

## THE MODIFIED LEVENBERG-MARQUARDT METHOD FOR NONLINEAR EQUATIONS WITH CUBIC CONVERGENCE

JINYAN FAN

ABSTRACT. We propose a modified Levenberg-Marquardt method for nonlinear equations, in which not only a LM step but also an approximate LM step are computed at every iteration. To ensure the global convergence of the new method, a new kind of predicted reduction is introduced for the merit function when using the trust region technique. The cubic convergence of the modified LM method is proved under the local error bound condition which is weaker than nonsingularity. Numerical results show that the new method is very efficient and could save many calculations of the Jacobian.

### 1. INTRODUCTION

We consider the system of nonlinear equations

$$(1.1) \quad F(x) = 0,$$

where  $F(x) : R^n \rightarrow R^n$  is continuously differentiable. Due to the nonlinearity of  $F(x)$ , (1.1) may have no solutions. Throughout the paper, we assume that the solution set of (1.1) denoted by  $X^*$  is nonempty, and in all cases  $\|\cdot\|$  refers to the 2-norm.

The Newton method is one of the best known methods for nonlinear equations. At every iteration, it computes the trial step

$$(1.2) \quad d_k^N = -J_k^{-1}F_k,$$

where  $F_k = F(x_k)$  and  $J_k = F'(x_k)$  is the Jacobian. If  $J(x)$  is Lipschitz continuous and nonsingular at the solution, then the convergence of the Newton method is quadratic [4]. However, when  $F(x)$  is complicated or when  $n$  is very large, the computation of the Jacobian may be expensive. There are mainly two ways to save the Jacobian calculations. One is to use an approximation of the Jacobian which needs less computation. For example, Quasi-Newton methods are popular methods for nonlinear equations and usually have superlinear convergence under some suitable conditions. The other way is to make more use of the available Jacobian  $J_k$ . For example, the following well-known modified Newton method computes not only the Newton step

$$(1.3) \quad d_k^N = -J_k^{-1}F_k,$$

---

Received by the editor April 19, 2010 and, in revised form, September 20, 2010.

2010 *Mathematics Subject Classification*. Primary 65K05, 90C30.

*Key words and phrases*. Nonlinear equations, Levenberg-Marquardt method, local error bound, cubic convergence.

The author was supported by Chinese NSF grants 10871127, 10701056 and the Chenxing Program of SJTU.

but also the approximate Newton step

$$(1.4) \quad d_k^{MN} = -J_k^{-1}F(y_k) \quad \text{with } y_k = x_k + d_k^N$$

at every iteration. The available  $J_k$  is used in (1.4) rather than  $J(y_k)$ , in other words, which saves the Jacobian calculation. The convergence of the modified Newton method described above is cubic if  $J(x)$  is Lipschitz continuous and nonsingular at the solution.

Though the Newton method and the modified Newton method enjoy the favorable property of fast convergence,  $d_k^N$  and  $d_k^{MN}$  may be very large and not stable when  $J_k$  is near singular. Moreover, when  $J_k$  is singular,  $d_k^N$  and  $d_k^{MN}$  are not defined. To overcome the difficulties caused by the possible singularity of  $J_k$ , the Levenberg-Marquardt method [5, 6] computes the trial step by

$$(1.5) \quad d_k^{LM} = -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k,$$

where the LM parameter  $\lambda_k \geq 0$  is updated from iteration to iteration. Obviously, when  $J_k$  is nonsingular and  $\lambda_k = 0$ , then the LM step  $d_k^{LM}$  is reduced to the Newton step  $d_k^N$ . The LM method has the same quadratic convergence as the Newton method if the Jacobian is Lipschitz continuous and nonsingular at the solution.

However, the condition on the nonsingularity of the Jacobian is too strong. Recently, under the local error bound condition which is weaker than nonsingularity, Yamashita and Fukushima [14] showed that if the LM parameter is chosen as  $\lambda_k = \|F_k\|^2$ , then the LM method preserves the quadratic convergence under the local error bound condition. Fan and Yuan [3] chose  $\lambda_k = \|F_k\|^\delta$  with  $\delta \in [1, 2]$  and proved that the LM method still achieves the quadratic convergence under the same conditions. More general choices of the LM parameter have been given in [2].

Inspired by the modified Newton method, we present the modified Levenberg-Marquardt method in this paper. At every iteration, the modified LM method first solves the linear equations

$$(1.6) \quad (J_k^T J_k + \lambda_k I)d = -J_k^T F_k \quad \text{with } \lambda_k = \mu_k \|F_k\|^\delta, \delta \in [1, 2]$$

to obtain the LM step  $d_k$ , where  $\mu_k > 0$  is updated from iteration to iteration, then solves the linear equations

$$(1.7) \quad (J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k) \quad \text{with } y_k = x_k + d_k$$

to obtain the approximate LM step  $\hat{d}_k$ . We use  $(J_k^T J_k + \mu_k \|F_k\|^\delta I)$  instead of  $(J(y_k)^T J(y_k) + \mu_{k+1} \|F(y_k)\|^\delta I)$  in (1.7), which does not involve the calculation of  $J(y_k)$ . Since the decomposition of  $J_k^T J_k + \lambda_k I$  is available after solving (1.6), only the function value  $F(y_k)$  is needed to solve (1.7). So the cost of obtaining  $\hat{d}_k$  will be inexpensive.

We could use line search or the trust region technique to obtain the global convergence of the modified LM method. Since  $d_k$  is a decent direction of the merit function  $\|F(x)\|^2$  at  $x_k$  but  $d_k + \hat{d}_k$  may not be, we prefer to use the trust region technique. It is worth pointing out that the usual predicted reduction  $\|F_k\|^2 - \|F_k + J_k d_k\|^2$  is always nonnegative [1], however,  $\|F_k\|^2 - \|F_k + J_k(d_k + \hat{d}_k)\|^2$  may not have such a desirable property. To overcome this difficulty, we introduce a new kind of predicted reduction, which is proven to be always nonnegative and of crucial importance for the global convergence of the modified LM method.

As the modified Newton method has cubic convergence, it is interesting to investigate whether the modified LM method could also achieve the cubic convergence under the local error bound condition which is weaker than nonsingularity. Thus, we need to explore not only the LM step  $d_k$  but also the approximate LM step  $\hat{d}_k$ . While  $d_k$  preserves the attractive properties as presented in [3],  $\hat{d}_k$  depends on the properties of  $J_k^T F(y_k)$  which contribute to the cubic convergence of the modified LM method.

The paper is organized as follows: In Section 2, we first introduce the new definition of the predicted reduction for the merit function, then propose a modified LM algorithm for (1.1) by using the trust region technique. We show that the new algorithm preserves the same global convergence as the existing LM algorithms under some suitable conditions. In Section 3, both  $d_k$  and  $\hat{d}_k$  are deeply investigated. We prove that the modified LM algorithm has the same cubic convergence as the modified Newton method under the local error bound condition, which is weaker than nonsingularity. Finally, some numerical results of the new algorithm are reported in Section 4.

2. THE MODIFIED LM ALGORITHM AND ITS GLOBAL CONVERGENCE

2.1. **New definition of the predicted reduction and the algorithm.** We take

$$(2.1) \quad \Phi(x) = \|F(x)\|^2$$

as the merit function for (1.1). As described in Section 1, the modified LM method not only computes the LM step  $d_k$  by (1.6) but also computes the approximate LM step  $\hat{d}_k$  by (1.7) at every iteration.

Define the actual reduction of  $\Phi(x)$  at the  $k$ -th iteration as

$$Ared_k = \|F_k\|^2 - \|F(x_k + d_k + \hat{d}_k)\|^2.$$

We cannot define  $\|F_k\|^2 - \|F_k + J_k(d_k + \hat{d}_k)\|^2$  as the predicted reduction as usual, because it cannot be proven to be nonnegative, which is required for the global convergence in the trust region method. We need to develop a new kind of predicted reduction.

Note that the LM step  $d_k$  is the minimizer of the convex minimization problem:

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

If we let

$$(2.3) \quad \Delta_{k,1} = \|d_k\| = \|-(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\|,$$

then it can be verified that  $d_k$  is also a solution of the following trust region problem:

$$(2.4) \quad \min_{d \in R^n} \|F_k + J_k d\|^2 \quad \text{s.t.} \quad \|d\| \leq \Delta_{k,1}.$$

By the famous result given by Powell in [10], we know that

$$(2.5) \quad \|F_k\|^2 - \|F_k + J_k d_k\|^2 \geq \|J_k^T F_k\| \min \left\{ \|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\}.$$

Similarly to  $d_k$ ,  $\hat{d}_k$  is not only the minimizer of the problem

$$(2.6) \quad \min_{d \in R^n} \|F(y_k) + J_k d\|^2 + \lambda_k \|d\|^2 \triangleq \varphi_{k,2}(d),$$

but also the solution of the trust region problem

$$(2.7) \quad \min_{d \in R^n} \|F(y_k) + J_k d\|^2 \quad \text{s.t.} \quad \|d\| \leq \Delta_{k,2},$$

where

$$(2.8) \quad \Delta_{k,2} = \|\hat{d}_k\| = \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\|.$$

So we also have

$$(2.9) \quad \|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \geq \|J_k^T F(y_k)\| \min \left\{ \|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|} \right\}.$$

Now based on the inequalities (2.5) and (2.9), it is reasonable for us to define the new predicted reduction as

$$(2.10) \quad Pred_k = \|F_k\|^2 - \|F_k + J_k d_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2,$$

which satisfies

$$(2.11) \quad \begin{aligned} Pred_k \geq & \|J_k^T F_k\| \min \left\{ \|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\} \\ & + \|J_k^T F(y_k)\| \min \left\{ \|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|} \right\}, \end{aligned}$$

and it is always nonnegative. The ratio of the actual reduction to the predicted reduction

$$r_k = \frac{Ared_k}{Pred_k}$$

will be used in deciding whether to accept the trial step and how to adjust the LM parameter.

We present the modified LM algorithm as follows.

**Algorithm 2.1** (Modified LM algorithm for nonlinear equations).

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

Step 2. If  $\|J_k^T F_k\| = 0$ , then stop. Solve

$$(2.12) \quad (J_k^T J_k + \lambda_k I)d = -J_k^T F_k \quad \text{with} \quad \lambda_k = \mu_k \|F_k\|^\delta$$

to obtain  $d_k$  and set

$$y_k = x_k + d_k.$$

Solve

$$(2.13) \quad (J_k^T J_k + \lambda_k I)d = -J_k^T F(y_k)$$

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

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

Step 3. Compute  $r_k = Ared_k/Pred_k$ . Set

$$(2.14) \quad x_{k+1} = \begin{cases} x_k + s_k, & \text{if } r_k \geq p_0, \\ x_k, & \text{otherwise.} \end{cases}$$

Step 4. Choose  $\mu_{k+1}$  as

$$(2.15) \quad \mu_{k+1} = \begin{cases} 4\mu_k, & \text{if } r_k < p_1, \\ \mu_k, & \text{if } r_k \in [p_1, p_2], \\ \max\{\frac{\mu_k}{4}, m\}, & \text{if } r_k > p_2. \end{cases}$$

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

In fact, our modified LM algorithm can be regarded as a trust region algorithm in some sense. As we know, the trust region algorithm updates the trust region directly at every iteration, while our modified LM algorithm updates the parameter  $\mu_k$ , which in turn updates the values  $\Delta_{k,1}$  and  $\Delta_{k,2}$  by (2.3) and (2.8) implicitly. Many other papers also discuss the Levenberg-Marquardt method and the trust region method; please see [7, 8, 13, 15, 16] and the references therein for more details.

The main difference of Algorithm 2.1 from the general LM algorithm [1] is that an approximate LM step  $\hat{d}_k$  is computed at every iteration.  $\hat{d}_k$  may not be as good as  $-(J(y_k)^T J(y_k) + \mu_{k+1} \|F(y_k)\|^\delta I)^{-1} J(y_k)^T F(y_k)$ , but it does not need to calculate  $J(y_k)$ . Another difference is that a new kind of predicted reduction is used to guarantee the actual reduction to be always nonnegative.

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

**2.2. Global convergence of Algorithm 2.1.** To study the global convergence of Algorithm 2.1, 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

$$(2.16) \quad \|J(y) - J(x)\| \leq L_1 \|y - x\|, \quad \forall x, y \in R^n$$

and

$$(2.17) \quad \|F(y) - F(x)\| \leq L_2 \|y - x\|, \quad \forall x, y \in R^n.$$

By the Lipschitzness of the Jacobian, we have

$$(2.18) \quad \|F(y) - F(x) - J(x)(y - x)\| \leq L_1 \|y - x\|^2, \quad \forall x, y \in R^n.$$

Next we show that the sequence generated by Algorithm 2.1 converges to the stationary point of the merit function.

**Theorem 2.3.** *Under the conditions of Assumption 2.2, Algorithm 2.1 terminates in finite iterations or satisfies*

$$(2.19) \quad \lim_{k \rightarrow \infty} \|J_k^T F_k\| = 0.$$

*Proof.* We prove by contradiction. Suppose the theorem is not true, then there exist a positive  $\tau$  and infinitely many  $k$  such that

$$(2.20) \quad \|J_k^T F_k\| \geq \tau.$$

Let  $T_1, T_2$  be the sets of the indices as follows:

$$T_1 = \{k \mid \|J_k^T F_k\| \geq \tau\},$$

$$T_2 = \{k \mid \|J_k^T F_k\| \geq \frac{\tau}{2} \text{ and } x_{k+1} \neq x_k\}.$$

Then  $T_1$  is an infinite set. In the following, we will derive the contradictions whether  $T_2$  is finite or infinite.

Case (I).  $T_2$  is finite.

It follows from the definition of  $T_2$  that the set

$$T_3 = \{k \mid \|J_k^T F_k\| \geq \tau \text{ and } x_{k+1} \neq x_k\}$$

is also finite. Let  $\tilde{k}$  be the largest index of  $T_3$ . Then we know that  $x_{k+1} = x_k$  holds for all  $k \in \{k > \tilde{k} \mid k \in T_1\}$ . Define the indices set

$$T_4 = \{k > \tilde{k} \mid \|J_k^T F_k\| \geq \tau \text{ and } x_{k+1} = x_k\}.$$

Suppose  $k \in T_4$ . It is easy to see that  $\|J_{k+1}^T F_{k+1}\| \geq \tau$ . Moreover, we have  $x_{k+2} = x_{k+1}$ . Otherwise, if  $x_{k+2} \neq x_{k+1}$ , then  $k+1 \in T_3$ , which contradicts the fact that  $\tilde{k}$  is the largest index of  $T_3$ . Hence, we have  $k+1 \in T_4$ . By induction, we know that  $\|J_k^T F_k\| \geq \tau$  and  $x_{k+1} = x_k$  hold for all  $k > \tilde{k}$ .

It now follows from Step 3 of Algorithm 2.1 that  $r_k < p_0$  for all  $k > \tilde{k}$ , which implies

$$(2.21) \quad \mu_k \rightarrow +\infty \quad \text{and} \quad \lambda_k \rightarrow +\infty$$

due to (2.12), (2.15) and  $x_{k+1} = x_k$  for all  $k > \tilde{k}$ . Hence we have

$$(2.22) \quad d_k \rightarrow 0.$$

Moreover, it follows from (2.7), (2.18), (2.21) and the definition of  $\hat{d}_k$  that

$$(2.23) \quad \begin{aligned} \|\hat{d}_k\| &= \left\| -(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k) \right\| \\ &\leq \left\| (J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k \right\| + \left\| (J_k^T J_k + \lambda_k I)^{-1} J_k^T J_k d_k \right\| \\ &\quad + L_1 \|d_k\|^2 \left\| (J_k^T J_k + \lambda_k I)^{-1} J_k^T \right\| \\ &\leq \|d_k\| + \|d_k\| + \frac{L_1 L_2}{\lambda_k} \|d_k\|^2 \\ &\leq \bar{c} \|d_k\| \end{aligned}$$

holds for all sufficiently large  $k$ , where  $\bar{c}$  is a positive constant. Therefore, we have

$$(2.24) \quad \|s_k\| = \|d_k + \hat{d}_k\| \leq (1 + \bar{c}) \|d_k\|.$$

Furthermore, it follows from (2.11), (2.17), (2.20) and (2.24) that

$$(2.25) \quad \begin{aligned} |r_k - 1| &= \left| \frac{Ared_k - Pred_k}{Pred_k} \right| \\ &\leq \frac{\left| \|F(x_k + d_k + \hat{d}_k)\|^2 - \|F_k + J_k d_k\|^2 + \|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \right|}{\left\| J_k^T F_k \right\| \min \left\{ \|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\} + \|J_k^T F(y_k)\| \min \left\{ \|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|} \right\}} \\ &\leq \frac{\|F_k + J_k s_k\| O(\|s_k\|^2 + \|d_k\|^2) + O(\|s_k\|^4 + \|d_k\|^4) + \|F_k + J_k d_k\| O(\|d_k\|^2)}{\left\| J_k^T F_k \right\| \min \left\{ \|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\}} \\ &\leq \frac{\|F_k\| O(\|d_k\|^2) + \|J_k \hat{d}_k\| O(\|d_k\|^2)}{\|d_k\|} \\ &\rightarrow 0, \end{aligned}$$

which implies that  $r_k \rightarrow 1$ . In view of the updating rule of  $\mu_k$ , we know that there exists a positive constant  $\tilde{m} > m$  such that  $\mu_k < \tilde{m}$  holds for all sufficiently large  $k$ , which is a contradiction to (2.21). Hence the supposition (2.20) cannot be true while  $T_2$  is finite.

Case (II).  $T_2$  is infinite.

It follows from (2.11) and (2.17) that

$$\begin{aligned}
 \|F_1\|^2 &\geq \sum_{k \in T_2} (\|F_k\|^2 - \|F_{k+1}\|^2) \\
 &\geq \sum_{k \in T_2} p_0 \text{Pred}_k \\
 (2.26) \quad &\geq \sum_{k \in T_2} p_0 \left\{ \|J_k^T F_k\| \min \left\{ \|d_k\|, \frac{\|J_k^T F_k\|}{\|J_k^T J_k\|} \right\} \right. \\
 &\quad \left. + \|J_k^T F(y_k)\| \min \left\{ \|\hat{d}_k\|, \frac{\|J_k^T F(y_k)\|}{\|J_k^T J_k\|} \right\} \right\} \\
 &\geq \sum_{k \in T_2} \frac{p_0 \tau}{2} \min \left\{ \|d_k\|, \frac{\tau}{2L_2^2} \right\},
 \end{aligned}$$

which implies

$$(2.27) \quad \lim_{k \rightarrow \infty, k \in T_2} d_k = 0.$$

Then the definition of  $d_k$  gives

$$(2.28) \quad \lambda_k \rightarrow +\infty, \quad k \in T_2.$$

Similar to (2.23), there exists a positive  $\tilde{c}$  such that

$$(2.29) \quad \|\hat{d}_k\| \leq \tilde{c} \|d_k\|$$

holds for all sufficiently large  $k \in T_2$ . Thus we have

$$(2.30) \quad \|s_k\| = \|d_k + \hat{d}_k\| \leq (1 + \tilde{c}) \|d_k\|.$$

Hence, we obtain from (2.26) that

$$(2.31) \quad \sum_{k \in T_2} \|s_k\| = \sum_{k \in T_2} \|d_k + \hat{d}_k\| < +\infty.$$

Furthermore, it follows from (2.16) and (2.17) that

$$\sum_{k \in T_2} \left| \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| \right| < +\infty.$$

Since (2.20) holds for infinitely many  $k$ , there exists a large  $\hat{k}$  such that  $\|J_{\hat{k}}^T F_{\hat{k}}\| \geq \tau$  and

$$\sum_{k \in T_2, k \geq \hat{k}} \left| \|J_k^T F_k\| - \|J_{k+1}^T F_{k+1}\| \right| < \frac{\tau}{2}.$$

By induction, we see that  $\|J_k^T F_k\| \geq \tau/2$  holds for all  $k \geq \hat{k}$ . Then we deduce from (2.26)–(2.29) that  $\lim_{k \rightarrow \infty} x_k$  exists and

$$(2.32) \quad d_k \rightarrow 0, \quad \hat{d}_k \rightarrow 0.$$

It follows from the definition of  $d_k$  that

$$(2.33) \quad \mu_k \rightarrow +\infty.$$

Next, by the same analysis as (2.25) we know that

$$r_k \rightarrow 1.$$

Thus, there exists a positive constant  $\bar{m} > m$  such that  $\mu_k < \bar{m}$  holds for all sufficiently large  $k$ , which is a contradiction to (2.33). Therefore, the supposition (2.20) cannot be true when  $T_2$  is infinite.

Summarizing the analyses above, we know that (2.19) holds true.  $\square$

Actually, if we let

$$(2.34) \quad \tilde{Pred}_k = \|F_k\|^2 - \|F_k + J_k(d_k + \hat{d}_k)\|^2$$

as usual, we have

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

which means

$$(2.35) \quad \tilde{Pred}_k = Pred_k + \|\hat{d}_k\|O(\|d_k\|^2) = O(Pred_k).$$

So  $\tilde{Pred}_k$  can be used instead of  $Pred_k$  while computing.

### 3. CUBIC CONVERGENCE OF THE MODIFIED LM ALGORITHM

We assume that the sequence  $\{x_k\}$  generated by Algorithm 2.1 converges to the solution set  $X^*$  of (1.1) and lies in some neighbourhood of  $x^* \in X^*$ . We first give the relations among the LM step  $d_k$ , the approximate step  $\hat{d}_k$  and the distance from the current iterate  $x_k$  to the solution set. Then we show that the LM parameter is bounded above which is very important to the derivation of the cubic convergence of the algorithm.

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

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

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

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

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, for examples please see [14]. Hence, the local error bound condition is weaker than nonsingularity.

Due to the Lipschitzness of the Jacobian, we have

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

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

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



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

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

**3.1. Properties of the trial step  $s_k$ .** In this subsection, we use the singular value decomposition technique to investigate the properties of  $d_k$ ,  $\hat{d}_k$ , and hence  $s_k$ .

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  $\bar{\Sigma}_{k,1} = \text{diag}(\bar{\sigma}_{k,1}, \dots, \bar{\sigma}_{k,r})$  with  $\bar{\sigma}_{k,1} \geq \bar{\sigma}_{k,2} \geq \dots \geq \bar{\sigma}_{k,r} > 0$ , and correspondingly the 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  $\Sigma_{k,1} = \text{diag}(\sigma_{k,1}, \dots, \sigma_{k,r})$  with  $\sigma_{k,1} \geq \sigma_{k,2} \geq \dots \geq \sigma_{k,r} > 0$ , and  $\Sigma_{k,2} = \text{diag}(\sigma_{k,r+1}, \dots, \sigma_{k,r+q})$  with  $\sigma_{k,r} \geq \sigma_{k,r+1} \geq \sigma_{k,r+2} \geq \dots \geq \sigma_{k,r+q} > 0$ . In the following, if the context is clear, we neglect the subscription  $k$  in  $\Sigma_{k,i}$  and  $U_{k,i}$ ,  $V_{k,i}$  ( $i = 1, 2, 3$ ), and write  $J_k$  as

$$J_k = U_1 \Sigma_1 V_1^T + U_2 \Sigma_2 V_2^T.$$

The following lemma gives the relationship between the trial step  $s_k$  and the distance from  $x_k$  to the solution set.

**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*

$$(3.5) \quad \|s_k\| \leq c_2 \text{dist}(x_k, X^*)$$

holds for all sufficiently large  $k$ .

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

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

which implies that  $\bar{x}_k \in N(x^*, b_1)$ . It follows from (2.15) and (3.1) that the LM parameter  $\lambda_k$  satisfies

$$(3.6) \quad \lambda_k = \mu_k \|F_k\|^\delta \geq m c_1^\delta \|\bar{x}_k - x_k\|^\delta.$$

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

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

which together with  $1 \leq \delta \leq 2$  gives

$$(3.7) \quad \|d_k\| \leq \tilde{c}_2 \|\bar{x}_k - x_k\|,$$

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

It follows from (3.3) and the definition of  $\hat{d}_k$  that

$$\begin{aligned} \|\hat{d}_k\| &= \|-(J_k^T J_k + \lambda_k I)^{-1} J_k^T F(y_k)\| \\ (3.8) \quad &\leq \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T F_k\| + \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T J_k d_k\| \\ &\quad + L_1 \|d_k\|^2 \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\ &\leq 2\|d_k\| + L_1 \|d_k\|^2 \|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\|. \end{aligned}$$

Using the SVD of  $J_k$ , we have

$$\begin{aligned} &\|(J_k^T J_k + \lambda_k I)^{-1} J_k^T\| \\ (3.9) \quad &= \left\| (V_1, V_2, V_3) \begin{pmatrix} (\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 & \\ & & 0 \end{pmatrix} \begin{pmatrix} U_1^T \\ U_2^T \\ U_3^T \end{pmatrix} \right\| \\ &\leq \left\| \begin{pmatrix} (\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 & & \\ & (\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 & \\ & & 0 \end{pmatrix} \right\| \\ &\leq \left\| \begin{pmatrix} \Sigma_1^{-1} & \\ & \lambda_k^{-1} \Sigma_2 \end{pmatrix} \right\|. \end{aligned}$$

By the theory of matrix perturbation [12] and the Lipschitzness of  $J_k$ , we have

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

which yields

$$(3.10) \quad \|\Sigma_1 - \bar{\Sigma}_1\| \leq L_1 \|\bar{x}_k - x_k\| \quad \text{and} \quad \|\Sigma_2\| \leq L_1 \|\bar{x}_k - x_k\|.$$

Since  $\{x_k\}$  converges to the solution set  $X^*$ , we assume that  $L_1 \|\bar{x}_k - x_k\| \leq \bar{\sigma}_r/2$  holds for all sufficiently large  $k$ . Then it follows from (3.10) that

$$(3.11) \quad \|\Sigma_1^{-1}\| \leq \frac{1}{\bar{\sigma}_r - L_1 \|\bar{x}_k - x_k\|} \leq \frac{2}{\bar{\sigma}_r};$$

moreover, we have from (3.6) that

$$(3.12) \quad \|\lambda_k^{-1} \Sigma_2\| = \frac{\|\Sigma_2\|}{\mu_k \|F_k\|^\delta} \leq \frac{L_1}{m c^\delta} \|\bar{x}_k - x_k\|^{1-\delta}.$$

The above inequalities together with (3.8) and (3.9) imply that there exists a positive  $\tilde{c}_2$  such that

$$(3.13) \quad \|\hat{d}_k\| \leq 2\|d_k\| + \tilde{c}_2\|d_k\|^2\|\bar{x}_k - x_k\|^{1-\delta}$$

holds for all sufficiently large  $k$ . Hence we obtain from (3.7) that

$$(3.14) \quad \|s_k\| = \|d_k + \hat{d}_k\| \leq \|d_k\| + \|\hat{d}_k\| \leq c_2\|\bar{x}_k - x_k\|$$

for some  $c_2 > 0$ . The proof is completed.  $\square$

**3.2. Boundedness of the LM parameter.** The updating rule of  $\{\mu_k\}$  indicates that  $\{\mu_k\}$  is bounded below. Next we show that  $\mu_k$  is also bounded above.

**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*

$$(3.15) \quad \mu_k \leq M$$

holds for all sufficiently large  $k$ .

*Proof.* First we prove that for all sufficiently large  $k$ ,

$$(3.16) \quad \|F_k\|^2 - \|F_k + J_k d_k\|^2 \geq \tilde{c}_3 \|F_k\| \min\{\|d_k\|, \|\bar{x}_k - x_k\|\},$$

$$(3.17) \quad \|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \geq \hat{c}_3 \|F(y_k)\| \min\{\|\hat{d}_k\|, \|\bar{y}_k - y_k\|\},$$

where  $\tilde{c}_3, \hat{c}_3$  are some positive constants.

To prove (3.16), we consider two cases. If  $\|\bar{x}_k - x_k\| \leq \|d_k\|$ , by (3.1), (3.3) and the fact that  $d_k$  is the solution of (2.4), we have

$$(3.18) \quad \begin{aligned} \|F_k\| - \|F_k + J_k d_k\| &\geq \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\| \\ &\geq c_1 \|\bar{x}_k - x_k\| + O(\|\bar{x}_k - x_k\|^2) \\ &\geq \tilde{c}_3 \|\bar{x}_k - x_k\| \end{aligned}$$

for some  $\tilde{c}_3 > 0$ . In the other case that  $\|\bar{x}_k - x_k\| > \|d_k\|$ , we have

$$(3.19) \quad \begin{aligned} \|F_k\| - \|F_k + J_k d_k\| &\geq \|F_k\| - \left\| F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} J_k(\bar{x}_k - x_k) \right\| \\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} (\|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|) \\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|} \cdot \tilde{c}_3 \|\bar{x}_k - x_k\| \\ &\geq \tilde{c}_3 \|d_k\|. \end{aligned}$$

Both (3.18) and (3.19) indicate that

$$\begin{aligned} \|F_k\|^2 - \|F_k + J_k d_k\|^2 &= (\|F_k\| + \|F_k + J_k d_k\|)(\|F_k\| - \|F_k + J_k d_k\|) \\ &\geq \tilde{c}_3 \|F_k\| \min\{\|d_k\|, \|\bar{x}_k - x_k\|\}. \end{aligned}$$

Similarly, if  $\|\bar{y}_k - y_k\| \leq \|\hat{d}_k\|$ , then by (3.1)–(3.3) and the fact that  $\hat{d}_k$  is the solution of (2.7), we have

$$\begin{aligned} &\|F(y_k)\| - \|F(y_k) + J_k \hat{d}_k\| \\ &\geq \|F(y_k)\| - \|F(y_k) + J_k(\bar{y}_k - y_k)\| \\ (3.20) \quad &\geq \|F(y_k)\| - \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| - \|(J_k - J(y_k))(\bar{y}_k - y_k)\| \\ &\geq c_1 \|\bar{y}_k - y_k\| + O(\|\bar{y}_k - y_k\|^2) + O(\|d_k\| \|\bar{y}_k - y_k\|) \\ &\geq \hat{c}_3 \|\bar{y}_k - y_k\| \end{aligned}$$

for some  $\hat{c}_3 > 0$ . In the other case where  $\|\bar{y}_k - y_k\| > \|\hat{d}_k\|$ , we have

$$\begin{aligned} &\|F(y_k)\| - \|F(y_k) + J_k \hat{d}_k\| \\ &\geq \|F(y_k)\| - \|F(y_k) + \frac{\|\hat{d}_k\|}{\|\bar{y}_k - y_k\|} J_k(\bar{y}_k - y_k)\| \\ (3.21) \quad &\geq \frac{\|\hat{d}_k\|}{\|\bar{y}_k - y_k\|} (\|F(y_k)\| - \|F(y_k) + J_k(\bar{y}_k - y_k)\|) \\ &\geq \frac{\|\hat{d}_k\|}{\|\bar{y}_k - y_k\|} \cdot \hat{c}_3 \|\bar{y}_k - y_k\| \\ &\geq \hat{c}_3 \|\hat{d}_k\|. \end{aligned}$$

Combining (3.20) and (3.21), we obtain

$$\begin{aligned} &\|F(y_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 \\ &= (\|F(y_k)\| + \|F(y_k) + J_k \hat{d}_k\|)(\|F(y_k)\| - \|F(y_k) + J_k \hat{d}_k\|) \\ &\geq \hat{c}_3 \|F(y_k)\| \min\{\|\bar{y}_k - y_k\|, \|\hat{d}_k\|\}. \end{aligned}$$

Hence it follows from (3.3), (3.7), (3.16) and (3.17) that

$$\begin{aligned} (3.22) \quad |r_k - 1| &= \left| \frac{Ared_k - Pred_k}{Pred_k} \right| \\ &= \left| \frac{\|F(x_k + d_k + \hat{d}_k)\|^2 - \|F(y_k) + J_k \hat{d}_k\|^2 + \|F(y_k)\|^2 - \|F_k + J_k d_k\|^2}{Pred_k} \right| \\ &\leq \frac{\|F_k + J_k s_k\| O(\|s_k\|^2 + \|d_k\|^2) + O(\|s_k\|^4 + \|d_k\|^4) + \|F_k + J_k d_k\| O(\|d_k\|^2)}{\tilde{c}_3 \|F_k\| \min\{\|d_k\|, \|\bar{x}_k - x_k\|\} + \hat{c}_3 \|F(y_k)\| \min\{\|\bar{y}_k - y_k\|, \|\hat{d}_k\|\}} \\ &\leq \frac{\|F_k + J_k s_k\| O(\|s_k\|^2 + \|d_k\|^2) + O(\|s_k\|^4 + \|d_k\|^4) + \|F_k + J_k d_k\| O(\|d_k\|^2)}{\|\bar{x}_k - x_k\| \|d_k\|}. \end{aligned}$$

In view of (3.4), (3.7) and (3.13), we see that

$$(3.23) \quad \|F_k + J_k d_k\| \leq \|F_k\| \leq O(\|\bar{x}_k - x_k\|),$$

$$(3.24) \quad \begin{aligned} \|F_k + J_k s_k\| &\leq \|F_k + J_k d_k\| + \|J_k \hat{d}_k\| \\ &\leq \|F_k\| + L_2(2\|d_k\| + \bar{c}_2\|d_k\|^2)\|\bar{x}_k - x_k\|^{1-\delta} \\ &\leq O(\|\bar{x}_k - x_k\|) \end{aligned}$$

and

$$(3.25) \quad \begin{aligned} \|s_k\|^2 &= \|d_k + \hat{d}_k\|^2 \leq \|d_k\|^2 + 2\|d_k\|\|\hat{d}_k\| + \|\hat{d}_k\|^2 \\ &\leq \|d_k\|^2 + 2\|d_k\|(2\|d_k\| + \bar{c}_2\|d_k\|^2)\|\bar{x}_k - x_k\|^{1-\delta} \\ &\quad + (2\|d_k\| + \bar{c}_2\|d_k\|^2)\|\bar{x}_k - x_k\|^{1-\delta})^2 \\ &\leq 9\|d_k\|^2 + O(\|d_k\|^3\|\bar{x}_k - x_k\|^{1-\delta}) + O(\|d_k\|^4\|\bar{x}_k - x_k\|^{2-2\delta}). \end{aligned}$$

It then follows from (3.7) that

$$\frac{O(\|d_k\|^2)}{\|\bar{x}_k - x_k\|\|d_k\|} \leq O(1)$$

and

$$\frac{O(\|s_k\|^2)}{\|\bar{x}_k - x_k\|\|d_k\|} \leq O(1).$$

The two inequalities above together with (3.22)–(3.24) imply that

$$r_k \rightarrow 1.$$

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

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

$$(3.26) \quad \lambda_k = \mu_k \|F_k\|^\delta \leq L_2^\delta M \|\bar{x}_k - x_k\|^\delta,$$

which means the LM parameter is bounded.

In the remainder of this section, we will show that  $\{x_k\}$  converges to some solution of the nonlinear equations (1.1) cubically.

**3.3. Cubic convergence of Algorithm 2.1.** By the SVD of  $J_k$ , we compute

$$(3.27) \quad d_k = -V_1(\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 U_1^T F_k - V_2(\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 U_2^T F_k,$$

$$(3.28) \quad \hat{d}_k = -V_1(\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 U_2^T F(y_k),$$

and

$$(3.29) \quad \begin{aligned} F_k + J_k d_k &= F_k - U_1 \Sigma_1 (\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 U_1^T F_k - U_2 \Sigma_2 (\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 U_2^T F_k \\ &= \lambda_k U_1 (\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F_k + \lambda_k U_2 (\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F_k + U_3 U_3^T F_k, \end{aligned}$$

$$(3.30) \quad \begin{aligned} F(y_k) + J_k \hat{d}_k &= F(y_k) - U_1 \Sigma_1 (\Sigma_1^2 + \lambda_k I)^{-1} \Sigma_1 U_1^T F(y_k) - U_2 \Sigma_2 (\Sigma_2^2 + \lambda_k I)^{-1} \Sigma_2 U_2^T F(y_k) \\ &= \lambda_k U_1 (\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F(y_k) + \lambda_k U_2 (\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F(y_k) + U_3 U_3^T F(y_k). \end{aligned}$$

In order to obtain the cubic convergence rate of the algorithm, not only a better estimation of  $\hat{d}_k$  than (3.13) is required but also the estimations of  $\|F_k + J_k d_k\|$

and  $\|F(y_k) + J_k \hat{d}_k\|$  are required. The following two lemmas give the estimations of  $\|U_1^T F_k\|$ ,  $\|U_2^T F_k\|$ ,  $\|U_3^T F_k\|$  as well as  $\|U_1^T F(y_k)\|$ ,  $\|U_2^T F(y_k)\|$ ,  $\|U_3^T F(y_k)\|$ , which are involved in (3.29) and (3.30).

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

- (a)  $\|U_1 U_1^T F_k\| \leq L_2 \|\bar{x}_k - x_k\|$ ;
- (b)  $\|U_2 U_2^T F_k\| \leq 2L_1 \|\bar{x}_k - x_k\|^2$ ;
- (c)  $\|U_3 U_3^T F_k\| \leq L_1 \|\bar{x}_k - x_k\|^2$ ,

where  $L_1, L_2$  are given in (3.2) and (3.4), respectively.

*Proof.* Result (a) follows immediately from (3.4).

Let  $\bar{s}_k = -J_k^+ F_k$ , where  $J_k^+$  is the pseudo-inverse of  $J_k$ . It is easy to see that  $\bar{s}_k$  is the least squares solution of  $\min \|F_k + J_k s\|$ , so we have from (3.3) that

$$\|U_3 U_3^T F_k\| = \|F_k + J_k \bar{s}_k\| \leq \|F_k + J_k(\bar{x}_k - x_k)\| \leq L_1 \|\bar{x}_k - x_k\|^2.$$

Let  $\tilde{J}_k = U_1 \Sigma_1 V_1^T$  and  $\tilde{s}_k = -\tilde{J}_k^+ F_k$ . Since  $\tilde{s}_k$  is the least squares solution of  $\min \|F_k + \tilde{J}_k s\|$ , it follows from (3.3) and (3.10) that

$$\begin{aligned} \|(U_2 U_2^T + U_3 U_3^T) F_k\| &= \|F_k + \tilde{J}_k \tilde{s}_k\| \\ &\leq \|F_k + \tilde{J}_k(\bar{x}_k - x_k)\| \\ &\leq \|F_k + J_k(\bar{x}_k - x_k)\| + \|(\tilde{J}_k - J_k)(\bar{x}_k - x_k)\| \\ &\leq L_1 \|\bar{x}_k - x_k\|^2 + \|U_2 \Sigma_2 V_2^T(\bar{x}_k - x_k)\| \\ &\leq L_1 \|\bar{x}_k - x_k\|^2 + L_1 \|\bar{x}_k - x_k\| \|\bar{x}_k - x_k\| \\ &\leq 2L_1 \|\bar{x}_k - x_k\|^2. \end{aligned}$$

Due to the orthogonality of  $U_2$  and  $U_3$ , we get result (b).  $\square$

**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)  $\|U_1 U_1^T F(y_k)\| \leq c_4 \|\bar{x}_k - x_k\|^2$ ;
- (b)  $\|U_2 U_2^T F(y_k)\| \leq c_5 \|x_k - \bar{x}_k\|^3$ ;
- (c)  $\|U_3 U_3^T F(y_k)\| \leq c_6 \|\bar{x}_k - x_k\|^3$ ,

where  $c_4, c_5, c_6$  are positive constants.

*Proof.* It follows from (3.11), (3.26), (3.29) and Lemma 3.4 that

$$\begin{aligned} \|F_k + J_k d_k\| &\leq \frac{4L_2^{1+\delta} M}{\bar{\sigma}_r^2} \|\bar{x}_k - x_k\|^{1+\delta} + (2L_1 + L_1) \|\bar{x}_k - x_k\|^2 \\ &\leq \left( \frac{4L_2^{1+\delta} M}{\bar{\sigma}_r^2} + 3L_1 \right) \|\bar{x}_k - x_k\|^2, \end{aligned}$$

which together with (3.3) and (3.7) imply that

$$\begin{aligned} \|F(y_k)\| &= \|F(x_k + d_k)\| \\ &\leq \|F_k + J_k d_k\| + L_1 \|d_k\|^2 \\ &\leq \left( \frac{4L_2^{1+\delta} M}{\bar{\sigma}_r^2} + 3L_1 + L_1 \tilde{c}_2^2 \right) \|\bar{x}_k - x_k\|^2 \\ &= c_4 \|\bar{x}_k - x_k\|^2, \end{aligned}$$

where  $c_4 = 4L_2^{1+\delta}M\bar{\sigma}_r^{-2} + 3L_1 + L_1\tilde{c}_2^2$ . So we have

$$(3.31) \quad \|U_1U_1^T F(y_k)\| \leq \|F(y_k)\| \leq c_4\|\bar{x}_k - x_k\|^2.$$

Thus the local error bound condition yields

$$(3.32) \quad \|\bar{y}_k - y_k\| \leq c_1^{-1}\|F(y_k)\| \leq c_7\|\bar{x}_k - x_k\|^2,$$

where  $c_7 = c_1^{-1}c_4$ .

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), (3.7) and (3.32) that

$$(3.33) \quad \begin{aligned} \|U_3U_3^T F(y_k)\| &= \|F(y_k) + J_k\bar{p}_k\| \\ &\leq \|F(y_k) + J_k(\bar{y}_k - y_k)\| \\ &\leq \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(J_k - J(y_k))(\bar{y}_k - y_k)\| \\ &\leq L_1\|\bar{y}_k - y_k\|^2 + L_1\|d_k\|\|\bar{y}_k - y_k\| \\ &\leq L_1c_7^2\|\bar{x}_k - x_k\|^4 + L_1\tilde{c}_2c_7\|\bar{x}_k - x_k\|^3 \\ &\leq c_6\|\bar{x}_k - x_k\|^3, \end{aligned}$$

where  $c_6 = L_1c_7^2 + L_1\tilde{c}_2c_7$ .

Let  $\tilde{J}_k = U_1\Sigma_1V_1^T$  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), (3.7) (3.10) and (3.32) that

$$(3.34) \quad \begin{aligned} \|(U_2U_2^T + U_3U_3^T)F(y_k)\| &= \|F(y_k) + \tilde{J}_k\tilde{p}_k\| \leq \|F(y_k) + \tilde{J}_k(\bar{y}_k - y_k)\| \\ &\leq \|F(y_k) + J(y_k)(\bar{y}_k - y_k)\| + \|(\tilde{J}_k - J(y_k))(\bar{y}_k - y_k)\| \\ &\leq L_1\|\bar{y}_k - y_k\|^2 + \|(J_k - J(y_k) - U_2\Sigma_2V_2^T)(\bar{y}_k - y_k)\| \\ &\leq L_1\|\bar{y}_k - y_k\|^2 + \|(J_k - J(y_k))(\bar{y}_k - y_k)\| + \|U_2\Sigma_2V_2^T(\bar{y}_k - y_k)\| \\ &\leq L_1\|\bar{y}_k - y_k\|^2 + L_1\|d_k\|\|\bar{y}_k - y_k\| + L_1\|\bar{x}_k - x_k\|\|\bar{y}_k - y_k\| \\ &\leq L_1c_7^2\|\bar{x}_k - x_k\|^4 + L_1\tilde{c}_2c_7\|\bar{x}_k - x_k\|^3 + L_1c_7\|\bar{x}_k - x_k\|^3. \end{aligned}$$

Combining (3.33) and (3.34), we know there exists a positive  $c_5 > 0$  such that

$$\|U_2U_2^T F(y_k)\| \leq c_5\|\bar{x}_k - x_k\|^3$$

due to the orthogonality of  $U_2$  to  $U_3$ . The proof is completed.  $\square$

Based on the results obtained above, we can show that  $\{x_k\}$  converges to some solution of (1.1) cubically.

It follows from (3.6), (3.10), (3.11), (3.28), (3.30) and Lemma 3.5 that

$$(3.35) \quad \begin{aligned} \|\hat{d}_k\| &= \|-V_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1U_1^T F(y_k) - V_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2U_2^T F(y_k)\| \\ &\leq \|\Sigma_1^{-1}\| \|U_1^T F(y_k)\| + \|\lambda_k^{-1}\Sigma_2\| \|U_2^T F(y_k)\| \\ &\leq \frac{2c_4}{\bar{\sigma}_r}\|\bar{x}_k - x_k\|^2 + \frac{L_1c_5}{mc_1^\delta}\|\bar{x}_k - x_k\|^{4-\delta} \\ &\leq c_8\|\bar{x}_k - x_k\|^2, \end{aligned}$$

where  $c_8 = 2c_4\bar{\sigma}_r^{-1} + L_1c_1^{-\delta}c_5m^{-1}$ , and

$$\begin{aligned}
(3.36) \quad & \|F(y_k) + J_k\hat{d}_k\| \\
&= \|\lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1} U_1^T F(y_k) + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1} U_2^T F(y_k) + U_3 U_3^T F(y_k)\| \\
&\leq \lambda_k \|\Sigma_1^{-2}\| \|U_1^T F(y_k)\| + \|U_2^T F(y_k)\| + \|U_3^T F(y_k)\| \\
&\leq \frac{4L_2^\delta M c_4}{\bar{\sigma}_r^2} \|\bar{x}_k - x_k\|^{2+\delta} + (c_5 + c_6) \|\bar{x}_k - x_k\|^3 \\
&\leq c_9 \|\bar{x}_k - x_k\|^3,
\end{aligned}$$

where  $c_9 = 4L_2^\delta M c_4 \bar{\sigma}_r^{-2} + c_5 + c_6$ . Thus, we have from (3.1)–(3.3), (3.7), (3.35) and (3.36) that

$$\begin{aligned}
c_1 \|\bar{x}_{k+1} - x_{k+1}\| &\leq \|F(x_{k+1})\| = \|F(y_k + \hat{d}_k)\| \\
&\leq \|F(y_k) + J(y_k)\hat{d}_k\| + L_1 \|\hat{d}_k\|^2 \\
&\leq \|F(y_k) + J_k\hat{d}_k\| + \|(J(y_k) - J_k)\hat{d}_k\| + L_1 \|\hat{d}_k\|^2 \\
&\leq c_9 \|\bar{x}_k - x_k\|^3 + L_1 \|d_k\| \|\hat{d}_k\| + L_1 \|\hat{d}_k\|^2 \\
&\leq c_9 \|\bar{x}_k - x_k\|^3 + L_1 \tilde{c}_2 c_8 \|\bar{x}_k - x_k\|^3 + L_1 c_8^2 \|\bar{x}_k - x_k\|^4,
\end{aligned}$$

which gives

$$(3.37) \quad \|\bar{x}_{k+1} - x_{k+1}\| \leq c_{10} \|\bar{x}_k - x_k\|^3,$$

where  $c_{10} = c_1^{-1}(c_9 + L_1 \tilde{c}_2 c_8 + L_1 c_8^2)$ . (3.37) implies that  $\{x_k\}$  converges to the solution set cubically.

Note that

$$(3.38) \quad \|\bar{x}_k - x_k\| \leq \|\bar{x}_{k+1} - x_{k+1}\| + \|s_k\|.$$

It follows from (3.37) that

$$(3.39) \quad \|\bar{x}_k - x_k\| \leq 2\|s_k\|$$

holds for all sufficiently large  $k$ . Hence we deduce from (3.37) and Lemma 3.2 that

$$(3.40) \quad \|s_{k+1}\| \leq O(\|s_k\|^3).$$

The above inequality indicates that  $\{x_k\}$  converges to some solution of (1.1) cubically, which is stronger than the convergence to the solution set.

We summarize our results in this section as follows:

**Theorem 3.6.** *Under the conditions of Assumption 3.1, the sequence generated by Algorithm 2.1 converges to some solution of (1.1) cubically.*



From the analyses above, we can see that under the local error bound condition which is weaker than nonsingularity, we have

$$\frac{\lambda_{k+1}}{\lambda_k^3} = \frac{\mu_{k+1}\|F_{k+1}\|^\delta}{(\mu_k\|F_k\|^\delta)^3} \leq O\left(\frac{\|x_{k+1} - \bar{x}_{k+1}\|^\delta}{\|x_k - \bar{x}_k\|^{3\delta}}\right) = O\left(\frac{\|s_{k+1}\|^\delta}{\|s_k\|^{3\delta}}\right) = O(1),$$

which implies that not only the LM parameter  $\{\lambda_k\}$  converges but also  $\{F_k\}$  converges cubically to zero as  $\{x_k\}$  converges cubically to the solution of (1.1).

#### 4. NUMERICAL RESULTS

We tested Algorithm 2.1 on some singular problems, and compared it with the general LM algorithm which does not solve the linear equations (2.13) and only computes the LM step  $d_k$ .

The test problems were created by modifying the nonsingular problems given by Moré, Garbow and Hillstom in [9], and have the same form as in [11],

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

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  $1 \leq k \leq n$ . Obviously,  $\hat{F}(x^*) = 0$  and

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

has rank  $n - k$ . A disadvantage of these problems is that  $\hat{F}(x)$  may have roots that are not roots of  $F(x)$ . We created two sets of singular problems, with  $\hat{J}(x^*)$  having rank  $n - 1$  and  $n - 2$ , by using

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

and

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

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  $n$ -th equations. (The first  $n$  equations in the standard problem are linear.)

We set  $p_0 = 0.0001, p_1 = 0.25, p_2 = 0.75, \mu_1 = 10^{-5}$  and  $\delta = 1$  for all the tests. The algorithm is terminated when the norm of  $J_k^T F_k$ , e.g., the derivative of  $\Phi$  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 of problems of rank  $n - 1$  are listed in Table 1, and the second set of 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 Hillstom in [9]; “NF” and “NJ” represent the numbers of function calculations, and Jacobian calculations, respectively; and “n.s. $x^*$ ?” gives a Y(yes) if the method converges to the same solution as the corresponding nonsingular problem, an N(no) otherwise. 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. Note that, for general nonlinear equations, the calculations of the Jacobian are usually  $n$  times of the function calculations. So we also presented the values “NF+NJ $\cdot n$ ” for comparisons of the total calculations. However, if the Jacobian is sparse, this kind of value does not mean much.

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

Problem	$n$	$x_0$	$(s_k = d_k)$		$(s_k = d_k + \tilde{d}_k)$	
			NF/NJ/NF+NJ* $n$	n.s. $x^*$ ?	NF/NJ/NF+NJ* $n$	n.s. $x^*$ ?
1	2	1	15/15/45	Y	21/11/43	Y
		10	17/17/51	Y	25/13/51	Y
		100	21/21/63	Y	29/15/59	Y
3	2	1	–		OF	
		10	35/18/71	Y	–	
4	4	1	16/16/80	Y	23/12/71	Y
		10	19/19/95	Y	27/14/83	Y
		100	22/22/110	Y	31/16/95	Y
5	3	1	8/8/32	N	11/6/29	N
		10	8/8/32	N	9/5/24	N
		100	8/8/32	N	11/6/29	N
6	31	1	44/25/819	N	75/15/540	N
8	10	1	8/8/88	Y	11/6/71	Y
		10	23/23/253	Y	33/17/203	Y
9	10	1	4/4/44	N	5/3/35	N
		10	7/7/77	N	11/6/71	N
		100	9/9/99	N	13/7/83	N
10	30	1	6/6/186	Y	7/4/127	Y
		10	7/7/217	Y	11/6/191	Y
		100	10/10/310	N	13/7/223	N
11	30	1	35/19/605	N	45/9/315	N
		10	60/46/1440	N	85/23/775	N
		100	45/35/1095	N	99/24/819	N
12	10	1	14/14/154	Y	19/10/119	Y
		10	16/16/176	Y	21/11/131	Y
		100	19/19/209	Y	27/14/167	Y
13	30	1	9/9/279	Y	13/7/223	Y
		10	14/14/434	Y	19/10/319	Y
		100	17/17/527	Y	25/13/415	Y
14	30	1	12/12/372	Y	17/9/ 287	Y
		10	18/18/558	Y	27/14/447	Y
		100	24/24/744	Y	35/18/575	Y

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

Problem	$n$	$x_0$	$(s_k = d_k)$		$(s_k = d_k + \hat{d}_k)$	
			NF/NJ/NF+NJ* $n$	n.s. $x^*$ ?	NF/NJ/NF+NJ* $n$	n.s. $x^*$ ?
1	2	1	11/11/33	N	15/8/31	N
		10	13/13/39	N	19/10/39	N
		100	17/17/51	N	23/12/47	N
3	2	1	35/25/85	N	57/17/91	N
		10	59/54/167	N	91/34/159	N
		100	25/18/61	N	51/15/81	N
4	4	1	14/14/70	N	19/10/59	N
		10	17/17/85	N	23/12/71	N
		100	20/20/100	N	29/15/89	N
5	3	1	13/13/52	Y	19/10/49	Y
		10	14/14/56	Y	19/10/49	Y
		100	24/18/78	Y	21/11/54	Y
6	31	1	3200/1793/58783	N	3201/801/28032	N
8	10	1	8/8/88	Y	11/6/71	Y
		10	23/23/253	Y	33/17/203	Y
9	10	1	4/4/44	N	5/3/35	N
		10	7/7/77	N	11/6/71	N
		100	10/10/110	N	13/7/83	N
10	30	1	6/6/186	Y	7/4/127	Y
		10	7/7/217	Y	11/6/191	Y
		100	10/10/310	N	15/8/255	N
11	30	1	19/9/289	N	41/9/311	N
		10	62/41/1292	N	65/17/575	N
		100	43/35/1093	N	87/23/777	N
12	10	1	14/14/154	Y	19/10/119	Y
		10	16/16/176	Y	21/11/131	N
		100	19/19/209	Y	27/14/167	N
13	30	1	9/9/279	Y	13/7/223	Y
		10	14/14/434	Y	19/10/319	Y
		100	17/17/527	Y	25/13/415	Y
14	30	1	12/12/372	Y	17/9/287	Y
		10	18/18/558	Y	27/14/447	Y
		100	24/24/744	Y	35/18/575	Y

From the tables, we can see that Algorithm 2.1 almost always outperforms the general LM algorithm, whether on the first singular test set or on the second test set. Though the function calculations of Algorithm 2.1 are more than those of the general LM algorithm, their Jacobian calculations are much less, which contributes to fewer calculations of Algorithm 2.1 than the general LM algorithm. The numerical results also support the faster local convergence result of Algorithm 2.1 obtained in Section 3 in some sense.

#### ACKNOWLEDGMENTS

The author would like to thank the referees for the helpful comments and suggestions, which have greatly improved the original version of the paper.

#### REFERENCES

1. J. Y. Fan, *A modified Levenberg-Marquardt algorithm for singular system of nonlinear equations*, Journal of Computational Mathematics, **21** (2003), 625–636. MR1999973 (2004g:65071)
2. J. Y. Fan and J. Y. Pan, *A note on the Levenberg-Marquardt parameter*, Applied Mathematics and Computation, **207** (2009), 351–359. MR2489105 (2010a:65080)
3. J. Y. Fan and Y. X. Yuan, *On the quadratic convergence of the Levenberg-Marquardt method without nonsingularity assumption*, Computing, **74** (2005), 23–39. MR2127319 (2006a:65079)
4. C. T. Kelley, *Iterative Methods for Optimization*, Frontiers in Applied Mathematics, 18, SIAM, Philadelphia, 1999. MR1678201 (2000h:90003)
5. K. Levenberg, *A method for the solution of certain nonlinear problems in least squares*, Quart. Appl. Math., **2** (1944), 164–166. MR0010666 (6:52a)
6. D. W. Marquardt, *An algorithm for least-squares estimation of nonlinear inequalities*, SIAM J. Appl. Math., **11** (1963), 431–441. MR0153071 (27:3040)
7. J. J. Moré, *The Levenberg-Marquardt algorithm: implementation and theory*, in: G. A. Watson, ed., Lecture Notes in Mathematics 630: Numerical Analysis, Springer-Verlag, Berlin, 1978, pp. 105–116. MR0483445 (58:3446)
8. J. J. Moré, *Recent developments in algorithms and software for trust region methods*, In: A. Bachem, M. Grötschel and B. Korte, eds., Mathematical Programming: The State of Art (Springer, Berlin, 1983) 258–287. MR717404 (85b:90066)
9. J. J. Moré, B. S. Garbow and K. H. Hillstom, *Testing unconstrained optimization software*, ACM Trans. Math. Software, **7** (1981), 17–41. MR607350 (83b:90144)
10. M. J. D. Powell, *Convergence properties of a class of minimization algorithms*, in: O. L. Mangasarian, R. R. Meyer and S. M. Robinson, eds., Nonlinear Programming 2, Academic Press, New York, 1975, pp. 1–27. MR0388270 (52:7128)
11. R. B. Schnabel and P. D. Frank, *Tensor methods for nonlinear equations*, SIAM J. Numer. Anal., **21** (1984), 815–843. MR760620 (86d:65066)
12. G. W. Stewart and J. G. Sun, *Matrix Perturbation Theory*, Academic Press, San Diego, CA, 1990. MR1061154 (92a:65017)
13. W. Y. Sun and Y. X. Yuan, *Optimization Theory and Methods: Nonlinear Programming*, Springer, New York, 2006. MR2232297 (2007c:90002)
14. N. Yamashita and M. Fukushima, *On the rate of convergence of the Levenberg-Marquardt method*, Computing (Supplement 15), (2001), 237–249. MR1874516 (2002k:65088)
15. Y. X. Yuan, *Trust region algorithms for nonlinear equations*, Information, **1** (1998), 7–20. MR1645695 (2000d:65090)
16. Y. X. Yuan, *A review of trust region algorithms for optimization*, In: J. M. Ball and J. C. R. Hunt, eds., ICM99: Proceedings of the Fourth International Congress on Industrial and Applied Mathematics, Oxford University Press, 2000, pp. 271–282. MR1824450 (2002a:90076)

DEPARTMENT OF MATHEMATICS, SHANGHAI JIAO TONG UNIVERSITY, SHANGHAI 200240, PEOPLE'S REPUBLIC OF CHINA

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