УДК 519.61

S. M. Shakhno[1], H. P. Yarmola[2], Yu. V. Shunkin[3]

CONVERGENCE ANALYSIS OF THE GAUSS-NEWTON-POTRA METHOD FOR NONLINEAR LEAST SQUARES PROBLEMS

S. M. Shakhno, H. P. Yarmola, Yu. V. Shunkin. *Convergence analysis of the Gauss-Newton-Potra method for nonlinear least squares problems*, Mat. Stud. **50** (2018), 211–221.

In this paper we study an iterative differential-difference method for solving nonlinear least squares problems with nondifferentiable residual function. We have proved theorems which establish the conditions of convergence, radius and the convergence order under Lipschitz and ω-conditions for the first-order derivatives of the differentiable part and for the first and second orders divided differences of the nondifferentiable part of the nonlinear function. The carried numerical experiments demonstrate the efficiency of the proposed method.

1. Introduction. Mathematical modeling of complex physical processes very often requires solving a nonlinear least squares problem. There are many methods for numerical solving this problem. The choice of method strongly depends on the properties of the problem, the computational complexity, the convergence order, the radius of convergence of the method, and so on. The basic method for its solving is the Gauss-Newton method [1, 2, 10, 11]. Its convergence order for problems with zero residual is quadratic but it requires calculation of derivatives of nonlinear functions. Moreover, this method cannot be applied for problems with nondifferentiable residual functions. As an option, we can use iterative-difference methods, which have similar computation efficiency and convergence order as the Gauss-Newton method, but do not require a calculation of the matrix of derivatives. They include the Secant type method, the Kurchatov type method and the Potra type method [1, 2, 5, 12, 14, 15]. There are no universal methods to solve successfully a wide range of such problems, therefore the problem of constracting new effective algorithms is relevant.

Some nonlinear problems can contain differentiable and nondifferentiable parts. In this case, the methods with decomposition of the nonlinear function can be used. This approach was applied and well recommended itself for solving of nonlinear equations [1, 2, 7, 9, 13, 18, 19]. These methods use the derivatives of the differentiable part of the function and the divided differences of the nondifferentiable part of the function. We apply this technique to constract methods for solving of nonlinear least squares problem with the nondifferentiable residual function.

Let us consider the nonlinear least squares problem [6, 16, 17]

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

2010 *Mathematics Subject Classification:* 49M15, 65F20, 65G99, 65H10.
Keywords: least squares problem; decomposition of operator; Gauss-Newton-Potra method; omega-conditions; divided differences.
doi:10.15330/ms.50.2.211-221

where the residual function $F+G$ is defined on $\mathbb{R}^p$ with its values on $\mathbb{R}^m$ and it is nonlinear by x; F is a continuously differentiable function; G is a continuous function, differentiability of which, in general, is not required. If $m=p$, then (1) reduces to a system of nonlinear equations.

For finding the solution of problem (1), we propose the method based on the Gauss-Newton method and the Potra type's method [15]

$$\begin{aligned} x_{k+1} &= x_k - (A_k^T A_k)^{-1} A_k^T (F(x_k)+G(x_k)),\ k=0,1,2,\dots, \\ A_k &= F'(x_k) + G(x_k,x_{k-1}) + G(x_{k-2},x_k) - G(x_{k-2},x_{k-1}). \end{aligned} \tag{2}$$

Here $F'(x_k)$ is a Fréchet derivative $F(x)$ at the point x_k, $G(x_k,x_{k-1})$, $G(x_{k-2},x_k)$, $G(x_{k-2},x_{k-1})$ are divided differences of the first order of the function $G(x)$ at the appropriate points [20]; x_0, x_{-1}, x_{-2} are given initial approximations. In case when $m=p$, this method reduces to the Newton-Potra methods ([13]).

In this article, we provide a local convergence analysis of the Gauss-Newton-Potra method (2) under classical Lipschitz conditions [4], which extend the convergence domain obtained in [17] and under weak ω-conditions ([3, 8, 9]), which do not required differentiability of the nonlinear function in the solution.

2. Convergence Analysis. Let us denote by $\Omega(x^*,r)=\{x\in D\subseteq\mathbb{R}^p\colon \|x-x^*\|<r\}$ as an open ball with the radius r ($r>0$) centred at x^*, D is an open convex subset of $\mathbb{R}^p$.

Sufficient conditions of the local convergence of the iterative process (2) are given in the following theorem.

Theorem 1. *Let the function $F+G:\mathbb{R}^p\to\mathbb{R}^m$, $m\ge p$, be continuous on a subset $D\subseteq\mathbb{R}^p$, where F is a continuously differentiable function. Assume that problem (1) has a solution $x^*\in D$ and the matrix $(A_*^T A_*)^{-1}$ exists, where $A_*=F'(x^*)+G(x^*,x^*)$, and $\|(A_*^T A_*)^{-1}\|\le B$. Suppose that the Fréchet derivative $F'(x)$ satisfies the Lipschitz conditions on D*

$$\|F'(x)-F'(x^*)\|\le L_0\|x-x^*\|, \tag{3}$$

$$\|F'(x)-F'(y)\|\le L\|x-y\|; \tag{4}$$

the function G has the first and the second order divided differences $G(\cdot,\cdot)$ and $G(\cdot,\cdot,\cdot)$ and

$$\|G(x,y)-G(u,v)\|\le M(\|x-u\|+\|y-v\|), \tag{5}$$

$$\|G(u,x,y)-G(v,x,y)\|\le N\|u-v\| \tag{6}$$

for each $x,y,u,v\in D$; L_0, L, M and N are non-negative numbers; $L_0\le L$.

Furthermore,

$$\|F(x^*)+G(x^*)\|\le\eta,\ \|A_*\|\le\alpha,\quad B(L_0+2M)\eta\le 1 \tag{7}$$

and $\Omega=\Omega(x^*,r_*)\subseteq D$, where the radius $r_*>0$ is the unique root of the equation

$$\begin{aligned} q(r) &= B\left[(\alpha+(L_0+2M)r+2Nr^2)((1/2\cdot L+M)r+4Nr^2)+\right.\\ &\left.+(L_0+2M+2Nr)\eta\right]+B\left[2\alpha+(L_0+2M)r+2Nr^2\right]\left[(L_0+2M)r+2Nr^2\right]-1=0. \end{aligned} \tag{8}$$

Then, for each x_0, x_{-1}, $x_{-2} \in \Omega$ the sequence $\{x_k\}$ generated by the method (2) is well-defined, located in Ω for each $k \geq 0$, and converges to x^*. Moreover, the following estimate holds for each $k \geq 0$

$$\begin{gathered}\|x_{k+1} - x^*\| \leq C_1\|x_k - x^*\| + C_2\|x_k - x^*\|^2 + C_3\|x_{k-1} - x^*\|\|x_{k-2} - x^*\| + C_4\|x_k - x^*\|^3 + \\ + C_5\|x_k - x^*\|\|x_{k-1} - x^*\|\|x_{k-2} - x^*\| + C_6\|x_k - x^*\|^2\|x_{k-1} - x^*\|\|x_{k-2} - x^*\| + \\ + C_7\|x_k - x^*\|\|x_{k-1} - x^*\|^2\|x_{k-2} - x^*\|^2,\end{gathered} \tag{9}$$

where

$$\begin{gathered}C_1 = g(r_*)\eta(L_0 + 2M),\ C_2 = g(r_*)\Big(\frac{\alpha L}{2} + \alpha M\Big),\ C_3 = 2g(r_*)\eta N, \\ C_4 = g(r_*)\Big(\frac{1}{2}L + M\Big)(L_0 + 2M),\ C_5 = 4g(r_*)\alpha N, \\ C_6 = g(r_*)(4L_0 + L + 10M)N,\ C_7 = 8g(r_*)N^2, \\ g(r) = B\Big[1 - B\big[2\alpha + (L_0 + 2M)r + 2Nr^2\big]\big[(L_0 + 2M)r + 2Nr^2\big]\Big]^{-1}.\end{gathered}$$

Proof. According to the intermediate value theorem, under conditions (7) the polynomial q has a positive root r_* on $[0; r]$ for a sufficiently large r. Since $q'(r) \geq 0$ for $r \geq 0$, this root is unique on $[0, r]$.

Let us choose arbitrary $x_0, x_{-1}, x_{-2} \in \Omega$ and denote

$$A_k = F'(x_k) + G(x_k, x_{k-1}) + G(x_{k-2}, x_k) - G(x_{k-2}, x_{k-1}).$$

Let $k = 0$. Then we obtain the following estimation

$$\begin{gathered}\|I - (A_*^T A_*)^{-1} A_0^T A_0\| = \\ = \|(A_*^T A_*)^{-1}(A_*^T(A_* - A_0) + (A_*^T - A_0^T)(A_0 - A_*) + (A_*^T - A_0^T)A_*)\| \leq \\ \leq B(\alpha\|A_* - A_0\| + \|A_*^T - A_0^T\|\|A_0 - A_*\| + \alpha\|A_*^T - A_0^T\|).\end{gathered} \tag{10}$$

Using conditions (3), (5), (6), we get

$$\begin{gathered}\|A_0 - A_*\| = \|F'(x_0) - F'(x^*) + G(x_0, x^*) - G(x^*, x^*) + G(x_0, x_{-1}) - G(x_0, x^*) + \\ + G(x_{-2}, x_0) - G(x_{-2}, x^*) + G(x_{-2}, x^*) - G(x_{-2}, x_{-1})\| \leq \\ \leq (L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|.\end{gathered} \tag{11}$$

Since for the Euclidean norm $\|A_* - A_0\| = \|A_*^T - A_0^T\|$, then from (10), (11) and the definition of r_* we get

$$\|I - (A_*^T A_*)^{-1} A_0^T A_0\| \leq B\left[2\alpha + (L_0 + 2M)r_* + 2Nr_*^2\right]\left[(L_0 + 2M)r_* + 2Nr_*^2\right] < 1. \tag{12}$$

According to the Banach lemma on invertible operator [11] and (12), it follows that $(A_0^T A_0)^{-1}$ exists and

$$\begin{gathered}\|(A_0^T A_0)^{-1}\| \leq g_0 = B\Big(1 - B\big[2\alpha + (L_0 + 2M)\|x_0 - x^*\| + \\ + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big] \times \\ \times \big[(L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\Big)^{-1} \leq\end{gathered}$$

$$\leq g(r_*) = B\Big(1 - B\big[2\alpha + (L_0 + 2M)r_* + 2Nr_*^2\big] \times \big[(L_0 + 2M)r_* + 2Nr_*^2\big]\Big)^{-1}.$$

Hence, x_1 is well-defined. Next, we can write

$$\|x_1 - x^*\| = \|x_0 - x^* - (A_0^T A_0)^{-1}(A_0^T(F(x_0) + G(x_0)) - A_*^T(F(x^*) + G(x^*)))\| \leq$$

$$\leq \| - (A_0^T A_0)^{-1}\| \Big\| -A_0^T\Big(A_0 - \int_0^1 F'(x^* + t(x_0 - x^*))dt -$$

$$-G(x_0, x^*)\Big)(x_0 - x^*) + (A_0^T - A_*^T)(F(x^*) + G(x^*))\Big\|.$$

Thus, by conditions (4), (5), (6) and inequalities

$$\Big\|A_0 - \int_0^1 F'(x^* + t(x_0 - x^*))dt - G(x_0, x^*)\Big\| = \Big\|F'(x_0) - \int_0^1 F'(x^* + t(x_0 - x^*))dt +$$

$$+G(x_0, x_{-1}) + G(x_{-2}, x_0) - G(x_{-2}, x_{-1}) - G(x_0, x^*)\Big\| \leq$$

$$\leq \frac{1}{2}L\|x_0 - x^*\| + M\|x_0 - x^*\| + 4N\|x_{-1} - x^*\|\|x_{-2} - x^*\|,$$

$$\|A_0\| \leq \|A_*\| + \|A_0 - A_*\| \leq$$

$$\leq \alpha + (L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|$$

we obtain

$$\|x_1 - x^*\| \leq B\big[\alpha + (L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\times$$

$$\times\big[(\frac{1}{2}L + M)\|x_0 - x^*\| + 4N\|x_{-1} - x^*\|\|x_{-2} - x^*\|\big]\|x_0 - x^*\| +$$

$$+\eta\big[(L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]/$$

$$/\Big[1 - B\big[2\alpha + (L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\times$$

$$\times\big[(L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\Big] \leq$$

$$\leq g_0\Big[\big[\alpha + (L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\times$$

$$\times\big[(\frac{1}{2}L + M)\|x_0 - x^*\| + 4N\|x_{-1} - x^*\|\|x_{-2} - x^*\|\big]\|x_0 - x^*\| +$$

$$+\eta\big[(L_0 + 2M)\|x_0 - x^*\| + N(\|x_{-2} - x^*\| + \|x_0 - x^*\|)\|x_{-1} - x^*\|\big]\Big] \leq$$

$$\leq g(r_*)\Big[\big[\alpha + (L_0 + 2M)r_* + 2Nr_*^2\big]\times$$

$$\times\Big[\Big(\frac{1}{2}L + M\Big)r_* + 4Nr_*^2\Big]r_* + (L_0 + 2M + 2Nr_*)r_*\eta\Big].$$

Let us suppose that $x_k \in \Omega$ for $k \geq 0$ and estimate (9) holds. We prove that $x_{k+1} \in \Omega$ and estimate (9) holds.

Using conditions (3), (5), (6), we get

$$\|I - (A_*^T A_*^T)^{-1} A_k^T A_k\| \leq$$

$$\leq B(2\alpha + (L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|)\times$$
$$\times((L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|)) \leq$$
$$\leq B\left[2\alpha + (L_0 + 2M)r_* + 2Nr_*^2\right]\left[(L_0 + 2M)r_* + 2Nr_*^2\right].$$

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

$$\|(A_k^T A_k)^{-1}\| \leq g_k = B\Big[1 - B\big[2\alpha + (L_0 + 2M)\|x_k - x^*\| +$$
$$+N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big] \times \big[(L_0 + 2M)\|x_k - x^*\| +$$
$$+N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]\Big]^{-1} \leq g(r_*).$$

Hence,

$$\|x_{k+1} - x^*\| \leq B\big[\alpha + (L_0 + 2M)(\|x_k - x^*\|) +$$
$$+N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]\times$$
$$\times\big[(\frac{1}{2}L + M)\|x_k - x^*\| + 4M\|x_{k-1} - x^*\|\|x_{k-2} - x^*\||\big]\|x_k - x^*| +$$
$$+\eta\big[(L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]/$$
$$/\Big[1 - B\big[2\alpha + (L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]\times$$
$$\times\big[(L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]\Big] \leq$$
$$\leq g_k\big[\alpha + (L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big]\times$$
$$\times\big[(\frac{1}{2}L + M)\|x_k - x^*\| + 4N\|x_{k-1} - x^*\|\|x_{k-2} - x^*\||\big]\|x_k - x^*| +$$
$$+\eta\big[(L_0 + 2M)\|x_k - x^*\| + N(\|x_{k-2} - x^*\| + \|x_k - x^*\|)\|x_{k-1} - x^*\|\big] \leq$$
$$\leq g_k\big[\alpha + (L_0 + 2M)\|x_k - x^*\| + 2N\|x_{k-2} - x^*\|\|x_{k-1} - x^*\|\big]\times$$
$$\times\Big[\Big(\frac{1}{2}L + M\Big)\|x_k - x^*\| + 4N\|x_{k-1} - x^*\|\|x_{k-2} - x^*\||\Big]\|x_k - x^*| +$$
$$+\eta\big[(L_0 + 2M)\|x_k - x^*\| + 2N\|x_{k-2} - x^*\|\|x_{k-1} - x^*\|\big] \leq$$
$$\leq g(r_*)\Big[\big[\alpha + (L_0 + 2M)r_* + 2Nr_*^2\big]\Big[\Big(\frac{1}{2}L + M\Big)r_* + 4Nr_*^2\Big]r_* + (L_0 + 2M + 2Nr_*)r_*\eta\Big]$$

and $x_{k+1} \in \Omega(x^*, r_*)$.

Thus, iterative process (2) is well-defined, $x_{k+1} \in \Omega(x^*, r_*)$ for $k \geq 0$ and estimate (9) holds for each $k \geq 0$.

Let's prove, that $x_k \to x^*$ for $k \to \infty$. Let's define functions a, b on $[0, r_*]$ as

$$a(r) = C_1 + C_2 r + C_4 r^2 + C_7 r^4, \quad b(r) = C_3 r + C_5 r^2 + C_6 r^3. \tag{13}$$

According to the choice of r_*, we get

$$a(r_*) \geq 0,\ b(r_*) \geq 0,\ a(r_*) + b(r_*) = 1. \tag{14}$$

Using estimate (9), the definition of the functions a, b and constants C_i $(i = 1, \ldots, 11)$, we get

$$\|x_{k+1} - x^*\| \leq C_1\|x_k - x^*\| + C_2\|x_k - x^*\|r_* + C_3\|x_{k-1} - x^*\|r_* +$$

$$+C_4\|x_k - x^*\|r_*^2 + C_5\|x_{k-1} - x^*\|r_*^2 + C_6\|x_{k-1} - x^*\|r_*^3 + C_7\|x_k - x^*\|r_*^4 = \\ = a(r_*)\|x_k - x^*\| + b(r_*)\|x_{k-1} - x^*\|. \quad (15)$$

In view of the proof in [1, 2], under conditions (13)–(15) the sequence $\{x_k\}$ converges to x^* for $k \to \infty$. □

In paper [17], we proved a similar theorem. But in this article we add new conditions (3) and (5). This allows to extend a convergence ball of method (2). Note that such an approach for extending the convergence domain was applied for the Gauss-Newton-Secant method in [6].

Corollary 1. *The convergence order of iterative method (2) with zero residual is equal to* 1.839....

Proof. If $\eta = 0$, we have the nonlinear least squares problem with zero residual in the solution. In this case constants $C_1 = 0,\ C_3 = 0$, and estimate (9) reduces to:

$$\|x_{k+1} - x^*\| \le (C_5 + C_6 r_* + C_7 r_*^2)\|x_k - x^*\|\|x_{k-1} - x^*\|\|x_{k-2} - x^*\| + \\ +(C_2 + C_4 r_*)\|x_k - x^*\|^2.$$

From the last estimate the assertion of the corollary follows. □

Let $G(x) \equiv 0$ in (1). Then $M = N = 0$ and constants $C_i = 0,\ i = 1, 3, 5, 6, 7$. The estimate (9) reduces to:

$$\|x_{k+1} - x^*\| \le (C_2 + C_4 r_*)\|x_k - x^*\|^2.$$

Thus, the convergence order of method (2) is quadratic.

Let $F(x) \equiv 0$ in (1). Then $L = L_0 = 0$ and constants $C_i = 0,\ i = 1, 3$. Estimate (9) takes the form

$$\|x_{k+1} - x^*\| \le (C_5 + C_6 r_* + C_7 r_*^2)\|x_k - x^*\|\|x_{k-1} - x^*\|\|x_{k-2} - x^*\| + \\ +(C_2 + C_4 r_*)\|x_k - x^*\|^2.$$

Thus, the convergence order of method (2) is 1.839....

Let us consider the Gauss-Newton-Potra local convergence theorem under ω-conditions. These conditions are weaker than Lipschitz conditions, moreover, the existence of the operator $\left[F'(x^*) + G(x^*, x^*)\right]^{-1}$ is not supposed.

Theorem 2. *Let the function $F + G\colon D \subseteq \mathbb{R}^p \to \mathbb{R}^m$ is continuous on a subset D, and F is continuously differentiable on this subset. Assume that problem (1) has a solution $x^* \in D$, $F(x^*) + G(x^*) = 0$, and a matrix $(A_*^T A_*)^{-1}$, where $A_* = F'(x^*) + G(\tilde{x}, x^*)$, exists for $\tilde{x}$ such, that $\|\tilde{x} - x^*\| = \delta > 0$ and $\|(A_*^T A_*)^{-1}\| \le B$. Suppose that Fréchet derivative F' satisfies the condition on D*

$$\|F'(x) - F'(y)\| \le \omega_0(\|x - y\|), \quad (16)$$

where $\omega_0\colon \mathbb{R}_+ \to \mathbb{R}_+$ is continuous non-negative function such, that continuous non-negative function $h\colon [0, 1] \to \mathbb{R}_+$ exists, and $\omega_0(tz) \le h(t)\omega_0(z)$ for $t \in [0, 1]$ and $z \in [0, \infty)$,

$T = \int_0^1 h(t)dt$, and the function G has divided differences of first order satisfying condition

$$\|G(x,y) - G(u,v)\| \le \omega_1(\|x-u\|, \|y-v\|), \tag{17}$$

where $\omega_1\colon \mathbb{R}_+ \times \mathbb{R}_+ \to \mathbb{R}_+$ is continuous non-negative function of two arguments.

Moreover, $\|A_*\| \le \alpha$, and there exists $r_* \in \mathbb{R}_+$ such that $\Omega = \Omega(x^*, r_*) \subseteq D$ and

$$p(r_*) + \tilde{p}(r_*) < 1, \tag{18}$$

where

$$p(r) = Bg(r,\delta)[2\alpha + g(r,\delta)],\ \tilde{p}(r) = Bq(r)[\alpha + g(r,\delta)],$$
$$g(r,\delta) = \omega_0(r) + \omega_1(r+\delta, r) + \omega_1(0,2r), \quad q(r) = T\omega_0(r) + \omega_1(0,r) + \omega_1(0,2r).$$

Then, for each x_0, x_{-1}, $x_{-2} \in \Omega$, the iterative process (2) is well-defined and generates the sequence $\{x_k\}$, $k = 0,1,\ldots$, located in Ω that converges to the solution x^*. Moreover, the following estimate holds for $k \ge 0$

$$\|x_{k+1} - x^*\| \le \frac{\tilde{p}(r_*)}{1 - p(r_*)}\|x_k - x^*\|. \tag{19}$$

Proof. Obviously, (18) implies $p(r_*) < 1$, $\tilde{p}(r_*) < 1$ and $\frac{\tilde{p}(r_*)}{1-p(r_*)} < 1$. The proof of this theorem is analogous to that of Theorem 1 by the method of mathematical induction.

Let us denote $A_k = F'(x_k) + G(x_k, x_{k-1}) + G(x_{k-2}, x_k) - G(x_{k-2}, x_{k-1})$. Assume that $x_{n+1} \in \Omega$ and estimate (19) holds for $n = 0, \ldots, k-1$. Let us prove that $x_{k+1} \in \Omega$ and estimate (19) holds.

Taking into account the previous calculations and conditions (16), (17), we can deduce

$$\begin{gathered}\|I - (A_*^T A_*^T)^{-1} A_k^T A_k\| \le B[2\alpha + \omega_0(\|x_k - x^*\|) + \omega_1(\|x_k - x^*\| + \|x^* - \tilde{x}\|, \|x_{k-1} - x^*\|) + \\ + \omega_1(\|x_{k-2} - x_{k-2}\|, \|x_k - x_{k-1}\|)][\omega_0(\|x_k - x^*\|) + \\ + \omega_1(\|x_k - x^*\| + \|x^* - \tilde{x}\|, \|x_{k-1} - x^*\|) + \omega_1(\|x_{k-2} - x_{k-2}\|, \|x_k - x_{k-1}\|)] \le \\ \le B[2\alpha + \omega_0(r_*) + \omega_1(r_* + \delta, r_*) + \omega_1(0, 2r_*)][\omega_0(r_*) + \omega_1(r_* + \delta, r_*) + \omega_1(0, 2r_*)] < 1.\end{gathered}$$

Hence, $\left(A_k^T A_k\right)^{-1}$ exists and

$$\begin{gathered}\|(A_k^T A_k)^{-1}\| \le p_k = \\ = B\{1 - B[2\alpha + \omega_0(\|x_k - x^*\|) + \omega_1(\|x_k - x^*\| + \|x^* - \tilde{x}\|, \|x_{k-1} - x^*\|) + \\ + \omega_1(0, \|x_k - x^*\| + \|x^* - x_{k-1}\|)] \times \\ \times [T\omega_0(\|x_k - x^*\|) + \omega_1(\|x_k - x^*\| + \|x^* - \tilde{x}\|, \|x_{k-1} - x^*\|) + \\ + \omega_1(0, \|x_k - x^*\| + \|x^* - x_{k-1}\|)]\}^{-1} \le p(r_*).\end{gathered}$$

So, the iteration x_{k+1} is well-defined and the following estimate holds

$$\begin{gathered}\|x_{k+1} - x^*\| = \|x_k - x^* - (A_k^T A_k)^{-1}[A_k^T(F(x_k) + G(x_k)) - A_*^T(F(x^*) + G(x^*))]\| \le \\ \le \| - (A_k^T A_k)^{-1}\| \, \|([-A_k^T(A_k - \int_0^1 F'(x^* + t(x_k - x^*))dt -\end{gathered}$$

$$-G(x_k,x^*))(x_k-x^*)+(A_k^T-A_*^T)(F(x^*)+G(x^*))]\| \le$$
$$\le p_k[\alpha+\omega_0(\|x_k-x^*\|)+\omega_1(\|x_k-\tilde{x}\|,\|x_{k-1}-x^*\|)+\omega_1(\|x_{k-2}-x_{k-2}\|,\|x_k-x_{k-1}\|)]\times$$
$$\times[T\omega_0(\|x_k-x^*\|)+\omega_1(\|x_k-x_k\|,\|x_{k-1}-x^*\|)+$$
$$+\omega_1(\|x_{k-2}-x_{k-2}\|,\|x_k-x_{k-1}\|)]\|x_k-x^*\| \le$$
$$\le p(r_*)[\alpha+\omega_0(\|x_k-x^*\|)+\omega_1(\|x_k-\tilde{x}\|,\|x_{k-1}-x^*\|)+\omega_1(0,\|x_k-x_{k-1}\|)]\times$$
$$\times[T\omega_0(\|x_k-x^*\|)+\omega_1(\|x_k-x_k\|,\|x_{k-1}-x^*\|)+\omega_1(0,\|x_k-x_{k-1}\|)]\|x_k-x^*\| \le$$
$$\le \frac{\tilde{p}(r_*)}{1-p(r_*)}\|x_k-x^*\| < \|x_k-x^*\| < r_*.$$

Hence, iterative process (2) is well-defined, $x_k \in \Omega$ and (19) holds for each $k \ge 0$. □

3. Numerical experiments. Let us compare the convergence rate of the *Gauss-Newton-Potra method* (2) with other methods for nonlinear least squares problems on several test cases. In particular, we compare it with the *Potra type's method* [15]:

$$\begin{aligned} &x_{k+1}=x_k-(A_k^TA_k)^{-1}A_k^T(F(x_k)+G(x_k)),\ k=0,1,2,...,\\ &A_k=F(x_k,x_{k-1})+F(x_{k-2},x_k)-F(x_{k-2},x_{k-1})+\\ &+G(x_k,x_{k-1})+G(x_{k-2},x_k)-G(x_{k-2},x_{k-1}), \end{aligned} \tag{20}$$

the *Gauss-Newton-Secant method* [16]:

$$\begin{aligned} &x_{k+1}=x_k-(A_k^TA_k)^{-1}A_k^T(F(x_k)+G(x_k)),\ k=0,1,2,...,\\ &A_k=F'(x_k)+G(x_k,x_{k-1}) \end{aligned} \tag{21}$$

and the *Secant type's method* [15]:

$$\begin{aligned} &x_{k+1}=x_k-(A_k^TA_k)^{-1}A_k^T(F(x_k)+G(x_k)),\ k=0,1,2,...,\\ &A_k=F(x_k,x_{k-1})+G(x_k,x_{k-1}). \end{aligned} \tag{22}$$

Let us denote $H(x)=F(x)+G(x)$ and $h(x)=\frac{1}{2}H(x)^TH(x)$.

Testing is carried out on nonlinear problems with a nondifferentiable function with zero and non-zero residual. The classic Gauss-Newton and Newton methods cannot be used for solving of such problems. The solution was obtained with the accuracy $\varepsilon=10^{-8}$. Calculations are performed until the following condition is satisfied

$$\|x_{k+1}-x_k\| \le \varepsilon.$$

Table 1 gives a number of iterations, which are needed for numerical solving of nonlinear least squares problems. Obtained results show that the method (2) usually has an advantage over other methods.

Example 1. ([16])

$$\begin{cases} H_1(x,y)=3x^2y+y^2-1+|x-1|,\\ H_2(x,y)=x^4+xy^3+|y|, \end{cases}$$

$$z^*=(x^*,y^*)\approx(0.8946554,0.3278265),\ h(z^*)=0.$$

Example 2. ([16])

$$\begin{cases} H_1(x,y) = 3x^2y + y^2 - 1 + |x-1|, \\ H_2(x,y) = x^4 + xy^3 + |y|, \\ H_3(x,y) = |x^2 - y|, \end{cases}$$

$$z^* = (x^*, y^*) \approx (0.7486280, 0.4303915), \ h(z^*) \approx 4.0469349 \cdot 10^{-2}.$$

Initial approximations (x_{-1}, y_{-1}) and (x_{-2}, y_{-2}) were chosen as follows

$$(x_{-1}, y_{-1}) = (x_0 - 10^{-4}, y_0 - 10^{-4}),$$
$$(x_{-2}, y_{-2}) = (x_0 - 2 \cdot 10^{-4}, y_0 - 2 \cdot 10^{-4}).$$

Example	(x_0, y_0)	Method			
		(2)	(20)	(21)	(22)
1	$(0.3, 0.9)$	9	9	10	14
	$(0.5, 0.5)$	10	14	11	15
	$(0, 2.7)$	10	14	11	17
2	$(1, 0.1)$	11	11	14	21
	$(1.5, 0)$	12	14	12	22
	$(0.55, 2.7)$	23	25	15	21

Table 1: Number of iterations for solving test problems.

Let us show that new conditions allow to increase the radius of the convergence domain. We use the Euclidean norm.

Example 3.

$$\begin{cases} H_1(x) = x^3 - 0.5x^2 - 1.75 + |x^2 - 2|, \\ H_2(x) = 2x^2 - 2 + |x+1|, \end{cases} \qquad x^* = 0.5, \ h(x^*) = 0.$$

Let us denote $D = [0.1, 0.9]$. Then we can write

$$F'(x) = \begin{pmatrix} 3x^2 - x \\ 4x \end{pmatrix}, \quad G(x,y) = \begin{pmatrix} -(x+y) \\ 1 \end{pmatrix}$$

and

$$F'(x) - F'(y) = \begin{pmatrix} (3(x+y) - 1)(x-y) \\ 4(x-y) \end{pmatrix}, \quad G(x,y) - G(u,v) = \begin{pmatrix} -(x-u+y-v) \\ 0 \end{pmatrix}.$$

So, $L_0 = 5.1224994$, $L = 5.9464275$, $M = 1$, $\alpha = 3.0923292$, $B = 0.1045752$, $\eta = 0$. A constant $N = 0$ because all elements of $G(u,x,y)$ are real numbers and do not depend on u, x, y.

We calculate radii r_{new} and r_{old} as positive solutions of equations

$$B(L_0 + 2M)(L_0 + L/2 + 3M)r^2 + B\alpha(2L_0 + L/2 + 5M)r - 1 = 0$$

and

$$B(L+2M)\Big(\frac{3}{2}L+3M\Big)r^2 + B\alpha\Big(\frac{5}{2}L+5M\Big)r - 1 = 0,$$

respectively, and get $r_{new} = 0.1416082$, $r_{old} = 0.1297157$. So, a convergence ball is extended.

4. Conclusions. In order to check the efficiency of the proposed method, a numerical experiment was conducted on problems with zero and non-zero residual. Taking into account the obtained results, it can be argued that combined methods, in particular, the Gauss-Newton-Potra method (2), as the Gauss-Newton-Secant method (21), usually converge faster than iterative-difference methods (20) and (22). Therefore, they are more effective for solving nonlinear least squares problems with a nondifferentiable residual function. Moreover, we verified the theoretical results and confirmed the expediency of new conditions of Theorem 1.

Acknowledgments. Authors are grateful to the unknown reviewer for useful remarks which helped them to improve the paper.

REFERENCES

1. I.K. Argyros, Á.A. Magrenán, A Contemporary Study of Iterative Methods, Elsevier (Academic Press), New York, NY, USA, 2018.
2. I.K. Argyros, Á.A. Magrenán, Iterative Methods and Their Dynamics with Applications: A Contemporary Study, CRC Press, 2011).
3. I.K. Argyros, S. Hilout, *Newton–Kantorovich approximations under weak continuity conditions*, J. Appl. Math. Comput., **37** (2011), 361–375.
4. I.K. Argyros, S. Hilout, *On an improved convergence analysis of Newton's method*, Appl. Math. Comput., **225** (2013), 372–386.
5. I.K. Argyros, H. Ren, *A derivative free iterative method for solving least squares problems*, Numer. Algorithms, **4(58)** (2011), 555–571.
6. I.K. Argyros, S. Shakhno, Yu. Shunkin, *Improved convergence analysis of Gauss-Newton-Secant method for solving nonlinear least squares problems*, Mathematics, **7** (2019), № 1, Article id: 99; doi: 10.3390/math7010099.
7. E. Catinas, *On some iterative methods for solving nonlinear equations*, Rev. Anal. Numer. Theor. Approx., **1(23)** (1994), 47–53.
8. M.A. Hernandez, M.J.Rubio, *A uniparametric family of iterative processes for solving nondifferentiable operators*, J. Math. Anal. Appl., **275** (2002), 821-834.
9. M.A. Hernandez-Veron, M.J. Rubio, *On the local convergence of a Newton–Kurchatov-type method for non-differentiable operators*, Appl. Math. Comput., **304** (2017), 1–9.
10. J.E. Dennis, R.B. Schnabel, Numerical Methods for Unconstrained Optimization and Nonlinear Equations, SIAM, Philadelphia, 1996.
11. J.M. Ortega, W.C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.
12. M.A. Ren, I.K. Argyros, *Local convergence of a secant type method for solving least squares problems*, AMC (Appl. Math. Comp.), **217** (2010), 3816–3824.
13. S.M. Shakhno, A.-V.I. Babjak, H.P. Yarmola, *Combined Newton-Potra method for solving nonlinear operator equations*, J. Appl. Comput. Math. (Kyiv), **3(120)** (2015), 170–178. (in Ukrainian)
14. S.M. Shakhno, O.P. Gnatyshyn, *Iterative-difference methods for solving nonlinear least-squares problem*, Progress in Industrial Mathematics at ECMI 98, Verlag B.G.Teubner GMBH, Stuttgart, (1999), 287–294
15. S.M. Shakhno, O.P. Gnatyshyn, *On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems*, AMC (Appl. Math. Comp.), **161** (2005), 253–264.

16. S. Shakhno, Yu. Shunkin, *One combined method for solving nonlinear least squares problems*, Visn. Lviv Univ. Ser. Appl. Math. Informatics, **25** (2017), 38–48. (in Ukrainian).
17. S. Shakhno, Yu. Shunkin, H.P. Yarmola, *Gauss-Newton-Potra method for nonlinear least squares problems with decomposition of operator*, XXXII Int. Conf. PDMU–2018 (Proceedings), (2018), 153–159.
18. S.M. Shakhno, H.P. Yarmola, *On the two-step method for solving nonlinear equations with nondifferentiable operator*, Proc. Appl. Math. Mech., **1** (2012), 617–618.
19. S.M. Shakhno, H.P. Yarmola, *Two-point method for solving nonlinear equations with nondifferentiable operator*, Mat. Stud., **36** (2011), 213–220. (in Ukrainian)
20. S. Ulm, *On generalized divided differences*, Proc. Acad. Sci. Estonian SSR. Physics. Mathematic, **16** (1967), 13–26. (in Russian)

Faculty of Applied Mathematics and Informatics
Ivan Franko National University of Lviv
1 Universyteska Str., Lviv, Ukraine
[1]stepan.shakhno@lnu.edu.ua
[2]halyna.yarmola@lnu.edu.ua
[3]yuriy.shunkin@lnu.edu.ua

Received 08.11.2018