Theorems on a Modified Newton Method for Solving Systems of Nonlinear Equations

Maitree Podisuk
Department of Mathematics and Computer Science Faculty of Science King Mongkut's Institute of Technology Chao-khun-tham Ladkrabang Ladkrabang Bangkok 10520 Thailand.

Received 21 Apr 1999
Accepted 16 May 2000

ABSTRACT
The modified Newton method for solving systems of nonlinear equations is one of the Newton-like methods. In this paper, results that will ensure the existence and uniqueness solutions to a system of nonlinear equations will be given. A second order convergence result is established.

KEYWORDS: modified Newton method, Newton-like methods, solving systems, nonlinear equations.

INTRODUCTION AND DEFINITIONS
The Newton method for solving a system of nonlinear equations

\[ f(x) = 0 \]

is the iteration equation

\[ x_{k+1} = x_k - (f'(x_k))^{-1}f(x_k), \quad k = 0, 1, \ldots. \]

where \( x \) is in \( \mathbb{R}^n \), \( f \) is a function from a subset of \( \mathbb{R}^n \) to a subset of \( \mathbb{R}^n \) and \( f'(x) \) is the Jacobian matrix of \( f(x) \).

The Newton-like method for solving (1) is the iteration equation

\[ x_{k+1} = x_k - (A(x_k))^{-1}f(x_k), \quad k = 0, 1, \ldots, \]

where \( A(x) \) is an invertible operator.

Podisuk \(^6\) introduced an iterative method which is one of Newton-like methods for solving (1) which has the form

\[ x_{k+1} = x_k - (H(x_k))^{-1}f(x_k), \quad k = 0, 1 \ldots. \]

where \( H \) is the Jacobian matrix of \( f(x) \).

N(x,t) = \{ y : y \in x, ||x-y|| < t \}

L(X,Y) = a real Banach space

\( f \in \text{Lip } K(D) \) if \( ||f(x)-f(y)|| \leq K ||x-y|| \) for some constant \( K \)

{\{x_k\} = \text{sequence of vectors in } \mathbb{R}^n}

\( B(x) = f'(x)A(x) \) and M(x) = f'(x)H(x).

Definition 1
Let \( t_0 \) and \( t \) be non-negative real numbers, \( g \) be a continuously differentiable real function on \([t_0, t_0 + t']\), and \( G \) a continuously differentiable operator on \( N(x_0, t') \subset X \) into \( X \). Then the equation \( t = g(t) \) will be said to majorize the equation \( x = G(x) \), or \( g \) majorizes \( G \), on \( N(x_0, t') \) if

\[ ||G(x_0) - x_0|| \leq g(t_0) - t_0 \quad \text{and} \quad ||G'(x)|| \leq g'(t) \quad \text{where} \quad ||x - x_0|| \leq t - t_0 < t \]

Definition 2
Let \( t_0 \) and \( t' \) be non-negative real numbers, \( g \) be a real function on \([t_0, t_0 + t']\), and \( G \) an operator sending \( N(x_0, t') \) into \( X \). Then the equation \( t = g(t) \) will be said to weakly majorize the equation \( x = G(x) \), or \( g \) weakly majorizes \( G \), on \( N(x_0, t') \) if

\[ ||G(x_0) - x_0|| \leq g(t_0) - t_0 \quad \text{and} \quad ||G'(x)|| \leq g(t) \quad \text{where} \quad ||x - x_0|| \leq t - t_0 < t' \]

LEMMA AND THEOREMS
We remark that Lemma 1 and Lemma 2 are known results in mathematical analysis and Lemma 3 is given by Ortega \(^4\).

Lemma 1. (The Banach Lemma)
Let \( j \in L(X,X) \) and \( \|j^{-1}\| \leq \delta < 1 \), then \( j^{-1} \) exists in...
Lemma 2.

Let \( f' \in \text{Lip}_K(D_0) \) and let \( x, y \in D_0 \), then

\[
|f(x) - f(y) - f'(x - y)| \leq \frac{1}{2} K ||x - y||^2.
\]

Lemma 3.

Let \( \{x_k\} \) be a sequence in \( X \) and \( \{t_k\} \) a sequence of non-negative real numbers such that

\[
||x_{k+1} - x_k|| \leq t_{k+1} - t_k, \quad k = 0, 1, ...
\]

and \( t_k \to t^* \). Under these conditions, there exists a point \( s \in X \) such that \( x_k \to s \)

\[
||s - x_k|| \leq t^* - t_k, \quad k = 0, 1, ...
\]

The following theorem is known as The Kantorovich Theorem. It is one of the fundamental theorems in numerical mathematics. The idea behind the proof of this theorem may be found in Kantorovich.3

Theorem 1. (The Kantorovich Theorem)

Let \( x_0 \) be in \( D_0 \) and let \( \Gamma_0 \in [f(x_0)] \) exist with \( \Gamma_0 f' \in \text{Lip}_K(D_0) \),

\[
||\Gamma_0 f(x_0)|| \leq \eta \quad \text{and} \quad h = K \eta \leq \frac{1}{2}.
\]

Define

\[
r_0(h) = \frac{1}{h} (1 - \sqrt{1 - 2h}) \eta, \quad r_1(h) = \frac{1}{h} (1 + \sqrt{1 - 2h}) \eta.
\]

Then if \( N(x_0, r_0(h)) \subset D_0 \), the sequence of iterates defined by Newton Method exists, remains in \( N(x_0, r_1(h)) \) and converges to \( s \) in \( N(x_0, r_1(h)) \) such that \( f(s) = 0 \).

Theorem 2.

If \( g \) majorizes \( G \) on \( N(x_0, t') \) and \( g \) has a fixed point in \( [t_0, t_0 + t'] \), then \( G \) has a fixed point \( s \) in \( N(x_0, t') \). Furthermore, \( x_{k+1} = G(x_k) \) and \( t_{k+1} = g(t_k) \), \( k = 0, 1, ... \), converges to \( s \) and \( t^* \) respectively with the real sequence majorizing the vector sequence.

Next lemma uses the weakly majorizing property and it is given by Dennis.1

Lemma 4.

If \( g(t) \in (t, t_0 + t') \) when \( t \in (t_0, t_0 + t') \), and \( g \) weakly majorizes \( G \) on \( N(x_0, t') \), then there are elements \( t^* \in [t_0, t_0 + t'] \), \( s \in N(x_0, t') \) such that

\[
x_{k+1} = G(x_k) \quad \text{and} \quad t^* = g(t_k), \quad k = 0, 1, ...
\]

converge to \( s \) and \( t^* \) respectively with the \( t \) sequence majorizing the \( x \) sequence.

Theorems 3-5 in this section are given by Dennis.1 These theorems give the convergence of the Newton-like method.

Theorem 3.

Let \( A \) be a function on \( N(x_0, r) \) such that \( A(x) \in L(X,Y) \) for each \( x \) and \( A(x) \) is invertible for each \( x \) in \( N(x_0, r) \) and that there is a real, nonvanishing, nonincreasing function \( a(t) \) on \( [0, r) \) such that

\[
||A^{-1}(x)|| \leq a(||x - x_0||)^{-1}.
\]

If \( f' \in \text{Lip}_K(N(x_0, r)) \) then if \( \sigma \geq 1 \) and \( \delta > 0 \) are real numbers such that

\[
a(t) + \sigma K t \]

is isotonic on \( (0, r) \), and

\[
||B(x)|| \leq a(||x - x_0||) + \sigma K ||x - x_0|| - \delta,
\]

for every \( x \in N(x_0, r) \)

then \( g(t) = t + (a(t))^{-1} (0.5 \sigma K t^2 - \delta t + a(0)) || (A(x_0))^{-1} f(x_0)|| \), weakly majorizes \( G(x) = x - (A(x))^{-1} f(x) \) on \( N(x_0, r) \).

Theorem 4.

Let \( f' \in \text{Lip}_K(N(x_0, r)) \) and \( (A(x_0))^{-1} \) exist and be bounded in the norm by \( (a(0))^{-1} \).
If $\|B(x_0)\| < a(0)$ and

\begin{align*}
\text{(10)} \quad h' &= K \| (A(x_0))^{-1} f(x_0) \| a(0)/(a(0) - \|B(x_0)\|)^2 \leq \frac{1}{2} \\
\text{(11)} \quad r_0' &= \frac{1}{K} (1 - \sqrt{1 - 2h'}) \|a(0) - \|B(x_0)\|| \leq r
\end{align*}

and

Then if $f$ has a unique zero $s \in \overline{N}(x_0, r_0')$, and

\[ x_{m+1} = x_m - (A(x_0))^{-1} f(x_m), \quad m = 0, 1, \ldots \]

converges to $s$ from any $x_0 \in \overline{N}(x_0, r)$ such that

\[ \|x_0 - x\| < r_0' = \frac{1}{K} (1 - \sqrt{1 - 2h'}) \|a(0) - \|B(x_0)\|| \].

If, in addition, $\sigma, \delta$ and $a$ satisfy the conditions of Theorem 3 and

\begin{align*}
\text{(12)} \quad h &= \frac{1}{\delta^2} \sigma K \| (A(x_0))^{-1} f(x_0) \| a(0) \leq \frac{1}{2} \\
\text{(13)} \quad r_0 &= \frac{1}{\sigma K} (1 - \sqrt{1 - 2h}) \delta < r \\
\text{(14)} \quad x_{m+1} &= x_m - (A(x_m))^{-1} f(x_m), \quad m = 0, 1, \ldots 
\end{align*}

converges to $s$.

Furthermore, $x_{m+1} = x_m - (A(x_m))^{-1} f(x_m)$ converges to $s$ from any $x_0 \in \overline{D} \cap N(x_0, r_1)$.

If, in addition, $\beta(\delta_0 + \eta_0) < 1$, $h = \frac{\beta K a}{1 - \beta \eta_0 - \beta \delta_0}$, and $N(x_0, r_0) \subset D$, then $x_{m+1} = x_m - (A(x_m))^{-1} f(x_m)$ converges to $s$.

**Main Results**

The following theorem of this section will ensure the convergence of the modified Newton method for solving a system of nonlinear equations which is a special kind of the Newton-like method.

**Theorem 6.**

Let $D$ be an open convex subset of the space $X$ and $f \in \text{Lip}_K(D)$. Assuming that $f(x)$ and $H(x)$ satisfy all the conditions of the previous theorems, then there exists a unique zero $s$ in $D$ so that for any point $x_0$ in $D$ the sequence $\{x_m\}$ where

\[ x_{m+1} = x_m - (H(x_m))^{-1} f(x_m) \]

converges to $s$.

**Proof**

By the results of the previous theorems, the existence and uniqueness of $s$ in $D$ is ensured and the sequence $\{x_m\}$ of points in $D$, where

\[ x_{m+1} = x_m - (H(x_m))^{-1} f(x_m) \]

for $x_0$ in $D$, converges to this unique points.

Theorem 7 is the last theorem of this section that
shows that the order of the convergence of the modified Newton method for solving a system of nonlinear equations which is of second order.

**Theorem 7.**

Let the conditions of Theorem 5 be satisfied and \( \delta_0 = 0 \), that is

\[
\|M(x)\| \leq \delta_1 \|x - x_0\|, \quad \forall x \in D.
\]

Then the order of the convergence of the method is equal to 2.

**Proof.**

Let \( Q = \sup (a(\|x-x_0\|))^{-1}, x \in N(x_0,r_0) \)

\[
P = Q \left( \frac{1}{2}K + \delta_1 \right)
\]

and \( e_k = \|s - x_k\| \) then

\[
e_{k+1} = \|s - x_{k+1}\|
= \|s - x_k + (H(x_k))^{-1}f(x_k)\|
= \|H(x_k))^{-1}f(s) + (H(x_k))^{-1}f(x_k) + (H(x_k))^{-1}H(x_k)(s - x_k)\|
\leq \|H(x_k))^{-1}||f(s) - f(x_k) + H(x_k)(s - x_k)\|
= \|H(x_k))^{-1}||f(s) - f(x_k) + f(x_k)(s - x_k) - f(x_k)(s - x_k) + H(x_k)(s - x_k)\|
\leq (\alpha(\|x_0 - x_k\|))^{-1} \left( \frac{1}{2}Ke_k^2 + \|f(x_k) - H(x_k)||e_k\|ight)
\leq (\alpha(\|x_0 - x_k\|))^{-1} \left( \frac{1}{2}Ke_k^2 + \|M(x_k)||e_k\|ight)
\leq (\alpha(\|x_0 - x_k\|))^{-1} \left( \frac{1}{2}Ke_k^2 + \delta_1 e_k^2\right)
\leq Q\left( \frac{1}{2}K + \delta_1 \right)e_k^2
\leq Pe_k^2.
\]

**Example**

We will use the above method to solve the following problem

\[
f(u, v) = \begin{bmatrix} u^2 + 2u + v^2 + v - 2 \\ u^2v + u^2v + 1 \end{bmatrix}
\]

which satisfies all the required conditions above.

Note that \( u = 1 \) and \( v = 1 \) is solution to this problem.

The iterative formulas are

\[
u_{m+1} = (2 - v_m)/(v^2 + 2)
u_{m+1} = (-u_{m-1} + (u^2 + 1).
\]

The iteration will be stopped when \( \|f(u_{m+1}, v_{m+1})\| < 0.0000001 \). In this example we have

\[
x = R^2, D = (0, 2)x(-2, 0), r_0 = 0.5, \delta_0 = 0, \delta_1 = 1, x_0 = (u_0, v_0),
\]

\[
f(u, v) = \begin{bmatrix} v^2 + 2uv + 1 \\ 2uv + u^2 + 1 \end{bmatrix}
\]

\[
f(u, v)^{-1} = \begin{bmatrix} 1 & -2uv + 1 \\ u^2 + 2uv - 4uv + 1 \end{bmatrix}
\]

\[
A(u, v) = \begin{bmatrix} v^2 & 0 \\ 0 & u^2 + 1 \end{bmatrix}, (A(u, v))^{-1} = \begin{bmatrix} 1 & 0 \\ -2uv + 1 & u^2 + 2 \end{bmatrix}
\]

\[
and B(u, v) = \begin{bmatrix} 0 & 2uv \\ 2uv & 0 \end{bmatrix}.
\]

**Table 1.** The following table 1 gives the results with different initial points.

<table>
<thead>
<tr>
<th>( u_0 )</th>
<th>( v_0 )</th>
<th>( r_0 )</th>
<th>( u_m )</th>
<th>( v_m )</th>
<th>number of iterations(m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.75</td>
<td>-0.75</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>19</td>
</tr>
<tr>
<td>1.25</td>
<td>-0.75</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>18</td>
</tr>
<tr>
<td>0.75</td>
<td>-1.25</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>17</td>
</tr>
<tr>
<td>1.25</td>
<td>-1.25</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>19</td>
</tr>
<tr>
<td>1.00</td>
<td>-0.75</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>18</td>
</tr>
<tr>
<td>1.00</td>
<td>-1.25</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>18</td>
</tr>
<tr>
<td>0.75</td>
<td>-1.00</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>18</td>
</tr>
<tr>
<td>1.25</td>
<td>-1.00</td>
<td>0.5</td>
<td>1</td>
<td>-1</td>
<td>18</td>
</tr>
</tbody>
</table>

The solution by the above method always converges to the point \((1, -1)\) for all the initial points in the domain D.
CONCLUSION

The order of convergence of Newton method is equal to 2. for the modified Newton method which is a Newton-like method together with the condition (19), the order of convergence is also 2. Without the condition (19) the order of convergence of the modified Newton method is linear.

REFERENCES