Available online at www.sciencedirect.com
SCIENCE DIRECT®
Journal of Computational and Applied Mathematics 191 (2006) 143–164
JOURNAL OF COMPUTATIONAL AND APPLIED MATHEMATICS
www.elsevier.com/locate/cam

Convergence behaviour of inexact Newton methods under weak Lipschitz condition

Jinhai Chen[a, b, *], Weiguo Li[b]

[a] Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
[b] School of Mathematics and Computational Sciences, University of Petroleum, Dongying 257061, Shandong Province, People's Republic of China

Received 13 September 2004

Abstract

Under weak Lipschitz condition, local convergence properties of inexact Newton methods and Newton-like methods for systems of nonlinear equations are established in an arbitrary vector norm. Processes with modified relative residual control are considered; the results easily provide an estimate of convergence ball for inexact methods. For a special case, the results are affine invariant. Some applications are given.

MSC: 65H10

Keywords: Systems of nonlinear equations; Inexact Newton methods; Weak Lipschitz condition; Affine invariant condition

1. Introduction

We consider the system of nonlinear equations:

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

* Corresponding author.
E-mail addresses: cjh_maths@sina.com, cjh_maths@yahoo.com.cn (J. Chen).

0377-0427/$ - see front matter © 2005 Elsevier B.V. All rights reserved.
doi:10.1016/j.cam.2005.03.076

where $f(x): R^n \to R^n$ is Fréchet differentiable. Let $f'(x)$ denote the Fréchet derivative of f at x. Inexact iterative procedures commonly used to solve (1.1) have the general forms:

For $k=0$ step 1 until convergence do

Find the step Δ_k which satisfies:

$$B_k\Delta_k = -f(x_k) + r_k \quad \text{where } \frac{\|r_k\|}{\|f(x_k)\|} \leqslant \eta_k. \tag{1.2}$$

Set $x_{k+1} = x_k + \Delta_k$,

where x_0 is a given initial guess, B_k is an $n \times n$ nonsingular matrix and $\{\eta_k\}$ is a sequence of forcing terms such that $0 \leqslant \eta_k \leqslant 1$. The process is inexact Newton method if $B_k = f'(x_k)$, the process is inexact modified Newton method if $B_k = f'(x_0)$, and it represents a inexact Newton-like method if $B_k = B(x_k)$ is an approximation of derivative $f'(x_k)$ (see [5,13–15]).

We remark that inexact methods include the class of Newton iterative methods (see [1,3–5,7,12–14]), where an iterative method is used to approximate the solution of linear systems (1.2).

For inexact Newton methods, local and rate of convergence properties can be characterized in terms of forcing sequence η_k. Let $\|\cdot\|$ denote any vector norm on R^n and matrix subordinate norm on $R^{n\times n}$. In [2] it is shown that, if the usual assumptions for Newton's method hold and η_k is uniformly less than 1, we can define a sequence $\{x_k\}$ linearly convergent to a solution x^* of (1.1) in the norm $\|y\|_* = \|f'(x^*)y\|$.

Recently, several authors (see [8,11]) have proposed applications of inexact methods in different fields of numerical analysis and pointed out difficulties in applying results of [2]. In fact, such results are norm-dependent and $\|\cdot\|_*$ is not computable. Then, they focused on the analysis of the stopping residual control $\|r_k\|/\|f(x_k)\| \leqslant \eta_k$ and its effect on convergence properties. Morini (see [13]) considered inexact methods where a scaled relative residual control was performed at each iteration; its iterative form is as follows:

For $k=0$ step 1 until convergence do

Find the step Δ_k which satisfies

$$B_k\Delta_k = -f(x_k) + r_k \quad \text{where } \frac{\|P_k r_k\|}{\|P_k f(x_k)\|} \leqslant \theta_k. \tag{1.3}$$

Set $x_{k+1} = x_k + \Delta_k$,

where P_k is an invertible matrix for each k. If $P_k = I$ for each k, (1.3) reduces to (1.2). It is worth noting that residuals of this form are used in iterative Newton methods if preconditioning is applied, and that P_k changes with index k if B_k does. But we also note that the results obtained in [13] cannot make us clearly see how big the radius of the convergence ball is.

Let x^* denote a solution of (1.1), $B(x,r)$ denote an open ball with radius r and center x, and let $\overline{B(x,r)}$ denote its closure. Under the hypothesis that $f'(x)$ satisfies the Lipschitz condition

$$\|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| \leqslant \int_{\tau\rho(x)}^{\rho(x)} L(u)\,\mathrm{d}u, \quad \forall x \in B(x^*, r), \tag{1.4}$$

where $\rho(x) = \|x - x^*\|$, $x^\tau = x^* + \tau(x - x^*)$, and L is a monotone function. Wang (see[16–18]) studied the convergence of the Newton's method.

In this paper, under weak Lipschitz condition, we continue to consider inexact methods where a scaled relative residual control was performed at each iteration. The results obtained are valid under widely used hypotheses on f and merge into the theories of Newton's and Newton-like methods in the limiting case of vanishing residuals, i.e., $\eta_k = 0$ for each k. The results also can make us clearly see how big the radius of the convergence ball is. Further, for a special case, such conditions of convergence are affine invariant and in agreement with the theory of [18].

2. Preliminaries

The condition on the function f

$$\|f(x) - f(x^\tau)\| \leqslant L\|x - x^\tau\|, \quad \forall x \in B(x^*, r), \tag{2.1}$$

where $x^\tau = x^* + \tau(x - x^*)$, $0 \leqslant \tau \leqslant 1$, is usually called radius Lipschitz condition in the ball $B(x^*, r)$ with constant L. Sometimes, if it is only required to satisfy

$$\|f(x) - f(x^*)\| \leqslant L\|x - x^*\|, \quad \forall x \in B(x^*, r), \tag{2.2}$$

we call it the center Lipschitz condition in the ball $B(x^*, r)$ with constant L. Furthermore, the L in the Lipschitz condition need not be a constant, but a positive integrable function. If this is the case, then (2.1) or (2.2) is replaced by

$$\|f(x) - f(x^\tau)\| \leqslant \int_{\tau\rho(x)}^{\rho(x)} L(u)\,\mathrm{d}u, \quad \forall x \in B(x^*, r), \quad 0 \leqslant \tau \leqslant 1, \tag{2.3}$$

or

$$\|f(x) - f(x^*)\| \leqslant \int_0^{\rho(x)} L(u)\,\mathrm{d}u, \quad \forall x \in B(x^*, r), \tag{2.4}$$

where $\rho(x) = \|x - x^*\|$. At the same time, the corresponding 'Lipschitz condition' is referred to as having the L average.

By Banach's theorem (see [9,10,14]), the following result can be obtained directly.

Lemma 2.1 (*See Wang [18]*). *Suppose that f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists and $f'(x^*)^{-1}f'$ satisfies the center Lipschitz condition with the L average*:

$$\|f'(x^*)^{-1}f'(x) - I\| \leqslant \int_0^{\rho(x)} L(u)\,\mathrm{d}u, \quad \forall x \in B(x^*, r), \tag{2.5}$$

where L is positive integrable function. Let r satisfy

$$\int_0^r L(u)\,\mathrm{d}u \leqslant 1. \tag{2.6}$$

Then $f'(x)$ is invertible in this ball and

$$\|f'(x)^{-1}f'(x^*)\| \leqslant \left(1-\int_0^{\rho(x)} L(u)\,\mathrm{d}u\right)^{-1}. \tag{2.7}$$

Proof. In fact, suppose $E = I - f'(x^*)^{-1}f'(x)$, $\forall x \in B(x^*, r)$, and a sequence $\{S_k\}_{k=0}^{\infty}$ defined by

$$S_k = I + E + \cdots + E^k.$$

From (2.5) and (2.6), we obtain $\|E\| \leqslant \int_0^{\rho(x)} L(u)\,\mathrm{d}u < \int_0^r L(u)\,\mathrm{d}u \leqslant 1$, and note that, for all m, n $(m > n)$,

$$\|S_m - S_n\| = \left\|\sum_{k=n}^{m-1} E^k\right\| \leqslant \sum_{k=n}^{m-1} \|E\|^k.$$

These imply $\{S_k\}_{k=0}^{\infty}$ is a Cauchy sequence and convergent.

By the equality:

$$(I-E)(I+E+\cdots+E^{k-1}) = (I+E+\cdots+E^{k-1})(I-E) = I - E^k,$$

we have

$$(I-E)\left(\sum_{k=0}^{\infty} E^k\right) = \left(\sum_{k=0}^{\infty} E^k\right)(I-E) = I,$$

i.e., $I - E = f'(x^*)^{-1}f'(x)$ is invertible, and

$$\|f'(x^*)^{-1}f'(x)\| = \left\|\sum_{k=0}^{\infty} E^k\right\| \leqslant \sum_{k=0}^{\infty} \|E\|^k = \frac{1}{1-\|E\|} \leqslant \frac{1}{1-\int_0^{\rho(x)} L(u)\,\mathrm{d}u}.$$

This shows the validity of (2.7). □

Lemma 2.2. *Suppose that f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists.*

If $f'(x^*)^{-1}f'$ satisfies the radius Lipschitz condition with the L average:

$$\|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| \leqslant \int_{\tau\rho(x)}^{\rho(x)} L(u)\,\mathrm{d}u, \quad \forall x \in B(x^*, r), \quad 0 \leqslant \tau \leqslant 1, \tag{2.8}$$

where $x^\tau = x^* + \tau(x - x^*)$, $\rho(x) = \|x - x^*\|$ and L is positive integrable function, then we have

$$\int_0^1 \|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\|\rho(x)\,\mathrm{d}\tau \leqslant \int_0^{\rho(x)} L(u)u\,\mathrm{d}u, \tag{2.9}$$

and

$$\|f'(x)^{-1}f(x)\|\leqslant\rho(x)+\frac{\int_0^{\rho(x)}L(u)u\,\mathrm{d}u}{1-\int_0^{\rho(x)}L(u)\,\mathrm{d}u}. \tag{2.10}$$

If $f'(x^*)^{-1}f'$ satisfies the center Lipschitz condition with the L average:

$$\|f'(x^*)^{-1}f'(x^\tau)-I\|\leqslant\int_0^{\tau\rho(x)}L(u)\,\mathrm{d}u,\quad\forall x\in B(x^*,r),\quad 0\leqslant\tau\leqslant 1, \tag{2.11}$$

where L is positive integrable function, then we have

$$\int_0^1\|f'(x^*)^{-1}f'(x^\tau)-I\|\rho(x)\,\mathrm{d}\tau\leqslant\int_0^{\rho(x)}L(u)(\rho(x)-u)\,\mathrm{d}u, \tag{2.12}$$

and

$$\|f'(y)^{-1}f(x)\|\leqslant\frac{\|x-x^*\|+\int_0^{\rho(x)}L(u)(\rho(x)-u)\,\mathrm{d}u}{1-\int_0^{\rho(y)}L(u)\,\mathrm{d}u},\quad\forall x,y\in B(x^*,r). \tag{2.13}$$

Proof. The Lipschitz conditions (2.8) and (2.11), respectively, imply that

$$\int_0^1\|f'(x^*)^{-1}(f'(x)-f'(x^\tau))\|\rho(x)\,\mathrm{d}\tau\leqslant\int_0^1\int_{\tau\rho(x)}^{\rho(x)}L(u)\,\mathrm{d}u\rho(x)\,\mathrm{d}\tau=\int_0^{\rho(x)}L(u)u\,\mathrm{d}u,$$

$$\int_0^1\|f'(x^*)^{-1}f'(x^\tau)-I\|\rho(x)\,\mathrm{d}\tau\leqslant\int_0^1\int_0^{\tau\rho(x)}L(u)\,\mathrm{d}u\rho(x)\,\mathrm{d}\tau=\int_0^{\rho(x)}L(u)(\rho(x)-u)\,\mathrm{d}u.$$

This proves (2.9) and (2.12).

Note that

$$f'(y)^{-1}f'(x^\tau)=I-f'(y)^{-1}f'(x^*)f'(x^*)^{-1}(f'(y)-f'(x^\tau)) \tag{2.14}$$

and

$$f'(y)^{-1}f(x)=f'(y)^{-1}(f(x)-f(x^*))=\int_0^1 f'(y)^{-1}f'(x^\tau)\,\mathrm{d}\tau(x-x^*), \tag{2.15}$$

where $x^\tau=x^*+\tau(x^*-x)$.

If $f'(x^*)^{-1}f'$ satisfies the radius Lipschitz condition, by (2.7) and (2.9) we obtain

$$\begin{aligned}\|f'(x)^{-1}f(x)\| &= \left\|\int_0^1 (I - f'(x)^{-1}f'(x^*)f'(x^*)^{-1}(f'(x) - f'(x^\tau)))\,\mathrm{d}\tau(x - x^*)\right\| \\ &\leqslant \left(1 + \|f'(x)^{-1}f'(x^*)\| \int_0^1 \|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\|\,\mathrm{d}\tau\right)\|x - x^*\| \\ &\leqslant \|x - x^*\| + \frac{\int_0^{\rho(x)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x)} L(u)\,\mathrm{d}u}.\end{aligned}$$

This proves (2.10).

If $f'(x^*)^{-1}f'$ satisfies the center Lipschitz condition, when $x, y \in B(x^*, r)$, we have

$$\begin{aligned}\|f'(y)^{-1}f(x)\| &= \left\|\int_0^1 (I - f'(y)^{-1}f'(x^*)f'(x^*)^{-1}(f'(y) - f'(x^\tau)))\,\mathrm{d}\tau(x - x^*)\right\| \\ &\leqslant \left(1 + \|f'(y)^{-1}f'(x^*)\| \int_0^1 \|f'(x^*)^{-1}(f'(y) - f'(x^\tau))\|\,\mathrm{d}\tau\right)\|x - x^*\| \\ &\leqslant \left(1 + \|f'(y)^{-1}f'(x^*)\| \int_0^1 \|f'(x^*)^{-1}(f'(y) - f'(x^*))\|\,\mathrm{d}\tau\right. \\ &\quad \left. + \int_0^1 \|f'(x^*)^{-1}(f'(x^*) - f'(x^\tau))\|\,\mathrm{d}\tau\right)\|x - x^*\| \\ &\leqslant \|x - x^*\| + \frac{\int_0^{\rho(y)} L(u)\,\mathrm{d}u\|x - x^*\| + \int_0^{\rho(x)} L(u)(\rho(x) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(y)} L(u)\,\mathrm{d}u} \\ &\leqslant \frac{\|x - x^*\| + \int_0^{\rho(x)} L(u)(\rho(x) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(y)} L(u)\,\mathrm{d}u}.\end{aligned}$$

This proves (2.13). □

Lemma 2.3. *Let*

$$h(t) = \frac{1}{t^\alpha}\int_0^t L(u)u^{\alpha-1}\,\mathrm{d}u, \quad \alpha \geqslant 1, \quad 0 \leqslant t \leqslant r, \tag{2.16}$$

where $L(u)$ is a positive integrable function and nondecreasing monotonically in $[0, r]$. Then $h(t)$ is nondecreasing with respect to t.

Proof. In fact, by the monotonicity of L, $\alpha\geqslant 1$, we obtain

$$\begin{aligned} h(t_2)-h(t_1) &= \left(\frac{1}{t_2^\alpha}\int_0^{t_2} - \frac{1}{t_1^\alpha}\int_0^{t_1}\right) L(u)u^{\alpha-1}\,\mathrm{d}u \\ &= \left(\frac{1}{t_2^\alpha}\int_{t_1}^{t_2} + \left(\frac{1}{t_2^\alpha}-\frac{1}{t_1^\alpha}\right)\int_0^{t_1}\right) L(u)u^{\alpha-1}\,\mathrm{d}u \\ &\geqslant L(t_1)\left(\frac{1}{t_2^\alpha}\int_{t_1}^{t_2} + \left(\frac{1}{t_1^\alpha}-\frac{1}{t_2^\alpha}\right)\int_0^{t_1}\right) u^{\alpha-1}\,\mathrm{d}u \\ &= L(t_1)\left(\frac{1}{t_2^\alpha}\int_0^{t_2} - \frac{1}{t_1^\alpha}\int_0^{t_1}\right) u^{\alpha-1}\,\mathrm{d}u = 0 \end{aligned}$$

for $0<t_1<t_2$. Thus, $h(t)=(1/t^\alpha)\int_0^t L(u)u^{\alpha-1}\,\mathrm{d}u$ is nondecreasing with respect to t. □

Lemma 2.4. *Let*

$$\varphi(t)=\frac{1}{t^2}\int_0^t L(u)(\alpha t-u)\,\mathrm{d}u, \quad \alpha\geqslant 1, \quad 0\leqslant t\leqslant r, \tag{2.17}$$

where $L(u)$ is a positive integrable function and nondecreasing monotonically in $[0,r]$. Then $\varphi(t)$ is nondecreasing monotonically with respect to t.

Proof. In fact, since L is a nondecreasing function, when $0<t_1<t_2\leqslant r$, $\alpha\geqslant 1$, we have

$$\begin{aligned} \varphi(t_2)-\varphi(t_1) &= \frac{1}{t_2^2}\int_0^{t_2} L(u)(\alpha t_2-u)\,\mathrm{d}u - \frac{1}{t_1^2}\int_0^{t_1} L(u)(\alpha t_1-u)\,\mathrm{d}u \\ &= (\alpha-1)\left(\frac{1}{t_2}\int_0^{t_2} L(u)\,\mathrm{d}u - \frac{1}{t_1}\int_0^{t_1} L(u)\,\mathrm{d}u\right) \\ &\quad + \frac{1}{t_2^2}\int_0^{t_2} L(u)(t_2-u)\,\mathrm{d}u - \frac{1}{t_1^2}\int_0^{t_1} L(u)(t_1-u)\,\mathrm{d}u \\ &= (\alpha-1)\left(\frac{1}{t_2}\int_0^{t_2} L(u)\,\mathrm{d}u - \frac{1}{t_1}\int_0^{t_1} L(u)\,\mathrm{d}u\right) \\ &\quad + \left(\frac{1}{t_1^2}-\frac{1}{t_2^2}\right)\int_0^{\frac{t_1}{2}} \left(L\left(\frac{t_1}{2}+u\right)-L\left(\frac{t_1}{2}-u\right)\right) u\,\mathrm{d}u \\ &\quad + \frac{(t_2-t_1)^2}{2t_2^2 t_1}\int_0^{t_1}(L(t_1)-L(u))\,\mathrm{d}u \\ &\quad + \frac{1}{t_2^2}\int_{t_1}^{t_2}(L(u)-L(t_1))(t_2-u)\,\mathrm{d}u \geqslant 0. \end{aligned}$$

Thus, $\varphi(t)=(1/t^2)\int_0^t L(u)(\alpha t-u)\,\mathrm{d}u$ is nondecreasing monotonically with respect to t. □

3. Convergence of inexact Newton methods with scale residual control

First we examine inexact Newton and modified inexact Newton methods that correspond to $B_k = f'(y_k)$ with $y_k = x_k$ and $y_k = x_0$, for each k, respectively.

Theorem 3.1. *Suppose x^* satisfies* (1.1), *f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists and $f'(x^*)^{-1} f'(x)$ satisfies the radius Lipschitz condition with L average*:

$$\| f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| \leqslant \int_{\tau\rho(x)}^{\rho(x)} L(u)\,\mathrm{d}u, \quad 0 \leqslant \tau \leqslant 1, \tag{3.1}$$

where $x^\tau = x^* + \tau(x - x^*)$, $\rho(x) = \|x - x^*\|$, and L is nondecreasing. Assume $B_k = f'(x_k)$, $\forall k$ in (1.3), *$\upsilon_k = \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k\, \mathrm{cond}(P_k f'(x_k))$ with $\upsilon_k \leqslant \upsilon < 1$. Let $r > 0$ satisfy*

$$(1 + \upsilon)\frac{\int_0^r L(u)u\,\mathrm{d}u}{r(1 - \int_0^r L(u)\,\mathrm{d}u)} + \upsilon \leqslant 1. \tag{3.2}$$

Then inexact Newton method is convergent for all $x_0 \in B(x^*, r)$ and

$$\|x_{k+1} - x^*\| \leqslant \left((1 + \upsilon)\frac{\int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_k) + \upsilon\right)\|x_k - x^*\|, \tag{3.3}$$

where

$$q = (1 + \upsilon)\frac{\int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} + \upsilon \tag{3.4}$$

is less than 1.

Proof. Arbitrarily choosing $x_0 \in B(x^*, r)$, where r satisfies (3.2), then q determined by (3.4) is less than 1. In fact, by the monotonicity of L and Lemma 2.3, we have

$$\begin{aligned} q &= (1 + \upsilon)\frac{\int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_0) + \upsilon \\ &< (1 + \upsilon)\frac{\int_0^r L(u)u\,\mathrm{d}u}{r^2(1 - \int_0^r L(u)\,\mathrm{d}u)} r + \upsilon \leqslant 1. \end{aligned}$$

Now if $x_k \in B(x^*, r)$, we have by (1.3)

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - f'(x_k)^{-1}(f(x_k) - f(x^*)) + f'(x_k)^{-1} r_k \\ &= f'(x_k)^{-1} f'(x^*) \int_0^1 f'(x^*)^{-1}(f'(x_k) - f'(x^\tau))(x_k - x^*)\,\mathrm{d}\tau + f'(x_k)^{-1} P_k^{-1} P_k r_k \end{aligned}$$

where $x^\tau = x^* + \tau(x_k - x^*)$. Hence, by Lemmas 2.1 and 2.2 and condition (3.1) we obtain

$$\begin{aligned} \|x_{k+1} - x^*\| &\leqslant \|f'(x_k)^{-1} f'(x^*)\| \int_0^1 \|f'(x^*)^{-1}(f'(x_k) - f'(x^\tau))\| \cdot \|x_k - x^*\|\,\mathrm{d}\tau \\ &\quad + \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f(x_k)\| \\ &\leqslant \frac{1}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \int_0^1 \int_{\tau\rho(x_k)}^{\rho(x_k)} L(u)\,\mathrm{d}u\,\rho(x_k)\,\mathrm{d}\tau \\ &\quad + \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k) f'(x_k)^{-1} f(x_k)\| \\ &\leqslant \frac{\int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + \theta_k \operatorname{cond}(P_k f'(x_k)) \left(\|x_k - x^*\| + \frac{\int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \right) \\ &\leqslant (1 + \upsilon_k) \frac{\int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + \upsilon_k \|x_k - x^*\|. \end{aligned}$$

Taking $k = 0$ above, we obtain $\|x_1 - x^*\| \leqslant q\|x_0 - x^*\| < \|x_0 - x^*\|$. Hence, $x_1 \in B(x^*, r)$, this shows that (1.3) can be continued an infinite number of times. By mathematical induction, all x_k belong to $B(x^*, r)$ and $\rho(x_k) = \|x_k - x^*\|$ decreases monotonically. Therefore, for all $k \geqslant 0$, we have

$$\begin{aligned} \|x_{k+1} - x^*\| &\leqslant (1 + \upsilon_k) \frac{\int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{\rho(x_k)^2 (1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 + \upsilon_k \rho(x_k) \\ &\leqslant \left((1 + \upsilon) \frac{\int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k) + \upsilon \right) \rho(x_k). \end{aligned}$$

Thus (3.3) follows. □

Theorem 3.2. *Suppose* x^* *satisfies* (1.1), *f has a continuous derivative in* $B(x^*, r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1} f'(x)$ *satisfies the center Lipschitz condition with L average*:

$$\|f'(x^*)^{-1} f'(x^\tau) - I\| \leqslant \int_0^{\tau\rho(x)} L(u)\,\mathrm{d}u, \quad 0 \leqslant \tau \leqslant 1, \tag{3.5}$$

where $x^\tau = x^* + \tau(x - x^*)$, $\rho(x) = \|x - x^*\|$, *and L is nondecreasing. Assume* $B_k = f'(x_0)$, $\forall k$ *in* (1.3), $\upsilon_k = \theta_k \|(P_0 f'(x_0))^{-1}\| \cdot \|P_0 f'(x_0)\| = \theta_k \operatorname{cond}(P_0 f'(x_0))$ *with* $\upsilon_k \leqslant \upsilon < 1$. *Let* $r > 0$ *satisfy*

$$(1+\upsilon)\frac{\int_0^r L(u)(r-u)\,\mathrm{d}u}{r(1-\int_0^r L(u)\,\mathrm{d}u)} + \frac{\upsilon + \int_0^r L(u)\,\mathrm{d}u}{1-\int_0^r L(u)\,\mathrm{d}u} \leqslant 1. \tag{3.6}$$

Then modified inexact Newton method is convergent for all $x_0 \in B(x^*, r)$ *and*

$$\|x_{k+1} - x^*\| \leqslant \left((1+\upsilon_k)\frac{\int_0^{\rho(x_0)} L(u)(\rho(x_0)-u)\,\mathrm{d}u}{\rho(x_0)^2(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_k) + \frac{\upsilon_k + \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}{1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}\right)\|x_k - x^*\|, \tag{3.7}$$

where

$$q = (1+\upsilon)\frac{\int_0^{\rho(x_0)} L(u)(\rho(x_0)-u)\,\mathrm{d}u}{\rho(x_0)(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} + \frac{\upsilon + \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}{1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \tag{3.8}$$

is less than 1.

Proof. Arbitrarily choosing $x_0 \in B(x^*, r)$, where r satisfies (3.6), then q determined by (3.8) is less than 1. In fact, by the monotonicity of L and Lemma 2.4, we have

$$\begin{aligned} q &= (1+\upsilon)\frac{\int_0^{\rho(x_0)} L(u)(\rho(x_0)-u)\,\mathrm{d}u}{\rho(x_0)^2(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_0) + \frac{\upsilon + \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}{1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \\ &< (1+\upsilon)\frac{\int_0^r L(u)(r-u)\,\mathrm{d}u}{r^2(1-\int_0^r L(u)\,\mathrm{d}u)}r + \frac{\upsilon + \int_0^r L(u)\,\mathrm{d}u}{1-\int_0^r L(u)\,\mathrm{d}u} \leqslant 1. \end{aligned}$$

Now if $x_k \in B(x^*, r)$, we have by (1.3)

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - f'(x_0)^{-1}(f(x_k) - f(x^*)) + f'(x_0)^{-1} r_k \\ &= f'(x_0)^{-1} f'(x^*) \int_0^1 f'(x^*)^{-1}(f'(x_0) - f'(x^\tau))(x_k - x^*)\,\mathrm{d}\tau + f'(x_0)^{-1} P_k^{-1} P_k r_k, \end{aligned}$$

where $x^\tau = x^* + \tau(x_k - x^*)$. Hence, by Lemmas 2.1 and 2.2 and condition (3.5) we obtain

$$\begin{aligned}
\|x_{k+1} - x^*\| \leqslant & \|f'(x_0)^{-1} f'(x^*)\| \int_0^1 \|f'(x^*)^{-1}(f'(x_0) - f'(x^\tau))\| \cdot \|x_k - x^*\| \,\mathrm{d}\tau \\
& + \theta_k \|(P_0 f'(x_0))^{-1}\| \cdot \|P_0 f(x_0)\| \\
\leqslant & \|f'(x_0)^{-1} f'(x^*)\| \int_0^1 (\|f'(x^*)^{-1}(f'(x_0) - f'(x^*))\| \\
& + \|f'(x^*)^{-1}(f'(x^*) - f'(x^\tau))\|) \\
& \cdot \|x_k - x^*\| \,\mathrm{d}\tau + \theta_k \|(P_0 f'(x_0))^{-1}\| \|P_0 f(x_0)\| \\
\leqslant & \frac{1}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \int_0^1 \left(\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u + \int_0^{\tau\rho(x_k)} L(u)\,\mathrm{d}u \right) \rho(x_k)\,\mathrm{d}\tau \\
& + \theta_k \|(P_0 f'(x_0))^{-1}\| \cdot \|P_0 f'(x_0) f'(x_0)^{-1} f(x_k)\| \\
\leqslant & \frac{\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u \rho(x_k) + \int_0^{\rho(x_k)} L(u)(\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} + \theta_k \operatorname{cond}(P_0 f'(x_0)) \\
& \times \left(\frac{1 + \int_0^{\rho(x_k)} L(u)(\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \right) \|x_k - x^*\| \\
\leqslant & \frac{\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u \rho(x_k) + \int_0^{\rho(x_k)} L(u)(\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} + \upsilon_k \\
& \times \left(\frac{\|x_k - x^*\| + \int_0^{\rho(x_k)} L(u)(\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \right).
\end{aligned}$$

Taking $k=0$ above, we obtain $\|x_1 - x^*\| \leqslant q\|x_0 - x^*\| < \|x_0 - x^*\|$. Hence, $x_1 \in B(x^*, r)$, this shows that (1.3) can be continued an infinite number of times. By mathematical induction, all x_k belong to $B(x^*, r)$ and $\rho(x_k) = \|x_k - x^*\|$ decreases monotonically. Therefore, for all $k = 0, 1, \ldots,$ we have

$$\begin{aligned}
\|x_{k+1} - x^*\| \leqslant & (1 + \upsilon_k) \frac{\int_0^{\rho(x_k)} L(u)(\rho(x_k) - u)\,\mathrm{d}u}{\rho(x_k)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 + \frac{\upsilon_k + \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \rho(x_k) \\
\leqslant & \left((1 + \upsilon) \frac{\int_0^{\rho(x_0)} L(u)(\rho(x_0) - u)\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k) + \frac{\upsilon + \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u} \right) \rho(x_k).
\end{aligned}$$

Thus (3.7) follows. □

Theorems 3.1 and 3.2 give an estimate of the radii of convergence ball for inexact Newton method and modified inexact Newton methods, respectively. In particular, for $\upsilon = 0$, the estimate for the radius of convergence ball for Newton's method is given by

$$\frac{\int_0^r L(u)u\,\mathrm{d}u}{r(1-\int_0^r L(u)\,\mathrm{d}u)} \leqslant 1$$

which can be found in [18]. Then, we can conclude that vanishing residuals, Theorem 3.1 merges into the theory of Newton's method.

A result analogous to Theorems 3.1 and 3.2 can also be proven for inexact Newton-like methods where $B_k = B(x_k)$ approximates $f'(x_k)$.

Theorem 3.3. *Suppose x^* satisfies* (1.1), *f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists and $f'(x^*)^{-1}f'$ satisfies the radius Lipschitz condition* (3.1) *with L is nondecreasing. Let $B(x)$ be an approximation to the $f'(x)$ for all $x \in B(x^*, r)$, $B(x)$ is invertible and*

$$\|B(x)^{-1}F'(x)\| \leqslant \omega_1, \quad \|B(x)^{-1}F'(x) - I\| \leqslant \omega_2. \tag{3.9}$$

Where $\upsilon_k = \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k \operatorname{cond}(P_k f'(x_k))$ with $\upsilon_k \leqslant \upsilon < 1$. Let $r > 0$ satisfy

$$(1+\upsilon)\frac{\omega_1 \int_0^r L(u)u\,\mathrm{d}u}{r(1-\int_0^r L(u)\,\mathrm{d}u)} + \omega_2 + \omega_1\upsilon \leqslant 1. \tag{3.10}$$

Then inexact Newton-like method is convergent for all $x_0 \in B(x^*, r)$ and

$$\|x_{k+1} - x^*\| \leqslant \left((1+\upsilon)\frac{\omega_1 \int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_k) + \omega_2 + \omega_1\upsilon\right)\|x_k - x^*\|, \tag{3.11}$$

where

$$q = (1+\upsilon)\frac{\omega_1 \int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} + \omega_2 + \omega_1\upsilon \tag{3.12}$$

is less than 1.

Proof. Arbitrarily choosing $x_0 \in B(x^*, r)$, where r satisfies (3.10), then q determined by (3.12) is less than 1. In fact, by the monotonicity of L and Lemma 2.3 , we have

$$\begin{aligned} q &= (1+\upsilon)\frac{\omega_1 \int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_0) + \omega_2 + \omega_1\upsilon \\ &< (1+\upsilon)\frac{\omega_1 \int_0^r L(u)u\,\mathrm{d}u}{r^2(1-\int_0^r L(u)\,\mathrm{d}u)}r + \omega_2 + \omega_1\upsilon \leqslant 1. \end{aligned}$$

Now if $x_k \in B(x^*, r)$, we have by (1.3)

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - B_k^{-1}(f(x_k) - f(x^*)) + B_k^{-1} r_k \\ &= x_k - x^* - \int_0^1 B_k^{-1} f'(x^\tau)\,\mathrm{d}t\,(x_k - x^*) + B_k^{-1} P_k^{-1} P_k r_k \\ &= -B_k^{-1} f'(x_k) \int_0^1 f'(x_k)^{-1} f'(x^*)(f'(x^*)^{-1}(f'(x_k) - f'(x^\tau)))(x_k - x^*)\,\mathrm{d}\tau \\ &\quad + B_k^{-1}(f'(x_k) - B_k)(x_k - x^*) + B_k^{-1} P_k^{-1} P_k r_k, \end{aligned}$$

where $x^\tau = x^* + \tau(x_k - x^*)$. Hence, by Lemmas 2.1and 2.2 and condition (3.1) we obtain

$$\begin{aligned} \|x_{k+1} - x^*\| &\leqslant \|B_k^{-1} f'(x_k)\| \int_0^1 \|f'(x_k)^{-1} f'(x^*)\| \|f'(x^*)^{-1}(f'(x_k) - f'(x^\tau))\| \cdot \|x_k - x^*\|\,\mathrm{d}\tau \\ &\quad + \|B_k^{-1}(f'(x_k) - B_k)\| \cdot \|x_k - x^*\| + \theta_k \|B_k^{-1} P_k^{-1}\| \|P_k f(x_k)\| \\ &\leqslant \frac{\omega_1}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \int_0^1 \int_{\tau\rho(x_k)}^{\rho(x_k)} L(u)\,\mathrm{d}u\,\rho(x_k)\,\mathrm{d}\tau + \omega_2 \rho(x_n) \\ &\quad + \theta_k \|B_k^{-1} f'(x_k)\| \|(P_k f'(x_k))^{-1}\| \|P_k f'(x_k)\| \left\| \int_0^1 f'(x_k)^{-1} f(x_k)\,\mathrm{d}\tau(x_k - x^*) \right\| \\ &\leqslant \frac{\omega_1 \int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + \omega_2 \rho(x_k) + \omega_1 \upsilon_k \left(\rho(x_k) + \frac{\int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \right) \\ &\leqslant (1 + \upsilon_k) \frac{\omega_1 \int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + (\omega_2 + \omega_1 \upsilon_k)\rho(x_k). \end{aligned}$$

Taking $k=0$ above, we obtain $\|x_1 - x^*\| \leqslant q\|x_0 - x^*\| < \|x_0 - x^*\|$. Hence, $x_1 \in B(x^*, r)$, this shows that (1.3) can be continued an infinite number of times. By mathematical induction, all x_n belong to $B(x^*, r)$ and $\rho(x_k) = \|x_k - x^*\|$ decreases monotonically. Therefore, for all $k \geqslant 0$, we have

$$\begin{aligned} \|x_{k+1} - x^*\| &\leqslant (1 + \upsilon_k) \frac{\omega_1 \int_0^{\rho(x_k)} L(u)u\,\mathrm{d}u}{\rho(x_k)^2 (1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 + (\omega_2 + \omega_1 \upsilon_k)\rho(x_k) \\ &\leqslant \left((1 + \upsilon) \frac{\omega_1 \int_0^{\rho(x_0)} L(u)u\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k) + \omega_2 + \omega_1 \upsilon \right) \rho(x_k). \end{aligned}$$

Thus (3.11) follows. □

Remark 3.1. The results we proved state inverse proportionality between $\mathrm{cond}(P_k B_k)$ and each forcing term θ_k. Such conditions are sufficient for convergence, and may be overly restrictive for the upper bounds on $\{\theta_k\}$, if $P_k B_k$ are bad conditioned matrices. But we can choose appropriate P_k to precondition B_k which can lead to relaxation on the forcing terms. The more properties about scaling residual control can be found in [13].

4. Applications

In the study of the Newton's method, the assumption that the derivative is Lipschitz continuous is considered traditional. In this section, we will apply the obtained results to some concrete cases. By taking L as a constant, the following corollaries are obtained under Lipschitz conditions (3.1) and (3.5) directly.

Corollary 4.1. *Suppose x^* satisfies* (1.1), *f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists and $f'(x^*)^{-1}f'(x)$ satisfies the radius Lipschitz condition with L average.*

$$\|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| \leqslant (1-\tau)L\|x - x^*\|, \quad 0 \leqslant \tau \leqslant 1, \tag{4.1}$$

where $x^\tau = x^* + \tau(x - x^*)$, $\rho(x) = \|x - x^*\|$. Assume $B_k = f'(x_k)$, $\forall k$ in (1.3), *$\upsilon_k = \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k \,\mathrm{cond}(P_k f'(x_k))$ with $\upsilon_k \leqslant \upsilon < 1$. Let $r > 0$ satisfy*

$$r = \frac{2(1-\upsilon)}{L(3-\upsilon)}. \tag{4.2}$$

Then inexact Newton method is convergent for all $x_0 \in B(x^*, r)$,

$$q = \upsilon + \frac{L\|x_0 - x^*\|(1+\upsilon)}{2(1 - L\|x_0 - x^*\|)} < 1, \tag{4.3}$$

and inequality (3.3) *holds.*

Corollary 4.2. *Suppose x^* satisfies* (1.1), *f has a continuous derivative in $B(x^*, r)$, $f'(x^*)^{-1}$ exists and $f'(x^*)^{-1}f'(x)$ satisfies the center Lipschitz condition with L average*:

$$\|f'(x^*)^{-1}f'(x^\tau) - I\| \leqslant \tau L \|x - x^*\|, \quad 0 \leqslant \tau \leqslant 1, \tag{4.4}$$

where $x^\tau = x^* + \tau(x - x^*)$, $\rho(x) = \|x - x^*\|$, and L is nondecreasing. Assume $B_k = f'(x_0)$, $\forall k$ in (1.3), *$\upsilon_k = \theta_k \|(P_0 f'(x_0))^{-1}\| \cdot \|P_0 f'(x_0)\| = \theta_k \,\mathrm{cond}(P_0 f'(x_0))$ with $\upsilon_k \leqslant \upsilon < 1$. Let $r > 0$ satisfy*

$$r = \frac{2(1-\upsilon)}{L(5+\upsilon)}. \tag{4.5}$$

Then modified inexact Newton method is convergent for all $x_0 \in B(x^*, r)$,

$$q = \frac{L\|x_0 - x^*\| + \upsilon}{1 - L\|x_0 - x^*\|} + \frac{L\|x_0 - x^*\|(1+\upsilon)}{2(1 - L\|x_0 - x^*\|)} < 1, \tag{4.6}$$

and inequality (3.7) *holds.*

Corollary 4.3. *Suppose* x^* *satisfies* (1.1), f *has a continuous derivative in* $B(x^*, r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1}f'$ *satisfies the radius Lipschitz condition* (4.1) *with L is positive number. Assume* $B(x)$ *be an approximation to the* $f(x)$ *for all* $x \in B(x^*, r)$, $B(x)$ *satisfies condition* (3.9), $\upsilon_k = \theta_k\|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k \operatorname{cond}(P_k f'(x_k))$ *with* $\upsilon_k \leqslant \upsilon < 1$. *Let* $r > 0$ *satisfy*

$$r = \frac{2(1 - \upsilon\omega_1 - \omega_2)}{L(2 + \omega_1 - \upsilon\omega_1 - 2\omega_2)}. \tag{4.7}$$

Then inexact Newton-like method is convergent for all $x_0 \in B(x^*, r)$,

$$q = \upsilon\omega_1 + \omega_2 + \frac{L\|x_0 - x^*\|\omega_1(1+\upsilon)}{2(1 - L\|x_0 - x^*\|)} < 1, \tag{4.8}$$

and the inequality (3.11) *holds.*

Remark 4.1. The results of Corollary 4.1 can be found in [13], and if taking $\upsilon = 0$ in Corollary 4.1, Wang (see [15,16,18]) also gave the similar results of Newton's method. But it seems that Corollaries 4.2 and 4.3 have not appeared in the literature.

5. Convergence under weaker Lipschitz condition

In this section, we will consider the system of nonlinear equations (1.1) under weaker Lipschitz condition.

In Section 3, we studied the inexact Newton method and inexact Newton-like method under condition (3.1). In fact, similar to Theorem 3.2, we can also give the convergence of the inexact Newton method and inexact Newton-like method with Lipschitz condition (3.5).

Theorem 5.1. *Suppose* x^* *satisfies* (1.1), f *has a continuous derivative in* $B(x^*, r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1}f'$ *satisfies the radius Lipschitz condition* (3.5) *with L is nondecreasing. Let* $B(x)$ *be an approximation to the* $f'(x)$ *for all* $x \in B(x^*, r)$, $B(x)$ *is invertible and*

$$\|B(x)^{-1}F'(x)\| \leqslant \omega_1, \quad \|B(x)^{-1}F'(x) - I\| \leqslant \omega_2. \tag{5.1}$$

Where $\upsilon_k = \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k \operatorname{cond}(P_k f'(x_k))$ *with* $\upsilon_k \leqslant \upsilon < 1$. *Let* $r > 0$ *satisfy*

$$\omega_1 \frac{\int_0^r L(u)(2r + (\upsilon - 1)u)\,\mathrm{d}u}{r(1 - \int_0^r L(u)\,\mathrm{d}u)} + \omega_2 + \omega_1 \upsilon \leqslant 1. \tag{5.2}$$

Then inexact Newton-like method is convergent for all $x_0 \in B(x^*, r)$ *and*

$$\|x_{k+1} - x^*\| \leqslant \left(\omega_1 \frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0) + (\upsilon - 1)u)\,\mathrm{d}u}{\rho(x_0)^2(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k) + \omega_2 + \omega_1 \upsilon \right) \|x_k - x^*\|, \tag{5.3}$$

where

$$q = \omega_1 \frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0) + (\upsilon - 1)u)\,\mathrm{d}u}{\rho(x_0)(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} + \omega_2 + \omega_1 \upsilon \tag{5.4}$$

is less than 1.

Proof. Arbitrarily choosing $x_0 \in B(x^*, r)$, where r satisfies (5.2), then q determined by (5.4) is less than 1. In fact, by the monotonicity of L and Lemmas 2.3 and 2.4, we have

$$\begin{aligned} q &= \omega_1 \frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0) + (\upsilon - 1)u)\,\mathrm{d}u}{\rho(x_0)^2(1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_0) + \omega_2 + \omega_1 \upsilon \\ &< \omega_1 \frac{\int_0^r L(u)(2r + (\upsilon - 1)u)\,\mathrm{d}u}{r^2(1 - \int_0^r L(u)\,\mathrm{d}u)} r + \omega_2 + \omega_1 \upsilon \leqslant 1. \end{aligned}$$

Now if $x_k \in B(x^*, r)$, we have by (1.3)

$$\begin{aligned} x_{k+1} - x^* &= x_k - x^* - B_k^{-1}(f(x_k) - f(x^*)) + B_k^{-1} r_k \\ &= x_k - x^* - \int_0^1 B_k^{-1} f'(x^\tau)\,\mathrm{d}t\,(x_k - x^*) + B_k^{-1} P_k^{-1} P_k r_k \\ &= -B_k^{-1} f'(x_k) \int_0^1 f'(x_k)^{-1} f'(x^*)(f'(x^*)^{-1}(f'(x_k) - f'(x^\tau)))(x_k - x^*)\,\mathrm{d}\tau \\ &\quad + B_k^{-1}(f'(x_k) - B_k)(x_k - x^*) + B_k^{-1} P_k^{-1} P_k r_k, \end{aligned}$$

where $x^\tau = x^* + \tau(x_k - x^*)$. Hence, by Lemmas 2.1 and 2.2 and condition (3.5) we obtain

$$\begin{aligned}\|x_{k+1}-x^*\| \leqslant\ & \|B_k^{-1} f'(x_k)\| \int_0^1 \|f'(x_k)^{-1} f'(x^*)\| \|f'(x^*)^{-1}(f'(x_k) - f'(x^\tau))\| \cdot \|x_k - x^*\| \,\mathrm{d}\tau \\ & + \|B_k^{-1}(f'(x_k) - B_k)\| \cdot \|x_k - x^*\| + \theta_k \|B_k^{-1} P_k^{-1}\| \|P_k f(x_k)\| \\ \leqslant\ & \|B_k^{-1} f'(x_k)\| \cdot \|f'(x_k)^{-1} f'(x^*)\| \int_0^1 (\|f'(x^*)^{-1}(f'(x_k) - f'(x^*))\| \\ & + \|f'(x^*)^{-1}(f'(x^*) - f'(x^\tau))\|) \cdot \|x_k - x^*\| \,\mathrm{d}\tau \\ & + \|B_k^{-1}(f'(x_k) - B_k)\| \cdot \|x_k - x^*\| + \theta_k \|B_k^{-1} P_k^{-1}\| \|P_k f(x_k)\| \\ \leqslant\ & \frac{\omega_1}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \int_0^1 \left(\int_0^{\rho(x_k)} L(u)\,\mathrm{d}u + \int_0^{\tau\rho(x_k)} L(u)\,\mathrm{d}u \right) \rho(x_k)\,\mathrm{d}\tau + \omega_2 \rho(x_n) \\ & + \theta_k \|B_k^{-1} f'(x_k)\| \|(P_k f'(x_k))^{-1}\| \|P_k f'(x_k)\| \left\| \int_0^1 f'(x_k)^{-1} f(x_k)\,\mathrm{d}\tau (x_k - x^*) \right\| \\ \leqslant\ & \omega_1 \frac{\int_0^{\rho(x_k)} L(u)(2\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + \omega_2 \rho(x_k) + \omega_1 \upsilon_k \left(\rho(x_k) + \frac{\int_0^{\rho(x_k)} L(u) u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} \right) \\ \leqslant\ & \omega_1 \frac{\int_0^{\rho(x_k)} L(u)(2\rho(x_k) - u)\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + \omega_1 \upsilon_k \frac{\int_0^{\rho(x_k)} L(u) u\,\mathrm{d}u}{1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u} + (\omega_2 + \omega_1 \upsilon_k) \rho(x_k).\end{aligned}$$

Taking $k = 0$ above, we obtain $\|x_1 - x^*\| \leqslant q \|x_0 - x^*\| < \|x_0 - x^*\|$. Hence, $x_1 \in B(x^*, r)$, this shows that (1.3) can be continued an infinite number of times. By mathematical induction, all x_n belong to $B(x^*, r)$ and $\rho(x_k) = \|x_k - x^*\|$ decreases monotonically. Therefore, for all $k \geqslant 0$, we have

$$\begin{aligned}\|x_{k+1}-x^*\| \leqslant\ & \omega_1 \frac{\int_0^{\rho(x_k)} L(u)(2\rho(x_k) - u)\,\mathrm{d}u}{\rho(x_k)^2 (1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 + \omega_1 \upsilon_k \frac{\int_0^{\rho(x_k)} L(u) u\,\mathrm{d}u}{\rho(x_k)^2 (1 - \int_0^{\rho(x_k)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 \\ & + (\omega_2 + \omega_1 \upsilon_k) \rho(x_k) \\ \leqslant\ & \omega_1 \frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0) - u)\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 + \omega_1 \upsilon \frac{\int_0^{\rho(x_0)} L(u) u\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k)^2 \\ & + (\omega_2 + \omega_1 \upsilon) \rho(x_k) \\ \leqslant\ & \left(\omega_1 \frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0) + (\upsilon - 1) u)\,\mathrm{d}u}{\rho(x_0)^2 (1 - \int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)} \rho(x_k) + \omega_2 + \omega_1 \upsilon \right) \rho(x_k).\end{aligned}$$

Thus (5.3) follows. □

If taking $B(x)^{-1}=F'(x)$, i.e., $w_1=1$, $w_2=0$ in Theorem 5.1, we obtain the inexact Newton method under condition (3.5) immediately.

Theorem 5.2. *Suppose* x^* *satisfies* (1.1), *f has a continuous derivative in* $B(x^*,r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1}f'(x)$ *satisfies the radius Lipschitz condition with L average*:

$$\|f'(x^*)^{-1}(f'(x)-f'(x^\tau))\|\leqslant\int_{\tau\rho(x)}^{\rho(x)}L(u)\,\mathrm{d}u,\quad 0\leqslant\tau\leqslant 1, \tag{5.5}$$

where $x^\tau=x^*+\tau(x-x^*)$, $\rho(x)=\|x-x^*\|$, *and L is nondecreasing. Assume* $B_k=f'(x_k)$, $\forall k$ *in* (1.3), $\upsilon_k=\theta_k\|(P_kf'(x_k))^{-1}\|\cdot\|P_kf'(x_k)\|=\theta_k\,\mathrm{cond}(P_kf'(x_k))$ *with* $\upsilon_k\leqslant\upsilon<1$. *Let* $r>0$ *satisfy*

$$\frac{\int_0^r L(u)(2r+(\upsilon-1)u)\,\mathrm{d}u}{r(1-\int_0^r L(u)\,\mathrm{d}u)}+\upsilon\leqslant 1. \tag{5.6}$$

Then inexact Newton method is convergent for all $x_0\in B(x^*,r)$ *and*

$$\|x_{k+1}-x^*\|\leqslant\left(\frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0)+(\upsilon-1)u)\,\mathrm{d}u}{\rho(x_0)^2(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}\rho(x_k)+\upsilon\right)\|x_k-x^*\|, \tag{5.7}$$

where

$$q=\omega_1\frac{\int_0^{\rho(x_0)} L(u)(2\rho(x_0)+(\upsilon-1)u)\,\mathrm{d}u}{\rho(x_0)(1-\int_0^{\rho(x_0)} L(u)\,\mathrm{d}u)}+\upsilon \tag{5.8}$$

is less than 1.

In particular, taking L as a constant, we can obtain the following corollaries.

Corollary 5.1. *Suppose* x^* *satisfies* (1.1), *f has a continuous derivative in* $B(x^*,r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1}f'$ *satisfies the center Lipschitz condition* (4.4) *where L is positive number. Assume* $B(x)$ *be an approximation to the* $f(x)$ *for all* $x\in B(x^*,r)$, $B(x)$ *satisfies condition*(5.1),$\upsilon_k=\theta_k\|(P_kf'(x_k))^{-1}\|\cdot\|P_kf'(x_k)\|=\theta_k\,\mathrm{cond}(P_kf'(x_k))$ *with* $\upsilon_k\leqslant\upsilon<1$. *Let* $r>0$ *satisfy*

$$r=\frac{2(1-\upsilon\omega_1-\omega_2)}{L(2+3\omega_1-\upsilon\omega_1-2\omega_2)}. \tag{5.9}$$

Then inexact Newton-like method is convergent for all $x_0\in B(x^*,r)$,

$$q=\upsilon\omega_1+\omega_2+\frac{L\|x_0-x^*\|\omega_1(3+\upsilon)}{2(1-L\|x_0-x^*\|)}<1, \tag{5.10}$$

and the inequality (5.3) *holds.*

Corollary 5.2. *Suppose* x^* *satisfies* (1.1), *f has a continuous derivative in* $B(x^*, r)$, $f'(x^*)^{-1}$ *exists and* $f'(x^*)^{-1} f'(x)$ *satisfies the center Lipschitz condition* (4.4). *Assume* $B_k = f'(x_k)$, $\forall k$ *in* (1.3), $\upsilon_k = \theta_k \|(P_k f'(x_k))^{-1}\| \cdot \|P_k f'(x_k)\| = \theta_k \operatorname{cond}(P_k f'(x_k))$ *with* $\upsilon_k \leqslant \upsilon < 1$. *Let* $r > 0$ *satisfy*

$$r = \frac{2(1-\upsilon)}{L(5-\upsilon)}. \tag{5.11}$$

Then inexact Newton method is convergent for all $x_0 \in B(x^*, r)$,

$$q = \upsilon + \frac{L\|x_0 - x^*\|(3+\upsilon)}{2(1 - L\|x_0 - x^*\|)} < 1, \tag{5.12}$$

and inequality (5.7) *holds.*

Remark 5.1. The results in this section are all new under the center Lipschitz condition. Especially, Theorems 5.1, 5.2 and Corollaries 5.1, 5.2 improve the convergence conclusions of Newton's method [15,16,18] in the limiting case of vanishing residuals, i.e., $\theta_k = \omega_0 = 0$, $\omega_1 = 1$. The following example shows that the convergence result under the center Lipschitz condition is an essential improvement.

Example 5.1. Define a function

$$f(x) = \int_0^x \left(1 + \tau \cos\frac{\pi}{\tau}\right) \mathrm{d}\tau, \quad \forall x \in R.$$

Then

$$f'(x) = \begin{cases} 1 + x \cos\frac{\pi}{x}, & x \neq 0; \\ 1, & x = 0. \end{cases}$$

It is clear that $x^* = 0$ is a zero of f and $f'(x)$ satisfies

$$\|f'(x^*) f'(x) - I\| = \left| x \cos\frac{\pi}{x} \right| \leqslant |x - x^*|, \quad \forall x \in R.$$

It follows from Theorems 5.1 and 5.2 that for any $x_0 \in B(x^*, 2/5)$

$$\|x_{k+1} - x_*\| \leqslant \left(\frac{3|x_0|}{2(1-|x_0|)}\right) \|x_k - x_*\|^2, \quad k = 0, 1, 2, \ldots.$$

However, there is no positive integrable function L such that (3.5) or (4.4) is satisfied. In fact, notice that

$$\|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| = \left| x \cos\frac{\pi}{x} - \tau x \cos\frac{\pi}{\tau x} \right| = \frac{1}{n},$$

for $x = \frac{1}{n}$, $\tau = 2n/(2n+1)$ and $n = 1, 2, \ldots$. Thus, if there was a positive integrable function L such that (3.5) holds in the ball $B(x^*, r)$ for some $r > 0$, it follows that there exists some $n_0 > 1$ such that

$$\int_0^r L(u)\,\mathrm{d}u \geqslant \sum_{k=n_0}^{+\infty} \int_{2n/(2n+1)}^{\frac{1}{n}} L(u)\,\mathrm{d}u \geqslant \sum_{k=n_0}^{+\infty} \frac{1}{k} = +\infty,$$

which is a contradiction.

Table 1
Estimated radii

υ	Estimates (4.2) $r = 2(1-\upsilon)/L(3-\upsilon)$	Estimates (4.5) $r = 2(1-\upsilon)/L(5+\upsilon)$	Estimates (5.2) $r = 2(1-\upsilon)/L(5-\upsilon)$
0.9	0.14285714	0.05084746	0.07317073
0.7	0.39130435	0.15789474	0.20930233
0.5	0.60000000	0.27272727	0.33333333
0.4	0.69230769	0.33333333	0.39130435
0.2	0.85714286	0.46153846	0.50000000
0	1.00000000	0.60000000	0.60000000

Remark 5.2. Theorems 3.1–3.3, 5.1 and 5.2 give us a perceptive apprehension of convergence for inexact Newton methods. Giving approximately estimated radii of convergence ball helps to ensure that inexact Newton methods converge rapidly and is also important to choose an initial iterative point inexact Newton methods. Fortunately, only a rough estimation of radii for convergence ball is enough to do those. The following example will show us how to give estimated radii of convergence ball.

Example 5.2. Define a function

$$f(x) = \begin{cases} -x + \frac{1}{3}x^2, & 0 \leqslant x \leqslant 1; \\ -x - \frac{1}{3}x^2, & -1 \leqslant x < 0. \end{cases}$$

Then

$$f'(x) = \begin{cases} -1 + \frac{2}{3}x, & 0 \leqslant x \leqslant 1; \\ -1 - \frac{2}{3}x, & -1 \leqslant x < 0. \end{cases}$$

Obviously, $x^* = 0$ is a zero of f and $f'(x)$ satisfies

$$\|f'(x^*)^{-1}(f'(x) - f'(x^\tau))\| = \tfrac{2}{3}(1-\tau)|x - x^*|.$$

Hence, when using inexact methods and taking different υ, we get the results of the estimated radii in Table 1.

Moreover, we should choose initial iterative point $x_0 < r$, which can ensure rapid convergence. In fact, by taking $x_0 = 1$, and $\upsilon = 0$ in inexact methods, we have

$$\begin{aligned} x_1 &= x_0 - [f'(x_0)]^{-1} f(x_0) \\ &= x_0 - \frac{1}{-1 + \frac{2}{3}x_0}\left(-x_0 + \frac{1}{3}x_0^2\right) \\ &= x_0 - 2x_0 = -1, \end{aligned}$$

and $x_n = (-1)^n, \quad n = 1, 2, \ldots$. That is to say, x_0 on the boundary of the convergence ball can make the inexact methods fails.

6. Discussion for assumptions

In the previous three sections, Lipschitz assumptions (3.1), (3.5), (4.1) and (4.4) have been used. We remark that the conditions (3.1), (3.5), (4.1) and (4.4) are all affine invariants, as it is insensitive with respect to transformations of the mapping $f(x)$ of the form: $f(x) \longrightarrow Af(x), A$ an invertible matrix, as long as the same affine transformation is also valid for $B(x)$ (see e.g. inexact Newton method, modified inexact Newton method and inexact Newton-like methods).

Since Newton's iterates are affine invariant, in [6,19] convergence conditions were determined in affine invariant terms. With the $P_k = B_k^{-1}$ proposed in this paper, we point out that Theorems 3.1–3.3, 5.1 and 5.2 represent an affine convergence analysis of inexact Newton methods.

Under affine invariant Lipschitz condition:

$$\|f'(x^*)^{-1}(f'(x) - f'(y))\| \leqslant L\|x - y\|, \quad \forall x, y \in B(x^*, r), \tag{6.1}$$

the convergence analysis of inexact Newton method is given in [13]. It follows from Example 5.1 that the conditions (3.1), (3.5), (4.1) and (4.4) are essentially weaker than the condition (6.1). That is to say, under weaker affine invariant Lipschitz condition, Theorems 3.1–3.3, 5.1 and 5.2 show the convergence analysis of inexact Newton methods. Hence, Theorems 3.1–3.3, 5.1 and 5.2 really extend the results in [13,15,16,18] and expand the application fields of inexact Newton methods and Newton's method.

Acknowledgements

The authors thank the referees for their careful reading and valuable suggestions which considerably improved the original manuscript of this paper.

References

[1] Y. Chen, D.Y. Cai, Inexact overlapped block Broyden methods for solving nonlinear equations, Appl. Math. Comput. 136 (2003) 215–228.
[2] R.S. Dembo, S.C. Eisenstat, T. Steihaug, Inexact Newton methods, SIAM J. Numer. Anal. 19 (1982) 400–408.
[3] J.E. Dennis, On Newton-like methods, Numer. Math. 11 (1968) 324–330.
[4] J.E. Dennis, On the Kantorovich hypothesis for Newton's method, SIAM J. Numer. Anal. 6 (1969) 493–507.
[5] J.E. Dennis, R.B. Schnabel, Numerical methods for unconstrained optimization and nonlinear equations, Prentice-Hall, Englewood Cliffs, NJ, 1983.
[6] P. Deuflhard, G. Heindl, Affine invariant convergence theroem for Newton methods and extension to related methods, SIAM J. Numer. Anal. 16 (1979) 1–10.
[7] E. Galligani, The Newton-arithmetic mean method for the solution of systems of nonlinear equations, Appl. Math. Comput. 134 (2003) 9–34.
[8] K.R. Jackson, The numerical solution of large systems of stiff IVPs for ODEs, Appl. Numer. Math. 20 (1996) 5–20.
[9] L.V. Kantorovich, G.P. Akilov, Functional Analysis, second ed., Pergamon Press, Oxford, 1982.
[10] R. Kress, Numerical Analysis, Springer, New York, 1998.
[11] J.M. Martinez, L. Qi, Inexact Newton methods for solving nonsmooth equations, J. Comput. Appl. Math. 60 (1995) 127–145.
[12] G.J. Miel, Majorizing sequences and error bounds for iterative methods, Math. Comp. 34 (1980) 185–202.
[13] B. Morini, Convergence behaviour of inexact Newton method, Math. Comp. 68 (1999) 1605–1613.
[14] J.M. Ortega, W.C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[15] J.F. Traub, H. Wozniakowski, Convergence and complexity of Newton iteration, J. Assoc. Comput. Math. 29 (1979) 250–258.
[16] X. Wang, The Convergence on Newton's method, Kexue Tongbao (A Special Issue of Mathematics, Physics & Chemistry) 25 (1980) 36–37.
[17] X. Wang, Convergence of Newton's method and inverse function in Banach space, Math. Comp. 68 (1999) 169–186.
[18] X. Wang, Convergence of Newton's method and uniqueness of the solution of equations in Banach space, IMA J. Numer. Anal. 20 (2000) 123–134.
[19] T.J. Ypma, Affine invariant convergence results for Newton's method, BIT 22 (1982) 108–118.