International Journal of Pure and Applied Mathematics
Volume 28 No. 2 2006, 177-186

A NEW FUNCTION-BASED MULTI-STEP QUASI-NEWTON METHOD

Issam A.R. Moghrabi
Department of Computer Science
Faculty of Science
Beirut Arab University
P.O. Box 11-5020, Beirut, LEBANON
e-mail: i_moghrabi@yahoo.com

Abstract: Multi-step quasi-Newton methods for unconstrained optimization were introduced by the authors ([7, 8]). At each step of the iterative process, these methods employ two polynomials, one to define a path interpolating recent iterates in the variable-space and the other to approximate the gradient as the path is followed. Numerical experiments described in [7] strongly indicated that several multi-step methods yield substantial computational gains over the standard (one-step) BFGS method. In this paper, we consider how to modify the structure of such methods to provide a more general model of the gradient with the intention of improving the approximation. The model is exploited in utilizing readily computed function values in updating the Hessian approximation. The results of numerical experiments on the new methods are reported and compared with those produced by existing methods.

AMS Subject Classification: 65K10
Key Words: unconstrained optimization, quasi-Newton methods, multi-step methods

1. Introduction

We address the problem of the numerical determination of an unconstrained minimum of the given objective function $f(\underline{x})$, where $\underline{x} \in R^n$ and f is assumed to be twice continuously differentiable with gradient and Hessian denoted by $\underline{g}$

Received: December 21, 2005

and G, respectively. The algorithms we consider here do not require G to be explicitly available and are iterative in character, generating from $\underline{x}_0$ (a given starting point) a sequence $\{\underline{x}_i\}$ of approximations to the desired minimum. In quasi-Newton methods, each Hessian $G(\underline{x}_i)$ is approximated by a corresponding matrix B_i. A new estimate, $\underline{x}_{i+1}$, is computed from the current one, $\underline{x}_i$, through use of a direction $\underline{p}_i$ which is determined by analogy with Newton's method:

$$\underline{p}_i = -B_i^{-1}\underline{g}(\underline{x}_i); \tag{1}$$

for example, we might execute a line-search along the ray $\{\underline{x}(t) : \underline{x}(t) \equiv \underline{x}_i + t\underline{p}_i, t \geq 0\}$. Before proceeding to the next iteration, it is necessary to obtain an estimate B_{i+1} of the Hessian $G(\underline{x}_{i+1})$, so that the process described above may be repeated from $\underline{x}_{i+1}$. Standard (one-step) quasi-Newton methods determine B_{i+1} by using a formula of the type

$$B_{i+1} = B_i + U(B_i, \underline{s}_i, \underline{y}_i), \tag{2}$$

where the term $U(B_i, \underline{s}_i, \underline{y}_i)$ is usually of rank one or two and where

$$\underline{s}_i \stackrel{\text{def}}{=} \underline{x}_{i+1} - \underline{x}_i, \tag{3}$$

$$\underline{y}_i \stackrel{\text{def}}{=} \underline{g}(\underline{x}_{i+1}) - \underline{g}(\underline{x}_i), \tag{4}$$

(in (2), we draw special attention to the fact that the term $U(B_i, \underline{s}_i, \underline{y}_i)$ depends only on the most recent step-vectors, $\underline{s}_i$ and $\underline{y}_i$, and not on any previous ones). The best-known and most widely-used example of such a formula is the BFGS update (Broyden [10], Fletcher [9], Goldfarb [5], Shanno [12])):

$$B_{i+1} = B_i - \frac{B_i\underline{s}_i\underline{s}_i^T B_i}{\underline{s}_i^T B_i \underline{s}_i} + \frac{\underline{y}_i\underline{y}_i^T}{\underline{s}_i^T\underline{y}_i}. \tag{5}$$

This formula produces a matrix B_{i+1} which satisfies the secant equation (as do most other such formulae):

$$B_{i+1}\underline{s}_i = \underline{y}_i. \tag{6}$$

The relevance of this equation is that it constitutes (Ford and Moghrabi [7, 8]) an approximation to the Newton equation

$$G(\underline{x}_{i+1})\underline{x}'(\tau^*) = \underline{g}'(\underline{x}(\tau^*)), \tag{7}$$

satisfied by $G(\underline{x}_{i+1})$ itself. In equation (7), $\underline{x} \equiv \underline{x}(\tau)$ (where $\tau \in R$) is a continuously differentiable curve in R^n passing through $\underline{x}_{i+1}$, while τ^* is the value of τ such that $\underline{x}(\tau^*) = \underline{x}_{i+1}$. The actual derivation by this means of the Secant equation from the Newton equation is described in [8].

Ford and Moghrabi [8] proposed a "multi-step" approach to the issue of computing approximations to the Hessian matrix. In this new approach, $\underline{x}(\tau)$ is chosen to be a vector polynomial form in τ of degree m which interpolates the

$m+1$ most recent iterates $\{\underline{x}_{i-m+k+1}\}_{k=0}^{m}$. Correspondingly, an interpolating polynomial approximation $\hat{\underline{g}}(\tau)$ is constructed for $\underline{g}(\underline{x}(\tau))$. This approximation has degree m and interpolates the known gradient values $\{\underline{g}(\underline{x}_{i-m+k+1})\}_{k=0}^{m}$. By means of this technique, a new approximation

$$B_{i+1}\underline{r}_i = \underline{w}_i \tag{8}$$

to the Newton equation is obtained (see [8]). In (8), the vectors $\underline{r}_i$ and $\underline{w}_i$ are defined by

$$\underline{r}_i \stackrel{\text{def}}{=} \sum_{j=0}^{m-1} \underline{s}_{i-j}\{\sum_{k=m-j}^{m} \mathcal{L}'_k(\tau_m)\} = \underline{x}'(\tau_m), \tag{9}$$

$$\underline{w}_i \stackrel{\text{def}}{=} \sum_{j=0}^{m-1} \underline{y}_{i-j}\{\sum_{k=m-j}^{m} \mathcal{L}'_k(\tau_m)\} \approx \underline{g}'(\underline{x}(\tau_m)), \tag{10}$$

where $\mathcal{L}_j(\tau)$ is the j-th Lagrange polynomial of degree m corresponding to the real numbers $\{\tau_k\}_{k=0}^{m}$, so that $\mathcal{L}_j(\tau_j)=1$ and $\mathcal{L}_j(\tau_i)=0$ for $i\neq j$ (the abscissae $\{\tau_k\}_{k=0}^{m}$ are the values of τ associated with the iterates $\{\underline{x}_{i-m+k+1}\}_{k=0}^{m}$ on the path $\underline{x}(\tau)$:

$$\underline{x}(\tau_k) = \underline{x}_{i-m+k+1}, \text{ for } k=0,1,\ldots,m.) \tag{11}$$

It is now a simple matter to construct matrices satisfying the new condition (8), in place of (6). For example, we may use an updating formula such as (5), where $\underline{s}_i$ and $\underline{y}_i$ are replaced by $\underline{r}_i$ and $\underline{w}_i$, respectively. We note that such an approach employs not only the most recent step-vectors $\underline{s}_i$ and $\underline{y}_i$, but also previous ones (see (9) and (10)).

Ford and Moghrabi [7, 8] reported that extensive numerical experimentation had shown that methods for which $m=2$ generally performed better than those for which $m=3$. In the remainder of this paper, therefore, we will consider only two-step methods.

2. A Rational Form for $\hat{\underline{g}}$

We wish to explore the possibility of obtaining a more accurate approximation to $\underline{g}'(\underline{x}(\tau_m))$ to use in updating B_i. To this end, instead of using a polynomial approximation to $\underline{g}(\underline{x}(\tau))$, we investigate the use of a rational form:

$$\hat{\underline{g}}(\tau,\theta) \equiv \underline{q}(\tau)/(1+\theta\tau), \tag{12}$$

where $\underline{q}(\tau)$ is a (vector) quadratic form in τ (note, however, that we still continue to use a "plain" quadratic form for $\underline{x}(\tau)$). Because we are considering two-step methods, the curve $\hat{\underline{g}}(\tau,\theta)$ is still required to interpolate the three most recent values of $\underline{g}$, for a set (to be specified in more detail below) of distinct values

$\{\tau_j\}_{j=0}^2$:

$$\underline{\hat{g}}(\tau_j,\theta) = \underline{g}(\underline{x}_{i-j+1}), \text{ for } j = 0,1,2. \tag{13}$$

We now stipulate (as for the methods **A1**, **A2** and **A3** described in [7]) that the set $\{\tau_j\}_{j=0}^2$ has been chosen such that

$$\tau_1 = 0, \tag{14}$$

and we write

$$\tau_1 - \tau_0 = -\tau_0 \stackrel{\text{def}}{=} \rho_{i-1} > 0; \quad \tau_2 - \tau_2 = \tau_2 \stackrel{\text{def}}{=} \rho_i > 0, \tag{15}$$

where, for example, the quantities ρ_{i-1} and ρ_i could be defined (as they are in method **A1**, see [7], where other ways of defining ρ_{i-1} and ρ_i are also discussed) by

$$\rho_{i-1} = \| \underline{s}_{i-1} \|_2; \quad \rho_i = \| \underline{s}_i \|_2 \,. \tag{16}$$

Differentiating equation (12) with respect to τ, we obtain

$$\underline{\hat{g}}'(\tau,\theta) \equiv [\underline{q}'(\tau) - \theta\underline{\hat{g}}(\tau,\theta)]/(1+\theta\tau). \tag{17}$$

Now, since $\underline{q}(\tau) \equiv (1+\theta\tau)\underline{\hat{g}}(\tau,\theta)$ is (by definition) a quadratic form in τ, we can express it in its "Lagrangian" form:

$$\begin{aligned}\underline{q}(\tau) \;\; &\equiv \;\; \frac{\tau(\tau-\tau_0)}{\tau_2(\tau_2-\tau_0)}(1+\theta\tau_2)\underline{g}_{i+1} \\ &\quad + \frac{(\tau-\tau_0)(\tau-\tau_2)}{\tau_2\tau_0}\underline{g}_i - \frac{\tau(\tau-\tau_2)}{\tau_0(\tau_2-\tau_0)}(1+\theta\tau_0)\underline{g}_{i-1},\end{aligned}$$

using the fact that $\tau_1 = 0$. Therefore,

$$\begin{aligned}\underline{q}'(\tau) \;\; &\equiv \;\; \frac{(2\tau+\rho_{i-1})}{\rho_i(\rho_{i-1}+\rho_i)}(1+\theta\rho_i)\underline{g}_{i+1} \\ &\quad - \frac{(2\tau+\rho_{i-1}-\rho_i)}{\rho_{i-1}\rho_i}\underline{g}_i + \frac{(2\tau-\rho_i)}{\rho_{i-1}(\rho_{i-1}+\rho_i)}(1-\theta\rho_{i-1})\underline{g}_{i-1}\end{aligned} \tag{18}$$

(using equations (15)), from which it follows that

$$\underline{q}'(0) = \mu^{-1}\{\delta^{-1}(1+\theta\rho_i)\underline{g}_{i+1} - (\delta^{-1}-\delta)\underline{g}_i - \delta(1-\theta\rho_{i-1})\underline{g}_{i-1}\}, \tag{19}$$

where we define

$$\delta \stackrel{\text{def}}{=} \rho_i/\rho_{i-1}; \quad \mu \stackrel{\text{def}}{=} \rho_i + \rho_{i-1}. \tag{20}$$

Therefore, using (17),

$$\begin{aligned}\underline{q}'(0) - \theta\underline{\hat{g}}(0,\theta) &= \underline{q}'(0) - \theta\underline{g}_i \\ &= \mu^{-1}\{\delta^{-1}\underline{y}_i + \delta\underline{y}_{i-1} + \theta\rho_{i-1}\underline{y}_i - \theta\rho_i\underline{y}_{i-1}\} = \underline{\hat{g}}'(0,\theta).\end{aligned} \tag{21}$$

We now define

$$\phi(\tau) \stackrel{\text{def}}{\equiv} f(\underline{x}(\tau)). \tag{22}$$

Then

$$\phi'(\tau) \;\; \equiv \;\; \underline{x}'(\tau)^T \underline{g}(\underline{x}(\tau)); \tag{23}$$

$$\phi''(\tau) \equiv \underline{x}'(\tau)^T \underline{g}'(\underline{x}(\tau)) + \underline{x}''(\tau)^T \underline{g}(\underline{x}(\tau)). \quad (24)$$

We focus attention on the curvature of ϕ at $\tau = 0$, corresponding to $\underline{x}_i$:

$$\phi''(0) \equiv \underline{x}'(0)^T \underline{g}'(\underline{x}(0)) + \underline{x}''(0)^T \underline{g}_i \approx \underline{x}'(0)^T \hat{\underline{g}}'(0, \theta) + \underline{x}''(0)^T \underline{g}_i. \quad (25)$$

Now, since $\underline{x}(\tau)$ is a quadratic form in τ, it is easy to show that

$$\begin{aligned} \underline{x}'(0) &= \mu^{-1}\{\delta^{-1}\underline{s}_i + \delta \underline{s}_{i-1}\}; \\ \underline{x}''(\tau) &\equiv \mu^{-1}\{(2/\rho_i)\underline{s}_i - (2/\rho_{i-1})\underline{s}_{i-1}\}, \quad \forall \tau. \end{aligned}$$

Therefore, using (21) and (25), we obtain

$$\begin{aligned} \phi''(0) \approx\ & \mu^{-2}\{\delta^{-1}\underline{s}_i + \delta \underline{s}_{i-1}\}^T \{\delta^{-1}\underline{y}_i + \delta \underline{y}_{i-1} + \theta\rho_{i-1}\underline{y}_i - \theta\rho_i \underline{y}_{i-1}\} \\ & + \mu^{-1}\{(2/\rho_i)\underline{s}_i - (2/\rho_{i-1})\underline{s}_{i-1}\}^T \underline{g}_i. \end{aligned} \quad (26)$$

We define

$$\lambda_{ij} \stackrel{\text{def}}{=} \underline{s}_i^T \underline{y}_j; \quad \sigma_{ij} \stackrel{\text{def}}{=} \underline{s}_i^T \underline{g}_j. \quad (27)$$

Then (25) becomes

$$\begin{aligned} \phi''(0) \approx\ & \mu^{-2}\{\delta^{-2}\lambda_{ii} + \lambda_{i,i-1} + \lambda_{i-1,i} + \delta^2 \lambda_{i-1,i-1} + 2(1+\delta^{-1})\sigma_{ii} \\ & -2(1+\delta)\sigma_{i-1,i} + \hat{\theta}(\delta^{-1}\underline{s}_i + \delta \underline{s}_{i-1})^T (\delta^{-1}\underline{y}_i - \underline{y}_{i-1})\}, \end{aligned} \quad (28)$$

on defining

$$\hat{\theta} \stackrel{\text{def}}{=} \theta \rho_i. \quad (29)$$

3. The New Algorithm

The algorithm proposed here derives from

$$\int_{\tau_1}^{\tau 2} \phi'(t)dt = f_{i+1} - f_i. \quad (30)$$

The expression for $\phi'(t)$ is given as

$$\begin{aligned} \phi'(t, v) = \mu^{-1} \Bigg[& \frac{(3t + 3\rho_{i-1}t + \rho_{i-1}t)(1 + v\rho_i)}{\mu \rho_i^2 (1 + vt)} \sigma_{i,i+1} \\ & - \frac{2t^3 + (3\rho_{i-1} - 2\rho_i)t^2 + (\rho_i^2 - 3\rho_i\rho_{i-1})t - \rho_i \rho_{i-1}^2}{\rho_i^2 \rho_{i-1}(1 + vt)} \sigma_{ii} \\ & + \frac{(1 - v\rho_{i-1})(2t^3 - \rho_{i-1}t^2 - \rho_{i-1}^2 t)}{\rho_i \rho_{i-1} \mu (1 + vt)} \sigma_{i,i-1} \\ & - \frac{(2t^3 + (2\rho_{i-1} - \rho_i)t^2 - \rho_i\rho_{i-1}t)(1 + v\rho_i)}{\rho_i \rho_{i-1} \mu (1 + vt)} \sigma_{i-1,i+1} \end{aligned}$$

$$+\frac{2t^3+(2\rho_{i-1}-3\rho_i)t^2-(3\rho_i\rho_{i-1}+\rho_i^2)t+\rho_i^2\rho_{i-1}}{\rho_{i-1}^2\rho_i(1+vt)}\sigma_{i-1,j}$$
$$\left.-\frac{(2t^3-(2\rho_{i-1}+\rho_i)t^2+\rho_i\rho_{i-1}t)(1-v\rho_{i-1})}{(1+vt)\rho_{i-1}^2\mu}\sigma_{i-1,i-1}\right]. \quad (31)$$

Thus, from (30) and (31) we obtain

$$\mu^{-1}\left\{\kappa_1\int_0^{t_2}\frac{2t^3+3\rho_{i-1}t^2+\rho_{i-1}^2t}{1+vt}dt\right.$$
$$\left.-k2\int_0^{t_2}\frac{2t^3+(3\rho_{i-1}-2\rho_i)t^2+(\rho_i^2-3\rho_i\rho_{i-1})t-\rho_i\rho_{i-1}^2}{(1+vt)}dt\right\}$$
$$+\kappa_3\int_0^{t_2}\frac{2t^3-\rho_{i-1}t^2-\rho_{i-1}^2t}{1+vt}dt-\kappa_4\int_0^{t_2}\frac{2t^3-(2\rho_{i-1}\rho_i)t^2+\rho_i\rho_{i-1}t}{1+vt}dt$$
$$+\kappa_5\int_0^{t_2}\frac{2t^3-(2\rho_{i-1}-3\rho_i)t^2-(3\rho_i\rho_{i-1}+\rho_i^2)t+\rho_i^2\rho_{i-1}}{(1+vt)}dt$$
$$-\kappa_6\int_0^{t_2}\frac{2t^3-(2\rho_{i-1}+\rho_i)t^2+\rho_i\rho_{i-1}t}{1+vt}dt=f_{i+1}-f_i\,, \quad (32)$$

where $\kappa_i \in R$ $(i = 1..6)$ are given by

$$\kappa_1=\frac{1+v\rho_i}{\mu\rho_i}\sigma_{i,i+1}\,, \qquad \kappa_2=\frac{\sigma_{ii}}{2\rho_i\rho_{i-1}}\,,$$
$$\kappa_3=\frac{1-\rho_{i-1}}{\rho_i\rho_{i-1}\mu}\sigma_{i,i-1}\,, \qquad \kappa_4=\frac{1+v\rho_i}{\rho_i\rho_{i-1}\mu}\,,$$
$$\kappa_5=\frac{\sigma_{i-1,i}}{\rho_{i-1}^2\rho_i}\,, \qquad \kappa_6=\frac{1-v\rho_{i-1}}{\rho_{i-1}^2\mu}\sigma_{i-1,i-1}\,.$$

If we now define the family of the functions

$$\gamma_i(v)=\int_0^{t_2}\frac{t^i}{1+vt}dt \quad (33)$$

for $i = 0, 1, 2, 3$.

We now have

$$\gamma_0(v)=\frac{\ln(|1+v\rho_i|)}{v}\,,$$
$$\gamma_1(v)=\frac{v\rho_i-\ln(|1+v\rho_i|)}{v^2}\,,$$

$$\gamma_2(v) = \frac{v^2\rho_i^2/2 + \ln(|1+v\rho_i|)}{v^3},$$

$$\gamma_3(v) = \frac{(\rho^3/3)v^3 - (\rho_i^2/2)v^2 + \rho_i v - \ln(|1+v\rho_i|)}{v^4}.$$

Therefore equation (32) can be written (in compact form) as

$$\begin{aligned} &2(\kappa_{-}\kappa_2 + \kappa_3 - \kappa_4 + \kappa_5 - \kappa_6)\gamma \\ &+ \{\rho_{i-1}(3\kappa_1 - 3\kappa_2 - \kappa_3 - 2\kappa_4 + 2\kappa_5 + 2\kappa_6) + \rho_i(2\kappa_2 + \kappa_4 + 3\kappa_5 + \kappa_6)\}\,\gamma_2 \\ &\quad + \left\{\rho_{i-1}^2(\kappa_1 - \kappa_3) + \rho_i\rho_{i-1}(3\kappa_2 + \kappa_4 - 3\kappa_5 + 1) - \kappa_5\rho_i^2\right\}\gamma_1 \\ &\qquad + (\rho_{i-1}\rho_i^2 - \rho_i\rho_{i-1}^2)\gamma_0 - \mu(f_{i+1} - f_i) = 0\,. \quad (34) \end{aligned}$$

A solution of the non-linear equation (34) provides a value for which is then used in computing the quantity $\underline{\hat{g}}'(\tau_2,\theta)$ that is used in updating the Hessian approximation.

4. Numerical Experiments

The new algorithm developed in Section 3 was compared (initially) with the standard (one-step) BFGS method. The distances ρ_i and ρ_{i-1} were defined by equations (15). We present below a sample of the results obtained (the results of more extensive tests on a larger set of functions are reported in Moghrabi [2]). Eleven test functions were used, each with either one or two starting-points, giving a total of twenty problems. Several of the functions used in the tests were taken from the set described by Moré, Garbow and Hillstrom [4] and can be used with varying dimension – in such cases, the tests were carried out on a range of appropriate dimensions and the results were summed, since lack of space prevents a complete tabulation of the individual figures. Information on the functions and the starting-points used, together with details of the actual implementation of such algorithms, may be found in Ford and Moghrabi [9].

The results of this first set of experiments are presented in Table 1. For each problem, the number of function/gradient evaluations required to solve the problem is given, followed by the number of iterations in brackets.

As a further test, the new method was compared with the two-step method **A1F** introduced in [9], which also uses function values to modify a two-step accumulative method (in the case of **A1F**, this is accomplished by making the path $\underline{x}(\tau)$ a rational function and determining the additional parameter by the reduction in the function-value). The results of this second comparison are given in Table 2.

Problem	**BFGS**	**NEW**
1(a)	542(530)	431(419)
1(b)	1653(1293)	1412(1074)
2(a)	631(612)	549(506)
2(b)	825(806)	770(710)
3(a)	159(122)	200(134)
3(b)	162(156)	291(104)
4(a)	807(386)	798(446)
4(b)	637(507)	813(553)
5(a)	634(596)	542(460)
5(b)	2390(2331)	1843(1645)
6(a)	2300(1786)	1991(1098)
6(b)	2136(2001)	1719(1527)
7(a)	280(226)	265(222)
7(b)	1967(965)	1859(922)
8(a)	5099(2272)	4670(1937)
8(b)	2822(1969)	2780(1655)
9(a)	1034(987)	780(718)
9(b)	1779(1559)	1673(1419)
10	132(126)	130(124)
11	1505(1262)	1133(841)
TOTALS	27494(20492)	24649(16487)

Table 1:

5. Summary and Conclusions

A new multi-step quasi-Newton optimization method has been constructed. The method is based upon the use of a rational form to approximate the behaviour of the gradient as a quadratic path in the variable-space is traversed. It determines the remaining parameter in the rational form (after the interpolatory conditions have been satisfied) by means of readily available values of the objective function. Numerical tests show that, while the new method exhibits substantial improvement in performance when compared with the BFGS method, it is slightly inferior to the two-step method **A1F** introduced in [9]. We therefore tentatively conclude that methods based directly upon function-values (as **A1F** is) are likely to be more successful than those utilizing function-values

Problem	**A1F**	NEW
1(a)	443(429)	431(419)
1(b)	1487(1101)	1412(1074)
2(a)	485(465)	549(506)
2(b)	679(652)	770(710)
3(a)	174(129)	200(134)
3(b)	108(97)	291(104)
4(a)	765(344)	798(446)
4(b)	569(463)	813(553)
5(a)	549(485)	542(460)
5(b)	1973(1877)	1843(1645)
6(a)	1912(1092)	1991(1098)
6(b)	1772(1607)	1719(1527)
7(a)	266(225)	265(222)
7(b)	1936(917)	1859(922)
8(a)	4696(1892)	4670(1937)
8(b)	2765(1819)	2780(1655)
9(a)	735(680)	780(718)
9(b)	1623(1379)	1673(1419)
10	130(124)	130(124)
11	1050(718)	1133(841)
TOTALS	24117(16495)	24649(16487)

Table 2:

indirectly by means of estimates.

References

[1] R. Brent, *Algorithms for Minimization without Derivatives*, Prentice-Hall, Englewood Cliffs, New Jersey (1973).

[2] C.G. Broyden, The convergence of a class of double-rank minimization algorithms - Part 2: The new algorithm, *J. Inst. Math. Applic.*, **6** (1970), 222-231.

[3] J.E. Dennis, R.B. Schnabel, Least change secant updates for quasi-Newton methods, *SIAM Review*, **21** (1979), 443-459.

[4] R. Fletcher, A new approach to variable metric algorithms, *Comput. J.*, **13** (1970), 317-322.

[5] R. Fletcher, *Practical Methods of Optimization*, Second Edition, Wiley, New York (1987).

[6] J.A. Ford, R.A. Ghandhari, On the use of function-values in unconstrained optimisation, *J. Comput. Appl. Math.*, **28** (1989), 187-198.

[7] J.A. Ford, R.A. Ghandhari, Efficient utilisation of function-values in quasi-Newton methods, *Colloquia Mathematica Societatis János Bolyai*, **59** (1990), 49-64.

[8] J.A. Ford, R.A. Ghandhari, On the use of curvature estimates in quasi-Newton methods, *J. Comput. Appl. Math.*, **35** (1991), 185-196.

[9] J.A. Ford, I.A. Moghrabi, Alternative parameter choices for multi-step quasi-Newton methods, *Optimization Methods and Software*, **2** (1993), 357-370.

[10] J.A. Ford, I.A. Moghrabi, Using function-values in multi-step quasi-Newton methods, *J. Comput. Appl. Math.*, **66** (1996), 201-211.

[11] J.A. Ford, A.F. Saadallah, A rational function model for unconstrained optimization, *Colloquia Mathematica Societatis János Bolyai*, **50** (1986), 539-563.

[12] D. Goldfarb, A family of variable metric methods derived by variational means, *Math. Comp.*, **24** (1970), 23-26.

[13] D.F. Shanno, Conditioning of quasi-Newton methods for function minimization, *Math. Comp.*, **24** (1970), 647-656.

[14] D.F. Shanno, K.H. Phua, Matrix conditioning and nonlinear optimization, *Math. Programming*, **14** (1978), 149-160.