The Algorithms of Broyden-CG for
Unconstrained Optimization Problems

Mohd Asrul Hery Ibrahim
Faculty of Entrepreneurship and Business
Universiti Malaysia Kelantan (UMK)
Kampus Kota, Malaysia

Mustafa Mamat
Faculty of Informatics and Computing
Universiti Sultan Zainal Abidin (UNISZA)
Kampus Tembila, Malaysia

Leong Wah June
Department of Mathematics
Faculty of Science, University Putra Malaysia (UPM)
Malaysia

Azfi Zaidi Sofi
Faculty of Science and Information Technology,
Kolej Universiti Islam Antarabangsa Selangor (KUIS)
Malaysia

Copyright © 2014 Mohd Asrul Hery Ibrahim 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.

Abstract

The conjugate gradient method plays an important role in solving large-scaled problems and the quasi-Newton method is known as the most efficient method in solving unconstrained optimization problems. Therefore, in this paper, the new hybrid
method between the conjugate gradient method and the quasi-newton method for solving optimization problem is suggested. The Broyden family formula is used as an approximation of Hessian in the hybrid method and the quasi-Newton method. Our numerical analysis provides strong evidence that our Broyden-CG method is more efficient than the ordinary Broyden method. Furthermore, we also prove that new algorithm is globally convergent and gratify the sufficient descent condition.

**Keywords:** Broyden method, conjugate gradient method, search direction, global convergent

1 Introduction

Consider the unconstrained optimization problems:

\[ \min_{x \in \mathbb{R}^n} f(x) \]  

(1)

and let \( f : \mathbb{R}^n \rightarrow \mathbb{R} \) be continuously differentiable. The Broyden’s family method is an iterative method. On the \( i \)th iteration, an approximation point \( x_i \) and the \((i+1)\)th iteration of \( x \) is given by

\[ x_{i+1} = x_i + \alpha_i d_i \]

(2)

where the search direction, \( d_i \) is calculated by

\[ d_i = -B_i^{-1}g(x_i) \]

(3)

which \( g_i \) is a gradient of \( f \). The search direction must satisfy the relation \( g_i^T d_i < 0 \), which guarantee that \( d_i \) is a descent direction of \( f(x) \) at \( x_i \) [1, 2]. Then the step size, \( \alpha_i \) in (2) was obtained using the Armijo line search as suggested by [3] such as:

\[ s > 0, \beta \in (0, 1), \text{ and } \alpha_i = \max \{s, s\beta, s\beta^2, \ldots\} \]

(4)

such that

\[ f(x) - f(x + \alpha_i d) \geq \sigma \alpha_i^T g_i \]

\[ i = 0, 1, 2, \ldots \] Then, the sequence of \( \{x_i\}_{i=0}^\infty \) is converged to the optimal point, \( x^* \), which minimizes [4]. The updated Hessian approximation formula in (3), require \( B_i \) positive definite and satisfying the quasi-Newton equation

\[ B_{i+1} s_i = x_i \]

(5)

where
The algorithms of Broyden-CG

\[ s_i = \alpha d \]
\[ y_i = g_{i+1} - g_i \]

(6)

The Broyden’s algorithm for unconstrained optimization problem uses the matrices \( B_i \) which is updated by the formula

\[
B_{i+1} = B_i - \left( \frac{B_i s_i y_i^T}{s_i^T B_i y_i} \right) + \frac{y_i y_i^T}{s_i^T y_i} + \phi \left( s_i^T B_i s_i \right) v_i v_i^T,
\]

(7)

where \( \phi \) is a scalar and

\[
v_i = \begin{bmatrix}
y_i \\
\frac{s_i^T y_i}{s_i^T B_i y_i} \\
\frac{s_i^T B_i s_i}{s_i^T B_i y_i}
\end{bmatrix}.
\]

This algorithm satisfy the quasi-Newton equation (7). The choice of the parameter \( \phi \) is important, since it can greatly affect the performance of the method [5]. When \( \phi = 1 \) in equation (7), we obtain the DFP algorithm and \( \phi = 0 \), we get the BFGS algorithm. But, [2, 6] extended his result to \( \phi \in (0,1] \). Based on [7], the Broyden’s algorithm is one of the most efficient algorithm for solving the unconstrained optimization problem.

This paper is organised as follows. In Section 2, we elaborate the new algorithm and the convergence analysis. An explanation about the numerical results is also given in Section 3 using the performance profile’s figure. The paper ends with a short conclusion in section 4.

2 The Broyden-CG Algorithm

The modification of the quasi-Newton method based on a hybrid method has already been presented by previous researchers. One of the studies is a hybridization of the quasi-Newton and Gauss-Siedel methods, designed in solving the system of linear equations in [8]. Luo et al. [9] 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 [4] combine the quasi-Newton methods and Cauchy descent method to solve unconstrained optimization problems, which is known 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 in [10-13]. 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
\[ d_i = \begin{cases} -B^{-1}_i g_i & i = 0 \\ -B^{-1}_i g_i + \eta(-g_i + \beta_i d_{i-1}) & i \geq 1 \end{cases} \] (8)

where \( \eta > 0 \) and \( \beta_i = (g_i^T g_{i-1}) / (g_i^T d_{i-1}) \).

Hence, the complete algorithms for the Broyden method and the Broyden-CG method will be organized in Algorithms 2.1 and 2.2, respectively.

**Algorithm 2.1: Broyden Method**

**Step0.** Given a starting point \( x_0 \) and \( B_0 = I_n \). Choose values for \( s, \beta \) and \( \sigma \) and set \( i = 1 \).

**Step1.** Terminate if \( \|g(x_{i+1})\| < 10^{-6} \) or \( i \geq 10000 \).

**Step2.** Calculate the search direction by (3)

**Step3.** Calculate the step size \( \alpha_i \) by (4)

**Step4.** Compute the difference between \( s_i = x_i - x_{i-1} \) and \( y_i = g_i - g_{i-1} \).

**Step5.** Update \( B_{i-1} \) by (7) to obtain \( B_i \).

**Step6.** Set \( i = i + 1 \) and go to Step 1.

**Algorithm 2.2: Broyden-CG Method**

**Step0.** Given a starting point \( x_0 \) and \( B_0 = I_n \). Choose values for \( s, \beta \) and \( \sigma \) and set \( i = 1 \).

**Step1.** Terminate if \( \|g(x_{i+1})\| < 10^{-6} \) or \( i \geq 10000 \).

**Step2.** Calculate the search direction by (8)

**Step3.** Calculate the step size \( \alpha_i \) by (4)

**Step4.** Compute the difference between \( s_i = x_i - x_{i-1} \) and \( y_i = g_i - g_{i-1} \).

**Step5.** Update \( B_{i-1} \) by (7) to obtain \( B_i \).

**Step6.** Set \( i = i + 1 \) and go to Step 1.

Based on Algorithms 2.1 and 2.2, 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 \( c_i > 0 \) such that
\[ g_i^T d_i \leq c_i \|g_i\|^2 \] (9)
for all \( i \geq 0 \), then the search directions satisfy the sufficient descent condition
The algorithms of Broyden-CG

which can be proved in Theorem 2.2. Hence, we need to make a few assumptions based on the objective function.

Assumption 2.1

H1: The objective function $f$ is twice continuously differentiable.

H2: The level set $L$ is convex. Moreover, positive constants $c_1$ and $c_2$ exist, satisfying
\[ c_1 \|z\|^2 \leq z^TF(x)z \leq c_2 \|z\|^2, \]
for all $z \in \mathbb{R}^n$ and $x \in L$, where $F(x)$ is the Hessian matrix for $f$.

H3: 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 2.1 (see [2, 6])

Let $\{B_i\}$ be generated by the Broyden’s family formula (7), where $B_i$ is symmetric and positive definite, and where $y_i^T s_i > 0$ for all $i$. Furthermore, assume that $\{s_i\}$ and $\{y_i\}$ are such that
\[ \frac{\|(y_i - G_i)s_i\|}{\|s_i\|} \leq \varepsilon_i \quad (10) \]
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
\[ \lim_{i \to \infty} \frac{\|(B_i - G_i)d_i\|}{\|d_i\|} = (11) \]
and the sequence $\{\|B_i\|\}, \{\|B_i^{-1}\|\}$ are bound.

Theorem 2.2

Suppose that Assumption 2.1 and Theorem 2.1 hold. Then condition (9) holds for all $i \geq 0$.

Proof:

From (9), we see that
Based on Powell [14], $g_i^T g_{i-1} \geq \varepsilon \|g_i\|^2$ with $\varepsilon = (0,1)$, then

$$g_i^T d_i = -g_i^T B_i g_i + \eta \left( -\|g_i\|^2 + \varepsilon \|g_i\|^2 \right).$$

$$\leq -\lambda_i \|g_i\|^2 + (-\eta + \eta\varepsilon) \|g_i\|^2$$

$$\leq c_1 \|g_i\|^2,$$

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

3 Numerical Analysis

In this section, we use the test problem considered in Andrei [15], Michalewicz [16] and More et al. [17] in Table 1 to analyse the improvement of the Broyden-CG method compared with the Broyden method. Each of the test problems are tested with dimensions varied from 2 to 1,000 variables. This represents a total of 180 test problems. As suggested by [17], for each of the test problems, the initial point, $x_0$ will further subtract from the minimum point. In doing so, this leads us to test the global convergence properties and the robustness of our method. For the Armijo line search, we use $s=1$, $\beta=0.5$ and $\sigma=0.1$. The stopping criteria we use are $\|g_i\| \leq 10^{-6}$ and the number of iterations exceeds its limit, which is set to be 10,000. In our implementation, the numerical tests were performed on an Acer Aspire with a Windows 7 operating system and using Matlab 2012 to run the programming for both methods.
**TABLE (1):** A list of problem functions.

<table>
<thead>
<tr>
<th>TEST PROBLEM</th>
<th>N-DIMENSIONAL</th>
<th>SOURCES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Powell Badly Scaled</td>
<td>2</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Beale</td>
<td>2</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Biggs Exp6</td>
<td>6</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Chebyquad</td>
<td>4,6</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Colville Polynomial</td>
<td>4</td>
<td>Michalewicz [16]</td>
</tr>
<tr>
<td>Variably Dimensioned</td>
<td>4,8</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Freudenstein And Roth</td>
<td>2</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Goldstein Price polynomial</td>
<td>2</td>
<td>Michalewicz [16]</td>
</tr>
<tr>
<td>Himmelblau</td>
<td>2</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Penalty 1</td>
<td>2,4</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Extended Powell Singular</td>
<td>4,8</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Extended Rosenbrock</td>
<td>2,10,100,200,500,1000</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Trigonometric</td>
<td>6</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Watson</td>
<td>4,8</td>
<td>More et al. [17]</td>
</tr>
<tr>
<td>Six-Hump Camel Back Polynomial</td>
<td>2</td>
<td>Michalewicz [16]</td>
</tr>
<tr>
<td>Extended Shallow</td>
<td>2,4,10,100,200,500,1000</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Extended Strait</td>
<td>2,4,10,100,200,500,1000</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Scale</td>
<td>2</td>
<td>Michalewicz [16]</td>
</tr>
<tr>
<td>Raydan 1</td>
<td>2,4</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Raydan 2</td>
<td>2,4</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Diagonal 3</td>
<td>2</td>
<td>Andrei [15]</td>
</tr>
<tr>
<td>Cube</td>
<td>2,10,100,200</td>
<td>More et al. [17]</td>
</tr>
</tbody>
</table>

**FIGURE 1:** Broyden method versus Broyden-CG method in term of number of iterations
The performance results will be shown in Figures 1 and 2, respectively, using the performance profile introduced by Dolan and More [18]. The performance profile seeks to find how well the solvers perform relative to the other solvers on a set of problems. In general, \( P(\tau) \) is the fraction of problems with performance ratio \( \tau \), thus, a solver with high values of \( P(\tau) \) or one that is located at the top right of the figure is preferable.

Figures 1 and 2 show that the Broyden-CG method has the best performance since it can solve 84% of the test problems compared with the Broyden method (83%). Moreover, we can also say that the Broyden-CG is the fastest solver on approximately 64% of the test problems for iteration and 69% of CPU-time.

4 Conclusion

We have presented a new search direction for Broyden method for solving unconstrained optimization problems. The performance profile for a broad class of test problems show that the Broyden-CG method is efficient and robust in solving unconstrained optimization problem. We also note that as the size and complexity of the problem increase, greater improvements could be realised by our Broyden-CG method. Our future research will be to try the Broyden-CG with another formula for \( \beta_k \).

References


The algorithms of Broyden-CG


http://dx.doi.org/10.1007/bf01593790


http://dx.doi.org/10.1145/355934.355936

http://dx.doi.org/10.1007/s101070100263

Received: September 9, 2014; Published: November 20, 2014