



**HAL**  
open science

# AN ITERATIVE METHOD FOR SOLVING NONLINEAR LEAST SQUARES PROBLEMS WITH NONDIFFERENTIABLE OPERATOR

Roman Iakymchuk, Stepan Shakhno, Halyna Yarmola

► **To cite this version:**

Roman Iakymchuk, Stepan Shakhno, Halyna Yarmola. AN ITERATIVE METHOD FOR SOLVING NONLINEAR LEAST SQUARES PROBLEMS WITH NONDIFFERENTIABLE OPERATOR. 2018. hal-01709582

**HAL Id: hal-01709582**

**<https://hal.science/hal-01709582>**

Preprint submitted on 15 Feb 2018

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

УДК 519.6

S. M. SHAKHNO, R. P. IAKYMCHUK, H. P. YARMOLA

**AN ITERATIVE METHOD FOR SOLVING NONLINEAR LEAST SQUARES PROBLEMS WITH NONDIFFERENTIABLE OPERATOR**

S. M. Shakhno, R. P. Iakymchuk, H. P. Yarmola. *An iterative method for solving nonlinear least squares problems with nondifferentiable operator*, Mat. Stud. ? (201?), ?-?.

An iterative differential-difference method for solving nonlinear least squares problems is proposed and studied. The method uses the sum of the derivative of the differentiable part of the operator and the divided difference of the nondifferentiable part instead of computing Jacobian. We prove the local convergence of the proposed method and compute its convergence rate. Finally, we carry out numerical experiments on a set of test problems.

**1. Introduction.** Nonlinear least squares problem often arise while solving overdetermined systems of nonlinear equations, estimating parameters of physical processes by measurement results, constructing nonlinear regression models for solving engineering problems, etc. Effective methods for solving nonlinear least squares problems are the Gauss-Newton method and its modifications ([1, 4, 5, 6, 7]). However, in practice, calculation of derivatives could either be very difficult or impossible. For instance, functions can be too complex, then derivatives may be computed approximately, or only values of functions are given (obtained from experiments) at certain points but it is known that those functions are nonlinear. Hence, one can use the iterative-difference methods ([1, 2, 8, 13]) that do not require calculation of derivatives and yet approach the Gauss-Newton method in terms of the convergence rate and the number of iterations.

In case when the nonlinear function has a differentiable and a nondifferentiable parts, one can employ iterative-difference methods from [1, 2, 8, 13]. However, one would preferably build iterative methods that take into account properties of the problem to solve. This is the approach we would like to follow here. In particular, we can use only the derivative of the differentiable operator instead of the full Jacobian, which, in fact, does not exist. In general, the methods obtained using this approach converge slowly. There are some good and efficient examples [1, 3, 9, 14] that use a sum of the derivative of the differentiable part of the operator and the divided difference of the nondifferentiable part instead of the Jacobian, however for solving nonlinear equations. In this work, we propose to follow this approach and design a novel combined method for solving nonlinear least squares problems. This method is based on the Gauss-Newton method, which is employed for the differentiable part of the operator, and the Secant type's method, which relies upon divided differences for the nondifferentiable part. We study the local convergence of this method under the classic and generalized Lipschitz conditions. In the latter, we use some positive integrable

---

2010 *Mathematics Subject Classification*: 49M15, 65B05, 65H10, 65K10.

*Keywords*: nonlinear least squares problem; differential-difference method; divided differences; radius of convergence; residual; error estimates.

function instead of the Lipschitz constant. The generalized Lipschitz conditions for divided differences were introduced in [11, 12]. To note, these conditions were successfully applied to study the convergence of the combined Newton-Secant method ([9]) and the two-step combined method ([10]) for nonlinear equations with nondifferentiable operator. On a set of test problems, we show performance results of the derived method and compare these results against the Secant type's method ([8, 13]) and the Gauss-Newton type's method.

**2. Formulation of the problem.** Let us consider the nonlinear least squares problem

$$\min_{x \in \mathbb{R}^p} \frac{1}{2} (F(x) + G(x))^T (F(x) + G(x)), \quad (1)$$

where residual function  $F + G: \mathbb{R}^p \rightarrow \mathbb{R}^m$  ( $m \geq p$ ) is nonlinear by  $x$ ,  $F$  is a continuously differentiable function,  $G$  is a continuous function, differentiability of which, in general, is not required.

For finding the solution of the problem (1), we propose the following modification of the Gauss-Newton method combined with the Secant type's method

$$x_{n+1} = x_n - (A_n^T A_n)^{-1} A_n^T (F(x_n) + G(x_n)), n = 0, 1, \dots, \quad (2)$$

where  $A_n = F'(x_n) + G(x_n, x_{n-1})$ ,  $F'(x_n)$  is a Fréchet derivative of  $F(x)$ ;  $G(x_n, x_{n-1})$  is a divided difference of the first order of function  $G(x)$  ([15]) at points  $x_n, x_{n-1}$ ;  $x_0, x_{-1}$  are given.

Setting  $A_n = F'(x_n)$ , for solving the problem (1), from (2) we get an iterative Gauss-Newton type's method

$$x_{n+1} = x_n - (F'(x_n)^T F'(x_n))^{-1} F'(x_n)^T (F(x_n) + G(x_n)), n = 0, 1, \dots \quad (3)$$

For  $m = n$ , the problem (1) turns into a system of nonlinear equations

$$F(x) + G(x) = 0.$$

In this case, the method (2) is transformed into the combined Newton-Secant method ([3, 14])

$$x_{n+1} = x_n - (F'(x_n) + G(x_n, x_{n-1}))^{-1} (F(x_n) + G(x_n)), n = 0, 1, \dots$$

and the method (3) into the Newton type method for solving nonlinear equations ([18])

$$x_{n+1} = x_n - (F'(x_n))^{-1} (F(x_n) + G(x_n)), n = 0, 1, \dots$$

**3. Analysis of the local convergence of the combined method (2).** Let us, at first, consider some auxiliary lemmas needed to obtain the main results.

**Lemma 1.** Let  $e(t) = \int_0^t E(u) du$ , where  $E$  is an integrable and positive nondecreasing function on  $[0, T]$ . Then,  $e(t)$  is monotonically increasing with respect to  $t$  on  $[0, T]$ .

**Lemma 2** ([6, 17]). Let  $h(t) = \frac{1}{t} \int_0^t H(u) du$ , where  $H$  is an integrable and positive nondecreasing function on  $[0, T]$ . Then  $h(t)$  is nondecreasing with respect to  $t$  on  $(0, T]$ .

Additionally,  $h(t)$  at  $t = 0$  is defined as  $h(0) = \lim_{t \rightarrow 0} \left( \frac{1}{t} \int_0^t H(u) du \right)$ .

**Lemma 3** ([16]). *Let  $s(t) = \frac{1}{t^2} \int_0^t S(u)u du$ , where  $S$  is an integrable and positive nondecreasing function on  $[0, T]$ . Then  $s(t)$  is nondecreasing with respect to  $t$  on  $(0, T]$ .*

The local convergence and its rate for the iterative process (2) are studied in the theorem below. We use the Euclidean norm. Note that for the Euclidean norm  $\|A - B\| = \|A^T - B^T\|$ , where  $A, B \in \mathbb{R}^{m \times p}$ .

**Theorem 1.** *Let  $F + G: \mathbb{R}^p \rightarrow \mathbb{R}^m$  be continuous on an open convex subset  $D \subset \mathbb{R}^p$ , and  $F$  is a continuously differentiable function,  $G$  is a continuous function. Suppose that the problem (1) has a solution  $x^* \in D$ , and the inverse operation  $(A_*^T A_*)^{-1} = [(F'(x^*) + G(x^*, x^*))^T \times (F'(x^*) + G(x^*, x^*))]^{-1}$  exists, such that  $\|(A_*^T A_*)^{-1}\| \leq B$ .*

*On the subset  $D$ , the Fréchet derivative  $F'$  satisfies the radius Lipschitz condition with  $L$  average*

$$\|F'(x) - F'(x^\tau)\| \leq \int_{\tau\rho(x)}^{\rho(x)} L(u)du, \quad x^\tau = x^* + \tau(x - x^*), \quad 0 \leq \tau \leq 1, \quad (4)$$

*the function  $G$  has the first order divided difference  $G(x, y)$ , and*

$$\|G(x, y) - G(u, v)\| \leq \int_0^{\|x-u\|+\|y-v\|} M(u)du \quad (5)$$

*for all  $x, y, u, v \in D$ ,  $\rho(x) = \|x - x^*\|$ ;  $L, M$  are positive nondecreasing functions on  $[0, 2R]$ ,  $R > 0$ .*

*Furthermore,*

$$\|F(x^*) + G(x^*)\| \leq \eta, \quad \|F'(x^*) + G(x^*, x^*)\| \leq \alpha; \quad \frac{B}{R} \left( \int_0^R L(u)du + \int_0^{2R} M(u)du \right) \eta < 1$$

*and  $\Omega = \Omega(x^*, r_*) = \{x: \|x - x^*\| < r_*\} \subseteq D$ , where  $r_*$  is the unique positive zero of the function  $q$  given by*

$$q(r) = B \left[ \left( \alpha + \int_0^r L(u)du + \int_0^{2r} M(u)du \right) \left( \int_0^r L(u)du + \int_0^r M(u)du \right) + \left( \frac{1}{r} \int_0^r L(u)du + \frac{1}{r} \int_0^{2r} M(u)du \right) \eta \right] + B \left[ 2\alpha + \int_0^r L(u)du + \int_0^{2r} M(u)du \right] \left[ \int_0^r L(u)du + \int_0^{2r} M(u)du \right] - 1.$$

*Then, for  $x_0, x_{-1} \in \Omega$ , the iterative process  $\{x_n\}$ ,  $n = 0, 1, \dots$ , generated by (2), is well defined, remains in  $\Omega$ , and converges to  $x^*$ . Moreover, the following error estimates hold for all  $n \geq 0$*

$$\|x_{n+1} - x^*\| \leq C_1 \|x_{n-1} - x^*\| + C_2 \|x_n - x^*\| + C_3 \|x_{n-1} - x^*\| \|x_n - x^*\| + C_4 \|x_n - x^*\|^2, \quad (6)$$

*where*

$$g(r) = B \left[ 1 - B \left( 2\alpha + \int_0^r L(u)du + \int_0^{2r} M(u)du \right) \left( \int_0^r L(u)du + \int_0^{2r} M(u)du \right) \right]^{-1};$$

$$C_1 = g(r_*) \frac{1}{2r_*} \int_0^{2r_*} M(u)du \eta; \quad C_2 = g(r_*) \left( \frac{1}{r_*} \int_0^{r_*} L(u)du + \frac{1}{2r_*} \int_0^{2r_*} M(u)du \right) \eta;$$

$$C_3 = g(r_*) \left( \alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right) \frac{1}{r_*} \int_0^{r_*} M(u) du;$$

$$C_4 = g(r_*) \left( \alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right) \frac{1}{r_*} \int_0^{r_*} L(u) u du.$$

*Proof.* According to l'Hôpital's rule we get

$$\lim_{r \rightarrow 0} \frac{1}{r} \int_0^r L(u) du = \lim_{r \rightarrow 0} \frac{L(r)}{1} = L(0), \quad \lim_{r \rightarrow 0} \frac{1}{r} \int_0^{2r} M(u) du = \lim_{r \rightarrow 0} \frac{M(2r)}{1} = M(0).$$

Taking into account Lemma 1 for a sufficiently small  $\eta$ ,  $q(0) = B(L(0) + M(0))\eta - 1 < 0$ . With a sufficiently large  $R$ , the inequality  $q(R) > 0$  holds. Taking into account the intermediate value theorem, the function  $q$  has a positive zero on  $(0, R)$  denoted by  $r_*$ . Moreover, this zero is the only one on  $(0, R)$ .

Indeed, according to Lemma 2, the function  $(\frac{1}{r} \int_0^r L(u) du + \frac{1}{r} \int_0^{2r} M(u) du)\eta$  is non-decreasing with respect to  $r$  on  $(0, R]$ . By Lemma 1, functions  $\int_0^r L(u) du$ ,  $\int_0^r M(u) du$ , and  $\int_0^{2r} M(u) du$  are monotonically increasing on  $[0, R]$ . Also, by Lemma 3, the function  $\int_0^r L(u) u du = r^2 (\frac{1}{r^2} \int_0^r L(u) u du)$  is monotonically increasing with respect to  $r$  on  $(0, R]$ . Therefore,  $q(r)$  is monotonically increasing on  $(0, R]$ . Thus, the graph of function  $q(r)$  crosses the positive  $r$ -axis only once on  $(0, R)$ .

We denote  $A_n = F'(x_n) + G(x_n, x_{n-1})$ . Let  $n = 0$ . By assumption,  $x_0, x_{-1} \in \Omega$ , we obtain the following estimation

$$\begin{aligned} & \|I - (A_*^T A_*)^{-1} A_0^T A_0\| = \|(A_*^T A_*)^{-1} (A_*^T A_* - A_0^T A_0)\| = \\ & = \|(A_*^T A_*)^{-1} (A_*^T (A_* - A_0) + (A_*^T - A_0^T) (A_0 - A_*) + (A_*^T - A_0^T) A_*)\| \leq \\ & \leq \|(A_*^T A_*)^{-1}\| (\|A_*^T\| \|A_* - A_0\| + \|A_*^T - A_0^T\| \|A_0 - A_*\| + \|A_*^T - A_0^T\| \|A_*\|) \leq \\ & \leq B(\alpha \|A_* - A_0\| + \|A_*^T - A_0^T\| \|A_0 - A_*\| + \alpha \|A_*^T - A_0^T\|). \end{aligned} \quad (7)$$

Using conditions (4) and (5), we get

$$\begin{aligned} & \|A_0 - A_*\| = \|(F'(x_0) + G(x_0, x_{-1})) - (F'(x^*) + G(x^*, x^*))\| = \\ & = \|F'(x_0) - F'(x^*) + G(x_0, x_{-1}) - G(x^*, x^*)\| \leq \\ & \leq \|F'(x_0) - F'(x^*)\| + \|G(x_0, x_{-1}) - G(x^*, x^*)\| \leq \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du, \end{aligned} \quad (8)$$

where  $\rho_k = \rho(x_k)$ . Then, from the inequality (7) and the equation  $q(r)=0$ , we obtain

$$\begin{aligned} & \|I - (A_*^T A_*)^{-1} A_0^T A_0\| \leq B \left[ 2\alpha + \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \times \\ & \quad \times \left[ \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \leq \\ & \leq B \left[ 2\alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] = \\ & = 1 - B \left\{ \left[ \alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \int_0^{r_*} L(u) u du + \int_0^{r_*} M(u) du \right] + \right. \\ & \quad \left. + \frac{1}{r_*} \left[ \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \eta \right\} < 1. \end{aligned} \quad (9)$$

Next, from (7), (8), (9), and the Banach lemma ([2]) follows that  $(A_0^T A_0)^{-1}$  exists and

$$\begin{aligned} \|(A_0^T A_0)^{-1}\| &\leq g_0 = B \left\{ 1 - B \left[ 2\alpha + \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \right. \\ &\quad \left. \times \left[ \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \right\}^{-1} \leq \\ &\leq g(r_*) = B \left\{ 1 - B \left[ 2\alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \right\}^{-1}. \end{aligned}$$

Hence,  $x_1$  is correctly defined. Next, we will show that  $x_1 \in \Omega$ .

Using the fact  $A_*^T(F(x^*) + G(x^*)) = (F'(x^*) + G(x^*, x^*))^T(F(x^*) + G(x^*)) = 0$ ,  $x_0, x_{-1} \in \Omega$  and the choice of  $r_*$ , we get the estimate as follows

$$\begin{aligned} \|x_1 - x^*\| &= \|x_0 - x^* - (A_0^T A_0)^{-1} [A_0^T(F(x_0) + G(x_0)) - A_*^T(F(x^*) + G(x^*))]\| \leq \\ &\leq \| - (A_0^T A_0)^{-1} \| \| - A_0^T [A_0 - \int_0^1 F'(x_* + t(x_0 - x^*)) dt - \\ &\quad - G(x_0, x^*)] (x_0 - x^*) + (A_0^T - A_*^T)(F(x^*) + G(x^*)) \| . \end{aligned}$$

From here, considering the inequalities

$$\begin{aligned} &\left\| A_0 - \int_0^1 F'(x_* + t(x_0 - x^*)) dt - G(x_0, x^*) \right\| = \\ &= \left\| F'(x_0) - \int_0^1 F'(x_* + t(x_0 - x^*)) dt + G(x_0, x_{-1}) - G(x_0, x^*) \right\| = \\ &= \left\| \int_0^1 [F'(x_0) - F'(x_* + t(x_0 - x^*))] dt + G(x_0, x_{-1}) - G(x_0, x^*) \right\| = \\ &= \left\| \int_0^1 [F'(x_0) - F'(x_0^t)] dt + G(x_0, x_{-1}) - G(x_0, x^*) \right\| \leq \\ &\leq \int_0^1 \int_{t\rho_0}^{\rho_0} L(u) du dt + \int_0^{\rho_{-1}} M(u) du = \int_0^{\rho_0} L(u) u du + \int_0^{\rho_{-1}} M(u) du \leq \\ &\leq \frac{1}{r_*^2} \int_0^{r_*} L(u) u du \rho_0^2 + \frac{1}{r_*} \int_0^{r_*} M(u) du \rho_{-1}, \\ \|A_0\| &\leq \|A_*\| + \|A_0 - A_*\| \leq \alpha + \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du, \end{aligned}$$

we get

$$\begin{aligned} \|x_1 - x^*\| &\leq g_0 \left\{ \left[ \alpha + \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \times \right. \\ &\quad \left. \times \left[ \int_0^{\rho_0} L(u) u du + \int_0^{\rho_{-1}} M(u) du \right] \|x_0 - x^*\| + \eta \left[ \int_0^{\rho_0} L(u) du + \int_0^{\rho_0 + \rho_{-1}} M(u) du \right] \right\} \leq \\ &\leq g_0 \left\{ \left[ \alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \frac{1}{r_*^2} \int_0^{r_*} L(u) u du \rho_0^2 + \right. \right. \\ &\quad \left. \left. + \frac{1}{r_*} \int_0^{r_*} M(u) du \rho_{-1} \right] \|x_0 - x^*\| + \eta \left[ \frac{1}{r_*} \int_0^{r_*} L(u) du \rho_0 + \frac{1}{2r_*} \int_0^{2r_*} M(u) du (\rho_0 + \rho_{-1}) \right] \right\} < \end{aligned}$$

$$\begin{aligned} &< g(r_*) \left\{ \left[ \alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \int_0^{r_*} L(u) u du + \int_0^{r_*} M(u) du \right] + \right. \\ &\quad \left. + \frac{1}{r_*} \left[ \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \eta \right\} r_* = p(r_*) r_* = r_*, \end{aligned}$$

where

$$\begin{aligned} p(r) = g(r) &\left\{ \left[ \alpha + \int_0^r L(u) du + \int_0^{2r} M(u) du_* \right] \left[ \int_0^r L(u) u du + \right. \right. \\ &\quad \left. \left. + \int_0^r M(u) du \right] + \frac{1}{r} \left[ \int_0^r L(u) du + \int_0^{2r} M(u) du \right] \eta \right\}. \end{aligned}$$

Therefore,  $x_1 \in \Omega$  and the estimate (6) holds for  $n = 0$ .

Let us assume that  $x_n \in \Omega$  for  $n = 0, 1, \dots, k$  and the estimate (6) holds for  $n = 0, 1, \dots, k - 1$ , where  $k \geq 1$  is an integer. We shall show:  $x_{n+1} \in \Omega$ , and that the estimate (6) holds for  $n = k$ .

We define

$$\begin{aligned} &\|I - (A_*^T A_*^T)^{-1} A_k^T A_k\| = \|(A_*^T A_*)^{-1} (A_*^T A_* - A_k^T A_k)\| = \\ &= \|(A_*^T A_*)^{-1} (A_*^T (A_* - A_k) + (A_*^T - A_k^T) (A_k - A_*) + (A_*^T - A_k^T) A_*)\| \leq \\ &\leq \|(A_*^T A_*)^{-1}\| (\|A_*^T\| \|A_* - A_k\| + \|A_*^T - A_k^T\| \|A_k - A_*\| + \|A_*^T - A_k^T\| \|A_*\|) \leq \\ &\leq B(\alpha \|A_* - A_k\| + \|A_*^T - A_k^T\| \|A_k - A_*\| + \alpha \|A_*^T - A_k^T\|) \leq \\ &\leq B \left[ 2\alpha + \int_0^{\rho_k} L(u) du + \int_0^{\rho_k + \rho_{k-1}} M(u) du \right] \left[ \int_0^{\rho_k} L(u) du + \int_0^{\rho_k + \rho_{k-1}} M(u) du \right] \leq \\ &\leq B \left[ 2\alpha + \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] \left[ \int_0^{r_*} L(u) du + \int_0^{2r_*} M(u) du \right] < 1. \end{aligned}$$

Consequently,  $(A_k^T A_k)^{-1}$  exists and

$$\begin{aligned} \|(A_{k+1}^T A_{k+1})^{-1}\| &\leq g_k = B \left\{ 1 - B \left[ 2\alpha + \int_0^{\rho_k} L(u) du + \int_0^{\rho_k + \rho_{k-1}} M(u) du \right] \times \right. \\ &\quad \left. \times \left[ \int_0^{\rho_k} L(u) du + \int_0^{\rho_k + \rho_{k-1}} M(u) du \right] \right\}^{-1} \leq g(r_*). \end{aligned}$$

Therefore,  $x_{k+1}$  is correctly defined and the following estimate holds

$$\begin{aligned} \|x_{k+1} - x_*\| &= \|x_k - x_* - (A_k^T A_k)^{-1} [A_k^T (F(x_k) + G(x_k)) - A_*^T (F(x_*) + G(x_*))]\| \leq \\ &\leq \| - (A_k^T A_k)^{-1} \| \left\| -A_k^T \left[ A_k - \int_0^1 F'(x_* + t(x_k - x_*)) dt - \right. \right. \\ &\quad \left. \left. - G(x_k, x_*) \right] (x_k - x_*) + (A_k^T - A_*^T) (F(x_*) + G(x_*)) \right\| \leq \\ &\leq \| - (A_k^T A_k)^{-1} \| \left\| -A_k^T \left[ A_k - \int_0^1 F'(x_* + t(x_k - x_*)) dt - \right. \right. \\ &\quad \left. \left. - G(x_k, x_*) \right] (x_k - x_*) + (A_k^T - A_*^T) (F(x_*) + G(x_*)) \right\| \leq \\ &\leq g_k \left\{ \left[ \alpha + \int_0^{\rho_k} L(u) du + \int_0^{\rho_k + \rho_{k-1}} M(u) du \right] \left[ \int_0^{\rho_k} L(u) u du + \int_0^{\rho_{k-1}} M(u) du \right] \|x_k - x_*\| + \right. \end{aligned}$$

$$\begin{aligned}
 & +\eta \left[ \int_0^{\rho_k} L(u)du + \int_0^{\rho_k + \rho_{k-1}} M(u)du \right] \leq g(r_*) \left\{ \left[ \alpha + \int_0^{r_*} L(u)du + \int_0^{2r_*} M(u)du \right] \times \right. \\
 & \quad \times \left[ \frac{1}{r_*^2} \int_0^{r_*} L(u)udu \rho_k^2 + \frac{1}{r_*} \int_0^{r_*} M(u)du \rho_{k-1} \right] \|x_k - x^*\| + \\
 & \quad \left. + \eta \left[ \frac{1}{r_*} \int_0^{r_*} L(u)du \rho_k + \frac{1}{2r_*} \int_0^{2r_*} M(u)du (\rho_k + \rho_{k-1}) \right] \right\} < p(r_*)r_* = r_*.
 \end{aligned}$$

This proves that  $x_{k+1} \in \Omega$  and the estimate (6) for  $n = k$ .

Thus, by the induction the iterative process (2) is correctly defined,  $x_n \in \Omega$  for all  $n \geq 0$ , and the estimate (6) holds for all  $n \geq 0$ .

It remains to prove that  $x_n \rightarrow x^*$  for  $n \rightarrow \infty$ .

Let us define the functions  $a$  and  $b$  on  $[0, r_*]$  as

$$\begin{aligned}
 a(r) &= g(r) \left\{ \left[ \alpha + \int_0^r L(u)du + \int_0^{2r} M(u)du \right] \left[ \int_0^r L(u)udu + \int_0^r M(u)du \right] + \right. \\
 & \quad \left. + \left[ \frac{1}{r} \int_0^r L(u)du + \frac{1}{2r} \int_0^{2r} M(u)du \right] \eta \right\}; \\
 b(r) &= g(r) \frac{1}{2r} \int_0^{2r} M(u)du \eta.
 \end{aligned} \tag{10}$$

According to the choice of  $r_*$ , we get

$$a(r_*) \geq 0, \quad b(r_*) \geq 0, \quad a(r_*) + b(r_*) = 1. \tag{11}$$

Using the estimate (6), the definition of the functions  $a$ ,  $b$  and constants  $C_i$  ( $i = 1, 2, 3, 4$ ), we get

$$\begin{aligned}
 \|x_{n+1} - x^*\| &\leq C_1 \|x_{n-1} - x^*\| + (C_2 + C_3 r_* + C_4 r_*) \|x_n - x^*\| = \\
 &= a(r_*) \|x_n - x^*\| + b(r_*) \|x_{n-1} - x^*\|.
 \end{aligned} \tag{12}$$

According to the proof in [8], under the conditions (10)–(12), the sequence  $\{x_n\}$  converges to  $x^*$  for  $n \rightarrow \infty$ . This completes the proof of Theorem 1.  $\square$

**Corollary 1.** *The convergence order of the iterative process (2) for the problem (1) with zero residual equals to  $\frac{1+\sqrt{5}}{2}$ .*

If  $\eta = 0$ , we have the nonlinear least squares problem with zero residual in solution. Then, the constants  $C_1 = 0$  and  $C_2 = 0$ , and the estimate (6) takes the form

$$\|x_{n+1} - x^*\| \leq C_3 \|x_{n-1} - x^*\| \|x_n - x^*\| + C_4 \|x_n - x^*\|^2.$$

This inequality can be written as

$$\|x_{n+1} - x^*\| \leq (C_3 + C_4) \|x_{n-1} - x^*\| \|x_n - x^*\|.$$

From here, we can write an equation for determining the convergence order as follows

$$t^2 - t - 1 = 0.$$

Hence, the positive root,  $t^* = \frac{1+\sqrt{5}}{2}$ , of the latter is the order of convergence of the iterative process (2).

In case  $G(x) \equiv 0$  in (1), we obtain the following consequence.



**Corollary 2.** *The convergence order of the iterative process (2) for the problem (1) with zero residual is quadratic.*

Indeed, if  $G(x) \equiv 0$ , then  $C_3 = 0$  and the estimate (6) takes the form

$$\|x_{n+1} - x^*\| \leq C_4 \|x_n - x^*\|^2,$$

which indicates quadratic convergence rate of process (2).

If  $L$  and  $M$  are constants, we study the convergence of the iterative process (2) in the theorem below, which is derived from Theorem 1.

**Theorem 2.** *Let  $F + G: \mathbb{R}^p \rightarrow \mathbb{R}^m$  be continuous on an open convex subset  $D \subset \mathbb{R}^p$ , moreover  $F$  is a continuously differentiable,  $G$  is a continuous function on  $D$ . Suppose that the problem (1) has a solution  $x^* \in D$ , and the inverse operation  $(A_*^T A_*)^{-1} = [(F'(x^*) + G(x^*, x^*))^T \times (F'(x^*) + G(x^*, x^*))]^{-1}$  exists, such that  $\|(A_*^T A_*)^{-1}\| \leq B$ .*

*For all  $x, y, u, v \in D$ , the Fréchet derivative  $F'$  satisfies the classical Lipschitz condition*

$$\|F'(x) - F'(y)\| \leq L\|x - y\|$$

*and the function  $G$  has a first order divided difference  $G(x, y)$  that satisfies*

$$\|G(x, y) - G(u, v)\| \leq M(\|x - u\| + \|y - v\|).$$

*Furthermore,*

$$\|F(x^*) + G(x^*)\| \leq \eta, \quad \|F'(x^*) + G(x^*, x^*)\| \leq \alpha;$$

$$B(L + 2M)\eta < 1$$

*and  $\Omega = \Omega(x^*, r_*) = \{x: \|x - x^*\| < r_*\} \subseteq D$ , where*

$$r_* = \frac{4(1 - BT\eta)}{5BT\alpha + \sqrt{25B^2T^2\alpha^2 + 24BT^2(1 - BT\eta)}}, \quad T = L + 2M.$$

*Then, for  $x_0, x_{-1} \in \Omega$ , the iterative process  $\{x_n\}$ ,  $n = 0, 1, \dots$ , generated by (2), is well defined, remains in  $\Omega$  and converges to  $x^*$  such that the following error estimate holds for all  $n \geq 0$ :*

$$\|x_{n+1} - x^*\| \leq C_1 \|x_{n-1} - x^*\| + C_2 \|x_n - x^*\| + C_3 \|x_{n-1} - x^*\| \|x_n - x^*\| + C_4 \|x_n - x^*\|^2,$$

*where*

$$\begin{aligned} g(r) &= B[1 - B(2\alpha + (L + 2M)r)(L + 2M)r]^{-1}; \\ C_1 &= g(r_*)M\eta; \quad C_2 = g(r_*)(L + M)\eta; \quad C_3 = g(r_*)(\alpha + (L + 2M)r_*)M; \\ C_4 &= g(r_*)(\alpha + (L + 2M)r_*)\frac{L}{2}. \end{aligned}$$

The proof of Theorem 2 is analogous to the proof of Theorem 1.

**4. The results of numerical experiments.** On several test examples we compare the convergence rate of the combined method (2), the Gauss-Newton type's method (3), and the Secant type's method for solving the nonlinear least squares problem ([8, 13])

$$x_{n+1} = x_n - \left(A_n^T A_n\right)^{-1} A_n^T (F(x_n) + G(x_n)),$$

$$A_n = F(x_n, x_{n-1}) + G(x_n, x_{n-1}), \quad n = 0, 1, \dots \quad (13)$$

Testing of methods was performed on nonlinear problems with nondifferentiable operator and both zero and non-zero residual in the solution. The classic Gauss-Newton and Newton's methods are inapplicable for solving such problems.

Let us denote  $H(x) \equiv F(x) + G(x)$  and  $h(x) = \frac{1}{2}(F(x) + G(x))^T(F(x) + G(x))$ . Below we list several test examples.

**Example 1.**  $p = 1, \quad m = 1;$

$$H(x) = x^2 + |x| = 0, \quad x^* = 0, \quad H(x^*) = 0.$$

**Example 2.**  $p = 1, \quad m = 1;$

$$H(x) = \sin x^2 + |x^3| = 0, \quad x^* = 0, \quad H(x^*) = 0.$$

**Example 3.** ([3, 14])  $p = 2, \quad m = 2;$

$$\begin{cases} H_1(x_1, x_2) = 3x_1^2x_2 + x_2^2 - 1 + |x_1 - 1|, \\ H_2(x_1, x_2) = x_1^4 + x_1x_2^3 - 1 + |x_2|, \end{cases}$$

$$(x_1^*, x_2^*) \approx (0.89465537, 0.32782652), \quad h(x^*) = 0.$$

**Example 4.**  $p = 3, \quad m = 4;$

$$\begin{cases} H_1(x_1, x_2, x_3) = x_3^2(1 - x_2) - x_1x_2 + |x_2 - x_3^2|, \\ H_2(x_1, x_2, x_3) = x_3^2(x_1^3 - x_1) - x_2^2 + |3x_2^2 - x_3^2 + 1|, \\ H_3(x_1, x_2, x_3) = 6x_1x_2^3 + x_2^2x_3^2 - x_1x_2^2x_3 + |x_1 - x_2 + x_3|, \\ H_4(x_1, x_2, x_3) = |2x_1 + x_2 + x_3/10|, \end{cases}$$

$$(x_1^*, x_2^*, x_3^*) = (-1, 2, 3), \quad h(x^*) = 0.45 \cdot 10^{-1}.$$

Obviously, the solution  $x^* = 0$  is a point of nondifferentiability of the function  $H(x)$  from Example 1 and Example 2.

Table 1 shows the results of the convergence of the investigated methods to the solution of the equation  $H(x) = 0$  with the accuracy  $\varepsilon = 10^{-8}$  for several initial approximations. Calculations were carried out to fulfillment of the condition  $\|x_{n+1} - x_n\| \leq \varepsilon$ . The additional initial point  $x_{-1}$  we calculated by setting  $x_{-1}$  to  $x_0 - 0.0001$ . In Table 1, the symbol '–' indicates the lack of method's convergence.

As can be seen from Table 1, the combined method (2) works efficiently when the solution is at the point of nondifferentiability of the nonlinear operator.

**5. Conclusions.** Based on the theoretical studies and the numerical experiments, comparison of the obtained results, we can argue that the combined differential-difference method (2) converges faster than the Gauss-Newton type method (3) and the Secant type. Moreover, the method (2) has a high order of convergence  $(1 + \sqrt{5})/2$  in case of zero residual as well as does not require calculation of derivatives for the nondifferentiable operator. Therefore, the proposed combined method (2) is an effective alternative for solving nonlinear least squares problems with nondifferentiable operator.

Table 1: Number of iterations for solving test problems.

Example	The initial approximation $x_0$	Method		
		Gauss–Newton type’s (3)	Secant type’s (13)	Combined method (2)
1	-0.01; 0.01	–	4	3
	-1; 1	–	8	6
	-10; 10	–	12	9
2	-0.01; 0.01	20	28	20
	-1; 1	24	38	29
	-10; 10	–	46	37
3	(1,0)	18	7	7
	(3,1)	21	12	10
	(0.5, 0.5)	21	15	10
4	(-0.5,2.3,3.5)	142	11	10
	(-1.5,2.5,3.5)	131	10	8
	(-10, 20, 30)	128	23	17

## REFERENCES

1. Argyros I.K., *Convergence and Applications of Newton-type Iterations* Springer-Verlag, New York, 2008.
2. Argyros I.K., Ren H., *A derivative free iterative method for solving least squares problems*, Numerical Algorithms, **58** (2011), 555–571.
3. Catinas E., *On some iterative methods for solving nonlinear equations*, Revue d’Analyse Numérique et de Theorie de l’Approximation, **23** (1994), 47–53.
4. Dennis J.E., Schnabel R.B., *Numerical Methods for Unconstrained Optimization and Nonlinear Equations*, SIAM, Philadelphia, 1996.
5. Iakymchuk R.P., Shakhno S.M., Yarmola H.P., *Convergence analysis of a two-step modification of the Gauss-Newton method and its Applications*, Journal of Numerical and Applied Mathematics, **3** (2017), №126, 61–74.
6. Li C., Zhang W., Jin X., *Convergence and Uniqueness Properties of Gauss-Newton’s Method*, Comput. Math. Appl., **47** (2004), 1057–1067.
7. Ortega J.M., Rheinboldt W.C., *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, Stateplace, New York, 1970.
8. Ren H., Argyros I.K., *Local convergence of a secant type method for solving least squares problems*, AMC (Appl. Math. Comp.), **217** (2010), 3816–3824.
9. Shakhno S.M., *Convergence of combined Newton-Secant method and uniqueness of the solution of nonlinear equations*, Scientific Journal of Ternopil National Technical University, **1** (2013), 243–252. (in Ukrainian)
10. Shakhno S.M., *Convergence of the two-step combined method and uniqueness of the solution of nonlinear operator equations*, Journal of Computational and Applied Mathematics, **261** (2014), 378–386.
11. Shakhno S.M., *On the Secant method under the generalized Lipschitz conditions for the divided difference operator*, Proc. Appl. Math. Mech., **1** (2007), 2060083–2060084.
12. Shakhno S.M., *Secant method under the generalized Lipschitz conditions for the first-order divided differences*, Matem. Visnyk NTSh., **4** (2007), 296–303. (in Ukrainian)
13. Shakhno S.M., Gnatyshyn O.P., *On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems*, AMC (Appl. Math. Comp.), **161** (2005), 253–264.

14. Shakhno S.M., Mel'nyk I.V., Yarmola H.P., *Convergence analysis of combined method for solving nonlinear equations*, J. Math. Sci., **212** (2016), 16–26.
15. Ulm S., *On generalized divided differences*, Proceedings of the Academy of Sciences of the Estonian SSR. Physics. Mathematics, **16** (1967), 13–26. (in Russian)
16. Wang X., *Convergence of Newton's method and uniqueness of the solution of equations in Banach space*, IMA J. Numer. Anal., **20** (2000) 123–134.
17. Wang X., Li C., *Convergence of Newton's method and uniqueness of the solution of equations in Banach space II*, Acta Mathematica Sinica, English Series, **19** (2003), 405–412.
18. Zabrejko P.P., Nguen D.F., *The majorant method in the theory of Newton-Kantorovich approximations and the Pták error estimates*, Numer. Funct. Anal. Optim., **9** (1987), 671–686.

Ivan Franko National University of Lviv, Lviv, Ukraine  
stepan.shakhno@lnu.edu.ua  
halyna.yarmola@lnu.edu.ua

KTH Royal Institute of Technology, Stockholm, Sweden  
riakymch@kth.se

*Received ?*