Vol. 2 No. 1 JOURNAL OF COMPUTATIONAL MATHEMATICS January 1984

ON ADVANCE OF SECOND ORDER QUASI-NEWTON METHODS*

PAN PING-QI (潘平奇)

(Nanjing Technological College of Forest Products, Nanjing, China)

1. Introduction

The key to the quasi-Newton methods for solving nonlinear equations in several variables or optimization problems lies in how the approximation to the Jacobian or Hessian matrix (or its inverse) is determined.

We denote by R^n and $R^{n\times n}$ the real n-dimensional linear space of all column vectors and that of all real square matrices of order n, respectively. The mapping with domain D in R^n and range in R^n is denoted by $F: D\subset R^n\to R^n$, which is the left of the system of equations or the gradient of the objective function.

Let $\hat{x}$ and $\tilde{x}$ be two distinct points in D and denote

$$\hat{F}=F(\hat{x}),\quad \tilde{F}=F(\tilde{x}),\quad B=F'(\hat{x}), \tag{1.1}$$

where $F'(x)$ is the Jacobian matrix of F at x, and

$$\delta=\tilde{x}-\hat{x},\quad \gamma=\tilde{F}-\hat{F}. \tag{1.2}$$

It is our purpose to approximate the Jacobian matrix of F at $\tilde{x}$.

Evidently, the linear mapping

$$L(x)=\tilde{F}+\tilde{B}(x-\tilde{x}), \tag{1.3}$$

where $\tilde{B}\in R^{n\times n}$ and

$$\tilde{B}\delta=\gamma \tag{1.4}$$

satisfies

$$L(\hat{x})=\hat{F},\ L(\tilde{x})=\tilde{F}. \tag{1.5}$$

So, if $\|\delta\|$ is small enough, $L(x)$ can be regarded as a reasonable approximation to $F(x)$ at $\tilde{x}$ and $\tilde{B}$ as an approximate Jacobian matrix of F at the same point. As a matter of fact, we have

Theorem 1. *Suppose that $F:D\subset R^n\to R^n$ is Fréchet-differentiable at $\tilde{x}\in \text{int}\,(D)$ and that $L(x)$ is determined by (1.3) and (1.4). Then, for any $t\in R^1$, which is non-zero and sufficiently small,*

$$L(\tilde{x}+t\delta)-F(\tilde{x}+t\delta)=tE(t,\delta), \tag{1.6}$$

where $\lim\limits_{\delta\to 0} E(t,\delta)/\|\delta\|=0$, $\lim\limits_{t\to 0} E(t,\delta)=\bar{E}(\delta)$ *and* $\lim\limits_{\delta\to 0}\bar{E}(\delta)/\|\delta\|=0.$

Proof.

$$F(\tilde{x}+t\delta)=\tilde{F}+F'(\tilde{x})(t\delta)-E_1(t\delta), \tag{1.7}$$

where

$$\lim_{t\delta\to 0} E_1(t\delta)/\|t\delta\|=0.$$

$$\hat{F}=\tilde{F}-F'(\tilde{x})\delta-\bar{E}(\delta), \tag{1.8}$$

* Received January 5, 1983.

where $$\lim_{\delta\to 0} \bar{E}(\delta)/\|\delta\| = 0.$$

By (1.3), (1.7), (1.4) and (1.8), we have

$$L(\tilde{x}+t\delta)-F(\tilde{x}+t\delta)=t\tilde{B}\delta-tF'(\tilde{x})\delta+E_1(t\delta)$$
$$=t(\tilde{F}-\hat{F}-F'(\tilde{x})\delta)+E_1(t\delta)=t[\bar{E}(\delta)+E_1(t\delta)/t]. \tag{1.9}$$

If we let

$$E(t, \delta)=\bar{E}(\delta)+E_1(t\delta)/t, \tag{1.10}$$

then

$$\lim_{\delta\to 0} E(t, \delta)/\|\delta\|=\lim_{\delta\to 0} \bar{E}(\delta)/\|\delta\|+\lim_{\delta\to 0} E_1(t\delta)/(t\|\delta\|)=0$$

and

$$\lim_{t\to 0} E(t, \delta)=\bar{E}(\delta)+\lim_{t\to 0} E_1(t\delta)/t=\bar{E}(\delta).$$

The theorem is then proved.

In view of the above, we regard (1.3) as the first order approximation to F at $\tilde{x}$. Let $\tilde{B}$ be nonsingular and $\tilde{H}$ be its inverse. Then (1.4) is equivalent to

$$\delta=\tilde{H}\gamma. \tag{1.11}$$

Equation (1.4) or (1.11) has been central to the development of quasi-Newton methods, and therefore it is often termed the quasi-Newton equation. But in view of Theorem 1, we call it first order quasi-Newton equation, and all of the known updates, which can be derived from it and used in quasi-Newton methods, are called, correspondingly, the first order.

Since there is no B or H in (1.4) or (1.11), $\tilde{B}$ or $\tilde{H}$ relates only to γ and δ rather than to B or H. Thus $\tilde{B}=\gamma q^T/q^T\delta$, where $q\in R^n$ and $q^T\delta\neq 0$, satisfies (1.4) and is independent of B. In fact, B or H in the first order updates can be, theoretically, replaced by any other positive definite matrices (for optimization problems).

In this paper, many new updates are derived by means of new approximations to F and shown to be of second order in a certain sense. They have not only higher precision than the old ones but also the same simplicity; moreover, B or H there cannot be replaced by other matrices, and therefore the approximate Jacobian (or Hessian) matrix or its inverse at the initial point must be used as the initial matrix, or theoretically so to say the least. In addition, it is much more significant that the so-called second order quasi-Newton equation derived later must become the new starting point of quasi-Newton methods.

2. One-Reduction Matrices of δ

Let us introduce the new concept needed in the ensuing sections:

Definition. *Let $\delta\in R^n$, $A\in R^{n\times n}$ is a one-reduction matrix of δ if*

$$\delta^T A\delta=1. \tag{2.1}$$

Two examples

(1) $$U(P, S)=\frac{1+\delta^T S\delta}{\delta^T P\delta}P-S, \quad \forall P, S\in R^{n\times n} \text{ and } \delta^T P\delta\neq 0. \tag{2.2}$$

(2) $$V(P, q)=\frac{1-\delta^T q}{\delta^T P\delta}P+\frac{qq^T}{\delta^T q}, \quad \forall P\in R^{n\times n} \text{ and } \delta^T P\delta\neq 0,$$
$$q\in R^n \text{ and } \delta^T q\neq 0. \tag{2.3}$$

The one-reduction matrix has the following simple properties:

(i) There is not any one-reduction matrix for a zero vector.

(ii) All one-reduction matrices are nonnegative definite.

(iii) If $\alpha_i \in R^1$ and A_i is a one-reduction matrix of δ, $i=1, 2, \cdots, m$, then $\sum_i \alpha_i A_i / \sum_i \alpha_i$ is also a one-reduction matrix of δ when $\sum_i \alpha_i \neq 0$. Otherwise $\delta^T(\sum_i \alpha_i A_i)\delta = 0$.

(iv) If A is a one-reduction matrix of δ, so is A^T.

(v) If a real symmetric matrix A is a one-reduction matrix of δ and its greatest and smallest eigenvalues are λ_1 and λ_n respectively, then

$$\lambda_n \leqslant 1/\delta^T\delta \leqslant \lambda_1.$$

3. Rank-One Updates

For the quadratic mapping

$$Q(x) = \hat{F} + B(x-\hat{x}) + (x-\hat{x})^T A(x-\hat{x})(\gamma - B\delta), \tag{3.1}$$

where A is an arbitrary one-reduction matrix of δ, we have, clearly,

$$Q(\hat{x}) = \hat{F}, \quad Q(\tilde{x}) = \tilde{F}, \quad Q'(\hat{x}) = B. \tag{3.2}$$

Comparing (3.2) with (1.5), we can reason that if $\|\delta\|$ is small enough, $Q(x)$ is a better approximation to F at $\tilde{x}$ than (1.3), and so the Jacobian matrix of $Q(x)$ at $\tilde{x}$ is a better approximation to that of F at $\tilde{x}$. (3.1) may be reduced to

$$Q(x) = \tilde{F} + [B + (\gamma - B\delta)\delta^T(A+A^T)](x-\tilde{x}) + (x-\tilde{x})^T A(x-\tilde{x})(\gamma - B\delta), \tag{3.3}$$

from which we obtain at once the general rank-one update

$$\tilde{B} = B + (\gamma - B\delta)\delta^T(A+A^T), \tag{3.4}$$

where A is an arbitrary one-reduction matrix of δ.

By the Sherman-Morrison formula the corresponding update for the inverse Jacobian matrix can also be obtained:

$$\tilde{H} = H + \frac{(\delta - H\gamma)\delta^T(A+A^T)H}{\delta^T(A+A^T)H\gamma - 1}. \tag{3.5}$$

Using various one-reduction matrices of δ as used in (3.4) or (3.5), we will obtain correspondingly various new updates. For example, let $A = U(P, S)$; (3.4) and (3.5) are, respectively, reduced to

$$\tilde{B} = B + (\gamma - B\delta)\delta^T\left[\frac{1+\delta^T S\delta}{\delta^T P\delta}(P+P^T) - (S+S^T)\right] \tag{3.6}$$

and

$$\tilde{H} = H + \frac{(\delta - H\gamma)\delta^T G H}{\delta^T G H\gamma - 1}, \tag{3.7}$$

where

$$G = (1+\delta^T S\delta)(P+P^T) - \delta^T P\delta(S+S^T).$$

If we let $A = V(P, q)$, they are reduced to

$$\tilde{B} = B + (\gamma - B\delta)\delta^T\left[\frac{1-\delta^T q}{\delta^T P\delta}(P+P^T) + \frac{2qq^T}{\delta^T q}\right] \tag{3.8}$$

and

$$\tilde{H} = H + \frac{(\delta - H\gamma)\delta^T G H}{\delta^T G H\gamma - \delta^T q \cdot \delta^T P\delta}, \tag{3.9}$$

where $$G=\delta^T q(1-\delta^T q)(P+P^T)+2(\delta^T P\delta)qq^T.$$

Furthermore, we put forward some updates as a special case:

(i) In (3.6) and (3.7), letting $P=B$ and $P=H$ respectively and $S=0$, we have

$$\tilde{B}=B+\frac{(\gamma-B\delta)\delta^T(B+B^T)}{\delta^T B\delta} \tag{3.10}$$

and

$$\tilde{H}=H+\frac{(\delta-H\gamma)\delta^T(H+H^T)H}{\delta^T(H+H^T)H\gamma-\delta^T H\delta}. \tag{3.11}$$

(ii) In (3.6) and (3.7), letting $S=0$ and $P=qq^T$, $\forall q\in R^n$ with $q^T\delta\neq 0$, we have

$$\tilde{B}=B+\frac{2(\gamma-B\delta)q^T}{\delta^T q} \tag{3.12}$$

corresponding to Broyden's rank-one update and

$$\tilde{H}=H+\frac{2(\delta-H\gamma)q^T H}{q^T(2H\gamma-\delta)}, \tag{3.13}$$

which can be, evidently, further specified when concrete vectors are used as q; thus if $q=\gamma-B\delta$, then we obtain the symmetric rank-one update

$$\tilde{B}=B+\frac{2(\gamma-B\delta)(\gamma-B\delta)^T}{\delta^T(\gamma-B\delta)}. \tag{3.14}$$

Of course, (3.10—3.14) may also be regarded as specific forms of (3.8) or (3.9); and in passing, it should be noted that only when the same A is used and B is nonsingular, are (3.4) and (3.5) invertible to each other, and thus (3.10) and (3.11) are not so.

Optimization often requires that the updates used are symmetric, but, $\tilde{B}$ in (3.6) and (3.8), except in (3.14), are generally non-symmetric even if B is. We next derive a symmetric one by Powell's symmetrization technique.

In (3.6) let B be a symmetric matrix, $P=W$, where W is an arbitrary matrix in $R^{n\times n}$ with $\delta^T W\delta\neq 0$ and $S=0$, and let $W^{(0)}=W$. At each iteration k we have

$$E^{(k)}=B+\frac{(\gamma-B\delta)\delta^T(W^{(k)}+W^{(k)T})}{\delta^T W^{(k)}\delta} \tag{3.15}$$

and

$$W^{(k+1)}=(E^{(k)}+E^{(k)T})/2. \tag{3.16}$$

Under certain conditions, straightforward algebraic operation gives that the limit of the sequence $\{W^{(k)}\}$ is the symmetric matrix

$$\tilde{B}=B+(2\gamma\gamma^T-\gamma\delta^T B-B\delta\gamma^T)/\delta^T\gamma. \tag{3.17}$$

It is somewhat surprising and interesting that (3.17) is independent of $W^{(0)}$ and satisfies

$$(2-\delta^T B\delta/\delta^T\gamma)\gamma=\tilde{B}\delta \tag{3.18}$$

which is also satisfied by the parts of Huang's family for Jacobian matrices, namely

$$\tilde{B}=B-\frac{B\delta(\mu B^T\delta+\gamma)^T}{(\mu B^T\delta+\gamma)^T\delta}+\rho\cdot\frac{\gamma(\varphi B^T\delta+\gamma)^T}{(\varphi B^T\delta+\gamma)^T\delta}, \tag{3.19}$$

where $\rho=2-\delta^T B\delta/\delta^T\gamma$, μ and φ are arbitrary parameters with $\mu\neq-\gamma^T\delta/\delta^T B\delta$ and $\varphi\neq-\gamma^T\delta/\delta^T B\delta$.

4. Rank-Two Updates

For the quadratic mapping

$$Q(x)=\hat{F}+B(x-\hat{x})+(x-\hat{x})^TA(x-\hat{x})\gamma-(x-\hat{x})^T\bar{A}(x-\hat{x})B\delta, \tag{4.1}$$

where A and $\bar{A}$ are any one-reduction matrices of δ, we have the same (3.2), and when $\|\delta\|$ is small enough the Jacobian matrix $\tilde{B}$ of (4.1) at $\tilde{x}$ can be regarded as a better approximation to that of F at the same point.

(4.1) can be reduced to

$$Q(x)=\tilde{F}+[B+\gamma\delta^T(A+A^T)-B\delta\delta^T(\bar{A}+\bar{A}^T)](x-\tilde{x})$$
$$+(x-\tilde{x})^TA(x-\tilde{x})\gamma-(x-\tilde{x})^T\bar{A}(x-\tilde{x})B\delta, \tag{4.2}$$

from which we obtain at once the general rank-two update

$$\tilde{B}=B+\gamma\delta^T(A+A^T)-B\delta\delta^T(\bar{A}+\bar{A}^T), \tag{4.3}$$

where A and $\bar{A}$ are arbitrary one-reduction matrices of δ.

Of course, the corresponding update for the inverse Jacobian matrix could be derived by using the Sherman-Morrison formula twice in succession. However we do not do so because the update would be quite complicated and inapplicable, and in practice, Gill and Murray's implementation of quasi-Newton methods has already made known that Jacobian matrix is a better choice than its inverse.

Similarly, by choosing specific matrices as A and $\bar{A}$, diverse updates can be derived. Here we cite one example.

Let $A=U(P,\ S)$ and $\bar{A}=V(K,\ q)$. From (4.3) it follows that

$$\tilde{B}=B+\gamma\delta^T\Big[\frac{1+\delta^TS\delta}{\delta^TP\delta}(P+P^T)-(S+S^T)\Big]$$
$$-B\delta\delta^T\Big[\frac{1-\delta^Tq}{\delta^TK\delta}(K+K^T)+\frac{2qq^T}{\delta^Tq}\Big], \tag{4.4}$$

which is still quite general. And in (4.4), if B is symmetric and

$$P=\gamma\gamma^T,\ S=\varphi B,\ K=B,\ q=\varphi\gamma, \tag{4.5}$$

where $\varphi\in R^1$ and $\varphi\neq 0$, then

$$\tilde{B}=B+2\gamma\Big[\frac{(1+\varphi\delta^TB\delta)\gamma^T}{\gamma^T\delta}-\varphi\delta^TB\Big]-2B\delta\Big[\frac{(1-\varphi\gamma^T\delta)\delta^TB}{\delta^TB\delta}+\varphi\gamma^T\Big], \tag{4.6}$$

which is a symmetric rank-two update and corresponds to Broyden's family. Further, let

$$\varphi=1/\gamma^T\delta \quad\text{and}\quad \varphi=0;$$

from (4.6) it follows respectively that

$$\tilde{B}=B+2(\mu\gamma\gamma^T-B\delta\gamma^T-\gamma\delta^TB)/\gamma^T\delta, \tag{4.7}$$

where $\mu=1+\delta^TB\delta/\gamma^T\delta$, and

$$\tilde{B}=B+\frac{2\gamma\gamma^T}{\gamma^T\delta}-\frac{2B\delta\delta^TB}{\delta^TB\delta}, \tag{4.8}$$

which correspond to DFP and BFGS updates.

5. Second Order Quasi-Newton Equation

Now we introduce the second order quasi-Newton equation.

It is easy to verify that the $\widetilde{B}$ in the general rank-two update (4.3) satisfies

$$\widetilde{B}\delta = 2\gamma - B\delta, \tag{5.1}$$

and so does the $\widetilde{B}$ in the rank-one update (3.4). (In fact, (3.4) can be regarded as a specific case of (4.3) with $A = \overline{A}$.) We have the following result in this respect.

Theorem 2. *Suppose that F: $D \subset R^n \to R^n$ is twice continuously Fréchet-differentiable on $D_0 = \text{int}\,(D)$ and that $\hat{x}$, $\tilde{x} \in D_0$ and*

$$Q(x) = \widetilde{F} + \widetilde{B}(x - \tilde{x}) + (x - \tilde{x})^T A (x - \tilde{x})\gamma - (x - \tilde{x})^T \overline{A}(x - \tilde{x})B\delta, \tag{5.2}$$

where $\widetilde{B}$ satisfies (5.1), A and $\overline{A}$ are arbitrary one-reduction matrices of δ. Then, for any $t \in R^1$ which is non-zero and sufficiently small,

$$Q(\tilde{x} + t\delta) - F(\tilde{x} + t\delta) = tR(t, \delta), \tag{5.3}$$

where $\lim\limits_{\delta \to 0} R(t, \delta)/\|\delta\|^2 = 0$, $\lim\limits_{t \to 0} R(t, \delta) = \overline{R}(\delta)$ and $\lim\limits_{\delta \to 0} \overline{R}(\delta)/\|\delta\|^2 = 0$.

Proof.

$$F(\tilde{x} + t\delta) = \widetilde{F} + tF'(\tilde{x})\delta + \frac{1}{2} F''(\tilde{x})(t\delta)(t\delta) - R_1(t\delta), \tag{5.4}$$

where $F'(x)$ and $F''(x)$ denote the first and second Fréchet-derivatives at $x \in D_0$ respectively (the same below), and $\lim\limits_{t\delta \to 0} R_1(t\delta)/(t^2\|\delta\|^2) = 0$.

$$\hat{F} = \widetilde{F} - F'(\tilde{x})\delta + \frac{1}{2} F''(\tilde{x})\delta\delta - R_2(\delta), \tag{5.5}$$

where $\lim\limits_{\delta \to 0} R_2(\delta)/\|\delta\|^2 = 0$.

$$\widetilde{F} = \hat{F} + B\delta + \frac{1}{2} F''(\hat{x})\delta\delta + R_3(\delta), \tag{5.6}$$

where $\lim\limits_{\delta \to 0} R_3(\delta)/\|\delta\|^2 = 0$.

By (5.2), (5.4), (5.1), (5.5) and (5.6) we have

$$\begin{aligned}
&Q(\tilde{x} + t\delta) - F(\tilde{x} + t\delta) \\
&= t\widetilde{B}\delta - tF'(\tilde{x})\delta + t^2\Big[\gamma - B\delta - \frac{1}{2} F''(\tilde{x})\delta\delta\Big] + R_1(t\delta) \\
&= t(2\gamma - B\delta) - tF'(\tilde{x})\delta + t^2\Big[\gamma - B\delta - \frac{1}{2} F''(\tilde{x})\delta\delta\Big] + R_1(t\delta) \\
&= t(\gamma - B\delta) + t[\gamma - F'(\tilde{x})\delta] + t^2(\gamma - B\delta) - \frac{t^2}{2} F''(\tilde{x})\delta\delta + R_1(t\delta) \\
&= (t + t^2)(\gamma - B\delta) + t\Big[-\frac{1}{2} F''(\tilde{x})\delta\delta + R_2(\delta)\Big] - \frac{t^2}{2} F''(\tilde{x})\delta\delta + R_1(t\delta) \\
&= (t + t^2)\Big[\gamma - B\delta - \frac{1}{2} F''(\tilde{x})\delta\delta\Big] + tR_2(\delta) + R_1(t\delta) \\
&= (t + t^2)R_3(\delta) + (t + t^2)[F''(\hat{x})\delta\delta - F''(\tilde{x})\delta\delta]/2 + tR_2(\delta) + R_1(t\delta),
\end{aligned} \tag{5.7}$$

where

$$\|F''(\hat{x})\delta\delta - F''(\tilde{x})\delta\delta\| = \|[F''(\hat{x}) - F''(\tilde{x})]\delta\delta\| \leqslant \|F''(\hat{x}) - F''(\tilde{x})\| \cdot \|\delta\|^2. \tag{5.8}$$

By the assumption that $F''(x)$ is continuous on D_0 it is known that

$$F''(\hat{x})\delta\delta - F''(\tilde{x})\delta\delta = R_4(\delta), \tag{5.9}$$

where $\lim_{\delta\to 0} R_4(\delta)/\|\delta\|^2 = 0$.

By (5.7), (5.9) and letting

$$R(t,\ \delta) = R_1(t\delta)/t + R_2(\delta) + (1+t)R_3(\delta) + \frac{1+t}{2} R_4(\delta), \tag{5.10}$$

we have

$$Q(\tilde{x}+t\delta) - F(\tilde{x}+t\delta) = tR(t,\ \delta). \tag{5.11}$$

And from (5.10),

$$\lim_{\delta\to 0} R(t,\ \delta)/\|\delta\|^2 = 0, \tag{5.12}$$

and

$$\lim_{t\to 0} R(t,\ \delta) = R_2(\delta) + R_3(\delta) + R_4(\delta)/2. \tag{5.13}$$

Hence Theoren 2 is proved.

In view of Theorem 2, (5.2) can be regarded as the second order approximation to F at $\tilde{x}$, and (5.1) is called the second order quasi-Newton equation. Finding the $\tilde{B}$ that satisfies (5.1), we can obtain a lot of new updates which are, correspondingly, called second order, and those derived in the previous sections may serve as examples.

6. Numerical Results

Limited trials in unconstrained optimization have been carried out to compare our methods with the conventional quasi-Newton methods. We use double precision FORTRAN program based upon Gill and Murrays' implementation of quasi-Newton methods with Himmelblau's convergence criterion and the procedure QUAD for the line search. B is factorized by using Algorithm 6.7.2 twice in succession with $\sigma^{(1)}$ defined by (6.7.25) (See [11]). ES-1022b computer is used.

Tables 1—6, where NI, NF, NG, NC and F are the same as in [11], give results for Problems 1—6 of Appendix B of [11] respectively. We emphasize that Algorithms 1—6 use the same FORTRAN program; the differences lie only in that the first three algorithms (i. e. 1, 2 and 3) use BFGS update (6.7.17) of [11] and the rest use our update (4.8), i. e.

$$\tilde{B} = B + \frac{2\gamma\gamma^T}{\gamma^T\delta} + \frac{2agg^T}{g^T\delta}. \tag{6.1}$$

In Algorithms 1, 2, 4 and 5 the unit matrix I is chosen as the initial matrix B_0; besides, in Algorithms 2 and 5 the initial matrix is transformed into $B_0 = (\delta_0^T\gamma_0/\delta_0^T\delta_0)I$ before it is updated. In Algorithms 3 and 6 a numerical approximation of the Hessian matrix of the objective function at the initial point is chosen as the initial matrix.

By comparing the sums of the entries under NC in Tables 1—6 we see that for each of Problems 1—6, Algorithm 4 is better than Algorithm 1, Algorithm 5 is better than Algorithm 2 except for Problem 6, and Algorithm 6 is superior to the others except for Problems 2 and 6. We conclude that, at least for Problems 1—6, our update is preferable to the conventional one. Of course, a larger sample of problems or differences of details in programming might yield a different result.

Table 1

A comparison in problem 1

Algorithm	NI	NF	NG	NC	F
1	27	177	28	233	0.44D-15
2	31	193	32	257	0.21D-13
3	32	196	35	266	0.13D-15
4	24	158	25	208	0.12D-14
5	27	177	28	233	0.27D-16
6	21	150	24	198	0.47D-13

Table 2

A comparison in problem 2

Algorithm	NI	NF	NG	NC	F
1	51	295	52	503	0.14D-10
2	43	247	44	423	0.38D-13
3	82	427	87	775	0.35D-12
4	41	235	42	403	0.29D-12
5	34	211	35	351	0.13D-12
6	63	327	68	599	0.15D-12

Table 3

A comparison in problem 3

Algorithm	NI	NF	NG	NC	F
1	62	300	63	552	0.41D-12
2	85	405	86	749	0.74D-13
3	50	235	55	455	0.66D-12
4	49	259	50	459	0.85D-15
5	74	355	75	655	0.26D-19
6	47	237	52	445	0.61D-14

Table 4

A comparison in problem 4

Algorithm	NI	NF	NG	NC	F
1	81	352	82	680	0.37D-19
2	97	408	98	800	0.28D-20
3	49	228	54	444	0.14D-19
4	50	242	51	446	0.52D-20
5	50	246	51	450	0.31D-20
6	26	142	31	266	0.37D-20

Table 5

A comparison in problem 5

Algorithm	NI	NF	NG	NC	F
1	27	159	28	243	0.15D-11
2	28	173	29	260	0.41D-12
3	23	131	27	212	0.76D-12
4	20	134	21	197	0.14D-12
5	23	149	24	221	0.54D-15
6	16	103	20	163	0.70D-16

Table 6

A comparison in problem 6

Algorithm	NI	NF	NG	NO	F
1	45	221	46	681	0.16D-10
2	38	195	39	585	0.92D-12
3	66	308	77	1078	0.80D-11
4	45	217	46	677	0.11D-12
5	42	213	43	643	0.19D-11
6	53	254	64	894	0.13D-10

7. Conclusion

Even though we have paid partioular attention to the updates corresponding to the existing first order ones, it does not mean that others must be bad. In order to judge what is good and what is bad, further theoretical analyses and numerical tests are necessary. At all events, the use of the second order quasi-Newton equation as the new starting point seems to have inspiring prospects.

References

[1] 席少霖,非线性最优化计算方法近况,应用数学与计算数学, 1982 年第 2 期, 1—14.

[2] D. Jacobs, The State of the Art in Numerical Analysis, Academic Press, 1977, 237—256.

[3] 朱凤石,赵瑞安,求函数极小值的变尺度法,计算机应用与应用数学, 1975 年第 9 期, 1—16.

[4] 何旭初,最优化方法,科学出版社, 1978.

[5] 王德人,非线性方程组解法与最优化方法,人民教育出版社, 1979.

[6] C. G. Broydon, Quasi-Newton methods and application to function minimization, *Math. Comp.*, **21** (1967), 368—381.

[7] P. E. Gill, W. Murray, Quasi-Newton methods for unconstrained optimization, *Journal of the Institute of Mathematics and its Applications*, 9, 1972, 91—108.

[8] M. J. D. Powell, A New Algorithm for Unconstrained Optimization, See "Rosen, J. B., Nonlinear Programming", Academic Press, New York, 1972.

[9] J. M. Ortega, W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, 1970.

[10] J. E. Dennis, J. J. More, Quasi-Newton methods, motivation and theory, *S. I. A. M. Review*, **19** (1977), 46—89.