https://doi.org/10.26682/sjuod.2022.25.2.51
Journal of University of Duhok., Vol. 25, No.2(Pure and Engineering Sciences), Pp 573-578, 2022

A NEW CONJUGATE GRADIENT WITH GLOBAL CONVERGES FOR NONLINEAR PROBLEMS

SALAH GAZI SHAREEF
Dept. of Mathematics, Faculty of Science, University of Zakho, Kurdistan Region-Iraq

(Received: May 16, 2022; Accepted for Publication: December 4, 2022)

ABSTRACT

The conjugate gradient(CG) method is one of the most popular and well-known iterative strategies for solving minimization problems, it has extensive applications in many domains such as machine learning, neural networks, and many other fields, partly because to its simplicity in algebraic formulation and implementation in codes of computer and partially due to their efficiency in solving large scale unconstrained optimization problems. Fletcher/Reeves (C, 1964) expanded the concept to nonlinear problems. In 1964, and this is widely regarded as the first algorithm of nonlinear conjugate gradient. Since then, other conjugate gradient method versions have been proposed. In this paper and in section one, we derive a new conjugate gradient for solving nonlinear minimization problems based on parameter of Perry. In section two we will satisfy some conditions like descent and sufficient descent conditions. In section three , we will study the global convergence of new suggestion. We present numerical findings in the fourth part to demonstrate the efficacy of the suggestion technique. Finally, we provide a conclusion.

KEYWORDS: Nonlinear Optimization, Algorithm of Conjugate Gradient, property of the Descent, property of the Sufficient Descent and Global Converges Properties.

1. INTRODUCTION

Consider the following nonlinear minimization problem bellow:

Min $f(x);\ \ x \in R^n$ (1.1)

Where; $f: R^{n*1} \to R^{1*1}$ is a continuously differentiable real-valued function

The methods of the form

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

Used for solving nonlinear conjugate gradient methods (1.1)

This iterative method is starting with an initial $x_1 \in R^n$,

where , v_k is the difference between points $x_{k+1} - x_k = \alpha_k d_k$; the stepsize α_k is calculated by one dimensional line search and d_k is search direction. There are two search directions , the first search direction is the direction of the steepest descent method, which is

$d_1 = -g_1$ (1.3)

The search direction for the next iteration is set to:

$d_{k+1} = -\, g_{k+1} + \beta_k\, d_k$ (1.4)

Where; $g(x_k) = \nabla f(x_k)$ and β_k is scalar. There are a lots of basic formulas of β_k are suggested , like Hestenes,Stiefel (HS)(Steifel, 1952), PolakRibiere,Polyak (PRP)(E. Polak, 1969), Fletcher ,Reeves(FR)(Reeves, 1964), Dai , Yuan (DY) (Dai, Y.H., and Yuan, 1999), Dai, Liao (Y.H. Dai, 2001) ,Perry(Perry, 1978), Liu , Storey (Liu Y., and Storey, 1991), and (CD)(R., 1987)which are shown below:

$$\beta_k^{HS} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{d_k^T(g_{k+1} - g_k)} \quad (1.5)$$

$$\beta_k^{PRP} = \frac{g_{k+1}^T(g_{k+1} - g_k)}{\|g_k\|^2} \quad (1.6)$$

$$\beta_k^{FR} = \frac{\|g_{k+1}\|^2}{\|g_k\|^2} \quad (1.7)$$

$$\beta_k^{DY} = \frac{\|g_{k+1}\|^2}{d_k^T y_k} \quad (1.8)$$

$$\beta_k^{DL} = \frac{g_{k+1}^T(y_k - t v_k)}{d_k^T y_k}\ , \quad t > 0 \quad (1.9)$$

$$\beta_k^{Perry} = \frac{g_{k+1}^T(y_k - v_k)}{d_k^T y_k} \quad (1.10)$$

$$\beta_k^{LS} = \frac{g_{k+1}^T y_k}{-d_k^T g_k} \quad (1.11)$$

$$\beta_k^{CD} = \frac{\|g_{k+1}\|^2}{-d_k^T g_k} \quad (1.12)$$

where $y_k = g_{k+1} - g_k$ and The symbol $\|.\|$ is used to represent the Euclidean norm of vectors. symbol. The property of the global of Fletcher R.(FR) method, Polak R. Polyak(PRP) method, Hestenes S.(HS)method, Dai Y.(DY) method, Conjugate D.(CD) method and Liu,S.(LS) method can see(E.G. Birgin, 2001)

(J. Sun, 2001)(L. Zhang, W. Zhou, 2006)(Raydan, 1997)(R., 1987)(Liu Y., and Storey, 1991).
Also, many parameters are suggested, for example, Hager , Zhang(W, Hager, 2005) suggested another conjugate gradient algorithm , called CG- DESCENT method. Zhang Li et al. (Zhang L., Zhou W.J., 2006a)(Zhang L., Zhou W.J., 2006b) also suggested some modification conjugate gradient methods. You can see (Hussein A. Kh., 2019)(Hussein A. Kh., 2020)(Alaa L. I., Muhammad A. S., 2019).

2. DERIVATIVE OF NEW FORMULA β_k^{NEW}

The idea for finding a new conjugate gradient algorithm to find the minimum of the unconstrained problems is to use a new vector as follows,
Consider the vector bellow

$$y_k^* = g_{k+1} + (1-\delta)\left(\left(\frac{g_{k+1}}{\gamma}\right) - \mu g_{k+1}\right) \quad (2.1)$$

Where,

$$\delta \in (0,1), \qquad \mu = 0.1$$

$$\gamma = \frac{2\sqrt{\omega}}{\|v_k\|}\,(1 + \|x_{k+1}\|)$$

and ω is the machine error
We replace the vector y_k in the numerator of (1.10) by y_k^*, we get

$$\beta_k^{NEW} = \frac{g_{k+1}^T(y_k^* - v_k)}{d_k^T y_k} \quad (2.2)$$

Or

$$\beta_k^{NEW} = \frac{g_{k+1}^T\left(g_{k+1} + (1-\delta)\left(\left(\frac{g_{k+1}}{\gamma}\right) - \mu g_{k+1}\right) - v_k\right)}{d_k^T y_k}$$

Here, we get

$$\beta_k^{NEW} = \frac{\|g_{k+1}\|^2 + (1-\delta)\frac{\|g_{k+1}\|^2}{\gamma} - (1-\delta)\mu\|g_{k+1}\|^2 - g_{k+1}^T v_k}{d_k^T y_k} \quad (2.3)$$

2.1 Outlines of the New Method

Step1: Choose x_1 and $\varepsilon = \frac{1}{10^5}$.
Step2: Set $d_1 = -g_1$; $If\ \|g_1\| \le \varepsilon$, then stop, $g_k = \nabla f(x_k)$; Set Choose = 1 .
Step3: Find the steplength $\alpha_k > 0$, satisfying the below conditions (Wolfe condition)

$$f(x_k + \alpha_k d_k) - f(x_k) \le q_1 \alpha_k g_k^T d_k$$

$$|g_{K+1}^T d_k| \le q_2 |g_k^T d_k|$$

where, $0 < q_1 < q_2 < 1$.

Step4: Calculate $x_{k+1} = x_k + \alpha_k d_k$
$g_{k+1} = \nabla f(x_{k+1})$; $If\ \|g_{k+1}\| \le \varepsilon$, then stop.
Step5: Calculate β_k^{NEW} by (2.3)
Step6: Evaluate $d_{k+1} = -g_{k+1} + \beta_k^{NEW} d_k$
Step7: If $|g_{k+1}^T g_k| > 0.2 * \|g_{k+1}\|^2$, then, go to step2 .
Else
$k = k + 1$ and goto step3.

2.2 Descent Property and Sufficient Descent Property of the New Algorithm

Theorem 1:- If (1.2) gives the sequence $\{x_k\}$, then the descent condition is satisfied by the equation (1.4) using new gradient algorithm (2.3); $d_{k+1}^T * g_{k+1} \le 0$ with exact line search and inexact line search.
Proof :- From equations (1.4) and (2.3) we get,

$$d_{k+1} = -g_{k+1} + \left(\frac{\|g_{k+1}\|^2 + (1-\delta)\frac{\|g_{k+1}\|^2}{\gamma} - (1-\delta)\mu\|g_{k+1}\|^2 - g_{k+1}^T v_k}{d_k^T y_k}\right) d_k \quad (2.4)$$

Implies that

$$d_{k+1} = -g_{k+1} + \left(\frac{\|g_{k+1}\|^2}{d_k^T y_k} + \frac{(1-\delta)}{\gamma}\frac{\|g_{k+1}\|^2}{d_k^T y_k} - (1-\delta)\mu\frac{\|g_{k+1}\|^2}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\right) d_k \quad (2.5)$$

Multiplying both sides of the above equation by g_{k+1} from right , we have

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} + \frac{(1-\delta)}{\gamma} * \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - (1-\delta)\mu \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - \alpha_k \frac{(d_k^T g_{k+1})^2}{d_k^T y_k}$$

This implies that

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \left(1 + \frac{(1-\delta)}{\gamma}\right)\frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - (1-\delta)\mu \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - \alpha_k \frac{(d_k^T g_{k+1})^2}{d_k^T y_k} \quad (2.6)$$

Since $\left(1 + \frac{(1-\delta)}{\gamma}\right) > (1-\delta)\mu$, we can write the above equation as follows

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \partial \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - \alpha_k \frac{(d_k^T g_{k+1})^2}{d_k^T y_k} \quad (2.7)$$

Where, $\partial = (1 + \frac{(1-\delta)}{\gamma}) - (1-\delta)\mu$, which is positive.

Here , we take two cases , the first case if the stepsize is determined by an exact line search; which is $d_k^T g_{k+1} = 0$, then, we get $d_{k+1}^T g_{k+1} \le 0$.

The second case if we have inexact line search which is $d_k^T g_{k+1} \ne 0$.

By mathematical induction, from the first search direction, we get $d_1^T g_1 = -\|g_1\|^2 \le 0$, and we suppose that it is true for case k that is mean $d_k^T g_k \le 0$. To prove case $k+1$

We know that DY parameter is satisfy the condition of the descent; then , the first term and second term of (2.7) are less than or equal to zero , and the third term clearly is les that zero, so we have

$$d_{k+1}^T g_{k+1} = -\|g_{k+1}\|^2 + \partial \frac{\|g_{k+1}\|^2}{d_k^T y_k} d_k^T g_{k+1} - \alpha_k \frac{\left(d_k^T g_{k+1}\right)^2}{d_k^T y_k} \le 0 \quad \blacksquare \qquad (2.9)$$

Theorem2:- Suppose that $\{x_k\}$ is produced by (1.2), then, the equation (1.4) with new equation (2.3) satisfies the sufficient descent condition,

$$d_{k+1}^T g_{k+1} \le -C\|g_{k+1}\|^2.$$

Proof:- From equation (2.7)

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

Since the parameter of DY is satisfy the descent property, then we get the following inequality

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

Or

$$d_{k+1}^T g_{k+1} \le -\alpha_k \frac{(d_k^T g_{k+1})^2}{d_k^T y_k} * \frac{\|g_{k+1}\|^2}{\|g_{k+1}\|^2} \qquad (2.10)$$

let $C = \alpha_k \frac{(d_k^T g_{k+1})^2}{d_k^T y_k} * \frac{1}{\|g_{k+1}\|^2}$

Then we have

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

2.3 Convergence Analysis

we assume that:

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

2) The function f is continuously differentiable and its gradient is Lipchitz continuous in a neighborhood of B, i.e. there is a constant $M > 0$ such that

$$\|g(x) - g(x_k)\| \le M\|x - x_k\|, \forall x, x_k \in \Omega \qquad (2.11)$$

There is a constant under these assumptions on f, $\rho \ge 0$; such that $\|g(x)\| \le \rho, \quad \forall x \in B$.

Lemma1: We suppose the assumptions (1) and (2) are holds and consider the equations (1.2) , (1.3) and (1.4), where d_k is a descent direction and α_k is calculated by the strong Wolfe condition.

$$f(x_k + \alpha_k d_k) - f(x_k) \le q_1 * \alpha_k g_k^T d_k \qquad (2.12)$$

$$\left|g_{K+1}^T d_k\right| \le q_2 g_k^T d_k \qquad (2.13)$$

If

$$\sum_{k\ge 1} \frac{1}{\|d_k\|^2} = \infty \qquad (2.14)$$

Then

$$\lim_{k\to\infty} inf\|g_k\| = 0 \qquad (2.15)$$

See (Dai, Y.H., and Yuan, 1999).

If f is a function that is uniformly convex,, there is a constant $\vartheta > 0$ such that:

$$\left(g(x) - g(y)\right)^T (x - y) \ge \vartheta\|x - y\|^2 \in \Omega \qquad (2.16)$$

We can rewrite (2.16) in the following manner:

$$y_k^T v_k \ge \vartheta\|v_k\|^2 \qquad (2.17)$$

Theorem 3: Suppose the above assumptions are holds and that f is a function that is uniformly convex. Then the equations (1.2), (1.4) with new method (2.3) where d_k satisfies the descent condition and α_k is obtained by the strong Wolfe conditions (2.12) and (2.13) satisfies property of the global convergence.

That is mean $\lim_{k\to\infty} inf\|g_{k+1}\| = 0$

Proof: From (1.4) and (2.3), we get

$$d_{k+1} = -g_{k+1} + \beta_k^{New} d_k \qquad (2.18)$$

$$\left|\beta_k^{New}\right| = \left|\frac{\|g_{k+1}\|^2 + (1-\delta)\frac{\|g_{k+1}\|^2}{\gamma} - (1-\delta)\mu\|g_{k+1}\|^2 - g_{k+1}^T v_k}{d_k^T y_k}\right| \qquad (2.19)$$

Or

$$\left|\beta_k^{New}\right| = \left|\frac{\|g_{k+1}\|^2}{d_k^T y_k} + \frac{(1-\delta)\|g_{k+1}\|^2}{\gamma\, d_k^T y_k} - \frac{(1-\delta)\mu\|g_{k+1}\|^2}{d_k^T y_k} - \frac{g_{k+1}^T v_k}{d_k^T y_k}\right| \qquad (2.20)$$

Since $g_{k+1}^T d_k \le d_k^T y_k$, then,

$$\left|\beta_k^{New}\right| \le \left|\frac{\|g_{k+1}\|^2}{d_k^T y_k}\right| + \left|\frac{(1-\delta)\|g_{k+1}\|^2}{\gamma\, d_k^T y_k}\right| + \left|\frac{(1-\delta)\mu\|g_{k+1}\|^2}{d_k^T y_k}\right| + |\alpha_k| \qquad (2.21)$$

Since $y_k^T d_k \ge \frac{\vartheta\|v_k\|^2}{\alpha_k}$, then , we can write equation (2.21) as follows

$$\left|\beta_k^{New}\right| \le \frac{\alpha_k\|g_{k+1}\|^2}{\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\|g_{k+1}\|^2}{\gamma\,\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\mu\|g_{k+1}\|^2}{\vartheta\|v_k\|^2} + \alpha_k \qquad (2.22)$$

Here, we get

$$\left|\beta_k^{New}\right| \le \frac{\alpha_k\rho^2}{\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\rho^2}{\gamma\,\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\mu\rho^2}{\vartheta\|v_k\|^2} + \alpha_k \qquad (2.23)$$

Now, since,

$$\|d_{k+1}\| \le \|g_{k+1}\| + \left|\beta_k^{New}\right|\|d_k\|$$

Then,

$$\|d_{k+1}\| \le \rho + \left(\frac{\alpha_k\rho^2}{\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\rho^2}{\gamma\,\vartheta\|v_k\|^2} + \frac{\alpha_k(1-\delta)\mu\rho^2}{\vartheta\|v_k\|^2} + \alpha_k\right)\|d_k\| \qquad (2.24)$$

Implies that

$$\|d_{k+1}\| \le \rho + \left(\frac{\rho^2}{\vartheta\|v_k\|} + \frac{(1-\delta)\rho^2}{\gamma\,\vartheta\|v_k\|} + \frac{(1-\delta)\mu\rho^2}{\vartheta\|v_k\|} + \|v_k\|\right) \qquad (2.25)$$

Since, $\|v_k\| = \|x - x_k\|$, $D = \max\{\|x - x_k\|\}, \forall x, x_k \in R\}$

Here,

$$\|d_{k+1}\| \le \rho + \left(\frac{\rho^2}{\vartheta D} + \frac{(1-\delta)\rho^2}{\gamma\,\vartheta D} + \frac{(1-\delta)\mu\rho^2}{\vartheta D} + D\right) \qquad (2.26)$$

So, inequality (2.26) becomes

$$\|d_{k+1}\| \le \rho + \left(\frac{\rho^2}{\vartheta D} + \frac{(1-\delta)\rho^2}{\gamma\,\vartheta D} + \frac{(1-\delta)\mu\rho^2}{\vartheta D} + D\right) = \varphi \qquad (2.27)$$

Then,

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

And

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

3. NUMERICAL RESULTS

In this part, we present detailed numerical findings of a variety of problems applying a new method. We compare the new algorithm with standard Conjugate Gradient algorithm(Perry). The comparative tests contain nonlinear unconstrained problems (a popular testing function) with different dimensions$n =$ 4,100,500,1000,3000 and 5000. FORTRAN 90 is the programming language used, the stopping condition is $\|g_{k+1}\| \le 10^{-5}$. Tables (i) and (ii) show the number of functions (NOF) and iterations (NOI). Results in tables (i) and (ii) showed that the our method is superior to standard Conjugate Gradient methods (Perry), with respect to the NOF and NOI .

Table (i) : Comparing the numerical results of the algorithms (Perry and New Algorithm)

Test function	Dim.	Algorithm of Perry		New algorithm	
		NOI	NOF	NOI	NOF
Powell	4	35	89	30	79
	100	43	105	34	90
	500	43	105	34	90
	1000	45	120	34	91
	3000	46	122	34	92
	5000	46	122	34	92
Rosen	4	30	83	27	90
	100	30	83	20	59
	500	30	83	22	64
	1000	30	83	16	51
	3000	30	83	20	60
	5000	30	83	17	53
Miele	4	34	113	21	76
	100	46	169	14	46
	500	52	198	30	117
	1000	58	229	44	190
	3000	58	229	37	152
	5000	64	261	35	146
Wolfe	4	11	24	16	33
	100	49	99	43	87
	500	52	105	47	96
	1000	70	141	49	100

	3000	170	351	146	309
	5000	166	350	140	296
Wood	4	30	68	27	63
	100	30	68	28	64
	500	30	68	27	62
	1000	30	68	28	65
	3000	30	68	28	64
	5000	30	68	28	64
Cubic	4	12	35	16	50
	100	13	37	11	33
	500	13	37	10	30
	1000	13	37	10	31
	3000	13	37	10	30
	5000	13	37	10	31
Non-Digonal	4	24	64	20	62
	100	29	79	19	56
	500	29	79	27	85
	1000	F	F	24	74
	3000	29	79	21	66
	5000	30	81	24	77
G-Edger	4	5	14	5	14
	100	5	14	5	14
	500	6	16	5	14
	1000	6	16	5	14
	3000	6	16	5	14
	5000	6	16	5	14
Total		1750	4682	1342	3650

Table(ii):- Percentage comparison of the algorithms(Algorithm of Perry and New Algorithm)

	Algorithm of Perry	New Algorithm
NOI	100 %	76.686 %
NOF	100 %	77.958 %

4. CONCLUSION

For nonlinear unconstrained minimization problems, a new conjugate gradient algorithm proposed. We have proved the descent condition of the proposed method , also the sufficient descent condition, moreover global convergence property. Numerical tests were done on problems with low and high dimensionality, and comparisons were done between different test functions. The new method has proven its efficiency through results in tables (i) and (ii).

5.REFERENCES

Alaa L. I., Muhammad A. S., S. G. S. (2019). A New Conjugate Gradient Coefficient for Unconstrained Optimization Based On Dai-Liao. *Science Journal of University of Zakho*, *7*(1), 34–36.

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

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

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

E.G. Birgin, J. M. M. (2001). A spectral conjugate gradient method for unconstrained optimization. *Appl. Math. Optim.*, *43*, 117–128.

Hussein A. Kh., S. G. S. (2019). A New Parameter Conjugate Gradient Method Based on Three Terms Unconstrained Optimization. *General Letters in Mathematics*, *7*(1), 39–44.

Hussein A. Kh., S. G. S. (2020). A New conjugate gradient method for unconstrained optimization problems with descent property. *Gen. Lett. Math.*, *9*(2).

J. Sun, J. Z. (2001). Convergence of conjugate gradient methods without line search. *Annals Operations Research*, *103*, 161–173.

J.C. Gilbert, J. N. (1992). Global Convergence properties of conjugate gradient methods for optimization. *SIAM J. Optim., 2*, 21–24.

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

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

Perry, A. (1978). A Modified Conjugate Gradient Algorithm. *JSTOR*, *26*(6), 1073–1078.

R., F. (1987). *Practical Methods of Optimization vol.1: Unconstrained Optimization*. New York: Jhon Wiley & Sons.

Raydan, M. (1997). The Barzilain and Borwein gradient method for the large unconstrained minimization problem. *SIAM J. Optim.*, *7*, 26–33.

Reeves, R. F. and C. M. (1964). Function minimization by conjugate gradients. *Computer Journal*, *7*, 149–154.

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

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

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

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

Zhang L., Zhou W.J., L. D. H. (2006b). Global convergence of a modified Fletcher-Reeves conjugate method with Armijo type line search. *Numerische Mathematik*, *104*, 561-572.