*I. J. Mathematical Sciences and Computing,* **2022, 2, 1-10**
Published Online June 2022 in MECS (http://www.mecs-press.org/)
DOI: 10.5815/ijmsc.2022.02.01

# A Hybrid Spectral Conjugate Gradient Method with Global Convergence

**Jing Li**
School of Mathematics and Information Science, Henan Polytechnic University, Henan, China
E-mail: lijing960223@163.com

**Shujie Jing**
School of Mathematics and Information Science, Henan Polytechnic University, Henan, China
E-mail: jsj-jjj@hpu.edu.cn

Received: 30 September 2021; Revised: 21 October 2021; Accepted: 10 November 2021; Published: 08 June 2022

**Abstract:** The spectral conjugate gradient (SCG) method is one of the most commonly used methods to solve large- scale nonlinear unconstrained optimization problems. It is also the research and application hot spot of optimization theorists and optimization practitioners. In this paper, a new hybrid spectral conjugate gradient method is proposed based on the classical nonlinear spectral conjugate gradient method. A new parameter $\beta^{LW}$ is given. Under the usual assumptions, the descending direction independent of any line search is generated, and it has good convergence performance under the strong Wolfe line search condition . On a set of test problems, the numerical results show that the algorithm is effective.

**Index Terms:** Unconstrained optimization, Strong Wolfe line search, Descending condition, Spectral conjugate gradient method, Global convergence.

## 1. Introduction

Optimization theory is a young subject with a strong application. Both the basic unconstrained optimization problems and emerging evolutionary algorithm problems have been applied in real life. For the nonlinear programming problems in engineering technology, the iterative numerical solution is the most typical and common. Most of these algorithms are combined with computers. They can not only calculate complex nonlinear programming problems but also have high speed and precision. According to the certainty of the forward direction in the iterative process, this kind of method can be divided into two categories: random search method and deterministic method.

The random search method is a method that people imitate to solve problems according to its principle inspired by the laws of nature. The search process of this kind of method mainly uses the function value information rather than the gradient information of the function. For example, Anes A A et al. [1] used particle swarm optimization to explore the expansion effect of neural network imaging technology, Hussain A et al. [2] used natural selection, crossover and mutation rules in genetic algorithm to optimize numerical functions, Mamdouh M et al. [3] realized path optimization of Internet of things navigation system based on machine learning, Sharma N et al [4] instilled the global optimization of regional routing protocol with a new firefly algorithm

The deterministic method is divided into pattern search algorithm and gradient method according to the degree of using function information. The pattern search algorithm is mainly suitable for the situation of few variables, complex structure of objective function ,and difficult calculation of gradient function. In addition to the function value information, the gradient algorithm uses the gradient information of the current point or the generated point in the iterative process. Therefore, gradient algorithm generally has a fast convergence speed, and it is easier to establish the theoretical properties of the algorithm. Deterministic methods generally have fast convergence speed and good quadratic convergence, but they are slightly deficient in global convergence and stability, and deterministic methods have great shortcomings in computational complexity and storage consumption. For large-scale nonlinear programming problems, the running time of each iteration step of the algorithm will be relatively long, which will affect the overall efficiency of the algorithm.

As the basis of nonlinear optimization theory research, unconstrained optimization theory and method are widely used in many fields in real life. With the advent of the big data era, the dimension of optimization problems increases sharply, which brings development space for solving large-scale unconstrained optimization problems.

Our main research goal is to solve large-scale unconstrained optimization problems. Steepest descent method, Newton method, quasi-Newton method and conjugate gradient method are common solutions. Among them, the conjugate gradient method is widely used because of its fast convergence speed and small storage. As the extension and extension of the conjugate gradient method, the spectral conjugate gradient method can more effectively adjust the structure of the search direction, which has become a research hotspot in recent years. How to adjust the parameter structure more effectively is a topic of continuous discussion. We aim to obtain the spectral conjugate gradient method with fewer gradient iterations.

## 2. Basic Knowledge

The spectral conjugate gradient (SCG) method is one of the most commonly used methods for solving large-scale unconstrained optimization problems. The SCG method was first proposed by Birgin and Martinez [5] in combination with the conjugate gradient (CG) method and the spectral gradient method. For classical CG methods, please refer to the literature [6,7,8,9,10,11]. Now, consider the following nonlinear unconstrained optimization problem.

$$\min \ f(x) \tag{1}$$

where $x \in R^n$ is the decision variable and the objective function $f : R^n \to R$ is continuously differentiable. The CG method is the most commonly used method to solve the problem of formula (1), CG method plays an important role in solving practical problems because of its small storage space and fast convergence speed. The general form of iteration is:

$$x_{k+1} = x_k + \alpha_k d_k, k \geq 1 \tag{2}$$

Where the step factor $\alpha_k$ is a positive parameter and the search direction $d_k$ is defined by

$$d_k = \begin{cases} -g_k, & k = 1, \\ -\theta_k g_k + \beta_k d_{k-1}, & k > 1. \end{cases} \tag{3}$$

The scalar $\beta_k$ is the conjugate direction control parameter and $g_k = \nabla f(x_k)$ is the gradient of $f$ at a point $x_k$.

Among them, the spectral parameter

$$\theta_k = \frac{s_{k-1}^T s_{k-1}}{s_{k-1}^T y_{k-1}} , \tag{4}$$

the conjugate direction control parameter

$$\beta_k = \frac{g_k^T (\theta_k y_{k-1} - s_{k-1})}{d_{k-1}^T y_{k-1}}, \tag{5}$$

However, the search direction $d_k$ of the spectral conjugate gradient method does not satisfy the descent condition

$$g_k^T d_k < 0. \tag{6}$$

We know that the search direction satisfies sufficient descent, which is an important condition to judge the global convergence. In addition, for SCG, the setting of spectral parameters and CG parameters is worth discussing, which affects not only the numerical performance but also the global convergence.

Therefore, many scholars constantly modify the search direction of the SCG method, trying to converte or approximate to some descent directions. Inspired by Andrei's[12] additional assumption about the boundedness of spectral parameters, Deng [13]proposed an improved the SCG algorithm for Nonconvex unconstrained optimization problems. The SCG direction was approaching the quasi-Newton search direction. Under certain mild conditions, global convergence was established. Dong[14] applied Perry's [15]idea and proposed a scale symmetric Perry conjugate gradient method with a restart program based on scaling technology and restart strategy. The memoryless BFGS method and scaling method were used respectively. Under Wolfe line search, the global convergence of uniformly convex and nonconvex functions was proved. Zhang[16] proposed an adaptive scaling parameter to overcome the shortcomings of original selection in Birgin and Martinez [5] and established the global convergence of the SCG method and new parameters for convex and non-convex objective functions.

Other scholars had obtained an extended conjugate condition with some versions of the section equation, For example, Peyghami et al[17] Proposed an improved secant equation, which applies a new updated Yabe and Takano rules as an adaptive version of conjugate gradient parameters. Using this modified crosscutting equation, Nezhadhosein [18] proposed a new modified descent spectrum nonlinear conjugate gradient method. Using the high-order accuracy and sufficient descent conditions of approximating the second-order curvature information of the objective function, the global convergence of the method for uniformly convex functions and general functions was proved. Sheekoo[19] used an improved trinomial HS method constructed by Baluch et al[20]. By adding $\mu\left|g_{n+1}^{T}d_{n}\right|$ to the denominator of $\beta_k^{HS}$, the appropriate parameter $\varphi$ was expressed by (SCGBZA). Under the usual assumptions, the descent property and global convergence of SCGBZA were proved.

The hybridization of the above two methods has become more and more research hotspot. That is the descent condition and the improved conjugate condition. By adjusting the parameters, the appropriate spectral coefficients and conjugate direction control parameters are obtained. The SCG method with sufficient descent properties is obtained as far as possible, and the global convergence is proved under general assumptions. The theoretical properties are guaranteed to illustrate the feasibility of the method. Then, the numerical results show the effectiveness of the method.

Next, inspired by the relevant literature, we construct a new parameter and algorithm. The theoretical properties and numerical results show that the algorithm is feasible.

## 3. New Algorithm and Its Descent

Alhawarat et al [21] proposed an improved version of Polak and Ribiere's [22]CG formula for the bounded Lipschitz problem. The CG method based on the PRP formula and Strong Wolfe Powell(SWP) Line Search in reference [23] is not globally convergent. Alhawarat et al proposed a nonnegative CG formula with new restart conditions $\beta_k^{AZPRP}$. The new formula is for parameters $\beta_k^{PRP}$ further modifications. When the new parameter adopted (SWP) line search or Weak Wolfe Powell(WWP) line search, it had global convergence, and the proof of sufficient descent condition of SWP Line search was given. The calculation $\beta_k$ was as follows:

$$\beta_k^{AZPRP}=\begin{cases}\dfrac{\|g_k\|^2-g_k^Tg_{k-1}}{\|g_{k-1}\|^2}, & \text{if } \|g_k\|^2>\left|g_k^Tg_{k-1}\right|,\\ \dfrac{\|g_k\|^2-\mu_k\left|g_k^Tg_{k-1}\right|}{\|g_{k-1}\|^2}, & \text{if } \|g_k\|^2>\mu_k\left|g_k^Tg_{k-1}\right|,\\ 0, & \text{otherwise.}\end{cases} \tag{7}$$

In the above formula

$$\mu_k=\frac{\|x_k-x_{k-1}\|}{\|y_{k-1}\|}, \tag{8}$$

Obviously,

$$0\le\beta_k^{AZPRP}\le\frac{\|g_k\|^2}{\|g_{k-1}\|^2}=\beta_k^{FR}, \tag{9}$$

Therefore, according to the argument put forward in [24], $\beta_k^{AZPRP}$ will inherit $\beta_k^{FR}$ all the advantages and features. It is proved that

$$c-2\le\frac{g_k^Td_k}{\|g_k\|^2}\le -c\ , \tag{10}$$

and

$$\sum_{k=0}^{\infty}\frac{\left(g_k^T d_k\right)^2}{\left\|d_k\right\|^2} < +\infty \ , \tag{11}$$

which shows that the sufficient descent condition and global convergence are satisfied.

Based on the idea of Wei[25] and Dai[26], Salih et al[27] proposed a modified parameter that satisfies the sufficient descent condition. The new parameter $\beta_k^{YHM} \geq 0$ had sufficient descent properties and satisfied the global convergence under strong Wolff line search. The calculation $\beta_k$ was as follows:

$$\beta_k^{YHM} = \begin{cases} \dfrac{g_k^T\left(g_k - g_{k-1}\right)}{\left\|g_{k-1}\right\|^2}, & \text{if } 0 \leq g_k^T g_{k-1} \leq \left\|g_k\right\|^2, \\ \dfrac{g_k^T\left(g_k - \dfrac{\left\|g_k\right\|}{\left\|g_{k-1}\right\|} g_{k-1}\right)}{\left\|g_{k-1}\right\|^2}, & \text{otherwise.} \end{cases} \tag{12}$$

It is proved that

$$\left(c-2\right)\left\|g_k\right\|^2 \leq g_k^T d_k \leq -c\left\|g_k\right\|^2, \tag{13}$$

and

$$\sum_{k=1}^{\infty}\frac{\left(g_k^T d_k\right)^2}{\left\|d_k\right\|^2} < \frac{L}{1-\sigma}\sum_{k=1}^{\infty}\left(-\alpha_k g_k^T d_k\right) < +\infty \ , \tag{14}$$

which shows that the sufficient descent condition and global convergence are satisfied.

Now, combined with the basic ideas of the above documents, we suggest a spectral conjugate gradient method named LW method, the parameter $\beta_k^{LW}$ is defined as follows:

$$\beta_k^{LW} = \begin{cases} \dfrac{\left\|g_k\right\|^2 - g_k^T g_{k-1}}{\max\left\{\left\|g_{k-1}\right\|^2, d_{k-1}^T \ y_{k-1} \ , \ -g_{k-1}^T \ d_{k-1}\right\}} \ , & \text{if } 0 < g_k^T g_{k-1} < \left\|g_k\right\|^2, \\ \dfrac{\left\|g_k\right\|^2}{\max\left\{\left\|g_{k-1}\right\|^2, d_{k-1}^T \ y_{k-1} \ , \ -g_{k-1}^T \ d_{k-1}\right\}} \ , & \text{otherwise.} \end{cases} \tag{15}$$

The search direction of this method is defined as follows:

$$d_k = \begin{cases} -g_k, & \text{if } \ k=1, \\ -\theta_k g_k + \beta_k^{LW} d_{k-1}, & \text{if } \ k>1. \end{cases} \tag{16}$$

where

$$\theta_k = 1 + \beta_k^{LW}\frac{d_{k-1}^T g_k}{\left\|g_k\right\|}, \tag{17}$$

In this work, we use line search called strong Wolfe line search condition (SWL)is defined by:

$$\begin{cases} f\left(x_k + \alpha_k d_k\right) \leq f\left(x_k\right) + \delta\alpha_k g_k^T d_k \\ \left|g_{k+1}^T d_k\right| \leq \sigma\left|g_k^T d_k\right| \end{cases}, \tag{18}$$

The corresponding algorithm is constructed below.

| Algorithm 2.1(LW Algorithm) |
|---|
| Input. Given an initial value $x_1 \in R^n$ , parameter $0<\delta<\sigma<1, \varepsilon>0$ . |
| Step1.Set $k=1$ , $d_1=-g_1$ . |
| Step2.If $\|g_k\|<\varepsilon$ , stop. |
| Step3.Determine the step length $\alpha_k$ satisfying(18) |
| Step4. Use the $x_{k+1}=x_k+\alpha_k d_k$ loop for the next iteration, $k=k+1$. |
| Step5. Calculate $g_{k+1}$ , if $\|g_{k+1}\|<\varepsilon$ , stop. |
| Step6. Use (15), (16) and (17) to calculate $\beta_k$ , $d_k$ and $\theta_k$ respectively. Turn Step3. |

**Lemma 2.1** Consider the sequence $\{g_k\}$ and $\{d_k\}$ generated by Algorithm 2.1.Then for any $k \geq 1$, the sufficient descent condition satisfy

$$g_k^T d_k=-\|g_k\|^2 \tag{19}$$

**Proof:** If $k=1$, then the search direction $d_k$ is given by $d_k=-g_k$,
so there is

$$d_1=-g_1, g_k^T d_k=g_1^T d_1=-\|g_1\|^2, \tag{20}$$

Otherwise, the search direction $d_k$ is given by

$$d_k=-\left(1+\beta_k^{LW} \frac{d_{k-1}^T g_k}{\|g_k\|}\right) g_k+\beta_k^{LW} d_{k-1}, \tag{21}$$

So, we get

$$\begin{aligned} g_k^T d_k &=-\left(1+\beta_k^{LW} \frac{g_k^T d_{k-1}}{\|g_k\|^2}\right)\|g_k\|^2+\beta_k^{LW} g_k^T d_{k-1} \\ &=-\|g_k\|^2 \end{aligned} . \tag{22}$$

Therefore, this method satisfies sufficient descent conditions for any k .

**Lemma 2.2** Let the sequences $\{g_k\}$ and $\{d_k\}$ be generated by Algorithm 2.1.Then for any $k \geq 1$, the relation $0<\beta_k^{LW}<\frac{\|g_k\|^2}{\|g_{k-1}\|^2}$ holds.

**Proof :** According to the definition of $\beta_k^{LW}$ in formula (15), if we want to prove that $\beta_k^{LW}>0$ is true, we only need to prove that $\max\left\{\|g_{k-1}\|^2, d_{k-1}^T y_{k-1}, -g_{k-1}^T d_{k-1}\right\}>0$ is true, that is to prove $d_{k-1}^T y_{k-1}>0$ , $-g_{k-1}^T d_{k-1}>0$ . We can easily obtain

$$d_{k-1}^T y_{k-1} \geq(1-\sigma)\|g_{k-1}\|^2, \tag{23}$$

Because of $0<\sigma<1$ , we can get the following inequality from equation (18) and (19)

$$d_{k-1}^T y_{k-1}>0, \tag{24}$$

Through the descent condition (19), we can get

$$-g_{k-1}^T d_{k-1}=\|g_{k-1}\|^2, \tag{25}$$

So

$$-g_{k-1}^T d_{k-1} > 0, \tag{26}$$

Therefore, we can easily get $\beta_k^{LW} > 0$ through the definition of formula (15), formula (24) and formula (26). In addition, if

$$0 < g_k^T g_{k-1} < \|g_k\|^2, \tag{27}$$

then

$$\begin{aligned} \beta_k^{LW} &= \frac{\|g_k\|^2 - g_k^T g_{k-1}}{\max\left\{\|g_{k-1}\|^2, d_{k-1}^T y_{k-1}, -g_{k-1}^T d_{k-1}\right\}} \\ &\le \frac{\|g_k\|^2}{\max\left\{\|g_{k-1}\|^2, d_{k-1}^T y_{k-1}, -g_{k-1}^T d_{k-1}\right\}} \end{aligned}. \tag{28}$$

Because of

$$\max\left\{\|g_{k-1}\|^2, d_{k-1}^T y_{k-1}, -g_{k-1}^T d_{k-1}\right\} \ge \|g_{k-1}\|^2, \tag{29}$$

So

$$\beta_k^{LW} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \tag{30}$$

that is, Lemma 2. 2 can be proved.

The proof of Lemma 2.1 shows that the SCG method constructed by Algorithm 2.1 meets the sufficient descent Condition(19). The establishment of the formula in Lemma 2.2 ensures that we prove that the algorithm meets the global convergence in the next section.

## 4. Global Convergence

To establish the convergence properties of the newly proposed formula, the following assumption is required.

**Assumption H ([28])**

(H1) The objective function $f(x)$ has a lower bound on the level set $\Omega = \left\{x \in R^n \mid f(x) \le f(x_1)\right\}$, where $x_1$ is the initial iteration point.

(H2) The objective function $f(x)$ is continuously differentiable in a certain neighborhood $\Lambda$ of the level set $\Omega$, and the gradient function $g(x)$ satisfies Lipschitz continuity, that is, there is a constant $L > 0$ which makes $\|g(x) - g(y)\| \le L\|x - y\|, \forall x, y \in \Lambda$ hold.

The following lemma is called the Zoutendijk condition[29], which can be used to analyze the global convergence property of the SCG method.

**Lemma 3.1** If Assumption H hold, search direction $d_k$ satisfies $g_k^T d_k < 0$, then sequences $\{g_k\}$ and $\{d_k\}$ generated by algorithm 1 satisfy Zoutendijk condition, then $\sum_{k\ge1} \frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} < +\infty$ holds.

**Theorem 3.1** If Assumption H hold, the sequence $\{g_k\}$ generated by Algorithm 2.1.Then

$$\liminf_{k\to\infty} \|g_k\| = 0. \tag{31}$$

**Proof:** Assume that (31) does not hold. There must be a constant $\gamma > 0$ ,make that

$$\|g_k\| \geq \gamma, k \geq 1. \tag{32}$$

Derived from formula (16)

$$d_k + \theta_k g_k = \beta_k^{LW} d_{k-1} \tag{33}$$

Square both sides at the same time

$$\left(d_k + \theta_k g_k\right)^T \left(d_k + \theta_k g_k\right) = \left(\beta_k^{LW}\right)^2 \|d_{k-1}\|^2, \tag{34}$$

Expanded

$$\|d_k\|^2 + 2\theta_k d_k^T g_k + \theta_k^2 \|g_k\|^2 = \left(\beta_k^{LW}\right)^2 \|d_{k-1}\|^2, \tag{35}$$

Transposition of terms

$$\|d_k\|^2 = \left(\beta_k^{LW}\right)^2 \|d_{k-1}\|^2 - 2\theta_k d_k^T g_k - \theta_k^2 \|g_k\|^2, \tag{36}$$

Divide both sides by $\left(g_k^T d_k\right)^2$ at the same time

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} = \left(\beta_k^{LW}\right)^2 \frac{\|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + 2\frac{\theta_k}{\|g_k\|^2} - \frac{\theta_k^2}{\|g_k\|^2}, \tag{37}$$

Because of

$$0 < \beta_k^{LW} < \frac{\|g_k\|^2}{\|g_{k-1}\|^2}, \tag{38}$$

So

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \leq \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \cdot \frac{\|d_{k-1}\|^2}{\left(g_k^T d_k\right)^2} + 2\frac{\theta_k}{\|g_k\|^2} - \frac{\theta_k^2}{\|g_k\|^2}, \tag{39}$$

combined with formula (19), it can be obtained

$$\begin{aligned} \frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} &\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + 2\frac{\theta_k}{\|g_k\|^2} - \frac{\theta_k^2}{\|g_k\|^2} \\ &= \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} - \frac{1}{\|g_k\|^2}\left(\theta_k^2 - 2\theta_k + 1 - 1\right) \\ &= \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} - \frac{(\theta_k - 1)^2}{\|g_k\|^2} + \frac{1}{\|g_k\|^2} \\ &\leq \frac{\|d_{k-1}\|^2}{\|g_{k-1}\|^4} + \frac{1}{\|g_k\|^2}. \end{aligned} \tag{40}$$

When $k = 1$,

$$\frac{\|d_1\|^2}{\left(g_1^T d_1\right)^2} = \frac{1}{\|g_1\|^2}. \tag{41}$$

Therefore, it can be obtained from the above inequality

$$\frac{\|d_k\|^2}{\left(g_k^T d_k\right)^2} \le \sum_{i=1}^{k} \frac{1}{\|g_i\|^2} \le \frac{k}{\gamma^2}. \tag{42}$$

Therefore

$$\frac{\left(g_k^T d_k\right)^2}{\|d_k\|^2} \ge \frac{\gamma^2}{k} = +\infty. \tag{43}$$

Contradictions with Lemma 3.1 above $\liminf_{k\to\infty} \|g_k\| = 0$. Thus, the proof is complete.

The proof of the above theorem shows that the constructed algorithm meets the Zoutendijk condition, that is, the algorithm has global convergence. In the previous section, it is proved that the algorithm meets the sufficient descent condition. That is, the theoretical properties are proved. In the next section, we will discuss the numerical effect of the new algorithm.

## 5. Numerical Results

To verify the numerical effect of algorithm 2.1, a set of test questions in the test function set [30] are tested, and the test results are compared with the AZPRP method proposed by Alhawarat [21]. Among them, the parameter value of Algorithm 2.1 is $\delta = 0.0001, \sigma = 0.05$, and the parameters of AZPRP is $\delta = 0.0001, \sigma = 0.1$ and $\sigma = 0.4$.

Table 1 Numerical experimental results

| Function | Dim | AZPRP PNI/PNF/PNG | LW PNI/PNF/PNG |
|---|---|---|---|
| Diagonal1 | 10 | 18/63/45 | 18/49/30 |
| BDQRTIC | 50 | 102/402/132 | 57/548/80 |
| Gen.White&Holst | 50 | 1174/5189/1679 | 1237/7058/1986 |
| Ext Quad Exponential EP1 | 50 | 3/13/4 | 2/11/3 |
| FHess1 | 50 | 404/1160/546 | 248/936/517 |
| ARWHEAD | 100 | 7/27/10 | 6/26/12 |
| Raydan1 | 200 | 5/175/88 | 4/175/90 |
| Raydan2 | 200 | 3/5/5 | 3/5/5 |
| Ext.White&Holst | 400 | 43/203/96 | 35/201/84 |
| Biggsb1 | 500 | 500/1007/1508 | 500/1007/1508 |
| Ext Quad penalty QP1 | 500 | 9/45/29 | 6/29/18 |
| Quadratic QF2 | 500 | 221/896/230 | 214/1062/224 |
| Quadratic QF1 | 500 | 122/489/241 | 122/536/362 |
| FHess2 | 500 | 4276/25496/4530 | 886/5546/1208 |
| CUBE | 500 | 5898/26124/8561 | 1640/8513/2833 |
| Ext Penalty | 500 | 15/65/23 | 12/51/17 |
| LIARWHD | 500 | 26/112/32 | 16/88/30 |
| Perturbed Quad | 500 | 122/489/123 | 122/609/239 |
| ENGVAL1 | 500 | 21/45/22 | 19/52/31 |
| SINQUAD | 800 | 118/737/429 | 116/720/327 |
| Ext PSC1 | 1000 | 9/21/11 | 7/23/14 |
| NONDQUAR | 1000 | 1691/4045/3497 | 1008/3741/3465 |
| FLETCHER | 1000 | 2986/13688/4702 | 2014/13356/4340 |
| Ext.Tridiagonal2 | 1000 | 26/54/28 | 26/54/28 |
| Gen.Tridiagonal1 | 1000 | 20/60/24 | 18/55/21 |
| Gen PSC1 | 1000 | 17/93/90 | 13/91/82 |
| Gen.Tridiagonal2 | 1000 | 32/105/41 | 37/126/48 |

The results of the test questions are shown in Table 1. All the algorithms are coded in MATLAB R2019a. Function represents the test problem; Dim represents the dimension of the test problem; PNI represents the number of iterations in the algorithm; PNF represents the calculation times of the function in the test problem; PNG represents the calculation times of the gradient in the test problem. When $\|g_k\| \le 10^{-5}$ or the maximum number of iterations exceeds 10000, the iteration is stopped

Algorithm 2.1 is superior to algorithm AZPRP in the number of iterations and number of gradient calculations, it improves the computational efficiency. Therefore, the new algorithm proposed in this paper is effective. This is an improved algorithm based on AZPRP and YHM.

## 6. Conclusion

Although the research of the conjugate gradient method has achieved rich results, it is still worthy of in-depth discussion to combine the Newton method with the conjugate ladder method with strong convergence to improve the efficiency of the algorithm and establish its convergence theory. In this work, we prove the sufficient descent property of the newly proposed SCG method through Strong Wolfe-Powell line search. The numerical results show that LW algorithm outperforms AZPRP conjugate gradient method in terms in the number of iterations and the number of gradient calculation. The theoretical properties and numerical calculations are excellent.

There is still much room for exploration on how to reasonably select the spectral parameters and conjugate parameters to fully combine the computational advantages of the two methods, and how to design the combination form of the hybrid conjugate gradient method and determine the calculation method of the combination parameters. In addition, we will also study small changes to the new formula and line search method to improve the speed and reduce the number of gradient evaluations required to solve the test function.

## References


[1] Abbas A A , Hewahi N M , Exploring the Effect of Imaging Techniques Extension to PSO on Neural Networks, International Journal of Image, Graphics and Signal Processing, 12(2)(2020)9-18.

[2] Hussain A , Yousaf S , Muhammad, et al, An Efficient Genetic Algorithm for Numerical Function Optimization with Two New Crossover Operators, International Journal of Mathematical Sciences and Computing, 4(4) (2018)1-17.

[3] Mamdouh M , Ezzat M , H Hefny, Optimized Planning of Resources Demand Curve in Ground Handling based on Machine Learning Prediction, International Journal of Intelligent Systems and Applications, 13(1)(2021)1-16.

[4] Sharma N, Batra U , Zafar S , Inculcating Global Optimization in ZRP through Newfangled Firefly Algorithm, International Journal of Computer Network and Information Security, 11(2)(2019)43-51.

[5] Birgin E G, Martinez J M, A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Optim, 43 (2)(2001) 117-128.

[6] Fletcher R, Reeves C M, Function minimization by conjugate gradients, The computer journal, 7(2)(1964) 149-154.

[7] Polak E , Ribière G, Note sur la convergence de mèthodes de directions conjugèes, Rev. Fran. Inform. Rech. Operationelle, 16 (3)(1969) 35-43.

[8] Polyak B T, The conjugate gradient method in extreme problems, USSR Comput. Math. Phys,1969,9(1969): 94-112.

[9] Dai Y H, Yuan Y X, A nonlinear conjugate gradient method with a strong global Convergence property, SIAM J. Optim , 10(1999) 177-182.

[10] Hanger W W, Zhang H, A new conjugate gradient method with guaranteed descent and efficient line search, Siam Optim , 16(1)(2005) 170-192.

[11] Liu Y L, Storey C S, Efficient generalized conjugate gradient algorithms, Journal of Optimization Thory and Applications,69, (1991) 129-137.

[12] Andrei N, New accelerated conjugate gradient algorithms as a modification of Dai–Yuan's computational scheme for unconstrained optimization, Journal of Computational & Applied Mathematics, 234(12) (2010)3397-3410.

[13] Deng S , Zhong W, Chen X , An Improved Spectral Conjugate Gradient Algorithm for Nonconvex Unconstrained Optimization Problems, Journal of Optimization Theory and Applications, 2013, 157(3).

[14] Dong Y L, Gen Q X, Symmetric Perry conjugate gradient method, Computational Optimization and Applications, 56(2)(2013) 317-341.

[15] Perry A , A Modified Conjugate Gradient Algorithm, Operations Research, 26(6) (1978)1073-1078.

[16] Zhang Y, Dan B, An efficient adaptive scaling parameter for the spectral conjugate gradient method, Optimization Letters, 10(1) (2016)119-136.

[17] Peyghami M R , Ahmadzadeh H , Fazli A , A new class of efficient and globally convergent conjugate gradient methods in the Dai–Liao family, Optimization Methods & Software, 30(4)(2015)1-21.

[18] Nezhadhosein S , A Modified Descent Spectral Conjugate Gradient Method for Unconstrained Optimization, Iranian Journal of Science and Technology, Transactions A: Science, (2020)1-12.

[19] Sheekoo A H , Al-Naemi G M ,Global Convergence Condition for a New Spectral Conjugate Gradient Method for Large-Scale Optimization, Journal of Physics: Conference Series, 1879(3)(2021)032001 .

[20] Baluch B, Salleh Z, Alhawarat A , A New Modified Three-Term Hestenes–Stiefel Conjugate Gradient Method with Sufficient Descent Property and Its Global Convergence, Journal of Optimization, 2018(6)(2018)1-13.

[21] Alhawarat A , Salleh Z , Mamat M , et al, An efficient modified Polak–Ribière–Polyak conjugate gradient method with global convergence properties, Optimization Methods & Software, (2016 ) 1-14.
[22] Polak E, Ribiere G, Note sur la convergence de méthodes de directions conjuguées, Esaim Mathematical Modelling & Numerical Analysis, 3(16)(1969)35-43
[23] Hager W W, Zhang H, A survey of nonlinear conjugate gradient methods, Pacific Journal of Optimization, 2(1)(2006)35-58.
[24] Gilbert J C, Nocedal J, Global convergence properties of conjugate gradient methods for optimization, Siam Journal on Optimization, 2(1)(1992)21-42.
[25] Wei Z,Yao S,Liu L,The convergence properties of some new conjugate gradient methods, Applied Mathematics& Computation, 183(2)(2006)1341-1350.
[26] Dai Z , F Wen, Another improved Wei–Yao–Liu nonlinear conjugate gradient method with sufficient descent property, Applied Mathematics & Computation, 218(14)(2012)7421-7430.
[27] Salih Y, Hamoda M A, Rivaie M, New hybrid conjugate gradient method with global convergence properties for unconstrained optimization, Malays J Comput Appl Math ,1(1)(2018) 29–38.
[28] Jian J B, Jiang X Z, Yin J H , Research progress in nonlinear conjugate gradient method, Journal of Yu lin Normal University (Natural Science),37(2)(2016) 3-10.
[29] Zoutendijk G , Nonlinear programming, computational methods, Integer & Nonlinear Programming, (1970)37-86.
[30] Andrei N, An unconstrained optimization test functions collection, Advanced Modeling and Optimization, 10(1)(2008)147-161.

## Authors' Profiles

**Jing Li** is now pursuing her master degree in the School of Mathematics and Information Science at Henan Polytechnic University, China. Her research interests include optimization theory in operational research. Her daily hobbies are studying cooking and surfing the Internet. She is good at cooking pineapple glutinous rice and amber walnut kernel.

**Shujie Jing** received a master of Science Degree in Computational Mathematics from Xi'an Jiaotong University in 1997. He is currently a lecture in the School of Mathematics and Information Science at Henan Polytechnic University, China. His research interests include optimization theory, geometric programming, convex analysis, numerical solution of nonlinear equations.

**How to cite this paper:** Jing Li, Shujie Jing," A Hybrid Spectral Conjugate Gradient Method with Global Convergence ", International Journal of Mathematical Sciences and Computing(IJMSC), Vol.8, No.2, pp. 1-10, 2022. DOI: 10.5815/ijmsc.2022.02.01