Global Convergence Condition for a New Spectral Conjugate Gradient Method for Large-Scale Optimization

Ahmed Hussien Sheekoo, Ghada M. al-Naemi

Department of Mathematics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq.
E-mail: drghadaalnaemi@uomosul.edu.iq

Abstract. The spectral conjugate gradient (SCG) method is an effective method to solve large-scale nonlinear unconstrained optimization problems. In this work, a new spectral conjugate gradient method is proposed with a strong Wolfe-Powell line search (SWP). The idea of the new one is using the β_{ZA} formula which is proposed by Baluch and et al., with suitable parameter φ denoted by (SCGBZA). Under the usual assumptions, the descent properties and overall global convergence of the proposed method (SCGBZA) are proved. The proposed method is numerically proven to be effective.

1. Introduction
Conjugate gradient (CG) and SCG methods are the most effective categories for solving large-scale nonlinear unconstrained optimization problems, this is because they have the advantage of fast convergence, low storage and simple iterations [1]. Now consider the nonlinear unconstrained optimization problems
\[ \min f(x), \quad x \in \mathbb{R}^n, \]  
where \( f: \mathbb{R}^n \to \mathbb{R} \) a smooth function, and its gradient vector is usually represented by \( g(x) = \nabla f(x) \). The initial point \( x_0 \in \mathbb{R}^n \) is usually calculated through iterative process. The new point calculated as follows:
\[ x_{n+1} = x_n + \gamma_n d_n, \quad n = 0,1,2,3,... \] (2)
and the direction
\[ d_{n+1} = \begin{cases} -g_{n+1}, & n = 0 \\ -g_{n+1} + \beta_n d_n, & n \geq 1 \end{cases} \] (3)
Where \( \beta_n \in \mathbb{R} \) is the parameter, and \( \gamma_n > 0 \) is the line search generated by inexact line search (ILS). In this work, we use (SWP) defined by:
\[ f(x_n + \gamma_n d_n) \leq f(x_n) + \tau g_n^T d_n \] (4)
\[ |g(x_n + \gamma_n d_n)^T d_n| \leq -\delta g_n^T d_n \] (5)
Generally, the above parameters $\tau$ and $\mu$ are required to satisfy $0 < \tau < 0.5 < \delta < 1$. In order to generate different (CG) methods, we have different choices for the parameter $\beta_n$. The most commonly used formula for parameters are Hestenes Stiefel method (HS) [2], Fletcher-Reeves method (FR) [3], Polak-Ribiere – Polyak method (PR) [4, 5], Conjugate – Descent method (CD) [6], Liu – Storey method (LS) [7] and Dai-Yuan method (DY) [8]. The parameters of these $\beta_n$ are as follows:

\[
\begin{align*}
\beta_n^{HS} &= \frac{\theta_{n+1}}{\gamma_{n+1}}, \\
\beta_n^{FR} &= \frac{\|g_{n+1}\|^2}{\|g_n\|^2} - 1, \\
\beta_n^{PR} &= \frac{\theta_{n+1}}{\gamma_{n+1}}, \\
\beta_n^{CD} &= \frac{\|g_{n+1}\|^2}{\gamma_{n+1}}, \\
\beta_n^{LS} &= \frac{\theta_{n+1}}{\gamma_{n+1}} - \frac{1}{\gamma_n}, \\
\beta_n^{DY} &= \frac{\|g_{n+1}\|^2}{\gamma_{n+1}}.
\end{align*}
\]

Respectively, where $\| \cdot \|$ is the standard Euclidean norm and $\gamma_n = g_{n+1} - g_n$.

Another way to solve the CG problem (1) is the SCG method. The main difference between the spectral gradient method and the gradient method is the calculation of the search direction [9]. The search direction of the spectral gradient method is defined by the following formula:

\[
d_n = -\varphi_n g_{n+1} + \beta_n d_n
\]

Where $\varphi_n$ is a parameter which is known as a spectral gradient. Observe that if $\varphi_n = 1$, then (6) reduced to (3).

SCG was first proposed by Barazilai and Borwein [10] in 1988. Further, Raydan [11] introduced the SCG method for potential large-scale unconstrained optimization. The main merit of this method is that only the gradient directions is used for each search, while the nonmonotone strategy can ensures global convergence. Birgin and Martines [12] concluded that their SCG method was globally convergent. However, there is actually no guarantee that the SCG method will generate descending direction. Therefore, Andrei [13] proposed a reduced proportion under the Wolfe line search. Based on the improved a CG algorithm which was proposed by Jiang et al., [14], an SCG method with sufficient descent feature based on the modified CG algorithm that proposed by Zhang et al. [15]. Many authors engaged in the development of the SCG methods, for more information, please refer to [16-21].

## 2. New algorithm and the descent property

The SCG method is obtained by combining the CG search direction and a scalar spectral parameter. Baluch et al. [22] in their paper, they constructed a modified three-term HS method in the following direction:

\[
d_{n+1} = -g_{n+1} + \beta_n^{BA} d_n - \varphi_n^{BA} \gamma_n
\]

Note that the suggested formula $\beta_n^{BA}$ actually modifies the classic $\beta^{HS}$ formula by adding $\mu \|g_{n+1}^T d_n\|$ in the denominator of HS, the proposed $\beta$ formula formulated as:

\[
\beta_n^{BA} = \frac{\theta_{n+1}}{\gamma_{n+1}^T - \gamma_n} \frac{\|g_{n+1}^T d_n\|}{\|g_{n+1}^T + \mu \|g_{n+1}^T d_n\|}
\]

with $\mu = 2$, and the parameter $\varphi$ formulated as:

\[
\varphi_n^{BA} = \frac{\theta_{n+1}^T d_n}{\gamma_{n+1}^T + \mu \|g_{n+1}^T d_n\|}
\]

If the line search is exact, then (8) reduce to the classical HS formula, and (7) reduce to (3).

In this work, we can summarize the direction defined as follows:

\[
d_{n+1} = \begin{cases} 
-\frac{g_{n+1}}{\gamma_{n+1}}, & n = 0 \\
-\varphi_n^{BA} g_{n+1} + \beta_n^{BA} d_n, & n \geq 1
\end{cases}
\]

Where $\beta_n^{BA}$ defined in (8), and $\varphi_n^{BA}$ it defined as follows:

\[
\varphi_n^{BA} = 1 - \frac{\theta_{n+1}^T d_n}{\gamma_{n+1}^T + \mu \|g_{n+1}^T d_n\|}
\]

with $\mu = 0.5$. If the line search is exact, $\varphi_n = 1$, (10) reduce to (3). The process of our proposed SCG method is described in the following algorithm:
3. Algorithm of SCG

Step 1: Choose an initial point \( x_1 \in \mathbb{R}^n, \epsilon \geq 0, \epsilon = 10^{-6}, \mu = 0.5, \) set 
\[ d_1 = -g_1, \quad n=1. \]

Step 2: Check convergence, if \( \|g_n\| \leq \epsilon \), then stop; otherwise, continue.

Step 3: Use (4) and (5) to determine the step size \( y_n = 0 \).

Step 4: Calculate the new point \( x_{n+1} \) using (2), \( g_{n+1} = g(x_{n+1}), \)
if \( \|g_{n+1}\| \leq \epsilon \), then stop; otherwise, continue.

Step 5: Calculate the direction \( d_{n+1} \) given in (10), \( \beta_n^{RZA} \) and \( \phi_n^{MBZA} \) are given in (8) and (11) respectively.

Step 6: If the Powel restart criteria
\[ \|g^T_{n+1} g_n\| \geq 0.2 \|g_{n+1}\|^2 \] (12)
is satisfied, set \( d_{n+1} = -g_{n+1} \) go to 3; otherwise, continue.

Step 7: Increase \( n \) and go to 3.

In order to ensure that the SCG algorithm proposed in (10) has sufficient descent property, which plays an important role in the global convergence analysis under the SWP in (4) and (5), we need the following assumptions:

4. Assumption (A).

1- \( f(x) \) is restricted from below on the level set \( \Psi = \{ x \in \mathbb{R}^n, f(x) \leq f(x_0) \} \), where \( x_0 \) is the starting point. i.e., there is a constant \( \eta > 0 \), which means \( \|x_n\| \leq \eta \ \forall x \in \Psi \) [23].

2- \( f(x) \) is continuously differentiable in a certain neighborhood \( N \) of \( \Psi \), and its gradient is Lipschitz continuous, i.e., there is a constant \( L > 0 \), such that
\[ \|g(x) - g(y)\| \leq L\|x - y\|, \forall x, y \in N. \]

Now using Assumption (A), there exists a positive constant \( (\tilde{\omega}, \tilde{\omega}, \tilde{\rho} \) and \( \rho \), such that:
\[ 0 < \tilde{\omega} \leq \|g_{n+1}\| \leq \omega, \quad \text{and} \quad 0 < \rho \leq \|g_n\| \leq \rho, \forall x \in \Psi \) [23].

Theorem (1): Suppose that assumption (A) holds. Assuming that the sequences \( \{g_n\} \) and \( \{d_n\} \) be generated by the algorithm SCG, and the step size \( y_n \) is obtained by SWP. Then, the proposed method has sufficient descent direction i.e.
\[ g^T_{n+1} d_{n+1} \leq -\nu \|g_{n+1}\|^2 \] (14)

Proof: Multiplying the both sides of the direction defined in (10) by \( g^T_{n+1} \) and substituting \( \beta_n^{RZA} \) and \( \phi_n^{MBZA} \) in (8) and (11) respectively, we get
\[ g^T_{n+1} d_{n+1} = -\left(1 - \frac{g^T_{n+1} d_n}{d_n y_{n+1} \|g_{n+1} \|^2} \right)\|g_{n+1}\|^2 + \frac{g^T_{n+1} d_n}{d_n y_{n+1} \|g_{n+1} \|^2} g^T_{n+1} d_{n+1} \] (15)

From (5), we have
\[ d_n y_{n} = d_n y_{n+1} - d_n g_n \geq -\frac{g^T_{n+1} d_n}{d_n y_{n+1} \|g_{n+1} \|^2} g^T_{n+1} d_{n+1} \]
i.e.
\[ d_n y_{n} \geq -(1 - \delta_1) g^T_{n+1} d_{n+1} . \] (17)

We have
\[ g^T_{n+1} y_{n} = g^T_{n+1} (g_{n+1} - g_n) = \|g_{n+1}\|^2 - g^T_{n+1} g_n. \]

By the one side of Powell restart criteria (12), we obtain
\[ g^T_{n+1} y_{n} \leq \|g_{n+1}\|^2 + 0.2 \|g_{n+1}\|^2 = 1.2 \|g_{n+1}\|^2. \] (18)

Using (16), (17) and (18) in (15), we get
\[ g_{n+1}^T d_{n+1} \leq - \left( 1 - \frac{-\delta_2 d_n^T g_n}{1 - \delta_1 d_n^T g_n - \mu \delta_1 d_n^T g_n} \| g_{n+1} \|^2 + \frac{1.2 \| g_{n+1} \|^2 \| g_n^T d_n \|}{(1 - \delta_1)(1 - d_n^T g_n) - \mu \delta_1 d_n^T g_n} \right) + \frac{1.2 \| g_{n+1} \|^2}{(1 - \delta_1)(1 - d_n^T g_n) - \mu \delta_1 d_n^T g_n} \| g_{n+1} \|^2 \]
\[ = - \left( 1 - \frac{1}{1 - (1+\mu)\delta_1} \| g_{n+1} \|^2 - \frac{1.2 \delta_2}{1 - (1+\mu)\delta_1} \| g_{n+1} \|^2 \right) \| g_{n+1} \|^2 \]
\[ = - \left( 1 - \frac{2.2 \delta_2}{1 - (1+\mu)\delta_1} \| g_{n+1} \|^2 \right). \]
\[ \Rightarrow g_{n+1}^T d_{n+1} \leq -\nu \| g_{n+1} \|^2, \text{ with } \nu = \left( 1 - \frac{1}{1 - (1+\mu)\delta_1} \right). \]

Therefore, the proposed algorithm satisfies the sufficient descent condition with SWP conditions.

5. The Global Convergence Property.

In this section, we will prove another important condition, called global convergence property. In the following lemma, we review the well-known Zoutendijk condition [24], which plays an important role in the proof of the global convergence analysis of SCG method.

Lemma [25]:
Let assumption (A) holds. Suppose any iteration method (2) and (3), and \( y_n \) is obtained by the SWP (4) and (5). If
\[ \sum_{n \geq 1} \frac{1}{\| d_n \|^2} = \infty, \] (20)
Then
\[ \lim_{n \to \infty} \inf \| g_n \| = 0. \] (21)

Theorem (2): Consider that assumption (A) is satisfied. The sequences \( \{ x_n \} \) and \( \{ d_n \} \) generated by the algorithm SCG, \( y_n \) is obtained by SWP and \( d_n \) is the descent direction. Then
\[ \lim_{n \to \infty} \inf \| g_n \| = 0 \]

Proof: As
\[ \beta_n^{BZA} = \frac{\delta_{n+1} (g_{n+1} - g_n)}{d_n y_n + \mu |g_{n+1}^T d_n|^2}. \]

The authors in [22] assumed that \( z = \left[ (1 - \delta_1) y_n z_1 + \mu \delta_1 z_1 \right] \), and where \( 0 < \delta_1 < 1, z_1 > 0 \) and \( \mu = 0.5 \), so \( z > 0 \). Therefore,
\[ |\beta_n^{BZA}| \leq \frac{\ell \| g_{n+1} \|}{z_2 \| d_n \|^2} = \frac{\ell \| g_{n+1} \|}{z_2} \] (22)

Now,
\[ |1 - \phi_n^{MBZA}| \leq \frac{1}{1 - \| g_{n+1} \|^2 \| d_n \|^2} \leq 1 + \frac{\ell \| g_{n+1} \|}{z_2 \| d_n \|^2} \]
\[ \leq 1 + \frac{\ell \| g_{n+1} \|}{z_2 \| d_n \|^2} \]
\[ \leq 1 + \frac{\ell \| g_{n+1} \|}{z_2 \| d_n \|^2} \]

Assume that \( A = \left( 1 + \frac{\omega}{z_b} + \frac{\ell}{z_2} \right) \)

Now, combining (22) and (23) with (10), we get
\[ \| d_{n+1} \| \leq \| 1 - \phi_n^{MBZA} \| \| g_{n+1} \| + |\beta_n^{BZA}| \| d_n \| \]
\[ \leq 1 + \frac{\ell \| g_{n+1} \|}{z_2 \| d_n \|^2} \]
\[
\frac{1}{\|d_{n+1}\|^2} \geq \frac{1}{B^2} \sum_{n \geq 1} 1 = +\infty
\]

By taking the summation to the both sides of (24), we get

\[
\liminf_{n \to \infty} \|g_n\| = 0. \blacksquare
\]

The new proposed algorithm has achieved global convergence.

6. Results and Discussion
We will report the results of several test functions in this section. Some test functions are selected to analyze the new method. These functions are considered from CUTEr [26] and Andrei [27]. Using SWP line search, a comparison was made between the new SCG method and the classical BZACG method based on the number of iterations (NOI) and number of function evaluation (NOF). Set \(\tau = 0.01\) and \(\delta = 0.7\), the stopping criterion of this algorithm was \(\|g_{n+1}\| \leq 10^{-6}\). If the number of iterations exceeds 600, the method is considered to have failed. All codes are written in double-precision FORTRAN 77 language and compiled into Visual (Fortran 6.6) (default compiler settings). Under Table 1, we have compiled the names of test functions used and the numerical results between the BZACG algorithm and the SCG algorithm.

| No. | Test Functions | N    | Classical algorithm | The SCG algorithm |
|-----|----------------|------|---------------------|-------------------|
|     |                | N    | NOI | NOF | NOI | NOF |
| 1   | Wood           | 1000 | 33  | 73  | 29  | 66  |
|     |                | 5000 | 34  | 75  | 29  | 66  |
|     |                | 10000 | 37 | 81  | 32  | 72  |
|     |                | 50000 | 38 | 83  | 32  | 72  |
|     |                | 100000 | 41 | 93  | 33  | 73  |
|     |                | 100000 | 14 | 44  | 14  | 44  |
|     |                | 50000 | 14  | 44  | 14  | 44  |
| 2   | Cubic          | 10000 | 14  | 44  | 14  | 44  |
|     |                | 50000 | 14  | 44  | 14  | 44  |
|     |                | 100000 | 14 | 44  | 14  | 44  |
|     |                | 100000 | 60 | 163 | 51  | 161 |
|     |                | 50000 | 60  | 163 | 51  | 161 |
| 3   | Extended Hiebert | 10000 | 60  | 163 | 51  | 161 |
|     |                | 50000 | 60  | 163 | 51  | 161 |
|     |                | 100000 | 60 | 163 | 51  | 161 |
|     |                | 100000 | 60 | 163 | 51  | 161 |
|     |                | 100000 | 600 | 1172 | 51  | 126 |
|     |                | 50000 | 600 | 1174 | 49  | 121 |
|     |                | 100000 | 600 | 1174 | 51  | 127 |
| 4   | Helical        | 50000 | 600 | 1174 | 60  | 144 |
|     |                | 100000 | 600 | 1174 | 75  | 176 |
| 5   | Powell         | 10000 | 42  | 126 | 39  | 116 |
| No. | Test Functions | N | Classical algorithm | The SCG algorithm |
|-----|----------------|---|---------------------|-------------------|
|     |                |   | NOI | NOF | NOI | NOF |
| 5000 |                |   | 42  | 126 | 39  | 116 |
| 10000 |               |   | 49  | 169 | 39  | 116 |
| 50000 |              |   | 57  | 190 | 45  | 135 |
| 100000 |             |   | 57  | 191 | 45  | 135 |
| 1000 |                 |   | 30  | 78  | 27  |  74 |
| 5000 |                 |   | 32  | 80  | 28  |  76 |
| 10000 |                |   | 36  | 92  | 29  |  78 |
| 100000 |              |   | 36  | 92  | 29  |  78 |
| 1000 |                 |   | 19  | 99  | 18  |  97 |
| 5000 |                 |   | 33  | 202 | 19  |  86 |
| 6 | Rosen | 10000 | 33 | 82 | 28 | 76 |
| 50000 |              |   | 36  | 92  | 29  |  78 |
| 100000 |             |   | 36  | 92  | 29  |  78 |
| 1000 |                 |   | 19  | 99  | 18  |  97 |
| 5000 |                 |   | 33  | 202 | 19  |  86 |
| 7 | Summ | 10000 | 40 | 356 | 34 | 240 |
| 50000 |              |   | 65  | 311 | 63  | 153 |
| 100000 |             |   | 72  | 377 | 69  | 286 |
| 1000 |                 |   | 188 | 634 | 187 | 560 |
| 5000 |                 |   | 387 | 1170 | 405 | 1292 |
| 8 | OSP | 10000 | 573 | 1790 | 567 | 1779 |
| 50000 |              |   | 580 | 2305 | 577 | 2002 |
| 100000 |             |   | 595 | 2287 | 580 | 2011 |
| 1000 |                 |   | 9   | 21  |  9  |  21 |
| 5000 |                 |   | 9   | 21  |  9  |  21 |
| 9 | DENISCHNB (CUTE) | 10000 | 9  | 21 | 9  | 21 |
| 50000 |              |   | 9   | 21  |  9  |  21 |
| 100000 |             |   | 9   | 21  |  9  |  21 |
| 1000 |                 |   | 86  | 334 | 77  | 305 |
| 5000 |                 |   | 93  | 376 | 84  | 346 |
| 10000 |                |   | 93  | 376 | 84  | 346 |
| 1000 |                 |   | 93  | 376 | 84  | 346 |
| 1 | Miele | 50000 | 107 | 460 | 98  | 430 |
| 100000 |            |   | 107 | 460 | 98  | 430 |
| 1000 |                 |   | 11  | 23  | 10  |  21 |
| 5000 |                 |   | 11  | 23  | 10  |  21 |
| 10000 |                |   | 12  | 25  | 11  |  23 |
| 50000 |              |   | 12  | 25  | 11  |  23 |
| 100000 |             |   | 12  | 25  | 12  |  23 |
| 1000 |                 |   | 60  | 121 | 50  | 104 |
| 5000 |                 |   | 151 | 306 | 60  | 126 |
| 10000 |                |   | 142 | 288 | 161 | 334 |
| 10000 |                |   | 113 | 234 | 166 | 342 |
| 1000 |                 |   | 86  | 334 | 77  | 305 |
| 5000 |                 |   | 93  | 376 | 84  | 346 |
| 10000 |                |   | 93  | 376 | 84  | 346 |
| 1 | Wolfe | 50000 | 169 | 344 | 168 | 347 |
| 100000 |            |   | 113 | 234 | 166 | 342 |
| 1000 |                 |   | 86  | 334 | 77  | 305 |
| 5000 |                 |   | 93  | 376 | 84  | 346 |
| 10000 |                |   | 93  | 376 | 84  | 346 |
| 1 | Miele | 50000 | 107 | 460 | 98  | 430 |
| 100000 |            |   | 203 | 606 | 189 | 582 |
| 1000 |                 |   | 256 | 756 | 151 | 452 |
| N o. | Test Functions     | N   | Classical algorithm | The SCG algorithm |
|------|-------------------|-----|---------------------|-------------------|
|      |                   |     | NOI | NOF | NOI | NOF |
| 1    | DIXMAANI          | 5000| 600 | 1780| 319 | 956 |
| 2    |                   | 10000| 600 | 1780| 438 | 1313|
| 3    |                   | 50000| 600 | 1780| 587 | 1643|
| 4    |                   | 100000| 600 | 1780| 530 | 1764|
| 5    | Ex-Freudenstein   | 1000| 9   | 22  | 8   | 21  |
| 6    |                   | 5000 | 600 | 120 | 8   | 21  |
| 7    |                   | 1000 | 600 | 240 | 8   | 21  |
| 8    |                   | 10000| 600 | 321 | 8   | 21  |
| 9    |                   | 5000 | 11  | 29  | 8   | 21  |
| 10   |                   | 10000| 600 | 321 | 8   | 21  |
| 11   |                   | 1000 | 10  | 32  | 7   | 18  |
| 12   |                   | 5000 | 6   | 18  | 6   | 18  |
| 13   |                   | 10000| 6   | 20  | 6   | 20  |
| 14   |                   | 1000 | 3   | 7   | 3   | 7   |
| 15   |                   | 5000 | 3   | 7   | 3   | 7   |
| 16   |                   | 1000 | 3   | 7   | 3   | 7   |
| 17   | Diagonal 4        | 50000| 4   | 10  | 4   | 10  |
| 18   |                   | 10000| 4   | 9   | 4   | 9   |
| 19   |                   | 1000 | 19  | 56  | 6   | 18  |
| 20   |                   | 5000 | 22  | 65  | 6   | 18  |
| 21   |                   | 10000| 22  | 65  | 6   | 18  |
| 22   | Recipe            | 50000| 22  | 65  | 6   | 18  |
| 23   |                   | 10000| 22  | 65  | 6   | 18  |
| 24   |                   | 1000 | 6   | 15  | 5   | 13  |
| 25   |                   | 5000 | 6   | 15  | 5   | 13  |
| 26   |                   | 10000| 7   | 17  | 6   | 15  |
| 27   |                   | 50000| 7   | 17  | 6   | 15  |
| 28   | DIXMAANA          | 10000| 7   | 17  | 6   | 16  |
| 29   |                   | 1000 | 12  | 29  | 10  | 25  |
| 30   | Shallwo           | 5000 | 12  | 29  | 10  | 25  |
| 31   |                   | 10000| 13  | 31  | 11  | 27  |

7. Conclusion
In this work, we prove the sufficient descent property and the global convergence property of the newly proposed spectral conjugate gradient method through strong Wolfe–Powell line search. The numerical results show that SCG algorithm outperforms BZA conjugate gradient method in terms in the number of iterations and the number of function evaluations.

Acknowledgments
The author expresses their gratitude and thanks to the encouragement and support of the College of Computer Sciences and Mathematics, the University of Mosul.

References
[1] M Mamat, I M Sulaiman, M Maulana, Sukono, and Z A Zakaria 2020 An Efficient Spectral Conjugate Gradient Parameter with Descent Condition for Unconstrained Optimization *Journal of Advance Research in Dynamical & Control Systems* 12 pp 2487-2493

[2] M R Hestenes and E Stiefel 1952 Methods of conjugate gradients for solving linear systems *Journal of research of the National Bureau of Standards* 49 pp 409-436

[3] R Fletcher and C M Reeves 1964 Function minimization by conjugate gradients *The computer journal* 7 pp 149-154

[4] E Polak and G Ribiere Note sur la convergence de méthodes de directions conjuguées 1969 *ESAIM: Mathematical Modelling and Numerical Analysis-Méthodisation Mathématique et Analyse Numérique* 3 pp 35-43

[5] B T Polyak 1969 The conjugate gradient method in extreme problems *USSR Computational Mathematics and Mathematical Physics* 9 pp 94-112

[6] R Fletcher 1987 *Practical methods of optimization* John and Sons, Chichester

[7] Y Liu and C Storey 1991 Efficient generalized conjugate gradient algorithms, part 1: theory *Journal of optimization theory and applications* 69 pp 129-137

[8] Y H Dai and Y Yuan 1999 A nonlinear conjugate gradient method with a strong global convergence property *SIAM Journal on optimization* 10 pp 177-182

[9] M Malik, M Mamat, S S Abas, I M Sulaiman and Sukono 2020 A new spectral conjugate gradient method with descent condition and global convergence property for unconstrained optimization *Journal of Mathematics and Computer Science* 10 pp 2053-2069

[10] J Barazilai and J M Borwein 1988 Two-point step size gradient methods *IMA Journal of Numerical Analysis* 8 pp 141-148

[11] M Raydan 1997 The Barazilai and Borwein gradient method for the large scale unconstrained minimization problem *SLAM Journal on Optimization* 7 pp 26-33

[12] E Birgin and M J Martines 2001 A spectral conjugate gradient method for unconstrained optimization *Applied Mathematical Optimization* 43 pp 117-128

[13] N Andrei 2007 Scaled conjugate gradient algorithms for unconstrained optimization *Computational Optimization and Applications* 38 pp 401-416

[14] H Jiang, S Deng, X Zheng and Z Wan 2012 Global gradient method *Journal of Applied Mathematics*

[15] L Zhang, W Zhou and D Li 2006 A descent modified Polak-Ribiere-Polyak conjugate method and its global convergence *IMA Journal of Numerical Analysis* 26 pp 629-640

[16] M Dawahdeh, I M Sulaiman, M Rivaie and M Mamat 2020 A new spectral conjugate gradient method with strong Wolfe-Powell line search *International Journal of Emerging Trends in Engineering Research* 8 pp 391-397

[17] J Lie and Y Jiang 2012 Global convergence of a spectral conjugate gradient method for unconstrained optimization *Hindawi Publishing Corporation Abstract and Applied Analysis* 12 Article ID 758287

[18] J K Liu, X Du and K Wang 2012 A mixed spectral CD-DY conjugate gradient method *Journal of Applied Mathematics* Article ID 569795-1

[19] G M Al-Naemi 2018 A Global Convergence of Spectral Conjugate Gradient Method for Large Scale Optimization *Journal of Education and Science for Pure Science* 27 pp 143-162

[20] U A Yakubu, M Mamat M A Mohamed and M Rivaie 2018 Secant condition free of a spectral PRP conjugate gradient method *International Journal of Engineering and Technology* 7 pp 325-328

[21] U A Yakubu, A Igudab, A V Mandarac and S Murtadal 2019 Scalar Parameter of a Spectral PRP Conjugate Gradient Method for Unconstrained Optimization *Malaysian Journal of Computing and Applied Mathematics* 2 pp 34-41
[22] B Baluch, Z Salleh and A Alhawarat 2018 A new modified three-term Heatenes-Stiefel conjugate gradient method with sufficient descent property and its global convergence Hindawi Journal of Optimization 2018 pp 1-13

[23] G M Al-Naemi 2013 Modified Nonlinear Conjugate Gradient Algorithms with Application in Neural Networks LAP LAMBERT Academic Publishing

[24] G Zoutendijk 1970 Nonlinear programming, computational methods Integer and nonlinear programming 143 pp 37-86

[25] Y Dai and L Z Liao 2001 New conjugacy conditions and related nonlinear conjugate gradient methods Applied Mathematics and Optimization 43 pp 87-101 DOI: 10.1007/s10492-0001-0001

[26] I Bongartz, A R Conn, N Gould and P L Toint 1995 ACM Transactions on Mathematical Software 21 pp 123-160

[27] N Andrei 2008 Advance Modelling and Optimization 10 pp 147-161