

Research Article

Scaled Diagonal Gradient-Type Method with Extra Update for Large-Scale Unconstrained Optimization

Mahboubeh Farid,¹ Wah June Leong,¹ Najmeh Malekmohammadi,² and Mustafa Mamat³

¹ Department of Mathematics, University Putra Malaysia, 43400 Serdang, Selangor, Malaysia

² Department of Mathematics, Islamic Azad University, South Tehran Branch, Tehran 1418765663, Iran

³ Department of Mathematics, Faculty of Science and Technology, University Malaysia Terengganu, 21030 Kuala Terengganu, Malaysia

Correspondence should be addressed to Mahboubeh Farid; mfarid7@gmail.com

Received 18 December 2012; Revised 26 February 2013; Accepted 26 February 2013

Academic Editor: Guanglu Zhou

Copyright © 2013 Mahboubeh Farid et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

We present a new gradient method that uses scaling and extra updating within the diagonal updating for solving unconstrained optimization problem. The new method is in the frame of Barzilai and Borwein (BB) method, except that the Hessian matrix is approximated by a diagonal matrix rather than the multiple of identity matrix in the BB method. The main idea is to design a new diagonal updating scheme that incorporates scaling to instantly reduce the large eigenvalues of diagonal approximation and otherwise employs extra updates to increase small eigenvalues. These approaches give us a rapid control in the eigenvalues of the updating matrix and thus improve stepwise convergence. We show that our method is globally convergent. The effectiveness of the method is evaluated by means of numerical comparison with the BB method and its variant.

1. Introduction

In this paper, we consider the unconstrained optimization problem

$$\min f(x), \quad x \in R^n, \quad (1)$$

where $f(x)$ is a continuously differentiable function from R^n to R . Given a starting point x_0 , using notations $g_k = g(x_k) = \nabla f(x_k)$ and B_k as an approximation to the Hessian $G_k = [\nabla^2 f(x_k)]$, the quasi-Newton-based methods for solving (1) are defined by the iteration

$$x_{k+1} = x_k - \alpha_k B_k^{-1} g_k, \quad k = 0, 1, 2, \dots, \quad (2)$$

where the stepsize α_k is determined through an appropriate selection. The updating matrix B_k is usually required to satisfy the quasi-Newton equation

$$B_k s_{k-1} = y_{k-1}, \quad (3)$$

where $s_{k-1} = x_k - x_{k-1}$ and $y_{k-1} = g_k - g_{k-1}$. One of the widely used quasi-Newton method to solve general nonlinear

minimization is the BFGS method, which uses the following updating formula:

$$B_{k+1} = B_k - \frac{B_k s_k s_k^T B_k}{s_k^T B_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}. \quad (4)$$

On the numerical aspect, this method supersedes most of the optimization methods; however, it needs $O(n^2)$ storage which makes it unsuitable for large-scale problems.

On the other hand, an ingenious stepsizes selection for gradient method was proposed by Barzilai and Borwein [1] in which the updating scheme is defined by

$$x_{k+1} = x_k - D_k^{-1} g_k, \quad (5)$$

where $D_k = (1/\alpha_k)I$ and $\alpha_k = s_{k-1}^T s_{k-1} / s_{k-1}^T y_{k-1}$.

Since that, the study of new effective methods in the frame of BB-like gradient methods becomes an interesting research topic for a wide range of mathematical programming; for example, see [2–10]. However, it is well known that BB method cannot guarantee a descent in the objective function at each iteration and the extent of the nonmonotonicity

depends in some way on the size of the condition number of objective function [11]. Therefore, the performance of BB method is greatly influenced by the condition of the problem (particularly, condition number of the Hessian matrix). Some new fixed stepsizes gradient-type methods of BB kind are proposed by [12–16] to overcome these difficulties. In contrast with the BB approach in which the stepsize is computed by means of a simple approximation of the Hessian in the form of scalar multiple of identity, these proposed methods consider approximation of the Hessian and its inverse in diagonal matrix form based on the weak secant equation and quasi-cauchy relation, respectively (for more details see [15, 16]). Though these diagonal updating methods are efficient, their performance can be greatly affected by solving ill-conditioned problems. Thus, there is room for improve on the quality of the diagonal updates formulation. Since methods as described in [15, 16] have useful theoretical and numerical properties, it is desirable to derive a new and more efficient updating frame for general functions. Therefore our aim is to improve the quality of diagonal updating when it is poor in approximating Hessian.

This paper is organized as follows. In the next section, we describe our motivation and propose our new-gradient type method. The global convergence of the method under mild assumption will be established in Section 3. Numerical evidence of the vast improvements due to the new approach is given in Section 4. Finally, conclusion is made in the last section.

2. Scaling and Extra Updating

Assume that B_k is positive definite, and let $\{y_k\}$ and $\{s_k\}$ be two sequences of n -vectors such that $y_k^T s_k > 0$ for all k . Because it is usually difficult to satisfy the quasi-Newton equation (3) with a nonsingular B_{k+1} of the diagonal form, one can consider satisfying it in some directions. If we project the quasi-Newton equation (3) (also called the secant equation), in a direction v such that $y_k^T v \neq 0$, then it gives

$$s_k^T B_{k+1} v = y_k^T v. \quad (6)$$

If $v = s_k$ is chosen, it leads to the so-called weak-secant relation,

$$s_k^T B_{k+1} s_k = y_k^T s_k. \quad (7)$$

Under this weak-secant equation, [15, 16] employ variational technique to derive updating matrix that approximates the Hessian matrix diagonally. The resulting update is derived to be the solution of the following variational problem:

$$\begin{aligned} \min \quad & \frac{1}{2} \|B_{k+1} - B_k\|_F^2 \\ \text{s.t.} \quad & s_k^T B_{k+1} s_k = s_k^T y_k, \end{aligned} \quad (8)$$

B_{k+1} is diagonal

and gives the corresponding solution B_{k+1} as follows:

$$B_{k+1} = B_k + \frac{(s_k^T y_k - s_k^T B_k s_k)}{\text{tr}(E_k^2)} E_k, \quad (9)$$

where $E_k = \text{diag}(s_{k,1}^2, s_{k,2}^2, \dots, s_{k,n}^2)$, $s_{k,i}$ is the i th component of the vector s_k , and tr denotes the trace operator.

Note that when $s_k^T y_k < s_k^T B_k s_k$, the resulting B_{k+1} is not necessarily positive definite and it is not appropriate for use within a quasi-Newton-based algorithm. Thus, it is desirable to propose a technique to measure the quality of B_k in terms of its Rayleigh quotient and try to find a way to improve ‘‘poor’’ quality B_k before calculating B_{k+1} . For this purpose, it will be useful to propose, at first quality a criterion to distinguish between poor, and acceptable quality of B_k .

Let us begin by considering the curvature of an objective function, f in direction s_k , which is represented by

$$s_k^T \bar{G}_k s_k = s_k^T y_k, \quad (10)$$

where $\bar{G}_k = \int_0^1 \nabla^2 f(x_k + ts_k) dt$ is the average Hessian matrix along s_k . Since it is not practical to compute the eigenvalue of the Hessian matrix in each iteration, we can estimate its relative size on the basis of the scalar

$$\rho_k = \frac{(s_k^T \bar{G}_k s_k / s_k^T s_k)}{(s_k^T B_k s_k / s_k^T s_k)} = \frac{s_k^T \bar{G}_k s_k}{s_k^T B_k s_k} = \frac{s_k^T y_k}{s_k^T B_k s_k}. \quad (11)$$

If $\rho_k > 1$, it implies that the eigenvalues of B_k approximated by its Rayleigh are relatively small compared to those of the local Hessian matrix at x_k . In this condition, we find that the strategy of extra update [17] seems to be useful for improving the quality of B_k by rapidly increasing its eigenvalues up to those of the actual Hessian relatively. This is done by updating B_k twice to obtain $\hat{B}_{k+1,2}$:

$$\hat{B}_{k+1,1} = B_k + \frac{(s_k^T y_k - s_k^T B_k s_k)}{\text{tr}(E_k^2)} E_k, \quad (12)$$

$$\hat{B}_{k+1,2} = \hat{B}_{k+1,1} + \frac{(s_{k-1}^T y_{k-1} - s_{k-1}^T \hat{B}_{k+1,1} s_{k-1})}{\text{tr}(E_{k-1}^2)} E_{k-1}, \quad (13)$$

and use it to obtain, finally, the updated B_{k+1} :

$$B_{k+1} = \hat{B}_{k+1,2} + \frac{(s_k^T y_k - s_k^T \hat{B}_{k+1,2} s_k)}{\text{tr}(E_k^2)} E_k. \quad (14)$$

On the other hand, when $\rho_k < 1$, it implies that the eigenvalue of B_k represented by its Rayleigh is relatively large and we have $s_k^T y_k - s_k^T B_k s_k < 0$. In this case, we should suggest a useful strategy to encounter this drawback. As we reviewed before, the updating scheme may generate nonpositive definite B_{k+1} when B_k has large eigenvalues relative to those possible values of \bar{G}_k , that is, when $s_k^T y_k < s_k^T B_k s_k$. On the contrary, this argument disappears when the eigenvalues of B_k are small (i.e., $s_k^T y_k > s_k^T B_k s_k$). This suggests that the scaling should be made to scale down B_k , that is, choosing $\rho_k < 1$ only when $s_k^T y_k - s_k^T B_k s_k < 0$ and take $\rho_k = 1$, whenever $s_k^T y_k > s_k^T B_k s_k$. Combining these two arguments, we choose the scaling parameter ρ_k such that

$$\gamma_k = \min(\rho_k, 1). \quad (15)$$

This scaling resembles the Al-Baali [18] scaling that is applied within the Broyden family. Because the value of γ_k is always < 1 , then by incorporating the scaling to B_k , it decreases the large eigenvalues of B_k constantly, and consequently we can keep positive definiteness of B_{k+1} (since $s_k^T \gamma_k > 0$), which is an important property in descent method. In this case, the following updating:

$$B_{k+1} = \gamma_k B_k + \frac{(s_k^T \gamma_k - \gamma_k s_k^T B_k s_k)}{\text{tr}(E_k^2)} E_k, \quad (16)$$

will be used. To this end, we have the following general updating scheme for B_{k+1} :

$$B_{k+1} = \begin{cases} \gamma_k B_k + \frac{(s_k^T \gamma_k - \gamma_k s_k^T B_k s_k)}{\text{tr}(E_k^2)} E_k; & \text{if } \rho_k \leq 1, \\ \widehat{B}_{k+1,2} + \frac{(s_k^T \gamma_k - s_k^T \widehat{B}_{k+1,2} s_k)}{\text{tr}(E_k^2)} E_k; & \text{if } \rho_k > 1, \end{cases} \quad (17)$$

where $\widehat{B}_{k+1,2}$ and γ_k are given by (13) and (15), respectively.

An advantage of using (17) is that the positive definiteness of B_{k+1} can be guaranteed in all iterations. This property is not exhibited in the other diagonal updating formula such as those in [15, 16]. Note that there is no extra storage required to impose our strategy and the cost of computing is also not increased significantly throughout the entire iteration. Now we can state the steps of our new diagonal-gradient method algorithm with the safeguarding strategy for monotonicity as follows.

2.1. ESDG Algorithm

Step 1. Choose an initial point $x_0 \in R^n$ and a positive definite matrix $B_0 = I$. Let $\theta \in (1, 2)$. Set $k := 0$.

Step 2. Compute g_k . If $\|g_k\| \leq \epsilon$, stop.

Step 3. If $k = 0$, set $x_1 = x_0 - g_0/\|g_0\|$.

Step 4. Compute $d_k = -B_k^{-1} g_k$, and calculate $\alpha_k > 0$ such that the following condition holds: $f(x_{k+1}) \leq f_k^{\max} + \sigma \alpha_k g_k^T d_k$ where $f_k^{\max} = \max\{f(x_k), f(x_{k-1})\}$ and $\sigma \in (0, 1)$ is a given constant.

Step 5. If $k \geq 1$, let $x_{k+1} = x_k - \alpha_k B_k^{-1} g_k$ and compute ρ_k and γ_k by (11) and (15), respectively. If $\rho_k < \theta$ then update B_{k+1} by (16).

Step 6. If $\rho_k \geq \theta$ then compute $\widehat{B}_{k+1,1}$ and $\widehat{B}_{k+1,2}$ by (12), (13), respectively, and then update as defined B_{k+1} (14).

Step 7. Set $k := k + 1$, and return to Step 2.

In Step 4, we employ the nonmonotone line search of [19, 20] to ensure the convergence of the algorithm. However, some other line search strategies may also be used.

3. Convergence Analysis

This section is devoted to study the convergence behavior of ESDG method. We will establish the convergence of the ESDG algorithm when applied to the minimization of a strictly convex function. To begin, we give the convergence result, which is due to Grippo et al. [21] for the step generated by the nonmonotone line search algorithm. Here and elsewhere, $\|\cdot\|$ denotes the Euclidean norm.

Theorem 1. *Assume that f is a strictly convex function and its gradient g satisfies the Lipschitz condition. Suppose that the nonmonotone line search algorithm is employed in a case that the steplength, α_k , satisfies*

$$f(x_{k+1}) \leq f_k^{\max} + \sigma \alpha_k g_k^T d_k, \quad (18)$$

where $f_k^{\max} = \max\{f(x_k), f(x_{k-1}), \dots, f(x_{k-m})\}$, with $m \leq k$ and $\sigma \in (0, 1)$, and the search direction d_k is chosen to obey the following conditions. There exist positive constants c_1 and c_2 such that

$$-g_k^T d_k \geq c_1 \|g_k\|^2, \quad \|d_k\| \leq c_2 \|g_k\|, \quad (19)$$

for all sufficiently large k . Then the iterates x_k generated by the nonmonotone line search algorithm have the property that

$$\liminf_{k \rightarrow \infty} \|g_k\| = 0. \quad (20)$$

To prove that the ESDG algorithm is globally convergent, it is sufficient to show that the sequence $\{\|B_k\|\}$ generated by (17) is bounded both above and below, for all finite k so that its associated search direction satisfies condition (19). Since B_k is diagonal, it is enough to show that each element of B_k , say $B_k^{(i)}$, $i = 1, \dots, n$, is bounded above and below by some positive constants. The following theorem gives the boundedness of $\{\|B_k\|\}$.

Theorem 2. *Assume that f is strictly convex function where there exist positive constants m and M such that*

$$m \|z\|^2 \leq z^T \nabla^2 f(x) z \leq M \|z\|^2, \quad (21)$$

for all $x, z \in R^n$. Let $\{\|B_k\|\}$ be a sequence generated by the ESDG method. Then $\|B_k\|$ is bounded above and below for all finite k , by some positive constants.

Proof. Let $B_k^{(i)}$ be the i th element of B_k . Suppose that B_0 is chosen such that $\omega_1 \leq B_0^{(i)} \leq \omega_2$, $i = 1, \dots, n$ where ω_1, ω_2 are some positive constants. It follows from (17) and the definition of γ in (15) that we have

$$B_1 = \begin{cases} \rho_0 B_0, & \text{if } \rho_0 \leq 1, \\ \widehat{B}_{1,2} + \frac{(s_0^T \gamma_0 - s_0^T \widehat{B}_{1,2} s_0)}{\text{tr}(E_0^2)} E_0, & \text{if } \rho_0 > 1, \end{cases} \quad (22)$$

where

$$\begin{aligned}\widehat{B}_{1,1} &= B_0 + \frac{(s_0^T y_0 - s_0^T B_0 s_0)}{\text{tr}(E_0^2)} E_0, \\ \widehat{B}_{1,2} &= \widehat{B}_{1,1} + \frac{(s_0^T y_0 - s_0^T \widehat{B}_{1,1} s_0)}{\text{tr}(E_0^2)} E_0.\end{aligned}\quad (23)$$

Moreover, by (21) and (11), we obtain

$$m \|s_k\|^2 \leq s_k^T y_k \leq M \|s_k\|^2, \quad \forall k. \quad (24)$$

Case 1. When $\rho_0 \leq 1$: by (24), one can obtain

$$\frac{m}{\omega_2} \leq \rho_0 = \frac{s_0^T y_0}{s_0^T B_0 s_0} \leq \frac{M}{\omega_1}. \quad (25)$$

Thus, it implies that $m\omega_1/\omega_2 \leq B_1^{(i)} = \rho_0 B_0^{(i)} \leq M\omega_2/\omega_1$.

Case 2. When $\rho_0 > 1$: from (3), we have

$$\widehat{B}_{1,1}^{(i)} = B_0^{(i)} + \frac{(s_0^T y_0 - s_0^T B_0 s_0)}{\text{tr}(E_0^2)} s_{0,i}^2. \quad (26)$$

Because $\rho_0 > 1$ also implies that $s_0^T y_0 - s_0^T B_0 s_0 > 0$, using this fact and (24) give

$$B_0^{(i)} \leq \widehat{B}_{1,1}^{(i)} \leq B_0^{(i)} + \frac{(M - \omega_1) \|s_0\|^2}{\text{tr}(E_0^2)} s_{0,i}^2. \quad (27)$$

Let $s_{0,M}$ be the largest component in magnitude of s_0 , that is, $s_{0,i}^2 \leq s_{0,M}^2$, for all i . Then it follows that $\|s_0\|^2 \leq n s_{0,M}^2$, and (27) becomes

$$\omega_1 \leq \widehat{B}_{1,1}^{(i)} \leq \omega_2 + \frac{n(M - \omega_1)}{\text{tr}(E_0^2)} s_{0,M}^4 \leq \omega_2 + n(M - \omega_1). \quad (28)$$

Using (28) and the same argument as previously mentioned, we can also show that

$$\omega_1 \leq \widehat{B}_{1,2}^{(i)} \leq \omega_2 + n(M - \omega_1) + n[M - (\omega_2 + n(M - \omega_1))]. \quad (29)$$

Hence in both cases, $B_1^{(i)}$ is bounded above and below, by some positive constants. Since the upper and lower bounds for $B_1^{(i)}$ are independent of k , respectively, we can proceed by using induction to show that $B_k^{(i)}$ is bounded, for all finite k . \square

4. Numerical Results

In this section we present the results of numerical investigation for ESDG method on different test problems. We also compare the performance of our new method with that of the BB method and that of MDGRAD method which is implemented using SMDQN of [22] with a same nonmonotone strategy as the ESDG method. Our experiments are

TABLE 1: Test problem and its dimension.

Problem	References
Extended Freudenstein and Roth, Extended Trigonometric,	
Broyden Tridiagonal, Extended Beale, Generalized Rosenbrock,	Moré et al. [24]
Extended Tridiagonal 2, Extended Himmelblau, Raydan 2, EG2,	
Extended Three Exponential Terms, Raydan 1, Generalized PSC1,	
Quadratic QF2, Generalized Tridiagonal 1, Perturbed Quadratic,	
Diagonal 2, Diagonal 3, Diagonal 5, Almost perturbed Quadratic,	
Hager, diagonal 4	Andrei [23]

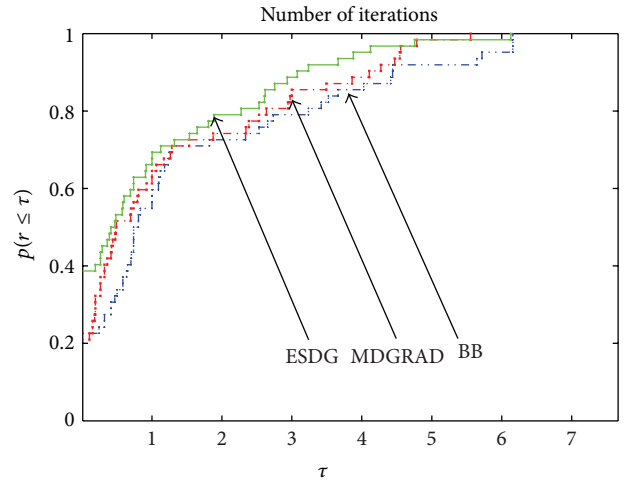


FIGURE 1: Performance profile based on iterations.

performed on a set of 20 nonlinear unconstrained problems with dimensions ranging from 10 to 10^4 (Table 1).

These test problems are taken from [23, 24]. The codes are developed with Matlab 7.0. All runs are performed on a PC with Core Duo CPU. For each test, the termination condition is $\|g(x_k)\| \leq 10^{-4}$. The maximum number of iterations is set to 1000.

Figures 1 and 2 show the efficiency of ESDG method when compared to MDGRAD and BB methods. Note that ESDG method increases the efficiency of Hessian approximation devoid of increasing the number of storages. Figure 2 also shows the implementation of the ESDG method with BB and MDGRAD methods using the CPU time as a measure. This figure shows that ESDG method is again faster than MDGRAD method in most problems and requires reasonable time to solve large-scale problems when compares to the BB method. Finally, we can conclude that our experimental comparisons indicate that our extension is very beneficial to the performance.

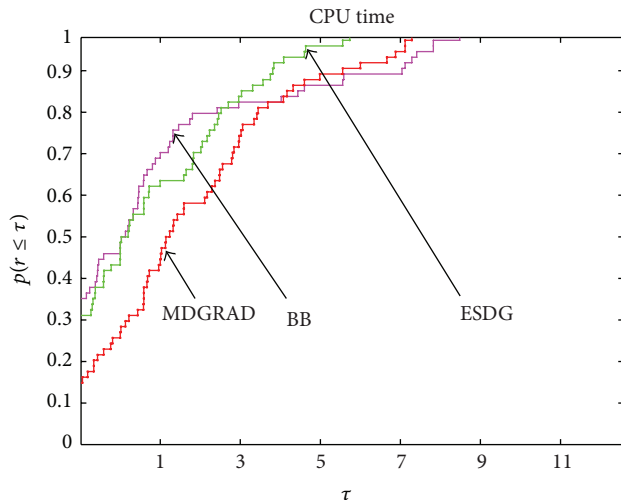


FIGURE 2: Performance profile based on CPU time per iteration.

5. Conclusion

We have presented a new diagonal gradient method for unconstrained optimization. Numerical study of the proposed method when compared with BB and MDGRAD methods is also performed. Based on our numerical experiments, we can conclude that ESDG method is significantly preferable compared to the BB and MDGRAD methods. Particularly, the ESDG method is proven to be a good option for large-scale problems when high-memory locations are required. In view of the remarkable performance of ESDG method, globally converged and with only $O(n)$ storage, we can expect that our proposed method would be useful for unconstrained large-scale optimization problems.

References

- [1] J. Barzilai and J. M. Borwein, "Two-point step size gradient methods," *IMA Journal of Numerical Analysis*, vol. 8, no. 1, pp. 141–148, 1988.
- [2] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex sets," *SIAM Journal on Optimization*, vol. 10, no. 4, pp. 1196–1211, 2000.
- [3] Y.-H. Dai and R. Fletcher, "Projected Barzilai-Borwein methods for large-scale box-constrained quadratic programming," *Numerische Mathematik*, vol. 100, no. 1, pp. 21–47, 2005.
- [4] Y.-H. Dai and L.-Z. Liao, "R-linear convergence of the Barzilai and Borwein gradient method," *IMA Journal of Numerical Analysis*, vol. 22, no. 1, pp. 1–10, 2002.
- [5] Y.-H. Dai, W. W. Hager, K. Schittkowski, and H. Zhang, "The cyclic Barzilai-Borwein method for unconstrained optimization," *IMA Journal of Numerical Analysis*, vol. 26, no. 3, pp. 604–627, 2006.
- [6] G. Frassoldati, G. Zanghirati, and L. Zanni, "New adaptive stepsize selections in gradient methods," *Journal of Industrial and Management Optimization*, vol. 4, no. 2, pp. 299–312, 2008.
- [7] M. Raydan, "On the Barzilai and Borwein choice of steplength for the gradient method," *IMA Journal of Numerical Analysis*, vol. 13, no. 3, pp. 321–326, 1993.
- [8] M. Raydan, "The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem," *SIAM Journal on Optimization*, vol. 7, no. 1, pp. 26–33, 1997.
- [9] B. Zhou, L. Gao, and Y. Dai, "Monotone projected gradient methods for large-scale box-constrained quadratic programming," *Science in China. Series A*, vol. 49, no. 5, pp. 688–702, 2006.
- [10] B. Zhou, L. Gao, and Y.-H. Dai, "Gradient methods with adaptive step-sizes," *Computational Optimization and Applications*, vol. 35, no. 1, pp. 69–86, 2006.
- [11] R. Fletcher, "On the Barzilai-Borwein method," Tech. Rep. NA/207, Department of Mathematics, University of Dundee, Scotland, UK, 2001.
- [12] M. Farid, W. J. Leong, and M. A. Hassan, "A new two-step gradient-type method for large-scale unconstrained optimization," *Computers & Mathematics with Applications*, vol. 59, no. 10, pp. 3301–3307, 2010.
- [13] M. Farid and W. J. Leong, "An improved multi-step gradient-type method for large scale optimization," *Computers & Mathematics with Applications*, vol. 61, no. 11, pp. 3312–3318, 2011.
- [14] M. Farid, W. J. Leong, and L. Zheng, "Accumulative approach in multistep diagonal gradient-type method for large-scale unconstrained optimization," *Journal of Applied Mathematics*, vol. 2012, Article ID 875494, 11 pages, 2012.
- [15] M. A. Hassan, W. J. Leong, and M. Farid, "A new gradient method via quasi-Cauchy relation which guarantees descent," *Journal of Computational and Applied Mathematics*, vol. 230, no. 1, pp. 300–305, 2009.
- [16] W. J. Leong, M. A. Hassan, and M. Farid, "A monotone gradient method via weak secant equation for unconstrained optimization," *Taiwanese Journal of Mathematics*, vol. 14, no. 2, pp. 413–423, 2010.
- [17] M. Al-Baali, "Extra updates for the BFGS method," *Optimization Methods and Software*, vol. 13, no. 3, pp. 159–179, 2000.
- [18] M. Al-Baali, "Numerical experience with a class of self-scaling quasi-Newton algorithms," *Journal of Optimization Theory and Applications*, vol. 96, no. 3, pp. 533–553, 1998.
- [19] E. G. Birgin, J. M. Martínez, and M. Raydan, "Inexact spectral projected gradient methods on convex sets," *IMA Journal of Numerical Analysis*, vol. 23, no. 4, pp. 539–559, 2003.
- [20] E. G. Birgin, J. M. Martínez, and M. Raydan, "Nonmonotone spectral projected gradient methods on convex," *Encyclopedia of Optimization*, pp. 3652–3659, 2009.
- [21] L. Grippo, F. Lampariello, and S. Lucidi, "A nonmonotone line search technique for Newton's method," *SIAM Journal on Numerical Analysis*, vol. 23, no. 4, pp. 707–716, 1986.
- [22] W. J. Leong, M. Farid, and M. A. Hassan, "Scaling on diagonal quasi-Newton update for large-scale unconstrained optimization," *Bulletin of the Malaysian Mathematical Sciences Society*, vol. 35, no. 2, pp. 247–256, 2012.
- [23] N. Andrei, "An unconstrained optimization test functions collection," *Advanced Modeling and Optimization*, vol. 10, no. 1, pp. 147–161, 2008.
- [24] J. J. Moré, B. S. Garbow, and K. E. Hillstom, "Testing unconstrained optimization software," *ACM Transactions on Mathematical Software*, vol. 7, no. 1, pp. 17–41, 1981.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

