J. Appl. Math. & Informatics Vol. **33***(2015), No. 1 - 2, pp. 127 - 143*
http://dx.doi.org/10.14317/jami.2015.127

A MODIFIED INEXACT NEWTON METHOD†

PENGZHAN HUANG* AND ABDURISHIT ABDUWALI

Abstract. In this paper, we consider a modified inexact Newton method for solving a nonlinear system $F(x) = 0$ where $F(x) : R^n \to R^n$. The basic idea is to accelerate convergence. A semi-local convergence theorem for the modified inexact Newton method is established and an affine invariant version is also given. Moreover, we test three numerical examples which show that the modified inexact scheme is more efficient than the classical inexact Newton strategy.

AMS Mathematics Subject Classification : 65H10, 49M15.
Key words and phrases : Modified inexact Newton method, Nonlinear equations, Semi-local convergence, Affine invariant.

1. Introduction

Consider the system of nonlinear equations

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

where $F(x) : D \subset R^n \to R^n$ is Fréchet differentiable. Let $F'(x)$ denote the Fréchet derivative of F at x.

Such equations (1.1) often arise in many important practical fields (e.g., physics and engineering, etc.). For example, input-output systems, least squares problems, finite difference or finite element problems, integral or differential equations, constrained function minimization, complementarity problems, variational inequalities, calculation of the load flows for power systems and solving initial or boundary value problems in ordinary or partial differential equations, etc.

Among all kinds of numerical methods for solving the nonlinear equations (1.1), Newton method [18, 24, 33] is the most classical one. In general, suppose that x_k is the current approximate solution; a new approximate solution x_{k+1} can be computed through the following general form:

Received March 5, 2014. Revised August 15, 2014. Accepted August 17, 2014. *Corresponding author. †This work was supported by the China Postdoctoral Science Foundation (Grant Nos. 2014T70954, 2013M530438).

Algorithm 1.1 : Newton method
1. Let $x_0 \in R^n$ be a given initial guess.
2. For $k = 0$ until convergence do.
    2.1. For the iteration x_k, find the step s_k satisfying
$$F'(x_k)s_k = -F(x_k). \tag{1.2}$$
    2.2. Set $x_{k+1} = x_k + s_k$.
    2.3. Set $k = k + 1$ and turn to 2.1.

If n is not too large, the Newton method is attractive because it converges rapidly from any sufficiently good initial data. However, Newton method has two disadvantages from the point of view of practical computation: one is that it requires computing Jacobian matrices, and the other is that it requires solving linear equations (1.2) exactly. Computing the exact solution using a direct method such as Gaussian elimination may be expensive if the Jacobian matrix is large and may not be justified when x_k is far from the exact solution x^*. In order to overcome the disadvantage of Newton method, using an iterative method and solving (1.2) approximately are reasonable. This approach was first considered by Dembo, Eisenstat and Steihaug in [11] (such a variant is the so-called inexact Newton method).

Algorithm 1.2 : Inexact Newton method
1. Let $x_0 \in R^n$ be a given initial guess.
2. For $k = 0$ until convergence do.
    2.1. Choose $\eta_k \in [0, 1)$.
    2.2. For the residual r_k and the iteration x_k, find the step s_k satisfying
$$F'(x_k)s_k = -F(x_k) + r_k, \tag{1.3}$$
    where
$$\frac{\|r_k\|}{\|F(x_k)\|} \le \eta_k. \tag{1.4}$$
    2.3. Set $x_{k+1} = x_k + s_k$.
    2.4. Set $k = k + 1$ and turn to 2.1.

Remark 1.1. In the above algorithm, s_k is the inexact Newton step and (1.4) is the inexact Newton condition. η_k is the forcing term for the k-th iteration step which may depend on x_k; taking $\eta_k \equiv 0$ gives the famous Newton method.

In typical applications, the choice of the forcing terms is critical to the efficiency of the method and can affect robustness as well. Usually, it is hard to choose a good sequence of forcing terms. In computational practice, several authors considered some concrete strategies. We list here the following:
1. The choice $\eta = 10^{-1}$ used by Elias, Coutinho and Martins [16].
2. The choice $\eta = 10^{-4}$ used by Cai, Gropp, Keyes and Tidriri [9].
3. The choice $\frac{1}{2^{k+1}}$ of Brown and Saad [7].
4. The choice $\eta = \min\{\frac{1}{k+2}, \|F(x_k)\|\}$ of Dembo and Steihaug [12].
5. Eisenstat and Walker [15] proposed two choices:

(1). Given $\eta_0 \in [0,1)$, choose

$$\eta_k = \frac{\|F(x_k) - F(x_{k-1}) - F'(x_{k-1})s_{k-1}\|}{\|F(x_{k-1})\|},\ k = 1, 2, \ldots,$$

or

$$\eta_k = \frac{|\ \|F(x_k)\| - \|F(x_{k-1}) + F'(x_{k-1})s_{k-1}\|\ |}{\|F(x_{k-1})\|},\ k = 1, 2, \ldots.$$

(2). Given $\gamma \in [0,1]$, $\alpha \in (1,2]$ and $\eta_0 \in [0,1)$, choose

$$\eta_k = \gamma \left(\frac{\|F(x_k)\|}{\|F(x_{k-1})\|} \right)^{\alpha},\ k = 1, 2, \ldots.$$

6. H.B. An, Z.Y. Mo and X.P. Liu [1] choosed forcing terms by the following way:

$$\eta_k = \begin{cases} 1 - 2p_1, & r_{k-1} < p_1, \\ \eta_{k-1}, & p_1 \le r_{k-1} < p_2, \\ 0.8\eta_{k-1}, & p_2 \le r_{k-1} < p_3, \\ 0.5\eta_{k-1}, & r_{k-1} \ge p_3, \end{cases} \quad k = 1, 2, \ldots,$$

where $0 < p_1 < p_2 < p_3 < 1$ are prescribed at first, and $p_1 \in (0, \frac{1}{2})$. In addition, assume that η_0 is given and $r_k = \frac{\|F(x_k)\| - \|F(x_k + s_k)\|}{\|F(x_k)\| - \|F(x_k) + F'(x_k)s(x_k)\|}$.

There are three types of convergence issues about inexact Newton method: global, local and semi-local convergence analysis. The first is the convergence analysis based on the whole domain, the second is the convergence analysis based on a neighborhood of the solution x^*, and the last is the convergence analysis based on a neighborhood of the initial guess x_0. Recently, several authors have studied the global convergence (see [14, 31]), local convergence (see [10, 20, 23, 34]) and semi-local convergence (see [2, 3, 4, 19, 22, 29, 30, 32, 36]) of inexact Newton method and proposed application in different fields [5, 21, 28].

After this method is established, some iteration methods are considered by many authors based on it. AIN (Accelerated Inexact Newton) method is presented by Fokkema, Sleijpen and Voest in [17]. They have shown how the classical Newton iteration scheme for nonlinear problems can be accelerated in a similar way as standard Richardson-type iteration schemes for linear equation. REGINN (Regularization based on Inexact Newton iteration) method is presented by Rieder in [25, 26, 27]. INM (Inexact Newton Multigrid) is considered by Brown, Vassilevski and Woodward in [6]. They have proved optimal-order and mesh-independent convergence of an inexact Newton method where the linear Jacobian systems are solved with multigrid techniques. Also, PIN (Preconditioned Inexact Newton) method is considered by Cai and Keyes in [8].

In this paper we consider a modified inexact Newton method. Our modification uses $F'(\hat{x}_k)s_k = -F(x_k) + r_k$ to replace $F'(x_k)s_k = -F(x_k) + r_k$, where $\hat{x}_k := x_k - F'(\hat{x}_{k-1})^{-1}F(x_k)$. The use of this method is particularly appropriate when an exact solution of Newton equation is difficult to obtain and/or when evaluating and preparing the Jacobian for the computation is costly, and this method has fast convergence as well.

The rest of this paper is organized as follows. In section 2, a modified inexact Newton algorithm is established. In section 3, we will present the semi-local convergence result for the given algorithm. Moreover, another theorem shows the affine invariance of the convergence of our proposed method. In section 4, we give three test problems using the algorithm presented in this paper to show its convergence properties and robustness. In the last section, some concluding remarks are made.

2. A modified inexact Newton method

In this section, we introduce a modified inexact Newton method.

From the Algorithm 1.2, we know that the inexact Newton iteration scheme is

$$F'(x_k)(x_{k+1} - x_k) = -F(x_k) + r_k,\ k = 0, 1, \dots. \tag{2.1}$$

Note that (2.1) can be rewritten as

$$x_{k+1} = x_k - F'(x_k)^{-1}F(x_k) + F'(x_k)^{-1}r_k,\ k = 0, 1, \dots, \tag{2.2}$$

if $F'(x)$ is invertible for every $x \in D$.

Consider one-dimensional nonlinear equation

$$f(x) = 0, \tag{2.3}$$

where $f(x) : R \to R$ and $f'(x)$ is invertible for every $x \in R$. It easily follows from (2.2) that the inexact Newton iteration scheme of (2.3) is

$$x_{k+1} = x_k - f'(x_k)^{-1}f(x_k) + f'(x_k)^{-1}r_k,\ k = 0, 1, \dots. \tag{2.4}$$

If $\eta_k \equiv 0$, then r_k in (2.4) is 0. Moreover, (2.4) yields the famous Newton iteration scheme as follows

$$x_{k+1} = x_k - f'(x_k)^{-1}f(x_k),\ k = 0, 1, \dots. \tag{2.5}$$

In fact, from the point of view of geometry, Newton iteration scheme (2.5) uses tangent line of point $(x_k, f(x_k))$ to approximate curved line. We claim that, if the tangent line of the point $\left(x_k - f'(x_k)^{-1}f(x_k), f\left(x_k - f'(x_k)^{-1}f(x_k)\right)\right)$ is used to replace the former one, then a new scheme can be obtained as follows

$$x_{k+1} = x_k - f'\left(x_k - f'(x_k)^{-1}f(x_k)\right)^{-1}f(x_k),\ k = 0, 1, \dots. \tag{2.6}$$

Furthermore by generalizing this method to n-dimensional case, we have

$$x_{k+1} = x_k - F'\left(x_k - F'(x_k)^{-1}F(x_k)\right)^{-1}F(x_k),\ k = 0, 1, \dots. \tag{2.7}$$

For simplicity, let $\hat{x}_k = x_k - F'(x_k)^{-1}F(x_k)$. Hence, (2.7) reduces to

$$x_{k+1} = x_k - F'(\hat{x}_k)^{-1}F(x_k),\ k = 0, 1, \dots. \tag{2.8}$$

However, we have to solve two inverse matrices at each iteration, which will cost a lot of computational time. Meanwhile, it is hard to get $\hat{x}_k$, because it is difficult to calculate $F'(x_k)^{-1}$ even if the scale of problems is medium. Here, we give a new scheme, which can use the information of the former iteration adequately and save CUP-time. Here, we take $\hat{x}_k = x_k - F'(\hat{x}_{k-1})^{-1}F(x_k)$ and $\hat{x}_0 = x_0$. Moreover, a modified inexact Newton algorithm is obtained.

Algorithm 2.1 : Modified inexact Newton method

1. Let $x_0 \in R^n$ be a given initial guess.
2. For $k = 0$ until convergence do.
   2.1. Choose $\eta_k \in [0, 1)$.
   2.2. For the residual r_k and the iteration x_k, find the step s_k satisfying

$$F'(\hat{x}_k)s_k = -F(x_k) + r_k, \tag{2.9}$$

where

$$\frac{\|r_k\|}{\|F(x_k)\|} \leq \eta_k. \tag{2.10}$$

   2.3. Set $x_{k+1} = x_k + s_k$.
   2.4. Set $k = k + 1$ and turn to 2.1.

Remark 2.1. In the above algorithm, s_k is the inexact Newton step and (2.10) is the modified inexact Newton condition. η_k is the forcing term for the k-th iteration step which may depend on x_k. Taking $\eta_k \equiv 0$ and $\hat{x}_k = x_k$ give the famous Newton method. Because $F'(\hat{x}_{k-1})^{-1}$ has already known in the previous iteration, it is easy to get $\hat{x}_k$.

3. Semi-local convergence analysis of modified inexact Newton method

In this section, we will present the semi-local convergence result for Algorithm 2.1 and an affine invariant version is also presented.

The following well-known results are useful for our theorems, one can find them in [24].

Definition 3.1. F' is Lipschitz continuous in D, if there exists $L \geq 0$ such that

$$\|F'(x) - F'(y)\| \leq L\|x - y\|, \tag{3.1}$$

for every $x, y \in D$.

Lemma 3.1. *Let $F : D \subset R^n \to R^n$ be Fréchet differentiable and F' be Lipschitz continuous satisfying* (3.1)*. Then,*

$$\|F(y) - F(x) - F'(x)(y - x)\| \leq \frac{L}{2}\|y - x\|^2, \tag{3.2}$$

for every $x, y \in D$.

Lemma 3.2. *Let $F : D \subset R^n \to R^n$ be Fréchet differentiable and F' be invertible for every $x \in D$ such that*

$$\|F'(z)^{-1}[F'(x) - F'(y)]\| \leq L\|x - y\|, \ \forall\ x, y, z \in D.$$

Then,

$$\|F'(z)^{-1}[F(y) - F(x) - F'(x)(y - x)]\| \leq \frac{L}{2}\|y - x\|^2,$$

for every $x, y, z \in D$.

Theorem 3.3. *Suppose $F : D \subset R^n \to R^n$ is Fréchet differentiable and F' is Lipschitz continuous in D. Let $x_0 \in D$ and $\|F(x_0)\| \le \eta$. Assume $\hat{x}_k = x_k - F'(\hat{x}_{k-1})^{-1}F(x_k)$ and $\|F'(\hat{x})^{-1}\| \le \gamma$. If*

$$\sigma = \frac{L\eta}{2}\left[(2\gamma + \gamma\eta_{max})^2 + \gamma^2\right] + \eta_{max} < 1, \tag{3.3}$$

$$S = \overline{S\left(x_0, \frac{\delta}{1-\sigma}\right)} \subset D, \tag{3.4}$$

where $\eta_{max} = \sup_k\{\eta_k\} < 1$, $L = \max\{L_1, L_2\}$ and $\delta = \gamma(1 + \eta_{max})\eta$. Then the sequence of modified inexact Newton method defined by (2.9) stays in S and converges to x^* which satisfies $F(x^*) = 0$.

Proof. Firstly, we will prove

$$\|F(x_k)\| \le \sigma^k\eta, \ k = 0, 1, \dots, \tag{3.5}$$

by induction.

For $k = 0$, (3.5) holds evidently.

Assume that (3.5) is true for $k \le m$.

Now, we prove the assertion for $k = m + 1$. Since

$$\begin{aligned}
\|F(x_{m+1})\| &= \|F(x_{m+1}) - F(x_m) - F'(\hat{x}_m)(x_{m+1} - x_m) + r_m\| \\
&= \|F(x_{m+1}) - F(\hat{x}_m) + F(\hat{x}_m) - F(x_m) \\
&\quad - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m + \hat{x}_m - x_m) + r_m\| \\
&\le \|F(x_{m+1}) - F(\hat{x}_m) - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m)\| \\
&\quad + \|F(x_m) - F(\hat{x}_m) - F'(\hat{x}_m)(x_m - \hat{x}_m)\| + \|r_m\|.
\end{aligned}$$

By Lemma 3.1, it follows that

$$\|F(x_{m+1}) - F(\hat{x}_m) - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m)\| \le \frac{L_1}{2}\|x_{m+1} - \hat{x}_m\|^2, \tag{3.6}$$

$$\|F(x_m) - F(\hat{x}_m) - F'(\hat{x}_m)(x_m - \hat{x}_m)\| \le \frac{L_2}{2}\|x_m - \hat{x}_m\|^2, \tag{3.7}$$

where L_1 and L_2 are Lipschitz constants.

Using (3.6) and (3.7), we have

$$\begin{aligned}
\|F(x_{m+1})\| &\le \frac{L}{2}\|x_{m+1} - \hat{x}_m\|^2 + \frac{L}{2}\|x_m - \hat{x}_m\|^2 + \|r_m\| \\
&= \frac{L}{2}\|x_{m+1} - x_m + F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 \\
&\quad + \frac{L}{2}\|F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 + \|r_m\| \\
&\le \frac{L}{2}\frac{\|F'(\hat{x}_m)^{-1}[r_m - F(x_m)] + F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2}{\|F(x_m)\|^2}\|F(x_m)\|^2
\end{aligned}$$

$$+ \frac{L}{2} \left\| F'(\hat{x}_{m-1})^{-1} \right\|^2 \|F(x_m)\|^2 + \eta_{max} \|F(x_m)\|.$$

Hence, it follows from the assumption of induction, (2.10) and (3.3) that

$$\begin{aligned} \|F(x_{m+1})\| &\leq \left[\frac{L}{2}(\gamma + \gamma\eta_{max} + \gamma)^2 \|F(x_m)\| + \frac{L}{2}\gamma^2 \|F(x_m)\| + \eta_{max} \right] \|F(x_m)\| \\ &\leq \left\{ \frac{L\eta}{2} \left[(2\gamma + \gamma\eta_{max})^2 + \gamma^2 \right] + \eta_{max} \right\} \|F(x_m)\| \\ &\leq \sigma\sigma^m \eta = \sigma^{m+1}\eta. \end{aligned}$$

This gives (3.5) and the induction is complete.

Indeed, by (2.9) and (3.5),

$$\begin{aligned} \|x_{k+1} - x_k\| &= \|F'(\hat{x}_k)^{-1}[r_k - F(x_k)]\| \leq \|F'(\hat{x}_k)^{-1}\| \|r_k - F(x_k)\| \\ &\leq \gamma \frac{\|F(x_k)\| + \|r_k\|}{\|F(x_k)\|} \|F(x_k)\| \leq \gamma(1 + \eta_{max}) \|F(x_k)\| \\ &\leq \gamma(1 + \eta_{max})\sigma^k \eta. \end{aligned}$$

By the definition of δ, we have

$$\|x_{k+1} - x_k\| \leq \sigma^k \delta. \tag{3.8}$$

Therefore, for $m \geq 0$,

$$\|x_{k+m} - x_k\| \leq \sum_{i=k}^{k+m-1} \|x_{i+1} - x_i\| \leq \sum_{i=k}^{k+m-1} \sigma^i \delta \leq \frac{\delta\sigma^k}{1-\sigma}. \tag{3.9}$$

Thus, by (3.3), $\{x_k\}$ is a Cauchy sequence and converges to x^* as $k \to +\infty$. Moreover,

$$\|x_m - x_0\| \leq \frac{\delta}{1-\sigma},$$

for any $m \geq 0$. In view of (3.4), we obtain $x_m \in S \subset D$, which implies that $x^* \in S \subset D$ as well.

Finally, we assert that

$$\|F(x^*)\| = \left\| \lim_{k \to +\infty} F(x_k) \right\| \leq \lim_{k \to +\infty} \sigma^k \eta = 0.$$

Hence, $F(x^*) = 0$. The proof is completed. □

It is well-known [13, 35] that many Newton-like methods are affine invariant in the sense that when they are used to solve the affinely transformed problem

$$\mathcal{F}(x) = 0, \ \mathcal{F} \equiv AF,$$

where A is any nonsingular $n \times n$ matrix.

Example. Consider the Newton iteration

$$x_{k+1} = x_k - [F(x_k)]^{-1} F(x_k), \ k = 0, 1, \ldots.$$

For F, we use affine transform $\mathcal{F} = AF$. Then

$$[\mathcal{F}'(x)]^{-1}\mathcal{F}(x) = [AF'(x)]^{-1}[AF(x)] = [F'(x)]^{-1}F(x).$$

Thus, we assert that the Newton iteration sequence $\{x_k\}$ is affine invariant.

Convergence conditions for affine invariant methods should themselves be invariant under the transformations of this form [13]. Now it is clear that even if the method (2.9) is affine invariant, the condition (2.10) is not affine invariant. In order to give semi-local convergence theorem for the modified inexact Newton method with affine invariant condition, we improve the method (2.9) with the above condition (2.10) as follows

$$F'(\hat{x}_k)s_k = -F(x_k) + F'(x_k)r_k,\ k = 0, 1, \ldots, \tag{3.10}$$

$$\frac{\|r_k\|}{\|F'(x_k)^{-1}F(x_k)\|} \le \nu_k,\ k = 0, 1, \ldots. \tag{3.11}$$

These improvements for inexact Newton method were proposed by Bai and Tong [4]. Here, we take them for our method.

The following theorem concerns the affine invariance of our algorithm.

Theorem 3.4. *Suppose $F : D \subset R^n \to R^n$ is Fréchet differentiable and the modified inexact Newton method is defined by* (3.10) *and* (3.11). *Assume F' is invertible for every $x \in D$ such that*

$$\|F'(z)^{-1}[F'(x) - F'(y)]\| \le L\|x - y\|,\ \forall\ x, y, z \in D.$$

Let $x_0 \in D$ and $\|F'(x_0)^{-1}F(x_0)\| \le \eta$. Suppose further that $\hat{x}_k = x_k - F'(\hat{x}_{k-1})^{-1}F(x_k)$, $\frac{\|F'(\hat{x})^{-1}F(x)\|}{\|F'(x)^{-1}F(x)\|} \le \mu$ and $\|F'(\hat{x})^{-1}F'(x)\| \le \gamma$. If

$$\sigma = \frac{L\eta}{2}\left[(2\mu + \gamma\nu_{max})^2 + \mu^2\right] + L\eta(\mu + \gamma\nu_{max}) + \nu_{max} < 1, \tag{3.12}$$

$$S = \overline{S\left(x_0, \frac{\delta}{1-\sigma}\right)} \subset D, \tag{3.13}$$

where $\nu_{max} = \sup_k\{\nu_k\} < 1$, $L = \max\{L_1, L_2\}$ and $\delta = \gamma(1 + \nu_{max})\eta$. Then the sequence of modified inexact Newton method stays in S and converges to x^* which satisfies $F(x^*) = 0$.

Proof. The proof is similar to the proof of Theorem 3.1. Firstly, we will prove

$$\|F'(x_k)^{-1}F(x_k)\| \le \sigma^k\eta,\ k = 0, 1, \ldots, \tag{3.14}$$

by induction.

For $k = 0$, (3.14) holds evidently.

Assume that (3.14) is true for $k \le m$.

Now, we prove the assertion for $k = m+1$. Let $T_{m+1} = \|F'(x_{m+1})^{-1}F(x_{m+1})\|$ for short. Since

$$\begin{aligned} T_{m+1} = \|&F'(x_{m+1})^{-1}[F(x_{m+1}) - F(x_m) - F'(\hat{x}_m)(x_{m+1} - x_m)] \\ &+ F'(x_{m+1})^{-1}F'(x_m)r_m\| \end{aligned}$$

$$
\begin{aligned}
&= \|F'(x_{m+1})^{-1}[F(x_{m+1}) - F(\hat{x}_m) + F(\hat{x}_m) - F(x_m) \\
&\quad - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m + \hat{x}_m - x_m)] + F'(x_{m+1})^{-1}F'(x_m)r_m\| \\
&\leq \|F'(x_{m+1})^{-1}[F(x_{m+1}) - F(\hat{x}_m) - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m)]\| \\
&\quad + \|F'(x_{m+1})^{-1}[F(x_m) - F(\hat{x}_m) - F'(\hat{x}_m)(x_m - \hat{x}_m)]\| \\
&\quad + \|F'(x_{m+1})^{-1}F'(x_m)r_m\|.
\end{aligned}
$$

By Lemma 3.2, it follows that

$$\|F'(x_{m+1})^{-1}[F(x_{m+1}) - F(\hat{x}_m) - F'(\hat{x}_m)(x_{m+1} - \hat{x}_m)]\| \leq \frac{L_1}{2}\|x_{m+1} - \hat{x}_m\|^2, \tag{3.15}$$

$$\|F'(x_{m+1})^{-1}[F(x_m) - F(\hat{x}_m) - F'(\hat{x}_m)(x_m - \hat{x}_m)]\| \leq \frac{L_2}{2}\|x_m - \hat{x}_m\|^2, \tag{3.16}$$

where L_1 and L_2 are constant.

Using (3.15) and (3.16),

$$
\begin{aligned}
T_{m+1} &\leq \frac{L}{2}\|x_{m+1} - \hat{x}_m\|^2 + \frac{L}{2}\|x_m - \hat{x}_m\|^2 \\
&\quad + \|F'(x_{m+1})^{-1}[F'(x_m) - F'(x_{m+1})]r_m\| + \|r_m\| \\
&= \frac{L}{2}\|x_{m+1} - x_m + F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 + \frac{L}{2}\|F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 \\
&\quad + L\|x_m - x_{m+1}\|\|r_m\| + \|r_m\| \\
&= \frac{L}{2}\| - F'(\hat{x}_m)^{-1}F(x_m) + F'(\hat{x}_m)^{-1}F'(x_m)r_m + F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 \\
&\quad + \frac{L}{2}\|F'(\hat{x}_{m-1})^{-1}F(x_m)\|^2 \\
&\quad + L\left\|-F'(\hat{x}_m)^{-1}F(x_m) + F'(\hat{x}_m)^{-1}F'(x_m)r_m\right\|\|r_m\| + \|r_m\|.
\end{aligned}
$$

Hence, it follows from the assumption of induction, (3.11) and (3.12) that

$$
\begin{aligned}
T_{m+1} &\leq \left\{\frac{L\eta}{2}\left[(2\mu + \gamma\nu_{max})^2 + \mu^2\right] + L\eta(\mu + \gamma\nu_{max}) + \nu_{max}\right\}\|F'(x_m)^{-1}F(x_m)\| \\
&\leq \sigma\sigma^m\eta = \sigma^{m+1}\eta.
\end{aligned}
$$

This gives (3.14) and the induction is complete.

Indeed, by (3.10) and (3.14)

$$
\begin{aligned}
&\|x_{k+1} - x_k\| \\
&= \|F'(\hat{x}_k)^{-1}[F'(x_k)r_k - F(x_k)]\| = \|F'(\hat{x}_k)^{-1}F'(x_k)[r_k - F'(x_k)^{-1}F(x_k)]\| \\
&\leq \gamma\frac{\|F(x_k)^{-1}F(x_k)\| + \|r_k\|}{\|F(x_k)^{-1}F(x_k)\|}\|F(x_k)^{-1}F(x_k)\| \leq \gamma(1 + \nu_{max})\|F(x_k)^{-1}F(x_k)\| \\
&\leq \gamma(1 + \nu_{max})\sigma^k\eta.
\end{aligned}
$$

By the definition of δ, we have

$$\|x_{k+1} - x_k\| \leq \sigma^k\delta. \tag{3.17}$$

Therefore, for $m \geq 0$,

$$\|x_{k+m} - x_k\| \leq \sum_{i=k}^{k+m-1} \|x_{i+1} - x_i\| \leq \sum_{i=k}^{k+m-1} \sigma^i \delta \leq \frac{\delta \sigma^k}{1-\sigma}. \tag{3.18}$$

Thus, by (3.12), $\{x_k\}$ is a Cauchy sequence and converges to x^* as $k \to +\infty$. Moreover,

$$\|x_m - x_0\| \leq \frac{\delta}{1-\sigma},$$

for any $m \geq 0$. In view of (3.13), we obtain $x_m \in S \subset D$, which implies that $x^* \in S \subset D$ as well.

Finally, we assert that

$$\|F(x^*)\| = \left\| \lim_{k \to +\infty} F(x_k) \right\| \leq \lim_{k \to +\infty} \sigma^k \eta = 0.$$

Hence, $F(x^*) = 0$. The proof is completed. □

Next, we consider the rate of convergence of the modified inexact Newton method. First, we give the following definition given by R.S. Dembo et al. [11].

Definition 3.2. Let $\{x_k\}$ be a sequence which converges to x^*. Then $x_k \to x^*$ with order at least q ($q > 1$) if

$$\|x_{k+1} - x^*\| = O(\|x_k - x^*\|^q) \quad \text{as } k \to \infty.$$

Theorem 3.5. *Under the assumptions of Theorem 3.1, if the modified inexact Newton iterates $\{x_k\}$ converge to x^*, then $x_k \to x^*$ with order at least 2 if and only if*

$$\|r_k\| = O(\|x_k - x^*\|^2) \quad \textit{as } k \to \infty.$$

Proof. Assume that $x_k \to x^*$ with order at least 2. Note that

$$\begin{aligned} r_k = {} & [F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)] - [F'(\hat{x}_k) - F'(x^*)](x_k - x^*) \\ & + [F'(x^*) + F'(\hat{x}_k) - F'(x^*)](x_{k+1} - x^*). \end{aligned}$$

Taking norms, we arrive at

$$\begin{aligned} \|r_k\| \leq {} & \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\| + \|F'(\hat{x}_k) - F'(x^*)\| \|x_k - x^*\| \\ & + [\|F'(x^*)\| + \|F'(\hat{x}_k) - F'(x^*)\|] \|x_{k+1} - x^*\|. \end{aligned}$$

Therefore, by Lemma 3.1, the continuity of F' at x^* and the assumption that $x_k \to x^*$ with order at least 2, we have

$$\|r_k\| = O(\|x_k - x^*\|^2) \quad \text{as } k \to \infty.$$

Conversely, assume that $\|r_k\| = O(\|x_k - x^*\|^2)$. Note that

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* + F'(\hat{x}_k)^{-1}[r_k - F(x_k)] \\ &= F'(\hat{x}_k)^{-1}[r_k + (F'(\hat{x}_k) - F'(x^*))(x_k - x^*) \\ &\quad - F(x_k) + F(x^*) + F'(x^*)(x_k - x^*)]. \end{aligned}$$

Taking norms, we arrive at

$$\|x_{k+1} - x^*\| \le \|F'(\hat{x}_k)^{-1}\|[\|r_k\| + \|F'(\hat{x}_k) - F'(x^*)\|\|x_k - x^*\| \\ + \|F(x_k) - F(x^*) - F'(x^*)(x_k - x^*)\|].$$

Therefore, by Lemma 3.1 and the assumption that $\|r_k\| = O(\|x_k - x^*\|^2)$, we have

$$\|x_{k+1} - x^*\| = O(\|x_k - x^*\|^2) \quad \text{as } k \to \infty.$$

□

4. Numerical experiments

In this section, we give three test problems using the algorithm presented in this paper to show its convergence property and robustness. The purpose of the first two problems are to show that Algorithm 2.1 is useful in the nonlinear case. By useful we mean that the algorithm presented has faster convergence. In the third problem, a nonlinear PDE is solved.

Example 1. Consider the following nonlinear equations:

$$\begin{cases} f_1 = x_1^3 + x_2 - 2, \\ f_2 = x_1 + 2x_2 - 3, \end{cases}$$

with $x^* = (1, 1)^{\mathrm{T}}$.

Take $\eta = 10^{-4}$ given by Cai, Gropp, Keyes and Tidriri [9]. Using the modified inexact Newton method, we can obtain the iterative solutions listed in Table 1 and 2 with initial guess $(-1, -1)^{\mathrm{T}}$ and $(510, 10^{21})^{\mathrm{T}}$, respectively.

Table 1. Comparison of iterative solutions of two algorithms with the initial guess $(-1, -1)^{\mathrm{T}}$

k	x_1^{IN}	x_2^{IN}	x_1^{MIN}	x_2^{MIN}
1	-0.6000	1.8000	0.7241	1.1379
2	0.1172	1.4414	0.8569	1.0715
3	-1.0969	2.0485	0.9678	1.0161
4	-0.6881	1.8440	0.9987	1.0007
5	-0.1646	1.5823	1.0000	1.0000
10	-1.2463	2.1231	1.0000	1.0000
20	0.9874	1.0063	1.0000	1.0000
22	1.0000	1.0000	1.0000	1.0000

In Table 1, a comparison of iterative solutions of two algorithms of Example 1 with the initial guess $(-1, -1)^{\mathrm{T}}$ is shown. Here, x_1^{IN} and x_2^{IN} denote the iterative solutions of inexact Newton method while x_1^{MIN} and x_2^{MIN} represent the iterative solutions of modified inexact Newton method. We have seen from Table 1 that the modified inexact Newton method has faster convergence.

In Table 2, a comparison of iterative solutions of two algorithms of Example 1 with the initial guess $(510, 10^{21})^{\mathrm{T}}$ is shown. We have seen from Table 2 that

Table 2. Comparison of iterative solutions of two algorithms with the initial guess $(510, 10^{21})^{\mathrm{T}}$

k	x_1^{IN}	x_2^{IN}	x_1^{MIN}	x_2^{MIN}
1	339.7500	0.0000	127.0000	0.0000
2	226.5003	-111.7502	31.7528	-14.3764
3	151.0007	-74.0004	7.9498	-2.4749
7	29.8317	-13.4159	0.9986	1.0007
8	19.8917	-8.4459	1.0000	1.0000
10	8.8541	-2.9271	1.0000	1.0000
18	1.0001	1.0000	1.0000	1.0000
19	1.0000	1.0000	1.0000	1.0000

two methods have a wide convergence domain and the modified inexact Newton method has faster convergence.

In Fig. 1, we present profiles for the history of absolute error for two algorithms of Example 1 with the initial guess $(-1, -1)^{\mathrm{T}}$. We have seen from Fig. 1 that the absolute error using modified inexact Newton method becomes small faster than that of inexact Newton method. It is observed that the proposed method performs better than the inexact Newton method.

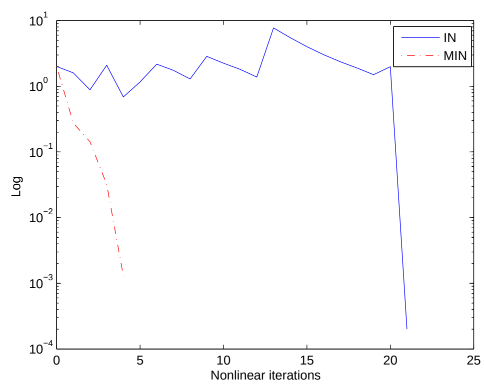

Fig. 1. History of absolute error for two algorithms with the initial guess $(-1, -1)^{\mathrm{T}}$.

In Fig. 2, we present profiles for history of absolute error for two algorithms of Example 1 with the initial guess $(510, 10^{21})^{\mathrm{T}}$. We have seen from Fig. 2 that the modified inexact Newton method may perform several times faster than the inexact Newton method.

Example 2. Consider the following nonlinear equations:

$$f_1 = 4x_1 - 4x_2^2,$$
$$f_i = -8x_{i-1}x_i + 8x_i^3 + 6x_i - 4x_{i+1}^2 - 2, \quad i = 2, 3, \ldots, 9,$$
$$f_{10} = -8x_9x_{10} + 8x_{10}^3 + 2x_{10} - 2.$$

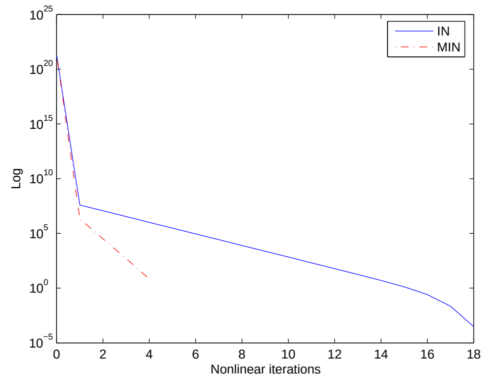

Fig. 2. History of absolute error for two algorithms with the initial guess $(510, 10^{21})^{\mathrm{T}}$.

Take $\eta = 10^{-4}$ given by Cai, Gropp, Keyes and Tidriri [9]. Using the modified inexact Newton method, we can obtain the iterative solutions. Here, we use $\|F(x_k)\| \leq 1.0e - 9$ as the stopping rule for this example.

Table 3. Results for Example 2

Method	No.Eq1	No.Eqk	No.Fd1	No.Fdk	No.it	CPU-time	Res
Presented method	10	20	1	2	109	0.7656	9.8920e-10
inexact Newton	10	10	1	1	322	0.7813	9.6020e-10

Results for Example 2 are listed in Table 3, where No.Eq1 denotes the number of equation in the first step, No.Eqk represents the number of equation in the k-th step ($k > 1$), No.Fd1 denotes the number of F' in the first step, No.Fdk represents the number of F' in the k-th step, No.it denotes the number of iteration and Res stands for the value of $\|F(x_k)\|$ when our stop rule is satisfied. As shown in Table 3, the number of iteration by the modified inexact Newton method is less than that of the classical inexact Newton method. Meanwhile, we have seen from Table 3 that the number of F' by the modified inexact Newton method in the k-th step is 2, but $F'(\hat{x}_{k-1})$ has already known in the previous iteration. Hence, the number of F' is indeed 1. Although the number of equation of the modified inexact Newton method is nearly twice of that of the classical one, the CPU-time of the former method is still less. All in all, our method indeed saves on computation.

Example 3. Consider one-dimensional Burgers' equation:

$$\begin{cases} \dfrac{\partial u}{\partial t} + u\dfrac{\partial u}{\partial x} = \nu\dfrac{\partial^2 u}{\partial x^2}\ , & 0 < x < 1,\ 0 < t \leq T, \\ u(x,0) = \sin(\pi x), & 0 \leq x \leq 1, \\ u(0,t) = u(1,t) = 0, & 0 < t \leq T, \end{cases} \tag{4.1}$$

where the positive number $\nu = \frac{1}{Re}$ is the coefficient of viscosity, and Re denotes the Reynolds number. This equation has an exact solution in the form of the infinite series

$$u(x,t) = 4\pi\nu \frac{\sum_{j=1}^{\infty} jI_j(\frac{1}{2\pi\nu})\sin(j\pi x)\exp(-j^2\pi^2\nu t)}{2\sum_{j=1}^{\infty} I_j(\frac{1}{2\pi\nu})\cos(j\pi x)\exp(-j^2\pi^2\nu t) + I_0(\frac{1}{2\pi\nu})}, \tag{4.2}$$

where $I_j(x)$ is the first type of the j-th modified Bessel function. When $j = 35$, it is used as an approximation to the infinite sum (4.2).

We solve (4.1) with finite difference method and the modified inexact Newton method. First, we discretize in space with centered difference to obtain a system of ordinary differential equations, which we write as

$$u_t = P(u),\ u(0) = u^0. \tag{4.3}$$

Then the nonlinear equations that should be solved for the implicit Euler method with a time step τ is

$$u^{n+1} - u^n = \tau P(u^{n+1}). \tag{4.4}$$

Moreover, one solves, at each time step, the nonlinear equations

$$F(U) = U - u^n - \tau P(U) = 0. \tag{4.5}$$

Then, we use the modified inexact Newton method to solve nonlinear equations (4.5).

Let $h = 1/M$ be the mesh width in space and set $x_i = ih$ for $i = 1, 2, \ldots, M-1$. Let $\tau = T/N$ be the mesh width in time and set $t_n = n\tau$ for $n = 1, 2, \ldots, N$. $u(x_i, t_n)$ is the approximate solution of $u(x,t)$. Discretization is on a 100×100 grid, so that $N = 100$ and $M = 100$. Hence, the spatial mesh width $h = 0.01$ and the time step $\tau = 0.01$. Take $\eta_{max} = 0.9$ used by Eisenstat and Walker in [15]. Here, GMRES(m) algorithm is used for linear systems and $m = 40$.

Table 4. Comparison of the numerical solution with the exact solution at different space points of Example 3 at $T = 0.1$ for $\nu = 0.1$

x	Numerical Solution	Exact solution	Absolute error $\lvert u^{\text{exact}} - u^{\text{num}}\rvert$
0.1	0.22339	0.22345	6.00e-05
0.2	0.43601	0.43580	2.10e-04
0.3	0.62603	0.62512	9.10e-04
0.4	0.78001	0.77772	2.30e-03
0.5	0.87696	0.87728	3.20e-04
0.6	0.90918	0.90425	4.90e-03
0.7	0.83796	0.83692	1.00e-03
0.8	0.64950	0.65731	7.80e-03
0.9	0.36471	0.36575	1.00e-03

In Table 4, a comparison of the numerical solution with the exact solution at different space points of (0, 1) for Example 3 at T = 0.1 and $\nu = 0.1$ is shown. It is observed that the proposed method gives sharp results. In order to show the physical behaviour of the given problem, we give surface plots of the computed solutions for different values of the coefficient of viscosity, ν in Figs. 3 and 4.

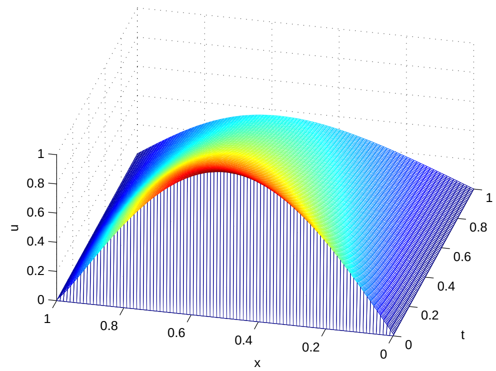

Fig. 3. Numerical solutions profiles of Example 3 for $\nu = 0.1$, $h = 0.01$ and $\tau = 0.01$.

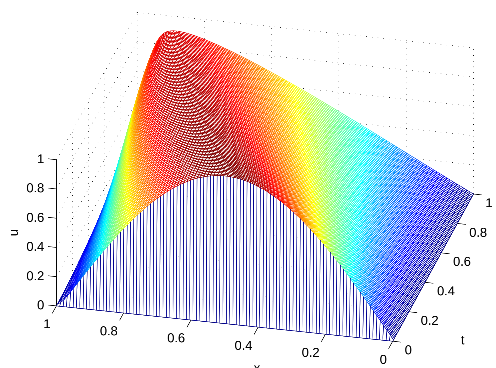

Fig. 4. Numerical solutions profiles of Example 3 for $\nu = 0.01$, $h = 0.01$ and $\tau = 0.01$.

5. Concluding remarks

The modified inexact Newton method for nonlinear equations has been presented. It is shown that the method given performs several times faster than the inexact Newton method. A semi-local convergence theorem for the modified inexact Newton method is studied and an affine invariant version is also presented. We then give three numerical examples which show that the modified inexact Newton scheme is more efficient than traditional inexact Newton strategy. Therefore, it is suggested to use the modified inexact Newton to get the numerical solution of the nonlinear equations effectively.

References

1. H.B. An, Z.Y. Mo and X.P. Liu, *A choice of forcing terms in inexact Newton method*, J. Comput. Appl. Math., **200** (2007), 47-60.
2. I.K. Argyros, *On the semilocal convergence of inexact Newton methods in Banach spaces*, J. Comput. Appl. Math., **228** (2009), 434-443.
3. I.K. Argyros, *A new convergence theorem for the inexact Newton methods based on assumptions involving the second Fréchet derivative*, Comput. Math. Appl., **37** (1999), 109-115.
4. Z.Z. Bai and P.L. Tong, *On the affine invariant convergence theorems of inexact Newton method and Broyden's method*, J. UEST China, **23** (1994), 535-539 (in Chinese).
5. Z.J. Bai, *Inexact Newton methods for inverse eigenvalue problems*, Appl. Math. Comput., **172** (2006), 682-689.
6. P.N. Brown, P.S. Vassilevski and C.S. Woodward, *On mesh-independent convergence of an inexact Newton-Multigrid algorithm*, SIAM J. Sci. Comput., **25** (2003), 570-590.
7. P.N. Brown and Y. Saad, *Hybird Krylov methods for nonlinear systems of equations*, SIAM J. Sci. Stat. Comput., **11** (1990), 450-481.
8. X.C. Cai and D.E. Keyes, *Nonlinearly preconditioned inexact Newton algorithms*, SIAM J. Sci. Comput., **24** (2002), 183-200.
9. X.C. Cai, W.D. Gropp, D.E. Keyes and M.D. Tidriti, *Newton-Krylov-Schwarz methods in CFD, in: Proceedings of the International Workshop on Numerical Methods for the Navier-Stokes Equations*, Vieweg, Braunschwieg, 1995.
10. J.H. Chen and W.G. Li, *Convergence behaviour of inexact Newton methods under weak Lipschitz condition*, J. Comput. Appl. Math., **191** (2006), 143-164.
11. R.S. Dembo, S.C. Eisenstat and T. Steihaug, *Inexact Newton methods*, SIAM J. Numer. Anal., **19** (1982), 400-408.
12. R.S. Dembo and T. Steihaug, *Truncated Newton algorithms for large-scale optimization*, Math. Program, **26** (1983), 190-212.
13. P. Deuflhard and G. Heindl, *Affine invariant convergence theorems for Newton's method and extensions to related methods*, SIAM J. Numer. Anal., **16** (1979), 1-10.
14. S.C. Eisenstat and H.F. Walker, *Globally convergent inexact Newton methods*, SIAM J. Optim., **4** (1994), 393-422.
15. S.C. Eisenstat and H.F. Walker, *Choosing the forcing term in an inexact Newton method*, SIAM J. Sci. Comput., **17** (1996), 16-32.
16. R.N. Elias, A.L.G.A. Coutinho and M.A.D. Martins, *Inexact Newton-type methods for the solution of steady incompressible viscoplastic flows with the SUPG/PSPG finite element formulation*, Comput. Methods Appl. Mech. Engrg., **195** (2006), 3145-3167.
17. D.R. Fokkema, G.L.G. Sleijpen and H.A.V.D. Vorst, *Accelerated inexact Newton schemes for large system of nonlinear equations*, SIAM J. Sci. Comput., **19** (1998), 657-674.
18. A. Galántai, *The theory of Newton's method*, J. Comput. Appl. Math., **124** (2000), 25-44.
19. X.P. Guo, *On semilocal convergence of inexact Newton methods*, J. Comput. Math., **25** (2007), 231-242.
20. C. Li and W.P. Shen, *Local convergence of inexact methods under the Hölder condition*, J. Comput. Appl. Math., **222** (2008), 544-560.
21. J.M. Martinez and L. Qi, *Inexact Newton methods for solving nonsmooth equations*, J. Comput. Appl. Math., **60** (1995), 127-145.
22. I. Moret, *A Kantorovich type theorem for inexact Newton methods*, Numer. Funct. Anal. Optim., **10** (1989), 351-365.

23. B. Morini, *Convergence behaviour of inexact Newton method*, Math. Comput., **68** (1999), 1605-1613.
24. J.M. Ortega and W.C. Rheinboldt, *Iterative Solution of Nonlinear Equations in Several Variables*, SIAM, Philadelphia, PA, 2000.
25. A. Rieder, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal., **43** (2005), 604-622.
26. A. Rieder, *On the regularization of nonliear ill-posed problems via inexact Newton iterations*, Inverse Probl., **15** (1999), 309-327.
27. A. Rieder, *On convergence rates of inexact Newton regularization*, Numer. Math., **88** (2001), 347-365.
28. J.N. Shadid, R.S. Tuminaro and H.F. Walker, *An inexact Newton method for fully coupled solution of the Navier-Stokes equations with heat and mass transport*, J. Comput. Phys., **137** (1997), 155-185.
29. W.P. Shen and C. Li, *Convergence criterion of inexact methods for operators with Hölder continuous derivatives*, Taiwan. J. Math., **12** (2008), 1865-1882.
30. W.P. Shen and C. Li, *Kantorovich-type convergence criterion for inexact Newton methods*, Appl. Numer. Math., **59** (2009), 1599-1611.
31. M.V. Solodov and B.F. Svaiter, *A globally convergent inexact Newton method for systems of monotone equations, Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods*, Kluwer Academic Publishers, 1998.
32. M. Wu, *A new semi-local convergence theorem for the inexact Newton methods*, Appl. Numer. Math., **200** (2008), 80-86.
33. T. Yamamoto, *Historical developments in convergence analysis for Newton's and Newton-like method*, J. Comput. Appl. Math., **124** (2000), 1-23.
34. T.J. Ypma, *Local convergence of inexact Newton methods*, SIAM J. Numer. Anal., **21** (1984), 583-590.
35. T.J. Ypma, *Affine invariant convergence for Newton's methods*, BIT, **22** (1982), 108-118.
36. J.L. Zhou, S.Y. Zhang, G.P. Yang and J.R. Tan, *A convergence theorem for the inexact Newton methods based on Hölder continuous Fréchet derivative*, Appl. Numer. Math., **197** (2008), 206-211.

Pengzhan Huang received his Ph.D. Degree from Xinjiang University. He is currently an associate professor at Xinjiang University. His research interests focus on numerical solutions for partial differential equations.

College of Mathematics and System Sciences, Xinjiang University, Urumqi 830046, P.R. China.
e-mail: `hpzh007@yahoo.com`

Abdurishit Abduwali received his Ph.D. from Okayama University of Science. He is currently a professor at Xinjiang University. His research interests focus on numerical solutions for partial differential equations.

College of Mathematics and System Sciences, Xinjiang University, Urumqi 830046, P.R. China.
e-mail: `rashit@xju.edu.cn`