Journal of Education and Science (ISSN 1812-125X), Vol: 29, No: 1, 2020 (325-335)

A New Modified Conjugate Gradient Method and Its Global Convergence Theorem

Aseel Moayad Qasim

Mathematics Department, College of Education of Pure Sciences, University of Mosul, Mosul, Iraq

E-mail: aseelqasim3@gmail.com

(Received March 24, 2013; Accepted June 25, 2013; Available online March 01, 2020)

DOI: 10.33899/edusj.2020.174696,

Abstract:

In this article, we try to proposed a new conjugate gradient method for solving unconstrained optimization problems, we focus on conjugate gradient methods applied to the non-linear unconstrained optimization problems, the positive step size is obtained by a line search and the new scalar to the new direction for the conjugate gradient method is derived from the quadratic function and Taylor series and by using quasi newton condition and Newton direction while deriving the new formulae. We also prove that the search direction of the new conjugate gradient method satisfies the sufficient descent and all assumptions of the global convergence property are considered and proved .in order to complete the benefit of our research we should take into account studied the numerical results which are written in FORTRAN language when the objective function is compared our new algorithm with HS and PRP methods on the similar set of unconstrained optimization test problems which is very efficient and encouragement numerical results.

Keywords: optimization, quadratic function, Taylor series, global convergence.

تطوير جديد لطريقة التدرج المترافق مع مبرهنات التقارب الشامل

أسيل مؤيد قاسم

قسم الرياضيات ، كلية التربية للعلوم الصرفة، جامعة الموصل، الموصل، العراق

الملخص

في هذا البحث حاولنا ان نقترح خوارزمية جديدة للتدرج المترافق لحل مسائل الامثلية غير المقيدة، وقد قمنا بالتركيز على طرائق التدرج المترافق المطبقة على مسائل الامثلية غيرالخطية وغير المقيدة , مع خطوة البحث الموجبة المستحصل عليها من خط البحث , كما تم اشتقاق المعلمة الجديدة للاتجاه الجديد لطرائق التدرج المترافق من الدالة التربيعية ومتسلسلة تايلر باستخدام شرط التركيب والشبيه بنيوتن واتجاه نيوتن في اشتقاق الصيغة الجديدة. كذلك قمنا ببرهنة ان اتجاه البحث لطريقة التدرج المترافق الجديدة تحقق الانحدار الكافي كما تم الاخذ بنظر الاعتبار برهان جميع فرضيات خاصية التقارب الشامل . ومن اجل استكمال الفائدة من

هذا البحث فقد اخذنا بنظر الاعتباردراسة النتائج العددية والتي كتبت بلغة فورتران وبمقارنة دالة الهدف للخوارزمية الجديدة مع طرائق هيستن وبولاك ريبير ولنفس المجموعة من مسائل الاختباروالتي كانت تمثل نتائج عددية كفوءة ومشجعة.

الكلمات المفتاحية: الامثلية،الدالة التربيعية، متسلسلة تايلر ، التقارب الشامل.

1. Introduction:

Conjugate Gradient methods (CG) contain a type of unconstrained optimization algorithms which are known by low memory requirements in strong and global convergence feathers.

This study focuses on conjugate gradient approaches for solving nonlinear unconstrained optimization problems [1].

Where $f : R^n \to R$ is a continuously differentiable function, that is bounded from below ,a nonlinear conjugate gradient method constructs a sequence using iteration x_k , $k \geq 1$ starting from an initial point $x_0 \in R$:

$$x_{k+1} = x_k + s_k \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(2)$$

$$s_k = \alpha_k d_k \;\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots(3)$$

Where d_k is a search direction, and the positive step size α_k is obtained by a line search such that $\alpha_k > 0$ .Often the sharpest descent path is used in the first iteration, namely $d_1 = -g_1$, and the other search direction can be defined recursively [2]:

$$d_{k+1} = -g_{k+1} + \beta_k d_k , \quad \ldots\ldots\ldots\ldots\ldots\ldots(4)$$

Here β_k is the CG updates parameter and $g_k = \nabla f(x_k)^T$ and different CG methods is correspond to different scalar β_k choices, If we let $\|.\|$ represent the Euclidean norm and define $y_k = g_{k+1} - g_k$, we may get a list of some options for the CG update parameter like thus :

$$\beta_k^{HS} = \frac{g_{k+1}^T y_k}{d_k^T y_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{Hestenes and Stiefel (1952) [3]}$$

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k{}^2\|} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{Fletcher and Reeves (1964)[4]}$$

$$\beta_k^{D} = \frac{g_{k+1}^T \nabla^2 f(x_k) d}{d_k^T \nabla^2 f(x_k) d_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{Daniel (1967)[5]}$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T y_k}{\|g_k\|^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{Polak Ribiere and Polyak (1969)[6]}$$

$$\beta_k^{CD} = -\frac{g_{k+1}^T g_{k+1}}{d_k^T g_k} \quad \ldots\ldots\ldots\ldots\ldots\ldots\text{Fletcher (1987)[7]}$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k}$$ ……………………..……..Liu and Storey(1991)[8]

$$\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k}$$ ………………..………………..Dai and Yuan(1999)[2]

$$\beta_k^{N} = (y_k - 2d_k \frac{\|y_k\|^2}{d_k^T y_k})^T \frac{g_{k+1}}{d_k^T y_k}$$ ……….……Hager and Zhang(2005)[9]

More detail see [10-12].

The global convergence qualities of CG algorithms are the topic of this research. we use β_k^{HS} , β_k^{PR} and compare it with our new β_k which will be derived later, the condition in (5) is used to avoid non-convergence in nonlinear functions that are utilized with inexact line search:

$$|g_{k+1}^T g_k| > 0.2\|g_{k+1}^T\|$$ …………………………… (5)

Which is called Powell restart condition because if the scalar β_k appears negative these strategies will restart the descent direction on the all iteration.

The search direction obtained by the new method at each iteration satisfies the sufficient descent condition as we prove. In order to grantee the global convergence of non-linear conjugate gradient methods for the CG line search we are often used Wolfe conditions, here we denote the Standard Wolfe line search which is mean that the step length α_k in equation (3) is obtained such that [13-15] :

$$f(x_k + \alpha_k d_k) - f(x_k) \le \rho \alpha_k g_k^T d_k ,$$ ……………………..……. (6)

$$|g_{k+1}^T d_k| < -\sigma\, g_k^T d_k$$ …... (7)

Where d_k is a descent direction and $0 < \rho \le \sigma < 1$. The strong Wolfe conditions consists of (6) and by rewrite the equation (7), see [16]:

$$g_{k+1}^T d_k \ge \sigma\, g_k^T d_k$$ ………………………………..(8)

but we are attention on shows whether there is a conjugate gradient method that converges under Wolfe's standard conditions or not. We develop a new formula β_k to show that this new conjugate gradient method is globally convergent if the classic Wolfe requirements (6) and (7) are met.

2: The Derivation of a new Scaled CG Method:

The derivation of most CG method are based, in some way, to the quadratic function and then generalized to non-quadratic functions by restart procedures. Hence, we may assume that our objective function $f(x)$ is a convex function, then the new method depends on the quadratic form:

$$f(x) = \frac{1}{2} x^T G x + b^T x + a$$ ……………(10)

Where G is the Hessian matrix and b = g and a is a constant, so that $y_k = g_{k+1} - g_k = Gs_k$ and by subsisting the value of $x_{k+1} = x_k + s_k$ in (10) we get [17] :

$$f(x_{k+1}) = \frac{1}{2}(x_k + s_k)^T G(x_k + s_k) + b^T(x_k + s_k) + a$$

$$f(x_{k+1}) = \frac{1}{2}(x_k + s_k)^T (Gx_k + Gs_k) + (b^T x_k + b^T s_k) + a$$

$$f(x_{k+1}) = \frac{1}{2}(x_k^T Gx_k + x_k^T Gs_k + s_k^T Gx_k + s_k^T Gs_k) + (b^T x_k + b^T s_k) + a$$

And if we assume that a= $1/2(s_k^T Gx_k + x_k^T Gs_k)$ then: $f_{k+1} = f_k + \frac{1}{2}s_k^T Gs_k + b^T s_k$ ……………………………………………….……...(11)

and when we compare (11) with Taylor series and note that (b = g) we get:

$$f_{k+1} = f_k + \frac{1}{2}s_k^T Gs_k + g_k^{\ T} s_k$$

$$2f_{k+1} = 2f_k + s_k^T Gs_k + 2g_k^{\ T} s_k$$

$$-2g_k^T s_k = 2(f_k - f_{k+1}) + s_k^T Gs_k$$

$s_k^T Gs_k = -2g_k^T s_k - 2(f_k - f_{k+1})$,using QN condition $Gs_k = y_k$

$s_k^T y_k = -2g_k^T s_k - 2(f_k - f_{k+1})$,now multiply with s_k ,

$\Rightarrow s_k^T s_k y_k = s_k\{-2g_k^T s_k - 2(f_k - f_{k+1})\}$, since $s_k^T s_k$ is a scalar we can divide both side with it $\Rightarrow$

$y_k = s_k \dfrac{\{-2g_k^T s_k - 2(f_k - f_{k+1})\}}{s_k^T s_k}$ now by multiply both side with s_k^T

$$Gs_k^T s_k = s_k^T s_k \frac{\{-2g_k^T s_k - 2(f_k - f_{k+1})\}}{s_k^T s_k}$$

$$G = \frac{-2g_k^T s_k - 2(f_k - f_{k+1})}{s_k^T s_k} * I_{nxn}$$

$$G^{-1} = \frac{s_k^T s_k}{-2g_k^T s_k - 2(f_k - f_{k+1})} * I_{nxn}$$

$$G^{-1} = -\frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})} * I_{nxn} \quad \text{……………………………(12)}$$

and since the direction is Newton direction then $d_{k+1} = -G^{-1} g_{k+1}$ then $d_{k+1} = \dfrac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})} g_{k+1}$

and when multiplying the direction with y_k^T we get:

$$y_k^T d_{k+1} = \frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})} y_k^T g_{k+1} \quad \text{………………………………(13)}$$

and since the new direction satisfying (4) so we have:

$y_k^T d_{k+1} = -y_k^T g_{k+1} + \beta_k y_k^T d_k$ ……………………....……..……(14)

From (13) & (14) we have:

$\frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})} y_k^T g_{k+1} = -y_k^T g_{k+1} + \beta_k y_k^T d_k$ then: $\beta_k y_k^T d_k = y_k^T g_{k+1}\left[1 + \frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})}\right]$ ……………………… (15)

$$\beta_k = \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1 + \frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})}\right] \quad \text{………………………….. (16)}$$

Then we have the new direction as $d_{k+1} = -g_{k+1} + \beta_k d_k$, and β_k as we explain in equation (16) which is derived from the quadratic function .

3. Outlines Of The New Algorithm:

Step(1): Initialize select $x_1 \in R^n$ and compute $f(x_1),\ g(x_1)$, consider $d_k = -g_k$ *and* $k = 1$.

Step (2): If $\|g_k\| \le \varepsilon$, stop. x_k is the optimal solution, else go to step (3).

Step (3): Compute α_k satisfying the Wolfe conditions (6), (7 (and update the variable $x_{k+1} = x_k + \alpha_k d_k$, Compute: $f_{k+1},\ g_{k+1}, y_k,\ s_k$.

Step (4): compute our new d_{k+1}. If Powell restart satisfied then set $d_{k+1} = -g_{k+1}$ else $d_{k+1} = d_k$ and set new α_k, $k = k+1$ go to step (2).

4. Descent Property of The New Algorithm:
Assumptions (4.1):

i. set $\xi = \{x \mid f(x) \le f(x_1)\}$ is bounded, to be specific, there exists a factor $B > 0$ such that $\|x\| \le B$ for all $x \in \xi$.

ii. In some neighborhood N of ξ we assume that $f(x)$ is continuously differentiable function $f(x)$, and the gradient is globally lipschitz continuous, this means there exist a factor $L > 0$ such that $\|\nabla f(x) - \nabla f(y)\| < L\|x - y\|$, $\forall x, y \in N$. [18]

Now we'll provide the following theorems, which guarantees the new algorithm's descent property:

Theorem (4.2):

Suppose $\{x_k\}$ is a sequences generated from the suppose algorithm if there exist a constant $\gamma > 0$ such that $\|g_{k+1}\| < \gamma$ for all k. then the search direction is descent direction for all k.

Proof: using induction method for k=1 we have $d_k = -g_k$ then $d_1^T g_1 < 0$, then we suppose that $g_k^T d_k < 0 \quad \forall \quad k \ge 2$ .Now, from $d_{k+1} = -g_{k+1} + \beta_k d_k$ and

$$\beta_k = \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1 + \frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})}\right]$$

And to check the descent for above direction for $k+1$, multiplying (4) by g_{k+1}^T then:

$$g_{k+1}^T d_{k+1} = -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1+\frac{s_k^T s_k}{2(g_k^T s_k + f_k - f_{k+1})}\right] g_{k+1}^T d_k$$

Using Wolfe condition (6) and since $s_k = \alpha_k d_k$ then we have:

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1+\frac{\alpha_k{}^2\|d_k\|^2}{2(\alpha_k\, g_k^T d_k - \rho\, \alpha_k\, g_k^T d_k)}\right] g_{k+1}^T d_k$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1+\frac{\alpha_k{}^2\|d_k\|^2}{2\alpha_k g_k^T d_k (1-\rho)}\right] g_{k+1}^T d_k$$

and since $g_k^T = -d_k$ then $g_k^T d_k = -\|g_k\|^2$ [19], then we have :

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1+\frac{\alpha_k{}^2\|d_k\|^2}{-2\alpha_k\|d_k\|^2(1-\rho)}\right] g_{k+1}^T d_k$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + \frac{y_k^T g_{k+1}}{y_k^T d_k}\left[1+\frac{\alpha}{2(\rho-1)}\right] g_{k+1}^T d_k$$

Assume $\rho \in (0,1/2)$, then $\frac{\alpha}{2(\rho-1)} < 0$ (very small and negative) and we assume that $\tau = \frac{\alpha}{2(\rho-1)}$ so we have:

$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + (\tau+1)\frac{y_k^T g_{k+1}}{y_k^T d_k} g_{k+1}^T d_k$ and since $g_{k+1}^T d_k = g_{k+1}^T d_k - g_k^T d_k + g_k^T d_k$ therefore

$$g_{k+1}^T d_k = y_k^T d_k + g_k^T d_k \text{ and since } g_k^T d_k < 0 \text{ then: } g_{k+1}^T d_k < y_k^T d_k$$

by substations this relation on the following equation we get:

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + (\tau+1)\frac{\|g_{k+1}\|^2}{g_{k+1}^T d_k}(g_{k+1}^T d_k)$$

$$g_{k+1}^T d_{k+1} \le -\|g_{k+1}\|^2 + (\tau+1)\|g_{k+1}\|^2$$

$$g_{k+1}^T d_{k+1} \le (\tau+1-1)\|g_{k+1}\|^2 \quad \Rightarrow \quad g_{k+1}^T d_{k+1} \le \tau\|g_{k+1}\|^2$$

assume τ = c and τ is negative as we proof , then $g_{k+1}^T d_{k+1} \le -c\|g_{k+1}\|^2$

which is descent direction for all k. Hence, the proof is completed by induction.

(4.3) Global Convergence Theorem:

We have the following lemma (4.4) and (4.5) for any conjugate Gradient method with the strong Wolfe line search, which were first found by zoutendijk [20] and Wolfe.

Journal of Education and Science (ISSN 1812-125X), Vol: 29, No: 1, 2020 (325-335)

Lemma (4.4): Assume that Assumptions i and ii, and the descent condition are true, and that α_k is obtained using the strong Wolfe line search [6]. Then there's

$$\sum_{k=1}^{\infty} -\alpha_k g_k^T d_k < \infty \qquad \text{………………….} \; (17)$$

Lemma (4.5): Consider any conjugate gradient method in the form (2) is gained by the strong Wolfe line search. Assume assumptions i ii, and the descent condition are true. Then the so-called Zoutendijk condition is true:

$$\sum_{k=1}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < +\infty \qquad \text{…………………….(18)}$$

Proof:

Since $d_{k+1} = -g_{k+1} + \beta_k d_k$

or $d_{k+1} + g_{k+1} = \beta_k \; d_k$

(Dai and Yuan, 1999).Taking the square of each said and noting:

$$\beta_k^2 = \left(\frac{y_k^T g_{k+1}}{y_k^T d_k}\right)^2 \left(1 + \frac{s_k^T s_k}{g_k^T s_k + f_k - f_{k+1}}\right)^2$$

$$\le \left(\frac{y_k^T g_{k+1}}{y_k^T d_k}\right)^2 \left(1 + \frac{\alpha^2 \|d_k\|}{(\alpha_k g_k^T d_k - \rho\alpha_k g_k^T d_k)}\right)^2$$

$$\le \left(\frac{y_k^T g_{k+1}}{y_k^T d_k}\right)^2 \left(1 + \frac{\alpha}{1-\rho}\right)^2 < \left(\frac{y_k^T g_{k+1}}{(\sigma-1) g_k^T d_k}\right)^2 \left(1 + \frac{\alpha}{1-\rho}\right)^2$$ let $y_k^T g_{k+1} < \|g_{k+1}\|^2 < \omega$ and let:

$$\left(\frac{1}{\sigma-1}\right)^2 \left(1 + \frac{\alpha}{1-\rho}\right)^2 = \gamma^2$$ then $\beta_k \le \left(\frac{\omega}{g_k^T d_k}\right)^2 \gamma^2 < \left(\frac{1}{g_k^T d_k}\right)^2$

$$\|d_{k+1}\|^2 = \frac{\|d_k\|^2}{(g_k^T d_k)^2} - 2 d_{k+1}^T g_{k+1} - \|g_{k+1}\|^2$$

Divide each term in above equation by $(d_{k+1}^T g_{k+1})^2$ and also using Wolfe condition (6) &(7) we get:

$$\frac{\|d_{k+1}\|^2}{(g_{k+1}^T d_{k+1})^2} \le \frac{\|d_k\|^2}{(g_k^T d_k)^2 (g_{k+1}^T d_{k+1})^2} - \frac{2}{d_{k+1}^T g_{k+1}} - \frac{\|g_{k+1}\|^2}{(d_{k+1}^T g_{k+1})^2}$$

$$= \frac{\|d_k\|^2}{(g_k^T d_k)^2 (g_{k+1}^T d_{k+1})^2} - \left(\frac{1}{\|g_{k+1}\|} + \frac{\|g_{k+1}\|}{d_{k+1}^T g_{k+1}}\right)^2 + \frac{1}{\|g_{k+1}\|^2}$$

$$\le \frac{\|d_k\|^2}{(g_k^T d_k)^2} + \frac{1}{\|g_{k+1}\|^2} \qquad \text{…………..…………………………..……(19)}$$

Because $\frac{\|d_1\|^2}{(g_1^T d_1)^2} = \frac{1}{\|g_1\|^2}$, then (19) shows that

$$\frac{\|d_k\|_2^2}{(g_k^T d_k)^2} \le \sum_{i=1}^{k} \frac{1}{\|g_i\|^2} \qquad \forall \quad k$$

……………..………….(20)

now if the theorem is not true , there exist a constant $c > 0$

s.t. $\|g_k\| \ge c \qquad \forall \quad k$ …………………………..…….(21)

Therefore it follow from (20),(21) that $\frac{\|d_k\|_2^2}{(g_k^T d_k)^2} \le \frac{k}{c^2}$ which implies that:

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} = \infty$$

……….…………………….…..…..(22)

Which is conflict with Zoutendijk theorem hence $\|g_k\| = 0$ therefore the algorithm is globally convergent.

6. Numerical experiments:

We now provide numerical tests comparing our novel approach to the HS and PRP algorithms on the same set of unconstrained optimization test functions, with the goal of determining which method is the most reliable and efficient for addressing any unconstrained optimization problem.

We investigated numerical trials with the same number of dimensions for each test function [21] for n=100, 300, 1000, 5000, 6000, and 10000.

All of the algorithms use the same line search and the same parameters. The number of iterations(No.I) and the number of function evaluations (No.F) are used to do the comparison, Our algorithms has stopped at once $\|g_k\| \le 10^{-5}$.

Table (6.1) Comparison algorithms for n=100,n=1000

problems	New n=100	PR N=100	HS n=100	New n=1000	PR N=1000	HS n=1000
	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)
Starit	8(18)	6(14)	6(14)	8(18)	6(14)	6(14)
Wolfe	44(89)	44(89)	205(411)	53(107)	64(129)	210(421)
Rosenbrok	26(68)	26(65)	26(68)	27(70)	27(67)	27(70)
Powell	41(112)	46(116)	55(130)	43(116)	57(150)	64(169)
Wood	78(162)	127(261)	98(202)	84(174)	146(299)	98(202)
Sum	12(63)	12(63)	12(63)	18(80)	22(104)	18(80)
cantral	37(261)	31(205)	43(278)	40(302)	41(332)	48(345)
Miele	67(208)	99(307)	52(159)	78(253)	101(322)	52(159)
Fred	9(24)	9(25)	10(27)	9(24)	10(27)	10(27)
Nondiagonal	27(66)	29(77)	27(66)	27(65)	30(78)	27(65)
Shallo	11(27)	11(28)	11(27)	11(27)	11(28)	11(27)
Cubic	16(44)	16(44)	16(44)	16(44)	16(44)	16(44)
Beal	13(31)	10(23)	11(27)	13(31)	11(27)	11(27)
Osp	48(165)	49(164)	48(165)	199(554)	149(568)	203(586)
Total	437(1338)	515(1481)	620(1681)	626(1865)	691(2189)	801(2263)

Table (6.2) Comparison algorithms for n=5000,n=10000

problems	New n=5000	PR N=5000	HS N=5000	New n=10000	PR n=10000	HS n=10000
	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)	No.I(No.F)
Wolfe	314(630)	255(512)	150(304)	261(525)	215(432)	305(613)
Rosenbrok	27(70)	27(67)	27(70)	27(70)	27(67)	27(70)
Powell	43(116)	57(150)	65(171)	43(116)	63(179)	65(171)
Central	111(321)	47(418)	51(390)	40(302)	50(466)	51(390)
Miele	81(273)	109(360)	54(174)	78(253)	109(360)	54(147)
Nondiagonal	27(65)	30(78)	27(65)	27(65)	30(78)	27(65)
Shallo	11(27)	11(28)	11(27)	11(27)	11(28)	11(27)
Beal	11(27)	11(27)	11(27)	11(27)	11(27)	11(27)
Sum	30(136)	19(86)	30(136)	47(181)	24(97)	47(181)
Osp	435(1353)	611(1859)	435(1353)	612(1979)	834(2546)	612(1979)
Cubic	16(44)	16(44)	16(44)	16(44)	16(44)	16(44)
Total	1106(3062)	1193(3629)	877(2761)	1173(3589)	1390(2590)	1226(3714)

Table (6.3) Comparison algorithms for n=300,n=6000

problems	New n=300	PR N=300	HS n=300	New n=6000	PR N=6000	HS n=6000
	No.I(No.F)	No.I(No.F)	No.I(No,F)	No.I(No.F)	No.I(No.F)	No.I(No.F)
Gpowell-3	18(39)	23(49)	30(62)	19(42)	23(49)	31(64)
wolfe	46(93)	53(107)	43(87)	204(414)	294(590)	275(554)
Cubic	16(44)	16(44)	16(44)	16(44)	16(44)	16(44)
Shallo	11(27)	11(28)	11(27)	11(27)	11(28)	11(27)
Beal	11(27)	11(27)	11(27)	11(27)	11(27)	11(27)
Edger	6(16)	7(18)	7(18)	6(16)	7(18)	7(18)
Total	108(246)	121(272)	118(265)	267(570)	362(756)	351(734)

7. conclusion:

we searched in this research for a new modification conjugate gradient method procedure which depends on derived the quadratic function . We have decided under our experiment that the global convergence for the suggested idea is state also the numerical experiment explained in Tables (6.1),(6.2) and (6.3) are the efficient of the proposed algorithm with respect to regular HS and PRP methods on average and according to the numbers of results.

8. References:

[1] R. Dehghani,M.M. Hosseini,N. Bidabadi," The modified quasi-Newton methods for solving unconstrained optimization problems", international journal of Numarical Modelling,2018,https://doi.org/10.1002/jnm.2459.

[2] Dai, Y. and Yuan, Y., "A Nonlinear conjugate gradient method with a strong global convergence property", SIAM journal of Optimization, no.10, pp.177-182,1999, https://doi.org/10.1137/S1052623497318992.

[3] M. R. Hestenes and E. Stiefel, "Methods of conjugate gradients for solving linear systems", Journal of Research of the National Bureau of Standards, vol. 49, no. 6, p. 409, Dec. 1952, Doi: 10.6028/jres.049.044.

[4] R. Fletcher, "Function minimization by conjugate gradients", The Computer Journal, vol. 7, no. 2, pp. 149–154,1964, doi: https://doi.org/10.1093/comjnl/7.2.149,

[5] Daniel J.W ," The conjugate gradient method for linear and nonlinear operator equations", SIAM J. Number. Anal.4, pp.10-26, 1967.

[6] E. Polak and G. Ribiere, "Note sur la convergence de méthodes de directions conjuguées", Revue française d'informatique et de recherche opérationnelle. Série rouge, vol. 3, no. 16, pp. 35–43, 1969, doi: 10.1051/m2an/196903r100351.

[7] Fletcher, R."Practical Methods Of Optimization" John Witley and sons ,Chichester , New York 1987.

[8] Liu, D. and Story, C. ,"Efficient Generalized Conjugate Gradient Algorithms". Part 1: Theory. Journal of Optimization Theory and Applications, 69, 129-137,1991, doi.org/10.1007/BF00940464.

[9] W. Hager, Hongchao Zhang," A new conjugate gradient method with guaranteed descent and an efficient line search", SIAM Journal on Optimization, 16, 170-192,2005, doi:10.1137/030601880.

[10] Snezana S.Djordjevic,"unconstrained Optimization Methods:Conjugate Gradient Methods and Trust –Region Methods",Applied Mathematics, 2019,doi.org/10.5772/intechopen.84374.

[11] S. S. Rao, "*Engineering optimization: theory and practice*." ,John Wiley & Sons, 2019.

[12] Serge Gratton , Vincent malmedy , Philippe L.Toint ," Using Approximate secant equation in limited memory methods for multilevel Unconstrained Optimization" ,comput optim Appl, vol 51,pp.967-979 ,2012.

[13] Joannis E. Livieris and Panagiotis Pitelas," A Descent Dai-Liao Conjugate Gradient Method Based on a Modified Secant Equation and its Global Convergence" , ISRN Computational Mathematics , Patras, Greece ,2012, https://doi.org/10.5402/2012/435495.

[14] L.Arman,Y.Xu,M.Rostami and F. Rahpeymaii, "Some three-term conjugate gradient methods for solving unconstrained optimization problems" ,Pacific jornal of optimization ,2019.

[15] j.Nocedal and S.J.Wright, "Numarical Optimization",Springer, New York, 2006.

[16] Jiankun Liu and Shouqiang Du," Modified Three-Term Conjugate Gradient Method and Its Applications", mathematical problems in engineering,2019, https://doi.org/10.1155/2019/5976595.

[17] Zhen-Jun shi and Jie Shen ,"Step - Size Estimation For Unconstrained Optimization Methods", volume 24,N.3,pp.399-416 , SBMAC,ISSN,0101-8205 , 2005.

[18] Wu C.Y and Chen G.Q ," New type of conjugate gradient algorithms for unconstrained optimization problems" , jurnal of system engineering and electronics , 21 (6) ,1000-1007 , 2010.

[19] Jjankum Liu and Shouqiang Du,"Modified Three_Term Conjugate Gradient Method and Its Applications", Mathematical Problems in Engineering, 2019.

[20] Guoyin Li, Chunming Tang , ZengxinWei ," New conjugacy condition and related new conjugate gradient methods for unconstrained optimization", Journal of Computational and Applied Mathematics 202 , pp .523 – 539 ,2007.

[21] G. Zoutendijk, “Nonlinear Programming, Computational Methods,” In: J. Abadie Ed., Integer and Nonlinear Programming, North-Holland, Amsterdam, pp. 37-86,1970.

[22] N. Andrei, “An unconstrained optimization test functions collection,” Adv. Model. Optim, vol. 10, no. 1, pp. 147-161,2008.