Good Characteristics of The New Spectral Conjugate Gradient Method for Unconstrained Optimization

Ahmed Hussien Sheekoo 1, Ghada M. Al-Naemi 2
1,2Department of Mathematics, College of Computer Science and Mathematics, University of Mosul, Mosul, Iraq.
E-mail: ahmed.csp106@student.uomosul.edu.iq, 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, we propose a new SCG method in which performance is numerically analyzed. We established the descent property and global convergence conditions based on assumptions through the strongWolfe-Powell line search. Numerical results were performed using benchmark functions widely used in many conventional functions to evaluate the efficiency of the proposed method.

Subject Classification: 90C30, 90C06, 65K05, 65K10.
Keywords: Spectral conjugate gradient, Unconstrained optimization, Global convergence, Sufficient descent condition, Strong Wolfe-Powell line search.

1. Introduction
The CG method is widely used for optimization due to its fast convergence speed, low storage capacity, and simple iterations [1]. CG- method is an iterative method for solving nonlinear unconstrained optimization problems. We can give a roughly clear definition of unconstrained optimization problem

\[ \min f(x), \quad x \in \mathbb{R}^n \]  

where \( f: \mathbb{R}^n \to \mathbb{R} \) is a continuous differential nonlinear function, and its gradient is represented by \( \nabla f(x_k) \). Start point \( x_0 \in \mathbb{R}^n \), and then calculate the follow-up point through an iterative process. The calculation formula for the new point is as follows:

\[ x_{k+1} = x_k + \alpha_k d_k, \quad \forall k \geq 0 \]  

(2)

Where \( \alpha_k \) is a positive step length calculated by performing some line search, and \( d_{k+1} \) is the search direction defined by the following:

\[ d_k = \begin{cases} -\nabla f(x_k), & k = 0 \\ -\nabla f(x_k) + \beta_k d_{k-1}, & k \geq 1 \end{cases} \]  

(3)
where $\beta_k \in R$ is a scalar. The six main forms of $\beta_k$ are in include Hestenes-Stiefel (HS) [2], Fletcher-Reeever (FR) [3], Polak-Ribiere-Polyak (PR) [4-5], Conjugate Descent-Fletcher (CD) [6], Liu-Storey (LS) [7], Dai-Yuan (DY) [8]. The parameters of these $\beta_k$ are as follows:

$$\beta_k^{HS} = \frac{g_k^T(g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}; \quad \beta_k^{FR} = \frac{g_k^T g_k}{||g_{k-1}||^2}; \quad \beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{||g_{k-1}||^2};$$

$$\beta_k^{CD} = -\frac{g_k^T g_k}{g_{k-1}^T d_{k-1}}; \quad \beta_k^{LS} = \frac{g_k^T y_{k-1}}{g_{k-1}^T d_{k-1}}; \quad \beta_k^{DY} = \frac{g_k^T g_k}{y_{k-1}^T d_{k-1}}.
$$

Where $y_{k-1} = g_k - g_{k-1}$, many authors have studied the convergence of CG-method under different line searches, and some have calculated the step size by exact line search (ILS). Others use another line search called strong Wolfe line search condition (SWL) is defined by:

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \rho \alpha_k g_k^T d_k,$$

$$|g(x_k + \alpha_k d_k)^T d_k| \leq -\sigma g_k^T d_k$$

(4)

Where $0 < \rho < 0.5 < \sigma < 1$.

Another well-known method that can be used to solve the problem (1) is the spectral CG-method (SCG) which was first proposed by Barzilai and Borwein [9]. The direction $d_{k+1}$ is defined by:

$$d_k = \begin{cases} g_k, & k = 0 \\ -\theta_k g_k + \beta_k d_{k-1}, & k \geq 1 \end{cases}$$

(5)

Where $\theta_k$ is the spectral gradient parameter? Later, Raydan proposed the SCG-method for large-scale unconstrained optimization problems in [10]. The attractive feature of the SCG-method is that it only needs gradient directions required in each line search to ensure global convergence. Surprisingly, in many known problems, the SCG-method outperforms the sophisticated CG method. Under some reasonable assumptions, Martinez and Birgin [11] concluded that their SCG-method is globally convergent. However, there is no to generate that the SCG-method will generate the correct direction. Therefore, Andrei [12] proposed the descending SCG-algorithm under Wolfe line search, Jiang et al. [13] based on the improved CG algorithm proposed by Zhang et al., the SCG-method with sufficient descent characteristics is designed by many authors [14-18].

2. A New SCG-method and the descent property

In this part, we will compute a new spectral parameter $\theta_k$. The search direction of SCG-method is usually as follows:

$$d_k = -\theta_k g_k + \beta_k d_{k-1}, \forall k \geq 1$$

(6)

Baluch and et al. [19] proposed the following formula:

$$\beta_k^{BZA} = \frac{g_k^T y_{k-1}}{(d_k^T y_{k-1} + \mu |g_k^T d_{k-1}|)}, \mu = 2$$

(7)

Now multiply both sides the second part of (6) by $y_{k-1}$, we will get

$$d_k^T y_{k-1} = -\theta_k g_k^T y_{k-1} + \beta_k d_{k-1}^T y_{k-1}.$$  

(8)
However, actual algorithms usually use adopt inexact line searches instead of exact line searches. Recently, Dai and Liao [20] replaced the conjugation condition 

\[ d_k^T g_{k-1} = g_k^T s_{k-1} \]

by the condition:

\[ \theta_k g_k^T y_{k-1} = t g_k^T s_{k-1} + \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|} d_{k-1}^T y_{k-1} \]

Substituting (7) and (9) in (8), we get:

\[ -t g_k^T s_{k-1} = \theta_k g_k^T y_{k-1} + \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|} d_{k-1}^T y_{k-1} \]

\[ \Rightarrow \theta_k g_k^T y_{k-1} = t g_k^T s_{k-1} + \frac{g_k^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|} d_{k-1}^T y_{k-1} \]

\[ \Rightarrow \theta_k = \frac{t g_k^T s_{k-1} + \frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|}}{\theta_k g_k^T y_{k-1}} \]  

We have \( s_{k-1} = \alpha_k d_{k-1} \), the above equation becomes

\[ \theta_k = \frac{t g_k^T s_{k-1} + \frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|}}{\theta_k g_k^T y_{k-1}} \]

\[ \Rightarrow \theta_k = \frac{t g_k^T s_{k-1} + \frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|}}{\theta_k g_k^T y_{k-1}} \]

\[ \Rightarrow 0 \leq \frac{t g_k^T s_{k-1} + \frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |g_k^T d_{k-1}|}}{\theta_k g_k^T y_{k-1}} \leq \mu = 0.01 \text{ and } t = 0.1 \]

Note that, if \( 1 < \theta_k < 0 \), then we put \( \theta_k = 1 \). i.e. (5) reduce to (3).

### The New Descent (SCG) Algorithm:

**Step 0:** Given an initial point \( x_0 \in R^n \), \( \varepsilon = 10^{-6} > 0, \mu = 0.01, t = 0.1 \).

Let \( d_0 = -g_0 \), set \( k = 1 \).

**Step 1:** If \( ||g_k|| \leq \varepsilon \), terminate. Otherwise, go to step 2.

**Step 2:** Determine a step length \( \alpha_k \) by (4).

**Step 3:** Generate new points through (2), calculate the gradient \( g_{k+1} = g(x_{k+1}) \), test

\[ ||g_{k+1}|| \leq \varepsilon, \text{ terminate; otherwise, continue.} \]

**Step 4:** Calculate the spectral parameter \( \theta_k^{new} \) by (10), if \( 1 < \theta_k < 0 \), put \( \theta_k = 1 \); Else, evaluate the conjugate parameter \( \theta_k^{RZA} \) represented by (7).

**Step 5:** The direction \( d_k \) defined in (5).

**Step 6:** If the Powel restart criteria

\[ ||g_k^T g_{k-1}|| \geq 0.2 ||g_k||^2 \]

is satisfied, set \( d_k = -g_k \) go to 2; otherwise, continue.

**Step 7:** Put \( k = k + 1 \) and go to 2.
Now we go to the proof of the descent condition of the algorithm for the proposed parameters:

**Theorem (1)**

Assume that SCG-method with search direction (6) and parameter $\beta_k^{BZA}$ defined in (7), and the step size $\alpha_k$ is obtained by SWL. The sufficient descent property

$$g_k^T d_k \leq -\eta \|g_k\|^2, \eta > 0, \forall k \geq 0$$  \hspace{1cm} (12)

**Proof:** To prove this assertion, we will use mathematical induction, if $k=0$, then $g_0^T d_0 = -\|g_0\|^2$. Therefore, condition (12) holds true. Now we assume that $k \geq 0$ is correct. Condition (12) also hold, now multiply both sides of (6) by $g_k^T$, we get

$$g_k^T y_{k-1} = g_k^T (g_k - g_{k-1}) = \|g_k\|^2 - g_k^T g_{k-1}$$  \hspace{1cm} (14)

Used one side of (11) in (14), we get

$$g_k^T y_{k-1} \leq \|g_k\|^2 + 0.2\|g_k\|^2$$  \hspace{1cm} (15)

And used another sides of (11) in (14), we get

$$g_k^T y_{k-1} \geq \|g_k\|^2 - 0.2\|g_k\|^2$$  \hspace{1cm} (16)

Since, $d_{k-1}^T g_{k-1} < 0$ it follows that $d_{k-1}^T g_k = y_{k-1}^T d_{k-1} + d_{k-1}^T g_{k-1} < y_{k-1}^T d_{k-1}$.

i.e. $d_{k-1}^T g_k < y_{k-1}^T d_{k-1}$  \hspace{1cm} (17)

By (17), $y_{k-1}^T d_{k-1} = (g_k - g_{k-1})^T d_{k-1} + g_{k-1}^T g_{k-1}$ and SWL, we get

$$-(1-\sigma)g_{k-1}^T d_{k-1} \leq y_{k-1}^T d_{k-1} \leq -(1+\sigma)g_{k-1}^T d_{k-1}.$$  \hspace{1cm} (18)

Since,

$$g_{k-1}^T d_{k-1} \leq -c\|g_{k-1}\|^2 \Rightarrow g_{k-1}^T d_{k-1} \geq c\|g_{k-1}\|^2$$  \hspace{1cm} (19)

Setting equation (11), (15-19) in equation (13)

$$g_k^T d_k \leq -\left(\frac{\alpha_k(-\sigma g_{k-1}^T d_{k-1})}{0.8\|g_k\|^2} - \frac{(1+\sigma)g_{k-1}^T d_{k-1}}{-\mu(1-\sigma)\|g_{k-1}^T d_{k-1}\|}\right)\|g_k\|^2 +$$

$$\frac{1.2\|g_k\|^2}{-(1-\sigma)g_{k-1}^T d_{k-1} + \mu g_{k-1}^T d_{k-1}}(-\sigma g_{k-1}^T d_{k-1})$$

$$\leq -\left(\frac{\alpha_k\sigma\|g_{k-1}\|^2}{0.8\|g_k\|^2} + \frac{(1+\sigma)}{(1-\sigma)\mu\|g_{k-1}^T d_{k-1}\|} + \frac{1.2\sigma}{(1-\sigma)\mu}\right)\|g_k\|^2$$
\[ = - \left( \frac{\nabla f(x_0)}{\nabla f(x_1)} \right)^2 \]

\[ 0 < \tilde{\gamma} < \| g_k \| \leq Y \& \ 0 < \tilde{\omega} \leq \| g_{k-1} \| \leq \omega. \quad [21]. \]

\[ g_k^T d_k \leq -\left( \frac{tb \sigma \omega^2}{0.8 \tilde{\gamma}^2} + \frac{1+2 \sigma}{1-(1+\mu)\sigma} \right) \| g_k \|^2 \]

We have, \( 0 < \alpha_k \leq b, \sigma = 0.9, \mu = 0.01, \Rightarrow 1 - (1 + \mu)\sigma > 0, t = 0.1, \)

Let \( \eta = \left( \frac{tb \sigma \omega^2}{0.8 \tilde{\gamma}^2} + \frac{1+2 \sigma}{1-(1+\mu)\sigma} \right) > 0, \)

so

\[ g_k^T d_k \leq -\eta \| g_k \|^2 < 0. \quad (21) \]

The new SCG-method satisfies the sufficient descent condition.

**Assumption (A).**

1- \( f(x) \) is bounded on the level set \( \Psi = \{ x \in R^i, f(x) \leq f(x_0) \} \), where \( x_0 \) is the starting point. i.e., there is a constant \( \tau > 0 \), which means \( \| x_k \| \leq \tau \forall x \in \Psi \).

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 \( \| g(x) - g(y) \| \leq L \| x - y \| \), \( \forall x, y \in N \).

The following Lemma is used to prove the global convergence which is proposed by Zoutendijk [22].

**Lemma (1) [22]**

Assume that Assumption (A) holds. Suppose a general iterative method (2), and the direction (6) is descent direction and the steplength \( \alpha_n \) is obtained by SWC (4). Then, Zoutendijk condition holds, i.e.

\[ \sum_{k=1}^{\infty} \frac{1}{\| d_k \|^2} < \infty \]

(23)

We can establish the global convergence of the proposed SCG-method based on Lemma (1).

**Theorem (2)**

Assume that assumption (A) holds, and let the sequence sequences \( \{ x_k \} \) and \( \{ d_k \} \) generated by the algorithm SCG, where \( \alpha_k \) is obtained by SWC (4). Then

\[ \lim \inf_{k \to \infty} \| g_k \| = 0 \]

(24)

Or

\[ \sum_{k=1}^{\infty} \frac{1}{\| d_k \|^2} < +\infty \]

(25)

**Proof:**

By contradiction, assume that the conclusion is not true. It must be proven that \( \| d_k \| \) it bounded above. from (6)

\[ \| d_k \| = \| -\theta_k g_k + \beta_k d_{k-1} \| \]
\[
\leq |\theta_k| \cdot \|g_k\| + |\beta_k| \cdot \|d_{k-1}\| \tag{26}
\]

\[
|\theta_k| = \left| - \left( \frac{\tau_k g_k^T d_{k-1}}{g_k^T g_{k-1}} + \frac{d_{k-1}^T y_{k-1}}{d_{k-1}^T y_{k-1} + \mu |\theta_k| d_{k-1}|} \right) \right| \tag{27}
\]

By using SWL, (16) and (18) in (27), we get

\[
\|d_k\| \leq \left| \frac{\tau_k g_k^T d_{k-1}}{0.8 \|g_k\|^2} + \frac{(1+\sigma)g_k^T d_{k-1}}{-\frac{1}{(1-\sigma) - \mu(\sigma)}} \right| \leq \left| \frac{\tau_k g_k^T d_{k-1}}{0.8 \|g_k\|^2} - \frac{(1+\sigma)}{\frac{1}{(1-\mu)\sigma}} \right| = B_1. \tag{28}
\]

Baluch and et al. [18], prove that \(|\beta_k| \leq A \frac{\|g_k\|}{\|d_{k-1}\|}\), where \(A = \frac{\mathcal{L}}{(1-\sigma)\alpha_k}\), \(\mathcal{L} \& c > 0\), \(\sigma = 0.5\) and \(\alpha_k > 0\), so that \(A > 0\).

\[
|\beta_k| \leq A \frac{\|g_k\|}{\|d_{k-1}\|} = B_2. \tag{29}
\]

Now, putting (28) and (29) in (26). We have

\[
\|d_k\| \leq B_1 \cdot \|g_k\| + B_2 \cdot \|d_{k-1}\| \\
\leq B_1 \cdot Y + B_2 \beta = \varphi. \\
\Rightarrow \sum_{k=1}^{\infty} \frac{1}{\|d_k\|^2} \geq \frac{1}{\varphi^2} \sum_{k=1}^{\infty} 1 = + \infty
\]

Where \(\varphi\) is an arbitrary scalar? This contradicts (25). Therefore, according to definition (2), we say that the method satisfies the global convergence condition.

**Results and Discussion:**

In this part, we will report the results of several test functions. Some test functions were selected to analyze the new method. These functions are considered from CUTEr [23], Andrei [24]. Using SWP line search, according to the sum of a number of iterations (SNOI) and the sum of a number of function evaluation (SNOF), the new SCG-method, the classic BZACG-method, and LS-method are compared. All symbols are written in double-precision FORTRAN 77 language and collected as visual format (F6.6). The new SCG-method is implemented using the SWP line search (4), with \(\rho = 0.001, \sigma = 0.9, \mu = 0.01\) and \(t = 0.1\), we tested 25 well-known test functions, the dimensions of which are (1000, 5000, 10000, 50000, and 100000). The stopping criterion of this algorithm is \(|g_{n+1}| \leq 10^{-6}\), or if \(NOI \geq 600\), the method is considered unsuccessful. It can be seen from the following Table (1) that the results obtained by the newly proposed method are better than the other methods that we pointed out in the table.
Table 1. the comparison between the SCG-method and the classical BZACG, PR-method and FR-method

| No. | SBZA-method  | Classical algorithm | BZA-method | PR-method | FR-method |
|-----|--------------|---------------------|------------|-----------|-----------|
|     | SNOI | SNOF | SNOI | SNOF | SNOI | SNOF | SNOI | SNOF | SNOI | SNOF |
| 1   | 130  | 285  | 150  | 358  | 150  | 390  | 150  | 390  |
| 2   | 159  | 363  | 297  | 629  | 159  | 363  | 147  | 334  |
| 3   | 55   | 135  | 55   | 135  | 55   | 140  | 59   | 143  |
| 4   | 275  | 859  | 1164 | 2894 | 336  | 1204 | 391  | 1532 |
| 5   | 25   | 57   | 74   | 153  | 25   | 55   | 25   | 55   |
| 6   | 80   | 220  | 95   | 255  | 80   | 220  | 80   | 225  |
| 7   | 10   | 27   | 15   | 37   | 10   | 30   | 10   | 30   |
| 8   | 27   | 69   | 27   | 69   | 28   | 71   | 28   | 71   |
| 9   | 30   | 79   | 30   | 75   | 30   | 79   | 29   | 77   |
| 10  | 20   | 50   | 20   | 50   | 20   | 50   | 20   | 50   |
| 11  | 40   | 115  | 47   | 131  | 1227 | 1195 | 3000 | 4780 |
| 12  | 150  | 315  | 165  | 356  | 205  | 1812 | 196  | 415  |
| 13  | 30   | 75   | 35   | 90   | 30   | 70   | 30   | 75   |
| 14  | 30   | 90   | 32   | 100  | 79   | 255  | 40   | 120  |
| 15  | 80   | 220  | 95   | 255  | 80   | 220  | 80   | 225  |
| 16  | 122  | 254  | 159  | 302  | 124  | 258  | 117  | 244  |
| 17  | 51   | 127  | 50   | 130  | 51   | 132  | 52   | 133  |
| 18  | 150  | 380  | 154  | 398  | 150  | 380  | 150  | 380  |
| 19  | 35   | 95   | 47   | 124  | 45   | 128  | 45   | 110  |
| 20  | 5    | 20   | 5    | 20   | 5    | 20   | 5    | 20   |
| 21  | 517  | 1036 | 557  | 1130 | 575  | 1167 | 550  | 1110 |
| 22  | 219  | 1142 | 208  | 1421 | 247  | 1396 | 229  | 1165 |
| 23  | 31   | 81   | 31   | 79   | 31   | 81   | 30   | 79   |
| 24  | 28   | 91   | 31   | 95   | 29   | 93   | 31   | 101  |
| 25  | 10   | 35   | 10   | 35   | 10   | 35   | 10   | 35   |
| Total | 2309 | 6220 | 3553 | 9321 | 3781 | 9844 | 5504 | 11899 |

The results of the summation for NOI and NOF and the comparison between the algorithms in the table (1) can be displayed by the following graphs:
Table 2. shows the performance percentage of the proposed methods SBZA-method relative to the FR, BZA, and PRP- methods.

Table 3. The percentage performance of the proposed methods

| Measures | SBZA&BZA-method | SBZA&PR-method | SBZA&FR-method |
|----------|-----------------|----------------|-----------------|
| NOI      | 35.02%          | 38.94%         | 58.05%          |
| NOF      | 33.27%          | 36.82%         | 47.73%          |

The following diagram is an explanation of Table 2.
Conclusions
In this work, a new spectral conjugate gradient method is proposed. An attractive feature of the proposed method is that it can generate sufficient descent conditions, regardless of the line search. Under strong Wolfe-Powell line search condition, the global convergence of the method has been established. Numerical results of SBZA, BZA, PRP, and FR methods.

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] Jardow, F.N. and Al-Naemi, G.M. 2021. A new hybrid conjugates gradient algorithm as a convex combination of MMWU and RMIL nonlinear problems, Journal of Interdisciplinary Mathematics. 24(3):637-655. DOI: 10.1080/09720502.2020.1815346.
[2] Hestenes, M. R. and Stiefel, E., 1952. Methods of conjugate gradients for solving linear systems, Journal of Research of the National Bureau of Standards, 49: 409–436. https://nvlpubs.nist.gov/nistpubs/jres/049/jresv49n6p409_A1b.pdf
[3] Fletcher, R. and Reeves, C. M. 1964. Function minimization by conjugate gradients, The Computer Journal, 7: 149–154. https://academic.oup.com/comjnl/article/7/2/149/335311
[4] Polak, E. and Ribiére, G. 1969. Note Sur la convergence de directions conjugates, Rev, Revue Française d'Informatique et de Recherche Opérationnelle, 3(16): 35–43. http://www.numdam.org/item?id=M2AN_1969__3_1_35_0
[5] Polyak, B. T. 1969. The conjugate gradient method in extreme problems, USSR Computational Mathematics and Mathematical Physics, 9(4): 94–112. https://www.researchgate.net/publication/222365587
[6] Fletcher, R. 1997. Practical Method of Optimization (I): Unconstrained Optimization, Wiley, New York, NY, USA, 2nd edition. https://www.wiley.com/en-ba
[7] Liu, Y. and Story, C. 1991. Efficient generalized conjugate gradient algorithms. I. Theory, Journal of Optimization Theory and Applications, 69(1): 129–137. https://link.springer.com/article/10.1007/BF00940464
[8] Dai, Y. H. and. Yuan, YA. 2000. Nonlinear conjugate gradient with a strong global convergence property, SIAM Journal on Optimization, 10(1): 177–182. https://doi.org/10.1137/S1052623497318992
[9] Barazilai, J. and Borwein, J. M. 1988. Two-point step size gradient methods, IMA Journal of Numerical Analysis, 8(1): 141-148. https://aip.scitation.org/doi/abs/10.1063/1.4952138
[10] Raydan, M. 1997. The Barzilai and Borwein gradient method for the large-scale unconstrained minimization problem, SIAM Journal on Optimization, 7(1): 26-33. https://doi.org/10.1137/S1052623494266365
[11] Birgin and M. J. 2001. Martines, A spectral conjugate gradient for unconstrained optimization. Applied Mathematical Optimization, 43: 117-128. https://doi.org/10.1007/s00245-001-0003-0
[12] And, A. 2007. Scaled conjugate gradient algorithm for unconstrained optimization. Computational Optimization and Application, 38(3):401-416.
[13] Jiang, H., Deng, S., Zheng, X. and Wan, Z. 2012. Global convergence of a modified spectral conjugate gradient method. *Journal of Applied mathematics*. **2012**: 1-13.

https://doi.org/10.10155/2012/641276

[14] Zhang, L., Zhou, W. and Li, D. 2006. A descent modified Polak-Ribiere-Polyak conjugate method and its global convergence. *IMA Journal of Numerical Analysis*, **26**: 629-640. https://doi.org/10.10109/IMANUM%2FDRL016

[15] Sheekoo, A.H. and Al-Naemi, G.M. Global convergence Condition for a New Spectral Conjugate Gradient Method for Large-Scale Optimization, to appear in IOP conference series in 2021.

[16] Sulaiman, M. I. Usman, A. Y. and Mustafa, M. 2020. Application of spectral conjugate gradient methods for solving unconstrained optimization problems. *An International Journal of Optimization and Control: Theories & applications*. **10**(2):198-205. https://doi.org/10.11121/ijocta.01.2020.0085.

[17] Jian, J., Yang, L., Jiang, X., Liu, P. and Liu, M. 2020. A spectral conjugate gradient method with descent property. *Mathematics*. **8**(280): 1-13.

[18] Malik, M., Mamat, M, Abas, S., Sulaiman, I. and Sukono. 2020. A new spectral conjugate gradient method with the decent condition and global convergence property for unconstrained optimization. Journal of Computer Science. 10(5): 2053-2069. http://doi.org/10.28919/jmcs/4829. ISSN: 1927-5307.

[19] Baluch, B., Salleh, and Alhawarat, Z. A. 2018. A New Modified Three-Term Hestenes–Stiefel Conjugate Gradient Method with Sufficient Descent Property and Its Global Convergence, *Journal of Optimization*. **2018**: 1-13. https://doi.org/10.1155/2018/5057096.

[20] Dai, Y. and Liao, L. Z. 2001. New conjugacy conditions and related nonlinear conjugate gradient methods. *Applied Mathematics and Optimization*. **43**: 87-101, 2001. https://link.springer.com/article/10.1007/s002450010019.

[21] Jardow, F.N and Al-Naemi, G.M. 2020. A New Hybrid Conjugate Gradient Algorithm for Unconstrained Optimization with Inexact line search, *Indonesian Journal of Electrical Engineering and Computer Science*. 20(2): 939-947. doi: 10.1155/ijeecs.v20.i2.pp939-947. ISSN: 2502-4752.

[22] Zoutendijk, G. 1970. Nonlinear programming, computational methods. *Integer and nonlinear programming*. **143**: 37-86. https://www.oalib.com/references/13789575.

[23] Bongartz, I., Conn, A. R., Gould, N. and Toint, P. L. 1995. ACM Transactions on Mathematical Software, **21**: 123-160. https://dl.acm.org/doi/10.1145/200979.201043.

[24] Andrei, N. 2008. An Unconstrained Optimization Test Functions Collection. *Adv. Model. Optim*. **10**(1): 147-161. http://www.apmath.spbu.ru/cnsa/pdf/obzor/.pdf.