AN APPLICATION OF A CERTAIN MEHOD OF DIFFERENTIAL DESCENT TO THE SOLUTION OF OPERATOR EQUATIONS

Otarov Amanbay

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

The construction of methods for solving applied problems is of несомненную актуальность In this case, the following requirements for the method are of particular importance: algorithmic simplicity and speed, estimation of the accuracy of an approximate solution; minimum a priori information about the desired solution; certain universality of the numerical algorithm. Bearing in mind a large number of specific mathematical problems (integral equations, boundary value problems for differential equations, etc.), it is convenient to study approximate methods immediately for some classes of equations, that is, in the form of operator equations. This work is a direct generalization and development of the method of connected differential descent [1,2,3] for solving systems of finite-dimensional equations as applied to solving operator equations considered in separable Banach spaces. The differential descent method based on the solution of the Cauchy problem was considered in [4]. S. M. Gerashenko [5] investigated the possibility of improving the convergence of differential descent methods. To dampen oscillations near the extremum point, an additional coefficient is introduced into the right side of the system of differential equations in order to increase the roughness of the system with respect to the calculation error. The so-called sliding mode is introduced into the search algorithm. The study of sliding mode differential descent methods was continued in [4,6]. In [6], the rate of convergence of such methods is studied. In articles by B. A. Galanov [1,2], S. I. Alyber and Ya. I. Alyber [7], the differential descent method was applied to solve a system of equations. The method makes it possible to obtain an exact solution.

Keywords: Recurrent sequence, Cauchy problem, scalar equation.

Karakalpak State University Otarova Jamila Amanbaevna Karakalpak State University

DOI: 10.31838/ecb/2023.12.s3.249

1. Introduction

B is the real separable Banach space; B^* is the conjugated with space B; $\|x\|_B$ is a norm of the element $x \in B$; H is a separable Hilbert space; (f, x) is a value of the inear element $x \in B$ with respect to x functional $f \in B^*$; $D(A)$ is a domain of definition, and $R(A)$ is the domain of values of the operator A; $l_p\,(p \geq 1)$ is a space of numerical sequences $x = (\xi_1, \xi_2, ..., \xi_k, ...)$, converges to a p- degree and with a norm $\|x\| = \left[\sum_{k=1}^{\infty} |\xi_k|^p\right]^{\frac{1}{p}}$; α_i - solution of the first i equation of the system (2); α_{ni} is an element belonging to α_i and corresponding to the original condition of x_n, $n = 0,1,2,...$ for the recurrence sequence of the Cauchy problem (6); x_{ni} approximating to α_{ni}; n_i is a number of segment divisions $\left[f_i(x_{ni-1}), 0\right]$; $h_i = f_i(x_{ni-1}) / n_i$ is a step of integration for the approximate intergration of the Cauchy problem (6).

2. Results

Let the equation be given as

$$P(x) = \theta, \qquad (1)$$

with twice continiously Frechet differentiable nonlinear operator of P, acting from B onto a real Hilbert space H with zero θ. By virtue of separability, in H there exists [8,9,10] the system $\{\varphi_i\}_1^{\infty}$ of elements φ_i, $i = 1,2,...$, satisfying the conditions: 1) system $\{\varphi_i\}_1^{\infty}$ is full in H; 2) elements $\varphi_1, \varphi_2, ...\varphi_i$ are linearly independent for any i. Then a system of scalar equations

$$f_i(x) = (P(x), \varphi_i) = 0, \quad i = 1,2,... \qquad (2)$$

is equavalent to the equation (1) and fucntional $f_i(x)$, $i = 1,2...$ defined by the left-hand side (2), have the domain of definition that conicides with the opereator P on this domain. In particular, if $\{\varphi_i\}_1^{\infty}$ is an orthonormal system of elements in H with properties 1) and 2), then functionals $f_i(x)$, $i = 1,2,...$ coincide with Fourier coefficients of the element $P(x) \in H$, $x \in B$. Then by the well-known Riesz – Fischer Theorem [7,9,14] the series $\sum_{i=1}^{\infty} (P(x), \varphi_i)^2 \sum_{i=1}^{\infty} = f_i^2(x)$ converges for any $x \in B$, i.e.:

$$f(x) = \left((P(x), \varphi_1), (P(x), \varphi_2), ...\right)$$

is an element of the space l_2 for all $x \in B$.

In (2), the elements of the system $\{\varphi_i\}_1^{\infty}$ are subject to the condition: 3) $\varphi_i \in D(grad\ f_i(x))$ is the domain of definition $f_i(grad\ f_i(x))$, i.e. elements of the system $\{\varphi_i\}_1^{\infty}$ are selected so that they satisfy certain boundary conditions that are natural for the functional $f_i(x)$. This latter condition ensures Frechet differentiability [8,9,10] of the functional $f_i(x)$ and allows to represent Frechet differential $f_i(x)$ in the form of:

$$Df_i(x,h) = \left(\left(P'_x\right)h, \varphi_i\right) = \left(\left(P'_x\right)^* \varphi_i, h\right),\ h \in B,$$

where P'_x is the Frechet derivative of the operator P at a point $x \in B$, and $\left(P'_X\right)^*$ is the operator conjugate with P'_x.

Therefore, we obtain the following expression for the gradient of the functional $f_i(x)$

$$f'_i = grad\ f_i(x) = \left(P'_x\right)^* \varphi_i,,\quad i = 1,2,...$$

The system of equations (2) can be represented in the following form too:

$$f(x) = \left(f_1(x), f_2(x),...\right) = 0,, \qquad (3)$$

where $x \in B$, $f(x) \in l_p\ (p \geq 1)$.

We shall assume that for all $x \in B$ the following conditions are satisfied:

a) the functionals функционалы $f_i(x), i = 1,2,...$ have continious Frechet derivatives up to the second order inclusive;

b) the operator $a_i^{i-1}(x)$ acting from B onto B and constructed from $f_j(x)$, $j = 1,2,...,i$ has continious Frechet derivatives;

c) the operator $a_i^{i-1}(x)$ is constructed so that

$$\left(f'_j, a_i^{i-1}\right) = 0,\quad j \leq i-1, \tag{4}$$

$$\left\| \frac{a_i^{i-1}}{\left(f'_i, a_i^{i-1}\right)} \right\|_B \leq K_i = const < \infty,\ f'_j = grad\ f_j(x) \in B^* \tag{5}$$

Under the assumptions made, the following theorem holds.

Theorem 1. The solutions $x_i(t_i)$, $i = 1,2,...$ of the recurrence sequence of the Cauchy problems

$$\frac{dx_i}{dt_i} = \frac{a_i^{i-1}(x)}{\left(f'_i, a_i^{i-1}\right)},,$$

$$x_i^0 = x_i\left(t_i = t_i^0\right) = x_{i-1}\left(t_{i-1} - 0\right) = \alpha_{0i-1},,$$

$$t_i^0 = f_i\left(\alpha_{0i-1}\right),\quad \alpha_{00} = x_0,\quad i = 1,2,... \tag{6}$$

are determined, at least, on such segments as $\left[t_i^0, \delta_i\right]$, что $0 \in \left[t_i^0, \delta_i\right]$ и $\alpha_{0i} = x_i(0)$.

We omit the proof of the theorem since it is analogous to the proof of the corresponding theorem in the works [1,2].

The conditions of the theorem guarantees the existance and uniqueness of solution for the Cauchy problem (6) for any i.

If for all $x \in B$ functionals $f_i(x), i = 1,2,...$ and operators $a_i^{i-1}(x)$, $i = 1,2,...$ from B onto B, constructed from $f_j(x)$, $j = 1,2,...$ are such that conditions are b) the theorem holds, and the right hand sides of the differential equations (6) satisfy a Lipschitz condition [8,11], then the solution $x_i(t_i)$, $i=1,2,...$ depends continiuosly on initial data, i.e. each problem is correct.

For the Cauchy problems (6) the first integrals are known for $i=n+1$:

$$f_j\left(x_{n+1}\left(t_{n+1}\right)\right)=0, \quad j\le n,$$
$$t_{n+1}-f_{n+1}\left(x_{n+1}\left(t_{n+1}\right)\right)=0. \tag{7}$$

It follows that the Cauchy problem (6) can be used for vanishing the residual of the i - th equation system (2) for zero residuals $(i-1)$ - of equations of this system. The alternative of the differential descent method being considered requires construction of the operators $a_i^{i-1}(x)\in B,\ x\in B$, $i=1,2,...$, ensuring the fulfillment of the conditions (4) and (5). Such operators can be constructed in many ways [3,12]. In particular, operators $\left\{a_i^{i-1}(x)\right\}_1^{\infty}$ can be constructed by the formula [12]:

$$a_i^j=a_i^{j-1}-a_j^{j-1}\frac{\left(f_j',\,a_i^{j-1}\right)}{\left(f',\,a_j^{j-1}\right)}, \qquad j<i\ , \tag{8}$$

where i, j are whole numbers.

Let $\left\{a_i^0(x)\right\}\subset B,,\ \ x\in B$ be a set of linearly independent elements, for which $\left(f_j',a_j^{j-1}\right)\neq 0,$ $j=1,2,...$. Then the formula (8) determines the elements $a_i^j(x)\in B$, $x\in B$ with properties [12]:

$$\left(f_k',a_i^j\right)=0,,\quad k\le j\ , \tag{9}$$

$$\left(f_i',\,a_i^j\right)\neq 0,,\quad j=i-1\ , \tag{10}$$

i.e. the superscript to the element $a_i^j(x)$ implies the property of (9), subscript implies the property of (10). If $a_i^0=f_i'\in B=B^*$, $i=1,2,...$, then to satisfy the properties (9), (10) linearly independent elements f_i', $i=1,2,...$ [8,9,13] are sufficient. And, if $f_i'\in B^*$, $i=1,2,...$ and T is a linear operator from B^* onto B, then we can put

$$a_i^0=T\ f_i', \quad i=1,2,... \tag{11}$$

Indeed, if for T there is bounded inverse operator T^{-1}, then the linear independece and completeness of the system $\left\{a_i^0=f_i'\right\}_1^{\infty}$ implies linear independence and completeness of the system (11) [10,14].

Instead of the system (2) a more general system of equations can be used:

$$\begin{aligned}
&\Phi_1(x)=f_1(x)=\left(P(x),\varphi_1\right)=0,\\
&\Phi_2(x)=f_2(x)-\Phi_2(x)\left(\varphi_1,\varphi_2\right)=0,\\
&\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\\
&\Phi_i(x)=f_i(x)-\sum_{k=1}^{i-1}\Phi_k(x)\left(\varphi_k,\varphi_i\right)=0,\quad i=3,4,...
\end{aligned} \tag{12}$$

If the system of elements $\left\{\varphi_i\right\}_1^{\infty}$ is orthogonal, then as a special case, we obtain the system (2).

It is easy to verify that the systems of equations (2) and (12) are equivalent: any solution of the system (2) satisfies the system (12) and vice versa.

In [14], it is shown that for $B = L_2(G)$ (G is the domain of variation of the variable point $Q, x(Q) \in B$) series, the coefficients of which are determined by the left-hand of the system (12) give a slightly better approximation to the function being approximated as compared to the Fourier series for the approximately arthonormal system $\{\varphi_i\}_1^\infty$.

In practice, the construction of a large number of orthonormalized elements is associated with significant computational difficulties [14], and in orthonormalization, approximately arthonormalized systems of elements are always obtained. Thence, the advantage of the system (12) over the system (2) becomes clear when using approximately arthonormal systems $\{\varphi_i\}_1^\infty$.

An approximate solution to the equation (1) implies the element $\alpha_{0m} \in B$, determined by the first m of the Cuachy problem from (6) and satisfying the truncated system (2), $i = 1, 2, ..., m$.

From a computational point of view, integration of the equation (6) suggests the use of some numerical integration formula [15,16] and in practice, the descent for solution to the system (2) occurs along the line determinedd by the approximate solution of the Cauchy problem (6) for $i = 1, 2, ..., m$. Here, the relations (7) can be used for the control of accuracy of the apprximate solution to the problem. Similarly to [2], iteration processes can be constructed. For example, using Euler's Method [15,16] with the step of $h_i = f_i(x_{ni-1})/n_i$ the error of which h_i (the equation (6) is solved with the initial conditions $t_{ni-1} = f_i(x_{ni-1})$, x_{ni-1} and it is supposed that $a_i^{i-1}(x)$, $f_i(x)$ a sufficient number of times are continuously differentiable B) an iteration method can be obtained

$$x_{n+1j} = x_{ni} - \frac{a_i^{i-1}(x_{ni-1})}{\left(f_i'(x_{ni-1}), a_i^{i-1}(x_{ni-1})\right)} h_i, \qquad (13)$$
$$n = 0,1,2,...; \quad i = 1,2,...,m$$

where $\alpha_{00} = x_{00} = x_0 \in B$ is an initial approximation to the solution of the system (2).

Applying various methods of numerical integration of the ordinary differential equations (methods of Runge-Kutta, Adama, etc.) to the equation (6), it is possible to obtain in a new way both many well-known and new iteration methods for solving the system of equation (2). Thus, the methods can be described analitically and subordinated to a single scheme.

If the equation is given

$$P(x) = Ax - g = \theta, \qquad (14)$$

where A is a linear operator acting from the dense set $D(A) \subset B$ onto the set $R(A) \subset H$ and having an bounded inverse operator A^{-1}, [8,9,11], $g \in R(A)$ is a fixed element, then

$$f_i(x) = (Ax - g, \varphi_i) = 0, \quad i = 1,2,... \qquad (15)$$

$f_i = A^*\varphi_i$, $i = 1, 2, ...$ Then, if operators a_i^{i-1}, $i = 1, 2, ...$ are calculated for $a_i^0 = f_i' = A^*\varphi_i$, $i = 1, 2, ...$ as per the formula (8), then the right-hand side of the differential equation (6) does not depend on x. Therefore, by integrating the equation (6) with consideration that a_i^{i-1} does not depend on x, to calculate the elements of the sequence α_{0i}, $i = 1, 2, ...$ we obtain the following formula:

$$\alpha_{0i} = \alpha_{0i-1} - \frac{f_i(\alpha_{0i-1})}{\left(A^*\varphi_i, a_i^{i-1}\right)} a_i^{i-1},, \quad i = 1,2,... , \qquad (16)$$

where $\alpha_{00} \in D(A)$, is an arbitrary element.

In exact calculation using the formula (16) at the m -th step, the element $\alpha_{0m} \in D(A)$ is obtained which is the solution of the first m of the equation system (15). If the calculations are approximate, then the process can be made iterative.

The recurrence sequence of the Cauchy problems (6) defines a sequence of elements α_{0i}, $i = 1, 2, \ldots$. However, the convergence condition for $i \to \infty$ to the zero of the residual $\| f(\alpha_{0i}) \|_{l_p}$, and elements α_{0i}, , $i = 1, 2, \ldots$. to the solution of the equation (1) remain unclear. Let's consider some of these conditions.

Let us say that $f(x)$ has the property of (l_p), if $f(\alpha_{0i}) \in W \subset l_p$, , $i = 1, 2, \ldots$ and for any $\varepsilon > 0$ and for all $\xi = (\xi_1, \xi_1, \ldots) \in W$, there exists n_ε such that

$$\sum_{n=N}^{\infty} |\xi_n| < \varepsilon^p, , \quad N \geq n_\varepsilon,$$

The property is trivial in case, $l_p = l_p^{(n)}$, [8,9] (l_p)

Let us prove the theorem.

Theorem 2. In order to

$$\lim_{i \to \infty} \| f(\alpha_{0i}) \|_{l_p} = 0,$$

it is necessary and sufficient that the operator $f(x)$ has the property of (l_p), .

Proof. The necessity of the theorem is obvious. Let us prove the sufficiency.

Let $f(x)$ c (l_p) be the property. Then for any $\varepsilon > 0$ there exists such n_ε, , that

$$\sum_{n=N}^{\infty} \left| f_n(\alpha_{0i}) \right|^p < \varepsilon^p, , \quad N \geq n_\varepsilon,$$

Therefore, for all $i \geq n_\varepsilon$, we have

$$\| f(\alpha_{oi}) \|_{l_p}^p = \sum_{j=1}^{\infty} \left| f_j(\alpha_{oi}) \right|^p =$$

$$= \sum_{j=1}^{n_\varepsilon} \left| f_j(\alpha_{oi}) \right|^p + \sum_{j=n_\varepsilon+1}^{\infty} \left| f_j(\alpha_{oi}) \right|^p < \varepsilon^p,$$

and there exixts such $i = n_\varepsilon$, , that the set $\overline{W} = \{ f(\alpha_{0i}) : i \geq n_\varepsilon \}$, is bounded. In combination with the (l_p), property, it means the set $\overline{W} \subset l_p$ is compact, i.e. the possibility of choosing such a subsequence $\alpha_{0ik} \in \overline{W}$, $k = 1, 2, \ldots$, for which

$$\lim_{k \to \infty} \| f(\alpha_{oik}) \|_{l_p} = 0, ,$$

which proves the theorem.

If in (2) $\{\varphi_i\}_1^\infty \subset H$ is an orthonormal system of elements, and P is such an operator that $f(x)$ has (l_2) property and satisfies the conditions of the theorem 1, then

$$\lim_{i\to\infty}\left\|P\left(\alpha_{oi}\right)\right\|_H = 0, \tag{16}$$

Hence, for operators P with bounded inverse, it follows that the sequence convergence $\{\alpha_{oi}\},\ i=1,2,...$ as well to the solution of the equation (1).

In particular, for the equation (14) from (16) it follows that

$$\lim_{i\to\infty}\alpha_{0i} = \alpha,,$$

where α is a solution of the equations (14) and (15).

Thus, Theorem 2 gives necessary and sufficient condition for the convergence to the zero of the residual $\left\|P\left(\alpha_{0i}\right)\right\|_H$, for $i\to\infty$. However, in the general case verification of the conditions of theorem 2 is difficult and is determined by the specific form of the system $\{\varphi_i\}_1^\infty$ and properties of the operator P.

In conclusion, we point out some of features of the considered method compared to other descent methods [7,8,17].

1. The solutions $x_i\left(t_i\right)$, in the recurrence sequence of the Cauchy problems (6) exist on the known finite intervals $\left[t_i^0,\ 0\right]$, and the first integrals are known for problems (6).

This allows to control the movement along the descent trajectory defined by the solutions of the problems (6) and creates some convinience in programming.

2. When implementing the ordinary sheme of the descent method [7,8,17], the minimum of the function of one independent variable is to be found at each step of the iterative process which requires additional computing work. The considered scheme relieves the computer man from this difficulty and it is distinguished by high algorithmicity [1,2,3].

3. Descent methods based on the idea of minimizing some real functional may not converge to a solution if we encounter a local minimum of this functional.

The nature of the considerations from which we proceeded when compiling system (2) shows that no connection of this equation (1) with variational problems has been used. Therefore, the functionals $f_i\left(x\right),,\ i=1,2,...$ of method (6) are not connected at all with variational problems and the above phenomenon, which is characteristic of known descent methods, is absent in the considered differential descent method.

4. The variety of methods for constructing operators $\left\{a_i^{i-1}\right\}_1^\infty$, [3,12] generates a variety of recurrence sequences of the problems (6) for each specific equation. In turn, various methods of numerical integration of the Cauchy problems (6) [15,16], generate a whole class of iterative processes each of which determines an approximate solution of a given equation and has its own region of convergence. Therefore, if one or another iterative process does not give the desired result, it is possible to vary the sequences of the Cauchy problems (6) and approximate methods for their solution.

3. References

Галанов Б.А. Об одном методе связного дифференциального спуска // В сб. «Математические методы в кибериетической технике», вып.1, Киев 1980, изд-во ИКАН УССР, с. 26-28.

Галанов Б.А. О сходимости одного непрерывного метода и его аппроксимаций // В сб. «Математические методы в кибернетической технике», вып. 7, Киев, 1970, изд-во ИКАН УССР, с. 31-40.

Отаров А.О. Численная реализация одного метода дифференциального спуска при решении нелинейных уравнений // Журнал «Вестник» КК филиала АН УзССР, г. Нукус, 1983, №2, с. 3-8.

Геращенко С. М., Применение теории автоматического регулирования к методу дифференциального спуска. Тр. II Респ. конференции математиков Белоруссии. Минск, Белорусск. ун-т, 1969, 179—191 (РЖМат, 1969, 7Б608).

Геращенко С. М., О выборе правых частей в системе дифференциальных уравнений градиентного метода. Дифференц. уравнения, 1967, 3, № 1!2, 2144-2150 (РЖМат, 1968, 85183).

Геращенко С. М., Сравнение по скорости сходимости некоторых модификаций дифференциального спуска. Дифференц. уравнения, 1970, б, № 10,1810—1817 (РЖМат, 1971, 3Б585).

Альбер С.И., Альбер Я.И. Применение метода дифференциального спуска для решения

нелинейных систем // ЖВМ и МФ, 1967, 7, №1, с. 14-32.
Канторович Л.В., Акилов Г.П. Функциональный анализ. – М.: Наука, 2004, 816 с.
Люстерник Л.А., Соболев В.И. Краткий курс функционального анализа. – М.: Высшая школа, 2009, 279 с.
Функциональный анализ / Под общей редакцией С.Г. Крейна. – М.: Наука, 1972, 544 с.
Красносельский М.А., Вайникко Г.М., Забрейко П.П., Рутицкий Я.Б., Стеценко В.Я. Приближенное решение операторных уравнений. – М.: Наука, 1969, 456 с.
Фаддеев Д.К., Фаддеева В. Н. Вычислительные методы линейной алгебры. – М.: Физматгиз, 2009, 736 с.
Михлин С.Г. Вариационные методы в математической физике. – М.: Наука, 1970, 512 с.
Алексидзе М.А. Решение граничных задач методом разложения по неортогональным функциям.-М.: Наука, 1978, 352 с.
Бахвалов Н.В., Жидков Н.П., Кобельков Г.М. Численные методы. – М.: Наука, 2021, 636 с.
Самарский А.А., Гулин А.В. Численные методы. – М.: Наука, 1989, 430 с.
Гавурин М.К. Нелинейные функциональные уравнения и непрерывные аналоги итерационных методов // Известия вузов, Математика, 1958, №5, с.18-31.