Comm. Korean Math. Soc. 9 (1994), No. 2, pp. 491–502

A QUASI-NEWTON METHOD USING DIRECTIONAL DERIVATIVES FOR NONLINEAR EQUATIONS

SUNYOUNG KIM

1. Introduction

Many problems arising in science and engineering require the numerical solution of a system of n nonlinear equations in n unknowns:

$$\text{(1)} \qquad \text{given } F : R^n \to R^n, \text{ find } x_* \in R^n \text{ such that } F(x_*) = 0.$$

Nonlinear problems are generally solved by iteration. Davidson [3] and Broyden [1] introduced the methods which had led to a large amount of research and a class of algorithm. This work has been called by the quasi-Newton methods, secant updates, or modification methods. Newton's method is the classical method for the problem (1) and quasi-Newton methods have been proposed to circumvent computational disadvantages of Newton's method.

A number of methods have been developed from Newton's method. The advantages of Newton's method are: the existence of a domain of attraction for a root insures stability for the iteration and its quadratic convergence to a root. On the other hand, in many problems, it requires a very good initial approximation to x_* in order to converge. Though its convergence rate is quadratic, it is not globally convergent. The more important disadvantage is that it requires the computation of the Jacobian matrix at every iteration, which involves n^2 function evaluations. This is very expensive operation for most functions. If the Jacobian is not analytically available, then approximating the Jacobian by a less expensive method is a very important issue.

Quasi-Newton methods use the concept of Newton's method but approximate the Jacobian matrix. Many of the successful quasi-Newton

Received November 30, 1993.

methods use updates which satisfy a linear constraint. The basic idea is to approximate the Jacobian matrix by a matrix B_{k+1} satisfying,

$$y_k = B_{k+1}(x_{k+1} - x_k), \tag{2}$$

where $y_k = F(x_{k+1}) - F(x_k)$. Equation (2) is called as the quasi-Newton equation and the methods that satisfy (2) are quasi-Newton methods. Of all the quasi-Newton methods that have been introduced, Broyden's method has been known as the most successful. It is very efficient for problems of small to medium size, since it does not require the evaluation of the Jacobian and has a relatively fast, q-superlinear rate of convergence ([4], [5]). However, it does not enjoy the self-correctiveness of Newton's method [5] and fails to converge much more often. Its convergence is unstable when the equations are highly nonlinear or sparse.

In this paper, we propose a method to improve the weakness of Broyden's method by considering two matrices for the update for the Jacobian and solving them together using generalized inverse.

Broyden's method will be briefly derived, in the next section. The motivation, derivation, algorithm, and convergence rate of the new method will be discussed in section 3. Numerical experiment will be followed.

We use R^n to denote n-dimensional real Euclidean space with the inner product, $\langle x, y\rangle = x^T y$, $\|\cdot\|$ stands for either l_2 vector norm $\|x\| = \langle x, x\rangle^{\frac{1}{2}}$, or for any matrix norm which is consistent with the l_2 vector norm in the sense that $\|Ax\| \le \|A\|\|x\|$ for $x \in R$ and any matrix, A of order n. In particular, the l_2 norm and the Frobenius norm are consistent with the l_2 vector norm, and the Frobenius norm is computed by

$$\|A\|_F^2 = tr(AA^T),$$

the weighted Frobenius norm is

$$\|A\|_{M,F} = \|AMA^T\|.$$

2. Broyden's method

Assume that $F : R^n \to R^n$ is continuously differentiable in an open convex set D and for given x in D and $s \ne 0$, the vector $\overline{x} = x + s$ belongs to D.

For given $\epsilon > 0$, there is a $\delta > 0$ such that

$$\|F(x) - F(\overline{x}) - F'(\overline{x})(x - \overline{x})\| \leq \epsilon\|x - \overline{x}\|$$

if $\|\overline{x} - x\| < \delta$, since F' is continuous at $\overline{x}$. Hence,

$$F(x) \approx F(\overline{x}) + F'(\overline{x})(x - \overline{x}). \tag{3}$$

If $\overline{B}$ denotes an approximation to $F'(\overline{x})$, and let $s = \overline{x} - x$ and $y = F(\overline{x}) - F(x)$, then, (3) becomes

$$y = \overline{B}s. \tag{4}$$

Now, suppose that we had an approximation B to $F'(x)$. Broyden reasoned that there really is no justification for having $\overline{B}$ differ from B on the orthogonal complement of s. This can be expressed as the requirement

$$\overline{B}z = Bz \ \text{ if } \langle z, s\rangle = 0. \tag{5}$$

Clearly (4) and (5) uniquely determine $\overline{B}$ from B and in fact [7],

$$\overline{B} = B + \frac{(y - Bs)s^T}{s^T s}. \tag{6}$$

From (6), Broyden's method can be used in an iterative method as follows.

Algorithm 2.1

Given $F : R^n \to R^n, x_0 \in R^n, B_0 \in R^{n\times n}$.
Do for $k = 0, 1, \ldots$:
Solve $B_k s_k = -F(x_k)$ for s_k,
$x_{k+1} := x_k + s_k$,
$y_k := F(x_{k+1}) - F(x_k)$,
$B_{k+1} := B_k + \dfrac{(y_k - B_k s_k)s_k^T}{s_k^T s_k}$.

It is known that Broyden's method has q-superlinear convergence rate and can compete with other algorithms when the Jacobian is difficult to evaluate. In Broyden's method, B_k depends on each B_j with $j < k$ and thus it may retain information which is irrelevant or even harmful; it is not self-correcting.

To improve the disadvantages of Broyden's method that we have discussed, we mainly investigate the variational problem:

$$\min\{\|J - B\|_{F,M} : J \in R^{n\times n}, Bs = y\}, \tag{7}$$

where J is the Jacobian matrix and B is an approximation to J and the norm is the weighted Frobenius norm with any weighting matrix M. This problem comes from the derivation of minimum change updates to the matrix [6], B, of approximate Jacobian matrices in quasi-Newton methods for solving nonlinear systems of equations. The problem, (7) is related to the bound deterioration Theorem [7] given as follows.

THEOREM 2.1. *Let $D \subseteq R^n$ be an open convex set containing $x, \overline{x}$, with $x \neq x_*$. Let $f : R^n \to R^n$, $J(x) \in Lip_\gamma(D)$, $B \in R^{n\times n}$, $\overline{B}$ defined by (6) for either the Frobenius or l_2 matrix norms,*

$$\|\overline{B} - J(\overline{x})\| \leq \|B - J(x)\| + \frac{3\gamma}{2}\|\overline{x} - x\|_2. \tag{8}$$

Furthermore, if $x_* \in D$ and $J(x)$ obeys the weaker Lipschitz condition

$$\|J(x) - J(x_*)\| \leq \gamma\|x - x_*\|, \text{ for all } x \in D,$$

then,

$$\|\overline{B} - J(x_*)\| \leq \|B - J(x_*)\| + \frac{\gamma}{2}(\|\overline{x} - x_*\|_2 + \|x - x_*\|_2). \tag{9}$$

Proof. [7].

If we get a tighter bound for $\|\overline{B} - J\|$ by (7), a faster convergence rate is obtained. In the next section, we propose a new quasi-Newton update and compare it with Broyden's update.

3. Proposed Method: An Update Using Directional Derivatives

There have been many methods for approximating the Jacobian matrix. A method using the directional derivatives is presented in order to handle problem (7), viz., approximate the Jacobian matrix more accurately. The directional derivatives in the steepest descent direction are utilized in the quasi-Newton equation when approximating the Jacobian since the steepest descent direction, by definition, is most rapidly decreasing direction.

If the analytical Jacobian is not available or expensive to compute, then initial Jacobians are evaluated numerically by forward difference. We will need the following lemma and the definition of the forward difference scheme given as follows.

LEMMA 3.1. *Let $F : R^n \to R^m$ be continuously differentiable in the open convex set $D \subset R^n$, $x \in D$, and let J be Lipschitz continuous at x in the neighborhood D, using a vector norm and the induced matrix operator norm and the constant γ. Then, for any $x + p \in D$,*

$$\|F(x+p) - F(x) - J(x)p\| \le \frac{\gamma}{2}\|p\|^2.$$

For any $u,\ v \in D$,

$$\|F(x+p) - F(x) - J(x)(v-u)\| \le \gamma\frac{\|v-x\| + \|u-x\|}{2}\|v-u\|. \tag{10}$$

Proof. [5].

DEFINITION 3.1. The element of an approximate Jacobian, b_{ij}, using the forward difference scheme is defined as

$$b_{ij} = \frac{f_i(x + he_j) - f_i(x)}{h},$$

where e_j denotes the jth unit vector. The approximation to j-th column of $J(x)$ is defined by

$$B_{\cdot j} = \frac{F(x + he_j) - F(x)}{h}.$$

Consider the steepest descent direction, ds, defined as,

$$ds = -B^T F(x),$$

where $F \in R^n$. Since $F(x)$ decreases most rapidly from x in the direction of the steepest descent direction, when approximating $F'(\overline{x})$, a Jacobian approximation at $\overline{x}$, the directional derivative, $F(\overline{x})'ds$ is considered and $\overline{x}$ is moved in the direction of ds.

We derive the method using directional derivatives as follows:

Let

$$ds = -B^T F(x).$$

If directional derivatives are used in approximating the Jacobian,

$$F'(\overline{x})ds \approx \frac{F(\overline{x} + hds) - F(\overline{x})}{h} \overset{\text{def}}{=} w. \tag{11}$$

We now use w as an approximation of $\overline{B}ds$. Therefore, the new ΔB, the update to the current Jacobian approximation B, can be obtained. From the quasi-Newton equation,

$$\begin{aligned} \overline{B}ds &= (B + \Delta B)ds \\ &= Bds + \Delta Bds. \end{aligned}$$

Using (11),

$$\Delta Bds = w - Bds \equiv u.$$

We propose a method which combines the above update and Broyden's update. Hence, we have two updates, two systems of equations to solve,

$$\Delta Bds = u$$

$$\Delta Bs = -F(\overline{x}).$$

Let $\overline{F}$ denote $F(\overline{x})$. Solving the above systems of equations for ΔB yields a rank-2 update. We note that (s, ds), $(F, u) \in R^{n\times 2}$. The update is:

$$\begin{aligned} \Delta B(s, ds) =& (\overline{F}, u), \\ \Delta B =& (\overline{F}, u)(s, ds)^+ \\ =& \frac{1}{s^T s ds^T s - (s^T s)^2}[ds^T s\overline{F} - s^T su]s^T \\ &+ [s^T s\overline{F} + s^T su]ds^T, \end{aligned} \tag{12}$$

where

$$(s,ds)^{+}=\left[\begin{pmatrix} s^T \\ ds^T \end{pmatrix}(s,ds)\right]^{-1}\begin{pmatrix} s^T \\ ds^T \end{pmatrix}.$$

(12) is used in iterations as follows:

Algorithm 3.2

$$\begin{aligned}
&\text{Given } F: R^n \to R^n, x_0 \in R^n, B_0 \in R^{n\times n}\\
&\text{Do for} k = 0,1,\ldots:\\
&\text{Solve } B_k s_k = -F(x_k) \text{ for } s_k\\
&\quad x_{k+1} := x_k + s_k\\
&\quad y_k := F(x_{k+1}) - F(x_k)\\
&\quad ds := -B^T F(x_k)\\
&\quad w := \frac{F(x_{k+1} + hods) - F(x_{k+1})}{h}\\
&\quad u := w - Bds
\end{aligned}$$

$$B_{k+1} := B_k + \frac{1}{s^T s ds^T s - (s^T s)^2}[ds^T s F_{k+1} - s^T s u]s^T$$

$$+ [F_{k+1} s^T s + u s^T s] ds^T$$

The above algorithm shows better performance than Broyden's method in numerical tests and we would like to show its convergent property theoretically.

Convergence Properties

The following theorem deals with the convergence rate of the above algorithm.

THEOREM 3.1. *Let $D \subseteq R^n$ be an open convex set containing $x, \overline{x}$, with $x \neq x_*$. Let $F: R^n \to R^n, J(x) \in Lip_\gamma(D), B \in R^{n\times n}, \overline{B}$ defined by (12). If $x_* \in D$ and $J(x)$ obeys the weaker Lipschitz condition, then for either the Frobenius or l_2 matrix norms,*

$$\|\overline{B} - J(x_*)\| \le \|B - J(x_*)\| + \frac{\mu\gamma}{2}(\|\overline{x} - x\|_2 + \|ds\|_2), \tag{13}$$

where $\mu = \frac{1}{(1-\frac{\mu_1}{\mu_2})^{\frac{1}{2}}}$, $\mu_1 = s^T s ds^T ds$, *and* $\mu_2 = (s^T ds)^2$.

Proof. The update is

$$\Delta B = (\overline{f}, u)(s, ds)^+$$

Let $(s, ds)^+ = \begin{pmatrix} a^T \\ b^T \end{pmatrix}$.

Then,

$$\|a^T\| = \frac{1}{(1-\frac{\mu_1}{\mu_2})^{\frac{1}{2}}} \frac{1}{\|s\|},$$

$$\|b^T\| = \frac{1}{(1-\frac{\mu_1}{\mu_2})^{\frac{1}{2}}} \frac{1}{\|ds\|}.$$

Let $J_* \equiv J(x_*)$. Subtracting J_* from both sides of $\overline{B} = B + \Delta B$,

$$\begin{aligned}
\overline{B} - J_* &= B - J_* + (\overline{F}, u) \begin{pmatrix} a^T \\ b^T \end{pmatrix} \\
&= B - J_* + (y - Bs)a^T + (w - Bds)b^T \\
&= B - J_* + (J_*s - Bs)a^T + (y - J_*s)a^T \\
&\quad + (J_*ds - Bds)b^T + (w - J_*ds)b^T \\
&= (B - J_*)[I - sa^T - dsb^T] + (y - J_*s)a^T + (w - J_*ds)b^T
\end{aligned}$$

Now,

$$\begin{aligned}
[I - sa^T - dsb^T] &= \left(I - (s, ds) \begin{pmatrix} a^T \\ b^T \end{pmatrix} \right) \\
&= [I - (s, ds)(s, ds)^+].
\end{aligned}$$

Since

$$[I - (s, ds)(s, ds)^+][I - (s, ds)(s, ds)^+] = [I - (s, ds)(s, ds)^+],$$

$[I - (s, ds)(s, ds)^+]$ is a projection and

$$\|[I - (s, ds)(s, ds)^+]\|_2 = 1. \tag{14}$$

Using Lemma 2.1,

$$
\begin{aligned}
&\|(y - J_*s)a^T + (w - J_*ds)b^T\| \\
&\qquad \le \|(y - J_*s)a^T\| + \|(w - J_*ds)b^T\| \\
&\qquad \le \frac{\gamma}{2}(\|s\|\|a^T\| + \|ds\|\|b^T\|) \\
&\qquad \le \frac{\gamma}{2}\mu(\|s\|_2 + \|ds\|_2) \\
&\qquad = \frac{\mu\gamma}{2}(\|\overline{x} - x\|_2 + \|ds\|_2).
\end{aligned} \tag{15}
$$

In view of (14) and (15), we have (13).

It is proved in [2] that $\lim_{k\to\infty} s = 0$ and so is $\lim_{k\to\infty} ds = 0$, therefore, Theorem 3.1 shows that the update produces a deteriorating bound for $\|\overline{B} - J_*\|$ at each iteration which implies q-linear convergence rate.

According to Dennis and Moré ([4], [8]) to get a q-superlinear convergence, we need only to prove that

$$
\lim_{k\to\infty} \frac{\|(B_k - F'(x^*))(x_{k+1} - x_k)\|}{\|x_{k+1} - x_k\|} = 0. \tag{16}
$$

The next theorem deals with q-superlinear convergence rate of the new method and uses the same assumption as the convergence theorem for Broyden's method.

THEOREM 3.2. *If there exists positive constants ϵ, δ such that if $\|x_0 - x_*\| \le \epsilon$ and $\|B_0 - B_*\| \le \delta$, then the sequence x_* generated by Algorithm 3.2 is well defined and converges q-superlinearly to x_*.*

Proof. Let $G = (s, ds)$ and $E = B - J_*$. Then, direct computation shows that

$$
\|EGG^+\|_F^2 = 2\|E\|_F^2.
$$

In view of

$$
\|E\|_F^2 = \|EGG^+\|_F^2 + \|E(I - GG^+)\|_F^2,
$$

and Theorem 3.2,

$$
\begin{aligned}
\|E_{k+1}\|_F &\le \|E_k(I - GG^+)\|_F + \frac{\gamma}{2}\mu(\|s\|_2 + \|ds\|_2) \\
&\le -\|E_k\|_F + +\frac{\gamma}{2}\mu(\|s\|_2 + \|ds\|_2) \qquad (17)\\
&\le \frac{\gamma}{2}\mu(\|s\|_2 + \|ds\|_2).
\end{aligned}
$$

This implies (16).

The bound $\|E_{k+1}\|$ in Broyden's method is

$$\|E_{k+1}\|_F \le \|E_k\|_F + \frac{\gamma}{2}(\|s\|_2). \qquad (18)$$

From (17) and (18), the proposed method has a smaller bound for $\|E_{k+1}\|$ than that of Broyden's since $\|s\|$ and $\|ds\|$ is negligible, which indicates that the proposed method has faster or at least same convergence rate as Broyden's. In next section, we will present numerical evidence showing efficiency of the proposed algorithm.

4. Numerical experiments and discussion

Numerical experiment with the proposed method and Broyden's methods are encouraging. The problems for unconstrained minimization, i.e., solving (1), from Moré et al ([10], [11]) were used to compare the number of iterations for each method using the same test functions and starting values. The problems number in [10] are used for easy reference: 21 for extended Resenbrock function, 22 for extended Powell's singular function, 26 for trigonometric function, 27 for Brown's almost linear function, 28 for discrete boundary value function, 29 for discrete integral equation function, 30 for Broyden's tridiagonal function, and 31 for Broyden's banded function. Broyden's banded function has been modified as follows: the original constant coefficients 2, 5 and 1 are changed to three parameters $w1$, $w2$ and $w3$, respectively. Varying the three parameters provides the problem with progressively worse scaling.

Results of computational experiment are given in Table 1. We used $n = 40$ and ftol $= 10^{-6}$. Initial Jacobians were evaluated numerically by finite differences. Two initial Newton iterations were used. It is evident from Table 1 that the proposed method converges faster than Broyden's Method in the number of iterations and it is also tested that the performance of the new method improves with increase in the size of the problems. For the problems that converge within a few of iterations, it takes the same number of iterations for the proposed method to converge because as shown Theorem 3.5, the smaller bound for E_k in the proposed method gives an advantage as the iteration proceeds.

Prob.	Broyden's Method	New Method
21	11	8
22	18	18
26	17	12
27	5	5
28	2	2
29	3	3
30	6	6
31(6)	11	9
31(12)	13	9
31(25)	16	12
31(50)	23	14
31(100)	Non Convergence	18
31(200)	60	22
31(400)	Non Convergence	29

Table 1. Broyden & new methods in no. of iterations

In Table 1, notice that Broyden's method failed for two problems while the new method succeeded in getting a solution. This is an example where Broyden's method lacks stability. The tests indicate that the new update shows less divergence.

The overhead of the new method is that it requires the calculation of the steepest descent direction and w at each iteration, this entails, respectively, additional $O(n^2)$ and n function evaluations. However, as indicated numerically and theoretically in the last section, the new update behaves much better for the problems which take many iterations to converge to a root and the impact of additional computation can be ignored because the new method takes fewer iterations to converge.

Concluding Remarks

We have presented a quasi-Newton type method employing directional derivatives for solving nonlinear equations and shown that the proposed method is locally and q-superlinearly convergent. It is also proved that the Jacobian approximation of the proposed method is better than that of Broyden's. Computational results show significant efficiency of the method.

References

1. C. G. Broyden, *A class of methods for solving nonlinear simultaneous equations*, Math. Comp. **19** (1969), 577–593.
2. C. G. Broyden, J. E. Dennis Jr. and J. J. Moré, *On the local and superlinear convergence of quasi-Newton methods*, J. Inst. Math. Appl. **12** (1973), 223–246.
3. W. C. Davidon, *Variable metric method for minimization*, Rep. ANL-5990 Rev. Argonne National Laboratories, Argonne, Illinois, 1959.
4. J. E. Dennis and J. J. Moré, *A characterization of superlinear convergence and its application to quasi-Newton methods*, Math. Comp. **28** (1974), 549–560.
5. ______, *Quasi-Newton methods, motivation and theory*, SIAM Rev. **19** (1977), 46–89.
6. J. E. Dennis and R. B. Schnabel, *Least change updates for quasi-Newton methods*, SIAM Rev. **21** (1979), 443–459.
7. ______, *Numerical methods for unconstrained optimization and nonlinear equations*, Prentice-Hall Inc., Englewood Cliffs, New Jersey, 1983.
8. Guangye Li, *Successive column correction algorithms for solving sparse nonlinear systems of equations*, Math. Programming **43** (1989), 187–207.
9. George W. Johnson and Nieve H. Austria, *A quasi-Newton method employing direct secant updates of matrix factorizations*, SIAM J. Numer. Anal **20** (1983), 315–415.
10. J. J. Moré, B. S. Garbow, and K. E. Hillstrom, *Fortran subroutines for testing unconstrained optimization software*, TOMS **7** (1981), 136–140.
11. ______, *Testing unconstrained optimization software*, ACM Trans. Math. Software **7** (1981), 17–41.
12. R. P. Tewarson, *A unified derivation of quasi-Newton methods for solving non-sparse and sparse nonlinear equations*, Computing **21** (1979), 113–125.

Department of Mathematics
Ewha Womans University
Seoul 120-750, Korea