
**International Journal of Operational Research**

ISSN online: 1745-7653 - ISSN print: 1745-7645
https://www.inderscience.com/ijor

# A new secant-like quasi-Newton method for unconstrained optimisation

Issam A.R. Moghrabi

**DOI:** 10.1504/IJOR.2021.10040598

**Article History:**

| | |
|---|---|
| Received: | 07 September 2020 |
| Accepted: | 28 March 2021 |
| Published online: | 12 January 2024 |

# A new secant-like quasi-Newton method for unconstrained optimisation

Issam A.R. Moghrabi

Accounting and MIS Department,
College of Business,
Gulf University for Science and Technology, Kuwait
Email: Moughrabi.i@gust.edu.kw

**Abstract:** The secant equation traditionally constitutes the basis of quasi-Newton methods, as the updated Hessian approximations satisfy the equation on each iteration. Modified versions of the secant relation have recently been the focus of several papers with encouraging outcomes. This paper continues with that idea where a secant-like modification that utilises nonlinear quantities in constructing the Hessian (or its inverse) approximation updates is derived. The technique takes advantage of data readily computed from the two most recent steps. Thus, it offers a substitute to the secant equation to produce better Hessian approximations that result in accelerated convergence to the objective function minimiser. The reported results provide adequate evidence to suggest that the proposed method is promising and deserves attention.

**Keywords:** quasi-Newton methods; secant-like methods; BFGS; unconstrained optimisation; multi-step methods.

**Reference** to this paper should be made as follows: Moghrabi, I.A.R. (2024) 'A new secant-like quasi-Newton method for unconstrained optimisation', *Int. J. Operational Research*, Vol. 49, No. 1, pp.65–84.

**Biographical notes:** Issam A.R. Moghrabi is a Professor of MIS/CS. He laid the foundations of the MBA at the GUST while serving as the MBA Director for the past five years. He is a Fulbright Scholar and obtained a post-doctoral research in Geographic Information Systems. His main research interests are in mathematical optimisation, management science and information retrieval and database systems. He serves as a referee for many well-known journals in his subjects of interest.

## 1 Introduction

The techniques considered in this paper are those used for solving unconstrained optimisation problems of the form:

$$\min f(x), \text{ where } f : R^n \rightarrow R.$$

quasi-Newton methods iteratively solve the above problem. For the current iteration $i$, the gradient of $f$ at $x_i$ is denoted as $g_i$, and the matrix $B_i$ is intended to approximate $G(x_i)$, the actual Hessian of $f$. The next approximation of the Hessian matrix is required to satisfy the so-called secant equation

$$B_{i+1}s_i = y_i, \tag{1}$$

for

$$y_i = g_{i+1} - g_i \text{ and } s_i = x_{i+1} - x_i.$$

The new iterate is defined as

$$x_{i+1} = x_i + \alpha_i p_i, \tag{2}$$

where $p_i$ the direction vector, obtained through solving

$$B_i p_i = -g_i \text{ or } p_i = -H_i g_i. \tag{3}$$

The matrix $H_i = B_i^{-1}$ is the inverse Hessian approximation that many prefer to maintain and update in order to avoid solving the system of equations resulting from using $B_i$ instead.

The step length $\alpha_i$ in equation (2) is computed to solve

$$\alpha_i = \min_{\alpha \in R^+} f\left(x_i + \alpha p_i\right),$$

employing some line search technique so as certain conditions are satisfied to guarantee convergence. The step length is generally measured in such a way that the following Powell-Wolfe conditions are met (Fletcher, 1970)

$$f\left(x_{i+1}\right) \leq f\left(x_i\right) + 10^{-4} s_u^T g_i \tag{4}$$

and

$$s_i^T g_{i+1} \geq 0.9 s_i^T g_i. \tag{5}$$

The quasi-Newton methods' basic concept is to copy Newton's method while requiring only the first partial derivatives of the function to be available. The derivation and then coding of the actual Hessian matrix is not required due to its susceptibility to human error and the demanding storage and evaluation requirements. The initial approximation to the Hessian is usually chosen as the identity matrix (or some scaled form) that is updated on a step-wise basis using the latest available step and gradient data. It has been proven that such methods display superlinear convergence given reasonable assumptions that need to hold on the objective functions (Dennis and Schnabel, 1979; Fletcher, 1970). The methods converge in at most n iterations on quadratic problems when the exact line searches are accurate (Broyden, 1970).

A well-known class of quasi-Newton methods is the Broyden family of updating formulae (Broyden, 1970; Fletcher, 1970). The most well-known members of the Broyden family of updates are the DFP, SR1, and the BFGS formulae. When compared numerically, the winner method of the three is the BFGS formula (Broyden, 1970; Fletcher, 1987). The BFGS update is given as

$$B_{i+1} = B_i + \frac{y_i y_i^T}{w_i^T r_i} - \frac{B_i s_i s_i^T B_i}{s_i^T B_i s_i}.$$

The global convergence of the BFGS update formula for convex objective functions has been established by several authors (see for example, Dai et al., 2002; Dennis and

Schnabel, 1979; Fletcher, 1987; Shanno and Phua, 1978; Shanno, 1978; Xiao et al., 2006; Yuan et al., 2017, 2018). Dai et al. (2002) show that when the line searches are inaccurate, the standard BFGS approach may not converge on non-convex functions (Wei et al., 2004).

The interestingness of the BFGS has been explored in a variety of applications and domains by many authors. The method has proven its viability as a serious contender in solving a variety of optimisation problems. BFGS has been regarded as an option in simulation optimisation studies (Yoon and Becker, 2020) to solve multi-objective simulation optimisation problems. The feasibility of using the BFGS formula to build a new local search heuristic rule useful in finding the nearest insert into the convex hull to solve the travelling salesman problem was proposed by Alipour and Razavi (2019). Another application where the method may prove useful in solving the unconstrained multiple-factor optimisation model (Cao et al., 2019). Woldu et al. (2020) utilise the BFGS in the development of a hybrid scaled conjugate gradient method that computes a search direction on which the step size is determined using a new modified non-monotone line search technique. Gondzio and Sobral (2019) research the effect of using quasi-Newton techniques in solving systems of nonlinear equations arising in interior point methods problems and applied to general quadratic programming. The cost of each iteration can be compared with the expense of calculating correctors in a conventional interior point iteration. Numerical studies stemming from applying the BFGS method show that the total number of matrix factorisations can be decreased. Mahdavi-Amiri and Sadaghiani (2020) propose and evaluate a quasi-Newton non-monotonous algorithm for highly convex multi-objective unconstrained optimisation. The algorithm enables a convex combination of recent function values to be reduced, all implemented within the BFGS quasi-Newton method of solving unconstrained problems of multi-objective optimisation. The numerical results mentioned indicate that fewer function evaluations are used by the quasi-Newton non-monotone algorithm than by the quasi-Newton monotone algorithm and other approaches.

## 2 Variants of the secant equation

This paper focuses on developing new quasi-Newton-like methods that derive from a variant of the secant equation (1). Much research has been developed to derive methods that numerically outperform the classical BFGS update and are globally convergent under reasonable assumptions. Such methods have proven, numerically, to be serious contenders to the traditional secant methods evaluated in terms of function/gradient evaluations in addition to the iteration count. There have been several approaches to deriving such methods originating from different perspectives, thus ending with different formulations (see Table 1). The success of these methods constitutes enough motivation for further pursuit on the derivation of similar techniques. Our approach relies on utilising more of the readily computed quantities into the updated matrix in the hope of developing the quality of the Hessian (or its inverse) approximation on each iteration.

This section gives a brief account of some of the methods that have introduced performance gains over the standard BFGS. The derivations of these algorithms are mainly based on motivating modifications of the secant equation (1). One particularly successful class of methods published in the literature is motivated by the idea of incorporating more of the data available at each iteration in the update of the Hessian approximation that would otherwise be discarded. Such data include, but are not limited to, the step vectors and the gradient difference vectors in (1) collected from the m recent iterations ($m > 1$) rather than just the most recent step. An example of such methods is the multi-step methods (Ford and Moghrabi, 1993, 1994, 1996). Those are introduced next.

If $X = \{x(\tau)\}$ represents a differentiable path in $R^n$, for $\tau \in R$, then, if the chain rule is applied to $g(x(\tau))$ to differentiate it with respect to $\tau$, one gets

$$G(x_{i+1})x'(\tau_m) = g'(x(\tau_m)). \tag{6}$$

In particular, if the path $X$ is chosen to encompass the most recent point $x_{i+1}$ [so that $x(\tau_m) = x_{i+1}$, say], then equation (6) defines a relationship (referred to as the 'Newton equation') that the Hessian matrix $G(x_{i+1})$ satisfies (see Al-Baali, 1985; Broyden, 1970; Dennis and Schnabel, 1979). The secant equation can be obtained from the so-called Newton Equation as a special case (Broyden, 1970). In Ford and Moghrabi (1994), $X$ is taken to be the polynomial that interpolates the $m + 1$ newest points $\{\underline{x}_{j-m+k+1}\}_{k=0}^{m}$. The vector $g'(x(\tau_m))$ is approximated by the vector polynomial differentiation ($\hat{g}(\tau)$, say) which interpolates the corresponding available gradient points $\{g(x_{i-m+k+1})\}_{k=0}^{m}$.

The scalar values $\{\tau_k\}_{k=0}^{m}$ have an association with the points $\{x_{i-m+k+1}\}_{k=0}^{m}$ on the curve $X = \{x(\tau)\}$:

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

Thus, if $B_{i+1}$ denotes some approximation to $G(x_{i+1})$ in equation (5) and

$$r_i \stackrel{\text{def}}{=} x'(\tau_m) = \sum_{j=0}^{m-1} \underline{s}_{i-j} \left\{ \sum_{k=m-j}^{m} \mathcal{L}_k'(\tau_m) \right\}; \tag{7}$$

$$w_i \stackrel{\text{def}}{=} \hat{g}'(\tau_m) = \sum_{j=0}^{m-1} y_{i-j} \left\{ \sum_{k=m-j}^{m} \mathcal{L}_k'(\tau_m) \right\} \approx g'(x(\tau_m)) \tag{8}$$

where $s_i$ and $y_i$ are as in equation (1) and $w_i$ provides an estimate to $g'(x(\tau_m))$. It is thus reasonable [by equation (6)] to require that the matrix $B_{i+1}$ satisfy a condition similar to the one in equation (6) as follows

$$B_{i+1} r_i = w_i. \tag{9}$$

One choice that was considered for the $\tau$ parameters [for $m = 2$ in equations (7) and (8)] is (Ford and Moghrabi, 1993)

$$\tau_0 = -\left(\|s_i\|_2 + \|s_{i-1}\|_2\right), \tau_2 = 0, \text{ and } \tau_1 = -\|s_i\|_2. \tag{10}$$

This choice is sensitive to the distances among the iterates in the space of the variables.

The multi-step BFGS Hessian update is given by

$$B_{i+1}^{MS} = B_i + \frac{w_i w_i^T}{w_i^T r_i} - \frac{B_i r_i r_i^T B_i}{r_i^T B_i r_i}. \tag{11}$$

The main merit of equation (11) is in its utilisation of several past step and gradient vectors rather than just using the latest single latest iteration vectors.

Some of the recent and well-known secant-like methods are motivated by the need to derive accelerated convergence variants of the classical quasi-Newton methods. Other modified secant equation methods are motivated by the concept that the classical secant relation uses only the most recent single gradient and step vectors in the updating process of the Hessian (or its inverse) approximation. In contrast, other readily computed data, such as function values, remain unexploited and that might, otherwise, prove valuable. In an attempt to better the 'quality' of the Hessian approximation matrix, several papers have considered variants of the classical secant relation (1) to utilise computed gradient, iteration difference vectors as well as the readily available function evaluations (see for instance, Ford and Moghrabi, 1996; Wei et al., 2004; Ortiz et al., 2019; Yuan et al., 2010, 2017, 2018; Yuan and Wei, 2010; Zhang et al., 1999). For example, Wei et al. (2006), through using Taylor's series, have introduced a modification to equation (1) as follows:

$$B_{i+1}s_i = y_i^*,$$

where $y_i^* = y_i + \dfrac{\theta_i}{\|s_i\|^2}$ and $\theta_i = 2(f_i - f_{i+1}) + (g_{i+1} + g_i)^T s_i$.

In a similar venue, Yuan and Wei (2010) derive an alternative replacement to $y_i^*$ given as:

$$y_i^* = y_i + \frac{\max(0, \theta_i)}{\|s_i\|^2}.$$

Although, in addition to the function values obtained from the latest iterate, such secant-like relationships use step and gradient vectors, the methods derived in Moghrabi (2017) use nonlinear interpolating polynomials and implement an entirely different secant-like equation. That secant relation utilises function values and other available data from the three most recent steps. To satisfy several secant-like conditions, they introduce the concept of making many updates on each iteration (see Al-Baali, 1985; Moghrabi, 2017). The updates are carried out in such a way that the first update observes (1) while subsequent updates satisfy

$$B_{i+1}^{(t+1)}u^t = v^t, t = 1, 2, m-1\ (m > 1)$$

where $u^t = r_{i-m+t}$ and $v^t = w_{i-m+t}$, for $r$ and $w$ as in equations (6) and (7).

Woldu et al. (2020) have developed a Hessian matrix that satisfies a modified secant condition of the type

$$B_{i+1}s_i = w_i,$$

where

$$w_i = y_i + \max(0, t_i)s_i$$

for

$$t_i = \frac{6(f_i - f_{i+1}) + 3\big(g(x_i + \alpha_i d_i, \varepsilon) + g(x_i, \varepsilon)\big)^T s_i}{s_i^T s_i} s_i$$

and for some chosen extra line search parameter $\varepsilon$ and integrated that matrix in the computation of scaled conjugate gradient search directions. The method has interesting convergence properties but with moderate practical impact.

Another secant-like equation is due to Deghani and Hosseini (2019), which they exploit to derive a modified BFGS update. The method is obtained using the Taylor series for the objective function $f(x)$ and utilises both the available function and gradient values. The method possesses better theoretical convergence features than the classical secant equation. The modified secant update is given by

$$B_{i+1}s_i = y_i + \frac{6(f_i - f_{i+1}) + 3(g(x_i + \alpha_i d_i) + g(x_i))^T s_i}{s_i^T s_i}.$$

Nakayama et al. (2019) propose a spectral-scaling secant condition used to develop a new memory less quasi-Newton method. The new method is based on the Broyden family. Both convex and pre-convex members of the family are taken into account. The new method is shown to satisfy sufficient descent conditions and possesses global convergence. The modified secant takes the form

$$B_{i+1}s_i = \gamma_i y_i,$$

where $\gamma_i > 0$ is some scaling parameter.

Similarly, Waziri et al. (2020) propose the following modified secant relation

$$B_{i+1}s_i = y_i + \left( \frac{6(f_i - f_{i+1}) + 3s_i^T (g_{i+1} + g_i)}{s_i^T u_i} \right) u_i$$

where $u_i$ is a vector chosen such that $s_i^T u_i \neq 0.$ This relationship is used as a basis to establish a new conjugate gradient methodology for solving nonlinear equation systems.

Deghani et al. (2019) employ Taylor's series to develop an updated secant relationship to get a more detailed estimate of the objective function's second curvature. The modified secant equation is incorporated in building a new BFGS form employed in solving unconstrained optimisation problems. Both gradient and function values are used in the proposed process while noting that the classical secant relationship uses only gradient values. Global convergence is established under suitable conditions and without requiring convexity of the function to be minimised. The proposed secant-like relation takes the form

$$B_{i+1}s_i = y_i + \delta_i \left( \frac{12(f_i - f_{i+1}) + s_i^T (5g_{i+1} + 7g_i) - \alpha_i^2 s_i^T g_i}{s_i^T s_i} \right) s_i,$$

for some chosen integer such that $\delta_i \geq 0$.

Faramarzi and Amini (2020) developed a modified conjugate gradient technique derived from a suggested modified secant relation. The new algorithm satisfies, independent of line search, the sufficient descent property. For both uniformly convex and general functions, the convergence properties of the proposed algorithm were studied. Numerical experiments demonstrate the suggested approach's superiority compared to several such methods in the same class. The secant-like relationship utilised is as follows:

$$B_{i+1}s_i = y_i + \left( \frac{\max\left(0, 6\left(f_i - f_{i+1}\right) + 3s_i^T\left(g_{i+1} + g_i\right)\right)}{s_i^T u_i} \right) u_i$$

where $u_i$ is a vector chosen such that $s_i^T u_i \neq 0$.

**Table 1** Some modifications of the secant equation

| *Author(s)* | *The derived secant variant* | *Ref.* |
|---|---|---|
| Powell | $B_{i+1}s_i = \vartheta_i y_i + (1 - \vartheta_i)B_i s_i$ | Fletcher (1987) |
| Zhang et al. | $B_{i+1}s_i = y_i + \frac{6(f_i - f_{i+1}) + 3(g_{i+1} + g_i)^T s_i}{s_i^T s_i} s_i$ | Zhang et al. (1999) |
| Wei et al. | $B_{i+1}s_i = y_i + \frac{2(f_i - f_{i+1}) + (g_{i+1} + g_i)^T s_i}{s_i^T s_i} s_i$ | Wei et al. (2006) |
| Li and Fukushima | $B_{i+1}s_i = y_i + \sigma_i s_i$, $\sigma_i < 10^{-6}$ | Li and Fukushima (2001) |
| Waziri et al. | $B_{i+1}s_i = y_i + \left( \frac{6(f_i - f_{i+1}) + 3s_i^T(g_{i+1} + g_i)}{s_i^T u_i} \right) u_i$ where $u_i$ is a vector chosen such that $s_i^T u_i \neq 0$. | Waziri et al. (2020) |
| Yuan et al. | $B_{i+1}s_i = y_i + \frac{\max(0, 2(f_i - f_{i+1}) + (g_{i+1} + g_i)^T s_i)}{s_i^T s_i} s_i$ | Yuan et al. (2017) |
| Deghani and Hosseini | $B_{i+1}s_i = y_i + \frac{6(f_i - f_{i+1}) + 3(g(x_i + \alpha_i d_i) + g(x_i))^T s_i}{s_i^T s_i}$ | Deghani and Hosseini (2019) |
| Woldu et al. | $B_{i+1}s_i = y_i + \max(0, t_i)s_i$, for $t_i = \frac{6(f_i - f_{i+1}) + 3(g(x_i + \alpha_i d_i, \varepsilon) + g(x_i, \varepsilon))^T s_i}{s_i^T s_i}$ for some chosen extra line search parameter $\varepsilon$. | Woldu et al. (2020) |
| Deghani et al. | $B_{i+1}s_i = y_i + \delta_i \left( \frac{12(f_i - f_{i+1}) + s_i^T(5g_{i+1} + 7g_i) - \alpha_i^2 s_i^T g_i}{s_i^T s_i} \right) s_i,$ for some chosen integer such that $\delta_i \geq 0$. | Deghani et al. (2019) |
| Caliciotti et al. | $H_{i+1}y_j = \rho_j s_j$ with $\begin{Bmatrix} \rho_j > 0, & \text{for } j < i, \\ \rho_j = 1, & \text{for } j = i. \end{Bmatrix}$ | Caliciotti et al. (2018) |

Caliciotti et al. (2018) have developed a novel preconditioned conjugate gradient method that utilises a symmetric rank one update that satisfies a weaker notion of the secant equation. They argue that the quasi-Newton update $H_{i+1}$, which in some way attempts to approximate the inverse Hessian matrix $G(x)^{-1}$, satisfies the following modified secant equation in all previous directions; that is, it results in the following modified secant equation;

$$H_{i+1}y_j = \rho_j s_j \text{ with } \begin{Bmatrix} \rho_j > 0, & \text{for } j < i, \\ \rho_j = 1, & \text{for } j = i. \end{Bmatrix}$$

Although Table 1 provides a brief account of the work done on developing modified secant equations, the discussion above provides more detail on the topic. Many other similar methods can be found in the literature for interested readers (Wei et al., 2006; Xiao et al., 2006; Yuan and Wei, 2010; Yuan et al., 2017, 2018, 2010; Zhang et al., 1999; Deghani et al., 2019; Das et al., 2019; Yoon and Bekker, 2020; Alipour and Razavi, 2019; Bolouri et al., 2020). The convergence properties for several of the methods cited here have been studied. For example, Yuan et al. (2018) prove global convergence under a less strict version of the Powell-Wolfe line search conditions (4) and (5).

A summary of selected methods is listed in Table 1.

We next examine a new method that is motivated by the same concept of incorporating data available from several of the latest iterations for deriving a new variant to the secant equation (1).

## 3 A new non-secant equation

This section focuses on the development of a new version of the secant equation (1) of the quasi-Newton type, hoping that this results in further numerical merits that methods in this class achieve over the classical BFGS update. We then need to prove that the new method is globally convergent under rational assumptions. It is believed that the success of such methods lies in the improved quality of the Hessian approximations they achieve at each step of the algorithm. There have been many ways to deriving these techniques from multiple viewpoints, thereby concluding with various formulations (see Table 1). Similar to what has been done in the above discussion on the nonlinear multi-step quasi-Newton methods [equations (5)–(10)], our derivation of the new non-secant relationship will also make use of the data computed during the three most recent cycles, thus choosing $m = 2$ in equation (6). The new secant equation will define the new condition that the updated Hessian (or its inverse) must satisfy. In particular, the iterates $x_{i-1}$, $x_i$ and $x_{i+1}$ are interpolated by a differentiable curve in $R^n$, namely $x(\tau)$, such that $x(\tau_0) = x_{i-1}$, $x(\tau_1) = x_i$ and $x(\tau_2) = x_{i+1}$. The corresponding objective function is given as $\varphi(\tau) \equiv f(x(\tau))$ and is modelled here using Taylor's expansion relation around the point $\tau_2$, corresponding to the most recent iterate $x_{i+1}$ as follows

$$\varphi(\tau) \cong \varphi(0) + \tau\varphi'(0) + \frac{1}{2}\tau^2\varphi''(0), \tag{12}$$

for $\tau$-values as in equation (10).

Using $\varphi(\tau) \equiv f(x(\tau))$, then the quantities in equation (12) are expressed as

$$\begin{aligned} &\varphi'(0) \equiv x'(0)^T g_{i+1}, \\ &\varphi''(0) \equiv x'(0)^T B_{i+1} x'(0) + x''(0)^T g_{i+1}, \end{aligned} \tag{13}$$

where $x'(\tau) \equiv \dfrac{dx}{d\tau}$ and $x''(\tau) \equiv \dfrac{d^2x}{d\tau^2}$.

Using the Lagrange representation to interpolate the iterates available from the previous three iterations, we have [obtained by setting $m = 2$ in equation (7)]

$$x'(\tau) = \mathcal{L}_2'(\tau)\left(s_i - \vartheta s_{i-1}\right) \tag{14}$$

and

$$x''(\tau) = \mathcal{L}_2''(\tau)\left(s_i - \delta s_{i-1}\right), \tag{15}$$

for

$$\vartheta = \frac{\delta^2}{1+2\delta}, \delta = \frac{\tau_2 - \tau_1}{\tau_1 - \tau_0},$$

and, for all $\tau$, the following quantities hold

$$\mathcal{L}_2'(\tau) = \frac{2-\tau-\tau_0-\tau_1}{\left(\tau_2-\tau_1\right)\left(\tau_2-\tau_0\right)}, \mathcal{L}_0'(\tau) = \frac{2\tau-\tau_1-\tau_2}{\left(\tau_0-\tau_1\right)\left(\tau_0-\tau_2\right)}, \tag{16}$$

$$\mathcal{L}_2''(\tau) = \frac{2}{\left(\tau_2-\tau_1\right)\left(\tau_2-\tau_0\right)}, \mathcal{L}_0''(\tau) = \frac{2}{\left(\tau_0-\tau_1\right)\left(\tau_0-\tau_2\right)}. \tag{17}$$

For the chosen $\tau$-values in equation (10), we obtain $\delta = \frac{\|s_i\|}{\|s_{i-1}\|}$, It is reasonable then that expression for $\delta$ is generalised by plugging in a scaling factor, $\gamma \geq 0$ (see Moghrabi, 2017) that gives a more straightforward, more convenient mechanism to switch to the standard one-step secant update method by setting $\gamma = 0$. Therefore,

$$\delta = \gamma \frac{\|s_i\|}{\|s_{i-1}\|}.$$

If $\tau$ in equation (12) is chosen to be $\tau_0$ and if the actual Hessian at $x_{i+1}$ is replaced by its approximation $B_{i+1}$, as is normally done in the standard quasi-Newton methods, one may require that

$$\tau_0^2 \varphi''(0) \cong 2\left[f_{i-1} - f_{i+1} - \tau_0 \varphi'(0)\right], \tag{18}$$

or equivalently,

$$f_{i-1} = f_{i+1} + \tau_0 x'(0)^T g_{i+1} + \frac{1}{2}\tau_0^2\left[x'(0)^T B_{i+1} x'(0) + x''(0)^T g_{i+1}\right], \tag{19}$$

for $x'(0)$ and $x''(0)$ as defined in equations (14) and (15), respectively. This gives

$$f_{i-1} = f_{i+1} + \tau_0\left[\frac{-\tau_0-\tau_1}{\tau_1\tau_0} s_i^T g_{i+1} + \frac{\tau_1}{\left(\tau_0-\tau_1\right)\tau_0}\right]$$
$$+\frac{1}{2}\tau_0^2\left[x'(0)^T B_{i+1} x'(0) + \left(\frac{2}{\tau_1\tau_0} s_i - \frac{2}{\left(\tau_0-\tau_1\right)\left(\tau_0\right)} s_{i-1}\right)^T g_{i+1}\right].$$

Now, from equations (13) and (19), we obtain

$$\begin{aligned} &\tau_0^2\left[\left(\frac{\left(-\tau_0-\tau_1\right)}{\tau_1\tau_0}\right)^2 \left(s_i - \vartheta s_{i-1}\right)^T B_{i+1}\left(s_i - \vartheta s_{i-1}\right) + \frac{2}{\tau_1\tau_0}\left(s_i - \delta s_{i-1}\right)^T g_{i+1}\right] \\ &= 2\left(f_{i-1} - f_{i+1} - \frac{-\tau_0-\tau_1}{\tau_1}\left(s_i - \vartheta s_{i-1}\right)^T g_{i+1}\right), \end{aligned} \tag{20}$$

for $\vartheta$ and $\delta$ are as defined in equations (14) and (15), respectively.

If we define

$$r_i \equiv s_i - \vartheta s_{i-1}\overline{r_i} \equiv s_i - \delta s_{i-1} \text{ and } \rho \equiv \frac{\tau_0 + \tau_1}{\tau_1},$$

then from equation (19) we have

$$\rho^2 r_i^T B_{i+1} r_i = 2\left[ f_{i-1} - f_{i+1} + \rho r_i^T g_{i+1} - \frac{\tau_0}{\tau_1} g_{i+1}^T \overline{r_i} \right]. \tag{21}$$

Equation (21) may be expressed as

$$u_i^T B_{i+1} u_i = u_u^T w_i + \mu_i \tag{22}$$

for

$$u_i \equiv \rho r_i,\ w_i \equiv y_i - \vartheta y_{i-1}$$

and

$$\mu_i = 2\left( f_{i-1} - f_{i+1} + u_i^T g_{i+1} - \frac{\tau_0}{\tau_1} g_{i+1}^T \overline{r_i} \right) - u_i^T w_i.$$

The relationship in equation (22) lends itself to proposing a new secant-variant equation of the form

$$B_{i+1} u_i = v_i \tag{23}$$

for $u_i$ as in equation (21) and $v_i \equiv w_i + \frac{\mu_i}{\|u_i\|^2} u_i$.

The computed search direction is downhill if $B_{i+1}$ is positive definite. By analogy with the standard secant equation (1), $B_{i+1}$ is positive definite if and only if $B_i$ is already so and $u_i^T v_i > 0$. As this cannot be guaranteed in this new formulation, equation (23) is replaced with

$$B_{i+1} u_i = \overline{v_i} \tag{24}$$

where

$$\overline{v_i} = v_i + \varepsilon_i u_i,\ \varepsilon_i = \gamma \|g_i\|^2 + \max\left( -\frac{u_i^T v_i}{\|u_i\|^2}, 0 \right),$$

for some positive constant $\gamma$. It is easy to show that relation (24) is an appropriate replacement since

$$\overline{v_i}^T u_i \geq \gamma \|g_i\|^2 \|u_i\|^2 > 0, \tag{25}$$

provides a guarantee to the positive definiteness of $B_{i+1}$.

The new BFGS algorithmic outline goes as follows:

*Algorithm NBFGS*

Input: $x_0 \in R^n$, and set $H_o = I$. Let iteration count $i = 0$.

Output: optimal solution

1 Stop if $\|g_i\| \leq \varepsilon$ (convergence threshold).

2 Compute $d_i = -H_i g_i$.

3 Minimise $f(x_i + \alpha p_i)$, to compute $\alpha_i$ such that conditions (5) are $\alpha \in R$ satisfied.

4 Let $x_{i+1} = x_i + \alpha_i p_i$. If equation (25) is satisfied, update $H_i$ using equation (10) with $w_i$ replaced by $\overline{v}_i$ and $r_i$ replaced by $u_i$ in equation (25), else let set $u_i = s_i$ and $\overline{v}_i = y_i$ in equation (25) and update such that equation (1) is satisfied.

5 Set $i = i + 1$ and go to 1.

## 4 Convergence properties

The analysis carried out here relies on the assumptions below:

1 The level set $D = \{x | f(x) < f(x_0)\}$ is bounded, for a starting point $x_0$.

2 The objective function $f$ is twice continuously differentiable on $D$ and in an open set $M$ containing $D$, there exists a constant $\mathfrak{z} > 0$ such that

$$\|g(x) - g(y)\| \leq \mathfrak{z}\|x - y\|, \text{ for all } x, y \in M. \tag{26}$$

Since $\{f_i\}$ is a diminishing sequence, the iterate sequence $\{x_i\}$ computed by the new algorithm is found in $D$, and there exists a constant $f^*$ such that:

$$\lim_{i \to \infty} f_i = f^*. \tag{27}$$

3 The objective function $f$ is uniformly convex, in that there are positive constants $m_1$ and $m_2$ such that

$$m_1 \|p\|^2 \leq d^T \nabla^2 f(x) d \leq m_2 \|p\|^2$$

holds for all $x \in D$ and $p \in R^n$.

*Theorem 1:* Let $\{x_i\}$ be generated by the algorithm NBFGS. Then we have

$$m_1 \|s_i\|^2 \leq s_i^T v_i \leq m_2 \|s_i\|^2,$$

$$\|u_i\| \leq (m_2 + \mathfrak{z})\|s_i\|$$

and

$$\sum_{i=1}^{\infty} -\alpha_i g_i^T p_i = \pi r^2.$$

*Proof:* Similar to the proof done in Hassan (2019).

*Theorem 2:* Let $f$ satisfy Assumptions 1 and 2 and $\{x_i\}$ be computed by algorithm NBFGS. Also, there are constants $k_1$ and $k_2$ such that

$$\|B_i s_i\| \le k_1 \|s_i\| \text{ and } s_i^T B_i s_i \ge k_2 \|s_i\|^2, \forall i. \tag{27}$$

Then the following holds

$$\liminf_{i\to\infty} \|g_i\| = 0. \tag{28}$$

*Proof:* By contradiction, assume that for small, non-negative constant $\varepsilon$, $\|g_i\| \ge \varepsilon$. Then, since $B_i s_i = -\alpha_i g_i$, $\|g_i\|^2 = \alpha_i^{-2} \|B_i s_i\|$. Thus, from Theorem 1 it follows that

$$\begin{aligned}
&\infty > \sum_{i=1}^{\infty} -\left(p_i^T g_i\right) \\
&\le \sum_{\forall i} \alpha_i^{-1}\left(s_i^T B_i s_i\right) \\
&= \sum_{\forall i} \alpha_i \frac{\|g_i\|^2}{\|B_i s_i\|^2}\left(s_i^T B_i s_i\right) \ge \sum_{\forall i} \tilde{\alpha}_i \varepsilon^2 k_2 k_1^{-2} = +\infty,
\end{aligned}$$

where $\tilde{\alpha}_i = \dfrac{(1-\sigma)m}{L}$, for $\sigma$ as in Assumption 3 and $\mathfrak{z}$ as in Assumption 2. The following inequality follows from assumption two and the Powell-Wolfe condition (5)

$$-(1-\sigma) g_i^T p_i \le \left(g_{i+1} - g_i\right)^T p_i \le \mathfrak{z}\alpha_i \|p_i\|^2$$

$$\alpha_i \ge \frac{-(1-\sigma) g_i^T p_i}{L\|p_i\|^2} = \frac{-(1-\sigma) p_i^T B_i p_i}{L\|p_i\|^2} \ge \frac{-(1-\sigma)m}{L} = \tilde{\alpha}_i.$$

We now proceed to prove that algorithm NBFGS has global convergence.

*Theorem 3:* Let $f$ satisfy Assumptions 1 and 2 above and $\{x_i\}$ generated by algorithm NBFGS. Then, the following holds

$$\liminf_{i\to\infty} \|g_i\| = 0. \tag{29}$$

*Proof:* As per Theorem 2, it suffices to prove that equation (29) holds for all $i$. By contradiction, assume this is not the case. Thus, there exists a positive constant $\mathfrak{z}$ such that

$$\|g_i\| > \mathfrak{z}, \forall i.$$

It is easy to show that

$$\|\overline{v}_i\| \le \theta \|s_i\|, \text{ for some positive constant } \theta. \tag{30}$$

From equation (25) and the fact that

$$\|u_i\| \ge \sigma\left(\|s_i\| - \varepsilon\|s_{i-1}\|\right) \ge \frac{1}{2}\sigma\|s_i\|,$$

it follows that

$$u_i^T \overline{v}_i \geq \gamma \mathfrak{z}^2 u_i^T u_i \geq \frac{\gamma}{2} \sigma \mathfrak{z}^2 s_i^T s_i. \tag{31}$$

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

$$\frac{\overline{v}_i^T \overline{v}_i}{u_i^T \overline{v}_i} \leq \rho,$$

for $\rho \equiv \dfrac{\theta^2}{\gamma \mathfrak{z}^2 \sigma / 2}$.

By Theorem 2, for the sequence $\{B_i\}$, there exist constants $k_1$ and $k_2$ such that equation (29) holds for all $i$. The proof is complete, based on the following [see Theorem 2.1 in Dai et al. (2002)].

If there are two positive constants $k_1$ and $k_2$ such that for all $i$

$$\frac{u_i^T \overline{v}_i}{u_i^T u_i} \geq k_1 \text{ and } \frac{\overline{v}_i^T \overline{v}_i}{u_i^T \overline{v}_i} \geq k_2,$$

such that for all positive integers $j$ (28) holds for no less than $[j/2]$ iterations of $i \in \{1, \ldots, j\}$.

## 5 Numerical tests and results

In this section, summaries of the numerical outcomes are tabulated. The test problems list is listed in Table 2. The chosen test set is for problems of variable dimensions, so that different test problem sizes are examined. The tested functions, and hence the presented results, correspond to problems classified based on the size/dimension category. Those are grouped into four categories, low ($2 \leq n \leq 20$), medium ($21 \leq n \leq 40$), moderately high ($41 \leq n \leq 1,000$), and high ($n > 1,000$). The results constitute the outcome of experiments conducted on 17 different test problems with dimensions varying from 2 to 100,000. Each listed problem has been tested using four distinct starting points. The total number of test problems obtained is 900. The test problems are extracted from Moré et al. (1981), Fletcher (1987), Tajadod et al. (2016) and Xiao et al. (2006).

The new algorithm NMBFGS is tested against that of Yuan et al. (2017) (see Table 1) and the multi-step BFGS (MSBFGS) in equation (8) (see Ford and Moghrabi, 1993, 1994, 1996). Yuan's method is used as the benchmark for the tests since it is, as per our tests, the most competitive of the techniques presented in Table 1. The summary of the overall results is reported in Table 3. Tables 4 to 7 present the scores obtained for small, medium, moderately large, and very large problems, respectively. As space precludes a detailed presentation of the figures obtained for each category separately, the results are summarised per dimension category. The scores reported indicate iteration, function/gradient evaluations count as well as the total execution times. The coding is done using C++ on a 64-bit machine with i7-3770, 3.4 GHz CPU.

**Table 2** Test problems and dimensions

| *ID* | *Problem* | *Dimension range* |
|---|---|---|
| 1 | Watson function | $3 \leq n \leq 31$ |
| 2 | Extended Rosenbrock | $2 \leq n \leq 10{,}000$, $n$ even |
| 3 | Extended Powell | $2 \leq n \leq 100{,}000$, $n$ mod 4 = 0 |
| 4 | Penalty function I | $2 \leq n \leq 1{,}000$ |
| 5 | Variably dimensioned function | $2 \leq n \leq 10{,}000$ |
| 6 | Trigonometric function | $2 \leq n \leq 10{,}000$ |
| 7 | Modified trigonometric function | $2 \leq n \leq 100{,}000$ |
| 8 | Broyden tridiagonal function | $2 \leq n \leq 1{,}000$ |
| 9 | Discrete boundary value function | $2 \leq n \leq 1{,}000$ |
| 10 | Oren and Spedicato power function | $2 \leq n \leq 10{,}000$ |
| 11 | Full set of distinct eigenvalues problem | $2 \leq n \leq 10{,}000$ |
| 12 | Tridiagonal function | $2 \leq n \leq 10{,}000$ |
| 13 | Wolfe function | $2 \leq n \leq 1{,}000$ |
| 14 | Diagonal Rosenbrock's function | $2 \leq n \leq 1{,}000$, $n$ even |
| 15 | Generalised shallow function | $2 \leq n \leq 1{,}000$, $n$ even |
| 16 | Powell singular | $n = 10{,}000$ |
| 17 | Helical valley | $n = 1{,}000$ |

**Table 3** Overall iteration, function evaluations count, and timing

| *Method* | *Evaluations* | *Iterations* | *Time (sec.)* | *Scores* |
|---|---|---|---|---|
| Yuan et al. | 50,456 | 40,492 | 24,026.44 | 187 |
| | 100% | 100% | 100% | 20.8% |
| MSBFGS | 45,084 | 34,294 | 23,212.543 | 316 |
| | 89.35% | 84.69% | 96.61% | 35.1% |
| *NMBFGS* | *41,313* | *30,419* | *21,434.622* | *397* |
| | *81.88%* | *75.12%* | *89.21%* | *44.1%* |

**Table 4** Iteration, function evaluations count, and timing-large problems

| *Method* | *Evaluations* | *Iterations* | *Time (sec.)* | *Scores* |
|---|---|---|---|---|
| Yuan et al. | 21,773 | 18,812 | 370.8 | 32 |
| | 100% | 100% | 100% | 23.6% |
| MSBFGS | 17,944 | 14,101 | 246.703 | 51 |
| | 82.41% | 74.96% | 66.53% | 33.9% |
| *NMBFGS* | *16,211* | *13,237* | *229.412* | *71* |
| | *74.45%* | *70.36%* | *61.87%* | *42.5%* |

For all the methods tested here, the computed iterate $x_{i+1}$ is computed from $x_i$ by employing a line search algorithm that applies safeguarded cubic interpolation with step-doubling (Yoon and Bekker, 2020). Once the conditions in equations (4) and (5) are satisfied, a new point, $x_{i+1}$, is accepted.

**Table 5** Iteration, function evaluations count and timing-moderately large problems

| *Method* | *Evaluations* | *Iterations* | *Time (sec.)* | *Scores* |
|---|---|---|---|---|
| Yuan et al. | 11,166 | 8,411 | 2,911.78 | 90 |
| | 100% | 100% | 100% | 23.6% |
| MSBFGS | 9,757 | 7,007 | 2,431.59 | 129 |
| | 87.38% | 83.31% | 83.51% | 33.9% |
| *NMBFGS* | *9,213* | *6,891* | *2,277* | *162* |
| | *82.51%* | *81.93%* | *78.20%* | *42.5%* |

**Table 6** Iteration, function evaluations count and timing-medium problems

| *Method* | *Evaluations* | *Iterations* | *Time (sec.)* | *Scores* |
|---|---|---|---|---|
| Yuan et al. | 10,213 | 9,618 | 16,213.66 | 31 |
| | 100% | 100% | 100% | 12% |
| MSBFGS | 9,172 | 8,879 | 11,401.19 | 99 |
| | 89.81% | 92.32% | 70.32% | 39% |
| *NMBFGS* | *7,811* | *6,001* | *10,001.13* | *124* |
| | *76.48%* | *62.39%* | *61.68%* | *49%* |

**Table 7** Iteration, function evaluations count and timing-small problems

| *Method* | *Evaluations* | *Iterations* | *Time (sec.)* | *Scores* |
|---|---|---|---|---|
| Yuan et al. | 8,345 | 4,390 | 9,342.67 | 34 |
| | 100% | 100% | 100% | 30.6% |
| MSBFGS | 8,211 | 4,307 | 9,133.06 | 37 |
| | 98.39% | 98.11% | 97.76% | 33.3 |
| *NMBFGS* | *8,078* | *4,290* | *8,927.08* | *40* |
| | *96.80%* | *97.72%* | *95.55%* | *36%* |

**Figure 1** Overall evaluations (see online version for colours)

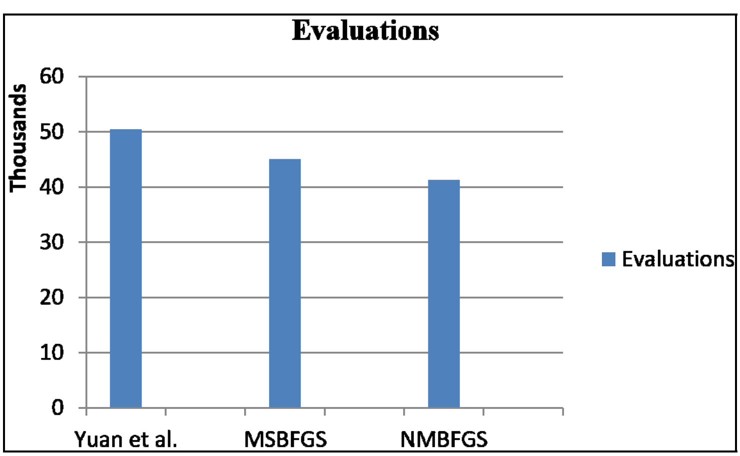


**Figure 2** Overall iterations (see online version for colours)

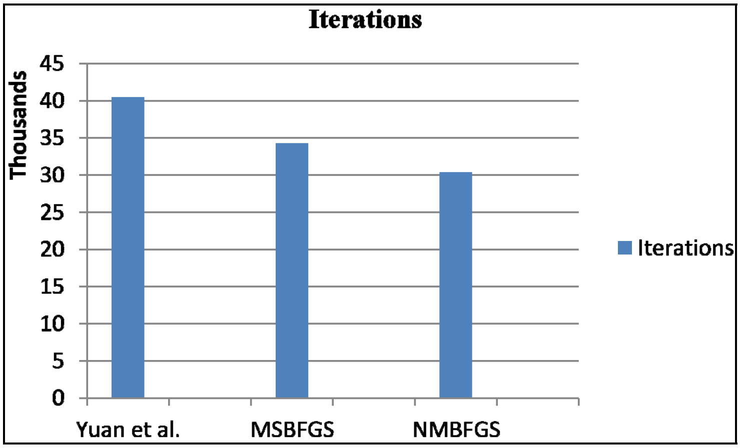
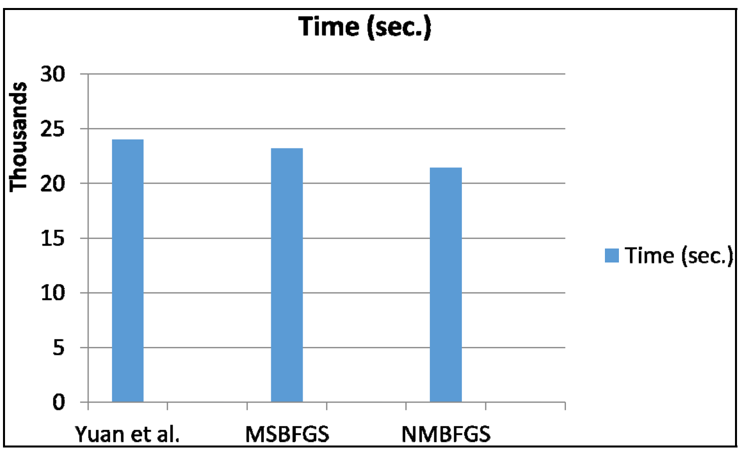



**Figure 3** Overall timings (see online version for colours)




For the standard BFGS method, it is well known that a necessary and sufficient condition for ensuring that the generated updated matrices $\{H_i\}$ are positive definite and hence, the computed search direction is downhill, is that $s_i^T y_i > 0$ (Broyden, 1970). By analogy, for the new method MSBFGS, condition (25) is imposed in order to guarantee that $\overline{v}_i^T u_i$ is 'sufficiently' positive to circumvent any potential numerical instability in the construction of $H_{i+1}$. Should condition (25) fail to hold, the algorithm reverts to using an MSBFGS iteration on that particular instance. An initial scaling is applied to the start inverse Hessian matrix estimate using the techniques introduced in Shanno and Phua (1978).

The preliminary investigation of the new method's numerical performance demonstrates its efficiency on the three evaluation criteria, namely, function and gradient evaluations, as well as the timings. The tabulated results reveal the overall performance merits of the new method as a winner over the other two algorithms on the tested problem of all sizes with a good percentage. The overall savings on the iteration count, function and gradient evaluations, and timings are about 25%, 19%, and 11%, respectively, compared with the benchmarked method (Yuan et al., 2017). The advantages of the new method are best observed as the dimension of the problem increases. On small problems, the gain is marginal (around 3% average on the three criteria), while on large problems, the gain is remarkable. The improvement is around 25% for evaluations and 30% for iteration count, and almost 40% for the timings. Overall, the new algorithm scored best on 44.1% of the 900 problems tested. The MSBFGS method scored best on 35.1% of those. Thus, the new method appears to be a serious contender to some of the well-known algorithms in the same class.

## 6 Conclusions

Secant-like methods built on modifications introduced to the classical secant relation (1) have been examined, and a new non-secant method is derived in this paper. The method utilises nonlinear quantities that incorporate data readily available from the latest three iterations in order to improve the quality of the computed step-wise Hessian approximations and consequently accelerate the convergence to a solution for a given problem. The new method's development is encouraged by several published approaches that motivate and reveal the viability of considering non-secant-based methods. Each of the derivations presented in this research line presents a logical justification for the modified secant equation they adopt (Ford and Moghrabi, 1993; Hassan, 2019; Li and Fukushima, 2001; Yoon and Bekker, 2020; Lewis et al., 2019). A base technique has been established in this work that encourages similar future research to continue. Stemming from the noticeable numerical performance as benchmarked against other similar methods, the new method further promises that the utilisation of data accumulated over the past iterations provides gains. This success constitutes a motivation for further future research that explores alternative techniques for deriving such methods.

The convergence behaviour of the new algorithm has been examined. The numerical scores reported above have shown promising outcomes that turn attention to the potentials of such methods. The scores provide evidence that data accumulated from previous iterations, such as function values, gradient points, and step vectors, pay off numerically and thus provide a motivation to exploit them further.

The performance gains reported play to the advantage of giving more attention to such type of methods that rely on tweaking the secant relation in equation (1). Future research ought to focus on modified secant methods for different choices of the parameters $\tau$ to assess the sensitivity of these methods' numerical behaviour to such choices. It is also worth studying these methods' capability to introduce comparable improvements when utilised in solving systems of nonlinear equations and an examination of the convergence characteristics in that domain. One limitation of this research worth mentioning is testing the new method on constrained problems applied to different managerial or engineering domains. This application constitutes a future venue that is worth exploring. The tests conducted so far on the derived algorithm have been benchmarked against two successful, well-known methods. Future work needs to consider testing more methods to compare the new algorithm's relative success on a broader spectrum.

Also, under investigation are similar methods for which the computed search direction $d_i$ is modified before carrying out the line search (see for example, Faramarzi and Amini, 2020). This idea appears to be rewarding. Also under consideration is a variant of the algorithm where, on every other iteration (or more frequently), the Hessian approximation update is skipped.

## References

Al-Baali, M. (1985) 'New property and global convergence of the Fletcher-Reeves method with inexact line searches', *IMA J. Numer. Anal.*, Vol. 5, No. 3, pp.122–124.

Alipour, M. and Razavi, S. (2019) 'A new local search heuristic based on nearest insertion into the convex hull for solving Euclidean TSP', *International Journal of Operational Research*, Vol. 34, No. 3, pp.409–429.

Bolouri, M., Ziari, S. and Ebrahimnejad, A. (2020) 'New approach for ranking efficient DMUs based on Euclidean norm in data envelopment analysis', *International Journal of Operational Research*, Vol. 37, No. 1, pp.85–104.

Broyden, C.G. (1970) 'The convergence of a class of double-rank minimization algorithms – part 2: the new algorithm', *J. Inst. Math. Applic.*, Vol. 6, No. 1, pp.222–231.

Caliciotti, A., Fasano, G. and Roma, M. (2018) 'Preconditioned nonlinear conjugate gradient methods based on a modified secant equation', *Applied Mathematics and Computation*, Vol. 318, No. C, pp.196–214, Elsevier.

Cao, Q., Griffin, T. and Li, X. (2019) 'A case of unconstrained multiple-factor optimisation with unknown function in the textile industry', *International Journal of Operational Research*, Vol. 34, No. 1, pp.54–65.

Dai, Y., Yuan, J. and Yuan, Y. (2002) 'Modified two-point step size gradient methods for unconstrained optimization', *Computational Optimization and Applications*, Vol. 22, No. 3, pp.103–109.

Das, S.K., Edalatpanah, S.A. and Mnadal, T. (2019) 'A new method for solving linear fractional programming problem with absolute value functions', *International Journal of Operational Research*, Vol. 36, No. 4, pp.455–466.

Deghani, R. and Hosseini, M. (2019) 'Using a modified secant equation for unconstrained optimization', *Journal of Mathematical Extension*, Vol. 13, No. 3, pp.103–116.

Deghani, R., Hosseini, M. and Bidabadi, N. (2019) 'The modified BFGS method with new secant relation for unconstrained optimization problems', *Computational Methods for Differential Equations*, Vol. 7, No. 1, pp.28–41.

Dennis, J.E. and Schnabel, R.B. (1979) 'Least change secant updates for quasi-Newton methods', *SIAM Review*, Vol. 21, No. 1, pp.443–459.

Faramarzi, P. and Amini, K. (2020) 'A modified conjugate gradient method based on a modified secant equation', *Journal of Mathematical Modeling*, Vol. 8, No. 1, pp.1–20, DOI: 10.22124/jmm.2019.14807.1343.

Fletcher, R. (1970) 'A new approach to variable metric algorithms', *Comput. J.*, Vol. 13, No. 4, pp.317–322.

Fletcher, R. (1987) *Practical Methods of Optimization*, 2nd ed., Wiley, New York.

Ford, J.A. and Moghrabi, I.A.R. (1993) 'Alternative parameter choices for multi-step quasi-Newton methods', *Optimization Methods and Software*, Vol. 2, No. 3, pp.357–370.

Ford, J.A. and Moghrabi, I.A.R. (1994) 'Multi-step quasi-Newton methods for optimization', *J. Comput. Appl. Math.*, Vol. 50, No. 1, pp.305–323.

Ford, J.A. and Moghrabi, I.A.R. (1996) 'Using function-values in multi-step quasi-Newton methods', *J. Comput. Appl. Math.*, Vol. 66, No. 2, pp.201–211.

Gondzio, J. and Sobral, F.N.C. (2019) 'Quasi-Newton approaches to interior point methods for quadratic problems', *Computational Optimization and Applications*, Vol. 74, No. 4, pp.93–120.

Hassan, B. (2019) 'A new quasi-Newton updating formula based on the new quasi-Newton equation', *Numerical Algebra Control and Optimization*, Vol. 4, No. 1, pp.63–72.

Lewis, M., Metcalfe, J. and Kochenberger, G. (2019) 'Robust optimisation of unconstrained binary quadratic problems', *International Journal of Operational Research*, Vol. 36, No. 4, pp.112–123.

Li, D. and Fukushima, M. (2001) 'A modified BFGS method and its global convergence in nonconvex minimization', *J. Comput. Appl. Math.*, Vol. 129, No. 1, pp.15–35.

Mahdavi-Amiri, F. and Sadaghiani, S. (2020) 'A superlinearly convergent non-monotone quasi-Newton method for unconstrained multiobjective optimization', *Optimization Methods & Software*, Vol. 35, No. 6, pp.1223–1247.

Moghrabi, I.A.R. (2017) 'Implicit extra-update multi-step quasi-newton methods', *Int. J. Operational Research*, Vol. 28, No. 3, pp.69–81.

Moré, J.J., Garbow, B.S. and Hillstrom, K.E. (1981) 'Testing unconstrained optimization software', *ACM Trans. Math. Softw.*, Vol. 7, No. 1, pp.17–41.

Nakayama, S., Narushima, Y. and Yabe, H. (2019) 'Memoryless quasi-Newton methods based on spectral-scaling Broyden family for unconstrained optimization', *Journal of Industrial & Management Optimization*, Vol. 15, No. 4, pp.1773–1793.

Ortiz, F., Simpson, J.R., Pignatiello, J.J. and Heredia-Langner, A. (2019) 'A genetic algorithm approach to multiple-response optimization', *Journal of Quality Technology*, Vol. 36, No. 4, pp.432–450.

Petra, C., Chiang, N. and Anitescu, M. (2019) 'A structured quasi-Newton algorithm for optimizing with incomplete Hessian information', *SIAM Journal on Optimization*, Vol. 29, No. 2, pp.1048–1075.

Shanno, D.F. (1978) 'On the convergence of a new conjugate gradient algorithm', *SIAM J. Numer. Anal.*, Vol. 15, No. 3, pp.1247–1257.

Shanno, D.F. and Phua, K.H. (1978) 'Matrix conditioning and nonlinear optimization', *Math. Programming*, Vol. 14, No. 1, pp.149–160.

Tajadod, M., Abedini, M., Rategari, A. and Mobin, M. (2016) 'A comparison of multi-criteria decision making approaches for maintenance strategy selection (a case study)', *Journal of Strategic Decision Sciences*, Vol. 7, No. 1, pp.51–56.

Waziri, M., Ahmed, K. and Sabi'u, J. (2020) 'A Dai-Liao conjugate gradient method via modified secant equation for system of nonlinear equations', *Arabian Journal of Mathematics*, Vol. 9, No. 5, pp.443–457.

Wei, Z., Li, G. and Qi, L. (2004) 'The superlinear convergence of a modified BFGS- type method for unconstrained optimization', *Comput. Optim. Appl.*, Vol. 29, No. 2, pp.315–332.

Wei, Z., Li, G. and Qi, L. (2006) 'New quasi-Newton methods for unconstrained optimization problems', *Math Program. Applied Mathematics and Computation*, Vol. 175, No. 3, pp.1156–1188.

Woldu, T., Zhang, H. and Fissuh, Y.H. (2020) 'A scaled conjugate gradient method based on new BFGS secant equation with modified nonmonotone line search', *American Journal of Computational Mathematics*, Vol. 10, No. 1, pp.1–22.

Xiao, Y.H., Wei, Z.X. and Zhang, L. (2006) 'A modified BFGS method without line searches for nonconvex unconstrained optimization', *Advances in Theoretical and Applied Mathematics*, Vol. 1, No. 3, pp.149–162.

Yoon, M. and Bekker, J. (2020) 'Multiobjective simulation optimisation on discrete sets: a literature review', *International Journal of Operational Research*, Vol. 39, No. 3, pp.364–405.

Yuan, G. and Wei, Z. (2010) 'Convex minimizations', *Comp. Optim. Appl.*, Vol. 47, No. 2, pp.237–255.

Yuan, G., Sheng, Z., Wang, B., Hu, W. and Li, C. (2018) 'The global convergence of a modified BFGS method for nonconvex functions', *Journal of Computational and Applied Mathematics*, Vol. 327, No. 1, pp.274–294.

Yuan, G., Wei, Z. and Lu, X. (2017) 'Global convergence of BFGS and PRP methods under a modified weak Wolfe-Powell line search', *Applied Mathematical Modelling*, Vol. 47, No. 4, pp.811–825.

Yuan, G., Wei, Z. and Wu, Y. (2010) 'Modified limited memory BFGS method with non-monotone line search for unconstrained optimization', *J. Korean Math. Soc.*, Vol. 47, No. 3, pp.767–788.

Zhang, J.Z., Deng, N.Y. and Chen, L.H. (1999) 'Quasi-Newton equation and related methods for unconstrained optimization', *JOTA*, Vol. 102, No. 4, pp.147–167.