ScienceAsia 26 (2000) : **243-247**

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 Chaokhuntaharn 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

(1) $f(x) = 0$

is the iteration equation

(2) $x^{k+1} = x^k-(f'(x^k))^{-1}f(x^k), \quad k = 0, 1, \dots .$

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

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

(3) $x^{k+1} = x^k - (A(x^k))^{-1}f(x^k), \quad k = 0,1, \dots,$

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

(4) $x^{k+1} = x^k - (H(x^k))^{-1}f(x^k), \quad k = 0, 1 \dots .$

$$\text{where } H = \begin{bmatrix} f_{11} & 0 & . & . & . & 0 \\ 0 & f_{22} & . & & & . \\ . & & . & . & & . \\ . & & . & . & . & . \\ . & & & . & . & 0 \\ 0 & & . & . & 0 & f_{nn} \end{bmatrix}, \quad f_{jj} = \frac{\partial f_j}{\partial x_j}, \quad j = 1, 2, \dots, n.$$

Throughout this paper,
$X \equiv R^n$, n-dimensional real vector space
$D \subset X$, the domain of f(x)
$D_0 \equiv$ closed convex subset of D
$N(x,t) = \{y : y \in x, ||x-y||<t\}$
$L(X,Y) =$ a real Banach space
$f \in Lip_K(D)$ if $||f(x)-f(y)|| \le K\, x-y||$ for some constant K
$\{x_k\} \equiv$ sequence of vectors in 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 $\overline{N}(x_0, t') \subset X$ 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

(5) $||G(x_0) - x_0|| \le g(t_0) - t_0$ and
(6) $||G'(x)|| \le g'(t)$ where $||x - x_0|| \le t - t_0 < t'$.

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 $x = G(x)$, g weakly majorizes G, if (5) holds and in addition

(7) $||G(G(x)) - G(x)|| \le g(g(t)) - g(t)$,
when $||x -x_0|| \le t - t_0 < t'$
and $||G(x) - x|| \le g(t) - t$.

LEMMAS 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 $||I - J|| \le \delta < 1$, then J^{-1} exists in

$L(X,X)$ and

$$\|J^{-1}\| \le (1-\delta)^{-1}.$$

Lemma 2.

Let $f' \in Lip_K(D_0)$ and let $x, y \in D_0$, then

$$\|f(x) - f(y) - f'(x - y)\| \le \frac{1}{2} K \|x - y\|^2.$$

Lemma 3.

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

$$\|x_{k+1} - x_k\| \le t_{k+1} - t_k, \quad k = 0,1, \ldots$$

and $t_k \to t^* < \propto$ Under these conditions, there exists a point $s \in X$ such that $x_k \to s$ and

$$\|s - x_k\| \le t^* - t_k, \quad k = 0, \ldots .$$

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 $\Gamma 0 \in [f'(x_0)]^{-1}$ exist with $\Gamma_0 f' \in Lip_K(D_0)$,

$$\|\Gamma_0 f(x_0)\| \le \eta \text{ and } h = K\eta \le \frac{1}{2}.$$

Define $r_0(h) = \frac{1}{h}(1-\sqrt{1-2h})\eta$

$$r_1(h) = \frac{1}{h}(1+\sqrt{1-2h})\eta.$$

Then if $N(x_0,r_0(h)) \subset D_0$, the sequence of iterates defined by Newton Method exists, remains in $N(x_0,r_0(h))$ and converges to s in $N(x_0,r_0(h))$ such that $f(s) = 0$. If $h < 1/2$, s is the only root in $N(x_0,r_1(h)) \cap D_0$, and if $h = 1/2$, s is unique in $N(x_0,r_1(h)) \cap D_0$. Furthermore, the sequence of the iterates satisfies the error bounds

$$\|s-x_m\| \le \frac{1}{h}(1/2^m)(1-\sqrt{1-2h})2^m \eta.$$

The following theorem is given by Kantorovich and Akilov.[3] This theorem together with Lemma 3 and Definition 1 give the convergence of the sequence $\{x_k\}$ in X when the sequence of $\{t_k\}$ converges.

Theorem 2.

If g majorizes G on $\overline{N}(x_0, t')$ and g has a fixed point in $[t_0, t_0 + t']$, then G has a fixed point s in $\overline{N}(x_0, t')$. Furthermore, $x_{k+1} = G(x_k)$ and $t_{k+1} = g(t_k)$, $k = 0, 1, \ldots$, 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 $t^* \in [t_0, t_0 + t']$, $s \in \overline{N}(x_0, t')$ such that

$$x_{k+1} = G(x_k)$$
and
$$t_{k+1} = g(t_k),\ k = 0, 1, \ldots$$

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

$$\|A^{-1}(x)\| \le a(\|x - x_0\|)^{-1}.$$

If $f' \in (Lip_k(\overline{N}(x_0,r))$ then if $\sigma \ge 1$ and $\delta > 0$ are real numbers such that

(8) $\quad a(t) + \sigma K t$

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

(9) $\quad \|B(x)\| \le a(\|x - x_0\|) + \sigma K\|x - x_0\| - \delta,$

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 $f' \in LipK(\overline{N}(x_0,r))$ and $(A(x_0))^{-1}$ exist and be bounded in the norm by $(a(0))^{-1}$.

If $\|B(x_0)\| < a(0)$ and

(10) $\quad h' = K\|(A(x_0))^{-1}f(x_0)\|a(0)/(a(0) - \|B(x_0)\|)^2 \le \frac{1}{2}$

and

(11) $\quad r_0' = \frac{1}{K}(1-\sqrt{1-2h'})(a(0) - \|B(x_0)\|) \le r$

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

$\|x_0' - x_0\| < r_1' = \frac{1}{K}(1-\sqrt{1-2h})(a(0) - \|B(x_0)\|).$

If, in addition, σ, δ and a satisfy the conditions of Theorem 3 and

(12) $\quad h = \frac{1}{\delta^2}\sigma K\|(A(x_0)^{-1}f(x_0)\|a(0) \le \frac{1}{2}$ and

(13) $\quad r_0 = \frac{1}{\sigma K}(1-\sqrt{1-2h})\,\delta < r$ then

(14) $\quad x_{m+1} = x_m - (A(x_m))^{-1}f(x_m), \quad m = 0, 1, \ldots,$

converges to s.

In the following theorem, we impose one more condition on A(x) and one condition on B(x) instead of $B(x_0)$.

Theorem 5.

Let $f' \in Lip_k(\overline{D})$ where $x_0 \in D$ and D is an open convex subset of X. Assume that

(15) $\quad \|(A(x_0))^{-1}f(x_0)\| \le \alpha$

(16) $\quad \|(A(x_0))^{-1}\| \le \beta$

(17) $\quad \|(A(x)-A(x_0)\| \le \eta_0+\eta_1\|x - x_0\|, \ \forall\, x \in D$

(18) $\quad \|(B(x) \le \delta_0+\delta_1\|x - x_0\|, \ \forall\, x \in D$

Then $\beta\delta_0 < 1, h' = \frac{\beta K\alpha}{(1-\beta\delta_0)^2} \le \frac{1}{2}$ and $N(x_0, r_0') \subset D$ where

$$r_0' = \frac{1-\sqrt{1-2h'}}{\beta K}(1-\beta\delta_0)$$

imply that f has a root $r \in \overline{N}(x_0, r_0')$ which is unique in $D \cap N(x_0, r_1')$ where

$$r_1' = \frac{1-\sqrt{1-2h'}}{\beta K}(1-\beta\delta_0).$$

Furthermore $\quad x'_{m+1} = x'_m - (A(x_0))^{-1}f(x'_m)$

converges to s from any $x'_0 \in \overline{D} \cap N(x_0, r'_1)$.

If, in addition, $\beta(\delta_0 + \eta_0) < 1$,

$$h = \frac{\sigma\beta K\alpha}{(1-\beta\eta_0-\beta\delta_0)^2} \le \frac{1}{2}$$

where $\sigma = \max\left(1, \frac{\delta_1+\eta_1}{K}\right)$, and $N(x_0, r_0) \subset D$,

$$r_0 = \frac{1-\sqrt{1-2h}}{\sigma\beta K}(1-\beta\eta_0-\beta\delta_0)$$

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 $f' \in LipK(\overline{D})$. 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

(19) $\quad ||M(x)|| \le \delta_1||x-x_0||, \ \forall\, x \in D.$

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

Proof.

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

$$P = Q\,(\tfrac{1}{2}K + \delta_1)$$

h

and $\quad e_k = ||s-x_k||$ then

$$
\begin{aligned}
ek+1 \quad &= ||s - x_{k+1}|| \\
&= ||s - x_k + (H(x_k))^{-1}f(x_k)|| \\
&= ||(H(x_k))^{-1}f(s) + (H(x_k))^{-1}f(x_k) + s - x_k|| \\
&= ||(H(x_k))^{-1}f(s) + (H(x_k))^{-1}f(x_k) + \\
&\quad (H(x_k))^{-1}H(x_k)(s - x_k)|| \\
&\le ||(H(x_k))^{-1}||\cdot||f(s) - f(x_k) + H(x_k)(s - x_k)|| \\
&= ||(H(x_k))^{-1}||\cdot||f(s) - f(x_k) + f'(x_k)(s - x_k) - \\
&\quad f'(x_k)(s - x_k) + H(x_k)(s - x_k)|| \\
&= ||(H(x_k))^{-1}||\cdot||f(s) - f(x_k) + f'(x_k)(s - x_k) - \\
&\quad (f'(x_k) - H(x_k)(s - x_k)|| \\
&\le (a(||x_0-x_k||))^{-1}\ (\frac{1}{2}Ke_k^2 + ||f'(x_k) - H(x_k)||\,e_k) \\
&= (a(||x_0-x_k||))^{-1}\ (\frac{1}{2}Ke_k^2 + ||M(x_k)||\,e_k) \\
&\le (a(||x_0-x_k||))^{-1}\ (\frac{1}{2}Ke_k^2 + \delta_1 e_k^2) \\
&\le\ Q(\frac{1}{2}K + \delta_1)e_k^{\ 2} \\
&\le Pe_k^{\ 2}.
\end{aligned}
$$

EXAMPLE

We will use the above method to solve the following problem

$$f(u, v) = \begin{bmatrix} uv^2 + 2u + v - 2 \\ u^2v + u + v + 1 \end{bmatrix}.$$

which satisfies all the required conditions above.

Note that u =1 and v =-1a is solution to this problem. The iterative formulas are

$$u_{m+1} = (2-v_m)/(v_m^2+2)$$

$$v_{m+1} = (-u_m-1)/(u_m^2+1).$$

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

$x = R^2$, $D = (0,2)x(-2,0)$, $D_0 = [0, 2]x[-2, 0]$, $r_0 = 0.5$, $\delta_0 = 0$, $\delta_1 = 1$, $x_0 = (u_0, v_0)$,

$$f'(u, v) = \begin{bmatrix} v^2 + 2 & 2uv + 1 \\ 2uv + 1 & u^2 + 1 \end{bmatrix},$$

$$f'(u, v))^{-1} = \frac{1}{u^2 + v^2 - 3u^2v^2 - 4uv + 1}\begin{bmatrix} u^2 + 1 & -2uv - 1 \\ -2uv - 1 & v^2 + 2 \end{bmatrix},$$

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

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

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

u_0	v_0	r_0	u_m	v_m	number of iterations(m)
0.75	-0.75	0.5	1	-1	19
1.25	-0.75	0.5	1	-1	18
0.75	-1.25	0.5	1	-1	18
1.25	-1.25	0.5	1	-1	19
1.00	-0.75	0.5	1	-1	18
1.00	-1.25	0.5	1	-1	18
0.75	-1.00	0.5	1	-1	18
1.25	-1.00	0.5	1	-1	18

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

1. Dennis JE, Jr (1970) 'On the Convergence of Newton-like Methods', appeared in *Numerical Methods for Nonlinear Algebraic Equations*, Edited by Philip Rabinowitz, Gordon and Breach Science Publishers, London, 163-81.
2. Jankowska J (1975) Multivariate Secant Method, Ph.D. Thesis, University of Warsaw, Poland.
3. Kantorovich LV and Akilov GP (1964) Functional Analysis in Normed Spaces, MacMillan, New York, N.Y.
4. Ortega JM (1968) The Newton-Kantorowich Theorem, *Amer Math Monthly* 75, 658-60.
5. Ostrowski AM (1966) Solution of Equations and Systems of Equations, second edition, Academic Press.
6. Podisuk M (1991) Modified Newton Method for Solving Systems of Nonlinear Equations, *Southeast Asia Bulletin of Mathematics* 15(2), 123-130.
7. Traub JF (1964) Iterative Methods for the Solution of Equations, Prentice-Hall, Inc.