r_0$, $z^0_j = g'(x_{k-1})\cdot s^0_j$, and r^0_{j-1} then with a sufficiently small $|h|$ evaluation

$$||p^{III}_K - p^H_k|| \leq (\varepsilon_k + C_3|h|\ \ ||g(x_{k-1})||)\,||g(x_{k-1})||/\sqrt{\delta_1}$$

is true, in which $C_3 > 0$ depends only on δ_1, δ_2, L, D.

In Theorems 1,2, and 3 the properties, which do not depend on λ_k, of vector p_k were analyzed. For specific ways of choosing λ_k the following statements are valid.

Theorem 4. Methods 1,2, and 3 within the frames of corresponding Theorems 1,2, and 3 with sufficiently small ε and $|h|$ ensure convergence of process 1, (2a), to the solution x^* of problem 1 from any initial point $x_0 \in E^n$. In this case convergence in the vicinity of x^* is quadratic, that is: $||x_k - x^*|| \leq C||x_{k-1} - x^*||^2$, where C is a certain constant.

Theorem 4 is valid not only relative to the way of choosing λ_k(2a), but it is true for all the ways proposed in [13].

Let conditions (5) and (6) with g substituted by f' be given for problem 2. If so, methods 1,2, and 3 with g substituted by f' are applicable for solving problem 2 as well and in this case, Theorem 1 remains correct, and Theorems 2 and 3 after putting $\sqrt{\delta_1} = \gamma_1$ and $\sqrt{\delta_2} = \gamma_2$ would be correct too.

Theorem 5. Methods 1,2, and 3, with f' instead of g, in the frames of the corresponding Theorems 1,2, and 3, ensure convergence of process (1), (2b) to the solution x^* of problem 2 from any initial point $x_0 \in E^n$, the rate of convergence being quadratic in the vicinity of x^*.

Theorem 5 is true as well for other ways of choosing λ_k which are described in [1, §8.3].

In the present work only three examples are given of DAA for solving problems 1 and 2 besides the examples from [2-6], but the field of DAA application might be sufficiently extended. In particular, for solving problem 1 in the frames of Theorem 1, one can apply DAA, for example, to the method of A-minimal iterations, to its binomial form, and to various forms of the s-step gradient method of steepest descent [9].

Memory volume required to realize the methods based on application of DAA to the method of A-minimal iterations and to its binomial form is equal to approximately 7n, just as in the case of methods 1,2, and 3.

It should be expected that the rate of convergence of sequence (1), (2a) of the problem 1 solution under the conditions of theorems 1 and 4 for these methods remains of the same high rate. Application of DAA to the s-step gradient method of steepest descent with the help of conjugate gradient scheme is only slightly different from method 1. Its specific feature is formulated like that: summation in the formula for p_k proceeds up to $s < m$. It is clear that at $s < m$ the rate of convergence of sequence (1) (2a) of the solution of problem 1 turns to be linear: $\|x_k - x^*\| \leq g \|x_{k-1} - x^*\|$, $0 < g < 1$. DAA is applicable for solving problem 2 under the restrictions of Theorem 2, for example to the method of columns orthogonalization [9], in which case matrix $A = g'(x)$ might as well be non-symmetric. These methods with g substituted by f' are also applicable for solving problem 2.

Note that the usage of DAA with the s-step gradient method of steepest descent for solving problem 2 is equivalent to the method of steepest descent at s = 1, and to a modification of discrete Newton method at s = m (method 1). In the authors' work [11, p. 160] a modification of the s-step method on the bases of DAA and method 1 for solving problem 2 in case of a nonconvex function f(x) is proposed. The essence of this modification as opposed to method 1 is in p_k, which is here expressed in that way:

$$p_k = \sum_{j=1}^{s} |a_j| s_j \quad ,$$

where $s < m$ or $s = m$. If the process of calculating p_k turns to be sufficiently stable, the direction of p_k would always be the direction of decreasing f(x) and the saddle point is not the point of attraction for it. Note also that the above s-step method for solving problems 1 and 2 opens the possibility to calculate p_k with variable number of steps s^k, $k = 1,2,\ldots,$. Taking small values $s^k < m$ at the beginning of the iteration process (1) and setting $s^k = m \leq n$ at high values of k, it is possible to realise an economic computational procedure for solving the above mentioned problems.

A deficiency of methods that use DAA for solving problems 1 and 2 is the errors in difference-approximation of derivatives. These errors can, to a certain degree, be controlled by the value of h. To this end it is possible to determine approximate values h_{opt} in the process of solving the above mentioned problems with the help of methods 1,2, and 3. It is to be done with the help of the following relations:

$$\text{a)}\quad h_{opt} = \operatorname*{argmin}_h \| g(x_{k-1} + hp_k) - g(x_{k-1}) + hg(x_{k-1}) \| / h$$

$$\text{b)}\quad h_{opt} = \operatorname*{argmin}_h \| f'(x_{k-1} + hp_k) - f'(x_{k-1}) + hf'(x_{k-1}) \| / h \qquad (9)$$

It is evident that the values under the sign of norm in (9) are difference-approximation of residual $R = A \cdot p_k + b$ for the system of equations (3). Another way of reducing the errors of difference approximations is, as was shown above, the use of a formula of higher level of accuracy in h instead of (4).

3. CONCLUDING REMARKS

Often in the course of solving a system of equations or an optimization problem information about eigenvalues of matrices $g'(x)$ and $f''(x)$ might become useful. It is known [9], that while solving problem 3 with the help of the methods of conjugate gradients and minimal iterations one can in parallel to the main process of calculations compute coefficients of the matrix A characteristic polinominal and, after determination of eigenvalues, it is not difficult to calculate eigenvectors of this matrix. Thus, by way of not too complex additional calculations in methods 1 and 3 it is possible to approximately determine eigenvalues and eigenvectors of matrices $g'(x_{k-1})$ and $f''(x_{k-1})$. The characteristic polinominal in methods 1 and 3 in this case is obtainable when $m = n$, which corresponds to the case of absence of multiple eigenvalues. As for non-symmetric matrices $g'(x)$ their characteristic polinomial can be obtained with the help of application of DAA to the method of orthogonalization of sequential iterations [9]. Of course, in all the discussed cases we shall receive approximate values of polinomial coefficients and eigenvectors and they would be the more accurate the less the value of $|h|$.

A number of numerical experiments were accomplished with the methods proposed in the present work. The results had shown, that the methods are not inferior to the discrete variant of Newton method in terms of the rate of divergency [8, p. 389, Algorithm A89].

REFERENCES

Birjukov, A.G. Annals of M.Ph.T.I., Series "Airphysics and Applied Mathemativs". Dolgoprudnyi, M.Ph.T.I., 1979, p. 188-92.

Birjukov, A.G. Collected Articles "Airphysics and Applied Mathematics". M.Ph.T.I., 1981, p. 99-101.

Birjukov, A.G. Proceedings of the Conference of M.Ph.T.I. 1977, Series "Airphysics and Applied Mathematics". Dolgoprudnyi, M.Ph.T.I. 1978, p. 125-28.

Brown, K.M. "Numerical Solution of Systems of Nonlinear Algebric Equations", ed. G.D. Byrne and C.A. Hall, New York-London, Acad. Press, 1973, pp. 281-348.

Burdakov, O.P. Doklady Acad. Nauk, 1980, v. 254, no. 3, p. 521-23.

Danilin, Y.M. Doklady Acad. Nauk, 1974, v. 218, no. 3, p.513-16.

Dennemeyer, R.F. and E.H. Mookini, I. Opt. Theory and Appl. 1975, v. 16, no. 1/2, p. 67-85.

Evtushenko, Y.G. Methods of Solving Extremal Problems and their Application in Optimization Systems. M., Nauka, 1982.

Faddeev, D.K. and V.N. Faddeva, Computation Methods of Linear Algebra. M.-L., Physmatgitz, 1963.

Fridman, V.M. Journal of Comp. Math. and Math. Phys., 1962, v. 2, p. 341-2.

Hestenes, M.R. Conjugate Direction Methods in Optimization. N.Y. Springer-Verlag, 1980.

Ortega, J. and W. Rheinboldt, Iterative solution of nonlinear Equations in Several Variables. Acad. Press, New York-London, 1970.

Pshenichyi, B.N. and N.W. Redkovskii, Journal of Computer Math. and Math. Phys., 1976, v. 16, no. 6, p. 1388-96.