Журнал обчисл. прикл. матем. Вип. 1, 2009, стор. 112-121

J. Numer. Appl. Math. No. 97, 2009, pp. 112-121

# ON A SECANT TYPE METHOD FOR NONLINEAR LEAST SQUARES PROBLEMS

UDC 519.6

S. M. SHAKHNO, O. P. GNATYSHYN AND R. P. IAKYMCHUK

ABSTRACT. In this paper the new Secant type method for nonlinear least squares problems is presented. We study the convergence of the proposed method under hypothesis that the divided differences of first order satisfy the generalized Lipschitz conditions. We obtain the radius of convergence ball, the uniqueness ball of the solution and rate of convergence of the method. Similar results under the generalized Hölder conditions are also presented. The results of numerical experiments, that represent the convergence analysis of Secant type method are given.

## 1. Introduction

Let us consider the nonlinear least squares problem:
find

$$\min f(x) := \frac{1}{2}F(x)^T F(x), \tag{1.1}$$

where $m \geq n$, residual function $F : \mathbb{R}^n \to \mathbb{R}^m$ is nonlinear by $x$.

For solving problem (1.1) we introduce the iterative Secant type method

$$x_{n+1} = x_n - (A_n^T A_n)^{-1} A_n^T F(x_n), \quad n = 0, 1, ..., \tag{1.2}$$

where $A_n = F(x_n, x_n + \alpha_n(x_{n-1} - x_n))$, $F(x, y)$ is the divided difference of first order for function $F(x)$ on the points $x$ and $y$, $\alpha_n \in [0, 1]$ and $\{\alpha_n\}$ is nonincreasing sequence, $x_{-1}$ and $x_0$ are given.

In particular, when $\alpha_n = 1$ from (1.2) we receive the Secant method [8], and when $\alpha_n = 0$ classic Gauss-Newton method [5].

In papers [3, 4, 10] the Newton, Gauss-Newton and Secant type methods are studied under the hypothesis that the derivative operator satisfies the generalized Lipschitz conditions, that is some integrable function is used instead of Lipschitz constant. The generalized Lipschitz conditions for divided differences are presented in [7] where the Secant method for solving nonlinear operator equations in Banach space is studied.

In this work we study the convergence of the Secant type method (1.2) for nonlinear least squares problems under generalized Lipschitz conditions. Let us note that Lipschitz condition with constant is the particular case of the generalized Lipschitz condition. We have shown, that under proper choice of parameter $\alpha_n$ method (1.2) converges faster than the Secant method, namely with rate $1 + p$. We have obtained the radius of convergence of the method. Similar results under the generalized Hölder condition are also presented. We have also carried out numerical experiments, that represent the convergence analysis of Secant type method (1.2).

† *Key words.* Nonlinear least squares problem, secant type method, generalized Lipschitz conditions, divided difference, convergence ball, uniqueness ball, rate of convergence.

## 2. Definitions and auxiliary lemmas

We state the definition of the divided difference of first order for operator $F$ [1].

**Definition 2.1** *Let $x, y$ are two fixed points in Euclidean space $\mathbb{R}^n$. The bounded linear operator $F(x, y)$, that maps from $\mathbb{R}^n$ into $\mathbb{R}^m$, is called the divided difference of first order for operator $F$ on the points $x$ and $y$, if it satisfies the following equality*

$$F(x, y)(x - y) = F(x) - F(y). \tag{2.1}$$

In the case, when $x = y$ we take that $F(x, x) = F'(x)$, where $F'(x)$ is the Jacobian matrix of nonlinear operator $F$ on point $x$.

Lipschitz condition in region $D \subset \mathbb{R}^n$ for the divided differences of first order has the form

$$\|F(x, y) - F(u, v)\| \le L(\|x - u\| + \|y - v\|), \tag{2.2}$$

where $x, y, u, v \in D$, $L$ is Lipschitz constant.

But $L$ in Lipschitz conditions does not have to be a constant, it also can be some positive integrable function. In this case condition (2.2) can be written in the form [7]

$$\|F(x, y) - F(u, v)\| \le \int_0^{\|x-u\|+\|y-v\|} L(z)dz, \tag{2.3}$$

where $x, y, u, v \in D$. Condition (2.3) is called the generalized Lipschitz condition that have the $L$ average.

Let $\mathbb{R}^{m\times n}$ is the set of all $m \times n$ matrixes $A$. Denote by $A^\dagger$ the pseudoinverse matrix by Moore-Penrose to $A$ and, if $A$ has full column rank, then $A^\dagger = (A^T A)^{-1} A^T$.

**Lemma 2.2** *[9,11] Assume that $A, E \in \mathbb{R}^{m\times n}$, $B = A + E$, $\|A^\dagger\|\|E\| < 1$, $rank(A) = rank(B)$, then*

$$\|B^\dagger\| \le \frac{\|A^\dagger\|}{1 - \|A^\dagger\|\|E\|}, \tag{2.4}$$

*and if $rank(A) = rank(B) = min\{m, n\}$, we can obtain*

$$\|B^\dagger - A^\dagger\| \le \frac{\sqrt{2}\|A^\dagger\|^2\|E\|}{1 - \|A^\dagger\|\|E\|}. \tag{2.5}$$

**Lemma 2.3** *[3] Assume that $A, E \in \mathbb{R}^{m\times n}(m \ge n)$, $B = A+E$, $\|EA^\dagger\| < 1$, $rank(A) = n$, then $rank(B) = n$.*

**Lemma 2.4** *[3] Let*

$$h(t) = \frac{1}{t^\alpha}\int_0^t L(u)u^{\alpha-1}du, \quad \alpha \ge 1, 0 \le t \le r, \tag{2.6}$$

*where $L(u)$ is positive integrable function and monotonically nondecreasing in [0,r]. Then $h(t)$ is nondecreasing with respect to $t$.*

## 3. Convergence of method (1.2)

**Theorem 3.1** *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuous in region $D \subseteq \mathbb{R}^n$. Assume that:*

1) *problem (1.1) has solution $x_*$ in region $\Omega(x_*, r) = \{x \in D : \|x - x_*\| < r\}$ and there exists $F'(x_*)$ and it has full column rank;*

*2) in region $\Omega(x_*, r)$ function $F(x)$ has the divided difference of the first order $F(x, y)$, which has full column rank and satisfies Lipschitz condition with the $L$ average :*

$$\|F(x,y) - F(u,v)\| \leq \int_0^{\|x-u\|+\|y-v\|} L(z)dz, \tag{3.1}$$

*where $x, y, u, v \in \Omega(x_*, r)$ and $L$ is nondecreasing;*

*3) $r$ satisfies inequality*

$$\frac{\beta \int_0^r L(z)dz}{1 - \beta \int_0^{2r} L(z)dz} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{2r} L(z)dz}{r(1 - \beta \int_0^{2r} L(z)dz)} \leq 1. \tag{3.2}$$

*Then method (1.2) converges for all $x_{-1}, x_0 \in \Omega$ such, that $\rho(x_{-1}) \leq \rho(x_0)$ and*

$$\|x_{n+1} - x_*\| \leq \frac{q_1}{\rho(x_0)} \|x_{n-1} - x_*\| \|x_n - x_*\| + q_0 \|x_{n-1} - x_*\|; \tag{3.3}$$

*where*

$$\rho(x) = \|x - x_*\|; \quad \alpha = \|F(x_*)\|; \quad \beta = \|[F'(x_*)^T F'(x_*)]^{-1} F'(x_*)^T\| \tag{3.4}$$

*and values*

$$\begin{aligned} q_0 &= \frac{\sqrt{2}\alpha\beta^2 \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz}{\rho(x_{-1})(1 - \beta \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz)}; \\ q_1 &= \frac{\beta \int_0^{(1-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz}{1 - \beta \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz} \end{aligned} \tag{3.5}$$

*are less than 1.*

*Proof.* Let choose arbitrary $x_{-1}, x_0 \in \Omega$, where $r$ satisfies (3.2), then $q_i (i = 0, 1)$, determined by (3.5), are less than 1. In fact, by the monotonicity of $L$ with lemma 2.4 function $\frac{1}{t} \int_0^t L(z)dz$ is nondecreasing by $t$. Then we have

$$\begin{aligned} q_0 = & \frac{\sqrt{2}\alpha\beta^2 \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz((2-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1}))}{\rho(x_{-1})(1 - \beta \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz)((2-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1}))} \\ & \leq \frac{\sqrt{2}\alpha\beta^2 \int_0^{2r} L(z)dz}{2r^2(1 - \beta \int_0^{2r} L(z)dz)}((2-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1})) \\ & \leq \frac{(2-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|}{2r} < 1; \end{aligned}$$

$$\begin{aligned} q_1 = & \frac{\beta \int_0^{(1-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz((1-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1}))}{(1 - \beta \int_0^{(2-\alpha_0)\rho(x_0)+\alpha_0\rho(x_{-1})} L(z)dz)((1-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1}))} \\ & \leq \frac{\beta \int_0^r L(z)dz}{r(1 - \beta \int_0^{2r} L(z)dz)}((1-\alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1})) \\ & \leq \frac{(1-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|}{r} < 1. \end{aligned}$$

Denote: $A_n = F(x_n, x_n + \alpha_n(x_{n-1} - x_n)), A_* = F'(x_*) = F(x_*, x_*)$. Then, under the assumptions of theorem, we obtain estimate

$$\begin{aligned}\|[A_*^T A_*]^{-1} A_*^T\| \|A_n - A_*\| &\le \beta \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz \\ &\le \beta \int_0^{(2-\alpha_n)\|x_n - x_*\| + \alpha_n \|x_{n-1} - x_*\|} L(z)dz \\ &\le \beta \int_0^{2r} L(z)dz < 1, \quad \forall x_{n-1}, x_n \in \Omega(x_*, r).\end{aligned}$$

By lemmas 2.2 and 2.3, and fact, that $A_n$ has full column rank, taking $B = F(x_n, x_n + \alpha_n(x_{n-1} - x_n)), A = F(x_*, y_*), E = B - A$ we obtain

$$\|[A_n^T A_n]^{-1} A_n^T\| \le \frac{\beta}{1 - \beta \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}, \quad \forall x_{n-1}, x_n \in \Omega(x_*, r);$$

$$\|[A_n^T A_n]^{-1} A_n^T - [A_*^T A_*]^{-1} A_*^T\| \le \frac{\sqrt{2}\beta^2 \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}{1 - \beta \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}.$$

Using (1.2), we can write

$$\begin{aligned}x_{n+1} - x_* = x_n - x_* - [A_n^T A_n]^{-1} A_n^T F(x_n) = [A_n^T A_n]^{-1} A_n^T (A_n(x_n - x_*) \\ -F(x_n) + F(x_*)) - [A_n^T A_n]^{-1} A_n^T F(x_*) + [A_*^T A_*]^{-1} A_*^T F(x_*).\end{aligned}$$

Hence,

$$\begin{aligned}\|x_{n+1} - x_*\| &\le \|[A_n^T A_n]^{-1} A_n^T\| \|A_n - F(x_n, x_*)\| \|x_n - x_*\| \\ &\quad + \|[A_*^T A_*]^{-1} A_*^T - [A_n^T A_n]^{-1} A_n^T\| \|F(x_*)\| \\ &\le \frac{\beta \int_0^{\|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}{1 - \beta \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz} \|x_n - x_*\| \\ &\quad + \frac{\sqrt{2}\alpha\beta^2 \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}{1 - \beta \int_0^{\|x_n - x_*\| + \|x_n + \alpha_n(x_{n-1}-x_n) - x_*\|} L(z)dz}.\end{aligned}$$

Taking $n = 0$ above, we obtain

$$\|x_1 - x_*\| \le (q_1 + q_0)\|x_0 - x_*\| < \|x_0 - x_*\| < r.$$

Hence, $x_1 \in \Omega(x_*, r)$. It follows that (1.2) can be continued an infinite number of times. By mathematical induction, all $\{x_n\}_{n \ge 0}$ belong to $\Omega(x_*, r)$ and $\rho(x_n)$ decreases monotonically. Thus, for all $n = 0, 1, ...$, we have

$$\begin{aligned}\|x_{n+1} - x_*\| &\le \frac{\beta \int_0^{(1-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})} L(z)dz \rho(x_n)}{1 - \beta \int_0^{(2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})} L(z)dz} \\ &\quad \times \frac{(1-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})}{(1-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})} \\ &+ \frac{\sqrt{2}\alpha\beta^2 \int_0^{(2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})} L(z)dz((2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1}))}{(1 - \beta \int_0^{(2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1})} L(z)dz)((2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1}))} \\ &\le \frac{\beta \int_0^{(1-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1})} L(z)dz((1-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1}))\rho(x_n)}{((1-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1}))(1 - \beta \int_0^{(2-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1})} L(z)dz)} \\ &+ \frac{\sqrt{2}\alpha\beta^2 \int_0^{(2-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1})} L(z)dz((2-\alpha_n)\rho(x_n) + \alpha_n \rho(x_{n-1}))}{((2-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1}))(1 - \beta \int_0^{(2-\alpha_0)\rho(x_0) + \alpha_0 \rho(x_{-1})} L(z)dz)} \\ &\le \frac{q_1}{\rho(x_{-1})} \rho(x_{n-1})\rho(x_n) + q_0 \rho(x_{n-1});\end{aligned} \tag{3.6}$$

So, inequality (3.3) is true. □

**Corollary 3.2** *Order of convergence of iterative process (1.2) in the case of zero residual is equal* $\frac{1+\sqrt{5}}{2}$.

So, the rate of convergence of method (1.2) is not less than the rate of convergence of Secant method for arbitrary $\alpha_n$.

## 4. The uniqueness ball for the solution of problem (1.1)

In this section we present investigation of the uniqueness ball for the solution of problem (1.1). The proof is carried out on the example of application of method (1.2) for solving problem (1.1).

**Theorem 4.1** *Let $x_*$ satisfies (1.1), $F(x)$ is continuous in $\Omega(x_*, r)$. Moreover, $F(x)$ has the divided difference of first order, and $F'(x_*) = F(x_*, x_*) = A_*$ has full column rank, which satisfies generalized Lipschitz condition with the $L$ average:*

$$\|F(x,y) - F(x_*, x_*)\| \le \int_0^{\rho(x)+\rho(y)} L(z)dz, \tag{4.1}$$

*where $x, y \in \Omega(x_*, r), \rho(x) = \|x - x_*\|$ and $L$ is nondecreasing. Let $r$ satisfies inequality*

$$\beta \int_0^r L(z)dz + \frac{\alpha\beta_0}{r} \int_0^{2r} L(z)dz \le 1, \tag{4.2}$$

*where $\alpha$ and $\beta$ are determined in (3.4),*

$$\beta_0 = \|[A_*^T A_*]^{-1}\|. \tag{4.3}$$

*Then problem (1.1) has unique solution $x_*$ in $\Omega(x_*, r)$.*

*Proof.* Let $x_0 \in \Omega, x_0 \neq x_*$ is also a solution (1.2). Then we have

$$[A_*^T A_*]^{-1} A_0^T F(x_0) = 0. \tag{4.4}$$

Hence,

$$\begin{aligned} x_0 - x_* = x_0 - x_* - [A_*^T A_*]^{-1} A_0^T F(x_0) = [A_*^T A_*]^{-1} A_*^T (A_*(x_0 - x_*) \\ -F(x_0) + F(x_*)) - [A_*^T A_*]^{-1} A_0^T F(x_0) + [A_*^T A_*]^{-1} A_*^T F(x_0) \\ = [A_*^T A_*]^{-1} A_*^T (A_* - F(x_0, x_*))(x_0 - x_*) + [A_*^T A_*]^{-1}(A_0^T - A_*^T)F(x_0). \end{aligned}$$

Under assumption (4.1) we obtain

$$\begin{aligned} \|x_0 - x_*\| \le \|[A_*^T A_*]^{-1} A_*^T\| \|A_* - F(x_0, x_*)\| \|x_0 - x_*\| \\ + \|[A_*^T A_*]^{-1}\| \|A_0^T - A_*^T\| \|F(x_0)\| \\ \le \beta \textstyle\int_0^{\rho(x_0)} L(z)dz\rho(x_0) + \alpha\beta_0 \int_0^{2\rho(x_0)} L(z)dz. \end{aligned}$$

Since $L(u) > 0$, it follows from lemma 2.4 that $\frac{1}{t^\alpha} \int_0^t L(u)u^{\alpha-1}du$ is monotonic nondecreasing by $t$. Then, taking into account (4.2) we get

$$\begin{aligned} \|x_0 - x_*\| \le \beta \textstyle\int_0^{\rho(x_0)} L(z)dz\rho(x_0) + \frac{\alpha\beta_0}{\rho(x_0)} \int_0^{2\rho(x_0)} L(z)dz\rho(x_0) \\ < \beta \textstyle\int_0^r L(z)dz\|x_0 - x_*\| + \frac{\alpha\beta_0}{r} \int_0^{2r} L(z)dz\|x_0 - x_*\| \\ \le \|x_0 - x_*\|. \end{aligned}$$

This is in contradiction with our assumption. Thus, it follows that $x_0 = x_*$. □

## 5. Corollaries

Let us consider the traditional approach to study of method (1.2), which is based on the fact that, divided differences of the first order satisfy Lipschitz condition (2.2). By taking $L$ to be a constant, the following corollaries are obtained under theorems 3.1 i 4.1.

**Corollary 5.1** *Suppose that*

1) *$x_*$ satisfies (1.1), function $F(x)$ is continuous in $\Omega(x_*, r)$, there exist Jacobian matrix $F'(x_*)$ and it has full column rank;*

2) *in region $\Omega(x_*, r)$ function $F(x)$ has the divided difference of the first order $F(x, y)$, which has full column rank and satisfy Lipschitz condition:*

$$\|F(x,y) - F(u,v)\| \leq L(\|x-u\| + \|y-v\|), \tag{5.1}$$

*where $x, y, u, v \in \Omega$, $L$ is positive number;*

3)

$$r = \frac{1 - 2\sqrt{2}\alpha\beta^2 L}{3\beta L}, \tag{5.2}$$

*where $\alpha, \beta$ are determined in (3.4)*

*Then method (1.2) converges for all $x_{-1}, x_0 \in \Omega$ such that $\rho(x_{-1}) \leq \rho(x_0)$, where $\alpha, \beta, \rho(x)$ are determined in (3.4). For*

$$\begin{aligned} q_0 &= \frac{\sqrt{2}\alpha\beta^2 L((2-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|)}{\|x_{-1} - x_*\|(1 - \beta L((2-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|)}; \\ q_1 &= \frac{\beta L((1-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|)}{1 - \beta L((2-\alpha_0)\|x_0 - x_*\| + \alpha_0\|x_{-1} - x_*\|} \end{aligned} \tag{5.3}$$

*inequality (3.3) holds.*

**Corollary 5.2** *Suppose that*

1) *$x_*$ satisfies (1.1), function $F(x)$ is continuous in $\Omega(x_*, r)$, there exist Jacobian matrix $F'(x_*)$ and it has full column rank;*

2) *in region $\Omega(x_*, r)$ function $F(x)$ has the divided difference of the first order $F(x, y)$, and it has full column rank and satisfies Lipschitz condition:*

$$\|F(x,y) - F(x_*, x_*)\| \leq L(\|x - x_*\| + \|y - x_*\|), \tag{5.4}$$

*where $x, y \in \Omega$ and $L$ is positive number;*

3)

$$r = \frac{\beta L}{1 - 2\alpha\beta_0 L}, \tag{5.5}$$

*where $\alpha, \beta, \beta_0$ are determined in (3.4) and (4.3).*

*Then problem (1.1) has unique solution $x_*$ in $\Omega(x_*, r)$.*

## 6. Convergence analysis under Hölder condition

We have studied the local convergence of method(1.2) under generalized Lipschitz conditions. We have also proved that sequence $x_n$ generated by method (1.2) converges to a solution $x_*$ of problem (1.1) superlinearly. In this section we will study the convergence of method (1.2) under generalized Hölder condition in the case, when $\alpha_n = O(\|x_n - x_*\|)$.

**Theorem 6.1** *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ is continuous in region $D \subseteq \mathbb{R}^n$. Suppose, that:*

1) *problem (1.1) has solution $x_*$ in $\Omega(x_*, r) = \{x \in D : \|x - x_*\| < r\}$ and there exist the Jacobian matrix $F^{'}(x_*)$ and it has full column rank;*
2) *in region $\Omega(x_*, r)$ function $F(x)$ has the divided difference of first order $F(x, y)$, which has full column rank and satisfies Hölder condition with $0 < p \le 1$:*

$$\|F(x,y) - F(u,v)\| \le \int_0^{\|x-u\|^p+\|y-v\|^p} L(z)dz, \tag{6.1}$$

*where $x, y, u, v \in \Omega(x_*, r)$ and $L$ is nondecreasing;*

3) *$r$ satisfies inequality*

$$\frac{\beta \int_0^{r^p} L(z)dz}{1 - \beta \int_0^{2r^p} L(z)dz} + \frac{\sqrt{2}\alpha\beta^2 \int_0^{2r^p} L(z)dz}{1 - \beta \int_0^{2r^p} L(z)dz} \le 1. \tag{6.2}$$

*Then method (1.2) converges for all $x_{-1}, x_0 \in \Omega$ such that $\rho(x_{-1}) \le \rho(x_0)$ and*

$$\begin{aligned} \|x_{n+1} - x_*\| \le q_1 &\frac{(1 - \alpha_n + O(1)\rho(x_{n-1}))^p}{\gamma_0^p} \|x_n - x_*\|^{p+1} \\ +q_0 &\frac{1 + (1 - \alpha_n + O(1)\rho(x_{n-1}))^p}{\rho(x_0)^p + \gamma_0^p} \|x_n - x_*\|^p, \end{aligned} \tag{6.3}$$

*where $\alpha$ and $\beta$ are determined in (3.4), $\gamma_0 = ((1 - \alpha_0)\rho(x_0) + \alpha_0\rho(x_{-1}))$ and values*

$$q_0 = \frac{\sqrt{2}\alpha\beta^2 \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz}{1 - \beta \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz}; q_1 = \frac{\beta \int_0^{\gamma_0^p} L(z)dz}{1 - \beta \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz} \tag{6.4}$$

*are less than 1.*

*Proof.* By theorem 3.1 it is easy to see that

$$\begin{aligned} \|x_{n+1} - x_*\| &\le \frac{\beta \int_0^{\gamma_n^p} L(z)dz \|x_n - x_*\|}{\gamma_n^p (1 - \beta \int_0^{\rho(x_n)^p+\gamma_n^p} L(z)dz} \gamma_n^p \\ &+ \frac{\sqrt{2}\alpha\beta^2 \int_0^{\rho(x_n)^p+\gamma_n^p} L(z)dz}{(\rho(x_n)^p + \gamma_n^p)(1 - \beta \int_0^{\rho(x_n)^p+\gamma_n^p} L(z)dz)} (\rho(x_n)^p + \gamma_n^p) \\ &\le \frac{\beta \int_0^{\gamma_0^p} L(z)dz \|x_n - x_*\|}{\gamma_0^p (1 - \beta \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz)} \gamma_n^p \\ &+ \frac{\sqrt{2}\alpha\beta^2 \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz}{(\rho(x_0)^p + \gamma_0^p)(1 - \beta \int_0^{\rho(x_0)^p+\gamma_0^p} L(z)dz)} (\rho(x_n)^p + \gamma_n^p) \\ &= q_1 \frac{(1 - \alpha_n + O(1)\rho(x_{n-1}))^p}{\gamma_0^p} \|x_n - x_*\|^{p+1} \\ &+ q_0 \frac{1 + (1 - \alpha_n + O(1)\rho(x_{n-1}))^p}{\rho(x_0)^p + \gamma_0^p} \|x_n - x_*\|^p, \end{aligned}$$

where $\gamma_n = ((1-\alpha_n)\rho(x_n) + \alpha_n\rho(x_{n-1}))$. So (6.3) holds. □

It follows from (6.3) that in the case of null residual in the solution ($\alpha = 0$) value $q_0 = 0$ and rate of convergence of the iterative process (1.2) is equal $1+p$. If $p = 1$, that corresponds with Lipschitz condition, we obtain the quadratic convergence as in Gauss-Newton method and it is higher than in Secant method ($(1+\sqrt{5})/2$). At the same time the amount of evaluations per one iteration are equal both in method (1.2) and Secant method.

## 7. Numerical experiments

In this section, we apply the studied above iterative Secant type method (1.2) to solve some test examples, proposed in [6], and compare the convergence behavior of this method under different values of parameter $\alpha_n$. As $\alpha_n \in [0,1]$, we will carry out evaluations with bound values of parameter,that is $\alpha_n = 1$ (Secant method) and $\alpha_n = 0$ (classical Gauss-Newton method), with constant values $\alpha_n$ from a unit interval and taking $\alpha_n$ as a variable value.

In evaluations we use the Euclidian norm $\|x\| = \sqrt{\sum_{i=1}^{n} x_i^2}$, and calculate the elements of the divided difference matrix as

$$F(x,y)_{i,j} = \frac{F_i(x_1,...,x_j,y_{j+1},...,y_n) - F_i(x_1,...,x_{j-1},y_j,...,y_n)}{x_j - y_j}, \quad i = 1,2,...,m, j = 1,2,...,n \tag{7.1}$$

The auxiliary initial approximations we evaluate as $x_{-1} = x_0 + 10^{-4}$. We find the solution with accuracy $\varepsilon = 10^{-8}$. And in tables we denote that $\Delta x_n = \|x_n - x_{n-1}\|$.

In the following tables 1-3 we present the results of numerical experiments. The comparison is conducted by the number of iterations performed to find the solution with a given accuracy.

Tabl. 1. *The number of iterations to receive the solution of test problems under the constant values $\alpha_n$*

| Example | $\alpha_n$ | | | | | |
|---|---|---|---|---|---|---|
| | 0 | 0.2 | 0.4 | 0.6 | 0.8 | 1 |
| Rosenbrock function | 2 | 3 | 3 | 3 | 3 | 3 |
| Wood function | 51 | 56 | 60 | 65 | 69 | 74 |
| Powell singular function | 12 | 13 | 13 | 14 | 19 | 16 |
| Box-3D Function | 6 | 6 | 7 | 7 | 7 | 8 |
| Freudenstein and Roth function | 30 | 12 | 21 | 57 | 29 | 19 |
| Kowalik and Osborne function | 10 | 10 | 12 | 21 | 20 | 16 |

Tabl. 2. *The number of iterations to receive the solution of test problems under the variable values* $\alpha_n$

| Example | $\alpha_n$ | | |
|---|---|---|---|
| | $10^{-2}\Delta x_n$ | $10^{-4}\Delta x_n$ | $\Delta x_n, if \Delta x_n < 1,$ $\frac{1}{\Delta x_n}, if \Delta x_n \geq 1$ |
| Rosenbrock function | 3 | 2 | 3 |
| Wood function | 51 | 51 | 53 |
| Powell singular function | 12 | 12 | 13 |
| Box-3D function | 6 | 6 | 6 |
| Freudenstein and Roth function | 13 | 19 | 9 |
| Kowalik and Osborne function | 10 | 15 | 11 |

Let us present more detailed results for Box-3D function ($n = 3, m = 15$) under different values of $\alpha_n$.

Tabl. 3. *Box-3D function residual at each of the iterations*

| Iterations | $\alpha_n = 0$ | $\alpha_n = 1$ | $\alpha_n = 10^{-2}\Delta x_n$ |
|---|---|---|---|
| 1 | 9.77530447E-02 | 9.77323734E-02 | 9.77323734E-02 |
| 2 | 6.28363713E-03 | 2.24070118E-02 | 8.68229361E-03 |
| 3 | 8.79598795E-05 | 1.93329610E-03 | 1.60820273E-04 |
| 4 | 2.63517737E-08 | 8.33557862E-05 | 9.89442135E-08 |
| 5 | 2.79495793E-15 | 5.37082031E-07 | 4.14185366E-14 |
| 6 | 2.41542623E-29 | 1.64970794E-10 | 6.97021661E-27 |
| 7 | | 3.36221280E-16 | |
| 8 | | 2.13441781E-25 | |

The above presented results in tables show that under successful choice of parameter $\alpha_n$ the proposed method (1.2) converges faster than Secant method and often practically is not worse than the Gauss-Newton method. Moreover, unlike the Gauss-Newton method, this method does not demand the analytically set derivatives.

## 8. Conclusion

In this paper for solving the nonlinear least squares problem the iterative Secant type method is proposed. It is proved that process, generated by this method converges and the superlinear order of convergence is established. It is studied the local convergence of the of the proposed method under generalized Lipschitz conditions and under Hölder conditions. On the basis of the carried out theoretical studies, practical calculations and comparison of the results it can be argued that the proposed modification with the successful choice of parameter $\alpha_n$ prevails the classic Secant method. The studied method is the effective method for solving the nonlinear least squares problems.

## Bibliography


1. Argyros I.K. On an Algorithm for Solving Nonlinear Operator Equation // Zeitschrift für Analysis und ihre Anwendungen. 1991. Vol. 10. N 1. P. 83-92.
2. Bartish M.Ja., Shcherbina Yu.M. On one difference method to solve non-linear operator equations // Dopov. AN URSR. Ser. A. 1972. Vol 7. P. 579-582. (in Ukrainian)
3. Chen J., Li W. Convergence of Gauss-Newton method and uniqueness of the solution // AMC (Appl. Math. Comp.) 2005. Vol. 170. P. 686-705.
4. Chen J., Shen Z. Convergence analysis of the secant type method // AMC (Appl. Math. Comp.) 2007. Vol. 188. P. 514-524.
5. Dennis J.M., Schnabel R.B. *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* New York: Prentice-Hall, 1983.
6. More J.J., Garbow B.S., Hillstrom K.E. Testing Unconstrained Optimization Software // ACM Transactions on Mathematical Software. 1981. Vol. 7. N 1. P. 17-41.
7. Shakhno S.M. Secant method under generalized Lipschitz conditions for the divided differences of the first order // Matem. Visn. NTSh. 2007. Vol. 4. P. 296-305. (in Ukrainian)
8. 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.) 2005 Vol. 161. P. 253-264.
9. Stewart G.W. On the continuity of the generalized inverse // SIAM J. Appl. Math. 1960. Vol. 17. P. 33-45.
10. Wang X. *Convergence of Newton's method and uniquiness of the solution of equations in Banach space* // IMA Journal of Numerical Analysis. 2000. Vol. 20. P. 123-134.
11. Wedin P.A. *Perturbation theory for pseudo-inverse* // BIT. 1973. Vol. 13. P. 217-232.


Ivan Franko National University of Lviv, 1, Universytets'ka Str., 79000, Lviv, Ukraine
*E-mail address*: shakhno@is.lviv.ua

Received 14.01.2009; revised 01.02.2009