

Analele Universității de Vest, Timișoara Seria Matematică – Informatică LV, 2, (2017), 33–49

Expanding the Applicability of Four Iterative Methods for Solving Least Squares Problems

Ioannis K. Argyros, Janak Raj Sharma, and Deepak Kumar

Abstract. The aim of this paper is to expand the applicability of four iterative methods for solving nonlinear least squares problems. The advantages obtained under the same computational cost as in earlier studies, include: larger radius of convergence, tighter error bounds on the distances involved and a better information on the location of the solution.

AMS Subject Classification (2010). 65G99; 65H10; 47H17; 49M15

Keywords. Local convergence, least squares problems, Gauss-Newton's method, Secant method, Three-point method, Kurchatov-type method

1 Introduction

Many problems in computational sciences and engineering can be written as nonlinear least squares

$$\min_{x \in \mathbb{R}^i} \frac{1}{2} F(x)^T F(x), \tag{1.1}$$

where $F : \Omega \subseteq \mathbb{R}^i \longrightarrow \mathbb{R}^j$ $(j \ge i)$, Ω is open convex and i, j are positive integers. The solution of problem (1.1) is denoted by x^* . The solution x^* is preferred in closed form but this is possible only in special cases. That explains why most solution methods are usually iterative. Let us consider the iterative method defined for each n = 0, 1, 2, ... by

$$x_{n+1} = x_n - [Q_n^T Q_n]^{-1} Q_n^T F(x_n), \qquad (1.2)$$

where $Q_n \in \mathcal{L}(\mathbb{R}^i, \mathbb{R}^j)$ is the space of bounded linear operators from \mathbb{R}^i into \mathbb{R}^i . Special choices of Q_n are considered in this study:

Secant method [2,3,10]: $Q_n = [x_n, x_{n-1}; F]$, where $[\cdot, \cdot; F] : \Omega^2 \to \mathcal{L}(\mathbb{R}^i, \mathbb{R}^j)$ is a divided difference of order one. The convergence order of the method is $\frac{1+\sqrt{5}}{2} = 1.618...$

Three-point method [10]: $Q_n = [x_n, x_{n-1}; F] + [x_{n-2}, x_n; F] - [x_{n-2}, x_{n-1}; F]$. The convergence order is 1.839....

Kurchatov-type method [2,11]: $Q_n = [2x_n - x_{n-1}, x_{n-1}; F]$. The convergence order of the method is two.

These methods are derivative free and are useful alternatives to the Gauss-Newton method obtained from (1.2), if we choose $Q_n = F'(x_n)$. The Gauss-Newton method has convergence order two and has been studied in [2–13] under various Lipschitz-type conditions. In the present paper, we study the local convergence of these methods under more precise Lipschitz-type conditions to obtain the following advantages (say A):

- (a_1) Larger radius of convergence leading to a wider choice of initial guesses.
- (a₂) Tighter error bounds on the distances $||x_n x^*||$ leading to the computation of fewer iterates to obtain a desired error tolerance.
- (a_3) A more precise location of the solution x^* .

The advantages (A) are obtained under the same computational cost as in earlier studies [2,3,10,11], since the computation of the Lipschitz constants in the old studies require the computation of the new Lipschitz constants as special cases. Hence, the applicability of these four methods to solve (1.1) is expanded.

The rest of paper is structured as follows: Sections 2-5 contain the local convergence of Secant method, three-point method, Kurchatov-type method and Gauss-Newton method, respectively.

2 Local convergence for the Secant method

The local convergence analysis in the case when $Q_n = [x_n, x_{n-1}; F]$ is based on the conditions (S):

- (s₁) $F: \Omega \subset \mathbb{R}^i \to \mathbb{R}^j$ is continuously differentiable.
- (s₂) There exist a solution x^* of problem (1.1) and $a > 0, b > 0, c \ge 0$ such that

$$||F'(x^*)|| \le c, \ ||F''(x^*)|| \le a,$$

$$(Q_*^T Q_*)^{-1} = [F'(x^*)^T F'(x^*)]^{-1} \text{ exists and } ||(Q_*^T Q_*)^{-1}|| \le b.$$

(s₃) There exist $K_1 \ge 0$, $K_2 \ge 0$ such that for $x, y \in \Omega$ divided difference of order one $[\cdot, \cdot; F]$ exists and

$$||[x, y; F] - F'(x^*)|| \le K_1 ||x - x^*|| + K_2 ||y - x^*||.$$

Define scalar polynomial h by

$$h(t) = b(K_1 + K_2)t^2 + 2ab(K_1 + K_2)t - 1.$$

Denote by r_0 , the only positive root of polynomial h is given by

$$r_0 = \frac{1}{(ab + \sqrt{a^2b^2 + b})(K_1 + K_2)}, \text{ for } K_1 + K_2 \neq 0.$$

Set: $\Omega_0 = \Omega \cap U(x^*, r_0).$

 (s_4) There exist $K_0 \ge 0, K \ge 0$ such that for each $x, y \in \Omega_0$

$$\|[x,y;\,F] - [x,x^*;\,F]\| \le K \|y - x^*\|$$

and

$$||[x, x^*; F] - F'(x^*)|| \le K_0 ||x - x^*||.$$

Then, we can show the following local convergence result for the Secant method under the (S) conditions and the preceding notation.

Theorem 2.1. Suppose that the conditions (S) hold. Then, sequence $\{x_n\}$ generated for x_{-1} , $x_0 \in U(x^*, r_0) - \{x^*\}$ by the secant method satisfies

$$\|x_{n+1} - x^*\| \le C_n^1 \|x_{n-1} - x^*\| + C_n^2 \|x_n - x^*\| + C_n^3 \|x_{n-1} - x^*\| \|x_n - x^*\|, \quad (2.1)$$

where

$$\begin{split} C_n^1 &= \frac{bcK}{1-p_n}, \ C_n^2 = \frac{bcK_0}{1-p_n}, \\ C_n^3 &= \frac{b(a+K_0\|x_n-x^*\|+K\|x_{n-1}-x^*\|)K}{1-p_n} \end{split}$$

and

$$p_n = b[2a(K_1||x_n - x^*|| + K_2||x_{n-1} - x^*||) + (K_1||x_n - x^*|| + K_2||x_{n-1} - x^*||)^2] < 1.$$

Moreover, in case of zero residual (i.e. c = 0), if abK < 1 and we choose $x_{-1}, x_0 \in U(x^*, r) - \{x^*\}$, then sequence $\{x_n\}$ remains in $U(x^*, r)$ for each $n = -1, 0, 1, 2, \ldots$ and converges to x^* with speed $\frac{1+\sqrt{5}}{2}$ so that

$$||x_{n+1} - x^*|| \le C ||x_{n-1} - x^*|| ||x_n - x^*||, \qquad (2.2)$$

for some $C \geq 0$, where r is the only root in $(0, r_0)$ of polynomial

$$g(t) = b(K_1 + K_2)^2 t^2 + b[2a(K_1 + K_2) + K(K_0 + K)]t + abK - 1.$$

Proof. We shall first show estimate (2.1). By hypothesis linear operator $Q_n \in \mathbb{R}^{j \times i}$ has a full column range and we can write

$$\|[I - (I - (F_*^{T}F_*)^{-1}Q_n^TQ_n)]^{-1}\| = \|(Q_n^TQ_n)^{-1}F_*^{T}F_*^{T}\|,$$
(2.3)

where $Q_n = [x_n, x_{n-1}; F], Q_* = F'(x^*) = F'_*$. We need the estimates

$$||F'_{*} - Q_{n}|| = ||F'^{T}_{*} - Q^{T}_{n}|| \text{ (for the Euclidean norm [9])}, \qquad (2.4)$$

and by (s_2) , (s_3) , (2.3), we get in turn that

$$\begin{aligned} \|I - (F_*^{T}F_*^{\prime})^{-1}Q_n^TQ_n\| &= \|(F_*^{T}F_*^{\prime})^{-1}(F_*^{T}F_*^{\prime} - Q_n^TQ_n)\| \\ &= \|(F_*^{T}F_*^{\prime})^{-1}(F_*^{T}(F_*^{\prime} - Q_n)) \\ &+ (F_*^{\prime T} - Q_n^T)(Q_n - F_*^{\prime}) + (F_*^{\prime T} - Q_n^T)F_*^{\prime})\| \\ &\leq \|(F_*^{\prime T}F_*^{\prime})^{-1}\|(\|F_*^{\prime T}\|\|F_*^{\prime} - Q_n\|) \\ &+ \|F_*^{\prime T} - Q_n^T\|\|Q_n - F_*^{\prime}\| + \|F_*^{\prime T} - Q_n^T\|\|F_*^{\prime}\|) \\ &\leq b(a\|F^{\prime}(x^*) - Q_n\| + \|F^{\prime T}(x^*) - Q_n^T\|\|Q_n - F^{\prime}(x^*)\| \\ &+ a\|F^{\prime T}(x^*) - Q_n^T\|) \\ &\leq b(a(K_1\|x_n - x^*\| + K_2\|x_{n-1} - x^*\|) \\ &+ a(K_1\|x_n - x^*\| + K_2\|x_{n-1} - x^*\|)) = p_n \\ &< b(2a(K_1 + K_2)r_0 + (K_1 + K_2)^2r_0^2) \\ &< g(r_0) + 1 = 0 + 1 = 1, \end{aligned}$$

$$(2.5)$$

by the definition of function g and the choice of r_0 .

36

It follows from (2.5) and the Banach lemma on invertible operators [9] that

$$\|(Q_n^T Q_n)^{-1} F_*'^T F_*'\| \le \frac{1}{1 - p_n}.$$
(2.6)

We also obtain by using (s_2) and (s_4) the estimates

$$\begin{aligned} \|Q_n - Q_*\| &= \|[x_n, x_{n-1}; F] - F'(x^*)\| \\ &= \|([x_n, x_{n-1}; F] - [x_n, x^*; F]) + [x_n, x^*; F] - F'(x^*)\| \\ &\leq \|[x_n, x_{n-1}; F] - [x_n, x^*; F]\| + \|[x_n, x^*; F] - F'(x^*)\| \\ &\leq K \|x_{n-1} - x^*\| + K_0 \|x_n - x^*\|, \end{aligned}$$
(2.7)

$$||Q_n|| = ||Q_* + (Q_n - Q_*)||$$

$$\leq ||Q_*|| + ||(Q_n - Q_*)||$$

$$\leq a + K ||x_{n-1} - x^*|| + K_0 ||x_n - x^*||$$
(2.8)

and

$$||Q_n - [x_n, x^*; F]|| = ||[x_n, x_{n-1}; F] - [x_n, x^*; F]||$$

$$\leq K ||x_{n-1} - x^*||.$$
(2.9)

Using the Secant method and estimates (2.4), (2.6)-(2.9), we obtain in turn that

$$\begin{aligned} \|x_{n+1} - x^*\| &= \|x_n - x^* - (Q_n^T Q_n)^{-1} (Q_n^T F(x_n) - Q_*^T F(x^*))\| \\ &\leq \| - (Q_n^T Q_n)^{-1} F'(x^*)^T F'(x^*)\| \| (F'(x^*)^T F'(x^*))^{-1} \\ &\times [-Q_n^T (Q_n - [x_n, x^*; F])(x_n - x^*) + (Q_n^T - Q_*^T) F(x^*)]\| \\ &\leq b \big((a + K_0 \|x_n - x^*\| + K \|x_{n-1} - x^*\|) K \|x_{n-1} - x^*\| \| \|x_n - x^*\| \\ &+ c (K_0 \|x_n - x^*\| + K \|x_{n-1} - x^*\|) \big) \big/ (1 - p_n), \end{aligned}$$
(2.10)

which shows (2.1) by the definition of C_n^1 , C_n^2 and C_n^3 . Moreover, if c = 0, then $C_n^1 = C_n^2 = 0$ and by the choice of r, we obtain (2.2) and

$$||x_{n+1} - x^*|| < ||x_n - x^*|| < r,$$
(2.11)

so $\lim_{n\to\infty} x_n = x^*$ and $x_{n+1} \in U(x^*, r)$.

Remark 2.1. Estimate (2.1) does not necessarily imply the convergence of sequence $\{x_n\}$ to x^* unless if c is sufficiently small, when we only obtain linear convergence. However, if c = 0 we showed such a convergence.

37

3 Local convergence for the three-point method

We consider the case $Q_n = [x_n, x_{n-1}; F] + [x_{n-2}, x_n; F] - [x_{n-2}, x_{n-1}; F]$ under the conditions (*T*):

- $(t_1) = (s_1).$
- $(t_2) = (s_2).$
- (t₃) There exist $K_0 \ge 0$, $K \ge 0$, $M \ge 0$ and divided differences of order one $[\cdot, \cdot; F]$ on Ω^2 and order two $[\cdot, \cdot, \cdot; F]$ on Ω^3 such that for each x, $y, u_1, u_2, u_3, u_4 \in \Omega$

$$\begin{aligned} \|[x, x^*; F] - F'(x^*)\| &\leq K_0 \|x - x^*\|, \\ \|[x, y; F] - [x, x^*; F]\| &\leq K \|y - x^*\|, \\ \|[u_1, u_3, u_4; F] - [u_2, u_3, u_4; F]\| &\leq M \|u_1 - u_2\|. \end{aligned}$$

Define scalar polynomial h by

$$h(t) = b[2a + (K_0 + K)t + 2Mt^2][K_0 + K + 2Mt]t - 1.$$

Denote by r_0 the only positive root of polynomial h. Then, we can show the following local convergence result for the three point method under the conditions (T).

Theorem 3.1. Suppose that the conditions (T) hold. Then, sequence $\{x_n\}$ generated for $x_{-2}, x_{-1}, x_0 \in U(x^*, r_0) - \{x^*\}$ by the three point method satisfies

$$||x_{n+1} - x^*|| \le \frac{q_n}{1 - p_n},\tag{3.1}$$

where

$$p_n = b[2a + (K_0 + K) ||x_n - x^*|| + M(||x_{n-2} - x^*|| + ||x_n - x^*||) ||x_{n-1} - x^*||]$$

× [(K_0 + K) ||x_n - x^*|| + M(||x_{n-2} - x^*|| + ||x_n - x^*||) ||x_{n-1} - x^*||]

and

$$q_{n} = b \big([a + (K_{0} + K) ||x_{n} - x^{*}|| + M(||x_{n-2} - x^{*}|| + ||x_{n} - x^{*}||) ||x_{n-1} - x^{*}||] \\ \times [M(||x_{n} - x^{*}|| + ||x_{n-1} - x^{*}||)(||x_{n} - x^{*}|| + ||x_{n-2} - x^{*}||) + K ||x_{n} - x^{*}||] ||x_{n} - x^{*}|| \\ + c((K_{0} + K) ||x_{n} - x^{*}|| + M(||x_{n-2} - x^{*}|| + ||x_{n} - x^{*}||)) ||x_{n-1} - x^{*}|| \big).$$

Moreover, if x_{-2} , x_{-1} , $x_0 \in U(x^*, r) - \{x^*\}$, then sequence $\{x_n\}$ remains in $U(x^*, r)$ for each $n = -2, -1, 0, 1, 2, \ldots$ and converges to x^* such that

$$||x_{n+1} - x^*|| \le c_1 ||x_n - x^*|| + c_2 ||x_n - x^*|| ||x_{n-1} - x^*|| ||x_{n-2} - x^*||$$

for some $c_1 \ge 0$, $c_2 \ge 0$, where r is the smallest zero in $(0, r_0)$ of function h defined by

$$h(t) = q(t) + p(t) - 1,$$

where

$$q(t) = b([a + (K_0 + K)t + 2Mt^2][4Mt + K]t + c((K_0 + K) + 2Mt)).$$

Proof. The proof follows the corresponding one in Theorem 2.1, using estimates (2.3), (2.5), (2.10) adjusted using the (t_3) conditions:

$$\begin{aligned} \|Q_n - F'(x^*)\| &= \|[x_n, x_{n-1}; F] + [x_{n-2}, x_n; F] - [x_{n-2}, x_{n-1}; F] - F'(x^*)\| \\ &\leq (K_0 + K) \|x_n - x^*\| + M \|x_{n-2} - x_n\| \|x_{n-1} - x^*\| \\ &\leq (K_0 + K) \|x_n - x^*\| + M (\|x_{n-2} - x^*\| + \|x_n - x^*\|) \|x_{n-1} - x^*\| \end{aligned}$$

$$\begin{aligned} \|Q_n - [x_n, x^*; F]\| &\leq M(\|x_n - x_{n-1}\| \|x_n - x_{n-2}\| + K \|x_n - x^*\| \\ &\leq M(\|x_n - x^*\| + \|x_{n-1} - x^*\|)(\|x_n - x^*\| + \|x_{n-2} - x^*\|) \\ &+ K \|x_n - x^*\|, \end{aligned}$$

$$\begin{aligned} \|Q_n\| &\leq \|Q_n - F'(x^*) + F'(x^*)\| \leq \|F'(x^*)\| + \|Q_n - F'(x^*)\| \\ &\leq a + (K_0 + K)\|x_n - x^*\| + M(\|x_{n-2} - x^*\| + \|x_n - x^*\|)\|x_{n-1} - x^*\|, \end{aligned}$$

so by (2.3)

$$\|I - (F_*'^T F_*')^{-1} Q_n^T Q_n\| \le b[2a + \|Q_n - F'(x^*)\|] \|Q_n - F'(x^*)\| \le p_n < p(r_0) = 1,$$
(3.2)

 \mathbf{SO}

$$\|(Q_n^T Q_n)^{-1} F_*'^T F_*'\| \le \frac{1}{1 - p_n}$$
(3.3)

and by (2.10)

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \left(\|(F'(x^*)^T F'(x^*))^{-1} [-Q_n^T (Q_n - [x_n, x^*; F])(x_n - x^*) + (Q_n^T - Q_*^T) F(x^*)] \| \right) / (1 - p_n) \\ &\leq \frac{q_n}{1 - p_n}, \end{aligned}$$

which shows (3.1). Then, we have by the choice of r and (3.1) that $x_n \in U(x^*, r)$ and $||x_{n+1} - x^*|| < ||x_n - x^*|| < r$, so $\lim_{n\to\infty} x_n = x^*$ and $x_{n+1} \in U(x^*, r)$.

Remark 3.1. It follows from (3.1) that for c = 0, there exist N and C = C(N) such that for each $n \ge N$,

$$||x_{n+1} - x^*|| \le C ||x_{n-2} - x^*|| ||x_{n-1} - x^*|| ||x_n - x^*||_{2}$$

so the order of convergence is the positive root of the equation $t^3 - t^2 - t - 1 = 0$, which is 1.839... Clearly, if $c \neq 0$ the order of convergence is only linear.

4 Local convergence for the Kurchatov method

The local convergence analysis in the case when $Q_n = [2x_n - x_{n-1}, x_{n-1}; F]$ is based on the conditions (V):

- $(v_1) = (s_1).$
- $(v_2) = (s_2).$
- (v_3) There exist $K_0 \ge 0$, $K \ge 0$, $M \ge 0$ such that for $x, y \in \Omega$, divided difference of order one $[\cdot, \cdot; F]$ exists and

$$\begin{aligned} \|[x, x^*; F] - F'(x^*)\| &\leq K_0 \|x - x^*\|, \\ \|[x, x^*; F] - F'(x)\| &\leq K \|x - x^*\|, \\ \|[2y - x, x; F] - F'(y)\| &\leq M \|y - x\|^2. \end{aligned}$$

Define scalar polynomials p and h by

$$p(t) = b[2a + 4Mt^{2} + (K_{0} + K)t][4Mt^{2} + (K_{0} + K)t]$$

and

$$h(t) = p(t) - 1. (4.1)$$

Denote by r_0 the smallest positive root of polynomial h. Then, we can show the following local convergence result for the Kurchatov method under the (V) conditions.

Theorem 4.1. Suppose that the conditions (V) hold and $U(x^*, 3r_0) \subseteq \Omega$. Then, sequence $\{x_n\}$ generated for x_{-1} , $x_0 \in U(x^*, r_0) - \{x^*\}$ by the Kurchatov method satisfies

$$||x_{n+1} - x^*|| \le \frac{q_n}{1 - p_n},\tag{4.2}$$

where

$$p_n = b [(2a + (K_0 + K) || x_n - x^* || + M || x_n - x_{n-1} ||^2) \\ \times ((K_0 + K) || x_n - x^* || + M || x_n - x_{n-1} ||^2)] \\ \le p(r_0)$$

and

$$q_{n} = b \big((a + (K_{0} + K) \| x_{n} - x^{*} \| + M \| x_{n} - x_{n-1} \|^{2}) \| x_{n} - x^{*} \| \\ \times [(K_{0} + K) \| x_{n} - x^{*} \| + M \| x_{n} - x_{n-1} \|^{2}] \\ + c ((K_{0} + K) \| x_{n} - x^{*} \| + M \| x_{n-1} - x^{*} \|^{2}) \big).$$

$$(4.3)$$

Moreover, if $U(x^*, 3r) \subseteq \Omega$, $b(K_0 + K)(a + c) < 1$ and we choose x_{-1} , $x_0 \in U(x^*, r_0) - \{x^*\}$, then sequence $\{x_n\}$ remains in $U(x^*, r_0)$ for each $n = -1, 0, 1, 2, \ldots$ so that (4.2) is satisfied, where r is the only positive root in $(0, r_0)$ of polynomial.

$$g(t) = b([a + (K_0 + K)t + 4Mt^2][K_0 + K + 4Mt]t + c(K_0 + K + 4Mt)) + p(t) - 1.$$
(4.4)

As the convergence does not exceed the quadratic one there exists $C^* > 0$ such that

$$||x_n - x_{n-1}||^2 \le C^* ||x_n - x^*||.$$
(4.5)

Then, (4.2) can be written as

$$||x_{n+1} - x^*|| \le C_1 ||x_n - x^*|| + C_2 ||x_n - x^*||^2,$$
(4.6)

for some $C_1 \ge 0$, $C_2 \ge 0$. In the case of zero residual (c = 0) the convergence order is quadratic.

Proof. We have h(0) = -1 < 0 and h(t) > 0 for sufficiently large t > 0. It then follows from the intermediate value theorem that function h has zeros in $(0, +\infty)$. Denote by r_0 the smallest such zero. Similarly, $g(0) = b(K_0 + K)(a + c) - 1 < 0$ and $g(r_0) > 0$. Denote by r the smallest zero of function g on $(0, r_0)$. Let $x, y \in U(x^*, \rho)$ for some $\rho > 0$. Then $||2y - x - x^*|| \le 2||y - x^*|| + ||x - x^*|| \le 3\rho$, so $2y - x \in U(x^*, \rho)$. Then $\rho = r_0$ or $\rho = r$ in what follows. The proof uses (2.3), (2.5) and (2.10) with adjusted estimates and (v_3) .

We have the estimates

$$||Q_n - F'(x_n)|| \le M ||x_n - x_{n-1}||^2,$$
(4.7)

$$||Q_n - [x_n, x^*; F]|| = ||([2x_n - x_{n-1}, x_{n-1}; F] - F'(x_n)) + (F'(x_n) - [x_n, x^*; F])|| \le ||[2x_n - x_{n-1}, x_{n-1}; F] - F'(x_n)|| + ||F'(x_n) - [x_n, x^*; F]|| \le M ||x_n - x_{n-1}||^2 + K ||x_n - x^*||,$$
(4.8)

$$\begin{aligned} \|Q_n - F'(x^*)\| &= \|[2x_n - x_{n-1}, x_{n-1}; F] - F'(x^*)\| \\ &\leq \|[2x_n - x_{n-1}, x_{n-1}; F] - [x_n, x^*; F]\| \\ &+ \|[x_n, x^*; F] - F'(x^*)\| \\ &\leq M \|x_n - x_{n-1}\|^2 + (K_0 + K) \|x_n - x^*\|, \end{aligned}$$
(4.9)

$$||Q_n|| \le a + (K_0 + K)||x_n - x^*|| + M||x_n - x_{n-1}||^2,$$
(4.10)

leading by (2.5) to

$$\|I - (F_*'^T F_*')^{-1} Q_n^T Q_n\| \le b[(2a + M \|x_n - x_{n-1}\|^2 + (K_0 + K) \|x_n - x^*\|) \times (M \|x_n - x_{n-1}\|^2 + (K_0 + K) \|x_n - x^*\|)] = p_n < p(r_0) = 1,$$
(4.11)

 \mathbf{SO}

$$\|(Q_n^T Q_n)^{-1} F_*'^T F_*'\| \le \frac{1}{1 - p_n}.$$
(4.12)

Moreover, by (2.10), (4.7)-(4.10) and (4.12), we obtain in turn that

$$\begin{aligned} \|x_{n+1} - x^*\| &\leq \left(b\left([a + (K_0 + K) \|x_n - x^*\| + M \|x_n - x^*\|^2 \right) \\ &\times [M \|x_n - x_{n-1}\|^2 + (K_0 + K) \|x_n - x^*\|] \\ &+ c[(K_0 + K) \|x_n - x^*\| + M \|x_n - x^*\|^2] \right) \right) / (1 - p_n) \\ &\leq \frac{q_n}{1 - p_n}, \end{aligned}$$

$$(4.13)$$

which shows (4.2). Furthermore, in view of (4.5), we obtain (4.6). \Box

Remark 4.1. The third estimate in (v_3) can be obtained if we suppose that a divided difference of order two $[\cdot, \cdot, \cdot; F]$ exists on Ω_0 such that

$$\|[u_1, u_2, u_3; F] - [u_4, u_2, u_3; F]\| \le M \|u_1 - u_4\|$$
(4.14)

for each $u_1, u_2, u_3, u_4 \in \Omega_0$. Then, using (4.14), we can have in turn that

$$\begin{aligned} \|Q_n - F'(x_n)\| &\leq \|[2x_n - x_{n-1}, x_{n-1}; F] - F'(x_n)\| \\ &= \|[2x_n - x_{n-1}, x_{n-1}; F] - [x_n, x_{n-1}; F] + [x_n, x_{n-1}; F] - F'(x_n)\| \\ &= \|([2x_n - x_{n-1}, x_{n-1}, x_n; F] - [x_n, x_{n-1}, x_n; F])(x_n - x^*)\| \\ &\leq M \|x_n - x_{n-1}\|^2, \end{aligned}$$
(4.15)

which is the third condition in (v_3) (or (4.7)). However, estimate (4.15) is weaker than (4.14) used in the literature [10, 11, 13].

5 Local convergence for the Gauss-Newton method

Let $\mathbb{R}^{i \times j}$ stand for the set of $i \times j$ matrix E, E^+ stand for the Moore-Penrose inverse of matrix E and if E has full rank (i.e., if rank $(E) = \min(i, j) = j$), then $E^+ = (E^T E)^{-1} E^T$. We shall use the standard auxiliary results [6,9]:

Lemma 5.1. Assume that $E_1, E \in \mathbb{R}^{i \times j}, E_2 = E_1 + E, ||E_1^+|| ||E|| < 1, rank(E_1) = rank(E_2), then$

$$||E_2^+|| \le \frac{||E_1^+||}{1 - ||E_1^+|| ||E||}.$$

Moreover, if rank $(E_1) = rank (E_2) = min (i, j)$, then

$$||E_2^+ - E_1^+|| \le \frac{\sqrt{2}||E_1^+||^2||E||}{1 - ||E_1^+|||E||}.$$

Lemma 5.2. Assume that $E_1, E \in \mathbb{R}^{i \times j}$ $(i \ge j), E_2 = E_1 + E, ||EE_1^+|| < 1, rank (E_1) = j$, then rank $(E_2) = j$.

Next, we present the local convergence analysis of the Gauss-Newton method in a slightly different way than in the previous sections, so we can compare the new analysis to the works of earlier studies using similar information [5,6].

Definition 5.1. We say that F' satisfies the center Lipschitz condition on Ω if there exists $L_0 \ge 0$ such that for each $x \in \Omega$, $||F'(x) - F'(x_0)|| \le L_0 ||x - x_0||$. Let $\beta > 0$. Define parameter ρ_0 by

$$\rho_0 = \frac{1}{\beta L_0}.\tag{5.1}$$

Set: $U_0 = U(x^*, \rho_0) \cap \Omega$. Notice that $U_0 \subseteq \Omega$.

Definition 5.2. We say that F' satisfies the restricted-Lipschitz condition on U_0 if there exists $L \ge 0$ such that for each $x \in U_0$

$$||F'(x) - F(x^{\theta})|| \le (1 - \theta)L||x - x^*||$$

for each $\theta \in [0, 1]$ and $x^{\theta} = x^* + \theta(x - x^*)$. Notice : $L = L(L_0)$. That is the construction of L depends on L_0 and U_0 .

In earlier studies [5, 6, 11, 13], the following definition is used.

Definition 5.3. We say that F' satisfies the Lipschitz condition on Ω if there exists $L_1 \geq 0$ such that for each $x \in \Omega$

$$||F'(x) - F'(x^{\theta})|| \le (1 - \theta)L_1||x - x^*||$$
 for each $\theta \in [0, 1]$.

Clearly, we have that

$$L_0 \le L_1,\tag{5.2}$$

$$L \le L_1, \tag{5.3}$$

hold, since $U_0 \subseteq \Omega$ and $\frac{L_1}{L_0}$ can be arbitrarily large [2–4].

In earlier studies only L_1 is used in the local convergence analysis of the Gauss-Newton method. However, in view of (5.2) and (5.3), the earlier results can be improved, if the more precise constants L_0 and L are used instead of L_1 (or L_1 and L_0). If one used with the old approach the Banach lemma on invertible operators and L_1 , then the estimate

$$\|(F'(x)^T F'(x))^{-1} F'(x)^T\| \le \frac{\beta}{1 - \beta L_1 \|x - x^*\|}$$
(5.4)

is obtained (for β to be precised later) instead of the more precise estimate really needed in the proof using L_0 :

$$\|(F'(x)^T F'(x))^{-1} F'(x)^T\| \le \frac{\beta}{1 - \beta L_0 \|x - x^*\|}.$$
(5.5)

Similarly, at the numerator of the estimates involved L, L_0 can be used instead of the less precise L_1 , L_2 , respectively leading to the advantages (A).

Next, we can show the main local convergence result for the Gauss-Newton method.

Theorem 5.3. Assume: vector x^* satisfies problem (1.1); F has a continuous derivative in Ω ; $F'(x^*)$ has full rank; F' satisfies the center-Lipschitz condition on Ω with constant L_0 ; F' satisfies the restricted-Lipschitz condition on U_0 with constant L. Then, for $x_0 \in U(x^*, \rho) - \{x^*\}$ sequence $\{x_n\}$ generated by the Gauss-Newton method is well defined, remains in $U(x^*, \rho)$ and converges to x^* , provided that

$$\sqrt{2}L_0 c\beta^2 < 1,\tag{5.6}$$

where

$$\rho = \frac{2(1 - \sqrt{2}L_0 c\beta^2)}{(2L_0 + L)\beta} < \rho_0, \tag{5.7}$$

 $c = ||F(x^*)||, \ \beta = ||[F'(x^*)^T F'(x^*)]^{-1} F'(x^*)^T||$ and

$$\lambda(L_0, L, \rho) = \lambda(\rho) = \frac{\beta L \rho}{2(1 - \beta L_0 \rho)} + \frac{\sqrt{2c\beta^2 L_0}}{1 - \beta L_0 \rho} \le 1.$$
(5.8)

Moreover, the following estimates hold

$$\|x_{n+1} - x^*\| \le \frac{\beta L}{2(1 - \beta L_0 \|x_0 - x^*\|)} \|x_n - x^*\|^2 + \frac{\sqrt{2}c\beta^2 L_0}{1 - \beta L_0 \|x_0 - x^*\|} \|x_0 - x^*\|$$
(5.9)

and

$$q_{0}(L_{0}, L, ||x_{0} - x^{*}||) = q_{0}$$

$$= \frac{\beta L ||x_{0} - x^{*}||}{2(1 - \beta L_{0} ||x_{0} - x^{*}||)} + \frac{\sqrt{2}c\beta^{2}L_{0}}{1 - \beta L_{0} ||x_{0} - x^{*}||} \in (0, 1).$$
(5.10)

Furthermore, if c = 0, then

$$||x_n - x^*|| \le q_0^{2^n - 1} ||x_0 - x^*|| \text{ for each } n = 1, 2, \dots$$
 (5.11)

Proof. We first show that $q_0 \in (0, 1)$. We have by (5.7) and (5.10) in turn that

$$q_{0} = \frac{\beta L}{2(1 - \beta L_{0} ||x_{0} - x^{*}||)} ||x_{0} - x^{*}|| + \frac{\sqrt{2}c\beta^{2}L_{0}}{1 - \beta L_{0} ||x_{0} - x^{*}||} < \frac{\beta L\rho}{2(1 - \beta L_{0}\rho)} + \frac{\sqrt{2}c\beta^{2}L_{0}}{1 - \beta L_{0}\rho} = 1$$
(5.12)

and

$$\begin{aligned} \|[F'(x^*)^T F'(x^*)]^{-1} F'(x^*)^T \| \|F'(x) - F'(x^*)\| &\leq \beta L_0 \|x - x^*\| \\ &\leq \beta L_0 \rho_0 < 1, \text{ for each } x \in U(x^*, \rho_0). \end{aligned}$$

By Lemma (2.1) and (2.2), we know $\forall x \in U(x^*, \rho), F'(x)$ has full rank and

$$\|[F'(x)^T F'(x)]^{-1} F'(x)^T\| \le \frac{\beta}{1 - \beta L_0 \|x - x^*\|} \text{ for each } x \in U(x^*, \rho_0)$$

and

$$\|[F'(x)^T F'(x)]^{-1} F'(x)^T - [F'(x^*) F'(x^*)]^{-1} F'(x^*)^T\| \le \frac{\sqrt{2\beta^2 L_0} \|x - x^*\|}{1 - \beta L_0 \|x - x^*\|}$$

for each $x \in U(x^*, \rho_0)$. Then, if $x_k \in U(x^*, \rho)$, we have by (1.2)

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - [F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T F(x_k). \\ &= [F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T [F'(x_k)(x_k - x^*) - F(x_k) + F(x^*)] \\ &+ [F'(x^*)^T F'(x^*)]^{-1} F(x^*)^T F'(x^*) - [F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T F(x^*). \end{aligned}$$

It follows that

$$\begin{aligned} \|x_{k+1} - x^*\| &= \|[F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T \| \| \int_0^1 [F'(x_k) - F'(x^* + \theta(x_k - x^*))] \\ &\times (x^* - x_k) d\theta \| + \|[F'(x^*)^T F'(x^*)]^{-1} F'(x^*)^T \\ &- [F'(x_k)^T F'(x_k)]^{-1} F'(x_k)^T \| \|F'(x^*)\| \\ &\leq \frac{\beta L \|x_k - x^*\|^2}{2(1 - \beta L_0 \|x_k - x^*\|)} + \frac{\sqrt{2}c\beta^2 L_0 \|x_k - x^*\|}{1 - \beta L_0 \|x_k - x^*\|}. \end{aligned}$$

Taking k = 0 above, we obtain $||x_1 - x^*|| \le q_0 ||x_0 - x^*|| < ||x_0 - x^*||$, so $x_1 \in U(x^*, \rho)$. Using mathematical induction, each x_k belongs to $U(x^*, \rho)$ and $||x_k - x^*||$ decreases monotonically. Hence, for each $k = 0, 1, \ldots$, we get

$$\begin{aligned} \|x_{k+1} - x^*\| &\leq \frac{\beta L \|x_k - x^*\|^3}{2(1 - \beta L_0 \|x_k - x^*\|)} + \frac{\sqrt{2}c\beta^2 L_0 \|x_k - x^*\|}{1 - \beta L_0 \|x_k - x^*\|} \\ &\leq \frac{\beta L \|x_0 - x^*\| \|x_k - x^*\|^2}{\|x_0 - x^*\|^2(1 - \beta L_0 \|x_0 - x^*\|)} + \frac{\sqrt{2}c\beta^2 L_0 \|x_k - x^*\|}{\|x_0 - x^*\|(1 - \beta L_0 \|x_0 - x^*\|)}. \end{aligned}$$

$$(5.13)$$

In particular, if c = 0, we have

$$||x_{k+1} - x^*|| \le \frac{\beta L ||x_0 - x^*||}{||x_0 - x^*||^2 (1 - \beta L_0 ||x_0 - x^*||)} ||x_k - x^*||^2$$

= $\frac{q_0}{||x_0 - x^*||} ||x_k - x^*||^2.$ (5.14)

Concerning the uniqueness of the solution.

Proposition 5.4. Suppose x^* satisfies (1.1), F has a continuous derivative in $U(x^*, \rho)$, $F'(x^*)$ has full rank and F' satisfies the center Lipschitz condition on Ω with constant L_0 . Let $\rho > 0$ satisfy

$$\rho = \mu(L_0) = \frac{2(1 - c\beta_0 L_0)}{\beta_0 L_0},\tag{5.15}$$

where c is given in Theorem 5.6 and

$$\beta_0 = \left\| [F'(x^*)^T F'(x^*)]^{-1} \right\|.$$
(5.16)

Then (1.1) has a unique solution x^* in $U(x^*, \rho)$.

Simply use L_0 instead of L_1 in the proof of Theorem 4.1 in [5].

- **Remark 5.1.** (a) The results in sections 2-4 improve the ones in [11], since K_0 , K, max $\{K_1, K_2\}$ are smaller than K^* used there.
 - (b) If $L_0 = L = L_1$, we obtain the results in [5] which in turn generalized earlier results in [1,9,11,12,14]. Moreover, if $L = L_1$, then the results reduce to the ones obtained by us in [2]-[4]. Otherwise, i.e. if

$$L_0 < L < L_1 \ (orL < L_0 < L_1)$$

(see [2–4] for example), then we obtain a larger radius of convergence, tighter error bounds on the distances $||x_n - x^*||$ and at least as precise information on the location of the solution x^* , since

$$\lambda(L_0, L) < \lambda(L_0, L_1) < \lambda(L_1, L_1),$$

$$q_0(L_0, L) < q_0(L_0, L_1) < q_0(L_1, L_1)$$

and

$$\mu(L_0) < \mu(L) < \mu(L_1).$$

These advantages are obtained under the same computational cost, since in practice the computation of the parameter L_1 requires the computation of parameters L_0 and L as special cases.

References

- S. Amat, M.A. Hernández, and N. Romero, Semilocal convergence of a sixth order iterative method for quadratic equations, *Appl. Numer. Math.*, 62, (2012), 833– 841.
- [2] I.K. Argyros and Á.A. Magreñán, Iterative Algorithms I, Nova Science Publishers Inc., New York, 2016.
- [3] I.K. Argyros and A.A. Magreñán, Iterative methods and their dynamics with applications, CRC Press, New York, 2017.
- [4] I.K. Argyros and Á.A. Magreñán, Iterative Algorithms II, Nova Science Publishers Inc., New York, 2017.
- [5] J. Chen and W. Li, Convergence of Gauss-Newton's method and uniqueness of the solution, Appl. Math. Comput., 170, (2005), 687–705.
- [6] W. M. Haübler, A Kantorovich convergence analysis for the Gauss-Newton method, *Numer. Math.*, 48, (1986), 119–125.
- [7] Á.A. Magreñán, Different anomalies in a Jarratt family of iterative root finding methods, Appl. Math. Comput., 233, (2014), 29–38.
- [8] A.A. Magreñán, A new tool to study real dynamics:, Appl. Math. Comput., 248, (2014), 29–38.
- [9] J.M. Ortega and W. C. Rheinboldt, Iterative solution of nonlinear equations in several variables, Academic Press, New York, 1970.
- [10] J.F. Potra, On an iterative algorithm of order 1.839... for solving nonlinear operator equations, Numer. Func. Anal. Opt., 7(1), (1985), 75–106.
- [11] S.M. Shankhno and O.P. Gnatyshyn, On an iterative algorithm of order 1.839... for solving the nonlinear least squares problems, *Appl. Math. Comput.*, 161, (2005), 253–264.
- [12] R. Ewing, K. Gross, and C. Martin, Newton's method estimates from data at one point, in: The merging of disciplines: New directions in Pure Applied and Computational Mathematics, Springer, New York, 1986.
- [13] G. W. Stewart, On the continuity of the generalized inverse, SIAM J. Math., 17, (1960), 33–45.
- [14] J.F. Traub and H. Wozniakowski, Convergence and complexity of Newton iteration, J. Assoc. Comput. Math., 29, (1979), 250–258.

Ioannis K. Argyros

Department of Mathematics Sciences Cameron University Lawton OK 73505 USA E-mail: iargyros@cameron.edu

Janak Raj Sharma Department of Mathematics Sant Longowal Institute of Engineering and Technology Longowal Punjab 148106 India E-mail: jrshira@yahoo.co.in

Deepak Kumar Department of Mathematics Sant Longowal Institute of Engineering and Technology Longowal Punjab 148106 India E-mail: deepak.babbi@gamil.com

Received: 14.08.2017 Accepted: 25.10.2017