Journal of Engineering and Applied Sciences 12 (17): 4504-4507, 2017
ISSN: 1816-949X

A Modified Search Direction of Broyden Family Method and its Global Convergence

[1]Mohd Asrul Hery Ibrahim, [2]Mustafa Mamat and [1]Azfi Zaidi Mohammad Sofi
[1]Faculty of Entrepreneurship and Business, Universiti Malaysia Kelantan,
Kampus Kota, Kota Bharu, Malaysia
[2]Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin Kampus Tembila,
Kuala Terenggongu, Malaysia

Abstract: In this study, a new search direction for Broyden family method is proposed for solving unconstrained optimization problems. The new search direction is develop by hybridization the two search direction in line search method known as quasi-Newton and conjugate gradient method under certain parameter. This method is popular as Broyden-CG method. The suggestion method has an attractive properties which is its search direction is sufficiently descent direction at every iteration. Under mild conditions, the researchers prove that the proposed method has global convergence.

Key words: Broyden family method, conjugate gradient method, search direction, global convergence, solving uncostrained optimiztion problem, iteration

INTRODUCTION

Consider, the unconstrained optimization problems:

$$\min_{x \in R^n} f(x) \tag{1}$$

and let f: $R^n \rightarrow R$ be continuously differentiable. The Broyden's method is an iterative method. On the i+1 th iteration, an approximation point x_i and the i+1th iteration of x is given by:

$$x_{i+1} = x_i + a_i d_i \tag{2}$$

where, the search direction, d_i is calculated by:

$$d_i = -B_i^{-1} g(x_i) \tag{3}$$

which g_i is a gradient of f. The search direction must satisfy the relation $g^T_i d_i < 0$ which guarantee that d_i is a descent direction of f (x) at x_i (Nocedal and Wright, 2006; Byrd and Nocedal, 1989). Then the step size, α_i in Eq. 2 was obtained using the Wolfe line search (Wolfe, 1971) which is $\alpha_i>0$ satisfying:

$$f(x_i) - f(x_i + a_i d_i) \geq -\delta a_i g_i^T d_i \tag{4}$$

And:

$$\sigma_1 g_i^T d_i, \leq (g(x_i + a_i d_i)^T d_i \leq -\sigma_2 g_i^T d_i$$

where, $0<\delta<\sigma_1<1$, $0\leq\sigma_2<+\infty$ are constants. Then, the sequence of $\{x^i\}_{i=0}^{\infty}$ is converged to the optimal point, x^* which minimizes (Han and Neumann, 2003). The updated Hessian approximation Eq. 3, require B_i positive definite and satisfying the quasi-Newton equation:

$$B_{i+1} s_i = y_i \tag{5}$$

Where:

$$s_i = a_i d_i \qquad y_i = g_{i+1} - g_i \tag{6}$$

The Broyden's algorithm for unconstrained optimization problem uses the matrices B_i which is updated by Eq. 7:

$$B_{i+1} = B_i - \left(\frac{B_i s_i s_i^T B_i}{s_i^T B_i y_i}\right) + \frac{y_i y_i^T}{s_i^T y_i} + \phi_i \left(s_i^T B_i s_i\right) v_i v_i^T \tag{7}$$

where, ϕ is a scalar and:

$$v_i = \left[\frac{y_i}{s_i^T} - \frac{B_i s_i}{s_i^T B_i s_i}\right]$$

This algorithm satisfy the quasi-Newton Eq. 7. The choice of the parameter ϕ is important, since it can greatly affect the perfomance of the method (Xu, 2003). When $\phi_i = 0$ in Eq. 7, we obtain the DFP algorithm and

Corresponding Author: Mohd Asrul Hery Ibrahim, Faculty of Entrepreneurship and Business, Universiti Malaysia Kelantan, Kampus Kota, Kota Bharu, Malaysia

$\phi_i = 0$ we get the BFGS algorithm. But, Byrd and Nocedal (1989) and Byrd *et al*. (1987) extended his result to $\phi \in (0, 1)$ based on Chong and Zak (2001) the Broyden's algorithm is one of the most efficient algorithm for solving the unconstrained optimization problem.

MATERIALS AND METHODS

The Broyden-CG algorithm: The modification of the quasi-Newton method based on a hybrid method has already been introduced by previous researchers. One of the studies is a hybridization of the quasi-Newton and Gauss-Siedel methods, aimed at solving the system of linear equations by Ludwig (2007). Luo *et al*. (2008) suggest the new hybrid method which can solve the system of nonlinear equations by combining the Quasi-Newton method with chaos optimization. Han and Newman (2003) combine the Quasi-Newton methods and Cauchy descent method to solve unconstrained optimization problems which is recognised as the Quasi-Newton-SD method.

Hence, the modification of the Quasi-Newton method by previous researchers spawned the new idea of hybridizing the classical method to yield the new hybrid method such as by Han and Neumann (2003) and Ibrahim *et al*. (2014). Hence, this study proposes a new hybrid search direction that combines the concept of search direction of the quasi-Newton and CG methods. It yields a new search direction of the hybrid method which is known as the Broyden-CG method. The search direction for the Broyden-CG method is:

$$d_i = \begin{cases} -B_i^{-1}g_i & i = 0 \\ -B_i^{-1}g_i + \eta(-g_i + \beta_i d_{i-1}) & i \geq 1 \end{cases} \quad (8)$$

where $\eta>0$ and $\beta_k = \frac{g_k^T(g_k - g_{k-1})}{g_k^T d_{k-1}}$.

Hence, the complete algorithm for the the Broyden-CG method will be arranged as follows (Algorithim 1):

Algorithm 1; Broyden-CG method:

Step 0. Given a starting point x_0 and $H_0 = I_n$ Choose values for s, β and σ set i = 1

Step 1. Terminate if $\|g(x_{i+1})\|<10^{-6}$ or $i\geq 10000$

Step 2. Calculate the search direction by Eq. 8

Step 3. Calculate the step size α_i by Eq. 4

Step 4. Compute the difference between $s_i = x_i-x_{i-1}$ and $y_i = g_i-g_{i-1}$

Step 5. Update H_{i-1} by Eq. 7 to obtain H_i

Step 6. Set i = i+1 and go to Step 1.

RESULTS AND DISCUSSION

Global convergence: Throughout this study, we assume that every search direction d_i satisfied the descent condition:

$$g_i^t d_i < 0$$

for all $i\geq 0$ If there exists a constant $c_1>0$ such that:

$$g_i^t d_i \leq c_i \|g_i\|^2 \quad (9)$$

for all $i\geq 0$ then the search directions satisfy the sufficient descent condition which can be proved in Theorem 3.2. Hence, we need to make a few assumptions based on the objective function:

Assumption 3.1:

- H_1: the objective function f is twice continuously differentiable
- H_2: the level set L is convex. Moreover, positive constants c_i and c_2 exist, satisfying:

$$c_1\|z\|^2 \leq z^T F(x)z \leq c_2\|z\|^2$$

for all $z\in R^n$ and $x\in L$ where f (x) is the Hessian matrix for f

- H_3: the Hessian matrix is Lipschitz continuous at the point x that is there exists the positive constant c_3 satisfying:

$$\|g(x) - g(x^*)\| \leq c_3\|x - x^*\|$$

for all x in a neighbourhood of x^*.

Theorem 3.1 (Byrd and Nocedal, 1989; Byrd *et al*., 1987): Let $\{B_i\}$ be generated by the Broyden family's Eq. 7 where B_i is symmetric and positive definite and where $y^T_i s_i>0$ for all i. Furthermore, assume that $\{s_i\}$ and $\{y_i\}$ are such that:

$$\frac{\|(y_i - G_*)s_i\|}{\|s_i\|} \leq \varepsilon_i \quad (10)$$

for some symmetric and positive definite matrix $G(x^*)$ and for some sequence $\{\varepsilon_i\}$ with the property $\sum_{i=1}^{\infty}\varepsilon_i < \infty$. Then:

$$\lim_{i\to\infty}\frac{\|(B_i - G_*)d_i\|}{\|d_i\|} = 0 \quad (11)$$

and the sequence $\{\|B_i\|\}$, $\{\|B_i^{-1}\|\}$ are bound.

Theorem 3.2: Suppose that Assumption 3.1 and Theorem 3.1 hold. Then condition Eq. 9 holds for all $i \geq 0$.

Proof: Equation 9 we see that:

$$\begin{aligned} g_i^T d_i &= -g_i^T B_i^{-1} g_i + \eta g_i^T \left(-g_i + \left((g_i - g_{i-1})^T g_i / g_i^T d_{i-1}\right) d_{i-1}\right) \\ &= -g_i^T B_i^{-1} g_i + \eta\left(-g_i^T g_i + \left((g_i - g_{i-1})^T g_i / g_i^T d_{i-1}\right) g_i^T d_{i-1}\right) \\ &= -g_i^T B_i^{-1} g_i + \eta\left(-g_i^T g_{i-1}\right) \end{aligned}$$

Based on Powell (1977), $g_i^T g_{i-1} \geq \varepsilon \|g_i\|^2$ with $\varepsilon = (0, 1)$ then:

$$\begin{aligned} g_i^T d_i &= -g_i^T B_i^{-1} g_i + \eta\left(\varepsilon \|g_i\|^2\right). \\ &\leq -\lambda_i \|g_i\|^2 - \eta\varepsilon \|g_i\|^2 \\ &\leq c_1 \|g_i\|^2, \end{aligned}$$

where, $c_i = -(\lambda_i + \eta\varepsilon)$ which is bound away from zero. Hence, $g_i^T d_i \leq c_1 \|g_i\|^2$ holds. The proof is completed.

Lemma 3.1: Under Assumption 3.1, positive constants ϖ_1 and ϖ_2 exist such that for any x_1 and any d_i with $g^T{}_i d_i < 0$, the step size α_i produced by Algorithm 1 will satisfy either:

$$f(x_i + a_i d_i) - f_i \leq -\varpi_1 \frac{\left(g_i^T d_i\right)^2}{\|d_i\|^2} \quad (12)$$

Or:

$$f(x_i + a_i d_i) - f_i \leq \varpi_1 g_i^T d_i$$

Proof: Suppose that $\alpha_i < 1$ which means that Eq. 4 failed for a step size $\alpha' \leq \alpha_i / \tau$:

$$f(x_i + \alpha_i' d_i) - f(x_i) \leq \varpi a' g_i^T d_i \quad (13)$$

Then, by using the mean value theorem, we obtain:

$$f(x_{i+1}) - f(x_i) = \bar{g}^T (x_{i+1} - x_i)$$

where $\bar{g} = \nabla f(\bar{x})$ for some $\bar{x} \in (x_i, x_{i+1})$. Now by the Cauchy-Schwartz inequality we get:

$$\begin{aligned} \bar{g}^T (x_{i+1} - x_i) &= g^T (x_{i+1} - x_i) + (\bar{g} - g_i)^T (x_{i+1} - x_i) \\ &= g^T (x_{i+1} - x_i) + \|\bar{g} - g_i\| (x_{i+1} - x_i) \\ &\leq g^T (x_{i+1} - x_i) + \mu \|x_{i+1} - x_i\|^2 \\ &\leq g^T (a' d_i) + \mu \|a' d\|^2 \\ &\leq g^T (a' d_i) + \mu \left(a' \|d\|\right)^2. \end{aligned}$$

Thus, from H_3:

$$(\varpi - 1) a' g_i^T d_i < a' (\bar{g} - g_i)^T d_i \leq M(a' \|d_i\|)^2$$

which implies that:

$$\alpha_i \geq t\alpha' > t(1-\varpi) \frac{-g_i^T d_i}{M(\alpha' \|d_i\|)^2}$$

Substituting this into Eq. 13 we have $f(x_i + \alpha'_i d_i) - f(x_i) \leq c_2 \frac{-g_i^T d_i}{(\alpha' \|d_i\|)^2}$

where, $c_2 = \tau (1-\varpi)/M$ which gives Eq. 12.

Theorem 3.3 (global convergence): Suppose that Assumption 3.1 and Theorem 3.1 hold. Then:

$$\lim_{i \to \infty} \|g_i\|^2 = 0$$

Proof: Combining descent property Eq. 9 and Lemma 3.1 gives:

$$\sum_{i=0}^{\infty} \frac{\|g_i\|^4}{\|d_i\|^2} < \infty \quad (14)$$

Hence, from Theorem 3.2 we can define that $\|d_i\| \leq -c\|g_i\|$. Then, Eq. 14 will be simplified as $\sum_{i=0}^{\infty} \|g_i\|^2 < \infty$. Therefore, the proof is completed.

CONCLUSION

In this study, we propose a new search direction of Broyden family method based on hybridization of quasi-Newton and conjugate gradient's search direction. Regarding theorem and lemma in previous section show that the search direction possesses the sufficient of descent condition. Hence, the global convergence of Broyden-CG method with Wolfe line search rule is proven. Our further interest is to try the suggested method to solve the large scale of unconstrained optimization problems in matlab programming and compared with classical Quasi-Newton methods in term of number of iteration, number of function evaluation and CPU-time.

REFERENCES

Byrd, R.H. and J. Nocedal, 1989. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. SIAM J. Numer. Anal., 26: 727-739.

Byrd, R.H., J. Nocedal and J. Y.X Yuan, 1987. Global convergence of a class of quasi-newton methods on Convex problems. SIAM J. Numer. Anal., 24: 1171-1190.

Chong, E.K.P. and S.H. Zak, 2001. An Introduction to Optimization. John Wiley and Sons, New York, pp: 365-433.

Han, L. and M. Neumann, 2003. Combining quasi-Newton and steepest descent directions. Intl. J. Appl. Math., 12: 167-171.

Ibrahim, M.A.H., M. Mamat and W.J. Leong, 2014. The hybrid BFGS-CG method in solving unconstrained optimization problems. Abstract Appl. Anal., 2014: 1-6.

Ludwig, A., 2007. The Gauss-Seidel-quasi-Newton method: A hybrid algorithm for solving dynamic economic models. J. Econ. Dyn. Control, 31: 1610-1632.

Luo, Y.Z., G.J. Tang and L.N. Zhou, 2008. Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-Newton method. Appl. Soft Comput., 8: 1068-1073.

Nocedal, J. and S.J. Wright, 2006. Numerical Optimization. Springer-Verlag, New York, USA., pp: 188-193.

Powell, M.J.D., 1977. Restart procedures for the conjugate gradient method. Mathe. Program., 12: 241-254.

Wolfe, P., 1969. Convergence conditions for ascent method. SIAM Rev., 11: 226-235.

Wolfe, P., 1971. Convergence conditions for ascent methods II: Some corrections. SIAM. Rev., 13: 185-188.

Xu, D.C., 2003. Global convergence of the broyden's class of quasi-newton methods with nonmonotone linesearch. Acta Math. Appl. Sinica, 19: 19-24.