New Dual Parameter Quasi-Newton Methods

Issam A.R. Moghrabi

***Abstract*—** We develop a framework model for the development of multi-step quasi-Newton methods which utilizes values of the objective function. The model developed here is constructed using iteration genereted data from the m+1 most recent iterates/gradient evaluations. The model hosts double free parameters which introduce a certain degree of flexibility. This permits the interpolating polynomials to exploit available computed function values which are otherwise discarded and left unused. Two new algorithms are derived for which function values are incorporated in the update of the inverse Hessian approximation at each iteration, in an attempt to accelerate convergence. The idea of incorporating function values is not new within the context of quasi-Newton methods but the presentation made in this paper constitutes a new approach for such algorithms. It has been shown in several earlier works that Including function values data in the update of the Hessian approximation numerically improves the convergence of Secant-like methods. The numerical scores of the new methods are reported with promising performance results.

***Keywords*—** Unconstrained optimization, quasi-Newton methods, multi-step methods, function value algorithms, Newton Method.

I. Introduction

THE "Newton Equation", which may be regarded as a generalization of the "Secant Equation" (Dennis and Schnabel [3,4]), is usually employed in the construction of quasi-Newton methods for optimization.

Let $f(\underline{x})$ be the objective function, where $\underline{x}\ \varepsilon\ R^n$, and let $\underline{g}$ and G denote the gradient and Hessian of f, respectively. Let $X = \{\underline{x}(\tau)\}$ denote a differentiable path in R^n, where $\tau\ \varepsilon\ R$. Then, upon applying the Chain Rule to $\underline{g}(\underline{x}(\tau))$ in order to determine its derivative with respect to τ, we obtain

$$G(\underline{x}(\tau))\underline{x}'(\tau) = \underline{g}'(\underline{x}(\tau)). \tag{1}$$

In particular, if we choose for the path X to pass through the most recent iterate $\underline{x}_{i+1}$ (so that $\underline{x}(\tau_m) = \underline{x}_{i+1}$, say), then equation (1) provides a condition (termed the "Newton Equation" in [1,3,4]) which the Hessian $G(\underline{x}_{i+1})$ must satisfy:

$$G(\underline{x}_{i+1})\underline{x}'(\tau_m) = \underline{g}'(\underline{x}(\tau_m)). \tag{2}$$

Therefore, if B_{i+1} denotes an approximation to $G(\underline{x}_{i+1})$, if

$$\underline{r}_i \stackrel{\text{def}}{=} \underline{x}'(\tau_m) \tag{3}$$

and $\underline{w}_i$ denotes an approximation to $\underline{g}'(\underline{x}(\tau_m))$, it is reasonable (by equation (2)) to require that B_{i+1} should satisfy a similar relation such as

$$B_{i+1}\underline{r}_i = \underline{w}_i. \tag{4}$$

(The derivation, in particular, of the Secant Equation from the Newton Equation is described in [2].) The relation in (4) defines the basis of the the Multi-step methods derived in [9,10,11]. In [9], it was proposed that X should be the vector polynomial which interpolates the $m+1$ most recent iterates $\{\underline{x}_{i-m+k+1}\}_{k=0}^{m}$ and that $\underline{w}_i$ should be obtained by constructing and differentiating the corresponding vector polynomial ($\underline{\hat{g}}(\tau)$, say) which interpolates the known gradient values $\{\underline{g}(\underline{x}_{i-m+k+1})\}_{k=0}^{m}$. Thus, the following explicit expressions for $\underline{r}_i$ and $\underline{w}_i$ may be derived [8,9,10,11]:

$$\begin{aligned}\underline{r}_i &= \underline{x}'(\tau_m)\\ &= \textstyle\sum_{j=0}^{m-1} \underline{s}_{i-j}\{\sum_{k=m-j}^{m} \mathcal{L}_k'(\tau_m)\};\end{aligned} \tag{5}$$

$$\begin{aligned}\underline{w}_i &= \underline{\hat{g}}'(\tau_m)\\ &= \textstyle\sum_{j=0}^{m-1} \underline{y}_{i-j}\{\sum_{k=m-j}^{m} \mathcal{L}_k'(\tau_m)\}\\ &\approx \underline{g}'\left(\underline{x}(\tau_m)\right),\end{aligned} \tag{6}$$

where

$$\underline{s}_i \stackrel{\Delta}{=} \underline{x}_{i+1} - \underline{x}_i, \tag{7}$$

$$\underline{y}_i \stackrel{\Delta}{=} \underline{g}(\underline{x}_{i+1}) - \underline{g}(\underline{x}_i) \tag{8}$$

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

$$\underline{x}(\tau_k) = \underline{x}_{i-m+k+1},\ \ for\ \ k = 0,1,\dots,m.) \tag{9}$$

We now stipulate, arbitrarily, that the set $\{\tau_j\}_{j=0}^{2}$ has been chosen such that

$$\tau_1 = 0, \tag{10}$$

and we write

$$\tau_1 - \tau_0 = -\tau_0 \stackrel{\text{def}}{=} \rho_{i-1} > 0;\ \ \tau_2 - \tau_1 = \tau_2 \stackrel{\text{def}}{=} \rho_i > 0, \tag{11}$$

Issam A.R. Moghrabi is with the Gulf University for Science and Technology, Kuwait. He is currently a professor with the Department of Management Information Systems. phone: +965-94453637; e-mail: moughrabi.i@ gust.edu.kw).

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

$$\rho_{i-1} = \| \underline{s}_{i-1} \|_2; \ \ \rho_i = \| \underline{s}_i \|_2. \tag{12}$$

In this paper, we will investigate a class of parameterized models for the path X. The free parameters in such models can be viewed as providing a means by which more information can be employed in updating the Hessian approximation (or its inverse), as in the methods derived in [8]. We describe, in the next sections, the non-linear model and then the particular technique (which essentially involves making use of the function-values at our disposal from the $m+1$ most recent iterations) that will be used in determining the free parameters. we finally present the numerical test results conducted on the methods and evaluate those results followed by conclusions.

II. The Nonlinear Model

We investigate here a model that embodies two free parameters (namely, ϑ_1 and ϑ_2), which allow us to specify that the model satisfies, simultaneously, more than one property. For example, we may require that the parameters are determined such that

$$\phi(\tau_2, \vartheta_1, \vartheta_2) = f_{i+1} \tag{13}$$

and

$$\phi(\tau_1, \vartheta_1, \vartheta_2) = f_i. \tag{14}$$

hold simustaneously.

Our model is defined by

$$\psi_i(\tau, \vartheta_1, \vartheta_2) \triangleq t^i / \propto (\tau, \vartheta_1, \vartheta_2) \tag{15}$$

for $i = 0,1,2,$ (for a 2-step method) and where

$$\propto (\tau, \vartheta_1, \vartheta_2) \triangleq 1 + \vartheta_1 \tau + \vartheta_2 \tau^2.$$

Alternatively, if we express this polynomial in its Lagrangian form, we obtain

$$u(\tau, \vartheta_1, \vartheta_2) \triangleq \frac{1}{\propto(\tau,\vartheta_1,\vartheta_2)} \tag{16}$$
$$\{\frac{\tau(\tau+\rho_{i-1})}{\rho_i(\rho_{i-1}+\rho_i)}[1+\vartheta_1\rho_i+\vartheta_2\rho_i^2]u_{i+1} - \frac{\tau+\rho_{i-1}}{\rho_{i-1}\rho_i}(\tau-\rho_i)u_i + \frac{\tau(\tau-\rho_i)}{\rho_{i-1}(\rho_{i-1}+\rho_i)}[1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2]u_{i-1}\}$$
$$\triangleq q(\tau,\vartheta_1,\vartheta_2)/\propto(\tau,\vartheta_1,\vartheta_2).$$

Now from (16), we obtain

$$u'(\tau,\vartheta_1,\vartheta_2) \triangleq \frac{1}{\propto(\tau,\vartheta_1,\vartheta_2)}[g'(\tau,\vartheta_1,\vartheta_2) \propto (\tau,\vartheta_1,\vartheta_2) - q(\tau,\vartheta_1,\vartheta_2)(\vartheta_1+2\vartheta_2\tau)] \tag{17}$$
$$= [g'(\tau,\vartheta_1,\vartheta_2) - (\vartheta_1+2\vartheta_2\tau)u(\tau,\vartheta_1,\vartheta_2)]/\propto(\tau,\vartheta_1,\vartheta_2)$$

From (17), it follows that $u'(\tau,\theta_1,\theta_2)$ at the three points, $\tau_0 = -\rho_{i-1}, \tau_1 = 0$ and $\tau_2 = \rho_i$ (see (11) and (12)), is given by the following expressions

$$u'(-\rho_{i-1},\vartheta_1,\vartheta_2) = \frac{q'(-\rho_{i-1},\vartheta_1,\vartheta_2)-(\vartheta_1-2\vartheta_2\rho_{i-1})u_{i-1}}{1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2}, \tag{18}$$

$$u'(0,\vartheta_1,\vartheta_2) = q'(0,\vartheta_1,\vartheta_2) - \vartheta_1 u_i, \tag{19}$$

and

$$u'(\rho_i,\vartheta_1,\vartheta_2) = \frac{q'(\rho_i,\vartheta_1,\vartheta_2)-(\vartheta_1+2\vartheta_2\rho_i)u_{i+1}}{1+\vartheta_1\rho_i+\vartheta_2\rho_i^2}. \tag{20}$$

Using (16), we derive

$$q'(\tau,\vartheta_1,\vartheta_2) = \left\{\frac{2\tau+\rho_{i-1}}{\rho_{i-1}}(1+\vartheta_1\rho_i+\vartheta_2\rho_i^2)u_{i+1} - \frac{(2\tau+\rho_{i-1}-\rho_i)\mu}{\rho_{i-1}\rho_i}u_i\right\} + \frac{2\tau-\rho_i}{\rho_{i-1}}[1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2]u_{i-1}\}\mu^{-1}, \tag{21}$$

from which we obtain the following quantities (for $\mu \triangleq \rho_i + \rho_{i-1}$, $\delta \triangleq - \rho_i/\rho_{i-1}$ and $\Delta u_j \triangleq u_{j+1} - u_j$):

$$q'(0,\vartheta_1,\vartheta_2) = \mu^{-1}\{-\delta^{-1}(1+\vartheta_1\rho_i+\vartheta_2\rho_i^2)u_{i+1} + (\delta^{-1}-\delta)u_i) + [1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2]\delta u_{i-1}$$
$$= \mu^{-1}\{-\delta^{-1}\Delta u_i - \delta\Delta u_{i-1} + \vartheta_2\rho_{i-1}\rho_i(\Delta u_{i-1}+\Delta u_i) + \vartheta_1\rho_{i-1}u_{i+1} + \vartheta_1\rho_i u_{i-1}\}, \tag{22}$$

$$q'(\rho_i,\vartheta_1,\vartheta_2) = \mu^{-1}\{(2-\delta^{-1})(1+\vartheta_1\rho_i+\vartheta_2\rho_i^2)u_{i+1} + (\delta-2+\delta^{-1})u_i - \delta(1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2)u_{i-1}$$
$$= \mu^{-1}\{2\Delta u_i - \delta^{-1}\Delta u_i + \delta\Delta u_{i-1} - \vartheta_1\rho_{i-1}u_{i-1} - \vartheta_1\rho_i u_{i+1} + \vartheta_2\rho_{i-1}\rho_i[\Delta u_{i-1}+\Delta u_i] + 2\vartheta_1\rho_i u_{i+1} + 2\vartheta_2\rho_i^2 u_{i+1} + \vartheta_1\rho_{i-1}u_{i+1} - \vartheta_1\rho_i u_{i-1}\}, \tag{23}$$

and

$$q'(-\rho_{i-1},\vartheta_1,\vartheta_2) = \mu^{-1}\{\delta^{-1}(1+\vartheta_1\rho_i+\vartheta_2\rho_i^2)u_{i+1} + (\delta-2+\delta^{-1})u_i + (\delta-2)(1-\vartheta_1\rho_{i-1}+\vartheta_2\rho_{i-1}^2)u_{i-1}$$
$$= \mu^{-1}\{\delta^{-1}\Delta u_i + (2-\delta)\Delta u_{i-1} - \vartheta_1\rho_{i-1}[\Delta u_{i-1}+\Delta u_i] - \vartheta_2\rho_{i-1}\rho_i u_{i+1} + \vartheta_1\rho_{i-1}u_{i-1} - 2\vartheta_2\rho_{i-1}^2 u_{i-1} + \vartheta_1\rho_i u_{i-1} - \vartheta_2\rho_{i-1}\rho_i u_{i-1} \tag{24}$$

We are, now. able to determine the quantities in (18), (19) and (20)

$$q'(0,\vartheta_1,\vartheta_2) - \vartheta_1 u_i = \mu^{-1}\{-\delta^{-1}\Delta u_i - \delta\Delta u_{i-1} + \vartheta_2\rho_{i-1}\rho_i[\Delta u_{i-1} + \Delta u_i] + \vartheta_1(\rho_{i-1}\Delta u_i - \rho_i\Delta u_{i-1})\} \quad (25)$$

and

$$q'(\rho_i,\vartheta_1,\vartheta_2) - (\vartheta_1 + 2\vartheta_2\rho_i)u_{i+1} = \mu^{-1}\{(2-\delta^{-1})\Delta u_i + \delta\Delta u_{i-1} + \vartheta_2\rho_{i-1}\rho_i[\Delta u_{i-1} + \Delta u_i] + \vartheta_1\rho_i[\Delta u_{i-1} + \Delta u_i]\} \quad (26)$$

and

$$q'(-\rho_{i-1},\vartheta_1,\vartheta_2) - (\vartheta_1 - 2\vartheta_2\rho_{i-1})u_{i-1} = \mu^{-1}\{\delta^{-1}\Delta u_i + (2-\delta)\Delta u_{i-1} - \vartheta_1\rho_{i-1}[\Delta u_{i-1} + \Delta u_i] - \vartheta_2\rho_{i-1}\rho_i[\Delta u_{i-1} + \Delta u_i]\}. \quad (27)$$

III. FUNCTION VALUE-BASED ALGORITHMS

A. *Algorithm Df1*

For this algorithm, the free parameters θ_1 and θ_2 and are determined via requiring the following relations

$$\phi'(0,\theta_1,\theta_2)[\tau_2 - \tau_0] = f(x_{i+1}) - f(x_{i-1}) \quad (28)$$

(for $\phi'(0,\theta_1,\theta_2) \triangleq x'(0,\theta_1,\theta_2)^T g_i$) and

$$x'(\tau_0,\theta_1,\theta_2)^T g_{i-1} = \omega'(\tau_0) \quad (29)$$

to hold simultanously and where $\omega(\tau)$ is the quadratic polynomial which interpolates the three most recent function values, f_{i-1} , f_i and f_{i+1}. That is,

$$\omega(\tau) = \alpha\tau^2 + \beta\tau + \gamma.$$

The coefficients α, β, and γ are given by

$$\gamma = f_i,$$
$$\beta = ((\delta^{-1} - \delta)f_i - \delta^{-1}f_{i+1} + \delta f_{i-1})/\mu,$$
$$\alpha = (f_{i+1} + (\delta - 1)f_i - \delta f_{i-1})/\mu\tau_2.$$

Now from (28), we obtain (using (25) and (28))

$$-\delta^{-1}\sigma_{ii} - \delta\sigma_{i-1,i} + \vartheta_2\rho_i\rho_{i-1}(\sigma_{ii} + \sigma_{i-1,i}) + \vartheta_1(\rho_{i-1}\sigma_{ii} - \rho_i\sigma_{i-1,i}) = f_{i+1} - f_{i-1}. \quad (30)$$

Also using (29) we obtain

$$\vartheta_1\{\rho_{i-1}(\sigma_{i,i-1} + \sigma_{i-1,i-1} - \mu\omega'(-\rho_{i-1}))\} + \vartheta_2\{\rho_{i-1}\mu\omega'(-\rho_{i-1}) + \rho_i\rho_{i-1}(\sigma_{i,i-1} + \sigma_{i-1,i-1})\} = (2-\delta)\sigma_{i-1,i-1} + \delta^{-1}\sigma_{i,i-1} - \mu\omega'(-\rho_{i-1}) \quad (31)$$

If we define the quantities

$$\eta \triangleq \rho_{i-1}\sigma_{i,i} - \rho_i\sigma_{i-1,i},$$
$$\zeta \triangleq \rho_{i-1}[\sigma_{i,i-1} + \sigma_{i-1,i-1} - \mu\omega'(-\rho_{i-1})]$$
$$\nu \triangleq \rho_i\rho_{i-1}(\sigma_{ii} + \sigma_{i-1,i})$$
$$\varepsilon \triangleq \rho_{i-1}\mu\omega'(-\rho_{i-1}) + \rho_i\rho_{i-1}(\sigma_{i,i-1} + \sigma_{i-1,i-1})$$
$$\pi \triangleq f_{i+1} - f_{i-1} + \delta^{-1}\sigma_{ii} + \delta\sigma_{i-1,i}$$

and

$$\lambda \triangleq (2-\delta)\sigma_{i-1,i-1} + \delta^{-1}\sigma_{i,i-1} - \mu\omega'(-\rho_{i-1})$$

and solve, simultaneously equations (28) and (29) for the unknowns and , we obtain the following expressions for the two parameters θ_1 and θ_2

$$\vartheta_1 = (\varepsilon\pi - \lambda\upsilon)/(\eta\varepsilon - \zeta\upsilon) \quad (32)$$

and

$$\vartheta_2 = (\eta\lambda - \zeta\pi)/(\varepsilon\eta - \lambda\zeta), \quad (33)$$

for a denominator that must be safeguarded numerically against vanishing. For this algorithm, we update the Hessian approximation to satisfy

$$B_{i+1}\{(2 - \delta^{-1} - \vartheta_2\rho_i\rho_{i-1} + \vartheta_1\rho_i)s_i + (\vartheta_1\rho_i - \vartheta_2\rho_i\rho_{i-1} + \delta)s_{i-1}\} = (2 - \delta^{-1} - \vartheta_2\rho_i\rho_{i-1} + \vartheta_1\rho_i)s_i + (\vartheta_1\rho_i - \vartheta_2\rho_i\rho_{i-1} + \delta)s_{i-1}. \quad (34)$$

B. *Algorithm Df2*

For this algorithm, the free parameters are determined by applying both

$$\int_{\tau_0}^{\tau_2}\phi'(\tau,\theta_1,\theta_2) = \phi(\tau_2,\theta_1,\theta_2) - \phi(\tau_0,\theta_1,\theta_2) \quad (35)$$

and

$$x'(\tau_2,\theta_1,\theta_2)^T g_{i+1} = \omega'(\tau_2) \quad (36)$$

(where ω is as defined in (29)) and requiring them to hold simultaneously. However, in order to reduce the complexity of the derivations. we will use

$$\int_{\tau_0}^{\tau_2}\phi'(\tau,\theta_1,\theta_2) \cong x'(0,\theta_1,\theta_2)^T g_i][\tau_2 - \tau_0]. \quad (37)$$

Now using (26), equation (36) takes the form

$$(2-\delta^{-1})\sigma_{i,i+1} + \delta\sigma_{i-1,i+1} - \vartheta_2\rho_i\rho_{i-1}(\sigma_{i,i+1} + \sigma_{i-1,i+1}) + \theta_1\rho_i(\sigma_{i,i+1} + \sigma_{i-1,i+1}) = (1 + \theta_1\rho_i + \vartheta_2\rho_i^2)\mu\omega'(\rho_i) \quad (38)$$

Solving the equations (36) and (37) for the parameters gives expressions (32) and (33). The update Hessian approximation matrix for this algorithm satisfies condition (34) as well.

IV. NUMERICAL EXPERIMENTS

The algorithms *Df1* and *Df2* developed above were first compared to one another. The results are reported in Table I. The better of the two algorithms is then compared to the standard BFGS and the best reported multi-step quasi-Newton algorithm in [10]. The results are presented in Table II. Eleven test functions were used, each with either one or two starting-points, giving a total of twenty problems. (Full details of the test functions and starting points may be found in [12].) Many of the functions employed in the tests may be used with varying dimensions. For variable dimension problems, the tests have been carried out on a range of suitably chosen dimensions that ensure the inclusion of different categories of problem sizes. The results are summed up since space precludes a tabulation of the individual figures for each dimension. The results reported are thus subtotals for each problem on a variety of dimensions and using two different starting points for each such dimension.

In all the methods considered here, the new point $\underline{x}_{i+1}$ was computed from $\underline{x}_i$ via a line-search algorithm which accepted the predicted point if the two standard stability conditions (Fletcher [5,6,7]) given below were satisfied and which, otherwise, used step-doubling and safeguarded cubic interpolation, as appropriate. A new iterate , x_{i+1} is accepted if it satisfies the following conditions [15]:-

$$f(x_{i+1}) \leq f(x_i) + 10^{-4} s_i^T g(x_i);$$
$$s_i^T g(x_{i+1}) \geq 0.9\{s_i^T g(x_i)\}.$$

It is easy to show (by analogy with standard theory for the BFGS method) that a necessary and sufficient condition for preserving positive-definiteness in the successive matrices $\{H_i\}$ is that $\underline{r}_i^T \underline{w}_i > 0$. In practice, we have imposed (in the implementations) the following requirement:-

$$r_i^T w_i > 10^{-4} \parallel r_i \parallel_2 \parallel w_i \parallel_2,$$

in order to ensure that $\underline{r}_i^T \underline{w}_i$ is "sufficiently" positive and thus avoid possible numerical instability in computing H_{i+1}. If this condition on $r_i^T w_i$ was not satisfied, the algorithm reverted to the choice $\theta_1 = \theta_2 = 0$. The initial inverse Hessian approximation was scaled using the scaling methods in [13,14].

The results of the numerical experiments presented in Tables I and II, show for each problem, the number of function/gradient evaluations required to solve the problem is given, followed (in brackets) by the number of iterations. A "† " is used to indicate the best score reported on a specific problem, for a given algorithm.

Table I. Comparison of Df1 and Df2

Problem	Df1	Df2
Watson (a)	463(446)	443(429)†
(b)	1992(1127)	1411(983)
Rosenbrock (a)	392(373)†	475(444)
(b)	444(422)†	672(634)
Ext. Powell (a)	345(134)	201(124)
(b)	320(105)	124(89)†
Penalty fn. (a)	773(344)	765(342)
(b)	505(384)†	577(450)
Trigonometric (a)	103(76)†	390(330)
(b)	153(139)†	733(668)
Broyden (a)	2291(1258)	1479(860)†
(b)	1814(1131)	1208(1050)
Wolfe (a)	398(281)	287(241)†
(b)	1941(639)†	2043(975)
Tridiagonal (a)	3214(2350)†	4918(1939)
(b)	2606(2247)†	3075(1974)
Powell (a)	1341(791)	780(667)†
(b)	2926(1618)	1735(1382)†
Sphere (a)	787(241)	155(136)†
(b)	256(181)†	532(391)
TOTALS	23064(14287)	22003(14108)

Table II. Comparison **of Df2 with BFGS and A1F**

Problem	**Df2**	**BFGS**	**A1F**
Watson (a)	443(429)	542(530)	443(429)
(b)	1411(983)	1653(1293)	1487(1101)
Rosenbrock (a)	475(444)	631(612)	485(465)
(b)	672(634)	825(806)	679(652)
Ext. Powell (a)	201(124)	159(122)	174(129)
(b)	124(89)	162(156)	108(97)
Penalty fn. (a)	765(342)	807(386)	765(344)
(b)	577(450)	637(507)	569(463)
Trigonometric (a)	390(330)	634(596)	549(485)
(b)	733(668)	2390(2331)	1973(1877)
Broyden (a)	1479(860)	2300(1786)	1912(1092)
(b)	1208(1050)	2136(2001)	1772(1607)
Wolfe (a)	287(241)	280(226)	266(225)
(b)	2043(975)	1967(965)	1936(917)
Tridiagonal (a)	4918(1939)	5099(2272)	4696(1892)
(b)	3075(1974)	2822(1969)	2765(1819)
Powell (a)	780(667)	1034(987)	735(680)
(b)	1735(1382)	1779(1559)	1623(1379)
Sphere (a)	155(136)	132(126)	130(124)
(b)	532(391)	1505(1262)	1050(718)
TOTALS	22003(14108)	27494(20492)	24117(16495)

V. CONCLUSION

A new model for the interpolating polynomial in multi-step methods is presented in this work. The model is used in constructing multi-step quasi-Newton methods which exploit iteration readily available data. The model includes two free parameters and it has been proven how those parameters may be determined by using available values of the objective function from the latest three iterates. It has been argued that much of the information computed at each version is renounced without making use of it [8]. Two algorithms are derived here that introduce a significant improvement, in numerical terms, over the standard (single-step) BFGS method and earlier successful quasi-Newton methods. The idea presented here is new in terms of the incorporation of free parameters that provide a tool to make use of any desired available data in the update of the Hessian (or its inverse) approximation to enhance the numerical performance of Secant-like methods.

In the new method for determining the parameterization of the interpolating curves in the two-step quasi-Newton methods, the parameters that influence the structure of the interpolating curve are obtained by minimizing a nonlinear equation at each iteration for algorithm *Df2*. The numerical results showed that such a price is justified by the significant improvement in numerical performance over the standard BFGS method. The method *Df2* improves slightly over *Df1*.

Future research might focus on issues like:

- Is there an optimal choice for the model and/or the curve parameters τ?
- Can these methods introduce similar improvements when applied to solving systems of non-linear equations?
- Further Study of the convergence properties of the methods.

References

[1] Al-Baali, M. "New property and global convergence of the Fletcher–Reeves method with inexact line searches", *IMAJ. Numer. Anal.*, 5, Feb. 1985, pp. 122–124.

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

[3] Byrd, R.H., Schnabel, R.B., and Shultz, G.A., "Parallel quasi-Newton methods for unconstrained optimization", *Math. Programing*, 42, April 1988, pp. 273-306.

[4] Dennis, J.E. and Schnabel, R.B., "Least change Secant updates for quasi-Newton methods", *SIAM Review*, 21, Feb. 1979, pp. 443-459.

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

[6] Fletcher, R., "A new approach to variable metric algorithms", *Comput. J.* , 13, April 1970, pp. 317-322.

[7] Fletcher, R. and Reeves, C., "Function minimization by conjugate gradients", *Computer J.*, 7, Jan. 1964, pp. 149–154.

[8] Ford, J.A. and Moghrabi I.A.R., "Using function-values in multi-step quasi-Newton methods", *J. Comput. Appl. Math*., 66, March 1996, pp. 201-211.

[9] Ford, J.A. and Moghrabi I.A.R., "Multi-step quasi-Newton methods for optimization", J. *Comput. Appl. Math*., 50, Jan. 1994, pp. 305-323.

[10] Ford, J.A. and Moghrabi I.A.R., "Alternative parameter choices for multi-step quasi-Newton methods", *Optimization Methods and Software*, 2, June 1993, pp. 357-370.

[11] Moghrabi, I.A.R., "Implicit extra-update multi-step quasi-newton methods", Int. J. Operational Research, 28, August 2017, pp. 69-81.

[12] Moré, J.J., Garbow, B.S., Hillstrom, K.E., "Testing unconstrained optimization software", *ACM Trans. Math. Softw*., **7**, April 1981, pp. 17–41.

[13] Shanno, D.F. and Phua, K.H., "Matrix conditioning and nonlinear optimization", *Math. Programming* ,14, May 1978, pp. 149-160.

[14] Shanno, D.F., "On the convergence of a new conjugate gradient algorithm", *SIAM J. Numer. Anal.*, 15, Jan. 1978, pp. 1247–1257.

[15] Wolfe, P., "Convergence conditions for ascent methods II: Some corrections", *SIAM Rev.*, 13, June 1971, pp. 185-188.

Issam A.R. Moghrabi is a Professor of M.I.S/C.S in the College of Business Administration at Gulf University for Science and Technology. He laid the foundations of the M.B.A Program at GUST while serving as the MBA Director for the past five years. He is a Fulbright Scholar and did a post-doctoral research in Geographic Information Systems. The author got his B.Sc. in Computer Science, with distinction, from the Lebanese American University, his M.Sc, with distinction, and Ph.D. from Essex University. His main research interests are in mathematical Optimization, Management Science and Information Retrieval and database systems [email: Moughrabi.i@gust.edu.kw].