symmetry

MDPI

Article

A New Adaptive Accelerated Levenberg–Marquardt Method for Solving Nonlinear Equations and Its Applications in Supply Chain Problems

Rong Li [1], Mingyuan Cao [1,2,*] and Guoling Zhou [1]

[1] School of Mathematics and Statistics, Beihua University, Jilin 132013, China
[2] School of Information Engineering, Hainan Vocational University of Science and Technology, Hainan 571126, China
* Correspondence: cmy0918@beihua.edu.cn

Abstract: In this paper, a new adaptive Levenberg–Marquardt method is proposed to solve the nonlinear equations including supply chain optimization problems. We present a new adaptive update rule which is a segmented function on the ratio between the actual and predicted reductions of the objective function to accept a large number of unsuccessful iterations and avoid jumping in local areas. The global convergence and quadratic convergence of the proposed method are proved by using the trust region technique and local error bound condition, respectively. In addition, we use the proposed algorithm to test on the symmetric and asymmetric linear equations. Numerical results show that the proposed method has good numerical performance and development prospects. Furthermore, we apply the algorithm to solve the fresh agricultural products supply chain optimization problems.

Keywords: accelerate Levenberg–Marquardt method; adaptive function; trust region technique; local error bound condition

Citation: Li, R.; Cao, M.; Zhou, G. A New Adaptive Accelerated Levenberg–Marquardt Method for Solving Nonlinear Equations and Its Applications in Supply Chain Problems. *Symmetry* **2023**, *15*, 588. https://doi.org/10.3390/sym15030588

Academic Editor: Theodore E. Simos

Received: 10 February 2023
Revised: 21 February 2023
Accepted: 22 February 2023
Published: 24 February 2023

1. Introduction

With the development of science and technology, more and more fields are involved in the solution of nonlinear equation problems, such as chemistry, mechanics, economics and product management [1–4]. For example, decentralized decision models in supply chain management and gas pressure volume models in physics can be converted into the following nonlinear equations

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

where $F(x) : \mathbb{R}^n \to \mathbb{R}^m$ is a continuously differentiable function. In particular, symmetric nonlinear equations with the Jacobian matrix symmetry also have a wide range of applications, such as the gradient mapping of unconstrained optimization problem, the Karush–Kuhn–Tucker (KKT) of equality constrained optimization problem, and other fields [5,6].

The steepest descent method, Newton method, quasi-Newton method, Gauss–Newton (GN) method are commonly used iterative methods for solving (1) [7–10]. The GN method is one of the most famous methods, when the Jacobian matrix is Lipschitz continuous and nonsingular at the solution of (1), the GN method has quadratic convergence. However, when the Jacobian matrix is singular or nearly singular, the GN method may not be well defined. In order to overcome this difficulty, the Levenberg–Marquardt (LM) method [11,12] for solving (1) was proposed. At the k-th iteration, the trial step is

$$d_k = -(J_k^{\mathrm{T}} J_k + \lambda_k I)^{-1} (J_k^{\mathrm{T}} F_k), \tag{2}$$

where $F_k = F(x_k)$, $J_k = J(x_k)$ is a Jacobian matrix of $F(x)$ at x_k, which may be a symmetric matrix or non-symmetric matrix, I is an identity matrix and the LM parameter $\lambda_k > 0$.

The LM method ensures the uniqueness of solution of (1), and it also has quadratic convergence if J_k is Lipschitz continuous, nonsingular at the solution, and λ_k is selected appropriately. In this sense, the update of the LM parameter has a great impact on the performance and efficiency of algorithm, many effective LM parameters have been proposed. Yamashita and Fukushima [13] chose the LM parameter as $\lambda_k = \|F_k\|^2$, and proved that the LM method had quadratic convergence under the local error bound condition and J_k is Lipschitz continuous at the solution. However, when $\{x_k\}$ is far away from the solution set, λ_k may be very large, which makes d_k very small and reduces the efficiency of algorithm; when $\{x_k\}$ is sufficiently close to the solution set, λ_k may be smaller than the machine epsilon and lose its role.

Based on these observations, Fan and Yuan [14] generalized the LM parameter in [13], and proved that the numerical results for choosing $\lambda_k = \|F_k\|$ is better than choosing $\lambda_k = \|F_k\|^2$. Fan [15] first introduced the regularization factor μ_k into the LM method and chose $\lambda_k = \mu_k\|F_k\|$, with numerical results showing that this choice of λ_k provides the best performance. However, when $\{x_k\}$ is far away from the solution set, the choice of both LM parameters does not provide good results. Therefore, to avoid this situation, Fan and Pan [16] chose the LM parameter as $\lambda_k = \mu_k\rho(x_k)$, in which μ_k is updated by a trust region technique. They defined $\rho(x_k)$ as a positive function of $\mathbb{R}^n \to \mathbb{R}_+$, i.e.,

$$\rho(x_k) = \begin{cases} \widetilde{\rho}(x_k), & \text{if } \widetilde{\rho}(x_k) \le 1, \\ 1, & \text{otherwise,} \end{cases}$$

where $\widetilde{\rho}(x_k) = O(\|F_k\|^\delta)$. This update strategy can obtain larger LM trial steps, so that the iterative sequence can quickly converge to the solution set when $\{x_k\}$ is far away from the solution set. Amini et al. [17] chose the LM parameter as

$$\lambda_k = \frac{\mu_k\|F_k\|}{1 + \|F_k\|}.$$

It is clear that when $\{x_k\}$ is far away from the solution set and $\|F_k\|$ is very large, $\frac{\|F_k\|}{1+\|F_k\|}$ is close to 1, so λ_k is close to μ_k. The choice of λ_k speeds up the efficiency of the algorithm more than previous LM parameters.

In addition to the above different choices of LM parameters, the introduction of adaptive technology also has a great impact on the LM method. As we all know, the ratio r_k between the actual and predicted reductions of the objective function reflects the degree to which the approximate quadratic model approaches the value function. To make more use of information about the ratio, Fan and Yuan [18] proposed an adaptive LM method by selecting $\lambda_{k+1} = \mu_{k+1}\|F_{k+1}\|^\delta$, $\mu_{k+1} = \mu_k q(r_k)$, $q(r_k)$ is a continuous non-negative function about r_k, and $\sigma \in (0, 2]$. The introduction of $q(r_k)$ avoids discontinuities when crossing the threshold $\frac{\mu_{k+1}}{\mu_k}$ of the ratio, and better numerical results can be obtained.

In fact, similar adaptive techniques have been proposed in the trust region algorithms. If r_k is sufficiently greater than 1, the iteration is too successful at this time, then we can reduce μ_k to a very small value. Then, the algorithm will continue to perform a large number of consecutive unsuccessful iterations. On the other hand, if $r_k \to -\infty$, d_k is a far-from-satisfactory trial step, then we can increase μ_k greatly. At this moment, the successive iteration points will be close to each other and the algorithm will converge slowly. Therefore, Hei [19] proposed an R-function by using an adaptive update strategy to update the trust region radius Δ_k, i.e., $\Delta_{k+1} = R(r_k)\Delta_k$. Furthermore, Walmag and Delhez [20] proposed a Λ-function to update the trust region radius, i.e., $\Delta_{k+1} = \Lambda(r_k)\Delta_k$, where Λ is a non-negative and bounded function about r_k. On this basis, Lu et al. [21] argued that the consistency between the model and the objective function is not good enough in too-successful iterations, so an L-function was proposed to update the trust region radius. They showed that the L-function contains some favorable features of the R-function and the Λ-function, and the method is more efficient in too-successful iterations. In this paper, we want to learn from the presentation of the L-function and provide a

new adaptive strategy to update the LM parameter. Our innovations mainly include the following:

♢ A new adaptive accelerated LM method is proposed, which can improve the consistency between the model and the objective function in too-successful iterations by using the ratio information of the actual reduction to the predicted reduction;

♢ The new algorithm can solve the situation in which the iterative sequence is far away from the optimal solution set, accept a large number of unsuccessful iterations and avoid jumping in local areas, thus improving the efficiency and stability of the algorithm;

♢ The new adaptive accelerated LM method has global convergence and quadratic convergence under local error bound.

The rest of this paper is organized as follows. In Section 2, we describe in detail a new adaptive accelerated LM method which makes full use of the ratio information. Furthermore, we demonstrate that the new algorithm has global convergence under the appropriate conditions and maintains quadratic convergence under local error bound condition. In Section 3, numerical results are given, indicating that the new algorithm is efficient. The conclusion is given in the last section.

2. Methodology

2.1. The Adaptive Accelerated Levenberg–Marquardt Method

In this section, our main aim is to discuss how to update the LM parameter to propose a new adaptive accelerated LM method. It is easy to see from (2) that d_k is the solution to the optimization problem

$$\min_{d\in\mathbb{R}^n} \|F_k + J_k d\|^2 + \lambda_k \|d\|^2 \doteq \psi_k(d). \tag{3}$$

If

$$\Delta_k = \| - (J_k^{\mathrm{T}} J_k + \lambda_k I)^{-1} (J_k^{\mathrm{T}} F_k)\|, \tag{4}$$

then d_k is also the solution of the subproblem

$$\begin{aligned} &\min_{d\in\mathbb{R}^n} \ \|F_k + J_k d\|^2 \doteq \varphi_k(d), \\ &\text{s.t.} \quad \|d\| \le \Delta_k. \end{aligned} \tag{5}$$

Therefore, the LM method can be regarded as a trust region method, which implicitly modifies the trust region radius Δ_k. The difference between the general trust region method and the LM method is that the LM method does not directly update the trust region radius, but updates the regularization factor μ_k.

We define the actual reduction and predicted reduction of the merit function $\|F_k\|^2$ at the k-th iteration as

$$Ared_k = \|F_k\|^2 - \|F(x_k + d_k)\|^2 \tag{6}$$

and

$$Pred_k = \varphi_k(0) - \varphi_k(d_k). \tag{7}$$

The ratio between the actual and predicted reductions of the objective function is defined by

$$r_k = \frac{Ared_k}{Pred_k}. \tag{8}$$

This ratio determines whether the trial step d_k is accepted. Here, we choose the LM parameter as

$$\lambda_{k+1} = \frac{\mu_{k+1}\|F_{k+1}\|}{1 + \|F_{k+1}\|}. \tag{9}$$

The usual empirical rules [22–25] of μ_{k+1} can be usually summarized as follows

$$\mu_{k+1} = \begin{cases} 4\mu_k, & \text{if } r_k < p_1, \\ \mu_k, & \text{if } p_1 \leq r_k \leq p_2, \\ \max\{\frac{\mu_k}{4}, m\}, & \text{if } r_k > p_2, \end{cases} \tag{10}$$

where $m > 0$ and $0 < p_1 < p_2 < 1$ are constants.

Iterations with r_k greater than p_2 are very successful iterations. In this case, it is usually assumed that the approximation of the model function to the objective function is accurate and μ_k should be reduced. However, at too-successful iterations, i.e., r_k is sufficiently greater than 1, the consistency between the model and the objective function is not good enough. Thus, we use an adaptive strategy to update the factor μ_{k+1}, i.e., $\mu_{k+1} = K(r_k)\mu_k$, where $K(r_k)$ is a function about r_k.

We construct $K(r_k)$ as follows:

$$K(r_k) = \begin{cases} \beta_1 + (\beta_2 - \beta_1)\exp(-(\frac{-r_k+p_1}{p_1})^2), & \text{if } r_k \leq p_1, \\ \beta_2, & \text{if } p_1 < r_k < p_2, \\ \frac{1-\beta_3\exp(p_2)}{1-\exp(p_2)} - \frac{(1-\beta_3)\exp(p_2)}{1-\exp(p_2)}\exp(-r_k + p_2) - \frac{1}{2}, & \text{if } r_k \geq p_2, \end{cases} \tag{11}$$

where $0 < \beta_2 < 1 < \beta_1 \leq \beta_3$ and $0 < p_1 < p_2 < 1$ are constants. Here, $K(r_k)$ satisfies the following properties

(1) $\lim\limits_{r_k \to -\infty} K(r_k) = \beta_1$;
(2) $\lim\limits_{r_k \to p_1} K(r_k) = \beta_2$;
(3) $\lim\limits_{r_k \to p_2} K(r_k) = \frac{1}{2}$;
(4) $\lim\limits_{r_k \to +\infty} K(r_k) = \frac{1-\beta_3\exp(p_2)}{1-\exp(p_2)} - \frac{1}{2}$.

If we obtain a satisfactory trial step d_k and ratio r_k, then we accept trial step d_k and reduce μ_k; otherwise, we reject trial step d_k and increase μ_k. At too-successful iterations, the actual reduction of the objective function obtained at iteration k is obviously greater than the predicted reduction. Although the current iteration allows the algorithm to progress towards the optimum, the approximation of the model function to the objective function is bad. Therefore, to avoid reducing μ_k too quickly, we use the K-function to update μ_k.

According to the properties of the K-function, the rate of μ_k reduction is the fastest when r_k is close to 1, i.e., when the model function provides an accurate local approximation of the objective function. The new idea we propose is to allow μ_k to be updated at a variable rate according to r_k, which would improve the efficiency and stability of the algorithm.

Based on the above analysis, we state a description of the new adaptive accelerated LM method (Algorithm 1) as follows.

In Algorithm 1, m is a given lower bound of the parameter μ_k. It is introduced to prevent the step from being too large when the sequence is near the solution.

Algorithm 1 NAALM.

0. Given $x_0 \in \mathbb{R}^n$, $\mu_0 > m > 0$, $0 \leq p_0 < p_1 < p_2 < 1$, $0 < \beta_2 < 1 < \beta_1 \leq \beta_3$, $\varepsilon > 0$. Let $k := 0$.

1. Compute F_k and J_k. If $\|J_k^{\mathrm{T}} F_k\| \leq \varepsilon$, stop. Otherwise, compute λ_k by (9).

2. Solving the following system

$$(J_k^{\mathrm{T}} J_k + \lambda_k I)d = -J_k^{\mathrm{T}} F_k \tag{12}$$

to determine d_k.

3. Compute $Pred_k$, $Ared_k$ and r_k by (6)–(8), respectively.

4. Set

$$x_{k+1} = \begin{cases} x_k + d_k, & \text{if } r_k \geq p_0, \\ x_k, & \text{if } r_k < p_0. \end{cases} \tag{13}$$

5. Choose μ_{k+1} as

$$\mu_{k+1} = \max\{m, K(r_k)\mu_k\}, \tag{14}$$

where $K(r_k)$ is given by (11). Set $k := k + 1$ and go to Step 1.

2.2. The Global Convergence

In this section, to obtain the global convergence of NAALM algorithm, we make the following assumption.

Assumption 1. *$F(x)$ is continuously differentiable, $F(x)$ and the Jacobian matrix $J(x)$ are Lipschitz continuous, i.e., there exist positive constants L_1 and L_2 such that*

$$\|J(y) - J(x)\| \leq L_1\|y - x\|, \ \ \forall\, x, y \in \mathbb{R}^n, \tag{15}$$

and

$$\|F(y) - F(x)\| \leq L_2\|y - x\|, \ \ \forall\, x, y \in \mathbb{R}^n. \tag{16}$$

Lemma 1. *Let d_k be computed by (12), then the inequality*

$$Pred_k \geq \|J_k^{\mathrm{T}} F_k\| \min\left\{\|d_k\|, \frac{\|J_k^{\mathrm{T}} F_k\|}{\|J_k^{\mathrm{T}} J_k\|}\right\} \tag{17}$$

holds for all $k \geq 0$.

Proof. From (7), for $\alpha \in [0, 1]$, we have

$$\begin{aligned} \|Pred_k\| &= \|F_k\|^2 - \|F_k + J_k d_k\|^2 \\ &\geq \|F_k\|^2 - \left\|F_k - J_k \frac{\alpha\|d_k\|}{\|J_k^{\mathrm{T}} F_k\|} J_k^{\mathrm{T}} F_k\right\|^2 \\ &\geq 2\alpha\|d_k\|\|J_k^{\mathrm{T}} F_k\| - \alpha^2\|d_k\|^2\|J_k^{\mathrm{T}} F_k\|, \end{aligned} \tag{18}$$

then

$$\begin{aligned} \|Pred_k\| &\geq \max_{0\leq\alpha\leq 1}\left(2\alpha\|d_k\|\|J_k^{\mathrm{T}} F_k\| - \alpha^2\|d_k\|^2\|J_k^{\mathrm{T}} F_k\|\right) \\ &\geq \|J_k^{\mathrm{T}} F_k\| \min\left\{\|d_k\|, \frac{\|J_k^{\mathrm{T}} F_k\|}{\|J_k^{\mathrm{T}} J_k\|}\right\}. \end{aligned} \tag{19}$$

The proof is complete. □

Theorem 1. *Under the conditions of Assumption 1, the sequence* $\{x_k\}$ *generated by NAALM algorithm satisfies*

$$\lim_{k\to\infty} \|J_k^{\mathrm{T}} F_k\| = 0. \tag{20}$$

Proof. If the theorem is not true, then there exist a positive τ and infinitely many k such that

$$\|J_k^{\mathrm{T}} F_k\| \geq \tau. \tag{21}$$

Let T_1, T_2 be the sets of all indices that satisfy

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

and

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

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

Case (I) T_2 is finite.

It follows from the definition of T_2 that the set

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

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 $k \in \{k > \tilde{k} | k \in T_1\}$. Define the indices set

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

Suppose $k \in T_4$. It is easy to see that $\|J_{k+1}^{\mathrm{T}} F_{k+1}\| \geq \tau$. 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 $\|J_k^{\mathrm{T}} F_k\| \geq \tau$ and $x_{k+1} = x_k$ hold for all $k > \tilde{k}$.

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

$$\mu_k \to +\infty \text{ and } \lambda_k \to +\infty, \tag{22}$$

due to (12)–(14) and $x_{k+1} = x_k$ for all $k > \tilde{k}$. Hence, we have

$$\lim_{k\to\infty} d_k = 0. \tag{23}$$

Furthermore, it follows from (21), (23) and Lemma 1 that

$$\begin{aligned} |r_k - 1| &= \left| \frac{Ared_k}{Pred_k} - 1 \right| \\ &= \left| \frac{\|F_k + J_k d_k\|^2 - \|F(x_k + d_k)\|^2}{Pred_k} \right| \\ &= \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{Pred_k} \\ &\leq \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{\|J_k^{\mathrm{T}} F_k\| \min\left\{ \|d_k\|, \frac{\|J_k^{\mathrm{T}} F_k\|}{\|J_k^{\mathrm{T}} J_k\|} \right\}} \\ &\leq \frac{O(\|d_k\|^2)}{\|d_k\|} \to 0, \end{aligned} \tag{24}$$

that is, $r_k \to 1$. In view of the updating rule of μ_k, we know that 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 (22). Hence, the supposition (21) cannot be true while T_2 is finite.

Case (II) T_2 is infinite.

It follows from Lemma 1 that

$$\begin{aligned} \|F_1\| &\geq \sum_{k\in T_2} (\|F_k\|^2 - \|F_{k+1}\|^2) \\ &\geq \sum_{k\in T_2} p_0 Pred_k \\ &\geq \sum_{k\in T_2} p_0 \|J_k^{\mathrm{T}} F_k\| \min\left\{ \|d_k\|, \frac{\|J_k^{\mathrm{T}} F_k\|}{\|J_k^{\mathrm{T}} J_k\|} \right\} \\ &\geq \sum_{k\in T_2} \frac{p_0 \tau}{2}, \end{aligned} \tag{25}$$

which gives

$$\sum_{k\in T_2} \|d_k\| < +\infty. \tag{26}$$

The above inequality, together with the Lipschitz conditions (15) and (16), implies that

$$\sum_{k\in T_2} \left| \|J_k^{\mathrm{T}} F_k\| - \|J_{k+1}^{\mathrm{T}} F_{k+1}\| \right| < +\infty. \tag{27}$$

Relation (27) and the fact that (21) holds for infinitely many k indicate that there exists a $\hat{k}$ with $\|J_{\hat{k}}^{\mathrm{T}} F_{\hat{k}}\| \geq \tau$ such that

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

By induction, we obtain that $\|J_k^{\mathrm{T}} F_k\| \geq \frac{\tau}{2}$ for all $k \geq \hat{k}$. This result and (26) mean that

$$\lim_{k\to\infty} \|d_k\| = 0. \tag{28}$$

It follows from (12) and (13) that $\mu_k \to +\infty$. By the same analysis as (24), we know that $\mu_k \to 1$. Hence, there exists a positive constant $\bar{m} > m$ such that $\mu_k < \bar{m}$ holds for all large k, which introduces a contradiction. Therefore, the supposition (21) cannot be true when T_2 is infinite. The proof is complete. □

2.3. Local Convergence

In this section, we will study the local convergence properties of the NAALM algorithm by using the singular value decomposition (SVD) technique. We assume that the sequence $\{x_k\}$ generated by the NAALM algorithm converges to the nonempty solution set X^* and lies in some neighborhood of $x^* \in X^*$. Firstly, we present some assumptions which the local convergence theory required.

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

$$c\mathrm{dist}(x, X^*) \leq \|F_k\|, x \in \mathbb{N}, \tag{29}$$

where $\mathrm{dist}(x, X^*)$ *is the distance from x to X^*.*

Assumption 2. *(i) $F(x)$ is continuously differentiable, and $J(x)$ is Lipschitz continuous on $N(x^*, b_1)$ with $b_1 < 1$, i.e., there exists a positive constant L_1 such that*

$$\|J(y) - J(x)\| \leq L_1\|y - x\|, \ \ \forall x, y \in N(x^*, b_1) = \{x | \|x - x^*\| \leq b_1\}. \tag{30}$$

(ii) $F(x)$ provides a local error bound on some neighborhood of $x^* \in X^*$, i.e., there exists a positive constant $c_1 > 0$ such that

$$\|F(x)\| \geq c_1 \text{dist}(x, X^*), \ \ \forall x \in N(x^*, b_1). \tag{31}$$

By the Lipschitzness of the Jacobian matrix proposed by (30), we have

$$\begin{aligned} \|F(y) - F(x) - J(x)(y - x)\| &= \left\| \int_0^1 J(x + t(y - x))(y - x)dt - J(x)(y - x) \right\| \\ &\leq \|y - x\| \int_0^1 \|J(x + t(y - x)) - J(x)\|dt \\ &\leq L_1\|y - x\|^2, \end{aligned} \tag{32}$$

and

$$\|F(y) - F(x)\| \leq L_2\|y - x\|, \forall x, y \in N(x^*, b_1), \tag{33}$$

where L_2 is a positive constant.

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

$$\|\bar{x}_k - x_k\| = \text{dist}(x_k, X^*), \forall x, y \in N(x^*, b_1). \tag{34}$$

To obtain the local convergence rate of x_k, we present some lemmas.

Lemma 2. *Under the conditions of Assumption 2, for all sufficiently large k, there exists a constant $c_2 > 0$ such that*

$$\|d_k\| \leq c_2\|\bar{x}_k - x_k\|. \tag{35}$$

Proof. According to (34), we have

$$\|\bar{x}_k - x^*\| \leq \|\bar{x}_k - x_k\| + \|x_k - x^*\| \leq 2\|x_k - x^*\| \leq b_1, \tag{36}$$

which means that $\bar{x} \in N(x^*, b_1)$. Following from (13),

$$\begin{aligned} \lambda_k = \frac{\mu_k\|F_k\|}{1 + \|F_k\|} &= \mu_k\left(1 - \frac{1}{1 + \|F_k\|}\right) \\ &\geq m\left(1 - \frac{1}{1 + c_1\|\bar{x}_k - x_k\|}\right) \\ &= \frac{mc_1\|\bar{x}_k - x_k\|}{1 + c_1\|\bar{x}_k - x_k\|}, \end{aligned} \tag{37}$$

and we have from (32) that

$$\|F_k + J_k(\bar{x}_k - x_k)\|^2 = \|F(\bar{x}_k) - F_k - J_k(\bar{x}_k - x_k)\|^2 \leq L_1^2\|\bar{x}_k - x_k\|^4. \tag{38}$$

As d_k is a minimizer of $\psi_k(d)$, we have

$$\begin{aligned}\|d_k\|^2 &\leq \frac{1}{\lambda_k}\varphi_k(d_k)\\ &\leq \frac{1}{\lambda_k}\varphi_k(\bar{x}_k - x_k)\\ &= \frac{1}{\lambda_k}(\|F_k + J_k(\bar{x}_k - x_k)\|^2 + \lambda_k\|\bar{x}_k - x_k\|^2)\\ &\leq \frac{1 + c_1\|\bar{x}_k - x_k\|}{mc_1\|\bar{x}_k - x_k\|}(L_1^2\|\bar{x}_k - x_k\|^4) + \|\bar{x}_k - x_k\|^2\\ &= O(\|\bar{x}_k - x_k\|^2),\end{aligned}$$

then there exists a constant $c_2 > 0$ such that $\|d_k\| \leq c_2\|\bar{x}_k - x_k\|$. The proof is completed. □

Lemma 3. *Under the conditions of Assumption 2, for all sufficiently large k, there exists a positive constant $M > m$ such that*

$$\mu_k \leq M. \tag{39}$$

Proof. First, we show that for sufficiently large k, the following inequality holds

$$Pred_k = \|F_k\|^2 - \|F_k + J_k d_k\|^2 \geq \min\left\{\frac{c_1}{2c_2}, \frac{c_1}{2}\right\}\|F_k\|\|d_k\|. \tag{40}$$

We consider two cases. In one case, if $\|\bar{x}_k - x_k\| \leq \|d_k\|$, then the definition of d_k and Assumption 2 imply that

$$\begin{aligned}\|F_k\| - \|F_k + J_k d_k\| &\geq \|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|\\ &\geq c_1\|\bar{x}_k - x_k\| - L_1\|\bar{x}_k - x_k\|^2\\ &\geq \frac{c_1}{2c_2}\|d_k\|.\end{aligned} \tag{41}$$

In the other case, if $\|\bar{x}_k - x_k\| > \|d_k\|$, then we have

$$\begin{aligned}\|F_k\| - \|F_k + J_k d_k\| &\geq \|F_k\| - \left\|F_k + \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}J_k(\bar{x}_k - x_k)\right\|\\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}(\|F_k\| - \|F_k + J_k(\bar{x}_k - x_k)\|)\\ &\geq \frac{\|d_k\|}{\|\bar{x}_k - x_k\|}\left(c_1\|\bar{x}_k - x_k\| - L_1\|\bar{x}_k - x_k\|^2\right)\\ &\geq \frac{c_1}{2}\|d_k\|.\end{aligned} \tag{42}$$

Inequalities (41) and (42), together with Lemma 2 show that

$$\begin{aligned}Pred_k &= (\|F_k\| + \|F_k + J_k d_k\|)(\|F_k\| - \|F_k + J_k d_k\|)\\ &\geq \|F_k\|(\|F_k\| - \|F_k + J_k d_k\|)\\ &\geq \min\left\{\frac{c_1}{2c_2}, \frac{c_1}{2}\right\}\|F_k\|\|d_k\|,\end{aligned} \tag{43}$$

which gives (40). Hence, it follows from (40), Assumption 2 and Lemma 2 that

$$\begin{aligned} |r_k - 1| &= \left| \frac{Ared_k}{Pred_k} - 1 \right| \\ &= \frac{\|F_k + J_k d_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{Pred_k} \\ &\leq \frac{\|F_k\| O(\|d_k\|^2) + O(\|d_k\|^4)}{O(\|F_k\| \|d_k\|)} \\ &= O(\|d_k\|) \to 0. \end{aligned}$$

Therefore, we have $r_k \to 1$, thus, there exists a constant $M > m$ such that $\mu_k \leq M$ for all large k. The proof is completed. □

Without generality, we assume $\text{rank}(J(x^*)) = r$ for all $\bar{x} \in N(x^*, b_1) \cap X^*$. Suppose the SVD of $J(\bar{x})$ is

$$J(\bar{x}) = [\bar{U}_1, \bar{U}_2] \begin{bmatrix} \bar{\Sigma}_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} \bar{V}_1^{\mathrm{T}} \\ \bar{V}_2^{\mathrm{T}} \end{bmatrix} = \bar{U}_1 \bar{\Sigma}_1 \bar{V}_1^{\mathrm{T}}, \tag{44}$$

where $\bar{\Sigma}_1 = \text{diag}(\bar{\sigma}_1, \bar{\sigma}_2, \ldots, \bar{\sigma}_r)$ with $\bar{\sigma}_1 \geq \bar{\sigma}_2 \geq \cdots \geq \bar{\sigma}_r > 0$ and $\bar{U} = [\bar{U}_1, \bar{U}_2]$, $\bar{V} = [\bar{V}_1, \bar{V}_2]$ are orthogonal matrices. Correspondingly, we consider SVD of $J(x_k)$ by

$$J(x_k) = [U_1, U_2, U_3] \begin{bmatrix} \Sigma_1 & 0 & 0 \\ 0 & \Sigma_2 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} V_1^{\mathrm{T}} \\ V_2^{\mathrm{T}} \\ V_3^{\mathrm{T}} \end{bmatrix} = U_1 \Sigma_1 V_1^{\mathrm{T}} + U_2 \Sigma_2 V_2^{\mathrm{T}}, \tag{45}$$

where $U = [U_1, U_2, U_3]$, $V = [V_1, V_2, V_3]$ are orthogonal matrixes, $\Sigma_1 = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_r)$ with $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ and $\Sigma_2 = \text{diag}(\sigma_{r+1}, \sigma_{r+2}, \cdots, \sigma_{r+q})$ with $\sigma_{r+1} \geq \sigma_{r+2} \geq \cdots \geq \sigma_{r+q} > 0$.

Lemma 4. *Under the conditions of Assumption 2, for all sufficiently large k, we have*

$(a) \|U_1 U_1^{\mathrm{T}} F_k\| \leq O(\|\bar{x}_k - x_k\|)$;
$(b) \|U_2 U_2^{\mathrm{T}} F_k\| \leq O(\|\bar{x}_k - x_k\|^2)$;
$(c) \|U_3 U_3^{\mathrm{T}} F_k\| \leq O(\|\bar{x}_k - x_k\|^2)$;
$(d) \|F_k + J_k d_k\| \leq O(\|\bar{x}_k - x_k\|^2)$.

Proof. The result (a) follows immediately from (16). By (15) and the theory of matrix perturbation [26], we have

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

which implies that

$$\|\Sigma_1 - \bar{\Sigma}_1\| \leq L_1 \|\bar{x}_k - x_k\| \text{ and } \|\Sigma_2\| \leq L_1 \|\bar{x}_k - x_k\|. \tag{46}$$

Let $s_k = -J_k^+ F_k$, where J_k^+ is the pseudo-inverse of J_k. It is easy to see that s_k is the least-squares solution of $\min \|F_k + J_k s\|$, so we obtain from (32) that

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

Let $\bar{J}_k = U_1\Sigma_1 V_1^{\mathrm{T}}$ and $\bar{s}_k = -\bar{J}_k^{+}F_k$. Since $\bar{s}_k$ is the least-squares solution of $\min \|F_k + \bar{J}_k s\|$, it follows from (32) that

$$\begin{aligned}\|(U_2U_2^{\mathrm{T}} + U_3U_3^{\mathrm{T}})F_k\| &= \|F_k + \bar{J}_k\bar{s}_k\| \\ &\leq \|F_k + \bar{J}_k(\bar{x}_k - x_k)\| \\ &\leq \|F_k + J_k(\bar{x}_k - x_k)\| + \|(\bar{J}_k - J_k)(\bar{x}_k - x_k)\| \\ &\leq L_1\|\bar{x}_k - x_k\|^2 + \|U_2\Sigma_2 V_2^{\mathrm{T}}(\bar{x}_k - x_k)\| \\ &\leq L_1\|\bar{x}_k - x_k\|^2 + L_1\|\bar{x}_k - x_k\|\|\bar{x}_k - x_k\| \\ &\leq O(\|\bar{x}_k - x_k\|^2).\end{aligned}$$

Due to the orthogonality of U_2 and U_3, we obtain the result (b).
Using (12) and (45), we obtain

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

and

$$\begin{aligned}F_k + J_k d_k &= F_k - U_1\Sigma_1(\Sigma_1^2 + \lambda_k I)^{-1}\Sigma_1 U_1^{\mathrm{T}}F_k - U_2\Sigma_2(\Sigma_2^2 + \lambda_k I)^{-1}\Sigma_2 U_2^{\mathrm{T}}F_k \\ &= \lambda_k U_1(\Sigma_1^2 + \lambda_k I)^{-1}U_1^{\mathrm{T}}F_k + \lambda_k U_2(\Sigma_2^2 + \lambda_k I)^{-1}U_2^{\mathrm{T}}F_k + U_3U_3^{\mathrm{T}}F_k.\end{aligned}$$

Following from (13) and (33), the LM parameter satisfies

$$\lambda_k = \frac{\mu_k\|F_k\|}{1 + \|F_k\|} \leq \mu_k\|F_k\| \leq ML_2\|\bar{x}_k - x_k\|.$$

Since $\{x_k\}$ converges to the solution set X^*, we assume that $L_1\|\bar{x}_k - x_k\| \leq \frac{\bar{\sigma}_r}{2}$ holds for all sufficiently large k. Then, it follows from (46) that

$$\|\Sigma_1^{-1}\| \leq \frac{1}{\bar{\sigma}_r - L_1\|\bar{x}_k - x_k\|} \leq \frac{2}{\bar{\sigma}_r}.$$

It then follows from Lemmas 3 and 4 that

$$\begin{aligned}\|F_k + J_k d_k\| &\leq \lambda_k\|\Sigma_1^{-2}\|\|U_1^{\mathrm{T}}F_k\| + \|U_2^{\mathrm{T}}F_k\| + \|U_3U_3^{\mathrm{T}}F_k\| \\ &\leq \frac{4L_2M\|\bar{x}_k - x_k\|^2}{\bar{\sigma}_r^2} + O(\|\bar{x}_k - x_k\|^2) + O(\|\bar{x}_k - x_k\|^2) \\ &= O(\|\bar{x}_k - x_k\|^2).\end{aligned} \tag{47}$$

The proof is completed. □

We can state the quadratic convergence result of the NAALM algorithm.

Theorem 2. *Let the sequence $\{x_k\}$ be generated by the NAALM algorithm, under Assumption 2, the sequence $\{x_k\}$ converges quadratically to a solution of nonlinear Equation (1).*

Proof. It follows from Assumption 2, Lemma 2 and (47) that

$$\begin{aligned}c_1\|\bar{x}_{k+1} - x_{k+1}\| &\leq \|F(x_{k+1})\| \\ &= \|F(x_k + d_k)\| \\ &\leq \|F_k + J_k d_k\| + O(\|d_k\|^2) \\ &= O(\|\bar{x}_k - x_k\|^2).\end{aligned} \tag{48}$$

On the other hand, it is clear that

$$\|\bar{x}_k - x_k\| = \mathrm{dist}(x_k, X^*) \leq \|\bar{x}_{k+1} - x_k\| \leq \|\bar{x}_{k+1} - x_{k+1}\| + \|d_k\|.$$

It follows from Lemma 2 that, for any sufficiently large k, we have

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

Therefore, $\|d_k\| = O(\|\bar{x}_k - x_k\|)$. This, along with (48), indicates that

$$\|d_{k+1}\| \leq O(\|d_k\|^2),$$

which implies that $\{x_k\}$ is quadratically convergent to a solution of set X^*. The proof is completed. □

3. Numerical Results

In this section, the numerical performance of NAALM algorithm will be listed. All codes were written in MATLAB R2016b on a PC with 1.19 GHz, 8.00 GB RAM, using Windows 11 operation system. In this section, we will expand on the following two aspects. On the one hand, the effectiveness of the NAALM algorithm is illustrated by comparing it with other algorithms on some test questions. On the other hand, it shows that the NAALM algorithm has good development prospects by applying the algorithm to a fresh agricultural products supply chain problem.

3.1. Some Singular Nonlinear Equations Problems

The test problems are constructed by modifying the nonsingular problems given by Moré et al. [27], which have the following form as [28]:

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

where $F(x)$ is the standard test function, $A \in \mathbb{R}^{n\times k}$ has full column rank with $0 \leq k \leq n$ and x is a solution of the equation $F(x) = 0$. According to the definition of $\hat{F}(x)$, we obtain

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

where $\hat{J}(x^*)$ is Jacobian matrix of $F(x)$ at x^* with rank $n-k$ and $\hat{F}(x^*) = 0$. In our test problems, some of $\hat{J}(x^*)$ are symmetric matrices and some are non-symmetric matrices. Note that some roots of $\hat{F}(x)$ may not be roots of $F(x)$. Similar to [28], we construct two sets of singular problems while $\hat{J}(x^*)$ have rank $n-1$ or $n-2$, by choosing

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

and

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & ... & 1 \\ 1 & -1 & 1 & -1 & ... & \pm 1 \end{bmatrix}^{\mathrm{T}} \in \mathbb{R}^{n\times 2}.$$

We test our NAALM algorithm on some singular nonlinear equations, and compare it with the self-adaptive Levenberg–Marquardt algorithm (SLM) proposed in [18]. The main differences between these two algorithms are in the updating rule of μ_k.

We set $p_0 = 10^{-4}$, $p_1 = \frac{1}{4}$, $p_2 = \frac{3}{4}$, $\beta_1 = \frac{5}{4}$, $\beta_2 = \frac{1}{3}$, $\beta_3 = \frac{6}{5}$, $m = 10^{-8}$, $\mu_0 = 10^{-2}$, for all the tests. All test methods are terminated when $\|J_k^{\mathrm{T}}F_k\| \leq 10^{-5}$. The algorithm is considered to fail when the number of iterations exceeds 500. Considering the global convergence of the algorithms, we run each test problem for five starting points, $-10x_0$, $-x_0$, x_0, $10x_0$ and $100x_0$, where x_0 is given by [28]. For n as a variable, we take $n = 500$, $n = 1000$, respectively.

The performance profile of two algorithms. including the number of iterations (NI), function evaluations (NF), gradient evaluations (NG) and CPU time (CPU), is analyzed using the profiles of Dolan and Moré [29]. Let Y and W be the set of methods and test problems, n_y, n_w be the number of methods and test problems, respectively. The performance profile $\psi : \mathbb{R} \to [0, 1]$ is for each $y \in Y$ and $w \in W$ defined that $a_{w,y} > 0$ is NI or NF or NG

or CPU required to solve problems w by method y. Furthermore, the performance profile is obtained by

$$\psi_y(\tau) = \frac{1}{n_w} size\{w \in W : log_2 r_{w,y} \leq \tau\},$$

where $\tau > 0$, $size\{\cdot\}$ is the number of the elements in a set, and $r_{w,y}$ is the performance ratio defined as

$$r_{w,y} = \frac{a_{w,y}}{min\{a_{w,y} : w \in W \text{ and } y \in Y\}}.$$

Generally, the method whose performance profile plot is on the top right will represent the best method.

As can be seen from Figure 1, the NAALM algorithm is better than the SLM algorithm in terms of the number of iterations, especially when $\tau > 2$, the curve of NAALM algorithm becomes stable, which indicates that NAALM algorithm can solve the problem only with fewer iterations. In terms of function evaluations, as shown in Figure 2, the NAALM algorithm curve in $\tau > 1.75$, it has reached a stable state, while SLM algorithm can reach a stable state only when the curve coincides with that of NAALM algorithm at $\tau > 2.75$; Figure 3 shows the performance diagram of the SLM algorithm and the NAALM algorithm in the Jacobian matrix. It can be seen that the NAALM algorithm can successfully solve test problems up to 98%, while SLM can only reach 94%, which shows that the NAALM algorithm can reduce the calculation times of the Jacobian matrix and save the calculation amount. Figure 4 shows the CPU time performance of the NAALM algorithm and the SLM algorithm. It can be seen from the figure that when $\tau < 4.5$, the curves of the NAALM algorithm and the SLM algorithm are similar, but when $\tau > 4.5$, both the NAALM algorithm and the SLM algorithm tend to be stable and coincide. Therefore, Figures 1–4 show that the accelerated version of the LM algorithm proposed in this paper can not only converge to the solution quickly, but also reduce the computation amount of the Jacobian matrix.

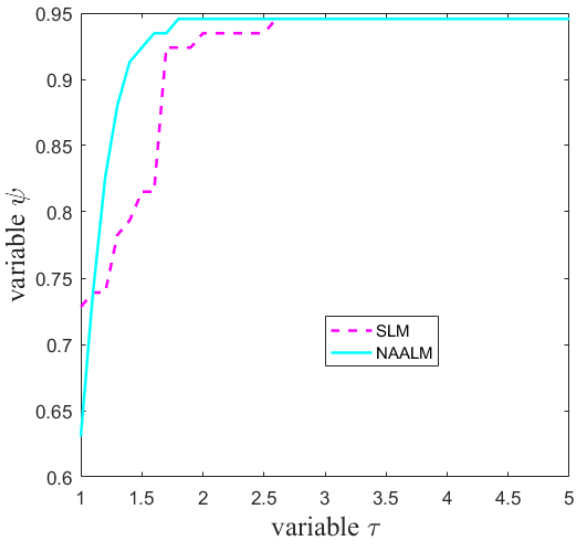

Figure 1. Performance profiles for the iterations.

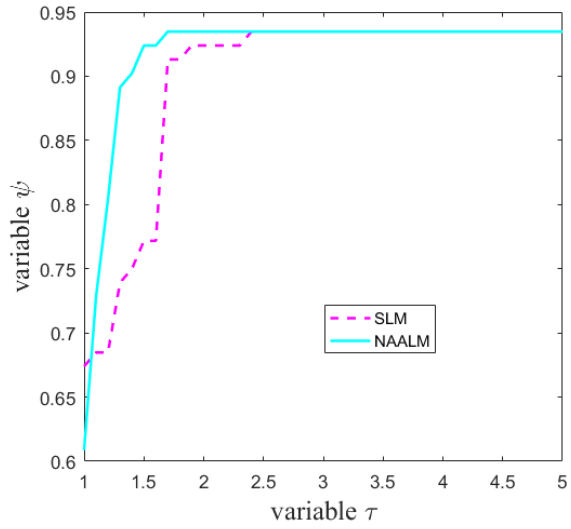

Figure 2. Performance profiles for the function evaluations.

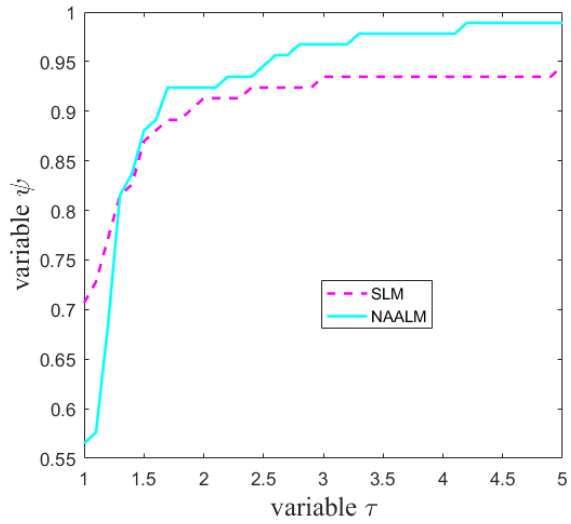

Figure 3. Performance profiles for the gradient evaluations.

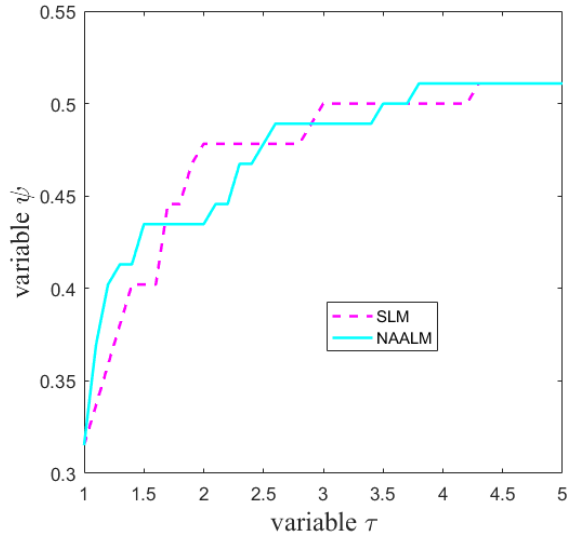

Figure 4. Performance profiles for the CPU time.

3.2. Supply Chain Optimization Problems

The security and stability of the supply chain has a great impact on promoting high-quality and sustainable development of the economy. Therefore, supply chain has been applied to many fields, such as low-carbon supply chain, manufacturing green supply chain, food trade supply chain. In recent years, with the improvement of living standards, the quality of fresh agricultural products has attracted widespread attention from consumers. In order to meet the demand of consumers for high quality and low price of fresh agricultural products, we use the NAALM algorithm to study how suppliers and retailers make decisions to maximize both their own profits and the total profit of the fresh agricultural products supply chain under the decentralized policy.

In this supply chain, as the leader of Stackelberg game, fresh agricultural product suppliers supply the same variety of ordinary fresh agricultural products (ofp) and green fresh agricultural products (gfp) to retailers as followers, while retailers sell them to consumers. Suppliers need to choose the optimal wholesale price strategy of two fresh agricultural products, and retailers need to choose the optimal retail price strategy of two fresh agricultural products and determine the order quantity of two fresh agricultural products by market demand.

Without considering the impact of emergencies, the market demand for fresh agricultural products is relatively stable, and it is only related to price and freshness. Due to the substitution of the same varieties of ofp and gfp, there is a competitive relationship in the demand market. Based on the demand function theory of alternative price competition, it is assumed that the demand function of two fresh agricultural products is as follows

$$q_i = a - b\frac{p_i}{\theta} + r\frac{p_j}{\theta}, i = 1, 2, j = 3 - i, \tag{49}$$

where q_1, q_2 represent the market demand of gfp and ofp, respectively, a represents the total potential market capacity of fresh agricultural products, p_1, p_2 represent the retail price of gfp and ofp, respectively, b is the price sensitivity coefficient, r is the competitive substitution coefficient of the two products, and it satisfies $b > r > 0$, $\theta(0 \leq \theta \leq 1)$is the freshness of fresh produce when it arrives at the retailer's store.

Under the decentralized policy, we regard suppliers and retailers as independent entities, and both with the goal of maximizing their respective interests. Now, the profit function of fresh agricultural products retailer is as follows

$$\max_{p_1,p_2} \pi_s = (p_1 - w_1)(a - b\frac{p_1}{\theta} + r\frac{p_2}{\theta}) + (p_2 - w_2)(a - b\frac{p_2}{\theta} + r\frac{p_1}{\theta}), \tag{50}$$

where w_1, w_2 represent the supply price of gfp and ofp, respectively, and the profit function of fresh agricultural products suppliers is as follows

$$\max_{w_1,w_2} \pi_R = (w_1 - \frac{c_1}{1-\beta})(a - b\frac{p_1}{\theta} + r\frac{p_2}{\theta} + (w_2 - \frac{c_2}{1-\beta})(a - b\frac{p_2}{\theta} + r\frac{p_1}{\theta}), \tag{51}$$

where β $(0 < \beta < 1)$ represents the quantity loss of fresh produce when it reaches the retailer's store, c_1, c_2 represents the unit production cost of gfp and ofp, respectively. Obviously, $p_1 > p_1 > 0$ and $c_1 > c_2 > 0$. We record the total profit of fresh agricultural products supply chain as follows:

$$\pi_T = \pi_s + \pi_R. \tag{52}$$

With reference to the setting of the parameters in the relevant literature [30], we set: $a = 50$, $b = 2$, $c_1 = 4$, $c_2 = 2$, $r = 1.5$, $\beta = 0.2$, $\theta = 0.85$. These values satisfy the theoretical proof in [30] and can guarantee that the optimal value has practical significance. Now, we transform the unconstrained optimization problem (51) into a nonlinear equation problem, and then choose different initial points and use the NAALM algorithm to solve the nonlinear equation problem.

As can be seen from Table 1, with certain parameters, the NAALM algorithm can be used to solve the optimization problem, so as to obtain the optimal pricing strategy with maximum profit in the supply chain led by suppliers under the decentralized policy. In addition, the global convergence and robustness of the NAALM algorithm are verified according to different initial values and the number of iterations.

Table 1. The optimal solution corresponding to different initial points by NAALM.

Initial Point	p_1	p_2	w_1	w_2	q_1	q_2	π_T
(1;1;1;1)	45.0266	43.7205	65.065	64.329	10.5818	13.2479	1.4587×10^3
(10;10;10;10)	44.9648	43.7460	64.991	64.358	10.5163	13.2607	1.4587×10^3
(30;30;30;30)	44.9920	43.7496	64.956	64.376	10.5825	13.3037	1.4587×10^3
(50;50;50;50)	45.0158	43.7525	65.062	64.372	10.5125	13.3036	1.4587×10^3
(100;100;100;100)	44.9751	43.7553	65.955	64.269	10.5222	13.2679	1.4587×10^3

4. Conclusions

We constructed a new function that makes full use of the ratio information to update LM parameters adaptively. Based on this new LM parameter, we presented an adaptive accelerated Levenberg–Marquardt method for solving nonlinear equations. Furthermore, we showed the global convergence analysis of the proposed algorithm. Furthermore, the quadratic convergence is also obtained under the local error bound condition. Numerical experiments demonstrated that our method has good numerical performance. In addition, the application of the NAALM algorithm to a supply chain problem showed that the new algorithm has a good application prospect. We further highlight that the proposed NAALM algorithm can be used in other fields, such as the symmetric system of nonlinear equations. It is vital to note that the method's convergence analysis in Hölderian local error bound condition will be taken into account in our future work.

Author Contributions: Conceptualization, R.L., M.C. and G.Z.; methodology, R.L., M.C. and G.Z.; software, R.L., M.C. and G.Z. validation, R.L., M.C. and G.Z.; formal analysis, R.L., M.C. and G.Z.; investigation, R.L., M.C. and G.Z.; resources, R.L., M.C. and G.Z.; data curation, R.L., M.C. and G.Z.; writing—original draft preparation, R.L., M.C. and G.Z.; writing—review and editing, R.L., M.C. and G.Z. All authors have read and agreed to the published version of the manuscript.

Funding: This work is supported by the key project of natural science foundation joint fund of Jilin Province (YDZJ202101ZYTS167, YDZJ202201ZYTS303); the project of education department of Jilin Province (JJKH20210030KJ, JJKH20230054KJ); the graduate innovation project of Beihua University (2022033, 2021002); the youth science and technology innovation team cultivation program of Beihua University.

Data Availability Statement: Not applicable.

Acknowledgments: The authors would like to thank the anonymous referees and editor for reading this paper carefully, providing valuable suggestions and comments, which grately improved the final version.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Musa, Y.B.; Waziri, M.Y.; Noor, M.A. An efficient method for solving system for nonlinear equation. *J. Math. Anal.* **2022**, *13*, 1–10.
2. Ribeiro, S.; Lopes, L.G. Overview and computational analysis of PSO variants for solving systems of nonlinear equations. *Commun. Intell. Syst.* **2022**, *461*, 1093–1105.
3. Ji, J.Y.; Wong, M.L. Decomposition-based multiobjective optimization for nonlinear equation systems with many and infinitely many roots. *Inf. Sci.* **2022**, *610*, 605–623. [CrossRef]
4. Artacho, F.J.A.; Fleming, R.; Vuong, P.T. Accelerating the DC algorithm for smooth functions. *Math. Program.* **2018**, *169B*, 95–118. [CrossRef]

5. Sabi'u, J.; Muangchoo, K.; Shah, A.; Abubakar, A.B.; Aremu, K.O. An inexact optimal hybrid conjugate gradient method for solving symmetric nonlinear equations. *Symmetry* **2021**, *13*, 1829. [CrossRef]
6. Sabi'u, J.; Muangchoo, K.; Shah, A.; Abubakar, A.B.; Jolaoso, L.O. A modified PRP-CG type derivative-free algorithm with optimal choices for solving large-scale nonlinear symmetric equations. *Symmetry* **2021**, *13*, 234. [CrossRef]
7. Niri, T.D.; Heydari, M.; Hosseini, M.M. Correction of trust region method with a new modified Newton method. *Int. J. Comput. Math.* **2022**, *97*, 1–15.
8. Bellavia, S.; Morini, B.; Rebegoldi, S. On the convergence properties of a stochastic trust-region method with inexact restoration. *Axioms* **2023**, *12*, 38. [CrossRef]
9. Zheng, L.; Chen, L.; Ma, Y.F. A variant of the Levenberg-Marquardt method with adaptive parameters for systems of nonlinear equations. *AIMS Math.* **2021**, *7*, 1241–1256. [CrossRef]
10. Yudin, N.E. Adaptive Gauss-Newton method for solving systems of nonlinear equations. *Dokl. Math.* **2021**, *104*, 293–296. [CrossRef]
11. Levenberg, K. A method for the solution of certain nonlinear problems in least squares. *Quart. Appl. Math.* **1944**, *2*, 164–166. [CrossRef]
12. Marquardt, D.W. An algorithm for least-squares estimation of nonlinear inequalities. *SIAM J. Appl. Math.* **1963**, *11*, 431–441. [CrossRef]
13. Yamashita, N.; Fukushima, M. On the Rate of Convergence of the Levenberg-Marquardt Method. *Computing* **2001**, *15*, 239–249.
14. Fan, J.Y.; Yuan, Y.X. *On the Convergence of a New Levenberg-Marquardt Method*; Report No. 005, AMSS; Chinese Academy of Sciences: Beijing, China, 2001.
15. Fan, J.Y. A modified Levenberg-Marquardt algorithm for singular system of nonlinear equation. *J. Comput. Math.* **2003**, *21*, 625–636.
16. Fan, J.Y.; Pan, J.Y. A note on the Levenberg-Marquardt parameter. *Appl. Mathe. Comput.* **2009**, *207*, 351–359. [CrossRef]
17. Amini, K.; Rostami, F.; Caristi, G. An efficient Levenberg-Marquardt method with a new LM parameter for systems of nonlinear equations. *Optimization* **2018**, *67*, 637–650. [CrossRef]
18. Fan, J.Y.; Yuan, Y.X. Convergence properties of a self-adaptive Levenberg-Marquardt algorithm under local error bound condition. *Comput. Optim. Appl.* **2006**, *34*, 47–62. [CrossRef]
19. Hei, L. A self-adaptive trust region algorithm. *J. Comput. Math.* **2003**, *21*, 229–236.
20. Walmag, J.M.B.; Delhez, E.J.M. A note on trust-region radius update. *Siam J. Optim.* **2005**, *16*, 548–562. [CrossRef]
21. Lu, Y.L.; Li, W.Y.; Cao, M.Y.; Yang, Y.T. A novel self-adaptive trust region algorithm for unconstrained optimization. *J. Appl. Math.* **2014**, *2014*, 1–8. [CrossRef]
22. Amini, K.; Rostami, F. A modified two steps Levenberg-Marquardt method for nonlinear equations. *J. Comput.* **2015**, *288*, 341–350. [CrossRef]
23. He, X.R.; Tang, J.Y. A smooth Levenberg-Marquardt method without nonsingularity condition for wLCP. *AIMS Math.* **2022**, *7*, 8914–8932. [CrossRef]
24. Fan, J.Y. Accelerating the modified Levenberg-Marquardt method for nonlinear equations. *Math. Comput.* **2014**, *83*, 1173–1187. [CrossRef]
25. Chen, L. A high-order modified Levenberg-Marquardt method for systems of nonlinear equations with fourth-order convergence. *Appl. Math. Comput.* **2016**, *285*, 79–93. [CrossRef]
26. Stewart, G.W.; Sun, J.G. *Matrix Perturbation Theory*; Academic Press: San Diego, CA, USA, 1990.
27. More, J.J. Recent developments in algorithms and software for trust region methods. *Math. Program.* **1983**, *85*, 258–287.
28. Schnabel, R.B.; Frank, P.D. Tensor methods for nonlinear equations. *SIAM J. Numer. Anal.* **1984**, *21*, 815–843. [CrossRef]
29. Dolan, E.; More, J.J. Benchmarking optimization software with performance profiles. *Math. Program.* **2022**, *91*, 201–213. [CrossRef]
30. Wen, H. Research on Profit Maximization Strategy of Fresh Agricultural Products Supply Chain under Different Dominated Subjects. Ph.D. Thesis, Huazhong Agricultural University, Wuhan, China, 2020.