A Modified Perry's Conjugate Gradient Method
Based on Powell's Equation for Solving Large-Scale Unconstrained Optimization

Mardeen Sh. Taher¹*, Salah G. Shareef²

¹Department of Mathematics, College of Science, Duhok University, Kurdistan Region, Iraq
²Department of Mathematics, College of Science, Zakho University, Kurdistan Region, Iraq

Received August 21, 2021; Revised September 22, 2021; Accepted October 28, 2021

Abstract It is known that the conjugate gradient method is still a popular method for many researchers who are focused in solving the large-scale unconstrained optimization problems and nonlinear equations because the method avoids the computation and storage of some matrices so the memory’s requirements of the method are very small. In this work, a modified of Perry conjugate gradient method which fulfills a global convergence with standard assumptions is shown and analyzed. The idea of new method is based on Perry method by using the equation which is founded via Powell in 1978. The weak Wolfe–Powell search conditions is used to choose the optimal line search, under the line search and suitable conditions we prove both descent and sufficient descent conditions. In particular, numerical results show that the new conjugate gradient method is more effective and competitive when compared to other of standard conjugate gradient methods including: - CG- Hestenes and Stiefel (H/S) method, CG -Perry method CG - Dai and Yuan (D/Y) method. The comparison is completed under a group of standard test problems with various dimensions from the CUTEst test library and the comparative performances of the methods are evaluated by total the number of iterations and the total number of function evaluations.

Keywords Conjugate Gradient, Wolfe Line Search, Descent and Sufficient Descent Conditions, Analysis of Convergence

1. Introduction

Conjugate gradient (CG) methods have been used to solve large-scale unconstrained optimization problem which is presented as:

\[
\min\{h(x) : x \in \mathbb{R}^n\}
\]

(1)

Where \( h(x) \) is smooth and twice continuously differentiable function over \( \mathbb{R}^n \), \( n \) is a number of variables for our problem. In general Conjugate gradient (CG) methods are generate a sequence \( x_i \) of estimates to the minimum of \( h(x) \), by iterative formula:

\[
x_{i+1} = x_i + \lambda_id_i
\]

(2)

However, \( \lambda_i > 0 \) is a steplength and must be achieved the following Wolfe–Powell search conditions:

\[
h(x_i + \lambda_id_i) - h(x_i) \leq \xi_1\lambda_1\nabla h_i^Td_i, \xi_1 \in \left(0, \frac{1}{2}\right)
\]

(3)

\[
\nabla h(x_i + \lambda_id_i)\nabla d_i \geq \xi_2\nabla h_i^Td_i, \xi_2 \in (\xi_1, 1)
\]

(4)

And using a search direction:

\[
\begin{cases}
  d_i = -g_i & i = 0 \\
  d_{i+1} = -g_{i+1} + \beta_id_i & i \geq 1
\end{cases}
\]

(5)

where \( \beta_i \) is a scalar determines the conjugate gradient methods (CG-methods), and we denote\( \nabla h(x_i) \) which is
gradient of \( h(x_i) \) by \( g_i \).

At current, the most effective formulas of \( \beta_i \) are the following:

\[
\beta_i^{H/S} = \frac{g_{i+1}(g_{i+1} - g_i)}{d_i^T(g_{i+1} - g_i)}
\]

\[
\beta_i^{F/R} = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i}
\]

\[
\beta_i^{PRP} = \frac{g_{i+1}^T(g_{i+1} - g_i)}{g_i^T g_i}
\]

\[
\beta_i^{Perry} = \frac{g_{i+1}^T(y_i - v_i)}{d_i^T(y_i)}
\]

\[
\beta_i^{DY} = \frac{g_{i+1}^T g_{i+1}}{d_i^T y_i}
\]

In which, the definition of \( \beta_i^{H/S} \) is due to Hestenes and Stiefel (H/S, 1952) [6], \( \beta_i^{F/R} \) is due to Fletcher and Reeves (F/R, 1964) [4,5], \( \beta_i^{PRP} \) is due to Polak-Ribiere-Polyak (PRP, 1969) [7], \( \beta_i^{Perry} \) is suggested by Perry(1978) [8] and in (1999) \( \beta_i^{DY} \) is established by Dai-Yuan [2]. For iterative mathematics the global convergence is one of requirement for it, and it mean by global convergence is any sequence generated via the iterative methods will either achieved after finite steps or consist of a subsequence which is converged to stationary point of the problem \( h(x) \) from a given initial point \( x_0 \), \( \lim_{\alpha \to \infty}(\nabla h(x_0 + \alpha a_i)) = 0 \) [11]. For Conjugate gradient the global convergence is investigated by researchers; Zoutendijk G. in 1970 proved that the Fletcher and Reeves method is global convergence in case that the line search is exact [13]. In Polak-Ribire method global convergence is described by Powell [9] when \( h(x) \) is strongly convex and the line search is also exact, but after that Powell demonstrated that the Polak-Ribire strategy with exact line search could circle infinitely without convergent to a required point. The same result applies to the Hestenes-Stiefel method.

Many efforts have been made in few recent years to design new formulas for conjugate gradient method which are not only satisfied global convergence but also improve numerical performance for method. Remainder the conjugate gradient methods have many application in real life. In our work, we found a new formula for CG-method which is satisfied the global convergence in section 3.3 and the new formula with weak Wolfe–Powell generate a descent direction at each iteration in section 3.2.

2. Derivation of the New Conjugacy Coefficient

In this section, we present a new formula for conjugate gradient method as a result improve the parameter of \( \beta_i^{Perry} \) in (6) by using the equation which is proposed by Powell in (1978)[10].

Since

\[
\beta_i^{Perry} = \frac{g_{i+1}^T(y_i - v_i)}{d_i^T(y_i)}
\]

The equation which is developed by Powell is:

\[
y_i^{powell} = y_i + (1 - \theta)(Gv_i - y_i)
\]

Where, \( y_i = g_{i+1} - g_i \) and \( v_i = x_{i+1} - x_i \)

Thus, we have the quasi-Newton equation:

\[
Gv_i = \delta_i y_i
\]

Now, suppose that

\[
\delta_i = \frac{||v_i||}{2\sqrt{(1 + ||x_{i+1}||)}} \tag{9}
\]

Where \( ||.||: R^n \rightarrow R \), is Euclidean norm, and \( \sigma \) is a machine accuracy, hence, from (8) and (9) we get:

\[
Gv_i = ||v_i|| \frac{y_i}{2\sqrt{(1 + ||x_{i+1}||)}}
\]

So, after putting (10) in (9) we obtain new equation \( y_i^* \).

\[
y_i^* = y_i + (1 - \theta)\left(||v_i|| \frac{y_i}{2\sqrt{(1 + ||x_{i+1}||)}} - y_i\right)
\]

\[
y_i^* = (1 - \theta)\left(||v_i|| \frac{y_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta y_i\right)
\]

The idea of our new formula is based on replacing the equation \( y_i \) in Perry formula (3) with the equation \( y_i^* \), and getting

\[
\beta_i^{new} = \frac{g_{i+1}^T\left(1 - \theta\right)\left(||v_i|| \frac{y_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta y_i\right) - v_i}{d_i^T\left(1 - \theta\right)\left(||v_i|| \frac{y_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta y_i\right)} - \beta_i^{new}
\]

Hence, we get the new parameter conjugacy:

\[
\beta_i^{new} = \frac{g_{i+1}^T v_i}{d_i^T y_i} - \frac{g_{i+1}^T v_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta d_i^T y_i \tag{14}
\]

we observe that, if the orthogonal condition is satisfied i.e. \( g_{i+1}^T g_i = 0 \) then the parameter \( \beta_i^{new} \) in (14) become as a following:

\[
\beta_i^{new} = \frac{g_{i+1}^T (g_{i+1} - g_i)}{d_i^T y_i} - \frac{g_{i+1}^T v_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta d_i^T y_i \tag{15}
\]

\[
\beta_i^{new} = \frac{||g_{i+1}||^2}{d_i^T y_i} - \frac{g_{i+1}^T v_i}{2\sqrt{(1 + ||x_{i+1}||)}} + \theta d_i^T y_i \tag{16}
\]
3. Algorithm, Descent Property and Convergence Analysis of New Method

3.1. The Algorithm of New Conjugate Gradient Method:

The New CG-method is summarized below by following steps:

Step 0: Let \( i = 0 \), select \( x_0 \in \mathbb{R}^n \) is an initial point, \( n \in \mathbb{Z} \) and \( \epsilon > 0 \).

Step 1: Test a criterion for stopping, if \( \|g_0\| < \epsilon \) then stop.

else go to step (5).

Step 2: Do cubic line search to calculate step length \( (\lambda_i) \) with Wolfe line searches (3) and (4).

Step 3: Determine \( \|g_{i+1}\| < \epsilon \) and \( x_{i+1} \) is defined in (15) or (16).

Step 4: Stop if \( \|g_{i+1}\| < \epsilon \) and \( x_{i+1} \) is a minimizer, else go to step (5).

Step 5: Calculate \( d_{i+1} = -g_{i+1} + \beta_i d_i \), \( \beta_i \) is defined in (18).

Step 6: If \( \|g_{i+1}\| > 0.2 \|g_{i+1}g_{i+1}\| \) , then go to step 1 else \( i := i + 1 \), and go to Step 2.

3.2. Descent and Sufficient Descent Conditions

To ensure that the new algorithm is convergence we needed to prove the following theorems:

**Theorem3.2.1:** Consider the sequences of \( \{x_i, d_i, g_i\} \) are generated by New CG method then the search direction \( \{d_i\} \) satisfies the descent property

\[
d^T_{i+1}g_{i+1} \leq 0 \tag{18}
\]

**proof:** (by induction.)

when \( i = 0 \), \( d_0 = -g_0 \) , so

\[
d_i^Tg_0 \leq -\|g_0\|^2. \tag{19}
\]

Assume the induction hypothesis that for a particular \( \lambda \) meaning \( d_i^Tg_i < 0 \) is true for any \( i > 0 \), now we prove (18) is hold at \( i + 1 \), from the definition of new direction we have

\[
d_{i+1} = -g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d_i \tag{20}
\]

Multiply both sides of (20) by \( g_i^T y_i \) from right to get,

\[
d^T_{i+1}g_{i+1} = -g^T_{i+1}g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i g_{i+1} \tag{21}
\]

\[
= -\|g_{i+1}\|^2 + \frac{g_i^T y_i}{\|v_i\|} d_i^T g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i g_{i+1} \tag{22}
\]

\[
= -\|g_{i+1}\|^2 + \frac{g_i^T y_i}{\|v_i\|} d_i^T g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i g_{i+1} \tag{23}
\]

Since, \( \lambda_i \) is parameter greater than zero so,

\[
d^T_{i+1}g_{i+1} = -\|g_{i+1}\|^2 + \frac{g_i^T y_i}{\|v_i\|} d_i^T g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i g_{i+1} \tag{24}
\]

it is clear that the first two terms of (24) refer to Hestenes and Stiefel method generates the descent directions, now we need to prove that the third term of (24) is less than or equal to zero.

Noticeably, all of \( \lambda_i \), \( (d^T_i g_{i+1})^2 \), \( \|v_i\| \), and \( \|x_{i+1}\| \) are positive and \( \theta \in (0,1) \). Noteworthy that

\[
d^T_i y_i = d^T_i (g_{i+1} - g_i) > (\xi_2 - 1) d^T_i g_i \tag{25}
\]

\((\xi_2 - 1) d^T_i g_i > 0\) therefor \( d^T_i y_i > 0\). So, the third term of (24) is also less than of zero.

In case, when the orthogonal property is satisfy, we see

\[
d_{i+1} = -g_{i+1} + \frac{\|g_{i+1}\|^2}{\|y_i\|} d^T_i y_i - \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i y_i \tag{25}
\]

Multiply both sides (25) by \( g^T_{i+1} \), we get

\[
d^T_{i+1}g_{i+1} = -\|g_{i+1}\|^2 + \frac{\|g_{i+1}\|^2}{\|y_i\|} d^T_i g_{i+1} + \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} d^T_i g_{i+1} \tag{26}
\]

The terms \(-\|g_{i+1}\|^2 + \frac{\|g_{i+1}\|^2}{\|y_i\|} d^T_i g_{i+1} \) proved is the descent property by Dai and Yuan [2] and we proved term \(- \frac{\left(g_i^T y_i \right) - \frac{g_i^T y_i}{(1-\theta)\|v_i\|}}{((1-\theta)\|v_i\| + \theta)\|y_i\|} < 0\) in above . So the descent property is satisfied when \( g^T_{i+1}g_i = 0 \) Hence, we obtain

\[
g^T_{i+1}d_{i+1} \leq 0 \tag{27}
\]

the proof is completed by induction.

**Theorem3.2.2:** - let the sequences of \( \{x_i\} \) and \( \{d_i\} \) are generated by our proposed method of modified CG satisfies the descent property for all new CG- method is
also hold the sufficient condition:
\[ d^T_{i+1}g_{i+1} < -c\lVert g_{i+1} \rVert^2, \forall i \geq 0 \text{ and } c > 0 \] (28)

**Proof:**

Observably, the Hestenes and Stiefel method generates descent directions so, the term \(-\lVert g_{i+1} \rVert^2 + \frac{\theta}{d^T_{i+1}g_{i+1}}d^T_{i+1}g_{i+1}\) in (24) is less than or equal to zero, consequently we can rewrite the (24) as following:
\[ d^T_{i+1}g_{i+1} \leq -\frac{\lambda_i(d^T_{i+1}g_{i+1})^2}{(1-\theta)(\frac{\lVert v_i \rVert}{\lVert y_i \rVert})\theta + \lambda_i d^T_{i+1}g_{i+1}} \] (29)

Multiply and divide the right side of (29) by \(\lVert g_{i+1} \rVert^2\), we obtain
\[ d^T_{i+1}g_{i+1} \leq -c\lVert g_{i+1} \rVert^2 \] (30)
\[ c = \frac{\lambda_i(d^T_{i+1}g_{i+1})^2}{(1-\theta)(\frac{\lVert v_i \rVert}{\lVert y_i \rVert})\theta + \lambda_i d^T_{i+1}g_{i+1}} \text{, and according to theorem (3.2.1), we see } c > 0 \text{.} \]
Therefore
\[ d^T_{i+1}g_{i+1} \leq -c\lVert g_{i+1} \rVert^2 \] (31)
Therefore, new method satisfied the sufficient descent condition.

### 3.3. Global Convergence Analysis for New Method:

In order to prove the global convergence of new nonlinear conjugate gradient methods, under the Wolfe line search. We will impose the following assumptions and lemma about the objective function \(h(x)\).

**Assumptions (3.3.1):** -

(i). Let \(h(x): R^n \to R\) be twice continuously differentiable, and the level set \(L_x = \{x : x \in R^n, \text{and } h(x) \leq h(\bar{x})\}\) is closed and bounded.

(ii). In some neighborhood \(B\) of \(L_x\), the gradient of \(h(x)\) is \(\nabla h(x)\) and satisfying Lipschitz condition, namely there exists a constant \(\ell > 0\) such that:
\[ \lVert \nabla a \rVert - \lVert \nabla b \rVert \leq \ell \lVert a - b \rVert \text{ for all } a \text{ and } b \in \ell \] (32)

Noticeably that the \(\{x_i\}\) is generating by new CG-method (14) which is in a bounded and under assumptions (3.3.1) on \(h(x)\), there exists a constant \(\gamma > 0\) such that \(\lVert \nabla h(x) \rVert \leq \gamma\), \(\forall x \in B\).

**Lemma (3.3.1):** let hypotheses (3.3.1) is satisfied and the \(\lambda_i\) is taken by the Wolfe conditions (3) and (4), if
\[ \sum_{i=1}^{\infty} \frac{1}{\lVert d_{i+1} \rVert^2} = \infty \] (33)
Then,\[ \lim_{i \to \infty} (\inf(\lVert g_{i+1} \rVert)) = 0 \] (34)

**Theorem (3.3.1):**

let \(x_i\) be generated by new CG-method (14) and function \(h(x)\) is uniformly convex, if assumptions (3.3.1) holds then we have
\[ \lim_{i \to \infty} (\inf(\lVert g_{i+1} \rVert)) = 0 \] (35)

**Proof:**

from (14), we have
\[ |\beta_1\|^2 = \frac{|\theta_1^T g_{i+1} - \theta_1^T v_i\|^2}{|\theta_1^T g_{i+1} - \theta_1^T v_i\|^2} \] (36)
Let \(\Gamma = ((1 - \theta)(\frac{\lVert v_i \rVert}{\lVert y_i \rVert})\theta + \lambda_i d^T_{i+1}g_{i+1})\)
so
\[ \beta_1\] becomes
\[ |\beta_1\|^2 = \frac{|\theta_1^T g_{i+1} - \theta_1^T v_i\|^2}{|\theta_1^T g_{i+1} - \theta_1^T v_i\|^2} \] (37)
we know that the \(\beta_1^T d_i < \beta_1^T v_i\) and using triangle and Cauchy–Schwarz inequalities in (37) and we obtain
\[ |\beta_1\|^2 \leq \frac{\lambda_i^2}{\Gamma^2} \] (38)
and return to hypothesis (ii), and using \(\lVert g_{i+1} \rVert \leq \gamma\), we get
\[ |\beta_1\|^2 \leq \frac{\gamma}{\Gamma} + \frac{\lambda_i}{\Gamma} \] (39)
Taking norm for both sides of new direction in (17):
\[ \lVert d_{i+1} \rVert = \lVert g_{i+1} + \beta_1\lVert d_i \rVert \] (40)
Apply preliminary of triangle inequality in (40), and we get
\[ \lVert d_{i+1} \rVert \leq \lVert g_{i+1} \rVert + \lVert \beta_1\lVert \lVert d_i \rVert \] (41)
From (38) and (41), we obtain
\[ \lVert d_{i+1} \rVert \leq \gamma + \left(\frac{\gamma}{\lVert d_i \rVert} + \frac{\lambda_i}{\Gamma}\right) \lVert d_i \rVert \] (42)
implies that
\[ \lVert d_{i+1} \rVert \leq \gamma + \left(\gamma + \frac{\lVert v_i \rVert}{\Gamma}\right) \lVert d_i \rVert \] (43)
and \(\lVert v_i \rVert \leq \| x_{i+1} - x_i \| \), suppose \(M = \max\{\| x_{i+1} - x_i \|, x_i \in R^n\}\)
\[ \lVert d_{i+1} \rVert \geq \| H \|, \| H \| = \gamma + \left(\gamma + \frac{M}{\Gamma}\right) \] (44)
Thus,
\[ \sum_{i=1}^{\infty} \frac{1}{\lVert d_{i+1} \rVert^2} > \sum_{i=1}^{\infty} \frac{1}{\lVert d_{i+1} \rVert^2} \] (45)
which indicates
\[ \sum_{i=1}^{\infty} \frac{1}{\lVert d_{i+1} \rVert^2} = \infty \] (46)
By Lemma (3.1.1), we obtain
\[ \lim_{i \to \infty} (\inf(\lVert g_{i+1} \rVert)) = 0 \] (47)
A Modified Perry's Conjugate Gradient Method
Based on Powell's Equation for Solving Large-Scale Unconstrained Optimization

Table 1. Comparing performance of the three method s (New formula $\beta_{i_{new}}$, $\beta_{i_{perry}}^{Perry}$ and classical Hestenes and Stiefel $\beta_{i_{H/S}}$)

| Test | N    | $\beta_{i_{new}}$ | $\beta_{i_{H/S}}$ | $\beta_{i_{perry}}^{Perry}$ |
|------|------|-------------------|-------------------|-----------------------------|
|      |      | NOI- NOF          | NOI- NOF          | NOI- NOF                    |
| Powell (3,-1,0,1) | 4    | 32-87             | 38-108            | 35-89                       |
|      | 100  | 39-114            | 40-122            | 43-105                      |
|      | 500  | 33-92             | 41-124            | 43-105                      |
|      | 1000 | 38-105            | 41-124            | 45-120                      |
|      | 3000 | 38-104            | 41-124            | 46-122                      |
|      | 5000 | 41-122            | 41-124            | 46-122                      |
| Wood (-3,-1,3,-1) | 4    | 27-62             | 30-68             | 30-68                       |
|      | 100  | 28-64             | 30-68             | 30-68                       |
|      | 500  | 28-64             | 30-68             | 30-68                       |
|      | 1000 | 28-64             | 30-68             | 30-68                       |
|      | 3000 | 28-64             | 30-68             | 30-68                       |
|      | 5000 | 28-64             | 30-68             | 30-68                       |
| Rosen (-1,2,1;...) | 4    | 30-83             | 30-83             | 30-83                       |
|      | 100  | 30-83             | 30-83             | 30-83                       |
|      | 500  | 30-83             | 30-83             | 30-83                       |
|      | 1000 | 30-83             | 30-83             | 30-83                       |
|      | 3000 | 30-83             | 30-83             | 30-83                       |
|      | 5000 | 30-83             | 30-83             | 30-83                       |
| Powell (0,1,2,..) | 4    | 15-33             | 16-36             | 16-36                       |
|      | 100  | 16-36             | 16-36             | 16-37                       |
|      | 500  | 16-36             | 16-36             | 16-37                       |
|      | 1000 | 16-36             | 16-36             | 16-37                       |
|      | 3000 | 16-36             | 16-36             | 16-37                       |
|      | 5000 | 16-36             | 16-36             | 16-37                       |
| Gcantrel (1,2,2,2) | 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                       |
| Cubic (-1,2,1;...) | 4    | 3-11              | 3-11              | 3-11                        |
|      | 100  | 14-81             | 14-81             | 15-84                       |
|      | 500  | 21-124            | 21-124            | 22-132                      |
|      | 1000 | 23-128            | 23-128            | 23-125                      |
|      | 3000 | 29-150            | 27-128            | 32-168                      |
|      | 5000 | 30-147            | 31-159            | 34-178                      |
| Sum (2;...) | 4    | 17-36             | 17-36             | 17-36                       |
|      | 100  | 483-1091          | 7678-15360        | 7735-15474                  |
|      | 500  | 565-1255          | 490-1093          | 6264-12573                  |
|      | 1000 | 526-1182          | 4005-8015         | 3391-6787                   |
|      | 3000 | 474-1072          | 6202-12409        | 515-1129                    |
|      | 5000 | 479-1068          | 481-1080          | 6357-12717                  |
| Dixin (+1;...) | 4    | 23-73             | 28-85             | 34-133                      |
|      | 100  | 35-121            | 33-114            | 46-169                      |
|      | 500  | 39-141            | 40-146            | 52-198                      |
|      | 1000 | 46-171            | 46-176            | 58-229                      |
|      | 3000 | 50-202            | 54-211            | 58-229                      |
|      | 5000 | 48-185            | 54-211            | 64-261                      |
| Miele (1,2,2,2) | 4    | 11-24             | 11-24             | 11-24                       |
|      | 100  | 49-99             | 49-99             | 49-99                       |
|      | 500  | 52-105            | 52-105            | 52-105                      |
|      | 1000 | 61-123            | 70-141            | 70-141                      |
|      | 3000 | 134-281           | 170-351           | 170-351                     |
|      | 5000 | 153-317           | 165-348           | 166-350                     |
| Wolfe (+1;...) | 4    | 6-26              | 8-44              | 8-46                        |
|      | 100  | 47-171            | 49-185            | 51-193                      |
|      | 500  | 105-314           | 112-353           | 107-327                     |
|      | 1000 | 154-460           | 156-473           | 156-473                     |
|      | 3000 | 196-597           | 197-602           | 197-602                     |
|      | 5000 | 256-774           | 256-774           | 256-774                     |
### Table 2. Comparing performance profiles of New method $\beta_{i}^{\text{new}}$ and Dai and Yuan method $\beta_{i}^{D/Y}$

| Test     | N   | $\beta_{i}^{\text{new}}$ | $\beta_{i}^{D/Y}$ |
|----------|-----|----------------------------|-------------------|
|          |     | NOI- NOF                   | NOI- NOF          |
| Powell   | 4   | 36-96                      | 50-128            |
| (-3,-1,0,1) | 100 | 37-97                      | 51-130            |
|          | 500 | 37-97                      | 51-130            |
|          | 1000| 37-97                      | 51-130            |
|          | 3000| 37-98                      | 52-132            |
|          | 5000| 37-98                      | 52-132            |
| Wood     | 4   | 27-63                      | 28-65             |
| (-3,-1,-3,-1) | 100 | 27-63                      | 28-65             |
|          | 500 | 27-63                      | 29-68             |
|          | 1000| 27-63                      | 29-68             |
|          | 3000| 29-68                      | 29-68             |
|          | 5000| 29-68                      | 29-68             |
| Rosen    | 4   | 27-63                      | 28-65             |
|          | 100 | 28-65                      | 28-65             |
|          | 500 | 28-65                      | 28-65             |
|          | 1000| 28-65                      | 28-65             |
|          | 3000| 30-69                      | 30-69             |
|          | 5000| 30-69                      | 30-69             |
| Geantrel | 4   | 14-85                      | 18-127            |
| (1,2,2,2) | 100 | 28-240                     | 20-153            |
|          | 500 | 16-109                     | 23-192            |
|          | 1000| 16-109                     | 23-192            |
|          | 3000| 17-125                     | 24-205            |
|          | 5000| 17-125                     | 24-205            |
| Dixon    | 4   | 13-28                      | 13-28             |
| (-1;...) | 100 | 460-994                    | 466-1021          |
|          | 500 | 430-927                    | 503-1085          |
|          | 1000| 451-977                    | 484-1048          |
|          | 3000| 432-937                    | 462-1005          |
|          | 5000| 456-989                    | 510-1115          |
| Miele    | 4   | 34-109                     | 36-115            |
| (1,2,2,2) | 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            |
| Wolfe    | 4   | 11-23                      | 11-23             |
| (-1;...) | 100 | 44-89                      | 45-91             |
|          | 500 | 49-99                      | 48-97             |
|          | 1000| 49-99                      | 52-105            |
|          | 3000| 112-237                    | 125-263           |
|          | 5000| 132-284                    | 159-327           |
| OSP      | 4   | 8-44                       | 8-44              |
| (1;...)  | 100 | 51-173                     | 52-180            |
|          | 500 | 136-427                    | 138-439           |
|          | 1000| 187-576                    | 196-607           |
|          | 3000| 387-1308                   | 421-1388          |
|          | 5000| 529-1806                   | 555-1857          |

### 4. Numerical Results

It is important to take note that the theoretical prove it is not adequate to show the effectiveness or robust of any conjugate gradient method. Thus, there would be need to study the method numerical by evaluate the performance method on group of test problems and evaluation the number of iteration or computation time (CPU time).

In this section, is recorded some numerical results by comparing the suggeston method with classical conjugate gradient methods. All of the methods are coded in Fortran95 and stopping the iteration when this statement $\|g_{i+1}\| < 10^{-5}$ is true. For testing we used the well-known nonlinear problems with dimension ranging between 4 to 5000 [1], all methods are using cubic fit mothod to find the steplength $\lambda_i > 0$ under conditions (3) and (4) with $\xi_1 = 0.001$ and $\xi_2 = 0.1$.

The following abbreviates in tables meaning:
N:- the number of variable in test problem.
NOI:- the number of iterations.
NOF: the number of function evaluations.

In table 1, the results show that the performance of the new method is better than the HS and Perry methods based on the total numbers of function evaluations and its gradient evaluations.

As a final note, table 2 demonstrates the performance of new method when the orthogonal condition is available $⟨g_{k+1}, g_k⟩ = 0$. During the numerical experiments outcome from comparing between New method $β_{new}$ and Dai and Yuan method $β^D/Y$ and see that the new method is top perform to Dai and Yuan method.

5. Conclusions

In this work, we present a modified Perry's conjugacy coefficient for conjugate gradient method. The proposed new method is apposite for solving large-scale unconstrained optimization problems because of its lower storage requirement. Furthermore, and under suitable conditions we show that the new method satisfy the descent, sufficient descent conditions and it is global convergent also the new method gives auspicious numerical results.

In future it can be used the new method to train the neural network in order to improve performance and in the image processing.

REFERENCES

[1] Andrei, Neculai. "An Unconstrained Optimization Test Functions Collection." Adv. Model. Optim 10.1: 147-161, 2008.

[2] Dai Y. H. and Yuan Y., "A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Propert". SIAM Journal on Optimization, vol. 10, no. 1, pp. 177–182, 1999.

[3] Dai, Y. and Yuan, Y. A "nonlinear conjugate gradient with a strong global convergence properties". SIAM Journal on Optimization, 10, 177-182. doi:10.1137/S1052623497318992

[4] Fletcher, R. "Practical Method of Optimization", Vol I: Unconstrained Optimization. 2nd Edition, Wiley, New York. 1997

[5] Fletcher, R.M., Reeves, C.M. "Function Minimization by Conjugate Gradients". Computer Journal, 7, 149-154. doi:10.1093/comjnl/7.2.149

[6] Hestenes, M.R., Stiefel, E.: "Methods of Conjugate Gradients for Solving Linear Systems", Journal of Research National Bureau of Standards. 49(6), 409–439, 1952.

[7] Liu, Dongyi, and Genqi Xu. "A Perry Descent Conjugate Gradient Method with Restricted Spectrum". Optimization Online, Nonlinear Optimization (unconstrained optimization):1-19.

[8] Perry, A.: "A Modified Conjugate Gradient Algorithm". Oper. Res., Tech. Notes 26(6), 1073–1078, 1978.

[9] Polak, E., Ribière, G.: "Note Sur La Convergence De Méthodes De Directions Conjuguées". Rev. Fr. Inform. Rech. Oper. 3(16), 35–43, 1969.

[10] Powell M. J. D.: "A Fast Algorithm for Nonlinearly Constrained Optimization Calculations", Lecture Notes in Mathematics, Vol. 630, 1, pp. 144-157, 1978.

[11] Sulaiman I. M., Mamat M., Waziri M. Y., Yakubu U. A., Malik M., "The Performance Analysis of a New Modification of Conjugate Gradient Parameter for Unconstrained Optimization Models," Mathematics and Statistics, Vol. 9, No. 1, pp. 16-23, 2021. DOI: 10.13189/ms.2021.090103.

[12] Wolfe P. : "Convergence conditions for ascent methods", SIAM Rev., 11, pp. 226–235,1969.

[13] Zoutendijk G., "Nonlinear Programming, Computational Methods in Integer and Nonlinear Programming", J. Abadie, ed., North-Holland, Amsterdam, pp. 37–86, 1970.