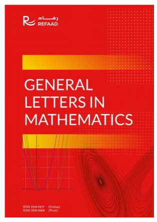
Available online at www.refaad.com
Gen.Lett. Math., 11(2) (2021), 36-45
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: https://www.refaad.com/Journal/Index/1
ISSN: 2519-9277 (Online) 2519-9269 (Print)



A New Coefficient of Conjugate Gradient Method with Global Convergence for Unconstrained Optimization Problems

Mardeen Sh. Taher[a,*], Salah G. Shareef[b]

[a] *Department of Mathematics, College of Science, Dohuk University, Kurdistan Region, Iraq.*

[b] *Department of Mathematics, College of Science, Zakho University, Kurdistan Region, Iraq.*

Abstract

In this article, we defined a new coefficient formula of the conjugate gradient method for solving non linear unconstrained optimization problems. The new formula β_k^{new} is type of line search and the idea of our work is to focus on modification the Perry's suggestion. We further show that global convergence result of new formula is recognized under Wolf-Powell line search. It is shown that the new CG coefficient satisfied sufficient descent conditions. In the end, numerical experiments with the collection of test functions show that the new β_k^{new} is more effective compared to some other standard formulas such as β_k^{H-S}, β_k^{Perry} and β_k^{D-Y}.

Keywords: Perry Conjugate Gradient, Wolfe line search, Global Convergence, Unconstrained optimization.
2010 MSC: MSC code1, MSC code2, more.

1. Introduction

There are several methods to find an optimum or near-optimum solution of unconstrained optimization problems that may ascend in fields such as technology, sciences, economics, and many more. Conjugate gradient (CG) method one of these methods and it plays a significant role to solve large scaled problems. This is because of low memory requirements as well as global convergence properties.The creation of the methods returns to 1952 when Hestenes and Stiefel introduced a CG method for solving a linear system of equations[5].In 1960, Fletcher and Reeves refined and developed the conjugate gradient method for solving the unconstrained nonlinear optimization problems[3].The improvement of this method non-stop, but it continues to present day.Therefore, there are various studies on the conjugate gradient methods, all of them are them focus on developing the CG parameter. Also in this paper a new modified CG parameter is proposed and analyzed .

Trying to solve the unconstrained minimization problem:-

$$\min\{f(x) : x \in R^n\} \tag{1.1}$$

Where $f(x)$ is twice continuously differentiable function over R^n, we are beginning with an initial point x_0 is a first approximation of the minimum point, and having found the new point x_{k+1} by searching along a decent direction d_k such that $d_k^T g_k \leqslant 0$,so that

*Corresponding author
Email addresses: mardinsh.tahir@uod.ac (Mardeen Sh. Taher), salah.shareef@uoz.edu.krd (Salah G. Shareef)

doi:10.31559/glm2021.11.2.3

Received 7 Nov 2021 : Revised : 25 Dec 2021 Accepted: 26 Jan 2022

$$x_{k+1} = x_k + \alpha_k d_k \tag{1.2}$$

However, α_k is a length step and fulfill the following Wolfe–Powell search conditions[11]: -

$$f(x_k + \alpha_k d_k) \leqslant f(x_k) + \delta_1 \alpha_k \nabla f_k^T d_k \tag{1.3}$$

$$\nabla f(x_k + \alpha_k d_k)^T d_k \geqslant \delta_2 \nabla f_k^T d_k \tag{1.4}$$

with $0 < \delta_1 < \delta_2 < 1$, and indicated $x_{k+1} - x_k$ by ν_k and $\nabla f(x_{k+1}) - \nabla f(x_k)$ by y_k.
we refer that the nonlinear conjugate gradient methods (CG-methods) are practical methods for finding the minimum value of large-dimensional functions because they do not need matrix storage. The basic idea of all CG-methods is calculated with a new direction by the following form:

$$\begin{cases} d_k = -g_k & k = 0 \\ d_{k+1} = -g_{k+1} + \beta_k d_k & k > 0 \end{cases} \tag{1.5}$$

In equation (1.5), $g = \nabla f(x)$ and g_k and g_{k+1} are gradient of $f(x)$ at the point x_k, x_{k+1} respectively, and β_k is a positive real number called the coefficient of conjugate gradient, several efforts have been completed in the few recent years to proposal new formulas of conjugate gradient methods which are take many alternative values, such as:

$$\beta_k^{H-S} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \tag{1.6}$$

$$\beta_k^{F-R} = \frac{g_{k+1}^T g_{k+1}}{g_k^T g_k} \tag{1.7}$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{g_k^T g_k} \tag{1.8}$$

$$\beta_k^{perry} = \frac{g_{k+1}^T (y_k - \nu_k)}{d_k^T y_k} \tag{1.9}$$

$$\beta_k^{D-Y} = \frac{g_{k+1}^T g_{k+1}}{d_k^T y_k} \tag{1.10}$$

$$\beta_k^{L-S} = -\frac{g_{k+1}^T y_k}{d_k^T g_k} \tag{1.11}$$

In which, the definition of β_k^{H-S} in (1.6) is due to Hestenes and Stiefel [5], β_k^{F-R} in (1.7) is due to Fletcher and Reeves [3], β_k^{PRP} in (1.8) is due to Polak-Ribiere-Polyak [9], β_k^{perry} in (1.9) is developed by Perry [8] , β_k^{D-Y} in (1.10) suggested by Dai and Yuan [6], and Liu and Storey founded β_k^{L-S} which is defined in (1.11) [7]. It is well known that the convergence condition is essential for each iterative methods .So, the study of the global convergence for Conjugate gradient methods is very important, in 1970, Zoutendijk G.[12], proved that the Fletcher and Reeves method is global convergence when the line search is exact . Powell studied the global convergence of Polak-Ribiere method and described the method as global convergence when the strongly convex condition and the line search are exact [10], but after that Powell established that the Polak-Ribiere method with exact line search could circle infinitely without convergent to a required point, the same result applies to the Hestenes-Stiefel method . Thus, some methods have strong properties of convergence but their practical performance is often very weak β_k^{F-S}. On the other hand some methods may not be convergent but have a good numerical performance like β_k^{PRP}. We will discuss the global convergence of new method and numerical performance in section 4 and 5.

2. The New Formula of CG-Coefficient β_k^{new}

In this section, we prepare a new formula for conjugate gradient method based on the formula of CG-Coefficient which is suggested by Perry of β_k^{perry}[8] by changing the CG update parameter of the HS conjugate gradient method in to (1.9). In this paper, we choose a suitable equation which was proposed by Powell in (1978) [10] and defined in as

$$\eta_k = (1-\theta)Gv_k + \theta y_k \tag{2.1}$$

here G is a symmetric matrix of second partial derivatives of function and $\theta \in (0,1)$ is a positive real number between 0 and 1.

Now, we suppose that

$$Gv_k = \frac{y_k}{\delta} \tag{2.2}$$

Let, $\delta = \frac{2\sqrt{\omega}}{||v_k||}(1+||x_{k+1}||)$, ω is a machine accuracy, and the function $||.|| \geqslant 0$ is the Euclidean norm of vectors.

By putting the value of δ in equation (2.2), therefor Gv_k can be taken as

$$Gv_k = ||v_k||\frac{y_k}{2\sqrt{\omega}(1+||x_{k+1}||)} \tag{2.3}$$

now we replace Gv_k in (2.1) by (2.3), and obtain the following

$$\eta_k = (1-\theta)\left(||v_k||\frac{y_k}{2\sqrt{\omega}(1+||x_{k+1}||)}\right) + \theta y_k \tag{2.4}$$

Change y_k in the numerator of (1.9), by η_k which is defined in (2.4), and get

$$\beta_k^{new} = \frac{g_{k+1}^T\left([(1-\theta)\left(||v_k||\frac{y_k}{2\sqrt{\omega}(1+||x_{k+1}||)}\right) + \theta y_k] - v_k\right)}{d_k^T y_k} \tag{2.5}$$

$$\beta_k^{new} = \frac{\left([(1-\theta)\left(||v_k||\frac{g_{k+1}^T y_k}{2\sqrt{\omega}(1+||x_{k+1}||)}\right) + \theta g_{k+1}^T y_k] - g_{k+1}^T v_k\right)}{d_k^T y_k} \tag{2.6}$$

Let $\mu = \left[(1-\theta)\left(\frac{||v_k||}{2\sqrt{\omega}(1+||x_{k+1}||)}\right) + \theta\right] \qquad \mu > 0$

Therefor the new formula is

$$\beta_k^{new} = \mu\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k} \tag{2.7}$$

we observe, if the orthogonal condition is satisfied i.e $\langle g_{k+1}, g_k\rangle = 0$, the β_k^{new} in (2.7) becomes

$$\beta_k^{new} = \mu\frac{g_{k+1}^T(g_{k+1}-g_k)}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k} \tag{2.8}$$

$$\beta_k^{new} = \mu\frac{||g_{k+1}||^2}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k} \tag{2.9}$$

It is very important for design a new formula of β_k^{new} the two conditions required to be satisfied, the frist one is the descent direction

$$d_{k+1}^T g_{k+1} \leqslant 0, \qquad \forall k \geqslant 0 \tag{2.10}$$

while the second one is the sufficient condition

$$d_{k+1}^T g_{k+1} < -\tau \|g_{k+1}\|^2, \quad \forall k \geqslant 0 \quad \text{and} \quad \tau > 0 \tag{2.11}$$

.

3. Generalized (The New CG- Algorithm)

- **Step 0:** Start with arbitrary initial point of solution $x_0 \in R^n, k = 0$, set $\epsilon > 0, n \in Z$.
- **Step 1:** Test if $\|g_k\| < \epsilon$ then stop, else $d_k = -g_k = -\nabla f(x_k)$ and go to step (2).
- **Step 2:** Using cubic line search to determine the size step α_k such that rules (3) and (4) are satisfied and generate the next iterate via $x_{k+1} = x_k + \alpha_k d_k$.
- **Step 3:** Test the optimality for the new point x_{k+1}, if $\|g_{k+1}\| < \epsilon$ then stope and x_{k+1} is a minimizer, otherwise compute $d_{k+1} = -g_{k+1} + \beta_k d_k$, β_k is defined in (2.7) or (2.9) and go to step 4.
- **Step 4:** if $|g_{k+1}^T g_k| > 0.2 g_{k+1}^T g_{k+1}$ then go to step 1, alse $k = k + 1$ and return to Step 2.

4. Convergence Analysis for New Algorithm

To ensure that the new method is converge we must show that both conditions in (2.10) and (2.11) are holds. In addition, the property of global convergence should be fulfilled .

4.1. Descent and Sufficient Conditions

The descent and sufficient decent Conditions are always assumed to achieve, considering that they are play very an important role for establishing the global convergence of conjugate gradient methods.

Theorem 4.1. *Consider the sequences of* $\{d_k\}$ *and* $\{g_k\}$ *are generated by new CG-method 3 then* (2.10) *is holds for all* $k > 0$*, i.e*

$$d_k^T g_k \leqslant 0 \tag{4.1}$$

Proof. By mathematical induction, we prove it,
at $k = 0$, $d_0 = -g_0$, we have

$$d_0^T g_0 \leqslant -\|g_0\|^2. \tag{4.2}$$

now we asssume that the conclusion (4.1) holds for $k \geqslant 0$ and $g_{k+1} neq 0$
Case(i) $g_{k+1}^T g_k \neq 0$,

$$d_{k+1} = -g_{k+1} + \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\right) d_k \tag{4.3}$$

multiply both sides of (4.3) by g_{k+1}^T from right and we get,

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\right) d_k^T g_{k+1}, \tag{4.4}$$

this leads to,

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} - \frac{g_{k+1}^T v_k}{d_k^T y_k} d_k^T g_{k+1}\right). \tag{4.5}$$

Thus,

$$d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 + \mu \frac{g_{k+1}^T y_k}{d_k^T y_k} d_k^T g_{k+1} - \frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k}. \tag{4.6}$$

It is known that the first two terms of equation (4.6) refer to Hestenes and Stiefel method which hold the descent condition, now we need to prove the third term of (4.6) is less than or equal to zero.

It is noted all of α_k, $(d_k^T g_{k+1})^2$, are positive and $d_k^T y_k = d_k^T(g_{k+1} - g_k) > (\delta_2 - 1)d_k^T g_k$, $(\delta_2 - 1)d_k^T g_k > 0$ therefor $d_k^T g_k > 0$.
which get to $g_{k+1}^T d_{k+1} \leqslant 0$.

Case (ii) when the orthogonal property $(g_{k+1}^T g_k = 0)$ is satisfied, then

$$d_{k+1} = -g_{k+1} + \Big(\mu \frac{||g_{k+1}||^2}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\Big) d_k, \tag{4.7}$$

multiply both sides (4.7) be g_{k+1}^T, we obtain

$$d_{k+1}^T g_{k+1} = -g_{k+1}^T g_{k+1} + \Big(\mu \frac{||g_{k+1}||^2}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\Big) d_k^T g_{k+1}, \tag{4.8}$$

$$d_{k+1}^T g_{k+1} = -||g_{k+1}||^2 + \mu \frac{||g_{k+1}||^2}{d_k^T y_k} d_k^T g_{k+1} - \frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k}. \tag{4.9}$$

Dai and Yuan are proved the term $-||g_{k+1}||^2 + \mu \frac{||g_{k+1}||^2}{d_k^T y_k} d_k^T g_{k+1}$ in (4.9) is descent condition and from above , we have $\frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k} > 0$. Therefor $d_{k+1}^T g_{k+1} \leqslant 0$. □

Theorem 4.2. *Consider the sequences of* $\{d_k\}$ *and* $\{g_k\}$ *are generated by new CG-Method 3 then* (2.11) *is satisfied,*

$$d_{k+1}^T g_{k+1} < -\tau ||g_{k+1}||^2, \qquad \forall k \geqslant 0 \tag{4.10}$$

Proof. It is seen from theorem 4.1 d_{k+1}^T is a descent direction which means that the two first terms of equation (4.3) are less than or equal to zero in two cases $g_{k+1}^T g_k = 0$ and $g_{k+1}^T g_k \neq 0$. So we have

$$d_{k+1}^T g_{k+1} \leqslant -\frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k} \tag{4.11}$$

Multiply and divide the term $\frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k}$ in (4.11) by $||g_{k+1}||^2$, we obtain

$$d_{k+1}^T g_{k+1} \leqslant -\frac{\alpha_k (d_k^T g_{k+1})^2 ||g_{k+1}||^2}{d_k^T y_k ||g_{k+1}||^2} \tag{4.12}$$

$$d_{k+1}^T g_{k+1} \leqslant -\left(\frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k ||g_{k+1}||^2}\right) ||g_{k+1}||^2. \tag{4.13}$$

We assume $\tau = \left(\frac{\alpha_k (d_k^T g_{k+1})^2}{d_k^T y_k ||g_{k+1}||^2}\right)$, and $\tau > 0$.
Since (4.13) becomes $g_{k+1}^T d_{k+1} \leqslant -\tau ||g_{k+1}||^2$, in this way the proof is completed. □

4.2. Global Convergence of New Conjugate Gradient Method

To study the convergent of new algorithm, we need to mention an important hypothesis, and lemma gives the Zoutendijk condition.

Hypotheses (H):

i) The set $\Omega = \{x : x \in R^n$, and $\psi(x) \leqslant \psi(\check{x})\}$ is closed and bounded on the initial point $\check{x}$.

ii) In some neighborhood $\aleph$ of Ω,the objective function $\psi(x)$ is continuously differentiable and its gradient $\psi'(x)$ is Lipschitz continuous, that means, there exists a constant $\zeta > 0$ such that :

$$\|\nabla\psi(\check{x}) - \nabla\psi(\check{y})\| \leqslant \zeta\|\check{x} - \check{y}\|, \qquad \forall\ \check{x} \text{ and } \check{y} \in \aleph \tag{4.14}$$

iii) objective function $\psi(x)$ is uniformly Convex and there is a constant number $\gamma > 0$ such that

$$(\nabla\psi(\check{x}) - \nabla\psi(\check{y}))^T(\check{x} - \check{y}) \leqslant \gamma\|\check{x} - \check{y}\|^2, \qquad \forall\ \check{x} \text{ and } \check{y} \in \aleph \tag{4.15}$$

under these Hypotheses (H) on $\psi(x)$, there exists a constant $\hbar > 0$ such that $\nabla\psi(\check{x}) \leqslant \hbar, \forall\ \check{x} \in \aleph$, and we give a useful lemma,it was essentially proved by Zoutendijk and Wolf [12, 10].

Lemma 4.3. *let $\check{x}$ is a starting point for which Hypotheses (H) is satisfied. Consider all techniques in the (1.2), suppose that a set of conjugate directions is defined in (1.5), and the step size α_k satisfies the standard Wolfe conditions of line search (1.3) and (1.4). If*

$$\sum_{k\geqslant 1} \frac{1}{\|d_{k+1}\|^2} = \infty. \tag{4.16}$$

Then , the following

$$\lim_{k\to\infty} (\inf\|g_{k+1}\|) = 0, \tag{4.17}$$

holds.

Theorem 4.4. *(This theorem gives the global convergence of new method)*
Consider the functionf(x) is uniformly convex and hypotheses (H) is satisfied. Let the sequence $\{x_k\}$ is generated by new method and the step size α_k is calculated by the weak Wolfe conditions of line search (1.3) and (1.4). Then the (4.17) is holds i.e.,

$$\lim_{k\to\infty} (\inf\|g_{k+}\|) = 0. \tag{4.18}$$

Proof. from (2.7) we have the direction:

$$d_{k+1} = -g_{k+1} + \Big(\mu\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\Big)d_k. \tag{4.19}$$

Taking norm of both sides of (4.19),

$$\|d_{k+1}\| = \left\|g_{k+1} + \Big(\mu\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\Big)d_k\right\|. \tag{4.20}$$

Apply preliminary of Cauchy–Schwarz and Triangle inequities to simplify the (4.20), and get

$$\|d_{k+1}\| \leqslant \|g_{k+1}\| + \left\|\Big(\mu\frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\Big)d_k\right\|. \tag{4.21}$$

By properties of norm, we have

$$\|d_{k+1}\| \leqslant \|g_{k+1}\| + \left| \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T \nu_k}{d_k^T y_k} \right) \right| \|d_k\| . \tag{4.22}$$

Return to Cauchy–Schwarz and use hypothesis (ii), we have seen

$$\left\|g_{k+1}^T y_k\right\| \leqslant \|g_{k+1}\| \, \|y_k\| \quad \text{and} \quad \|y_k\| \leqslant \imath \, \|\nu_k\| .$$

So, we see

$$\left| \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T \nu_k}{d_k^T y_k} \right) \right| \leqslant \mu \frac{\|g_{k+1}\| \, \|y_k\|}{\|d_k\| \, \|y_k\|} - \frac{\|g_{k+1}\| \, \|\nu_k\|}{\|d_k\| \, \|y_k\|}. \tag{4.23}$$

Note that in hypothesis (iii), $\|g_{k+1}\| < \hbar$, and the truth that $g_{k+1}^T y_k < d_k^T y_k$ so the equation (4.23) becomes as follows

$$\left| \left(\mu \frac{g_{k+1}^T y_k}{d_k^T y_k} - \frac{g_{k+1}^T \nu_k}{d_k^T y_k} \right) \right| \leqslant \mu \frac{\hbar}{\|d_k\|} + \alpha_k. \tag{4.24}$$

Now, put (4.24) in (4.22) we get

$$\|d_{k+1}\| \leqslant \hbar + \mu\hbar + \frac{\hbar}{\zeta}. \tag{4.25}$$

Suppose $\hbar + \mu\hbar + \frac{\hbar}{\zeta} = \mathcal{F}$, such that $\mathcal{F}$ is positive real number.

Therefor

$$\|d_{k+1}\| \leqslant \mathcal{F}. \tag{4.26}$$

So that we can write,

$$\sum_{k \geqslant 1} \frac{1}{\|d_{k+1}\|^2} \geqslant \sum_{k \geqslant 1} \frac{1}{\mathcal{F}^2} = \infty. \tag{4.27}$$

Thus,

$$\sum_{k \geqslant 1} \frac{1}{\|d_{k+1}\|^2} = \infty, \tag{4.28}$$

so, by applying Lemma 4.3, we conclude that (4.18) is proved, i.e., $\lim_{k\to\infty}(\inf\|g_{k+1}\|) = 0$, in this way the proof is achieved. □

5. Numerical Experiments and Discussions

The numerical experiments of the new method (β_k^{new}) it will be discussed and displayed in this section. All codes of methods are written in Fortran (95) with the stopping condition is $\|g_{k+1}\| < 10^{-5}$, we used for testing the well-known nonlinear problems from the CUTEr library with the number of variables each test problem is 4 ,100,500,1000, 3000 and 5000, [1]. Throughout the numerical results, we compare the performance of new method with the Hestenes and Stiefel and Perry methods, and in case of orthogonal we compare the new method with Dai and Yuan method[2]. Cubic fit technique is used to find the size step α_k under conditions (3) and (4) in this paper we used the values of δ_1 and δ_2 in (3) and (4) were chosen to be 0.001 and 0.1 respectively. The numerical results are recorded in Tables 1 and 2 where Name, N, NOI, and NOF represent the name of the test function, the variables of function , the number of iterations and the number of function evaluations, respectively.

Table 1 shows the performance of three methods (β_k^{new}, β_k^{perry} and β_k^{H-S}). We noted that the new method has the good result compared to other both methods (β_k^{perry} and β_k^{H-S}) based on number of iterations and the number of function evaluations, while Table 2 illustrates the result of new method when the orthogonal holds and seen the new method has the best performance when compared with and Dai and Yuan method β_k^{D-Y}.

6. Conclusion

Overall, different choices for the scalar coefficient β_k leads to different conjugate gradient methods .In this paper, a new type of coefficient in the conjugate gradient methods for solving unconstrained optimization problems is proposed .the numerical tests were carried out on low and high problems, and comparisons were made amongst different test functions. The new method has proven its efficiency through results in tables 1 and 2

References

[1] Andrei, Neculai. *An unconstrained optimization test functions collection*. Adv. Model. Optim **10**.1 (2008): 147-161. 5

[2] Dai, Yu-Hong, and Yaxiang Yuan. *A nonlinear conjugate gradiant method with a strong global convergence property*. SIAM Journal on optimization **10**.1 (1999): 177-182. 5 https://doi.org/10.1137/s1052623497318992

[3] Fletcher, Roger. *Practical methods of optimization*. John Wiley & Sons, 2013. 1, 1

[4] Fletcher, Reeves, and Colin M. Reeves. *Function minimization by conjugate gradients*. The computer journal **7** .2 (1964): 149-154. https://doi.org/10.1093/comjnl/7.2.149

[5] Hestenes, Magnus Rudolph, and Eduard Stiefel. *Methods of conjugate gradients for solving linear systems*. Vol. 49. No. 1. Washington, DC: NBS, 1952.

[6] Liu, D. Y., and G. Q. Xu. *A Perry descent conjugate gradient method with restricted spectrum*. Optimization Online, Nonlinear Optimization (unconstrained optimization) (2011): 1-19.

[7] Liu, Y., and C. Storey. *Efficient generalized conjugate gradient algorithms, part 1: theory*. Journal of optimization theory and applications **69**.1 (1991): 129-137. https://doi.org/10.1007/bf00940464

[8] Perry, Avinoam. *A modified conjugate gradient algorithm*. Operations Research **26**.6 (1978): 1073-1078. https://doi.org/10.1287/opre.26.6.1073

[9] Polak, Elijah, and Gerard Ribiere. *Note sur la convergence de méthodes de directions conjuguées*. ESAIM: Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique 3.R1 (1969): 35-43. 1, 1

[10] Powell, Michael JD. *A fast algorithm for nonlinearly constrained optimization calculations*. Numerical analysis. Springer, Berlin, Heidelberg, 1978. 144-157. 1

[11] Wolfe, Philip. *Convergence conditions for ascent methods*. SIAM review **11**.2 (1969): 226-235. https://doi.org/10.1137/1011036 1

[12] Zoutendijk, G. *Nonlinear programming, computational methods*. Integer and nonlinear programming (1970): 37-86. 1, 2

Table 1: Comparing performance of the three methods (β_k^{new} , β_k^{H-S} and β_k^{perry}).

Test	N	β_k^{new} NOI-NOF	β_k^{H-S} NOI-NOF	β_k^{perry} NOI-NOF
Powell (3,-1,0,1)	4	32-85	38-108	35-89
	100	35-102	40-122	43-105
	500	35-102	41-124	43-105
	1000	35-100	41-124	45-120
	3000	35-100	41-124	46-122
	5000	35-100	41-124	46-122
Wood (-3,-1,-3,-1)	4	28-66	30-68	30-68
	100	30-71	30-68	30-68
	500	27-63	30-68	30-68
	1000	30-73	30-68	30-68
	3000	28-66	30-68	30-68
	5000	28-67	30-68	30-68
Rosen (-1.2,1;..)	4	26-85	30-83	30-83
	100	20-59	30-83	30-83
	500	20-59	30-83	30-83
	1000	16-52	30-83	30-83
	3000	16-51	30-83	30-83
	5000	17-53	30-83	30-83
Powell (0,1,2;..)	4	15-32	16-36	16-36
	100	16-35	16-36	16-37
	500	16-35	16-36	16-37
	1000	16-35	16-36	16-37
	3000	16-35	16-36	16-37
	5000	16-35	16-36	16-37
Cubic (-1.2,1;..)	4	12-35	12-35	12-35
	100	13-37	13-37	13-37
	500	13-37	13-37	13-37
	1000	13-37	13-37	13-37
	3000	13-37	13-37	13-37
	5000	13-37	13-37	13-37
Miele (1,2,2,2)	4	23-60	28-85	34-133
	100	34-96	33-114	46-169
	500	38-110	40-146	52-198
	1000	39-111	46-176	58-229
	3000	52-160	54-211	58-229
	5000	62-199	54-211	64-261
Wolfe (-1;..)	4	14-29	11-24	11-24
	100	43-87	49-99	49-99
	500	45-92	52-105	52-105
	1000	47-96	70-141	70-141
	3000	116-248	170-351	170-351
	5000	147-310	165-348	166-350

Table 2: Comparing performance profiles of New method β_k^{new} and Dai and Yuan method β_k^{D-Y}.

Test	N	β_k^{new} NOI-NOF	β_k^{D-Y} NOI-NOF
Powell (-3,-1,0,1)	4	27-63	50-128
	100	28-65	51-130
	500	28-65	51-130
	1000	28-65	51-130
	3000	30-69	52-132
	5000	30-69	52-132
Wood (-3,-1,-3,-1)	4	27-63	28-65
	100	27-63	28-65
	500	27-63	29-68
	1000	29-67	29-68
	3000	29-67	29-68
	5000	29-67	29-68
Rosen (-1.2,1;..)	4	27-90	27-63
	100	20-59	28-65
	500	22-64	28-65
	1000	16-51	28-65
	3000	16-51	30-83
	5000	17-53	30-83
Miele (1,2,2,2)	4	34-109	36-115
	100	38-127	46-156
	500	33-111	53-188
	1000	57-207	60-222
	3000	46-160	66-257
	5000	63-249	66-257
Dixon (-1;..)	4	13-28	13-28
	100	523-1162	466-1021
	500	489-1094	503-1085
	1000	475-1061	484-1048
	3000	498-1107	462-1005
	5000	480-1076	510-1115