MATHEMATICS OF COMPUTATION
Volume 83, Number 287, May 2014, Pages 1173–1187
S 0025-5718(2013)02752-4
Article electronically published on August 8, 2013

ACCELERATING THE MODIFIED LEVENBERG-MARQUARDT METHOD FOR NONLINEAR EQUATIONS

JINYAN FAN

Abstract. In this paper we propose an accelerated version of the modified Levenberg-Marquardt method for nonlinear equations (see Jinyan Fan, Mathematics of Computation 81 (2012), no. 277, 447–466). The original version uses the addition of the LM step and the approximate LM step as the trial step at every iteration, and achieves the cubic convergence under the local error bound condition which is weaker than nonsingularity. The notable differences of the accelerated modified LM method from the modified LM method are that we introduce the line search for the approximate LM step and extend the LM parameter to more general cases. The convergence order of the new method is shown to be a continuous function with respect to the LM parameter. We compare it with both the LM method and the modified LM method; on the benchmark problems we observe competitive performance.

1. Introduction

The Levenberg-Marquardt method is one of the most well-known methods for nonlinear equations and nonlinear least squares [9, 10]. The first theoretical study of its local convergence was carried out on the nonsingularity condition. During many years, numerical useful suggestions were developed for the sake of its better performance [11, 12, 17, 19, 20].

In this paper we consider the nonlinear equations

$$F(x) = 0, \tag{1.1}$$

where $F(x) : R^n \to R^n$ is a continuously differentiable function. At every iteration, the LM method computes the trial step

$$d_k = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k, \tag{1.2}$$

where $F_k = F(x_k)$, $J_k = F'(x_k)$ is the Jacobian and λ_k is the LM parameter updated from iteration to iteration. Under the local error bound condition, which is weaker than nonsingularity, Yamashita and Fukushima showed in [18] that the LM method has quadratic convergence if the LM parameter is chosen as $\lambda_k = \|F_k\|^2$. (Here and in what follows, $\|\cdot\|$ refers to the 2-norm.) Noticing that when the sequence is close to the solution set, λ_k may be smaller than the machine precision which makes it lose its role, while when the sequence is far away from the solution set, λ_k may be very large which makes the LM step small and hence prevents the iterates from moving fast to the solution set. Therefore, Fan and Yuan [4] took the LM parameter as $\lambda_k = \|F_k\|$, and proved that the LM method still has quadratic

Received by the editor February 6, 2012 and, in revised form, August 8, 2012.
2010 *Mathematics Subject Classification.* Primary 65K05, 90C30.
Key words and phrases. Nonlinear equations, Levenberg-Marquardt method, local error bound.
The author was supported by Chinese NSF grants 10871127 and 11171217.

convergence under the same conditions; moreover, the numerical results on the same problems have been greatly improved. Actually, we have shown in [2] that if the LM parameter is chosen as $\lambda_k = \|F_k\|^\delta$ with $\delta \in (0, 2]$, then the convergence order of the LM method will be $\min\{1 + \delta, 2\}$ under the local bound condition. Other research on the LM method can be found in [5–8].

In [3], the modified LM method was proposed. At every iteration, not only a LM step d_k but also an approximate LM step

$$\hat{d}_k = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k) \quad \text{with} \quad y_k = x_k + d_k \tag{1.3}$$

are computed. Then the trial step is set as

$$s_k = d_k + \hat{d}_k. \tag{1.4}$$

We do not calculate $J(y_k)$ and just use J_k in (1.3), which could save many Jacobian calculations and be very useful when $F(x)$ is complicated or n is very large. Choosing the LM parameter $\lambda_k = \|F(x_k)\|^\delta$ with $\delta \in [1, 2]$, we showed that the modified LM method converges cubically under the local error bound condition.

In [3], attention was mainly paid to the convergence analysis of the modified LM method for $\delta \in [1, 2]$. The question is what convergence rate of the modified LM method could be if $\delta \in (0, 1)$. For the sake of completeness, in this paper we will discuss the modified LM method with

$$\lambda_k = \|F_k\|^\delta, \quad \delta \in (0, 2]. \tag{1.5}$$

Moreover, we will employ a line search for the approximate LM step for better numerical performance, that is, the trial step will be computed by

$$s_k = d_k + \alpha_k \hat{d}_k, \tag{1.6}$$

where α_k is the step size for $\hat{d}_k$. Hence, when $\alpha_k \equiv 0$, the above method reduces to the general LM method; while $\alpha_k \equiv 1$, it reduces to the modified LM method. We will show that the convergence order of the new LM method will be $\min\{1 + 2\delta, 3\}$ which is a continuous function of δ. The new method turns out to accelerate the original modified LM method.

The paper is organized as follows: In Section 2, we first give the motivation of line search for the approximate LM step, then propose the accelerated modified LM algorithm for (1.1). We show that the new algorithm preserves the same global convergence as the existing LM algorithms under suitable conditions. In Section 3, we derive the convergence order of the accelerated modified LM algorithm under the local error bound condition. Finally, some numerical results of the new algorithm are reported in Section 4.

2. The accelerated modified Levenberg-Marquardt algorithm

2.1. **The motivation.** We take

$$\Phi(x) = \|F(x)\|^2 \tag{2.1}$$

as the merit function for (1.1).

Since the LM step d_k given by (1.2) is not only the minimizer of the convex minimization problem,

$$\min_{d \in R^n} \|F_k + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_{k,1}(d), \tag{2.2}$$

but also the solution of the trust region problem,

$$\begin{aligned} &\min_{d\in R^n} \|F_k + J_k d\|^2 \\ &s.t.\ \ \|d\| \le \Delta_{k,1} \triangleq \|d_k\|, \end{aligned} \tag{2.3}$$

we obtain from Powell's famous result given in [14] that

$$\|F_k\|^2 - \|F_k + J_k d_k\|^2 \ge \|J_k^T F_k\| \min\left\{\|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|}\right\}. \tag{2.4}$$

Similarly, the approximate LM step $\hat{d}_k$ given by (1.3) is not only the minimizer of the problem

$$\min_{d\in R^n} \|F(y_k) + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_{k,2}(d), \tag{2.5}$$

but also the solution of the trust region problem

$$\begin{aligned} &\min_{d\in R^n}\ \|F(y_k) + J_k d\|^2 \\ &s.t.\ \ \|d\| \le \Delta_{k,2} \triangleq \|\hat{d}_k\|. \end{aligned} \tag{2.6}$$

So we also have

$$\|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \ge \|J_k^T F(y_k)\| \min\left\{\|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|}\right\}. \tag{2.7}$$

As we know, the trial step s_k in the modified LM method is set as the addition of d_k and $\hat{d}_k$. That is, the step size for the approximate LM step is a unit. However, if $\hat{d}_k$ is a descent direction of the merit function Φ at y_k, then more reduction of Φ at y_k could be expected. So we may perform a line search at y_k along $\hat{d}_k$ by solving the problem

$$\min_{\alpha>0} \|F(y_k + \alpha \hat{d}_k)\|^2, \tag{2.8}$$

which could be approximated by

$$\min_{\alpha>0} \|F(y_k) + \alpha J(y_k)\hat{d}_k\|^2. \tag{2.9}$$

Since we do not calculate $J(y_k)$ in order to save the Jacobian calculation, we replace $J(y_k)$ with J_k in (2.9) and approximate it by

$$\min_{\alpha>0} \|F(y_k) + \alpha J_k \hat{d}_k\|^2. \tag{2.10}$$

The above problem is equivalent to

$$\max_{\alpha>0} \|F(y_k)\|^2 - \|F(y_k) + \alpha J_k \hat{d}_k\|^2 \triangleq \phi(\alpha), \tag{2.11}$$

where

$$\phi(\alpha) = -\hat{d}_k^T J_k^T J_k \hat{d}_k \alpha^2 + 2\hat{d}_k^T (J_k^T J_k + \lambda_k I)\hat{d}_k \alpha \tag{2.12}$$

is a quadratic function of α, and attains its maximum at

$$\tilde{\alpha}_k = \frac{\hat{d}_k^T (J_k^T J_k + \lambda_k I)\hat{d}_k}{\hat{d}_k^T J_k^T J_k \hat{d}_k} = 1 + \frac{\lambda_k \hat{d}_k^T \hat{d}_k}{\hat{d}_k^T J_k^T J_k \hat{d}_k} > 1, \tag{2.13}$$

provided that $J_k\hat{d}_k \neq 0$. Noticing that $\tilde{\alpha}_k$ is always greater than 1, it is reasonable to think it is a good idea to perform a line search at y_k along $\hat{d}_k$ which may accelerate the method. As $\|F(y_k)+\alpha J_k \hat{d}_k\|^2$ is just an approximation of $\|F(y_k)+\alpha J(y_k)\hat{d}_k\|^2$

and $\|F(y_k + \alpha\hat{d}_k)\|^2$, on the other hand, $\tilde{\alpha}_k$ may be very large if $J_k\hat{d}_k$ is close to 0. So we bound it and compute the step size α_k for $\hat{d}_k$ by solving the problem

$$\max_{\alpha\in[1,\hat{\alpha}]} \|F(y_k)\|^2 - \|F(y_k) + \alpha J_k\hat{d}_k\|^2 \triangleq \phi(\alpha), \tag{2.14}$$

where $\hat{\alpha} > 1$ is a positive constant. Then we set the trial step as

$$s_k = d_k + \alpha_k\hat{d}_k. \tag{2.15}$$

We call this new method the accelerated modified LM method.

2.2. **The algorithm.** To ensure the global convergence of the accelerated modified LM method, we use the trust region technique to justify whether s_k is a good step and could be accepted or not.

We define the actual reduction of $\Phi(x)$ at the kth iteration as

$$\text{Ared}_k = \|F_k\|^2 - \|F(x_k + d_k + \alpha_k\hat{d}_k)\|^2, \tag{2.16}$$

and the new predicted reduction as

$$\text{Pred}_k = \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + \alpha_kJ_k\hat{d}_k\|^2. \tag{2.17}$$

Since α_k is the maximizer of (2.14), we always have

$$\|F(y_k)\|^2 - \|F(y_k) + \alpha_kJ_k\hat{d}_k\|^2 \geq \|F(y_k)\|^2 - \|F(y_k) + J_k\hat{d}_k\|^2, \tag{2.18}$$

which together with (2.4) and (2.7) implies

$$\begin{aligned}\text{Pred}_k =& \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + \alpha_kJ_k\hat{d}_k\|^2\\ \geq& \|F_k\|^2 - \|F_k + J_kd_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k\hat{d}_k\|^2\\ \geq& \|J_k^TF_k\|\min\left\{\|d_k\|, \frac{\|J_k^TF_k\|}{\|J_k^TJ_k\|}\right\}\\ &+ \|J_k^TF(y_k)\|\min\left\{\|\hat{d}_k\|, \frac{\|J_k^TF(y_k)\|}{\|J_k^TJ_k\|}\right\}.\end{aligned} \tag{2.19}$$

From above, we can see that the new predicted reduction is always greater than or at least equivalent to the prediction reduction defined in the modified LM (MLM) algorithm given in [3] (Algorithm MLM). Hence a greater actual reduction could be expected.

We now present the accelerated modified LM algorithm (AMLM) as follows.

Algorithm 2.1 (Algorithm AMLM).

Step 1. *Given* $x_1 \in R^n, \mu_1 > m > 0, 0 < p_0 \leq p_1 \leq p_2 < 1, 0 < \delta \leq 2, \hat{\alpha} > 1, k := 1.$

Step 2. *If* $\|J_k^TF_k\| = 0$, *then stop. Solve*

$$(J_k^TJ_k + \lambda_kI)d = -J_k^TF_k \qquad \textit{with} \quad \lambda_k = \mu_k\|F_k\|^\delta \tag{2.20}$$

to obtain d_k, *and set*

$$y_k = x_k + d_k.$$

Solve

$$(J_k^TJ_k + \lambda_kI)d = -J_k^TF(y_k) \tag{2.21}$$

to obtain $\hat{d}_k$, *and set*

$$s_k = d_k + \alpha_k\hat{d}_k, \tag{2.22}$$

where α_k *is the step size obtained by solving* (2.14).

Step 3. *Compute* $r_k = \text{Ared}_k/\text{Pred}_k$. *Set*

$$x_{k+1} = \begin{cases} x_k + s_k, & \textit{if } r_k \geq p_0, \\ x_k, & \textit{otherwise.} \end{cases} \tag{2.23}$$

Step 4. *Choose* μ_{k+1} *as*

$$\mu_{k+1} = \begin{cases} 4\mu_k, & \textit{if } r_k < p_1, \\ \mu_k, & \textit{if } r_k \in [p_1, p_2], \\ \max\{\frac{\mu_k}{4}, m\}, & \textit{if } r_k > p_2. \end{cases} \tag{2.24}$$

Set $k = k + 1$, *and go to Step 2.*

We require μ_k to be no less than a positive constant m to prevent the steps from being too large when the sequence is near the solution.

The differences of Algorithm AMLM from Algorithm MLM are that we introduce the step size α_k for the approximate LM step $\hat{d}_k$ and extend the LM parameter to more general cases. Since the left-hand side of the linear system (2.21) is identical to that of (2.20), the computational effort of solving (2.21) could be significantly reduced.

2.3. **The global convergence.** To study the global convergence of Algorithm AMLM, we make the following assumptions.

Assumption 2.2. *$F(x)$ is continuously differentiable, and both $F(x)$ and its Jacobian $J(x)$ are Lipschitz continuous, i.e., there exist positive constants L_1 and L_2 such that*

$$\|J(y) - J(x)\| \leq L_1\|y - x\| \qquad \forall x, y \in R^n \tag{2.25}$$

and

$$\|F(y) - F(x)\| \leq L_2\|y - x\| \qquad \forall x, y \in R^n. \tag{2.26}$$

By the Lipschitzness of the Jacobian, we have

$$\|F(y) - F(x) - J(x)(y - x)\| \leq L_1\|y - x\|^2 \quad \forall x, y \in R^n. \tag{2.27}$$

The global convergence analysis of Algorithm AMLM is the same as that of Algorithm MLM, so we omit it here and just state the convergence result as below.

Theorem 2.3. *Under the conditions of Assumption* 2.2, *Algorithm AMLM terminates in finite iterations or satisfies*

$$\lim_{k\to\infty} \|J_k^T F_k\| = 0. \tag{2.28}$$

To conclude this section, we will show the difference between the new predicted reduction (2.17) and the following one:

$$\text{APred}_k = \|F_k\|^2 - \|F_k + J_k(d_k + \alpha_k\hat{d}_k)\|^2, \tag{2.29}$$

which has been used frequently as usual. By simple calculations, we have

$$\begin{aligned} &\text{APred}_k - \text{Pred}_k \\ &= \|F_k + J_k d_k\|^2 - \|F(y_k)\|^2 + \|F(y_k) + \alpha_k J_k \hat{d}_k\|^2 - \|F_k + J_k(d_k + \alpha_k \hat{d}_k)\|^2 \\ &= 2\alpha_k \hat{d}_k^T J_k^T (F(y_k) - F_k - J_k d_k), \end{aligned}$$

which means

$$\text{APred}_k = \text{Pred}_k + \|\hat{d}_k\| O(\|d_k\|^2) = O(\text{Pred}_k). \tag{2.30}$$

Hence APred_k could be used instead of Pred_k while computing for simplicity.

3. Convergence rate of the accelerated modified LM algorithm

The solution set of (1.1) denoted by X^* is always assumed to be nonempty. In addition to this, the sequence $\{x_k\}$ generated by Algorithm AMLM is assumed to converge to X^* and lie in some neighbourhood of $x^* \in X^*$.

In this section, we will first give the relationships among the LM step d_k, the approximate LM step $\hat{d}_k$, and the distance from x_k to the solution set, which are almost the same as those shown in [3, Lemma 3.2], but the proof will be more brief than that in [3]. Then we will show that the LM parameter is bounded. Finally we will derive the convergence order of Algorithm AMLM.

The local convergence theory requires the following assumptions.

Assumption 3.1. *(a) $F(x)$ is continuously differentiable, and $\|F(x)\|$ provides a local error bound on some neighbourhood of $x^* \in X^*$, i.e., there exist positive constants $c_1 > 0$ and $b_1 < 1$ such that*

$$\|F(x)\| \geq c_1 \text{dist}(x, X^*) \qquad \forall x \in N(x^*, b_1) = \{x \mid \|x - x^*\| \leq b_1\}. \tag{3.1}$$

(b) The Jacobian $J(x)$ is Lipschitz continuous on $N(x^*, b_1)$, i.e., there exists a positive constant L_1 such that

$$\|J(y) - J(x)\| \leq L_1 \|y - x\| \quad \forall x, y \in N(x^*, b_1). \tag{3.2}$$

Note that, if $J(x)$ is nonsingular at a solution of (1.1), then the solution is an isolated one, so $\|F(x)\|$ provides a local error bound on its neighborhood. However, the converse is not necessarily true; please refer to [18] for examples. Hence, the local error bound condition is weaker than nonsingularity.

Due to the Lipschitzness of the Jacobian, we have

$$\|F(y) - F(x) - J(x)(y - x)\| \leq L_1 \|y - x\|^2 \quad \forall x, y \in N(x^*, b_1). \tag{3.3}$$

Moreover, there exists a constant $L_2 > 0$ such that

$$\|F(y) - F(x)\| \leq L_2 \|y - x\| \quad \forall x, y \in N(x^*, b_1). \tag{3.4}$$

In the following, we denote by $\bar{x}_k$ the vector in X^* that satisfies

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

3.1. Properties of $d_k, \hat{d}_k$ and s_k.

In this subsection, we use the singular value decomposition (SVD) technique to investigate the properties of d_k, $\hat{d}_k$, and hence s_k.

By the result given by Behling and Iusem in [1, Theorem 1], we assume without of generality that $\text{rank} J(\bar{x}) = r$ for all $\bar{x} \in N(x^*, b_1) \cap X^*$. Suppose the SVD of $J(\bar{x}_k)$ is

$$\begin{aligned} \bar{J}_k &= \bar{U}_k \bar{\Sigma}_k \bar{V}_k^T \\ &= (\bar{U}_{k,1}, \bar{U}_{k,2}) \begin{pmatrix} \bar{\Sigma}_{k,1} & \\ & 0 \end{pmatrix} \begin{pmatrix} \bar{V}_{k,1}^T \\ \bar{V}_{k,2}^T \end{pmatrix} \\ &= \bar{U}_{k,1} \bar{\Sigma}_{k,1} \bar{V}_{k,1}^T, \end{aligned}$$

where $\bar{\Sigma}_{k,1} = \operatorname{diag}(\bar{\sigma}_{k,1}, \ldots, \bar{\sigma}_{k,r})$ with $\bar{\sigma}_{k,1} \geq \bar{\sigma}_{k,2} \geq \cdots \geq \bar{\sigma}_{k,r} > 0$. The corresponding SVD of J_k is

$$\begin{aligned} J_k &= U_k \Sigma_k V_k^T \\ &= (U_{k,1}, U_{k,2}, U_{k,3}) \begin{pmatrix} \Sigma_{k,1} & & \\ & \Sigma_{k,2} & \\ & & 0 \end{pmatrix} \begin{pmatrix} V_{k,1}^T \\ V_{k,2}^T \\ V_{k,3}^T \end{pmatrix} \\ &= U_{k,1}\Sigma_{k,1}V_{k,1}^T + U_{k,2}\Sigma_{k,2}V_{k,2}^T, \end{aligned}$$

where $\Sigma_{k,1} = \operatorname{diag}(\sigma_{k,1}, \ldots, \sigma_{k,r})$ with $\sigma_{k,1} \geq \sigma_{k,2} \geq \cdots \geq \sigma_{k,r} > 0$, and $\Sigma_{k,2} = \operatorname{diag}(\sigma_{k,r+1}, \ldots, \sigma_{k,r+q})$ with $\sigma_{k,r} \geq \sigma_{k,r+1} \geq \sigma_{k,r+2} \geq \cdots \geq \sigma_{k,r+q} > 0$. In the following, if the context is clear, we neglect the subscript k in $\Sigma_{k,i}$, $U_{k,i}, V_{k,i} (i = 1, 2, 3)$, and write J_k as

$$J_k = U_1\Sigma_1V_1^T + U_2\Sigma_2V_2^T.$$

By the theory of matrix perturbation [16] and the Lipschitzness of J_k, we know that

$$\|\operatorname{diag}(\Sigma_1 - \bar{\Sigma}_1, \Sigma_2, 0)\| \leq \|J_k - \bar{J}_k\| \leq L_1\|\bar{x}_k - x_k\|,$$

which yields

$$\|\Sigma_1 - \bar{\Sigma}_1\| \leq L_1\|\bar{x}_k - x_k\| \quad \text{and} \quad \|\Sigma_2\| \leq L_1\|\bar{x}_k - x_k\|. \tag{3.5}$$

The inequalities in (3.5) play important roles in proving the following result.

Lemma 3.2. *Under the conditions of Assumption* 3.1*, if* $x_k, y_k \in N(x^*, b_1/2)$*, then there exists a constant* $c_2 > 0$ *such that*

$$\|s_k\| \leq c_2 \operatorname{dist}(x_k, X^*) \tag{3.6}$$

holds for all sufficiently large k*.*

Proof. Since $x_k \in N(x^*, b_1/2)$, we have

$$\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq 2\|x_k - x^*\| \leq b_1,$$

hence $\bar{x}_k \in N(x^*, b_1)$. It then follows from (3.1) that the LM parameter satisfies

$$\lambda_k = \mu_k\|F_k\|^\delta \geq mc_1^\delta\|\bar{x}_k - x_k\|^\delta. \tag{3.7}$$

As d_k is a minimizer of $\varphi_{k,1}(d)$, we have from (3.3) that

$$\begin{aligned} \|d_k\|^2 &\leq \frac{\varphi_{k,1}(d_k)}{\lambda_k} \\ &\leq \frac{\varphi_{k,1}(\bar{x}_k - x_k)}{\lambda_k} \\ &= \frac{\|F_k + J_k(\bar{x}_k - x_k)\|^2}{\lambda_k} + \|\bar{x}_k - x_k\|^2 \\ &\leq L_1^2c_1^{-\delta}m^{-1}\|\bar{x}_k - x_k\|^{4-\delta} + \|\bar{x}_k - x_k\|^2, \end{aligned}$$

which together with $0 < \delta \leq 2$ gives

$$\|d_k\| \leq \tilde{c}_2\|\bar{x}_k - x_k\|, \tag{3.8}$$

where $\tilde{c}_2 = \sqrt{L_1^2c_1^{-\delta}m^{-1} + 1}$ is a positive constant.

By the definition of $\hat{d}_k$ and (3.3), we obtain

$$
\begin{aligned}
\|\hat{d}_k\| &= \| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\| \\
&\le \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\| + \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T J_k d_k\| \\
&\quad + L_1\|d_k\|^2 \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\
&\le 2\|d_k\| + L_1\|d_k\|^2 \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\|.
\end{aligned}
\tag{3.9}
$$

Now using the SVD of J_k, we have

$$
\begin{aligned}
&\|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\
&= \left\| (V_1, V_2, V_3) \begin{pmatrix} (\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 & \\ & & 0 \end{pmatrix} \begin{pmatrix} U_1^T \\ U_2^T \\ U_3^T \end{pmatrix} \right\| \\
&\le \left\| \begin{pmatrix} (\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 & \\ & & 0 \end{pmatrix} \right\| \\
&\le \|(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1\| + \|\lambda_k^{-1}\Sigma_2\|.
\end{aligned}
\tag{3.10}
$$

Since

$$
\frac{\sigma_{k,i}}{\sigma_{k,i}^2 + \lambda_k} \le \frac{\sigma_{k,i}}{2\sigma_{k,i}\sqrt{\lambda_k}} = \frac{1}{2\sqrt{\lambda_k}}
$$

for any positive $\sigma_{k,i} (i = 1, \ldots, r)$, we have from (3.7) that

$$
\|(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1\| \le \frac{1}{2c_1^{\frac{\delta}{2}} m^{\frac{1}{2}}} \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}},
\tag{3.11}
$$

which together with (3.5), (3.8), and $\delta \in (0, 2]$ implies

$$
\begin{aligned}
\|\hat{d}_k\| &\le 2\|d_k\| + L_1\|d_k\|^2 \left(\frac{1}{2c_1^{\frac{\delta}{2}} m^{\frac{1}{2}}} \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}} + \frac{L_1}{mc_1^{\delta}} \|\bar{x}_k - x_k\|^{1-\delta}\right) \\
&\le 2\|d_k\| + \check{c}_2 \|d_k\|^2 \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}}
\end{aligned}
\tag{3.12}
$$

for some positive $\check{c}_2$. It then follows from (3.8) that

$$
\|\hat{d}_k\| \le \hat{c}_2 \|\bar{x}_k - x_k\|
\tag{3.13}
$$

holds for some positive constant $\hat{c}_2$. Therefore we have

$$
\|s_k\| = \|d_k + \alpha_k \hat{d}_k\| \le c_2 \|\bar{x}_k - x_k\|,
$$

where $c_2 = \tilde{c}_2 + \hat{\alpha}\hat{c}_2$. The proof is completed. □

3.2. **Boundedness of the LM parameter.** In this subsection, we will show $\{\mu_k\}$ is bounded above, which will play a key role in estimating $\|F_k + J_k d_k\|$ and $\|F(y_k) + \alpha_k J_k \hat{d}_k\|$ in the next subsection.

Lemma 3.3. *Under the conditions of Assumption* 3.1, *if* $x_k, y_k \in N(x^*, b_1/2)$, *then there exists a positive constant* $M > m$ *such that*

$$
\mu_k \le M
\tag{3.14}
$$

holds for all sufficiently large k.

Proof. Following the result given in [3, Lemma 3.3], we have

$$\|F_k\|^2 - \|F_k + J_k d_k\|^2 \geq O(\|F_k\|) \min\{\|d_k\|, \|\bar{x}_k - x_k\|\}, \tag{3.15}$$

$$\|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \geq O(\|F(y_k)\|) \min\{\|\hat{d}_k\|, \|\bar{y}_k - y_k\|\}, \tag{3.16}$$

which together with (2.18) gives

$$\begin{aligned} \|F(y_k)\|^2 - \|F(y_k) + \alpha_k J_k \hat{d}_k\|^2 \geq& \|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \\ \geq& O(\|F(y_k)\|) \min\{\|\hat{d}_k\|, \|\bar{y}_k - y_k\|\}. \end{aligned} \tag{3.17}$$

It then follows from (3.1), (3.3) and (3.8) that

$$\begin{aligned} &|r_k - 1| \\ &= \left| \frac{\text{Ared}_k - \text{Pred}_k}{\text{Pred}_k} \right| \\ &= \left| \frac{\|F(x_k + d_k + \alpha_k \hat{d}_k)\|^2 - \|F(y_k) + \alpha_k J_k \hat{d}_k\|^2 + \|F(y_k)\|^2 - \|F_k + J_k d_k\|^2}{\text{Pred}_k} \right| \\ &\leq \frac{\|F_k + J_k s_k\| O(\|s_k\|^2 + \|d_k\|^2) + O(\|s_k\|^4 + \|d_k\|^4) + \|F_k + J_k d_k\| O(\|d_k\|^2)}{\|F_k\| \min\{\|d_k\|, \|\bar{x}_k - x_k\|\} + \|F(y_k)\| \min\{\|\hat{d}_k\|, \|\bar{y}_k - y_k\|\}} \\ &\leq \frac{\|F_k + J_k s_k\| O(\|s_k\|^2 + \|d_k\|^2) + O(\|s_k\|^4 + \|d_k\|^4) + \|F_k + J_k d_k\| O(\|d_k\|^2)}{\|\bar{x}_k - x_k\| \|d_k\|}. \end{aligned} \tag{3.18}$$

Since d_k is the minimizer of (2.3), we have from (3.4) and (3.13) that

$$\|F_k + J_k d_k\| \leq \|F_k\| \leq O(\|\bar{x}_k - x_k\|),$$

$$\|F_k + J_k s_k\| \leq \|F_k + J_k d_k\| + \alpha_k \|J_k \hat{d}_k\| \leq \|F_k\| + L_2 \hat{\alpha} \|\hat{d}_k\| \leq O(\|\bar{x}_k - x_k\|);$$

moreover, it follows from (3.12) that

$$\begin{aligned} \|s_k\|^2 =& \|d_k + \alpha_k \hat{d}_k\|^2 \leq \|d_k\|^2 + 2\alpha_k \|d_k\| \|\hat{d}_k\| + \alpha_k^2 \|\hat{d}_k\|^2 \\ \leq& \|d_k\|^2 + 2\alpha_k \|d_k\| (2\|d_k\| + \check{c}_2 \|d_k\|^2 \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}}) \\ &+ \alpha_k^2 (2\|d_k\| + \check{c}_2 \|d_k\|^2 \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}})^2 \\ \leq& (1 + 4\hat{\alpha} + 4\hat{\alpha}^2) \|d_k\|^2 + O(\|d_k\|^3 \|\bar{x}_k - x_k\|^{-\frac{\delta}{2}}) + O(\|d_k\|^4 \|\bar{x}_k - x_k\|^{-\delta}). \end{aligned}$$

The above inequalities imply that

$$r_k \to 1.$$

Hence there exists a positive constant $M > m$ such that $\mu_k \leq M$ holds for all sufficiently large k. The proof is completed. ☐

Lemma 3.3 together with (3.4) indicates that the LM parameter satisfies

$$\lambda_k = \mu_k \|F_k\|^\delta \leq L_2^\delta M \|\bar{x}_k - x_k\|^\delta, \tag{3.19}$$

so the LM parameter is bounded above.

3.3. **Convergence order of Algorithm AMLM.** Based on the results obtained in the above two subsections, we are now ready to derive the convergence order of Algorithm AMLM. In this subsection, we will show that $\{x_k\}$ converges to some solution of the nonlinear equations (1.1) with the convergence order being $\min\{1 + 2\delta, 3\}$.

By the SVD of J_k, we compute

$$d_k = -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F_k - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F_k, \tag{3.20}$$

$$\hat{d}_k = -V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^T F(y_k), \tag{3.21}$$

$$\begin{aligned} & F_k + J_k d_k \\ &= F_k - U_1\Sigma_1(\Sigma_1^2+\lambda_k I)^{-1}\Sigma_1 U_1^T F_k - U_2\Sigma_2(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 U_2^T F_k \\ &= \lambda_k U_1(\Sigma_1^2+\lambda_k I)^{-1}U_1^T F_k + \lambda_k U_2(\Sigma_2^2+\lambda_k I)^{-1}U_2^T F_k + U_3 U_3^T F_k, \end{aligned} \tag{3.22}$$

and

$$\begin{aligned} & F(y_k) + J_k\hat{d}_k \\ &= F(y_k) - U_1\Sigma_1(\Sigma_1^2+\lambda_k I)^{-1}\Sigma_1 U_1^T F(y_k) - U_2\Sigma_2(\Sigma_2^2+\lambda_k I)^{-1}\Sigma_2 U_2^T F(y_k) \\ &= \lambda_k U_1(\Sigma_1^2+\lambda_k I)^{-1}U_1^T F(y_k) + \lambda_k U_2(\Sigma_2^2+\lambda_k I)^{-1}U_2^T F(y_k) + U_3 U_3^T F(y_k). \end{aligned} \tag{3.23}$$

To analyse the convergence rate, the estimations of the above quantities are needed. The following result can be found in [3].

Lemma 3.4 ([3]). *Under the conditions of Assumption* 3.1, *if* $x_k \in N(x^*, b_1/2)$, *then we have*

(a) $\|U_1U_1^T F_k\| \le O(\|\bar{x}_k - x_k\|)$,
(b) $\|U_2U_2^T F_k\| \le O(\|\bar{x}_k - x_k\|^2)$,
(c) $\|U_3U_3^T F_k\| \le O(\|\bar{x}_k - x_k\|^2)$.

Lemma 3.5. *Under the conditions of Assumption* 3.1, *if* $x_k, y_k \in N(x^*, b_1/2)$, *then we have*

(a) $\|U_1U_1^T F(y_k)\| \le O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}})$,
(b) $\|U_2U_2^T F(y_k)\| \le O(\|x_k - \bar{x}_k\|^{\min\{2+\delta,3\}})$,
(c) $\|U_3U_3^T F(y_k)\| \le O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}})$.

Proof. It follows from Lemma 3.4, (3.19), and (3.22) that

$$\begin{aligned} \|F_k + J_k d_k\| &\le \lambda_k\|\Sigma_1^2\|^{-1}\|U_1^T F_k\| + \|U_2U_2^T F_k\| + \|U_3U_3^T F_k\| \\ &\le O(\|\bar{x}_k - x_k\|^{1+\delta}) + O(\|\bar{x}_k - x_k\|^2) \\ &= O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}), \end{aligned}$$

which together with (3.3) and (3.8) gives

$$\|F(y_k)\| = \|F(x_k + d_k)\| \le \|F_k + J_k d_k\| + L_1\|d_k\|^2 = O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}).$$

So we have

$$\|U_1U_1^T F(y_k)\| \le \|F(y_k)\| \le O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}). \tag{3.24}$$

Moreover, the local error bound condition implies that

$$\|\bar{y}_k - y_k\| \le c_1^{-1}\|F(y_k)\| \le O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}). \tag{3.25}$$

Let $\bar{p}_k = -J_k^+ F(y_k)$. Then $\bar{p}_k$ is the least squares solution of $\min \|F(y_k) + J_k p\|$. By simple calculations, we deduce from (3.2), (3.3), (3.8), and (3.25) that

$$\begin{aligned} \|U_3 U_3^T F(y_k)\| &= \|F(y_k) + J_k \bar{p}_k\| \\ &\leq \|F(y_k) + J_k(\bar{y}_k - y_k)\| \\ &\leq \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(J_k - J(y_k))(\bar{y}_k - y_k)\| \\ &\leq L_1 \|\bar{y}_k - y_k\|^2 + L_1 \|d_k\| \|\bar{y}_k - y_k\| \\ &\leq O(\|\bar{x}_k - x_k\|^{\min\{2+2\delta,4\}}) + O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}}) \\ &= O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}}). \end{aligned} \tag{3.26}$$

Let $\tilde{J}_k = U_1 \Sigma_1 V_1^T$ and $\tilde{p}_k = -\tilde{J}_k^+ F(y_k)$. Since $\tilde{p}_k$ is the least squares solution of $\min \|F(y_k) + \tilde{J}_k p\|$, we deduce from (3.2), (3.3), (3.5), (3.8), and (3.25) that

$$\begin{aligned} &\|(U_2 U_2^T + U_3 U_3^T) F(y_k)\| \\ =& \|F(y_k) + \tilde{J}_k \tilde{p}_k\| \leq \|F(y_k) + \tilde{J}_k(\bar{y}_k - y_k)\| \\ \leq& \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(\tilde{J}_k - J(y_k))(\bar{y}_k - y_k)\| \\ \leq& L_1 \|\bar{y}_k - y_k)\|^2 + \|(J_k - J(y_k) - U_2 \Sigma_2 V_2^T)(\bar{y}_k - y_k)\| \\ \leq& L_1 \|\bar{y}_k - y_k)\|^2 + \|(J_k - J(y_k))(\bar{y}_k - y_k)\| + \|U_2 \Sigma_2 V_2^T (\bar{y}_k - y_k)\| \\ \leq& L_1 \|\bar{y}_k - y_k\|^2 + L_1 \|d_k\| \|\bar{y}_k - y_k\| + L_1 \|\bar{x}_k - x_k\| \|\bar{y}_k - y_k\| \\ \leq& O(\|\bar{x}_k - x_k\|^{\min\{2+2\delta,4\}}) + O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}}) \\ =& O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}}). \end{aligned} \tag{3.27}$$

Combining (3.26) and (3.27), we know that

$$\|U_2 U_2^T F(y_k)\| \leq O(\|\bar{x}_k - x_k\|^{\min\{2+\delta,3\}}),$$

due to the orthogonality of U_2 to U_3. The proof is completed. □

Based on the lemmas above, we obtain the convergence result of Algorithm AMLM as follows.

Theorem 3.6. *Under the conditions of Assumption* 3.1, *the sequence generated by Algorithm AMLM converges to some solution of* (1.1) *with the convergence order being* $\min\{1 + 2\delta, 3\}$.

Proof. It follows from (3.5), (3.7), (3.21), and Lemma 3.5 that

$$\begin{aligned} \|\hat{d}_k\| &= \| - V_1(\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 U_2^T F(y_k)\| \\ &\leq \|\Sigma_1^{-1}\| \|U_1^T F(y_k)\| + \|\lambda_k^{-1} \Sigma_2\| \|U_2^T F(y_k)\| \\ &\leq O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}) + O(\|\bar{x}_k - x_k\|^{\min\{3,4-\delta\}}) \\ &= O(\|\bar{x}_k - x_k\|^{\min\{1+\delta,2\}}), \end{aligned} \tag{3.28}$$

which is a better estimation than (3.13). Moreover, we have from (2.18), (3.19), (3.23), and Lemma 3.5 that

(3.29)
$$\begin{aligned}
&\|F(y_k)+\alpha_k J_k\hat{d}_k\| \\
&\le \|F(y_k)+J_k\hat{d}_k\| \\
&= \|\lambda_k U_1(\Sigma_1^2+\lambda_k I)^{-1}U_1^T F(y_k)+\lambda_k U_2(\Sigma_2^2+\lambda_k I)^{-1}U_2^T F(y_k)+U_3U_3^T F(y_k)\| \\
&\le \lambda_k\|\Sigma_1^{-2}\|\|U_1^T F(y_k)\|+\|U_2^T F(y_k)\|+\|U_3^T F(y_k)\| \\
&\le O(\|\bar{x}_k-x_k\|^{\min\{1+2\delta,2+\delta\}})+O(\|\bar{x}_k-x_k\|^{\min\{2+\delta,3\}}) \\
&= O(\|\bar{x}_k-x_k\|^{\min\{1+2\delta,3\}}).
\end{aligned}$$

Now combining (3.1)–(3.3), (3.8), (3.28), and (3.29), we obtain

(3.30)
$$\begin{aligned}
&c_1\|\bar{x}_{k+1}-x_{k+1}\| \\
&\le \|F(x_{k+1})\| = \|F(y_k+\alpha_k\hat{d}_k)\| \\
&\le \|F(y_k)+\alpha_k J(y_k)\hat{d}_k\| + L_1\alpha_k^2\|\hat{d}_k\|^2 \\
&\le \|F(y_k)+\alpha_k J_k\hat{d}_k\| + \alpha_k\|(J(y_k)-J_k)\hat{d}_k\| + L_1\alpha_k^2\|\hat{d}_k\|^2 \\
&\le \|F(y_k)+\alpha_k J_k\hat{d}_k\| + L_1\hat{\alpha}\|d_k\|\|\hat{d}_k\| + L_1\hat{\alpha}^2\|\hat{d}_k\|^2 \\
&\le O(\|\bar{x}_k-x_k\|^{\min\{1+2\delta,3\}})+O(\|\bar{x}_k-x_k\|^{\min\{2+\delta,3\}})+O(\|\bar{x}_k-x_k\|^{\min\{2+2\delta,4\}}) \\
&= O(\|\bar{x}_k-x_k\|^{\min\{1+2\delta,3\}}).
\end{aligned}$$

The above inequality implies that $\{x_k\}$ converges to the solution set X^* with the convergence order being $\min\{1+2\delta,3\}$.

Note that
$$\|\bar{x}_k-x_k\| \le \|\bar{x}_{k+1}-x_{k+1}\| + \|s_k\|.$$
It follows from (3.30) that
$$\|\bar{x}_k-x_k\| \le 2\|s_k\|$$
holds for all sufficiently large k. Hence we deduce from Lemma 3.2 that

$$\|s_{k+1}\| \le O(\|s_k\|^{\min\{1+2\delta,3\}}). \tag{3.31}$$

Therefore the sequence $\{x_k\}$ converges to some solution of (1.1) with the convergence order being $\min\{1+2\delta,3\}$. □

From Theorem 3.6, we can see that for any $\delta\in(0,1)$, the convergence order of Algorithm AMLM is $1+2\delta$, while for any $\delta\in[1,2]$, the convergence order is cubic. These results extend those we obtained in [3].

4. Numerical results

We tested Algorithm AMLM on some singular problems, and compared it with both Algorithm LM ($\alpha_k\equiv 0$) and Algorithm MLM ($\alpha_k\equiv 1$).

The test problems were created by modifying the nonsingular problems given by Moré, Garbow, and Hillstrom in [13], and have the same form as in [15]:
$$\hat{F}(x) = F(x) - J(x^*)A(A^TA)^{-1}A^T(x-x^*),$$

TABLE 1. Results on the first singular test set with $\text{rank}(F'(x^*)) = n - 1$

			Algorithm LM	Algorithm MLM	Algorithm AMLM
Problem	n	x_0	NF/NJ/NT/F/CPUT	NF/NJ/NT/F/CPUT	NF/NJ/NT/F/CPUT
8	1000	1	11/11/11011/1.63e-5/1.00	15/8/8015/1.10e-5/0.79	13/7/7013/1.41e-5/0.73
9	1000	1	1/1/1001/3.60e-5/0.05	1/1/1001/3.60e-5/0.06	1/1/1001/3.60e-5/0.06
		10	9/9/9009/3.48e-3/0.81	15/8/8015/4.72e-3/0.80	15/8/8015/3.03e-3/0.81
		100	16/16/16016/1.92e-4/1.51	29/15/15029/4.48e-5/1.57	27/14/14027/8.57e-5/1.48
10	1000	1	10/10/10010/1.27e-5/23.46	15/8/8015/1.53e-5/23.92	13/7/7013/6.83e-6/17.78
		10	15/15/15015/1.06e-5/35.19	23/12/12023/7.66e-6/35.44	21/11/11021/4.37e-6/28.10
		100	12/12/12012/5.20e-012/28.24	17/9/9017/1.13e-011/26.64	15/8/8015/1.37e-7/20.45
11	1000	1	13/8/8013/2.70e-4/0.98	25/7/7025/2.73e-4/1.08	77/19/19077/2.54e-4/3.23
		10	67/42/42067/2.83e-3/5.26	87/27/27087/9.97e-4/3.90	103/29/29103/1.02e-3/4.52
		100	39/29/29039/9.87e-3/3.26	79/25/25079/6.19e-3/3.57	81/22/22081/6.19e-3/3.52
12	1000	1	30/30/30030/3.87e-7/2.85	43/22/22043/1.43e-7/2.32	43/22/22043/1.43e-7/2.34
		10	32/32/32032/2.95e-7/3.06	45/23/23045/2.46e-7/2.43	45/23/23045/2.46e-7/2.43
13	1000	1	11/11/11011/6.38e-5/1.03	15/8/8015/1.30e-4/0.80	13/7/7013/1.91e-4/0.70
		10	16/16/16016/6.17e-5/1.50	23/12/12023/4.99e-5/1.23	21/11/11021/6.22e-5/1.14
		100	19/19/19019/1.10e-4/1.79	27/14/14027/1.13e-4/1.47	25/13/13025/1.40e-4/1.36
14	1000	1	13/13/13013/1.13e-5/1.40	19/10/10019/4.45e-6/1.22	17/9/9017/2.22e-5/1.09
		10	19/19/19019/7.37e-6/2.08	27/14/14027/4.85e-6/1.75	27/14/14027/4.53e-6/1.75
		100	24/24/24024/1.48e-5/2.62	35/18/18035/4.26e-6/2.25	35/18/18035/4.23e-6/2.28

where $F(x)$ is the standard nonsingular test function, x^* is its root, and $A \in R^{n\times k}$ has full column rank with $1 \le k \le n$. Obviously, $\hat{F}(x^*) = 0$ and

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

has rank $n - k$. We created two sets of singular problems with the rank of $\hat{J}(x^*)$ being $n-1$ and $n-2$ by choosing

$$A \in R^{n\times 1}, \qquad A^T = (1, 1, \dots, 1)$$

and

$$A \in R^{n\times 2}, \qquad A^T = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots & 1 \\ 1 & -1 & 1 & -1 & \cdots & \pm 1 \end{pmatrix},$$

respectively. Meanwhile, we made a slight alteration on the variable dimension problem, which has $n+2$ equations in n unknowns; we eliminate the $(n-1)$-th and nth equations. (The first n equations in the standard problem are linear.)

We set $p_0 = 0.0001,\ p_1 = 0.25,\ p_2 = 0.75,\ m = 10^{-8},\ \mu_1 = 1, \delta = 1$ for all the tests. The algorithm is terminated when the norm of $J_k^T F_k$, e.g., the derivative of Φ at x_k, is less than 10^{-5}, or when the number of the iterations exceeds $100(n+1)$.

The results for the first set problems with rank $n-1$ are listed in Table 1, and the second set with rank $n-2$ in Table 2. The third column of the table indicates that the starting point is $x_0, 10x_0$, and $100x_0$, where x_0 is suggested by Moré, Garbow, and Hillstrom in [13]; "NF" and "NJ" represent the numbers of function calculations and Jacobian calculations, respectively. Note that, for general nonlinear equations, the calculations of the Jacobian are usually n times of the function calculations, so we also showed the values "NT=NF+NJ*n" for comparisons of the total calculations, though this kind of value does not mean much if the Jacobian is sparse. We also presented a final value of the norm of the function and the CPU time comparisons as well. If the method failed to find the solution in $100(n+1)$ iterations, we denoted it by the sign "–", and if the iterations had underflows or overflows, we denoted it by OF.

TABLE 2. Results on the first singular test set with $\text{rank}(F'(x^*)) = n - 2$

			Algorithm LM	Algorithm MLM	Algorithm AMLM
Problem	n	x_0	NF/NJ/NT/F/CPUT	NF/NJ/NT/F/CPUT	NF/NJ/NT/F/CPUT
8	1000	1	11/11/11011/1.63e-5/1.01	15/8/8015/1.10e-5/0.81	13/7/7013/1.41e-5/0.72
9	1000	1	1/1/1001/3.60e-5/0.05	1/1/1001/3.60e-5/0.05	1/1/1001/3.60e-5/0.05
		10	9/9/9009/3.48e-3/0.84	15/8/8015/4.72e-3/0.83	15/8/8015/3.03e-3/0.83
		100	17/17/17017/4.79e-5/1.62	29/15/15029/5.62e-5/1.58	27/14/14027/9.61e-5/1.48
10	1000	1	10/10/10010/1.27e-5/21.33	15/8/8015/1.53e-5/18.33	13/7/7013/6.83e-6/18.74
		10	15/15/15015/1.06e-5/32.09	23/12/12023/7.66e-6/27.77	21/11/11021/4.37e-6/29.59
		100	17/13/13017/2.72e-5/28.78	17/9/9017/2.73e-5/20.83	15/8/8015/2.73e-5/21.56
11	1000	1	13/8/8013/2.70e-4/0.98	25/7/7025/2.73e-4/1.08	75/19/19075/2.54e-4/3.21
		10	66/46/46066/2.82e-3/5.49	81/24/24081/1.01e-3/3.64	105/32/32105/1.01e-3/4.76
		100	40/31/31040/9.84e-3/3.45	85/29/29085/5.89e-3/3.99	71/21/21071/5.89e-3/3.18
12	1000	1	30/30/30030/3.87e-7/2.90	43/22/22043/1.43e-7/2.34	43/22/22043/1.43e-7/2.39
13	1000	1	11/11/11011/6.38e-5/1.03	15/8/8015/1.30e-4/0.81	13/7/7013/1.91e-4/0.72
		10	16/16/16016/6.17e-5/1.53	23/12/12023/4.99e-5/1.25	21/11/11021/6.22e-5/1.14
		100	19/19/19019/1.10e-4/1.82	27/14/14027/1.13e-4/1.48	25/13/13025/1.40e-4/1.39
14	1000	1	13/13/13013/1.13e-5/1.40	19/10/10019/4.45e-6/1.23	17/9/9017/2.22e-5/1.12
		10	19/19/19019/7.37e-6/2.15	27/14/14027/4.85e-6/1.75	27/14/14027/4.53e-6/1.79
		100	24/24/24024/1.48e-5/2.65	35/18/18035/4.26e-6/2.26	35/18/18035/4.23e-6/2.31

From the tables, we can see that Algorithm AMLM almost always outperforms Algorithm MLM or at least performs as well as Algorithm MLM. Meanwhile Algorithm MLM always outperforms Algorithm LM, whether on the first singular test set or on the second test set. Not only the function calculations and the Jacobian calculations but also the CPU time of Algorithm AMLM are less than those of Algorithm LM and Algorithm MLM. The results indicate that the line search really accelerates the method and contributes a lot to the numerical performance. That would be great helpful for the real application of the method and especially useful for the large scale problems.

ACKNOWLEDGMENTS

The author would like to thank Professor Xinwei Liu at Hebei University of Technology, China, and the referees for the useful comments and suggestions on the paper.

REFERENCES

[1] Roger Behling and Alfredo Iusem, *The effect of calmness on the solution set of systems of nonlinear equations*, Math. Program. **137** (2013), no. 1-2, Ser. A, 155–165, DOI: 10.1007/s10107-011-0486-7. MR3010423

[2] Jinyan Fan and Jianyu Pan, *Convergence properties of a self-adaptive Levenberg-Marquardt algorithm under local error bound condition*, Comput. Optim. Appl. **34** (2006), no. 1, 47–62, DOI 10.1007/s10589-005-3074-z. MR2224974 (2007d:65054)

[3] Jinyan Fan, *The modified Levenberg-Marquardt method for nonlinear equations with cubic convergence*, Math. Comp. **81** (2012), no. 277, 447–466, DOI 10.1090/S0025-5718-2011-02496-8. MR2833503 (2012e:65091)

[4] Jin-yan Fan and Ya-xiang Yuan, *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption*, Computing **74** (2005), no. 1, 23–39, DOI 10.1007/s00607-004-0083-1. MR2127319 (2006a:65079)

[5] Andreas Fischer, *Local behavior of an iterative framework for generalized equations with nonisolated solutions*, Math. Program. **94** (2002), no. 1, Ser. A, 91–124, DOI 10.1007/s10107-002-0364-4. MR1953167 (2004b:90116)

[6] Andreas Fischer and Pradyumn K. Shukla, *A Levenberg-Marquardt algorithm for unconstrained multicriteria optimization*, Oper. Res. Lett. **36** (2008), no. 5, 643–646, DOI 10.1016/j.orl.2008.02.006. MR2459519 (2009i:90119)

[7] A. Fischer, P. K. Shukla, and M. Wang, *On the inexactness level of robust Levenberg-Marquardt methods*, Optimization **59** (2010), no. 2, 273–287, DOI 10.1080/02331930801951256. MR2765461 (2011k:65074)

[8] C. T. Kelley, *Iterative methods for optimization*, Frontiers in Applied Mathematics, vol. 18, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1999. MR1678201 (2000h:90003)

[9] Kenneth Levenberg, *A method for the solution of certain nonlinear problems in least squares*, Quart. Appl. Math. **2** (1944), 164–168. MR0010666 (6,52a)

[10] Donald W. Marquardt, *An algorithm for least-squares estimation of nonlinear parameters*, J. Soc. Indust. Appl. Math. **11** (1963), 431–441. MR0153071 (27 #3040)

[11] Jorge J. Moré, *The Levenberg-Marquardt algorithm: implementation and theory*, Numerical analysis (Proc. 7th Biennial Conf., Univ. Dundee, Dundee, 1977), Springer, Berlin, 1978, pp. 105–116. Lecture Notes in Math., Vol. 630. MR0483445 (58 #3446)

[12] J. J. Moré, *Recent developments in algorithms and software for trust region methods*, Mathematical programming: the state of the art (Bonn, 1982), Springer, Berlin, 1983, pp. 258–287. MR717404 (85b:90066)

[13] Jorge J. Moré, Burton S. Garbow, and Kenneth E. Hillstrom, *Testing unconstrained optimization software*, ACM Trans. Math. Software **7** (1981), no. 1, 17–41, DOI 10.1145/355934.355936. MR607350 (83b:90144)

[14] M. J. D. Powell, *Convergence properties of a class of minimization algorithms*, Nonlinear programming, 2 (Proc. Sympos. Special Interest Group on Math. Programming, Univ. Wisconsin, Madison, Wis., 1974), Academic Press, New York, 1974, pp. 1–27. MR0386270 (52 #7128)

[15] Robert B. Schnabel and Paul D. Frank, *Tensor methods for nonlinear equations*, SIAM J. Numer. Anal. **21** (1984), no. 5, 815–843, DOI 10.1137/0721054. MR760620 (86d:65066)

[16] G. W. Stewart and Ji Guang Sun, *Matrix perturbation theory*, Computer Science and Scientific Computing, Academic Press Inc., Boston, MA, 1990. MR1061154 (92a:65017)

[17] Wenyu Sun and Ya-Xiang Yuan, *Optimization theory and methods: Nonlinear programming*, Springer Optimization and Its Applications, vol. 1, Springer, New York, 2006. MR2232297 (2007c:90002)

[18] N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method*, Topics in numerical analysis, Comput. Suppl., vol. 15, Springer, Vienna, 2001, pp. 239–249, DOI 10.1007/978-3-7091-6217-0_18. MR1874516 (2002k:65088)

[19] Ya-xiang Yuan, *Trust region algorithms for nonlinear equations*, Information **1** (1998), no. 1, 7–20. MR1645695 (2000d:65090)

[20] Ya-Xiang Yuan, *Recent advances in numerical methods for nonlinear equations and nonlinear least squares*, Numer. Algebra Control Optim. **1** (2011), no. 1, 15–34, DOI 10.3934/naco.2011.1.15. MR2806290 (2012e:65112)

Department of Mathematics, and MOE-LSC, Shanghai Jiao Tong University, Shanghai 200240, People's Republic of China

E-mail address: jyfan@sjtu.edu.cn