Available online at www.refaad.com
Gen. Lett. Math., 7(2)2019, 79-86
Research Article

General Letters in Mathematics (GLM)

Journal Homepage: http://www.refaad.com/views/GLM/home.aspx
ISSN: 2519-9277 (Online) 2519-9269 (Print)

A new class of three-term conjugate gradient methods for solving unconstrained minimization problems

Alaa Luqman Ibrahim [a*], Salah Gazi Shareef [b]

[a,b] Department of Mathematics, Faculty of Science, University of Zakho, Zakho, Kurdistan Region, Iraq
Emails: [a] alaa.ibrahim@uoz.edu.krd, [b] salah.shareef@uoz.edu.krd

Abstract

Conjugate gradient (CG) methods which are usually generate descent search directions, are beneficial for large-scale unconstrained optimization models, because of its low memory requirement and simplicity. This paper studies the three-term CG method for unconstrained optimization. The modified a three-term CG method based on the formal $\boldsymbol{t}^*$ which is suggested by Kafaki and Ghanbari [11], and using some well-known CG formulas for unconstrained optimization. Our proposed method satisfies both (the descent and the sufficient descent) conditions. Furthermore, if we use the exact line search the new proposed is reduce to the classical CG method. The numerical results show that the suggested method is promising and exhibits a better numerical performance in comparison with the three- term (ZHS-CG) method from an implementation of the suggested method on some normal unconstrained optimization test functions.

Keywords: *Unconstrained optimization, three-term Conjugate gradient method, (descent and sufficient descent) conditions.*
2010 MSC No: 80M50, 90C26.

1. Introduction

In this paper, we are interested to resolving unconstrained optimization problem, particularly for large scale which is given by the form

$$min\, h(z) \qquad \forall\, z \in R^n \tag{1.1}$$

Where $h: R^n \rightarrow R$ and $h \in C^2$. Numerous similar professional fields of science that can revert to the above optimization problem (see, e.g., [7, 19]). A nonlinear conjugate gradient method is an iterative scheme that creates a sequence $\{z_i\}$ of an approximation to the solution of (1.1), using the repetition:

$$z_{i+1} = z_i + s_i, \;\; i = 0,1,2, \dots \;\;\;, \;\; s_i = \lambda_i d_i \tag{1.2}$$

Where $\lambda_i > 0$ is the step length, the present iterate is z_i, and the search direction d_i is designed by:

* Corresponding author
Email address: alaa.ibrahim@uoz.edu.krd (Alaa Luqman Ibrahim)
doi:10.31559/glm2019.7.2.4
Received Date: *20 Oct 2019*
Accepted Date: *25 Dec 2019*

$$d_i = \begin{cases} -g_i & if \quad i = 0 \\ -g_{i+1} + \beta_i \, d_i & if \quad i \geq 1 \end{cases} \tag{1.3}$$

Where $g_i = \nabla h\,(z_i)$ is the gradient and β_i is a scalar and called the conjugate gradient parameter. For example, Hestenes and Stiefel (HS) [9], Polak-Ribiere-Polyak (PRP) [14], Fletcher and Reeves (FR) [5], used an update parameter, respectively, given by:

$$\beta_i^{HS} = \frac{g_{i+1}^T y_i}{d_i^T y_i} \tag{1.4}$$

$$\beta_i^{PRP} = \frac{g_{i+1}^T y_i}{g_i^T g_i} \tag{1.5}$$

$$\beta_i^{FR} = \frac{g_{i+1}^T g_{i+1}}{g_i^T g_i} \tag{1.6}$$

Where $y_i = g_{i+1} - g_i$. The PRP method is very effective as regards numerical performance, but it failure as regards the global convergence for the general functions under Wolfe line search method and that is open problem many researcher want to solve it. It is worth noting that a recent work of Yuan et al. [16] proved the global convergence of PRP method under a modified Wolfe line search method for general functions. Al-Baali [2], Gilbert and Nocedal [8] and Hu and Storey [10] hinted that the sufficient descent property may be decisive for the global convergence of the CG methods including the PRP method.

The CG method have another important class is called the three-term conjugate gradient method in which the search direction is determined as a linear combination of g_i , s_i, and y_i as

$$d_i = -g_i - \tau_1 s_i + \tau_2 y_i \tag{1.7}$$

Where τ_1 and τ_2 are scalar. Among the generated three-term conjugate gradient methods in the literature we have the three-term conjugate methods proposed by Zhang et al. [17, 18] by considering a descent modified PRP and also a descent modified HS conjugate gradient method as

$$d_{i+1}^{ZPRP} = \begin{cases} -g_{i+1} & if \quad i = 0 \\ -g_{i+1} + \frac{g_{i+1}^T y_i}{g_i^T g_i} d_i - \frac{g_{i+1}^T d_i}{g_i^T g_i} y_i & if \quad i \geq 1 \end{cases} \tag{1.8}$$

and

$$d_{i+1}^{ZHS} = \begin{cases} -g_{i+1} & if \quad i = 0 \\ -g_{i+1} + \frac{g_{i+1}^T y_i}{d_i^T y_i} d_i - \frac{g_{i+1}^T d_i}{d_i^T y_i} y_i & if \quad i \geq 1 \end{cases} \tag{1.9}$$

In the same manner Nazareth [13] submit a computationally effective three-term nonlinear conjugate gradient method with the following search direction:

$$d_{i-1} = -y_i + \frac{y_i^T y_i}{y_i^T d_i} d_i - \frac{y_{i-1}^T y_i}{y_{i-1}^T d_{i-1}} d_{i-1} \tag{1.10}$$

The step-length λ_i in (1.2) is computed by carrying out a line search. The standard strong Wolfe line search [15] is to compute λ_i such that

$$h(z_i + \lambda_i d_i) \leq h(x_i) + \delta \lambda_i g_i^T d_i \tag{1.11}$$

$$|g(z_i + \lambda_i d_i)^T d_i| \leq -\sigma g_i^T d_i \tag{1.12}$$

However, the standard Wolfe condition (1.11) and

$$g(z_i + \lambda_i d_i)^T d_i \geq \sigma g_i^T d_i \quad , \text{ where .} \quad 0 < \delta < \sigma < 1 \tag{1.13}$$

The residual part of this paper is organized as follows: Section two, transact with the derivation of the new three-term conjugate gradient method (NTT-CG). In section three, we present prove of descent condition and sufficient descent condition of the (NTT-CG) method. The numerical results and discussion are reported in Section four. Finally, concluding the paper with final remarks in the last section.

2. Derivation A New Three-Term Conjugate Gradient Method (NTT-CG)

In this work, we transact with a modified form of the search direction (1.9). In 2014 Kafaki and Ghanbari [11] proposed a new value of t^* as follow:

$$t^* = \frac{\|y_i\|}{\|s_i\|} + \frac{s_i^T y_i}{\|s_i\|^2} \tag{2.1}$$

Now, let the parameter t_i is chosen here such that it presents the convex combination of first and second term of equation (2.1). Hence

$$t_i = \gamma \frac{\|y_i\|}{\|s_i\|} + (1-\gamma)\frac{s_i^T y_i}{\|s_i\|^2}, \quad \text{where } \gamma \in (0,1)$$

and by adding a parameter t_i on the third term of d_i were $i \geq 1$ with changed parameter in second term as one of classical formula of CG method. More precisely, the search direction of our method, named as (NTT-CG) method, defined by:

$$d_{i+1}^{NTT-CG} = \begin{cases} -g_{i+1} & if \quad i=0 \\ -g_{i+1} + \beta_i d_i - t_i\left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) y_i & if \quad i \geq 1 \end{cases} \tag{2.2}$$

Where $t_i = \gamma \frac{\|y_i\|}{\|s_i\|} + (1-\gamma)\frac{s_i^T y_i}{\|s_i\|^2}$ and the parameter β_i is given from normal CG method, in this paper we use the formula of Hestenes and Stiefel (HS) , Polak-Ribiere-Polyak (PRP) and Fletcher and Reeves (FR). Then the new search directions can be written as follows:

1. **If $\boldsymbol{\beta_i = \beta_i^{HS}}$, then**

$$d_{i+1}^{NTT-HS} = \begin{cases} -g_{i+1} & if \quad i=0 \\ -g_{i+1} + \beta_i^{HS} d_i - t_i\left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) y_i & if \quad i \geq 1 \end{cases} \tag{2.3}$$

2. **If $\boldsymbol{\beta_i = \beta_i^{PRP}}$, then**

$$d_{i+1}^{NTT-PRP} = \begin{cases} -g_{i+1} & if \quad i=0 \\ -g_{i+1} + \beta_i^{PRP} d_i - t_i\left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) y_i & if \quad i \geq 1 \end{cases} \tag{2.4}$$

3. **If $\boldsymbol{\beta_i = \beta_i^{FR}}$, then**

$$d_{i+1}^{NTT-FR} = \begin{cases} -g_{i+1} & if \quad i=0 \\ -g_{i+1} + \beta_i^{FR} d_i - t_i\left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) y_i & if \quad i \geq 1 \end{cases} \tag{2.5}$$

Remark 2.1: we can use another formula of CG method to obtain a new three term search direction see ([1, 4, 6, 12]).

Remark 2.2: Note that if $t_i = 0$, or the exact line search is used in equation (2.3), then the method reduces to the HS method. Also, if $t_i = 1$, then the method is precisely the ZHS method.

Remark 2.3: It is clear, if $t_i = 0$, or the line search is exact, then the direction generated by (2.4) and (2.5) reduces to PRP and FR methods respectively.

Remark 2.4: Obviously, γ, $(1-\gamma)$, $\|y_i\|, \|s_i\|^2$, $\|s_i\|$ and $s_i^T y_i$ are positive, therefore the parameter t_i is greater than or equal to zero.

2.1 Algorithm of (NTT-CG) Method

Step 1 : Given an initial point $z_0 \in R^n$ and set $d_0 = -g_0, i = 0$.

Step 2 : If $\|g_i\| = 0$ then stop, otherwise go to Step 3.

Step 3 : Determine the step size λ_i by using cubic line search to minimize $h(z_i + \lambda_i d_i)$.

Step 4 : Set $z_{i+1} = z_i + s_i$.

Step 5 : Compute g_{i+1}, if $\|g_{i+1}\| \leq 10^{-5}$ stop.

else go to Step 6.

Step 6 : Compute d_{i+1} by using one of the equations [(2.3), (2.4) or (2.5)].

Step 7 : If $\|g_{i+1}\|^2 \leq \frac{|g_i^T g_{i+1}|}{0.2}$ is satisfied go to step 2,

else $i = i + 1$ and go to step 3.

3. The Descent And Sufficient Descent Conditions Of The New Three-Term (NTT-CG) Method

An important feature for any minimization algorithm is the descent and the sufficient descent conditions. In this section, we examine the behaviour of decent and sufficient decent conditions of the (NTT-CG) method.

Lemma 3.1. The search direction d_{i+1} defined in equation (2.2) satisfies the descent condition

$$g_{i+1}^T d_{i+1} \leq 0 \tag{3.1}$$

Proof: From (2.2) we have if $i = 0$

$$, g_0^T d_0^{NTT-CG} = -\|g_0\|^2 \leq 0$$

Suppose that $g_i^T d_i^{NTT-CG} \leq 0 , \forall\, i.$

Now, we prove the present search direction is descent direction at the iteration $(i + 1)$.

$$d_{i+1}^{NTT-CG} = -g_{i+1} + \beta_i d_i - t_i (\frac{g_{i+1}^T d_i}{d_i^T y_i}) y_i \tag{3.2}$$

By multiplying equation (3.2) by g_{i+1}^T, we have

$$g_{i+1}^T d_{i+1}^{NTT-CG} = -\|g_{i+1}\|^2 + \beta_i g_{i+1}^T d_i - t_i \left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) g_{i+1}^T y_i \tag{3.3}$$

If $g_{i+1}^T d_i = 0$, then the equation (3.3) is satisfies the descent condition i.e.

$$g_{i+1}^T d_{i+1}^{NTT-CG} = -\|g_{i+1}\|^2 \leq 0$$

However, if $g_{i+1}^T d_i \neq 0$.We concludes that the first two terms of equation (3.3) are achieve the descent condition i.e.

$$, -\|g_{i+1}\|^2 + \beta_i\, g_{i+1}^T d_i \leq 0 \tag{3.4}$$

Where $\beta_i = (\beta_i^{HS}, \beta_i^{PR}\ and\ \beta_i^{FR})$, because the HS, PRP and FR methods satisfies the descent condition. Therefore equation (3.3) can be written as

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq -t_i (\frac{g_{i+1}^T d_i}{d_i^T y_i}) g_{i+1}^T y_i \tag{3.5}$$

Multiplying both sides by (-1), we have

$$-g_{i+1}^T d_{i+1}^{NTT-CG} \geq t_i (\frac{g_{i+1}^T d_i}{d_i^T y_i}) g_{i+1}^T y_i \tag{3.6}$$

By using the standard Wolfe condition, $g_{i+1}^T d_i \geq \sigma d_i^T g_i = -\sigma\|g_i\|^2$, then

$$-g_{i+1}^T d_{i+1}^{NTT-CG} \geq t_i (\frac{-\sigma\|g_i\|^2}{d_i^T y_i}) g_{i+1}^T y_i \tag{3.7}$$

Now, multiplying the equation (3.7) by(-1) , we get

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq t_i (\frac{\sigma\|g_i\|^2}{d_i^T y_i}) g_{i+1}^T y_i \tag{3.8}$$

Since the (Lipschize condition) $\|y_i\| \leq L\lambda_i \|d_i\|$ achieves the following inequality

$$g_{i+1}^T y_i \leq L\lambda_i g_{i+1}^T d_i \text{ , where } . L > 0$$

by using the above inequality we obtained:

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq t_i(\frac{\sigma\|g_i\|^2}{d_i^T y_i})L\lambda_i g_{i+1}^T d_i \quad (3.9)$$

Now, by using the strong Wolfe condition $g_{i+1}^T d_i \leq -\sigma d_i^T g_i$

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq -t_i(\frac{\|g_i\|^2}{d_i^T y_i})\sigma^2 L\lambda_i d_i^T g_i \quad (3.10)$$

We know that $g_{i+1}^T d_i \leq d_i^T y_i$, and by Wolfe condition we can write

$\sigma d_i^T g_i \leq d_i^T y_i$, and since $d_i^T g_i < 0$,
Implies $\frac{1}{d_i^T y_i} \geq \frac{1}{\sigma d_i^T g_i} \quad \Rightarrow \quad \frac{-1}{d_i^T y_i} \leq \frac{-1}{\sigma d_i^T g_i}$
So, we have

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq -t_i(\frac{\|g_i\|^2}{\sigma d_i^T g_i})\sigma^2 L\lambda_i d_i^T g_i \quad (3.11)$$

This implies that

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq -t_i L\lambda_i \sigma \|g_i\|^2 \quad (3.12)$$

Obviously, L, σ, λ_i, $\|g_i\|^2$ and t_i are positive. Hence, we obtain

$g_{i+1}^T d_{i+1}^{NTT-CG} \leq 0$

Lastly, we get to

$$g_{i+1}^T d_{i+1}^{NTT-CG} = -\|g_{i+1}\|^2 + \beta_i g_{i+1}^T d_i - t_i\left(\frac{g_{i+1}^T d_i}{d_i^T y_i}\right) g_{i+1}^T y_i \leq 0. \quad ■$$

Lemma 3.2. The search directions defined by (2.2) in which the parameter β_i is computed by one of the equations [(1.4), (1.5) and (1.6)] satisfy the sufficient descent condition,

$$g_{i+1}^T d_{i+1} \leq -K\|g_{i+1}\|^2 \quad (3.13)$$

Proof: It is evident that the first two terms of equation (3.3) are less than or equal to zero and from equation (3.12), we have

$$g_{i+1}^T d_{i+1}^{NTT-CG} \leq -[t_i L\lambda_i \sigma\left(\frac{\|g_i\|^2}{\|g_{i+1}\|^2}\right)] \|g_{i+1}\|^2 \quad (3.14)$$

This means that the equation (3.11) has been achieved where

$K = [t_i L\lambda_i \sigma\left(\frac{\|g_i\|^2}{\|g_{i+1}\|^2}\right)]$. ■

4. Numerical results and discussion

This section will report the numerical experiment of the modified three term CG $(NTT - CG)$ with $(ZHS - CG)$ methods. The tests include well-known nonlinear problems standard test functions [3], with size of the variable ranging from $10 \leq n \leq 5000$. The code of the algorithm is written by FORTRAN 95 language with given initial points. In Table (1) the numerical results show that the $(NTT - CG)$ method is more effective than $(ZHS - CG)$ method with respect to NOI and NOF, and we notate:
n: The variables number.
NOI: Number of iterations
NOF: Number of functions evaluation.
NTT-CG: New Three-Term Conjugate Gradient Methods

ZHS-CG: three-term conjugate methods proposed by Zhang et al. [18] with Hestenes and Stiefel (HS) parameter.

Table(1): The results for the (NTT-CG) and (ZHS-CG) methods on the tested problems

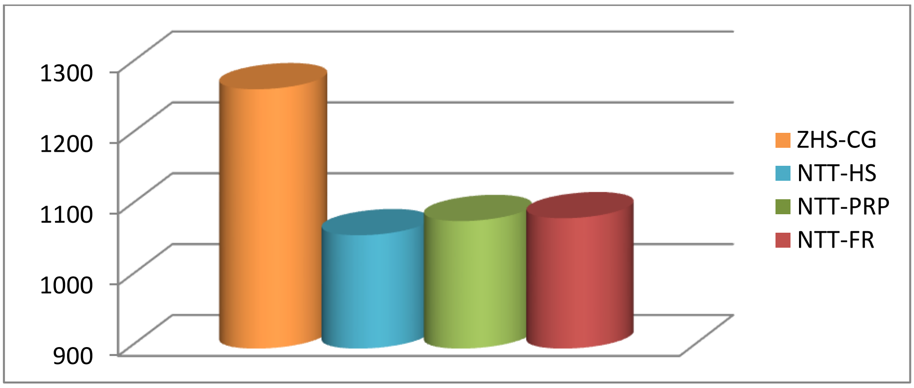
Test Function	n	$ZHS-CG$		$NTT-HS$		$NTT-PRP$		$NTT-FR$	
		NOI	NOF	NOI	NOF	NOF	NOF	NOI	NOF
Beal	10	11	28	11	26	11	26	11	26
	100	12	30	11	26	11	26	11	26
	500	12	30	12	28	11	26	12	28
	1000	12	30	12	28	11	26	12	28
	5000	12	30	12	28	11	28	12	28
Wood	10	29	67	23	56	26	63	24	57
	100	29	67	23	56	26	63	24	57
	500	29	67	26	62	26	63	25	60
	1000	29	67	26	62	26	63	25	60
	5000	29	67	28	66	27	65	25	60
Wolfe	10	32	65	33	67	32	65	28	57
	100	49	99	41	83	48	97	44	89
	500	51	103	43	87	51	103	43	87
	1000	59	119	46	93	51	103	46	93
	5000	159	331	105	224	157	329	129	270
Sum	10	7	44	7	45	7	44	7	45
	100	14	83	14	83	14	80	14	85
	500	20	109	19	111	19	105	20	104
	1000	25	114	14	77	24	114	22	110
	5000	35	184	32	154	29	136	28	132
Powell	10	38	100	30	90	24	71	30	90
	100	41	117	32	94	24	71	32	94
	500	43	121	35	109	24	71	35	109
	1000	43	121	35	109	24	71	35	109
	5000	43	121	35	109	24	71	35	109
Mile	10	46	157	41	129	35	113	41	129
	100	53	190	47	163	43	145	47	163
	500	53	190	48	165	49	177	48	165
	1000	57	209	54	200	50	179	54	200
	5000	60	226	55	202	56	212	55	202
Central	10	22	155	21	150	21	149	21	150
	100	26	204	21	150	22	163	21	150
	500	26	204	22	163	22	163	22	163
	1000	28	231	22	163	22	163	22	163
	5000	32	292	24	195	22	163	24	195
Total		1266	4372	1060	3653	1080	3607	1084	3693

Table(2): The percentage of improvement between the (NTT-CG) and (ZHS-CG) methods

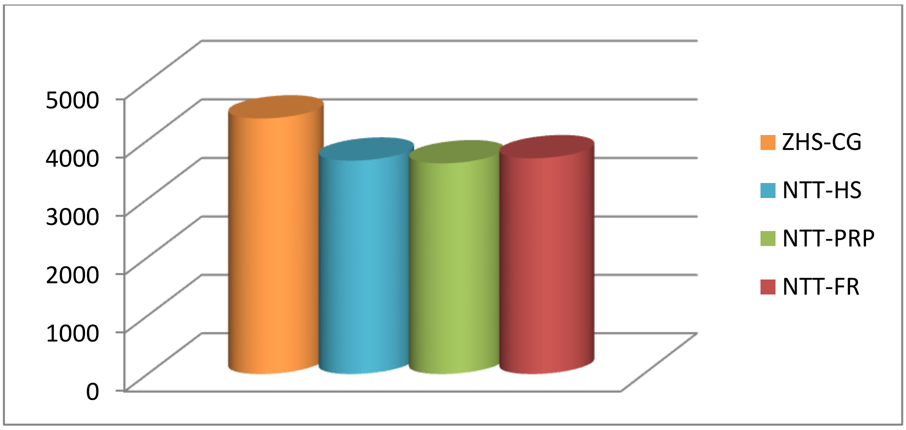
Tools	$ZHS-CG$	$NTT-HS$	$NTT-PRP$	$NTT-FR$
NOI	100%	83.7283%	85.3081%	85.6240%
NOF	100%	83.5544%	82.5023%	84.4694%

Table 2 shows the rate of improvement in the new method $(NTT-CG)$ with the standard method $(ZHS-CG)$. As we notice that NOI and NOF of the $(ZHS-CG)$ method are about 100%. That means, the new method has improvement as compared to standard method with 16.2717% in NOI and 16.4456% in NOF if we using the search direction (2.3), and 14.6919% in NOI and 17.4977% in NOF, when used the search direction (2.4). Lastly the rate of improvement in the $(NTT-FR)$ method is 14.376% in NOI and 15.5306% in NOF.

Indeed, we can see from Table 2 the $(NTT-HS)$ method is more efficient than $(NTT-PRP)$ and $(NTT-FR)$ with respect to NOI while the $(NTT-PRP)$ is more active than the $(NTT-HS)$ and $(NTT-FR)$ depending on NOF. Finally, it is obvious that the $(NTT-CG)$ is reliable method compare with the standard one $(ZHS-CG)$.



Figure(1): Shows the comparison between new methods and the (ZHS-CG) method according to the total number of iterations (NOI)



Figure(2): Shows the comparison between new methods and the (ZHS-CG) method according to the total number of functions (NOF)

5. Conclusion

In this paper, a modified three-term conjugate gradient method for solving nonlinear unconstrained optimization in formula (2.2) based on a formula t^* by using some well-known CG formulas (HS, PRP and FR) for unconstrained optimization is presented. The proposed method possesses descent and sufficient descent conditions also holds without any line search technique. The numerical results show that the proposed method is promising and more efficient than the normal method $(ZHS-CG)$ considered. In future we can use the $(NTT-CG)$ method for training the neural network in order to investigate the performance of its behaviour.

References:

[1] A. L. Ibrahim., M. A. Sadiq. & S. G. Shareef., *A New Conjugate Gradient Coefficient for Unconstrained Optimization Based On Dai-Liao, JUOZ ,* 7(1)(2019),34-36 https://doi.org/10.25271/sjuoz.7.1.525

[2] M. AlBaali, *Descent property and global convergence of Fletcher-Reeves method with inexact line search*, IMA. J.Numer.Anal., 5(1985), 121-124, https://doi.org/10.1093/imanum/5.1.121

[3] N. Andrei, *An Unconstrained Optimization Test Functions Collection*, Advanced Modeling and Optimization, 10 (2008), 147-161.

[4] Y. H. Dai & L.Z. Liao, *New conjugacy conditions and related nonlinear conjugate gradient methods*, Application Mathematical Optimization, 43(1)(2001), 87-101, https://doi.org/10.1007/s002450010019

[5] R. Fletcher & C.M. Reeves, *Function minimization by conjugate gradients*, The Computer Journal, 7(2)(1964), 149-154, https://doi.org/10.1093/comjnl/7.2.149

[6] R. Fletcher, *Practical methods of optimization unconstrained optimization*, John Wiley & Sons, New York, NY, USA, (1987).

[7] Z. Fu, X . Wu, C. Guan, X. Sun & K. Ren, *Towards efficient multi-keyword fuzzy search over encrypted outsourced data with accuracy improvement*, IEEE Trans. Inf. Forensics Secur, 11(12)(2016),2706-2716, https://doi.org/10.1109/tifs.2016.2596138

[8] J.C. Gilbert & J. Nocedal, *Global convergence properties of conjugate gradient methods for optimization*, SIAM J. Optim, 2(1)(1992), 21-42, https://doi.org/10.1137/0802003

[9] M. R. Hestenes & E. Stiefel, *Methods of conjugate gradients for solving linear systems*, Journal of Research of the National Bureau of Standards, 49(6)(1952), 409-436, https://doi.org/10.6028/jres.049.044

[10] Y.F. Hu & C. Storey, *Global convergence result for conjugate method*, J. Optim. Theory Appl, 71(1991), 399-405.

[11] S. B. Kafaki & R. Ghanbari, *The Dai-Liao nonlinear conjugate gradient method with optimal parameter choices*, European Journal of Operational Research, 234(3)(2014), 625-630, https://doi.org/10.1016/j.ejor.2013.11.012

[12] H. Khatab, S. Shareef, *A New Parameter Conjugate Gradient Method Based on Three Terms Unconstrained Optimization*, General Letters in Mathematics, 7(1)(2019), 39–44, doi:10.31559/glm2019.7.1.5

[13] L. Nazareth, *A conjugate direction algorithm without line searches*, Journal of Optimization Theory and Applications, 23(3)(1977), 373–387, https://doi.org/10.1007/bf00933447

[14] E. Polak & G. Ribiere, *Note surla convergence des méthodes de directions conjuguées.*, Revue française d'informatique et de recherche opérationnelle, 3(16)(1969), 35-43, https://doi.org/10.1051/m2an/196903r100351

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

[16] G. Yuan, Z. Wei & X. Lu, *Global convergence of the BFGS method and the PRP method for general functions under a modified weak Wolfe–Powell line search*, Appl. Math. Model, 47(2017), 811–825, doi: 10.1016/j.apm.2017.02.008.

[17] L. Zhang, W. Zhou & D-H. Li, *A descent modified Polak-Ribière-Polyak conjugate gradient method and its global convergence,* IMA J. Numer. Anal, 26(4)(2006), 629-640, https://doi.org/10.1093/imanum/drl016

[18] L. Zhang, W. Zhou & D. Li, *Some descent three-term conjugate gradient methods and their global convergence, Optim*. Methods Softw. 22(4)(2007), 697-711, https://doi.org/10.1080/10556780701223293

[19] Z. Zhou, Y. Wang, , QMJ. Wu, C-N. Yang & X. Sun, *Effective and efficient global context verification for image copy detection*, IEEE Trans. Inf. Forensics Secur, (2016), doi:10.1109/TIFS.2016.2601065.