========= MATHEMATICS =========

Some Properties of Smooth Convex Functions and Newton's Method

D. V. Denisov[a,*], **Academician of the RAS Yu. G. Evtushenko**[a,b,c,d,**], **and A. A. Tret'yakov**[b,e,f,***]

Abstract—New properties of convex infinitely differentiable functions related to extremal problems are established. It is shown that, in a neighborhood of the solution, even if the Hessian matrix is singular at the solution point of the function to be minimized, the gradient of the objective function belongs to the image of its second derivative. Due to this new property of convex functions, Newtonian methods for solving unconstrained optimization problems can be applied without assuming the nonsingularity of the Hessian matrix at the solution of the problem and their rate of convergence in argument can be estimated under fairly general assumptions.

The unconstrained minimization problem deals with functions $f(\cdot)$ of n real variables defined and sufficiently smooth in a neighborhood $U(x^*)$ of the minimizer. Throughout this paper, the set of minimizers of f is denoted by $X^* = \text{Arg min } f$ and is assumed to be nonempty. The necessary condition for x^* to be a minimizer of f is the equality $f'(x^*) = 0$, and the Hessian matrix of f at the minimizer is positive semidefinite. Newton's method has been studied in numerous works, for example, in [2−6]. An overview of relevant studies can be found in [1]. In this paper, we show that, even if the Hessian matrix is singular at the point x^*, the objective function gradient in a neighborhood of x^* belongs to the image of its second derivative and, hence, the Newtonian system for the descent direction is solvable at points of this neighborhood. This topological property of convexity and optimality provides a new insight into Newton-type methods and makes it possible to prove their rate of convergence without the restrictive assumption that the Hessian matrix is nonsingular at x^*. For unconstrained optimization problems, we establish properties that refine Lojasiewicz's inequality [7, 8]. Namely, the inequality is generalized to the entire spectrum of derivatives up to a certain order. We consider functions whose derivatives are bounded uniformly in argument and jointly in order in a neighborhood $U(x^*)$ of x^*, i.e., for this neighborhood, there exists a positive constant M such that the values of the derivatives of any order do not exceed M in absolute value. Additionally, we consider functions that are sufficiently smooth in a neighborhood of the minimizer, i.e., infinitely differentiable functions. Throughout this paper, without loss of generality, we assume that $f(x^*) = 0$ and use the notation $f^{(0)}(x) = f(x)$. The following result holds for functions of one variable.

Lemma 1. *If x^* is an isolated local minimizer of a function $f : R \to R$ that is sufficiently smooth in a neighborhood $U(x^*)$ and has derivatives that are bounded in $U(x^*)$ uniformly in argument and jointly in order, then there exists an even $2p$, $p = p_f \in \mathbb{N}$, such that*

$$f^{(2p)}(x^*) > 0, \quad f^{(k)}(x^*) = 0,$$

$$\frac{f^{(k)}(x)}{(x - x^*)^{2p-k}} \geq C_k, \quad k = 0,1,2,\ldots,2p-1,$$

for all $x \in \overset{\circ}{U}(x^*)$, where $C_k = C_{k,f}$ for $k = 0, 1, 2, \ldots, 2p - 1$ are positive constants independent of x.

[a] *Faculty of Computational Mathematics and Cybernetics, Lomonosov Moscow State University, Moscow, 119991 Russia*

[b] *Dorodnicyn Computing Centre, Federal Research Center "Computer Science and Control," Russian Academy of Sciences, Moscow, 119333 Russia*

[c] *Moscow Institute of Physics and Technology (National Research University), Dolgoprudnyi, Moscow oblast, 141700 Russia*

[d] *Moscow Aviation Institute (National Research University), Moscow, 125080 Russia*

[e] *System Research Institute, Polish Academy of Sciences, Warsaw, 01-447 Poland*

[f] *Faculty of Sciences, Siedlce University, Siedlce, 08-110 Poland*

*e-mail: dvden@cs.msu.ru

**e-mail: yuri-evtushenko@yandex.ru

***e-mail: tret@ap.siedlce.pl

Proof. Since x^* is a local minimizer of f, we have $f'(x^*) = 0$. First, we show that, under the conditions of Lemma 1, it is not possible that $f^{(k)}(x^*) = 0$ for any k, $k = 1, 2, \ldots$. Indeed, in this case, given a fixed point $x = x^* + t \in U(x^*)$, since the derivatives are bounded in a neighborhood of x^*, the Taylor formula with zero derivatives up to any fixed order k yields the inequality $|f'(x^* + \theta t)| \leq \dfrac{M(|\theta t|)^k}{k!}$, where M is the above-introduced the upper bound for the values of the derivatives of f in the indicated neighborhood and $\theta \in (0,1)$. Combining this result with the condition that the local minimizer x^* is isolated yields $f(x^* + t) = \displaystyle\int_0^1 f'(x^* + \theta t)d\theta \leq \dfrac{M}{(k+1)!}$. For sufficiently large k, this means a violation of the possible assumption $f(x) \neq 0$, $x \neq x^*$, which contradicts the fact that the local minimizer x^* is isolated. Therefore, there exists a finite $k > 1$ such that $f^{(k)}(x^*) \neq 0$ and $f^{(l)}(x^*) = 0$, $l = 1, 2, \ldots, k-1$. Moreover, k can only be even: $k = 2p$, $p \in \mathbb{N}$, and $f^{(2p)}(x^*) > 0$, since x^* is a local minimizer of f. Now, to complete the proof, we denote $\dfrac{1}{(2p+1)!} f^{(2p)}(x^*)$ by C_0. Then the first inequality in the lemma follows from the Taylor formula and the boundedness of the $(2p+1)$th derivative in the neighborhood $U(x^*)$. The Taylor expansion of $f^{(k)}(x)$ in a neighborhood of x^* yields the other inequalities of the lemma for any k, $k = 1, 2, \ldots, 2p-1$, if C_k denotes $\dfrac{1}{(2p+1-k)!} f^{(2p)}(x^*)$.

Corollary 1. *Under the conditions of Lemma* 1, *the function f is locally convex.*

Indeed, if $f^{(2)}(x^*) > 0$, then the second derivative remains positive in a small neighborhood of x^*, which means that $f(x)$ is locally convex. If $f^{(2)}(x^*) = 0$, then, as is shown in Lemma 1, $f^{(2p)}(x^*) > 0$ for some $p = p_f \in \mathbb{N}$, $p > 1$, and, additionally, $f^{(k)}(x^*) = 0$ for $k = 1, 2, \ldots, 2p-1$. Then $f^{(2)}(x) \geq C_2(x - x^*)^{2p-2}$ for any $x \in U(x^*)$, where C_2 is a positive constant determined in Lemma 1. This inequality also means that f is locally convex.

For a function of n variables, the kth derivative in the direction h at a point x is denoted by $f_h^{(k)}(x)$.

Corollary 2. *If $f: R^n \to R$ is a sufficiently smooth function whose derivatives are bounded uniformly in argument and jointly in order in a neighborhood $U(x^*)$ of an isolated minimizer x^*, then, for any $h \in R^n: \|h\| = 1$, there exists an even $2p_{h,f}$, $p_{h,f} \in \mathbb{N}$, such that $f^{(2p_{h,f})}(x^*) > 0$,*

$f^{(k)}(x^*) = 0$, *and* $\dfrac{f^{(k)}(x^* + th)}{t^{2p_{h,f} - k}} \geq C_k$, $k = 0, 1, 2, \ldots,$ $2p_{h,f} - 1$ *for all* $t \in (0, \delta_h]$, *where* $C_k = C_{k,h,f}$ *for* $k = 0, 1, 2, \ldots, 2p_{h,f} - 1$ *are positive constants independent of t, $x^* + \delta_h \in U(x^*)$; moreover, the function f is locally convex in the direction h.*

Here, the element $p = p_{h,f} \in \mathbb{N}$ is determined by the direction $h, \|h\| = 1$ and does not depend on small t for which $x^* + \delta_h \in U(x^*)$, but the value $\delta_h > 0$ depends on h and, in the general case, δ_h can be infinitesimal with respect to h. It will be shown in Lemma 2 below that, for a convex function f, we can guarantee the existence of an upper bound for $p_{h,f} \in \mathbb{N}$ and a positive lower bound for δ_h on the set of vectors $h: \|h\| = 1$.

Lemma 1 implies that the Newton operator $\psi(x) = x - f^{(2)}(x)^{-1} f'(x)$ is well defined in a sufficiently small deleted neighborhood of the solution. In the case of R^n of arbitrary dimension, Lemma 1 means that f is convex along any straight line passing through the point x^*, assuming that the function is sufficiently smooth in the intersection of the straight line and the neighborhood $U(x^*)$. Additionally, for any straight line passing through the point x^* along the vector h, we have $f_h^{(2p)}(x^*) \geq C_2$ for some exponent $2p$, $p = p_h \in \mathbb{N}$ and some constant $C_2 = C_{2,h} > 0$. Lojasiewicz's inequality guarantees that this inequality holds if the function is analytic in $U(x^*)$ for some p and C_2 independent of h. Due to the convexity of the function in $U(x^*)$, we assume that the function is only sufficiently smooth and obtain the "stability" of Lojasiewicz's inequality, which expands the applicability of Newton's method.

Lemma 2. *If $f: R^n \to R$ is a convex sufficiently smooth function whose derivatives are bounded uniformly in argument and jointly in order in a neighborhood $U(x^*)$ of an isolated minimizer x^*, then there exists a even exponent $2p$, $p = p_f \in \mathbb{N}$, a constant $C = C_f > 0$, and a number $\delta = \delta_f > 0$ such that*

$$f(x^* + th) \geq Ct^{2p}$$

for all $h: \|h\| = 1$, $t \in (0, \delta]$.

Proof. Let us prove the existence of $p = p_f \in \mathbb{N}$ such that, if $f_h^{(i)}(x^*) = 0$ for $i = 0, 1, 2, \ldots, 2p'(h) - 1$ and $f_h^{(2p'(h))}(x^*) > 0$, then $p'(h) \leq p$. Assume the opposite. Then, for some sequences $h_k: \|h_k\| = 1$, $t_k: t_k \to +0$, it is true that $f(x^* + t_k h_k) < C_k t^{m_k}$, where $C_k > 0$ is a bounded sequence and $m_k \in \mathbb{N}$ is an increasing sequence. Without loss of generality, we assume that $h_k \to h$ and $C_k \to C$ as $k \to \infty$. Let $\text{conv}\{h_k, h_{k+1}, h_{k+2}, \ldots, h_{k+n}\}$ be denoted by V_k. If $\dim V_k < n$, this set of vectors is mod-

ified by adding linearly independent increments of length on the order of $C_k t^{2p}$ ($C_k \to 0$, $k \to \infty$) to $n - \dim V_k$ vectors, after which we can assume that $\dim V_k = n$. Next, suppose that $h'_k \in \operatorname{int} V_k$, $h''_k = h + \alpha_k(h'_k - h)$, $\alpha_k \in (0,1)$, $\alpha_k = \inf\{\alpha > 0: h + \alpha(h'_k - h) \in V_k\}$. The convexity and continuity of f imply that $f(x^* + t_k h''_k) \le 2Ct^{m_k}$. On the other hand, Lemma 1 implies that, for some $p = p_h \in \mathbb{N}$, there exists $k \in \mathbb{N}$, $k \le p$, for which $f_h^{(2k)}(x^*) \ge C_2 > 0$ for some $C_2 > 0$. Since $h''_k \to h$ as $k \to \infty$, for sufficiently large indices k, we have $f(x^* + t_k h''_k) \ge C_2 t^{2k}$, which contradicts the assumption.

Corollary 3. *The proof of Lemma 2 implies the existence of an exponent $2p$, $p = p_f \in \mathbb{N}$, and constants $C_k = C_{k,f}$, $k = 1, 2, ..., 2p' - 1$, for which*

$$f_h^{(2p')}(x^*) > 0, \quad f_h^{(k)}(x^*) = 0, \quad k = 0, 1, ..., 2p' - 1,$$
$$f_h^{(k)}(x^* + th) \ge C_k t^{2p-k}, \quad 0 < t \le \delta,$$

moreover, $p' = p'(h) \le p$ for all h: $\|h\| = 1$.

For the point $x = x^* + th \in U(x^*)$, where $\|h\| = 1$, the basis $G = G(h) = \{g_1, g_2, ..., g_n\}$ in R^n and the corresponding index $q = q(h)$ are determined as follows. Consider the matrices

$$a_k = \frac{1}{(k-1)!} f^{(k)}(x^*)[h]^{k-2}, \quad k = 2, 3, ...$$

Then, for a sufficiently smooth function f at the point $x = x^* + th$, we have the relations

$$f^{(2)}(x)th = \sum_{k \ge 2}(k-1)a_k ht^{k-1},$$
$$f'(x) = \sum_{k \ge 2} a_k ht^{k-1}, \quad f(x) = \sum_{k \ge 2} \frac{a_k[h]^2 t^k}{k}. \tag{1}$$

Next the index $k_1 \ge 2$ is defined as the minimum among those indices k for which the one-dimensional space $L_1 = L_{k_1} = Lin\{a_k h\} \ne \{0\}$. The straight line L_1 is redenoted by L^1, the vector g_1 is defined as $\frac{a_{k_1}h}{\|a_{k_1}h\|}$, and the index $k_2 > k_1$ is defined as the minimum k for which $Lin\{a_k h\}$ is not contained in L^1. Then $L_{k_2} = Lin\{a_{k_2}h\}$, $\dim(L_{k_2} \oplus L^1) = 2$, the one-dimensional subspace L_2 is defined as $Pr_{(L^1)^\perp} L_{k_2}$, and the vector g_2 is defined as the normalized vector $Pr_{L_2} a_{k_2} h$. Here, $Pr_{L_2}(a_{k_2}h) \subseteq Pr_{L_2} Im a_{k_2}$. Next, the space L^2 is defined as $L^1 \oplus L_2$. For each $j = 3, 4, ..., q = q(h)$, in a similar fashion, we determine the straight lines $L_j = Pr_{(L^{j-1})^\perp} L_{k_j}$, $j = 3, 4, ..., q$ and the corresponding unit vectors g_j, $j = 1, 2, ..., q$.

Here, $q = q(h) = \dim Lin\{a_1 h, a_2 h, ...\} \le n$, the direct sum $L^{j-1} \oplus L_j$ is denoted as L^j and is a j-dimensional subspace of R^n. At the final stage, we construct a subspace L^q of dimension $q = q(h) \le n$. The orthogonal complement of L^q is H: $H = (L^q)^\perp$ for $q < n$ and $H = \{0\}$ for $q = n$. The basis $G(h)$ in R^n is defined as the set consisting of the constructed mutually orthogonal unit vectors $g_1, g_2, ..., g_q$ directed along the constructed orthogonal straight lines L_l, $l = 1, 2, ..., q$, and an arbitrary orthogonal basis $g_{q+1}, g_{q+2}, ..., g_n$ in H.

Theorem 1. *If $f: R^n \to R$ is a sufficiently smooth function whose derivatives are bounded uniformly in argument and jointly in order in a neighborhood $U(x^*)$ of an isolated minimizer x^*, then, for any fixed h: $\|h\| = 1$, the system*

$$f^{(2)}(x^* + th)s = f'(x^* + th) \tag{2}$$

is solvable with respect to s for sufficiently small t.

Proof. For any fixed vector h, the solution s of the system in the basis G is determined coordinatewise as follows. Let $Pr_H s = 0$. Then, according to (1), the coordinate s_q is determined by the condition $Pr_{L_q}((k_q - 1)a_{k_q} t^{k_q - 2}s) = Pr_{L_q} f'(x) = Pr_{L_q} \sum_{k \ge k_q} a_k ht^{k-1}$.

Specifically, $s_q = \frac{Pr_{L_q} ht}{(k_q - 1)} + o(t)$. Substituting the coordinate s_q into system (2) yields the coordinate $s_{q-1} = \frac{Pr_{L_{q-1}} ht}{(k_{q-1} - 1)} + o(t)$. Next, the coordinates s_l, $l = q - 2, q - 3, ..., 1$, are determined sequentially. Generally speaking, the solution of system (2) is not unique, but it follows from (1) that the coordinates of any solution s in the basis $G = G(h)$ are given by

$$s_j = \frac{Pr_{L_j} ht}{k_j - 1} + o(t), \quad j = 1, 2, ..., q. \tag{3}$$

Remark 1. To obtain a unique solution of system (2), we consider the problem

$$\|s\|^2 \to \min_s, \quad f^{(2)}(x)s = f'(x), \tag{4}$$

whose solution exists for each fixed h for all sufficiently small t under the assumptions made in Theorem 1. The length of the interval in t within which the solution exists depends on h.

Example 1. For the nonconvex function $f(x) = x_1^4 + (x_2 - x_1^2)^2$, Theorem 1 is violated for points of the parabola $x_2 = 4x_1^2$. Thus, the length of the interval on which Theorem 1 holds tends to zero as the vector h approaches $(1, 0)$.

Below, we show that, in the case of a convex function, there is a neighborhood with respect to t with a common radius for all vectors of the unit sphere in which problem (4) is solvable. The form of the objective function of problem (4) implies that its solution is given by $s = f^{(2)}(x)^+ f'(x)$, where $(\cdot)^+$ denotes a pseudoinverse operator on its image. The coordinates of the solution to this problem in the basis G satisfy the condition $Pr_H s = 0$.

Theorem 2. *Under the conditions of Lemma* 2, *for all x sufficiently close to x^*, problem* (4) *is solvable and its solution satisfies the conditions*

$$
\begin{aligned}
(f'(x), s) &\geq M_1 \|x - x^*\|^{2p}, \\
(f''(x)s, s) &\geq M_2 \|x - x^*\|^{2p},
\end{aligned}
\tag{5}
$$

where the exponent $2p$ is defined in Lemma 2 *and the constants $M_1 = M_{1,f}$ and $M_2 = M_{2,f} > 0$ do not depend on x.*

Proof. Lemma 2 and (1) imply that, for every $h: \|h\| = 1$, the numbers k_j determining the basis G satisfy the following condition: there is an index $l_1 \in \{k_1, k_1 + 1, \ldots, 2p\}$ for which $a_{l_1}[h]^2 \geq C > 0$, the exponent $2p$ is defined in Lemma 2 and is independent of h, and the constant $C > 0$ is also independent of h and is defined in Corollary 3. The last inequality is equivalent to the condition $\dfrac{a_{l_1}[h]^2}{l_1 - 1} \geq C_1 t^{2p}$ for small t. Denoting $\dfrac{C}{2p - 1}$ by M_1, we derive the inequality stated in the theorem. The second inequality is derived from Corollary 3 in a similar manner.

Corollary 4. *Under the conditions of Lemma* 2, *for a convex function f, the system*

$$
f^{(2)}(x)s = f'(x)
\tag{6}
$$

solvable with respect to s for all x sufficiently close to x^*, which is equivalent to the inclusion

$$
f'(x) \in \operatorname{Im} f^{(2)}(x)
\tag{7}
$$

irrespective of the rank of the matrix $f^{(2)}(x)$.

In what follows, the set $\{x \in R^n: f(x) \leq f(x^*) + \varepsilon\}$ is denoted by X_ε^* and the diameter of a set A is denoted by diam$A$.

Theorem 3. *Under the conditions of Lemma* 2, *there exists a positive integer p and a constant $\overline{C} < \infty$ for which*

$$
\operatorname{diam} X_\varepsilon^* < \overline{C} \varepsilon^{\frac{1}{2p}}.
\tag{8}
$$

Remark 2. The above-mentioned properties hold assuming that the set of minimizers of $f(x)$ is an isolated point. If this assumption is rejected, while the other assumptions of Theorem 1 are satisfied, then we have the representation

$$
\operatorname{Arg\,min} f = \{x^*\} + L,
\tag{9}
$$

where L is the eigensubspace of R^n determined by the condition $f^k(x^*)[h]^k = 0$ for any positive integer k.

Theorem 1 does not imply that $f^{(2)}(x)$ is positive definite in a neighborhood of the minimizer x^*. Nevertheless, it guarantees that Newton's method with a suitable initial point is applicable for finding a minimizer of a smooth function, since problem (1) makes it possible to obtain a transition vector in the iterative scheme without assuming that the objective function is convex. Moreover, in this case, we can obtain the monotonicity of Newton's method in argument and, under stronger assumptions, a linear rate of convergence. If the objective function is convex, then this property holds without making additional assumptions.

First, we use Theorem 1 to prove the monotonicity of Newton's scheme in argument. Specifically, the Newton operator is defined as $\psi(x, s) = x - f^{(2)}(x)^+ f'(x)$, where $f^{(2)}(x)^+ f'(x) = s$ is the solution of problem (4). The process of constructing the solution s in Theorem 1 and the remark to the theorem imply that, for any fixed $h: \|h\| = 1$, this operator is well defined for sufficiently small t without assuming that the function f is convex. To obtain a convergence estimate, the following property is naturally introduced in addition to the assumptions of Theorem 1.

Definition 1. The function f is said to be weakly regular at the point x^* if there exists a constant $d = d_f > 0$ such that $\max\limits_{j \leq q} |h_j| \geq d$ for any h, $\|h\| = 1$. Here, as before, $q = q(h) = \dim \operatorname{Lin}\{a_2 h, a_3 h, \ldots\}$ and h_j is the jth coordinate of the vector h in the basis G.

Theorem 4. *If the function f is weakly regular at the point x^*, then, under the conditions of Theorem* 1, *there exists $\lambda \in (0,1)$ such that*

$$
\|\psi(x^* + th, s) - x^*\| \leq \lambda t
$$

for all sufficiently small t.

Proof. For any solution of system (1), under the weak regularity assumption at the point x^*, there exists an index j for which $\|Pr_{L_{k_j}} h\| > d$. It follows that

$$
\begin{aligned}
\|th - s\| &\leq \left\| t \sum_{l=1, l \neq k_j}^{q} Pr_{L_l} h - Pr_{L_l} s \right\| + t \|Pr_H h\| \\
&\quad + |t Pr_{L_{k_j}} h - Pr_{L_{k_j}} s| \leq t(1 - \alpha) \\
&\quad + t\alpha \left(1 - \frac{1}{k_j - 1} \right) \leq t \left(1 - \frac{d}{k_j - 1} \right) = \lambda t,
\end{aligned}
$$

$$
\lambda = \lambda(h) = 1 - \frac{d}{k_j - 1} \in (0,1), \quad \alpha = \alpha(h) \geq d.
$$

Remark 3. The resulting inequality does not mean the linear convergence of Newton's method, since λ

involved in the estimate depends, generally speaking, on h. In the case of a convex function, the result holds in a stronger form, namely, $\|\psi(x,s) - x^*\| \leq \lambda\|x - x^*\|$ for all x from a sufficiently small neighborhood of x^*, since the following property is valid for convex functions.

Definition 2. The function f is said to be uniformly regular at the point x^* if there exists an index $m = m_f \in \mathbb{N}$ and a constant $d = d_f > 0$ such that there is $j \leq q : k_j \leq m$ for which $\|Pr_{L_j} h\| \geq d$.

Theorem 5. *If the function f is uniformly regular at the point x^*, then, under the conditions of Theorem 1, there exists $\lambda \in (0,1)$ such that $\|\psi(x,s) - x^*\| \leq \lambda\|x - x^*\|$ for all x from a sufficiently small neighborhood of x^*.*

Proof. Theorem 4 implies that the denominator $\lambda = \lambda(h)$ is bounded from above uniformly in h, $\|h\| = 1$:

$\lambda(h) = 1 - \dfrac{d}{k_j - 1} \leq \lambda' < 1$, where $j : k_j \leq m$. Such an index j exists, as was shown in Theorem 1. Additionally, the uniform regularity assumption implies that $\|h_H^\perp\| > d$ for some $d > 0$ independent of h, where h_A is the projection of the vector h onto the subspace A. This yields an h-independent estimate for λ.

Remark 4. Lemma 2 implies that a convex function is uniformly regular at the point x^*: $k_j = k_j(h) \leq m = 2p$, $\|h\| = 1$. Uniform regularity is a necessary and sufficient condition for Newton's method to converge linearly under the hypotheses of Theorem 1 without assuming that the function f is convex. Newton's iteration has the form

$$x_{k+1} = x_k - f^{(2)}(x_k)^+ f'(x_k),$$

and the convergence rate estimate of the method is linear:

$$\|x_{k+1} - x^*\| \leq \lambda\|x_k - x^*\|, \quad \lambda \in (0,1), \quad k = 0,1,2,\dots,$$

for an initial approximation that is sufficiently close to the solution x^* in the case when f is uniformly regular at x^*.

Corollary 5. *Under the conditions of Theorem 2, for a convex function, the rate of convergence of Newton's iteration is linear for any choice of the initial point x from a sufficiently small neighborhood of the solution. Indeed,*

Theorem 2 implies that a convex function is uniformly regular at the solution x^*, which implies a linear convergence estimate.

To ensure that Newton's method converges faster than linearly, we consider a modified operator $\psi_1(x,s)$ written coordinatewise in the basis G as $\psi_1(x,s)_j = x - (k_j - 1)s_j$, $j = 1,2,\dots,q$; $\psi_1(x,s)_H = x_H - g(x,h)$, where $g(x,h) : H \to H$ and $s = f^{(2)}(x)^+ f'(x)$. The operator $\psi_1(x,s)_H$ determines the modified Newton scheme $x_{k+1} = \psi_1(x_k, s_k)$. The following result is a consequence of Theorem 1.

Theorem 6. *Given a good initial approximation, the modified Newton scheme guarantees a superlinear convergence estimate if and only if the function $g(x,h)$ satisfies the condition $\|g(x,h) - h_H\| = o(t)$.*

REFERENCES

1. B. T. Polyak, Tr. Inst. Sist. Anal. Ross. Akad. Nauk **28**, 44−62 (2006).

2. B. Bamadio and K. A. Lebedev, Mezhdunar. Nauch.-Issled. Zh. **37** (6−2), 11−14 (2015).

3. V. I. Zabotin and Yu. A. Chernyaev, Comput. Math. Math. Phys. **58** (3), 322−327 (2018). https://doi.org/10.1134/S0965542518030144

4. D. Budzko, A. Cordero, and J. R. Torregrosa, SeMA J. **66** (1), 43−53 (2014). https://doi.org/10.1007/s40324-014-0020-y

5. Y. Nesterov, Math. Program. **112** (1), 159−181 (2008). https://doi.org/10.1007/s10107-006-0089-x

6. B. Polyak and A. Tremba, Optim. Methods Software **35** (6), 1272−1303 (2020). https://doi.org/10.1080/10556788.2019.1669154

7. T. H. Colding and W. P. Minicozzi, "Lojasiewicz inequalities and applications," arXiv:1402.5087. 2014.

8. S. Lojasiewicz, C. R. Acad. Sci. **246** (5), 683−686 (1958).

Translated by I. Ruzanova