Available online at sjuoz.uoz.edu.krd

journals.uoz.edu.krd

Science Journal of University of Zakho

Vol. 10, No. 4, pp. 169 –174, October-December 2022

p-ISSN: 2663-628X
e-ISSN: 2663-6298

A NEW MODIFIED CONJUGATE GRADIENT FOR NONLINEAR MINIMIZATION PROBLEMS

Hussein A. Khatab [a, *] and Salah G. Shareef [a]

[a] Dept. of Mathematics, Faculty of Science, University of Zakho, Zakho, Kurdistan Region - Iraq
(Hussein.khatab@uoz.edu.krd ; salah.shareef@uoz.edu.krd)

Received: *19 May, 2022 /* ***Accepted:*** *6 Aug., 2022 /* ***Published:*** *17 Oct., 2022* ***https://doi.org/10.25271/sjuoz.2022.10.4.933***

ABSTRACT:

The conjugate gradient is a highly effective technique to solve the unconstrained nonlinear minimization problems and it is one of the most well-known methods. It has a lot of applications. For large-scale and unconstrained minimization problems, conjugate gradient techniques are widely applied. In this paper, we will suggest a new parameter of conjugate gradient to solve the nonlinear unconstrained minimization problems, based on the parameter of Dai and Liao. We will study the property of the descent , the property of the sufficient descent and property of the global convergence of the new method. We introduce some numerical data to prove the efficacy of the our method.

KEYWORDS: Unconstrained Nonlinear minimization, Method of Conjugate Gradient, Descent property, property of the Sufficient Descent and property of the Global convergence.

1. INTRODUCTION

We are interested to consider the following nonlinear unconstrained minimization problem

$$\operatorname{Min} f(x)\ ,\ x \in R^n \tag{1}$$

where $f: R^n \to R$ is a continuously differentiable real-valued function.. A nonlinear unconstrained conjugate gradient algorithms to solve (1) are iterative technics and have the form

$$x_{m+1} = x_m + \alpha_m d_m \tag{2}$$

Which is starting from an initial guess $x_1 \in R^n$, where $v_m = x_{m+1} - x_m$, the positive step length α_m is calculated by one dimensional line search and d_m is the search direction . The conjugate gradient technics have two search directions: the first one for the first iteration is the same of the search direction of steepest descent technic , which has the form:

$$d_1 = -g_1 \tag{3}$$

and the other search direction has the form:

$$d_{m+1} = -g_{m+1} + \beta_m d_m \tag{4}$$

where $g_m = \nabla f(x_m)$ and β_m is a scalar. There are A lot of formulas of parameter β_m are suggested. The most well-know formulas of β_m are called Hestenes-Stiefel (H/S) (Hestenes, M.R. & Stiefel, 1952), Liu–Storey (Liu, Y. & Storey, 1991), Dai – Liao (Dai, Y.H. & Liao, 2001), Polak-Ribiere-Polyak (PRP) (Polak, E. & Ribiere, 1969), Dai - Yuan (D/Y) (Dai, Y.H. & Yuan, 1999) and Fletcher-Reeves (F/R) (Fletcher, R., & Reeves, 1964) which have the following forms:

$$\beta_m^{HS} = \frac{g_{m+1}^T y_m}{d_m^T y_m} \tag{5}$$

$$\beta_m^{LS} = \frac{g_{m+1}^T y_m}{-d_m^T g_m} \tag{6}$$

$$\beta_m^{DL} = \frac{g_{m+1}^T (y_m - t v_m)}{d_m^T y_m},\ t > 0 \tag{7}$$

$$\beta_m^{PRP} = \frac{g_{m+1}^T (g_{m+1} - g_m)}{\|g_m\|^2} \tag{8}$$

$$\beta_m^{DY} = \frac{\|g_{m+1}\|^2}{d_m^T (g_{m+1} - g_m)} \tag{9}$$

$$\beta_m^{FR} = \frac{\|g_{m+1}\|^2}{\|g_m\|^2} \tag{10}$$

where $y_m = g_{m+1} - g_m$. The symbol $\|.\|$ means the Euclidean norm. Many authors have researched the property of global convergence of the parameters FR, PRP, and HS without regular restarts, including Zoutendijk (Zoutendijk, 1970), Dai – Yuan (Dai , Y. H. & Yuan, 1996), Gilbert - Nocedal (Gilbert J. C. & Nocedal, 1992).

A lot of parameters of coefficient conjugate gradient methods are suggested. Hager - Zhang (Hager, W, & Zhang, 2005) suggested a new conjugate gradient and called CG-DESCENT method. Zhang Li et al. (Zhang L., Zhou W.J., 2006)(Zhang, L. , Zhou, W. & Li, 2006) also suggested some modified conjugate gradient methods. The convergence property of the nonlinear conjugate gradient method for unconstrained minimization problems has been researched by many authors such as (Dai, Y.H. , Han, J., Liu, G., Sun, D., Yin, H. & Yuan, 2000),(Grippo, L. & Lucidi, 1997),(Deng, N.Y. & Li, 1995),(Hu, Y.F. & Storey, 1991) and (Powell, 1977).

The following is the outline of the paper. We will derive a new modification of the conjugate gradient in section 2 together with present the algorithm. The property of the descent and sufficient descent of a proposed algorithm are verified in section 3. Section 4 shows the convergence analysis. Finally, the numerical results and conclusions will be presented.

* Corresponding author

2. DERIVATION OF NEW METHOD (β_m^{New})

A new parameter in this part is derived by using the equation:

$$y_m^* = y_m + \mu_m v_m.$$

Consider the equation bellow:

$$y_m^* = y_m + \mu_m v_m \quad (11)$$

Suppose that $\mu_m = \tau * \frac{N_m}{1+N_m}$, $N_m = \frac{y_m^T v_m}{\|v_m\|}$ and $\tau \in (0,1)$

We replace y_m in equation (7) by y_m^*, here we get a new parameter as follows:

$$\beta_m^{New} = \frac{g_{k+1}^T(y_m^* - t_m v_m)}{d_m^T y_m^*} \quad (12)$$

Implies that

$$\beta_m^{New} = \frac{g_{m+1}^T(y_m + \mu_m v_m - t_m v_m)}{d_m^T(y_m + \mu_m v_m)} \quad (13)$$

Here, we get

$$\beta_m^{New} = \frac{g_{m+1}^T(y_m + (\mu_m - t_m) v_m)}{d_m^T(y_m + \mu_m v_m)} \quad (14)$$

So ,

$$\beta_m^{New} = \frac{g_{m+1}^T y_m}{d_m^T(y_m + \mu_m v_m)} + \frac{(\mu_m - t_m) g_{m+1}^T v_m}{d_m^T(y_m + \mu_m v_m)} \quad (15)$$

Where $t_m = \frac{N_m}{1+N_m}$, $\mu_m = \tau * \frac{N_m}{1+N_m}$, $N_m = \frac{y_m^T v_m}{\|v_m\|}$ and $\tau \in (0,1)$

2.1 Algorithm of the New Method

Step 1 : Select x_1, $\varepsilon = 10^{-5}$ and $\tau \in (0,1)$

Step 2 : Set $d_1 = -g_1$, $g(x_m) = \nabla f(x_m)$. Set $m = 1$. $If\ \|g_1\| \le \varepsilon$, then stop.

Step 3 : Calculate the step length $\alpha_m > 0$ that satisfies the Wolfe line search requirements.

$$f(x_m + \alpha_m d_m) - f(x_m) < k_1 \alpha_m g_m^T d_m$$

$$|g_{m+1}^T d_m| < k_2 |g_m^T d_m|$$

where, $0 < k_1 < k_2 < 1$.

Step 4 : Calculate $x_{m+1} = x_m + \alpha_m d_m$

$g_{m+1} = \nabla f(x_{m+1})$,

$If\ \|g_{m+1}\| \le \varepsilon$, then stop.

Step 5 : Compute $\mu_m = \tau * \frac{N_m}{1+N_m}$,

$N_m = \frac{y_m^T v_m}{\|v_m\|}$ and β_m^{New} by (15)

Step 6 : Compute $d_{m+1} = -g_{m+1} + \beta_m^{New} d_m$

Step 7 : If $|g_{m+1}^T g_m| > 0.2\|g_{m+1}\|^2$, then go to step2 .
Else
$m = m + 1$, then, go to step 3.

3. DECENT AND SUFFICIENT DESCENT PROPERTIES OF THE NEW ALGORITHM

Theorem 1:- If equation (2) generates the sequence $\{x_m\}$, then the descent requirement is satisfied by the search direction in (4) with new conjugate gradient (15), i.e. $d_{m+1}^T g_{m+1} \le 0$ with two cases of line search , exact and inexact.

Proof:- From equations (4) and (15) we have,

$$d_{m+1} = -g_{m+1} + \left(\frac{g_{m+1}^T y_m}{d_m^T(y_m + \mu_m v_m)} + \frac{(\mu_m - t_m) g_{m+1}^T v_m}{d_m^T(y_m + \mu_m v_m)}\right) d_m \quad (16)$$

Multiplying both sides of equation (16) by g_{m+1} , to obtain

$$d_{m+1}^T g_{m+1} = -\|g_{m+1}\|^2 + \left(\frac{g_{m+1}^T y_m}{d_m^T(y_m + \mu_m v_m)} + \frac{(\mu_m - t_m) g_{m+1}^T v_m}{d_m^T(y_m + \mu_m v_m)}\right) d_m^T g_{m+1}$$

Since $v_m = \alpha_m d_m$, we get

$$d_{m+1}^T g_{m+1} = -\|g_{m+1}\|^2 + \frac{g_{m+1}^T y_m}{d_m^T y_m + \mu_m d_m^T v_m} d_m^T g_{m+1} + \frac{\alpha_m(\mu_m - t_m)}{d_m^T y_m + \mu_m d_m^T v_m} (d_m^T g_{m+1})^2 \quad (17)$$

We notice that if we use an exact line search then, we get $d_{m+1}^T g_{m+1} \le 0$.

If we use an inexact line search to calculate the step size, here, we will prove by mathematical induction, $d_0^T g_0 = -\|g_0\|^2 \le 0$, where $d_0 = -g_0$, and we assume that it is true for case m that means $d_m^T g_m \le 0$. To prove case $m + 1$, consider the equation

$$d_{m+1}^T g_{m+1} = -\|g_{m+1}\|^2 + \frac{g_{m+1}^T y_m}{d_m^T y_m + \mu_m d_m^T v_m} d_m^T g_{m+1} + \frac{\alpha_m(\mu_m - t_m)}{d_m^T y_m + \mu_m d_m^T v_m} (d_m^T g_{m+1})^2 \quad (18)$$

The right-hand side of the equation (18), the first term is clearly smaller than zero, the third term also is negative because $\mu_m < t_m$ and $d_m^T y_m > 0$. For the second term, we have two cases. Case one, if it is negative then the proof is done. The second case, if it is positive, we can write the equation (18) as follows

$$d_{m+1}^T g_{m+1} \le -\|g_{m+1}\|^2 + \frac{g_{m+1}^T y_m}{d_m^T y_m} d_m^T g_{m+1} + \frac{\alpha_m(\mu_m - t_m)}{d_m^T y_m + \mu_m d_m^T v_m} (d_m^T g_{m+1})^2 \quad (19)$$

Since the parameter of H/S satisfies the descent condition, the first and second terms from the right side of the inequality (19) are less than or equal to zero, and the third term also is negative because $\mu_m < t_m$.
Then, $d_{m+1}^T g_{m+1} \le 0$

Theorem 2:- If (2) generates $\{x_k\}$, then the search direction in (4) with the new conjugate gradient (15) satisfies the sufficient descent requirement.

$$d_{m+1}^T g_{m+1} \le -k_3 \|g_{m+1}\|^2.$$

Proof:- Multiplying both sides of equation (16) by g_{m+1} and $v_m = \alpha_m d_m$, then, we get:

$d_{m+1}^T g_{m+1} = -\|g_{m+1}\|^2 + \frac{g_{m+1}^T y_m}{d_m^T y_m + \mu_m d_m^T v_m} d_m^T g_{m+1} + \frac{\alpha_m(\mu_m - t_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2$ (20)

We know that the first and third terms of (20) are negative .

If the second term is negative, then, we have the following inequality:

$d_{k+1}^T g_{k+1} \le -\|g_{m+1}\|^2 + \frac{\alpha_m(\mu_m - t_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2$ (21)

Here, we can write (21) as follows

$$d_{k+1}^T g_{k+1} \le -(\|g_{m+1}\|^2 + (\frac{\alpha_m(t_m - \mu_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2) * \frac{\|g_{m+1}\|^2}{\|g_{m+1}\|^2}$$

Where $t_m - \mu_m > 0$

let $k_3 = \left(\|g_{m+1}\|^2 + \frac{\alpha_m(t_m - \mu_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2\right) * \frac{1}{\|g_{m+1}\|^2}$

so,

$$d_{m+1}^T g_{m+1} \le -k_3 \|g_{m+1}\|^2. \blacksquare$$

If the second term is positive, then we get (19), and we can write it as follows:

$$d_{m+1}^T g_{m+1} \le -\frac{\alpha_m(t_m - \mu_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2 \quad (22)$$

Because the parameter of (HS) satisfies the descent condition,

here we can write (22) as follows:

$$d_{m+1}^T g_{m+1} \le -(\frac{\alpha_m(t_m - \mu_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2) * \frac{\|g_{m+1}\|^2}{\|g_{m+1}\|^2}$$

Let $k_3 = (\frac{\alpha_m(t_m - \mu_m)}{d_m^T y_m + \mu_m d_m^T v_m}(d_m^T g_{m+1})^2) * \frac{1}{\|g_{m+1}\|^2}$

$$d_{m+1}^T g_{m+1} \le -k_3 \|g_{m+1}\|^2. \blacksquare$$

4. CONVERGENCE ANALYSIS OF THE NEW METHOD

Assume that:

a. The level set $S_1 = \{x : x \in R^{n*1}, f(x) \le f(x_1)\}$ is bounded, where x_1 represents the beginning point.

b. In a neighborhood Ω of S_1, f is continuously differentiable and its gradient g is Lipschitz continuously, namely, there exists a constant $L_1 > 0$ such that

$$\|g(x) - g(x_m)\| \le L_1 \|x - x_m\|, \forall x, x_m \in \Omega \quad (23)$$

There exists a constant $\rho \ge 0$ with the above hypotheses on f, such that $\|g(x)\| \le \rho,\ \forall x \in S_1$.

Some researchers proved that for any conjugate gradient method with strong Wolfe line search the following general result holds:

Lemma 1: Suppose that the assumptions (a) and (b) hold and consider any conjugate gradient (2) , (3) and (4), where d_m is a descent direction and α_m is calculated by the strong Wolfe conditions.

$$f(x_m + \alpha_m d_m) - f(x_m) \le k_1 \alpha_m g_m^T d_m \quad (23)$$

$$|g_{m+1}^T d_m| \le k_2 g_m^T d_m \quad (24)$$

If

$$\sum_{k \ge 1} \frac{1}{\|d_m\|^2} = \infty \quad (25)$$

Then

$$\lim_{k \to \infty} inf \|g_m\| = 0 \quad (26)$$

There exists a constant $\vartheta \ge 0$ such that for all $x, y \in \Omega$:

$$\left(g(x) - g(y)\right)^T (x - y) \ge \vartheta \|x - y\|^2 \quad (27)$$

if f is a uniformly convex function.

We can rewrite (27) in the following manner:

$$y_m^T v_m \ge \vartheta \|v_m\|^2 \quad (28)$$

Theorem 3: Assume that the hypotheses a and b hold and that f is a uniformly convex function. The new algorithm of the form (2), (4) and (15) where d_k satisfies the property of the descent and α_m is computed by the strong Wolfe conditions, (23) and (24) satisfies the property of the global convergence.

(i.e.) $\lim_{m \to \infty} inf \|g_m\| = 0$

Proof: Since

$$d_{m+1} = -g_{m+1} + \beta_m^{New} d_m \quad (29)$$

$|\beta_m^{New}| = \left|\frac{g_{m+1}^T y_m}{d_m^T (y_m + \mu_m v_m)} + \frac{(\mu_m - t_m) g_{m+1}^T v_m}{d_m^T (y_m + \mu_m v_m)}\right|$ Implies that

$$|\beta_m^{New}| \le \left|\frac{g_{m+1}^T y_m}{d_m^T y_m + \mu_m d_m^T v_m}\right| + \left|\frac{(\mu_m - t_m) g_{m+1}^T v_m}{d_m^T y_m + \mu_m d_m^T v_m}\right| \quad (30)$$

Since $d_m^T y_m \ge \frac{\vartheta}{\alpha_m} \|v_m\|^2$, then , we can write equation (30) as follows:

$$|\beta_m^{New}| \le \left|\frac{g_{m+1}^T y_m}{\frac{\vartheta}{\alpha_m}\|v_m\|^2 + \mu_m d_m^T v_m}\right| + \left|\frac{(\mu_m - t_m) g_{m+1}^T v_m}{\frac{\vartheta}{\alpha_m}\|v_m\|^2 + \mu_m d_m^T v_m}\right|$$

It implies that

$$|\beta_m^{New}| \le \frac{\|g_{m+1}\|\|y_m\|}{\left|\frac{\vartheta}{\alpha_m}\|v_m\|^2+\mu_m d_m^T v_m\right|} + \frac{|(\mu_m - t_m)|\|g_{m+1}\|\|v_m\|}{\left|\frac{\vartheta}{\alpha_m}\|v_m\|^2+\mu_m d_m^T v_m\right|}$$

Here we get

$$|\beta_m^{New}| \le \frac{\|g_{m+1}\|\|y_m\|}{\left|\frac{\vartheta}{\alpha_m}\|v_m\|^2+\frac{\mu_m}{\alpha_m}\|v_m\|^2\right|} + \frac{|(\mu_m - t_m)|\|g_{m+1}\|\|v_m\|}{\left|\frac{\vartheta}{\alpha_m}\|v_m\|^2+\frac{\mu_m}{\alpha_m}\|v_m\|^2\right|}$$

So, we have

$$|\beta_m^{New}| \le \frac{\|g_{m+1}\|L_1\|v_m\|}{\left|\frac{\vartheta+\mu_m}{\alpha_m}\right|\|v_m\|^2} + \frac{|(\mu_m - t_m)|\|g_{m+1}\|\|v_m\|}{\left|\frac{\vartheta+\mu_m}{\alpha_m}\right|\|v_m\|^2}$$

and

$$|\beta_m^{New}| \le \frac{\|g_{m+1}\|L_1}{\left|\frac{\vartheta+\mu_m}{\alpha_m}\right|\|v_m\|} + \frac{|(\mu_m - t_m)|\|g_{m+1}\|}{\left|\frac{\vartheta+\mu_m}{\alpha_m}\right|\|v_m\|}$$

Then

$$|\beta_m^{New}| \le \frac{\rho L_1}{w_2 * \|v_m\|} + \frac{w_1 L_1}{w_2 * \|v_m\|}$$

Where $w_1 = |(\mu_m - t_m)|$ and $w_2 = \left|\frac{\vartheta+\mu_m}{\alpha_m}\right|$

Here , $\|d_{m+1}\| \le \|g_{m+1}\| + |\beta_m^{New}|\|d_m\|$

so,

$\|d_{m+1}\| \le \rho + (\frac{\rho L_1}{w_2 * \|v_m\|} + \frac{w_1 L_1}{w_2 * \|v_m\|})\|d_m\|$ (31)

We know that $\|d_m\| = \frac{\|v_k\|}{\alpha_m}$, so, we get

$$\|d_{m+1}\| \le \rho + (\frac{\rho L_1}{w_2 * \alpha_m} + \frac{w_1 L_1}{w_2 * \alpha_m}) \quad (32)$$

The above inequality becomes

$$\|d_{m+1}\| \le \rho + (\frac{\rho L_1}{w_2 * \alpha_m} + \frac{w_1 L_1}{w_2 * \alpha_m}) = \varphi$$

Then,

$$\sum_{m\ge1}\frac{1}{\|d_{m+1}\|^2} \ge \sum_{m\ge1}\frac{1}{\varphi^2} = \sum_{m\ge1} 1 = \infty$$

And

$\sum_{m\ge1}\frac{1}{\|d_{m+1}\|^2} = \infty$ .By using lemma (1), we get $\lim_{m\to\infty} inf\|g_{m+1}\| = 0$

5. NUMERICAL RESULTS

We present the numerical results of a number of nonlinear problems using the new method in this part. We compare the new method with the standard conjugate gradient algorithms (HS) and (DL), the comparative tests contain nonlinear unconstrained problems (standard test function) with different dimensions $4 \le n \le 5000$. FORTRAN 90 is used to write all of the programs. The stopping condition is $\|g_{m+1}\| \le 10^{-5}$. The number of the functions (NOF) and the number of the iterations (NOI) are notably mentioned in tables 1 through 3. More experimental results in tables 1–3 show that the new method surpasses standard conjugate gradient methods (HS) and (DL) in respect of NOI and NOF.

Table 1: Comparative Performance of the algorithms (HS and New algorithm)

Test function	Dim.	Algorithm of HS		New algorithm β_k^{NEW}	
		NOI	NOF	NOI	NOF
GCentral	4	22	159	22	158
	10	22	159	22	158
	50	22	159	22	158
	100	22	159	22	158
	500	23	171	23	172
	1000	23	171	22	159
	2000	27	234	22	163
	3000	27	234	22	163
	5000	28	248	24	193
Miele	4	28	85	41	119
	10	31	102	47	139
	50	31	102	42	141
	100	33	114	48	171
	500	40	146	41	149
	1000	46	176	25	82
	2000	54	211	33	102
	3000	54	211	30	97
	5000	54	211	49	175
Wood	4	30	68	27	62
	10	30	68	28	64
	50	30	68	28	63
	100	30	68	27	62
	500	30	68	28	64
	1000	30	68	28	63
	2000	30	68	30	68
	3000	30	68	30	68
	5000	30	68	30	68
Powell	4	38	108	33	86
	10	38	108	33	86
	50	38	108	36	100
	100	40	122	36	100
	500	41	124	36	106
	1000	41	124	36	106
	2000	41	124	36	105
	3000	41	124	36	106
	5000	41	124	36	106
Dixon	4	17	36	17	36
	10	21	45	21	45
	50	4044	8090	466	1014
	100	7678	15360	511	1135
	500	499	1122	507	1137
	1000	2366	4737	492	1093
	2000	519	1159	461	1021
	3000	6202	12409	420	962
	5000	476	1068	504	1134
Extended Psc1	4	7	18	6	16
	10	6	16	6	16
	50	6	16	6	16
	100	7	18	6	16
	500	7	18	6	16
	1000	7	18	6	16
	2000	7	18	6	16
	3000	7	18	6	16
	5000	7	18	6	16
Fred	4	8	23	6	19
	10	8	23	6	18
	50	8	23	6	18

	100	8	23	6	18
	500	8	23	6	18
	1000	8	23	6	18
	2000	8	23	6	18
	3000	8	23	6	18
	5000	8	23	6	18
Non Diagonal	4	24	64	21	57
	10	26	72	27	72
	50	29	79	24	67
	100	29	79	24	67
	500	F	F	24	67
	1000	29	79	24	67
	2000	29	79	24	67
	3000	29	79	24	67
	5000	30	81	24	67
Rosen	4	30	83	30	83
	10	30	83	29	81
	50	30	83	29	80
	100	30	83	29	80
	500	30	83	27	75
	1000	30	83	27	75
	2000	30	83	27	75
	3000	30	83	27	75
	5000	30	83	25	71
Wolfe	4	11	24	11	23
	10	32	65	32	65
	50	47	95	47	95
	100	49	99	49	99
	500	52	105	52	105
	1000	70	141	69	139
	2000	65	136	62	130
	3000	170	351	174	358
	5000	165	348	158	330
Total		24356	51923	5758	14661

Table 2: Comparative Performance of the algorithms (DL and New algorithm)

Test function	Dim.	Algorithm of DL, where $t_k = \frac{N_k}{1+N_k}$		New algorithm β_k^{NEW}	
		NOI	NOF	NOI	NOF
GCentral	4	22	159	22	158
	10	22	159	22	158
	50	22	158	22	158
	100	22	158	22	158
	500	23	170	23	172
	1000	23	171	22	159
	2000	27	234	22	163
	3000	27	232	22	163
	5000	28	250	24	193
Miele	4	41	137	41	119
	10	51	172	47	139
	50	57	202	42	141
	100	37	129	48	171
	500	41	150	41	149
	1000	40	146	25	82
	2000	44	164	33	102
	3000	49	190	30	97
	5000	45	166	49	175
Wood	4	30	68	27	62
	10	30	68	28	64
	50	30	68	28	63
	100	30	68	27	62
	500	30	68	28	64
	1000	30	68	28	63
	2000	30	68	30	68
	3000	30	68	30	68
	5000	30	68	30	68
Powell	4	35	88	33	86
	10	35	88	33	86
	50	35	90	36	100
	100	42	105	36	100
	500	47	115	36	106
	1000	43	107	36	106
	2000	43	107	36	105
	3000	43	107	36	106
	5000	45	124	36	106
Dixon	4	17	36	17	36
	10	21	45	21	45
	50	3072	6146	466	1014
	100	7768	15540	511	1135
	500	496	1123	507	1137
	1000	573	1248	492	1093
	2000	489	1101	461	1021
	3000	2906	5817	420	962
	5000	496	1125	504	1134
Extended Psc1	4	7	18	6	16
	10	6	16	6	16
	50	6	16	6	16
	100	7	18	6	16
	500	7	18	6	16
	1000	7	18	6	16
	2000	7	18	6	16
	3000	7	18	6	16
	5000	7	18	6	16
Fred	4	8	23	6	19
	10	8	23	6	18
	50	8	23	6	18
	100	8	23	6	18
	500	8	23	6	18
	1000	8	23	6	18
	2000	8	23	6	18
	3000	8	23	6	18
	5000	8	23	6	18
Non Diagonal	4	24	64	21	57
	10	26	72	27	72
	50	29	79	24	67
	100	29	79	24	67
	500	29	79	24	67
	1000	F	F	24	67
	2000	29	79	24	67
	3000	29	79	24	67
	5000	30	81	24	67
Rosen	4	30	83	30	83
	10	30	83	29	81
	50	30	83	29	80
	100	30	83	29	80
	500	30	83	27	75
	1000	30	83	27	75
	2000	30	83	27	75
	3000	30	83	27	75
	5000	30	83	25	71
Wolfe	4	11	24	11	23
	10	32	65	32	65
	50	47	95	47	95
	100	49	99	49	99
	500	52	105	52	105
	1000	70	141	69	139
	2000	65	136	62	130
	3000	170	351	174	358
	5000	166	350	158	330
Total		18366	40040	5758	14661

Table 3: Percentage of Improving of the New Method

	Algorithm of HS	New Algorithm
NOI	100%	23.6409919527%
NOF	100%	28.2360418312%
	Algorithm of DL	New Algorithm
NOI	100%	31.3514102145%
NOF	100%	36.6158841159%

4. CONCLUSION

We proposed a new conjugate gradient method to find the minimum of the nonlinear unconstrained minimization problems. We proved the descent and sufficient descent conditions of the new method, and we also studied the global convergence of the proposed method. The numerical tests were conducted on problems with low and high dimensionality, with comparisons made between different test functions. The effectiveness of the new method is seen in tables 1, 2, and 3.

REFERENCES

Dai, Y.H. , Han, J., Liu, G., Sun, D., Yin, H. & Yuan, Y. X. (2000). Convergence properties of nonlinear conjugate gradient methods. *SIAM J. Optim., 10*(2), 345–358.

Dai, Y.H. & Liao, l. z. (2001). New conjugacy conditions and related nonlinear conjugate gradient methods. *Appl. Math. Optim*, *43*, 87–101.

Dai, Y.H. & Yuan, Y. (1999). A nonlinear conjugate gradient with a strong global convergence properties. *SIAM J. Optim.*, *10*(1), 177–182.

Dai , Y. H. & Yuan, Y. (1996). Convergence properties of the Fletcher-Reeves method. *IMA J. Numer. Anal.*, *16*(2), 155–164.

Deng, N.Y. & Li, Z. (1995). Global convergence of three terms conjugate gradient methods. *Method and Sofeware, 4*, 273–282.

Fletcher, R., and Reeves, C. (1964). Function minimization by conjugate gradients. *Comput. J.*, *7*, 149–154.

Gilbert J. C. & Nocedal, J. (1992). Global convergence properties of conjugate gradient methods for optimization. *SIAM. J. Optim.*, *2*(1), 21–42.

Grippo, L. & Lucidi, S. (1997). A global convergent version of the Polak-Ribiere conjugate gradient method. *Math. Prog., 78*, 375–391.

Hager, W, & Zhang, H. C. (2005). A new conjugate gradient method with guaranteed descent and an efficient line search. *SIAM J. Optim*, *16*, 170–192.

Hestenes, M.R. & Stiefel, E. (1952). Method of conjugate gradient for solving linear equations. *J.Res. Nat. Bur. Stand*, *49*, 409–436.

Hu, Y.F. & Storey, C. (1991). Global convergence result for conjugate gradient methods. *J. Optim. Theory Appl.*, *71*, 399–405.

Liu, Y. & Storey, C. (1991). Efficient generalized conjugate gradient algorithms, Part 1: Theory. *JOTA*, *69*, 129–137.

Polak, E. & Ribiere, G. (1969). Note sur la convergence de directions conjugees. *Rev. Francaise Inform. Recherche Operationelle 3*, *16*, 35–43.

Powell, M. J. D. (1977). Restart procedures of the conjugate gradient method. *Math. Prog.*, *2*, 241–254.

Zhang, L. , Zhou, W. & Li, D. (2006). Global convergence of a modified Fletcher-Reeves conjugate gradient method with Armijo-type line search. *Numer. Math.*, *104*, 561–572.

Zhang L., Zhou W.J., L. D. H. (2006). A descent modified Polak-Ribiere-Polyak conjugate gradient method and its global convergence. *IMA Journal of Numerical Analysis*, *26*, 629–640.

Zoutendijk, G. (1970). Nonlinear programming, computational methods, in Integer and Nonlinear Programming. *J. Abadie*, *North-Holl*, 37–86.