Accepted Manuscript

Three-steps modified Levenberg–Marquardt method with a new line search for systems of nonlinear equations

Keyvan Amini, Faramarz Rostami

PII: S0377-0427(15)00625-1
DOI: http://dx.doi.org/10.1016/j.cam.2015.12.013
Reference: CAM 10411

To appear in: *Journal of Computational and Applied Mathematics*

Received date: 9 July 2015

Please cite this article as: K. Amini, F. Rostami, Three-steps modified Levenberg–Marquardt method with a new line search for systems of nonlinear equations, *Journal of Computational and Applied Mathematics* (2015), http://dx.doi.org/10.1016/j.cam.2015.12.013

THREE-STEPS MODIFIED LEVENBERG-MARQUARDT METHOD WITH A NEW LINE SEARCH FOR SYSTEMS OF NONLINEAR EQUATIONS

Keyvan Amini*, Faramarz Rostami

Department of Mathematics, Faculty of Science, Razi University, Kermanshah, Iran

Abstract

Three steps modified Levenberg-Marquardt method for nonlinear equations introduced by Yang [18]. This method uses the addition of the Levenberg-Marquardt (LM) step and two approximate LM steps as the trial step at every iteration. Using trust region technique, the global and biquadratic convergence of the method is proved by Yang. The main aim of this paper is to introduce a new line search strategy while investigate the convergence properties of the method with this line search technique. Since the search direction of Yang method may be not a descent direction, standard line searches can not be used directly. In this paper we propose a new nonmonotone third order Armijo type line search technique which guarantees the global convergence of this method while we use an adaptive LM parameter. It is proved that the convergence order of the new method is biquadratic. Numerical results show the new algorithm is efficient and promising.

Keywords: Nonlinear equations, Levenberg-Marquardt method, Local error bound condition, Line search, Global and biquadratic convergence.

1. Introduction

We consider the nonlinear system of equations

$$F(x) = 0, \tag{1}$$

where $F : \mathbf{R}^n \to \mathbf{R}^n$ is a continuously differentiable function and $F'(x)$ is Lipschitz continuous. Due to the nonlinearity of $F(x)$, (1) may have no solution. Throughout the paper, we let that the solution set of (1) is nonempty and denote it by X^*. There are many algorithms for solving the problem (1), for example: Gauss-Newton method, Newton's method, trust region methods, quasi-Newton methods and etc. ([3-9], [16]). The Levenberg-Marquardt (LM) is one family of classical methods for solving problem (1). [10, 12]. This family, at every iteration, computes the trial step d_k by solving the linear system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F_k, \tag{2}$$

where $F_k = F(x_k)$ and $J_k = F'(x_k)$ is the Jacobian matrix of F at x_k and λ_k is a nonnegative regularized parameter. The LM step (2) is actually a modification of the Gauss-Newton step where parameter λ_k is introduced to prevent the steps from being too large when $J_k^T J_k$ is singular or nearly singular. It is clear that for $\lambda_k = 0$, the LM method is reduced to Newton step when J_k is nonsingular. The LM method has the quadratic convergence as the Newton method when the Jacobian matrix is nonsingular and Lipschitz continuous at the solution. In [17], Yamashita and Fukushima proved the quadratic convergence for LM

*Corresponding author
Email addresses: kamini@razi.ac.ir (Keyvan Amini*), fara.rostami58@gmail.com (Faramarz Rostami)

method with $\lambda_k = \|F(x_k)\|^2$ under the local error bound condition which is weaker than the nonsingularity of the Jacobian at the solution. In [8], Fan and Yuan showed that a similar result is satisfied when $\lambda_k = \|F(x_k)\|^\delta$ with $\delta \in [1, 2]$ where numerical results show the choice $\delta = 1$ has the best results. In [6], Fan introduced a modified Levenberg-Marquardt (MLM) method with cubic convergence. The MLM method in each iteration, firstly, obtains d_{1k} by solving (2), then with setting $y_k = x_k + d_{1k}$, solves the linear system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k), \tag{3}$$

to obtains d_{2k} and sets

$$s_k = d_{1k} + d_{2k}.$$

Fan showed that with choosing

$$\lambda_k = \mu_k \|F_k\|^\delta,$$

where $\mu_k > 0$ is updated at each iteration and $\delta \in [1, 2]$, the MLM method converges cubically under the local error bound condition. It is noticeable that in the k-iteration, this method doesn't need compute $J(y_k)$ and only uses J_k in (3) and so avoids of some Jacobian computation. Inspired by this fact, to reduce the cost of Jacobian computation, Yang in [18], used another approximation step by solving the following system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(z_k), \ \textit{with}\ z_k = y_k + d_{2k}, \tag{4}$$

and proposed the trail step as follows

$$d_k = d_{1k} + d_{2k} + d_{3k}. \tag{5}$$

The globally convergence of the method was described by using of a trust region technique. Under the local error bound condition, Yang also showed the new method has biquadratic convergence. It is noticeable that d_k defined by (5) is no longer a descent direction and so it is not easy to prove the global convergence of the method with standard line searches. The main purpose of this paper is to introduce a new nonmonotone line search and show this fact that the similar convergence properties can be proved as the trust region case under local error bound condition. Note that if d_k is a descent direction of merit function $\phi(x) = \frac{1}{2}\|F(x)\|^2$ at x_k, one can choose a steplength $\alpha_k > 0$ satisfying

$$\|F(x_k + \alpha d_k)\|^2 \leq \|F_k\|^2 + 2\sigma\alpha F_k^T J_k d_k, \tag{6}$$

where $\sigma \in (0, 1)$ is a given constant, the next iterate is then determined as

$$x_{k+1} = x_k + \alpha_k d_k.$$

It is clear that this line search is suitable for descent directions while direction d_k defined by (5) may be not necessarily a descent direction of the merit function ϕ. Therefore, the standard line search techniques such as (6) can not be used directly in this case. Li and Fukushima in [11], presented a nonmonotone line search for nonlinear equations, that is,

$$\|F(x_k + \alpha d_{1k})\|^2 \leq (1 + \epsilon_k)\|F_k\|^2 - \sigma_1\alpha^2\|d_{1k}\|^2 - \sigma_2\alpha^2\|F_k\|^2, \tag{7}$$

where σ_1 and σ_2 are positive constants and the positive sequence $\{\epsilon_k\}$ satisfies

$$\sum_{k=0}^{\infty} \epsilon_k < \infty. \tag{8}$$

Since the direction of MLM method contains two parts, motivated by (7), Zhou in [19], used a new nonmonotone second order Armijo type line search as follows

$$\|F(x_k + \alpha d_{1k} + \alpha^2 d_{2k})\|^2 \leq (1 + \epsilon_k)\|F_k\|^2 - \sigma_1\alpha^2\|d_{1k}\|^2 - \sigma_2\alpha^2\|d_{2k}\|^2 - \sigma_3\alpha^2\|F_k\|^2, \tag{9}$$

where σ_1, σ_2 and σ_3 are positive constants and the positive sequence $\{\epsilon_k\}$ satisfies (8). It is straightforward to see that as $\alpha \to 0^+$, the left-hand side of (9) goes to $\|F_k\|^2$, while the right-hand side tends to the positive constant $(1+\epsilon_k)\|F_k\|^2$. Thus, (9) is satisfied for all sufficiently small $\alpha > 0$. Hence, one can obtains α_k by means of a backtracking process. This line search can avoid the necessity of descent directions to ensure that each iteration is well defined. After determination α_k satisfying in (9), Zhou set

$$x_{k+1} = x_k + \alpha_k d_{1k} + \alpha_k^2 d_{2k}, \tag{10}$$

and proved that the cubic convergence of the MLM method is preserved under the local error bound condition. The numerical experiments on the Extended Rosenbrock problem and Extended Powell Singular problem with different starting points for different n values showed that this algorithm (Algorithm 2.1 in [19]) performs well as the MLM algorithm. Inspired by (7) and (9), since the direction d_k defined by (5) contains three parts, an extension of line search (9) is as follows:

$$\|F(x_k+\alpha d_{1k}+\alpha^2 d_{2k}+\alpha^3 d_{3k})\|^2 \leq (1+\epsilon_k)\|F_k\|^2-\sigma_1\alpha^2\|d_{1k}\|^2-\sigma_2\alpha^2\|d_{2k}\|^2-\sigma_3\alpha^2\|d_{3k}\|^2-\sigma_4\alpha^2\|F_k\|^2. \tag{11}$$

This line search contains some disadvantages, for example when α_k is small then α_k^2 and α_k^3 may be very small , so $\alpha_k^2 d_{2k}$ and $\alpha_k^3 d_{3k}$ may be smaller than the machine precision which causes to lose their roles. To overcome this disadvantage, we use a new nonmonotone third order Armijo type line search:
Let σ_1 and σ_2 are positive constants, γ is a very small positive constant, and d_{1k}, d_{2k} and d_{3k} are computed by solving (2), (3) and (4), respectively. We set

$$d_k = \begin{cases} d_{1k}+d_{2k}+d_{3k} & \text{if } F_k^T J_k(d_{1k}+d_{2k}+d_{3k}) \leq -\gamma, \\ d_{1k} & \text{otherwise,} \end{cases} \tag{12}$$

then we determine a positive steplength α_k so that the following line search hold for $\alpha = \alpha_k$:

$$\|F(x_k+\alpha d_k)\|^2 \leq (1+\epsilon_k)\|F_k\|^2 - \sigma_1\alpha^2\|d_k\|^2 - \sigma_2\alpha^2\|F_k\|^2, \tag{13}$$

where the positive sequence $\{\epsilon_k\}$ satisfies (8).
In this paper we also use the following LM parameter

$$\lambda_k = \mu\|F_k\|^{\delta_k} \quad with \quad \delta_k = \begin{cases} \frac{1}{\|F_k\|} & \text{if } \|F_k\| \geq 1, \\ 1 & \text{otherwise.} \end{cases} \tag{14}$$

where μ is a positive constant. This parameter is introduced by Amini and Rostami [1], to reduce some disadvantages of the papular parameter $\lambda_k = \mu_k\|F_k\|$ that is generally used in literatures. (For example see [6, 7, 8, 18, 19]). This choice even when $\|F_k\|$ is very large, doesn't permit λ_k to be large and so prevents the LM step to be small too. [1].

The rest of the paper is organized as follows. In Section 2, we describe the new algorithm in more details. In Section 3, we establish the global convergence of the proposed algorithm under some suitable conditions. In Section 4, the biquadratic convergence of the new algorithm is proved under the local error bound condition. Finally, some numerical experiments are given in Section 5.

2. The three steps modified Levenberg-Marquardt algorithm

In this section, we present the new algorithm.

Algorithm 2.1. The modified three-steps Levenberg-Marquardt with a new line search.

Input: $x_0 \in \mathbf{R}^n$, $\mu > 0$, $\gamma > 0$, $\varepsilon > 0$, $\epsilon > 0$, $\sigma_1, \sigma_2 > 0$, $r, \rho \in (0, 1)$ and the sequence $\{\epsilon_k\}$ satisfying in (8).
Step 0 Set $k := 0$.
Step 1 Compute $F_k = F(x_k)$ and $J_k = J(x_k)$.

If $\|J_k^T F_k\| \leq \varepsilon$, stop. Otherwise compute λ_k by (14).

Step 2

a) Obtain d_{1k} by solving the following linear system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F_k, \tag{15}$$

b) Solve the linear system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k), \tag{16}$$

to obtain d_{2k}, where $y_k = x_k + d_{1k}$.

c) Solve the linear system

$$(J_k^T J_k + \lambda_k I)d = -J_k^T F(z_k), \tag{17}$$

to obtain d_{3k}, where $z_k = y_k + d_{2k}$.

d) Set $d_k = d_{1k} + d_{2k} + d_{3k}$.

Step 3 If

$$\|F(x_k + d_k)\| \leq \rho \|F_k\|, \tag{18}$$

then take $\alpha_k = 1$ and go to step 5. Otherwise go to step 4.

Step 4 Set

$$d_k = \begin{cases} d_{1k} + d_{2k} + d_{3k} & \text{if } F_k^T J_k(d_{1k} + d_{2k} + d_{3k}) \leq -\gamma, \\ d_{1k} & \text{otherwise.} \end{cases} \tag{19}$$

Compute $\alpha_k = max\{1, r^1, r^2, ...\}\ with\ \alpha = r^i\ satisfying$

$$\|F(x_k + \alpha d_k)\|^2 \leq (1 + \epsilon_k)\|F_k\|^2 - \sigma_1 \alpha^2 \|d_k\|^2 - \sigma_2 \alpha^2 \|F_k\|^2, \tag{20}$$

where the positive sequence $\{\epsilon_k\}$ satisfies (8).

Step 5 Set $x_{k+1} = x_k + \alpha_k d_k$. Set k=k+1 and goto step 1.

Remark 2.1.

i) In step 4, when $F_k^T J_k(d_{1k} + d_{2k} + d_{3k}) \leq -\gamma$, it can be resulted that $d_{1k} + d_{2k} + d_{3k}$ is a suitable direction, so it is better that algorithm uses the direction $d_k = d_{1k} + d_{2k} + d_{3k}$.

ii) As $\alpha \to 0^+$, the left-hand side of (20), goes to $\|F_k\|^2$ while the right-hand side tends to the positive value $(1 + \epsilon_k)\|F_k\|^2$, thus (20) is satisfied for all sufficiently small $\alpha > 0$. This shows that Algorithm 2.1 is well defined.

3. Global convergence

In this section, we show Algorithm 2.1 is globally convergence. Firstly, we define

$$\Omega_0 = \{x \mid \|F(x)\| \leq e^{\epsilon/2}\|F_0\|\} \tag{21}$$

where ϵ is a positive constant such that

$$\sum_{k=0}^{\infty} \epsilon_k \leq \epsilon < \infty. \tag{22}$$

To study the global convergence of Algorithm 2.1, we need the following assumptions.

Assumption 3.1 There exists a neighborhood Ω of Ω_0 such that both $F(x)$ and its Jacobian $J(x)$ are Lipschitz continuous on it, i.e., there exists a positive constant L such that

$$\|F(y) - F(x)\| \leq L\|y - x\|, \quad \forall x, y \in \Omega, \tag{23}$$

and

$$\|J(y) - J(x)\| \leq L\|y - x\|, \quad \forall x, y \in \Omega. \tag{24}$$

Due to the lipschitz continuous of F, it is clear that

$$\|J(x)\| \leq L, \quad \forall x \in \Omega. \tag{25}$$

Now we can state two following lemmas that show the sequence $\{x_k\}$, generated by Algorithm 2.1, is belong to Ω_0 and the sequence $\|F_k\|$ is convergence.

Lemma 3.1. *([4, Lemma 3.3]). Let $\{a_k\}$ and $\{r_k\}$ be positive sequences satisfying $a_{k+1} \leq (1 + r_k)a_k + r_k$ and $\sum_{k=0}^{\infty} r_k < \infty$ Then $\{a_k\}$ converges.*

Lemma 3.2. *Let the sequence $\{x_k\}$ be generated by Algorithm 2.1, then*

(a) the sequence $\|F_k\|$ converges and $x_k \in \Omega_0$ for all $k \geq 0$.

(b) the sequence $\|F_k\|$ is bounded, that is, there exists a constant $M > 0$ such that

$$\|F_k\| \leq M, \quad \forall k \geq 0. \tag{26}$$

Proof. From (18)- (20), we have

$$\|F_{k+1}\|^2 \leq (1 + \epsilon_k)\|F_k\|^2,$$

this inequality together with (22) and Lemma 3.1 imply that $\{\|F_k\|^2\}$ and so $\{\|F_k\|\}$ are convergence. Moreover, from the last inequality, we deduce that

$$\|F_{k+1}\| \leq (1 + \epsilon_k)^{1/2}\|F_k\| \leq \cdots \leq \prod_{i=0}^{k}(1 + \epsilon_i)^{1/2}\|F_0\|,$$

this inequality along with arithmetic-geometric means inequality and (22) results

$$\|F_{k+1}\| \leq (\sum_{i=0}^{k} \frac{1}{k+1}(1 + \epsilon_i))^{\frac{k+1}{2}}\|F_0\| \leq (1 + \frac{\epsilon}{k+1})^{\frac{k+1}{2}}\|F_0\| \leq e^{\frac{\epsilon}{2}}\|F_0\|,$$

which means $x_k \in \Omega_0$ for all k. The proof of (a) is completed. Part (a) together with the definition of Ω_0 implies that the sequence $\|F_k\|$ is bounded. □

Lemma 3.3. *Let the sequence $\{x_k\}$ be generated by Algorithm 2.1. If (18) holds for infinite k, then $\|F_k\|$ is convergence to zero.*

Proof. Denote the index sets

$$H_j = \{k \leq j \mid (18) holds\}, \quad G_j = \{0, 1, \cdots, j\} \setminus H_j, \quad j = 1, 2, \cdots.$$

If (18) holds for infinite k, then as $j \to \infty$, $|H_j| \to \infty$, where $|H_j|$ is the number elements of H_j. From (18)-(20), we have

$$\begin{aligned}\|F_{k+1}\| &\leq (\prod_{i \in G_k}(1 + \epsilon_i)^{1/2} \prod_{i \in H_k} \rho)\|F_0\| \\ &= (\prod_{i \in G_k}(1 + \epsilon_i)^{1/2})\rho^{|H_k|}\|F_0\| \\ &\leq e^{\frac{\epsilon}{2}}\rho^{|H_k|}\|F_0\| \to 0 \quad as\ k \to \infty.\end{aligned}$$

So, $\|F_k\| \to 0$. □

The following theorem shows that the sequence $\{x_k\}$, generated by Algorithm 2.1, converges to a stationary point of the merit function.

Theorem 3.4. *Under the conditions of Assumption 3.1, Algorithm 2.1 either terminates in finite iterations or satisfies*

$$\liminf_{k\to\infty} \|J_k^T F_k\| = 0. \tag{27}$$

Proof. By contradiction, suppose there exist $\tau > 0$ and an integer $\hat{k}$ such that

$$\|J_k^T F_k\| \geq \tau, \qquad \forall\ k > \hat{k}. \tag{28}$$

This along with (25) implies that

$$\|F_k\| \geq L^{-1}\tau \tag{29}$$

holds for sufficiently large k. So, by lemma (3.3), the inequality (18) holds only for finite k. On the other hand the relations (20), (22) and (26) result

$$\sum_{k=0}^{\infty} \alpha_k^2 \|F_k\|^2 < \infty,$$

which implies $\lim_{k\to\infty} \alpha_k \|F_k\| = 0$. This relation together with (29) yields

$$\lim_{k\to\infty} \alpha_k = 0. \tag{30}$$

If $F_k^T J_k(d_{1k} + d_{2k} + d_{3k}) \leq -\gamma$, from the line search (20), we have

$$\begin{aligned}\|F(x_k + \bar{\alpha}_k d_k)\|^2 - \|F_k\|^2 &> -\bar{\alpha}_k^2(\sigma_1\|d_k\|^2 + \sigma_2\|F_k\|^2) + \epsilon_k\|F_k\|^2 \\ &> -\bar{\alpha}_k^2(\sigma_1\|d_k\|^2 + \sigma_2\|F_k\|^2),\end{aligned}$$

while $\bar{\alpha}_k = \frac{\alpha_k}{r}$. This inequality along with (23) concludes

$$\begin{aligned}\bar{\alpha}_k^2(\sigma_1\|d_k\|^2 + \sigma_2\|F_k\|^2) &> -(\|F(x_k + \bar{\alpha}_k d_k)\|^2 - \|F_k\|^2) \\ &= -2F_k^T(F(x_k + \bar{\alpha}_k d_k) - F_k) - \|F(x_k + \bar{\alpha}_k d_k) - F_k\|^2 \\ &\geq -2F_k^T(F(x_k + \bar{\alpha}_k d_k) - F_k) - L^2\bar{\alpha}_k^2\|d_k\|^2. \end{aligned} \tag{31}$$

On the other hand, by the mean value theorem, we have

$$F_k^T(F(x_k + \bar{\alpha}_k d_k) - F_k) = \bar{\alpha}_k F_k^T J_k d_k + F_k^T \int_0^1 (J(x_k + t\bar{\alpha}_k d_k) - J_k)\bar{\alpha}_k d_k dt.$$

Substituting d_k together with (24) and (26), result

$$\begin{aligned}F_k^T(F(x_k + \bar{\alpha}_k d_k) - F_k) &\leq \bar{\alpha}_k F_k^T J_k(d_{1k} + d_{2k} + d_{3k}) + \frac{1}{2}LM\bar{\alpha}_k^2\|d_k\|^2, \\ &\leq -\bar{\alpha}_k\gamma + \frac{1}{2}LM\bar{\alpha}_k^2\|d_k\|^2. \end{aligned} \tag{32}$$

Combining (31) and (32) results

$$\bar{\alpha}_k^2(\sigma_1\|d_k\|^2 + \sigma_2\|F_k\|^2) + L^2\bar{\alpha}_k^2\|d_k\|^2 > 2\bar{\alpha}_k\gamma - LM\bar{\alpha}_k^2\|d_k\|^2,$$

or equivalently

$$\bar{\alpha}_k((\sigma_1 + LM + L^2)\|d_k\|^2 + \sigma_2\|F_k\|^2) > 2\gamma.$$

So, it is cocluded that

$$\bar{\alpha}_k > \frac{2\gamma}{(\sigma_1 + LM + L^2)\|d_k\|^2 + \sigma_2\|F_k\|^2}. \tag{33}$$

Similarly, if $F_k^T J_k(d_{1k} + d_{2k} + d_{3k}) > -\gamma$ then the relations (19) and (20), imply

$$\begin{aligned}\|F(x_k + \bar{\alpha}_k d_{1k})\|^2 - \|F_k\|^2 &> -\bar{\alpha}_k^2(\sigma_1\|d_{1k}\|^2 + \sigma_2\|F_k\|^2) + \epsilon_k\|F_k\|^2\\ &\geq -\bar{\alpha}_k^2(\sigma_1\|d_{1k}\|^2 + \sigma_2\|F_k\|^2),\end{aligned}$$

where $\bar{\alpha}_k = \frac{\alpha_k}{r}$. This relation along with (23) concludes that

$$\begin{aligned}\bar{\alpha}_k^2(\sigma_1\|d_{1k}\|^2 + \sigma_2\|F_k\|^2) &> -(\|F(x_k + \bar{\alpha}_k d_{1k})\|^2 - \|F_k\|^2)\\ &= -2F_k^T(F(x_k + \bar{\alpha}_k d_{1k}) - F_k) - \|F(x_k + \bar{\alpha}_k d_{1k}) - F_k\|^2\\ &\geq -2F_k^T(F(x_k + \bar{\alpha}_k d_{1k}) - F_k) - L^2\bar{\alpha}_k^2\|d_{1k}\|^2. \end{aligned} \tag{34}$$

On the other hand, using (24), (26) and (15) results

$$\begin{aligned}F_k^T(F(x_k + \bar{\alpha}_k d_{1k}) - F_k) &= F_k^T J_k\bar{\alpha}_k d_{1k} + F_k^T\int_0^1 (J(x_k + t\bar{\alpha}_k d_{1k}) - J_k)\bar{\alpha}_k d_{1k}dt\\ &\leq F_k^T J_k\bar{\alpha}_k d_{1k} + \frac{1}{2}LM\bar{\alpha}_k^2\|d_{1k}\|^2\\ &= -\bar{\alpha}_k d_{1k}^T(J_K^T J_k + \lambda_k I)d_{1k} + \frac{1}{2}LM\bar{\alpha}_k^2\|d_{1k}\|^2. \end{aligned} \tag{35}$$

The relations (34) and (35) conclude that

$$\bar{\alpha}_k((\sigma_1 + LM + L^2)\|d_{1k}\|^2 + \sigma_2\|F_k\|^2) > 2d_{1k}^T(J_K^T J_k + \lambda_k I)d_{1k} > 2\lambda_k\|d_{1k}\|^2,$$

while the last inequality is induced from the positive definitive of $J_k^T J_k$. So, we have

$$\bar{\alpha}_k > \frac{2\lambda_k\|d_{1k}\|^2}{(\sigma_1 + LM + L^2)\|d_{1k}\|^2 + \sigma_2\|F_k\|^2}. \tag{36}$$

Now, let $J_k = U\Sigma V^T$ be the singular value decomposition (SVD) of J_k, where U, V are two orthogonal matrixes and $\Sigma = diag(\sigma_1, \sigma_2, ..., \sigma_n)$ with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_n \geq 0$. Then

$$\|(J_k^T J_k + \lambda_k I)^{-1}\| = \|V(\Sigma^2 + \lambda_k I)^{-1}V^T\| = \|(\Sigma^2 + \lambda_k I)^{-1}\| = \max_{i\in\{1,2...,n\}}(\sigma_i^2 + \lambda_k I)^{-1} \leq \lambda_k^{-1}. \tag{37}$$

This inequality together with (14), (15), (25) and (26) implies that

$$\begin{aligned}\|d_{1k}\| &= \|(J_k^T J_k + \lambda_k I)^{-1}J_k^T F_k\| \leq \|(J_k^T J_k + \lambda_k I)^{-1}\|\|J_k\|\|F_k\|\\ &\leq L\lambda_k^{-1}\|F_k\| = \frac{L}{\mu}\|F_k\|^{1-\delta_k}\\ &\leq \frac{L}{\mu}M_1, \end{aligned} \tag{38}$$

where $M_1 = max\{1, M\}$. Now, from (15), (16), (23), (37), and (25), we have

$$\begin{aligned}\|d_{2k}\| &= \|(J_k^T J_k + \lambda_k I)^{-1}J_k^T F(y_k)\|\\ &\leq \|(J_k^T J_k + \lambda_k I)^{-1}J_k^T(F(y_k) - F_k)\| + \|(J_k^T J_k + \lambda_k I)^{-1}J_k^T F_k\|\\ &\leq L^2\lambda_k^{-1}\|d_{1k}\| + \|d_{1k}\|\end{aligned}$$

$$= (\frac{L^2}{\mu\|F_k\|^{\delta_k}} + 1)\|d_{1k}\|. \tag{39}$$

If $\|F_k\| < 1$ then (14) implies $\delta_k = 1$. This along with (29) and (39) conclude that

$$\|d_{2k}\| \leq (\frac{L^3}{\mu\tau} + 1)\|d_{1k}\|, \tag{40}$$

holds for sufficiently large k.
On the other hand if $\|F_k\| \geq 1$ then $\delta_k = \frac{1}{\|F_k\|}$, so $\|F_k\|^{\delta_k} > 1$ and by (39), we have

$$\|d_{2k}\| \leq (\frac{L^2}{\mu} + 1)\|d_{1k}\|. \tag{41}$$

The relations (40) and (41) imply that

$$\|d_{2k}\| \leq (\frac{L^2}{\mu} M_2 + 1)\|d_{1k}\| \tag{42}$$

holds for sufficiently large k, where $M_2 = max\{1, \frac{L}{\tau}\}$.
In a similar way, from (14), (17), (25), (26) and (29), it is easily seen that

$$\|d_{3k}\| \leq (\frac{L^2}{\mu} M_2 + 1)^2\|d_{1k}\| \tag{43}$$

holds for sufficiently large k.
If $\liminf_{k\to\infty} \|d_{1k}\| = 0$, then we have from (15) and (25) that

$$\liminf_{k\to\infty} \|J_k^T F_k\| = \liminf_{k\to\infty} \|(J_k^T J_k + \lambda_k I)d_{1k}\| = 0,$$

which is a contradiction to (28). Hence there exists a constant $\beta > 0$ such that

$$\liminf_{k\to\infty} \|d_{1k}\| > \beta.$$

Consequently, we can deduce from the above inequality together with (38), (42), (43), (26), (33) and (36), that $\{\alpha_k\}$ is bounded away from zero. This contradicts (30) and the proof is completed. □

4. Biquadratic convergence.

In this section, we show that the convergence rate of Algorithm 2.1 is biquadratic. First, we prove some useful lemmas then we show $\delta_k = 1$ and $\alpha_k = 1$, for sufficiently large k. By these facts we can establish the biquadratic convergence of Algorithm 2.1 using completely same arguments as [18]. We need some assumptions.

Definition 4.1. *Let N be a subset of $\mathbf{R}^n$ such that $N \cap X^* \neq \phi$, we say that $\|F(x)\|$ provides a local error bound on N for (1), if there exists a positive constant $c > 0$ such that*

$$\|F(x)\| \geq c\, dist(x, X^*), \qquad \forall x \in N, \tag{44}$$

where $dist(x, X) = \inf_{y\in X} \|y - x\|$.

In the sequel, we denote $\bar{x}_k$, the vector in X^* satisfying

$$\|\bar{x}_k - x_k\| = dist(x_k, X^*). \tag{45}$$

Assumption 4.1
(a) There exists a solution $x^* \in X^*$ of (1).
(b) $F(x)$ and $J(x)$ are Lipschitz continuous on $N(x^*, b)$, i.e., there exists a positive constant $L > 0$ such that

$$\|F(y) - F(x)\| \leq L\|y - x\|, \quad \forall x, y \in N(x^*, b), \tag{46}$$

and

$$\|J(y) - J(x)\| \leq L\|y - x\|, \quad \forall x, y \in N(x^*, b), \tag{47}$$

where $0 < b < 1$ and $N(x^*, b) = \{x \in \mathbf{R}^n \quad | \quad \|x - x^*\| \leq b\}$.
(c) $\|F(x)\|$ provides a local error bound on $N(x^*, b)$ for (1).

It is clear that the Lipschitzness of the Jacobian results

$$\|F(y) - F(x) - J(x)(y - x)\| \leq L\|y - x\|^2, \quad \forall x, y \in N(x^*, b). \tag{48}$$

Lemma 4.1. *Let Assumption 4.1 hold, then we have*

$$\mu\Gamma \leq \lambda_k \leq L\mu\|\bar{x}_k - x_k\|,$$

where $\Gamma = min\{1, c\|\bar{x}_k - x_k\|\}$.

Proof. If $\|F_k\| < 1$ then $\delta_k = 1$ and $\lambda_k = \mu\|F_k\|$, in this case the local error bound condition (44) along with (46) and $F(\bar{x}_k) = 0$, conclude that

$$c\mu\|\bar{x}_k - x_k\| \leq \lambda_k \leq L\mu\|\bar{x}_k - x_k\|. \tag{49}$$

But if $\|F_k\| \geq 1$ then we have $0 < \frac{1}{\|F_k\|} < 1$, thus

$$1 = \|F_k\|^0 \leq \|F_k\|^{\frac{1}{\|F_k\|}} \leq \|F_k\|,$$

and so

$$\mu \leq \lambda_k \leq \mu\|F_k\|.$$

Using (46) and $F(\bar{x}_k) = 0$, we get

$$\mu \leq \lambda_k \leq L\mu\|\bar{x}_k - x_k\|. \tag{50}$$

If we set $\Gamma = min\{1, c\|\bar{x}_k - x_k\|\}$ from (49) and (50) we get

$$\mu\Gamma \leq \lambda_k \leq L\mu\|\bar{x}_k - x_k\|,$$

which completes the proof. □

Lemma 4.2. *Let Assumption 4.1 hold, then there exists a positive constant* $c_1 > 0$ *such that the inequality*

$$\|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \leq c_1 + \frac{L}{\mu\Gamma}\|\bar{x}_k - x_k\|$$

holds for all sufficiently large k, *where* $\Gamma = min\{1, c\|\bar{x}_k - x_k\|\}$.

Proof. According to the result given by Behling and Iusem in [2, Theorem 1] and without loss of generality, we can assume $rank(J(\bar{x})) = r \ \ for\,all \ \bar{x} \in N(x^*, b) \bigcap X^*$. Suppose the SVD of $J(\bar{x})$ is

$$J(\bar{x}) = [\bar{U}_1, \bar{U}_2] \begin{bmatrix} \bar{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{V}_1^T \\ \bar{V}_2^T \end{bmatrix} = \bar{U}_1 \bar{\Sigma}_1 \bar{V}_1^T,$$

where $\bar{U} = [\bar{U}_1, \bar{U}_2]$ and $\bar{V} = [\bar{V}_1, \bar{V}_2]$ are two orthogonal matrixes and $\bar{\Sigma}_1 = diag(\bar{\sigma}_1, \bar{\sigma}_2, ..., \bar{\sigma}_r)$ with $\bar{\sigma}_1 \geq \bar{\sigma}_2 \geq ... \geq \bar{\sigma}_r > 0$. Correspondingly, we consider the SVD of $J(x)$ by

$$J(x) = [U_1, U_2, U_3] \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^T \\ V_2^T \\ V_3^T \end{bmatrix} = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T, \tag{51}$$

where $U = [U_1, U_2, U_3]$ and $V = [V_1, V_2, V_3]$ are two orthogonal matrixes and $\Sigma_1 = diag(\sigma_1, \sigma_2, ..., \sigma_r)$ with $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_r > 0$, and $\Sigma_2 = diag(\sigma_{r+1}, \sigma_{r+2}, ..., \sigma_{r+q})$ *with* $\sigma_{r+1} \geq \sigma_{r+2} \geq ... \geq \sigma_{r+q} > 0.$
In the following, for clearness, we also neglect the subscription k in the decomposition of $J(x_k)$ and still write J_k as same as (51)

$$J_k = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T. \tag{52}$$

So, we have

$$\begin{aligned} \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| &= \|[V_1, V_2, V_3] \begin{bmatrix} (\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 & \\ & & 0 \end{bmatrix} \begin{bmatrix} U_1^T \\ U_2^T \\ U_3^T \end{bmatrix} \| \\ &\leq \| \begin{bmatrix} (\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 & \\ & & 0 \end{bmatrix} \| \\ &\leq \|\Sigma_1^{-1}\| + \|\lambda_k^{-1}\Sigma_2\|. \end{aligned} \tag{53}$$

By the theory of matrix perturbation [15] and the Lipschitzness of J_k, we have

$$\|diag(\Sigma_1 - \bar{\Sigma}_1, \Sigma_2, 0)\| \leq \|J_k - \bar{J}_k\| \leq L\|\bar{x}_k - x_k\|,$$

which yields

$$\|\Sigma_1 - \bar{\Sigma}_1\| \leq L\|\bar{x}_k - x_k\| \quad \textit{and} \quad \|\Sigma_2\| \leq L\|\bar{x}_k - x_k\|. \tag{54}$$

Since $\{x_k\}$ converges to the solution set X^*, we assume that $L\|\bar{x}_k - x_k\| \leq \frac{\bar{\sigma}_r}{2}$ holds for all sufficiently large k. Then it follows from (54) that

$$\|\Sigma_1^{-1}\| \leq \frac{1}{\bar{\sigma}_r - L\|\bar{x}_k - x_k\|} \leq \frac{2}{\bar{\sigma}_r}, \tag{55}$$

moreover, for sufficiently large k, we have from (54) and Lemma 4.1 that

$$\|\lambda_k^{-1}\Sigma_2\| = \frac{\|\Sigma_2\|}{\mu\|F_k\|^{\delta_k}} \leq \frac{L}{\mu\Gamma}\|\bar{x}_k - x_k\|. \tag{56}$$

If we set $c_1 = \frac{2}{\bar{\sigma}_r}$ then from (53), (55) and (56), we get

$$\|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \leq c_1 + \tfrac{L}{\mu\Gamma}\|\bar{x}_k - x_k\|. \qquad \square$$

The following lemma describes an important property for the directions d_{1k}, d_{2k} and d_{3k}.

Lemma 4.3. *Under the condition of Assumption 4.1, for sufficiently large k, we have*

(*a*) $\|d_{1k}\| = o(\|\bar{x}_k - x_k\|)$,
(*b*) $\|d_{2k}\| = o(\|\bar{x}_k - x_k\|)$,
(*c*) $\|d_{3k}\| = o(\|\bar{x}_k - x_k\|)$.

Proof. The proof of (a) and (b) is similar to Lemma 3.3 in [19]. Here we proof (c). From (17), (48), (16) and Lemma 4.2, we have

$$\begin{aligned}\|d_{3k}\| &= \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(z_k)\| \\ &\le \|(J_k^T J_k+\lambda_k I)^{-1} J_k^T (F(z_k)-F(y_k)-J_k d_{2k})\|+\|(J_k^T J_k+\lambda_k I)^{-1} J_k^T F(y_k)\|+\|(J_k^T J_k+\lambda_k I)^{-1} J_k^T J_k d_{2k}\| \\ &\le L\|d_{2k}\|^2 \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| + 2\|d_{2k}\| \\ &\le L(c_1 + \frac{L}{\mu\Gamma}\|\bar{x}_k - x_k\|)\|d_{2k}\|^2 + 2\|d_{2k}\|,\end{aligned}$$

by the last inequality and (b), we get

$$\|d_{3k}\| = o(\|\bar{x}_k - x_k\|). \tag{57}$$

The proof is completed. □

Lemma 4.4. *Under the condition of Assumption 4.1, for sufficiently large k, we have*

(a) $\|U_1 U_1^T F(z_k)\| = o(\|\bar{x}_k - x_k\|)$,

(b) $\|U_2 U_2^T F(z_k)\| = o(\|\bar{x}_k - x_k\|)$,

(c) $\|U_3 U_3^T F(z_k)\| = o(\|\bar{x}_k - x_k\|)$.

Proof. The proof is similar to Lemma 3.4 in [18]. □

The following result can be found in [6].

Lemma 4.5. *Under the condition of Assumption 4.1, for sufficiently large k, we have*

(a) $\|U_2 U_2^T F_k\| = o(\|\bar{x}_k - x_k\|)$,

(b) $\|F_k + J_k d_{1k}\| = o(\|\bar{x}_k - x_k\|)$,

(c) $\|F(y_k)\| = o(\|\bar{x}_k - x_k\|)$,

(d) $\|F(y_k) + J_k d_{2k}\| = o(\|\bar{x}_k - x_k\|)$.

Lemma 4.6. *Let Assumption 4.1 hold. Then for sufficiently large k, we have*

(a) $\|F(z_k)\| = o(\|\bar{x}_k - x_k\|)$,

(b) $\|F(z_k) + J_k d_{3k}\| = o(\|\bar{x}_k - x_k\|)$.

Proof. According to the definition of z_k, (47), Lemma 4.3 and Lemma 4.5 , we have

$$\begin{aligned}\|F(z_k)\| &= \|F(y_k + d_{2k})\| \\ &\le \|F(y_k) + J(y_k)d_{2k}\| + L\|d_{2k}\|^2 \\ &\le \|F(y_k) + J_k d_{2k}\| + \|J(y_k) - J_k\|\|d_{2k}\| + L\|d_{2k}\|^2 \\ &\le \|F(y_k) + J_k d_{2k}\| + L\|d_{1k}\|\|d_{2k}\| + L\|d_{2k}\|^2 \\ &= o(\|\bar{x}_k - x_k\|).\end{aligned}$$

If we set $\varphi_k(d) = \|F(z_k) + J_k d\|^2 + \lambda_k\|d\|^2$, it is obviously from (17) that d_{3k} is the minimizer of φ_k. So we have

$$\|F(z_k) + J_k d_{3k}\| \le \sqrt{\varphi_k(d_{3k})} \le \sqrt{\varphi_k(0)} = \|F(z_k)\| = o(\|\bar{x}_k - x_k\|), \tag{58}$$

where the last equality is deduced from part (a). □

Lemma 4.7 shows that the unit step is always accepted for sufficiently large k while Lemma 4.8 shows, for sufficiently large k, $\delta_k = 1$.

Lemma 4.7. *Let Assumption 4.1 hold. Then for sufficiently large k, we have $\alpha_k = 1$.*

Proof. From (48), (58), Lemma 4.3 and Lemma 4.5, we have

$$\begin{aligned}\|F(x_k+d_{1k}+d_{2k}+d_{3k})\| &\leq \|F(x_k+d_{1k}+d_{2k}+d_{3k})-F_k-J_k(d_{1k}+d_{2k}+d_{3k})\|+\|F_k+J_k(d_{1k}+d_{2k}+d_{3k})\| \\ &\leq L\|d_{1k}+d_{2k}+d_{3k}\|^2+\|F_k+J_kd_{1k}\|+\|F(y_k)+J_kd_{2k}\|+\|F(z_k)+J_kd_{3k}\|+\|F(y_k)\|+\|F(z_k)\| \\ &= o(\|\bar{x}_k-x_k\|).\end{aligned}$$

Therefore, there exist a positive sequence $\{r_k\}$ convergence to zero such that

$$\|F(x_k+d_{1k}+d_{2k}+d_{3k})\| \leq r_k\|\bar{x}_k-x_k\| \leq c^{-1}r_k\|F_k\|, \tag{59}$$

where the last inequality uses the error bound condition.
Since $r_k \to 0$ and $c\rho$ is positive constant, there exists $k_0 \in \mathbf{N}$ such that

$$r_k < c\rho, \qquad \forall k \geq k_0. \tag{60}$$

The relations (59)-(60) imply

$$\|F(x_k+d_{1k}+d_{2k}+d_{3k})\| \leq \rho\|F_k\|, \qquad \forall k \geq k_0.$$

So (18) holds for all sufficiently large k, which means $\alpha_k = 1$ for sufficiently large k and the proof is completed. □

Lemma 4.8. *Let Assumption 4.1 hold and the sequence $\{x_k\}$ is generated by Algorithm 2.1, then the set $D = \{k \in \mathbf{N} : \ \|F(x_k)\| \geq 1\}$ is finite.*

Proof. By contradiction, suppose the set D is infinite. This along with lemma 4.7 imply that there exists $k_1 \in \mathbf{N}$ such that

$$\|F_k\| \geq 1, \quad \forall\, k \geq k_1. \tag{61}$$

By theorem 3.4, the sequence $\{x_k\}$ is converge to x_*. Because F is Lipschitz continuous, by (46), there exists a constant L so that

$$\|F(x_k)\| = \|F(x_k)-F(x_*)\| \leq L\|x_k-x_*\|, \tag{62}$$

this inequality along with (61) concludes that

$$\|x_k-x_*\| \geq \frac{1}{L}, \quad \forall\, k \geq k_1,$$

which is a contradiction to the fact that $x_k \to x_*$. This shows the assumption is incorrect and the proof is completed. □

The following theorem shows the convergence rate of Algorithm 2.1 is biquadratic.

Theorem 4.9. *Under the condition of Assumption 4.1, the convergence rate of Algorithm 2.1 is biquadratic.*

Proof. From Lemmas 4.7 and 4.8, we deduce $\delta_k = 1$ and $\alpha_k = 1$, for sufficiently large k. Thus theorem can be established using completely same arguments as [18], so we omit the proof here. □

5. Numerical experiments

In this section, we report some numerical experiments to show that Algorithm 2.1 is an effective algorithm for solving nonlinear equations and in particular, it works quite well on singular test problems. In the sense we compare Algorithm 2.1 with four well known modified Levenberg-Marquardt algorithms, we use the following abbreviations for these algorithms while the used value for parameters are also mentioned.

MLMY: Modified Levenberg Marquardt algorithm introduced by Yang [18] while

$$p_0 = 10^{-4},\ p_1 = 0.25,\ p_2 = 0.75,\ \delta = 1,\ \mu_1 = 10^{-3},\ m = 10^{-6}.$$

MLMZ: Modified Levenberg Marquardt algorithm introduced by Zhou [19] while

$$\sigma_1 = \sigma_2 = \sigma_3 = 0.005,\ \rho = 0.8,\ r = 0.5,\ \delta = 1,\ \mu = 10^{-4},\ \epsilon_k = \frac{0.5^k}{10}.$$

MLM: Modified Levenberg Marquardt algorithm introduced by Fan [6] while

$$p_0 = 10^{-4},\ p_1 = 0.25,\ p_2 = 0.75,\ \mu_1 = 10^{-3},\ m = 10^{-6},\ \delta = 1.$$

AMLM: Accelerating Modified Levenberg Marquardt algorithm introduced by Fan [7] while

$$p_0 = 10^{-4},\ p_1 = 0.25,\ p_2 = 0.75,\ \mu_1 = 10^{-3},\ m = 10^{-6},\ \ \delta = 1,\ \tilde{\alpha} = 5.$$

MLMN: Algorithm 2.1 with

$$\sigma_1 = \sigma_2 = 0.005,\ \rho = 0.8,\ r = 0.5,\ \gamma = 10^{-16},\ \mu = 10^{-4},\ \epsilon_k = \frac{0.5^k}{10}.$$

In the sequel we compare Algorithm 2.1 with MLM, AMLM, MLMY and MLMZ on some singular test problems which similar to [14], are constructed by modifying the standard test problems given in [13] by the following form

$$\hat{F}(x) = F(x) - J(x^*)A(A^TA)^{-1}A^T(x - x^*),$$

where $F(x)$ is a standard test function, $A \in \mathbf{R}^{n\times k}$ has full column rank with $1 \le k \le n$ and x^* is a solution of the equation $F(x) = 0$. Obviously

$$\hat{J}(x^*) = J(x^*)(I - A(A^TA)^{-1}A^T),$$

has rank $n - k$ and $\hat{F}(x^*) = 0$. However, $\hat{F}(x)$ may have roots that are not roots of $F(x)$. We constructed a set of singular problems while $\hat{J}(x^*)$ has rank $n - 1$, by choosing

$$A = [1, 1, ..., 1]^T \ \in \mathbf{R}^{n\times 1},$$

All codes are written in MATLAB R2009 programming environment on a personal PC with 2.5 GHz, 4 GB RAM, using Windows 7 operation system. The algorithms are terminated when the number of iterations exceeds 500 or

$$\|J_k^T F_k\| \le 10^{-5}.$$

Table 1 list the numerical results for five algorithms on the five test problems with different starting points and different n values. All test problems are run for seven starting points $-100x_0,\ -10x_0,\ -x_0,\ x_0,\ 10x_0,\ 100x_0,\ 1000x_0$, where x_0 is suggested in [13]. In Table 1, "NF" and "NJ" represent the numbers of function evaluations, Jacobian evaluations and "NS?" returns Y(yes) or N(no) while "Y" shows the corresponding method is converged to x^* and "N" shows that it is converged to another solution. Besides, the sign "$-$" indicates that the number of iterations exceeds 500 or one method failures. Note that, for general nonlinear equations, the evaluations of the Jacobian are usually n times of the function evaluations. So, we use the values "$NT = NF + NJ * n$" for comparisons of the total evaluations. From Table 1, we see that Algorithm MLMN is competitive with the other methods, in the number of failures. Furthermore, the following observations can be resulted that show the new algorithm is efficient and performs well for singular problems.

1. The MLMN algorithm is the best algorithm, in terms of the number of total evaluations, among the considered algorithms for almost 77% of the test problems while MLM, AMLM, MLMY and MLMZ solve 16%, 17%, 53% and 19% of the problems in the least number of total evaluations, respectively.

2. For Helical valley and Extended Helical valley problems, MLMN algorithm could successfully find x^* in 8 cases of 14 cases, while MLM, AMLM, MLMY and MLMZ could successfully find x^* in 7, 6, 5 and 7 cases, respectively.

Conclusions

In this paper, we proposed a new three-steps modified Levenberg-Marquardt algorithm for solving systems of nonlinear equations. This algorithm uses a new nonmonotone line search and an adaptive LM parameter. Under suitable assumptions, the proposed algorithm is shown to be globally convergent. The biquadratic convergence of the new algorithm is also obtained under the local error bound condition, which is weaker than the nonsingularity at the Jacobian. Numerical experiments demonstrated that the developed algorithm outperforms the other similar algorithms.

References

References

[1] K. Amini and F. Rostami, *A new modified two steps Levenberg-Marquardt method for nonlinear equations,* J. Comput. Appl. Math. **288** (2015), 341-350.

[2] R. Behling and A. Iusem, *The effect of calmness on the solution set of systems of nonlinear equations,* Math. Program. **137** (2013), 155-165.

[3] C. G. Broyden, *Quasi-Newton methods and their applications to function minimization.* Math. Comp. **21** (1967), 577-593.

[4] J.E. Dennis, J.J. More, *A characterization of superlinear convergence and its applications to quasi-Newton methods.* Math. Comp. **28** (1974), 549-560.

[5] J. E. Dennis and R.B. Schnable, *Numerical Methods for Unconstrained Optimization and Nonlinear Equations.* Prentice-Hall, Englewood Cliffs, NJ, 1983.

[6] J. Fan, *The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence.* Math. Comp. **81** (2012), 447-466.

[7] J. Fan, *Accelerating the modified Levenberg-Marquardt method for nonlinear equations.* Math. Comp. **83** (2014), 1173-1187.

[8] J. Fan and Y. X. Yuan, *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption.* Computing **74** (2005), 23-39.

[9] C. T. Kelley, *Iterative Methods for Optimization.* Vol. **18**, SIAM, Philadelphia, 1999.

[10] K. Levenberg, *A method for the solution of certain nonlinear problems in least squares.* Quart. Appl. Math. **2** (1944), 164-168.

[11] D. Li, M. Fukushima, *A globally and superlinearly convergent GaussNewton-based BFGS method for symmetric nonlinear equations.* SIAM J. Numer. Anal. **37** (1999) 152-172.

[12] D. W. Marquardt, *An algorithm for least-squares estimation of nonlinear parameters.* J. Soc. Indust. Appl. Math. **11** (1963), 431-441.

[13] J. J. Moré, B. S. Garbow and K. E. Hillstrom, *Testing unconstrained optimization software.* ACM Trans. Math. Softw. **7** (1981), 17-41.

[14] R. B. Schnabel and P. D. Frank, *Tensor methods for nonlinear equations.* SIAM J. Numer. Anal. **21** (1984), 815-843.

[15] G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory.* Academic Press, San Diego, CA, 1990.

[16] Ph. L. Toint, *Nonlinear step size control, trust regions and regularization for unconstrained optimization.* Optim. Methods Softw. **28** (2013), 82-95.

[17] N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method.* Computing **15** (Suppl.) (2001), 237-249.

[18] X. Yang, *A higher-order Levenberg-Marquardt method for nonlinear equations.* Appl. Math. Comput. **219** (2013), 10682-10694.

[19] W. Zhou, *On the convergence of the modified Levenberg-Marquardt method with a nonmonotone second order Armijo type line search.* J. Comput. Appl. Math. **239** (2013), 152-161.

Table 1: Numerical results for singular nonlinear equations with rank $(F'(x^*)) = n-1$

			MLMZ		MLMY		MLM		AMLM		MLMN	
Problem	n	x_0	$NF/NJ/NT$	$NS?$	$NF/NJ/NT$	$NS?$	$NF/NJ/NT$	$NS?$	$NF/NJ/NT$	$NS?$	$NF/NJ/NT$	$NS?$
Rosenbrock	2	-100	23/12/47	Y	34/12/58	Y	27/14/55	Y	27/14/55	Y	37/13/63	Y
		-10	25/13/51	Y	28/10/48	Y	23/12/47	Y	23/12/47	Y	31/11/53	Y
		-1	19/10/39	Y	28/10/48	Y	21/11/43	Y	21/11/43	Y	19/7/33	Y
		1	21/11/43	Y	25/9/43	Y	21/11/43	Y	21/11/43	Y	25/9/43	Y
		10	25/13/51	Y	31/11/53	Y	25/13/51	Y	25/13/51	Y	31/11/53	Y
		100	29/15/59	Y	37/13/63	Y	29/15/59	Y	29/15/59	Y	37/13/63	Y
		1000	37/19/75	Y	43/15/73	Y	35/18/69	Y	35/18/69	Y	43/15/73	Y
Extended Rosenbrock	500	-100	59/22/11059	Y	349/84/42349	Y	313/110/55313	Y	363/114/57363	Y	40/14/7040	Y
		-10	27/14/7027	Y	202/53/26702	Y	25/13/6525	Y	25/13/6525	Y	34/12/6034	Y
		-1	25/13/6525	Y	34/12/6034	Y	25/13/6525	Y	25/13/6525	Y	25/9/4525	Y
		1	25/13/6525	Y	31/11/5531	Y	25/13/6525	Y	25/13/6525	Y	31/11/5531	Y
		10	29/15/7529	Y	37/13/6537	Y	29/15/7529	Y	29/15/7529	Y	34/12/6034	Y
		100	35/18/9035	Y	43/15/7543	Y	37/18/9037	Y	37/18/9037	Y	40/14/7040	Y
		1000	74/31/15574	Y	76/22/11076	Y	39/20/10039	Y	39/20/10039	Y	46/16/8046	Y
Powell singular	4	-100	23/12/71	Y	28/10/68	Y	23/12/71	Y	23/12/71	Y	28/10/68	Y
		-10	19/10/59	Y	22/8/55	Y	19/10/59	Y	19/10/59	Y	22/8/55	Y
		-1	13/7/41	Y	16/6/40	Y	13/7/41	Y	13/7/41	Y	16/6/40	Y
		1	13/7/41	Y	16/6/40	Y	13/7/41	Y	13/7/41	Y	16/6/40	Y
		10	19/10/59	Y	22/8/55	Y	19/10/59	Y	19/10/59	Y	22/8/55	Y
		100	23/12/71	Y	28/10/68	Y	23/12/71	Y	23/12/71	Y	28/10/68	Y
		1000	27/14/83	Y	34/12/82	Y	27/14/83	Y	27/14/83	Y	34/12/82	Y
Extended Powell singular	500	-100	25/13/6525	Y	31/11/5531	Y	25/13/6525	Y	25/13/6525	Y	31/11/5531	Y
		-10	19/10/5019	Y	25/9/4525	Y	19/10/5019	Y	19/10/5019	Y	25/9/4525	Y
		-1	15/8/4015	Y	19/7/3519	Y	15/8/4015	Y	15/8/4015	Y	19/7/3519	Y
		1	15/8/4015	Y	19/7/3519	Y	15/8/4015	Y	15/8/4015	Y	19/7/3519	Y
		10	19/10/5019	Y	25/9/4525	Y	19/10/5019	Y	19/10/5019	Y	25/9/4525	Y
		100	25/13/6525	Y	31/11/5531	Y	25/13/6525	Y	25/13/6525	Y	31/11/5531	Y
		1000	29/15/7529	Y	37/13/6537	Y	29/15/7529	Y	29/15/7529	Y	37/13/6537	Y
Powell badly	2	-100	-		-		-		-		-	
		-10	-		-		-		-		-	
		-1	-		-		-		-		-	
		1	-		-		-		-		-	
		10	-		-		-		-		-	
		100	-		-		-		-		1895/138/2169	N
		1000	-		-		-		-		3712/269/4252	N
Extended Powell badly	500	-100	241/99/49741	N	-		-		-		-	
		-10	113/34/17113	N	-		-		-		-	
		-1	74/19/9574	N	868/195/98368	N	-		-		13/5/2513	N
		1	16/8/4016	N	523/138/69523	N	-		-		10/4/2010	N
		10	-		-		-		-		20/7/3520	N
		100	-		-		-		-		16/6/3016	N
Wood	4	-100	31/16/95	Y	40/14/96	Y	31/16/95	Y	31/16/95	Y	40/14/96	Y
		-10	27/14/83	Y	34/12/82	Y	27/14/83	Y	27/14/83	Y	34/12/82	Y
		-1	21/11/65	Y	25/9/61	Y	21/11/65	Y	21/11/65	Y	25/9/61	Y
		1	23/12/71	Y	28/10/68	Y	23/12/71	Y	23/12/71	Y	28/10/68	Y
		10	27/14/83	Y	34/12/82	Y	27/14/83	Y	27/14/83	Y	34/12/82	Y
		100	31/16/95	Y	40/14/96	Y	31/16/95	Y	31/16/95	Y	40/14/96	Y
		1000	35/18/107	Y	43/15/103	Y	35/18/107	Y	35/18/107	Y	43/15/103	Y
Extended Wood	500	-100	35/18/9035	Y	43/15/7543	Y	33/17/8533	Y	33/17/8533	Y	43/15/7543	Y
		-10	29/15/7529	Y	37/13/6537	Y	29/15/7529	Y	29/15/7529	Y	37/13/6537	Y
		-1	23/12/6023	Y	28/10/5028	Y	23/12/6023	Y	23/12/6023	Y	28/10/5028	Y
		1	25/13/6525	Y	31/11/5531	Y	25/13/6525	Y	25/13/6525	Y	31/11/5531	Y
		10	29/15/7529	Y	37/13/6537	Y	29/15/7529	Y	29/15/7529	Y	37/13/6537	Y
		100	33/17/8533	Y	40/14/7040	Y	33/17/8533	Y	33/17/8533	Y	43/15/7543	Y
		1000	-		46/16/8046	Y	37/19/9537	Y	37/19/9537	Y	46/16/8046	Y
Helical valley	3	-100	11/6/29	N	25/9/52	Y	9/5/24	N	9/5/24	N	16/6/34	Y
		-10	17/9/44	Y	28/7/49	N	23/11/56	Y	23/11/56	Y	10/4/22	Y
		-1	1/1/4	Y	1/1/4	Y	1/1/4	Y	1/1/4	Y	1/1/4	Y
		1	11/6/29	N	10/4/22	N	11/6/29	N	11/6/29	N	13/5/28	N
		10	9/5/24	N	28/10/58	Y	11/6/29	N	11/6/29	N	28/10/58	Y
		100	21/11/54	Y	25/7/46	N	17/8/41	N	13/7/34	N	16/6/34	N
		1000	11/6/29	N	16/6/34	N	29/12/65	Y	27/12/63	Y	16/6/34	N
Extended Helical valley	501	-100	29/15/7544	Y	19/7/3526	N	25/13/6538	Y	25/13/6538	Y	19/7/3526	Y
		-10	19/7/3526	N	13/5/2518	N	11/6/3017	N	11/6/3017	N	13/5/2518	Y
		-1	1/1/502	Y	1/1/502	Y	1/1/502	Y	1/1/502	Y	1/1/502	Y
		1	11/6/3017	N	16/6/3022	N	13/7/3520	N	13/7/3520	N	13/5/2518	N
		10	11/6/3017	N	13/5/2518	N	11/6/3017	N	11/6/3017	N	34/12/6046	Y
		100	25/13/6538	Y	37/12/6049	Y	25/13/6538	Y	25/13/6538	Y	19/7/3526	N
		1000	31/16/4047	Y	34/10/5044	N	27/14/7041	Y	15/8/4023	N	19/7/3526	N