International Journal of Engineering & Technology, 7 (3.28) (2018) 325-328

International Journal of Engineering & Technology

Website: www.sciencepubco.com/index.php/IJET

Research paper

Secant Condition Free of a Spectral PRP Conjugate Gradient Method

Usman Abbas Yakubu[1, 2], Mustafa Mamat[2]*, Mohamad Afendee Mohamed[2], Sukono[3], Mohd Rivaie[4]

[1]*Department of Mathematics, Northwest University, Kano. Nigeria*
[2]*Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (UniSZA), Kuala Terengganu, Malaysia*
[3]*Fakultas MIPA, Universitas Padjadjaran, Bandung, Indonesia*
[4]*Department of Computer Science & Mathematics, Universiti Teknologi MARA (UiTM), Kuala Terengganu, Malaysia*
**Corresponding author E-mail: must@unisza.edu.my*

Abstract

Recently conjugate gradient method have been in practice widely to solve unconstrained minimization problems as a result of fewer storage locations and less computational expensive in dealing with the large-scale problems. In this work, we suggested a spectral PRP CG method without employing the secant condition and use some unconstrained problems with many variables to prove its sufficient descent as well as global convergence, the results is validated by apply exact line search.

Keywords: *Global convergence; exact line search; spectral CG; secant condition; sufficient descent property*.

1. Introduction

In the paper [1] Raydan have introduced a spectral gradient method (SGM) for a large-scale unconstrained optimization problems. Recently, Birgin and Martinez [3] purposed a spectral CG method and they computed their spectral parameter using standard secant equation as initially acquainted by Barzilai and Borwein [11]. A spectral CG method combines the advantages of a spectral parameter and classical CG direction by constructing a new search direction, see Du et al. [12], Zull et al. [13], Yakubu et al. [16], Abashar et al. [20], Abashar et al. [21] and Raydan et al. [1] for more details.

In this research we derived the spectral parameter $\theta_k$by using the classical CG direction and a scalar parameter β_k of Polak-Ribiére-Polyak (PRP) method without employing the secant condition. A new spectral PRP CG method is presented to solve large-scale unconstrained minimization problems and compare its performance with classical PRP, FR conjugate gradient methods. Therefore, we consider the general unconstrained optimization problem:

$$\min f(x), \qquad x \in R^n \tag{1}$$

where $f: R^n \rightarrow R$ is continuous as well as differentiable, g_k is a gradient vector of a function f and initial point $x_0 \in R^n$ is normally solved iteratively according to the recurrence expression

$$x_{k+1} = x_k + \gamma_k d_k\,, \qquad k = 0,1,2,3,4,\dots \tag{2}$$

where as x_k stand for a current iteration, $\gamma_k > 0$ represent a step size obtained by line search method called exact line search, given as

$$\gamma_k = arg \min_{\gamma>0} f(x_k + \gamma d_k) \tag{3}$$

also d_k is a search direction defined as

$$d_k = \begin{cases} -g_k, & if\ k = 0 \\ -g_k + \beta_k d_{k-1}, & if\ k \geq 1 \end{cases} \tag{4}$$

and $g_k = \nabla f(x)$, parameter $\beta_k \in R$ are the gradient vector and CG coefficient, respectively. A classical β_k of FR and PRP are given below:

$$\beta_k^{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \tag{5}$$

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} \tag{6}$$

where g_k and g_{k-1} in the above equations are gradient vectors of function f at points x_k , x_{k-1} respectively, and $\|.\|$ represent a Euclidian norm. Generally, consideration has been made on a global convergence of a CG methods above. In paper [5] G. Zoutendijk proved FR method on a line search (3) converge globally. Later, in paper [9] Powell disproved this assertion and noted clearly that the method have demonstrated a poor practical behaviour and its convergence is not global. PRP method is the best CG method amongst others but in general the convergence analysis for a nonlinear function is uncertain [14].

Nevertheless, in paper [6] J. C. Gilbert and Nocedal established a global convergence result of the PRP method by limiting a scalar β_k to be non-negative that is $\beta_k{}^{PRP} = max\{0, \beta_k{}^{PRP}\}$. However, spectral CG method is more effective in terms of a numerical execution than the other methods see also Raydan et al. [1], Du et al. [12], Zull et al. [13], Hu [7], Wu [4], Yakubu et al. [16], Abidin et al. [18], Mamat et al. [19], and Huang et al. [8].

In this piece of research work, we present a new spectral PRP CG method without using the secant condition and tested its performance with classical formula PRP and FR methods. The details of the method is presented in section 2, we demonstrate a sufficient descent condition and global convergence in section 3. Numerical result are also presented in section 4 based on CPU time and number of iterations. Conclusion follows in section 5.

2. Description of the New Method

A spectral conjugate gradient method proposed originally by Barzilai and Borwein [11], the direction is generated by $d_k = -\theta g_k + \beta_k s_{k-1}$ where $s_{k-1} = \gamma_{k-1} d_{k-1}$ and θ_k is a spectral parameter. Birgin and Martinez [3] purposed a spectral CG method and they computed spectral parameter θ_k using standard secant equation. The recent paper [13] N. Zull, encourage us to determine and projecting the new method. The direction is defined as

$$d_k = \begin{cases} -g_k, & if\ k = 0 \\ -\theta_k g_k + \beta_k^{PRP} d_{k-1}, & if\ k \geq 1 \end{cases} \quad (7)$$

From equation (7) above, we have

$$d_k = -\theta_k g_k + \beta_k^{PRP} d_{k-1} \rightarrow \quad d_k - \beta_k^{PRP} d_{k-1} = -\theta_k g_k$$

using the fact that $d_k = -g_k$ from (7) and substituting (6), we have

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

Note that $g_k^T g_{k-1} = 0$ implies that

$$\theta_k = 1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}} \quad (8)$$

where θ_k is a new spectral parameter computed without employing secant condition by using exact line search procedure.

Algorithm 2.1 (SPRP Method)

Step 1: Given a starting point $x_0 \in R^n$ set $k = 0$
Step 2: Calculate β_k as given in formula (6) above
Step 3: Calculate d_k as given in (7). If $\|g_k\| = 0$, then stop.
Step 4: Calculate γ_k as given in (3).
Step 5: Update the new point as given in the recurrence expression (2).
Step 6: If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| < \varepsilon$ then stop, otherwise go to step 1 with $k = k + 1$

3. Global convergence analysis

3.1. Sufficient descent condition

For a sufficient descent condition to hold, we consider

$$g_k^T d_k \leq -C\|g_k\|^2 \quad for\ k \geq 0\ and\ C > 0 \quad (9)$$

Theorem 3.1: *Suppose a CG method with search direction* (7) *and* β_k^{PRP} *given by* (5), *the condition* (9) *holds* $\forall\, k \geq 0$.

Proof: To prove the assertion, we proceed by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (9) satisfied as $k = 0$. Now, we assume it is true for $k \geq 0$. Condition (9) as well hold, from equation (7), multiplying both sides by g_{k+1}^T

$$g_{k+1}^T d_{k+1} = -\theta_k g_{k+1}^T g_{k+1} + \beta_k^{PRP} g_{k+1}^T d_k$$

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

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

It is known that $g_{k+1}^T d_k = 0$. Hence, (9) is true for $k + 1$. ■

3.2. Global convergence properties

The following assumptions applied on the objective function to psychoanalyze the global convergence of a general CG methods.

Assumptions 3.1

(i) A level set $\Omega = \{x \in R^n \mid f(x) \leq f(x_0)\}$ is bounded, the function f is continuously differentiable in a neighborhood N of the level set Ω and x_0 is a starting point.
(ii) $g(x)$ is globally Lipschitz continuous in N that is $\exists$ a constant $L > 0$, such that $\|g(x) - g(y)\| \leq L\|x - y\|$ for any $x, y \in N$.

Lemma 3.1: *Suppose the assumptions 3.1 holds and consider any recurrence expression* (2) *and direction* (7), γ_k *is found in equation* (3). *Then, Zoutendijk condition below holds.*

$$\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty \quad (10)$$

Proof of a lemma is in paper [5] G. Zoutendijk.

Theorem 3.2: *Suppose the assumptions 3.1 holds, for any* $\{x_k\}$, $\{d_k\}$ *be given as spectral PRP CG method,* γ_k *are determined by* (3) *and* β_k *can be found as in* (5). *Then*

$$\lim_{k\to\infty} \|g_k\| = 0 \quad (11)$$

Proof: From equation (7), we have $d_{k+1} + \theta_k g_{k+1} = \beta_k^{PRP} d_k$, and square both sides of the equation,

$$(d_{k+1} + \theta_k g_{k+1})^2 = \left(\beta_k^{PRP} d_k\right)^2$$

$$\|d_{k+1}\|^2 = \left(\beta_k^{PRP}\right)^2 \|d_k\|^2 - 2\theta_k g_{k+1}^T d_{k+1} - \theta_k^2 \|g_{k+1}\|^2 \quad (12)$$

Substituting (6) into (12), we obtain

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

$$\|d_{k+1}\|^2 = \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_k\|^2 - 2\theta_k g_{k+1}^T d_{k+1} - \theta_k^2 \|g_{k+1}\|^2 \quad (13)$$

Also, substituting $g_{k+1}^T d_{k+1} = -C\|g_{k+1}\|^2$ in (13) above

$$\|d_{k+1}\|^2 = \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_k\|^2 + 2C\theta_k \|g_{k+1}\|^2 - \theta_k^2 \|g_{k+1}\|^2$$

$$\|d_{k+1}\|^2 = \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_k\|^2 - \|g_{k+1}\|^2 (\theta_k^2 - 2C\theta_k) \quad (14)$$

Multiply both side of equation (14) by $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$, we have

$$\|d_{k+1}\|^2 \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}\left(\frac{\|g_k\|^4}{\|g_{k-1}\|^4}\|d_k\|^2 - \|g_{k+1}\|^2(\theta_k^2 - 2C\theta_k)\right)$$

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left((2C\theta_k - \theta_k^2) + \frac{\|g_k\|^4}{\|g_{k-1}\|^4}\|d_k\|^2\right) \quad (15)$$

Substituting (8) in (15) and note that for exact line search $g_{k+1}^T d_k = 0$, we get

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left((2C-1) + \frac{\|g_k\|^4}{\|g_{k-1}\|^4}\|d_k\|^2\right)$$

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \qquad (16)$$

From the Lemma 3.1, $\lim_{k\to\infty} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} < 0$. It implies that Theorem 3.2 does not hold true, then $\lim_{k\to\infty} \frac{(g_{k+1}^T d_{k+1})^2}{\|d_{k+1}\|^2} = \infty$ and from equation (16) this is true $\infty \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So, Theorem 3.2 is true for a sufficient large k. ■

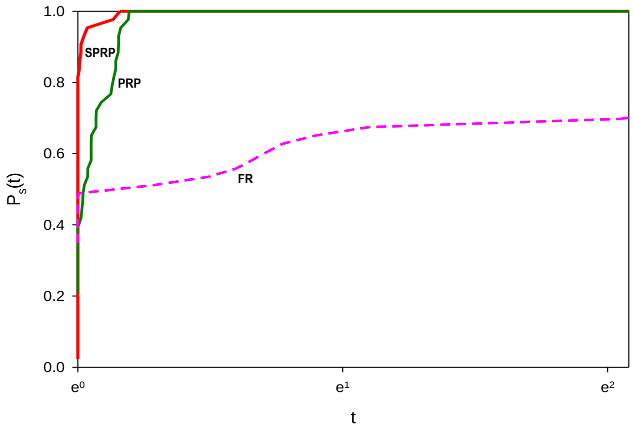

Fig. 1: Performance profile based on number of iterations

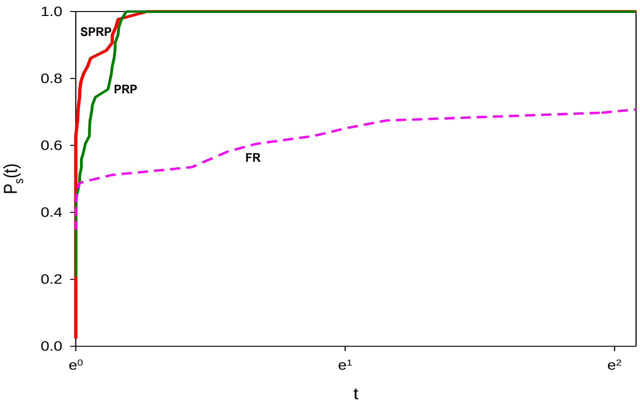

Fig. 2: Performance profile based on CPU time

Finally, a new method has absolute potentials when compared with the classical PRP and FR methods. Based on the comparison in the above figures, it shows that spectral PRP method and PRP method successfully reached the solution points, indeed the efficiency of the spectral PRP method is highly applaud.

4. Results and Discussion

In this section, we test the Algorithm 2.1 above and compared its performance with a classical FR and PRP methods. The comparisons are established on CPU time as well as number of iterations. The step size is found by using exact line search, the stopping criteria used for both methods is as suggested by Hillstrom and $\|g_k\| < \varepsilon$. A test problems in paper [15] N. Andrei were used, we also established a problems in Table 1 with four different initial values by MatlabR2012 subroutine programming using Intel® Core™ i5-3317U (1.7GHz with 4 GB (RAM)). In this research, we represent failure due to (i) memory requirement (ii) number of iterations exceed 1000.
The results shown in a Figure 1 and 2 are performance profile acquainted by paper [2] Dolan and More. Performances are established on CPU time and number of iterations respectively. Therefore, the highest value of a percentage of probability will be regarded as the best performing method and also the method that reached the top foremost is considered most sophisticated upon all the methods.

Table.1: Test Problems functions

Functions	Dimensions	Initial Points
Trecanni	2	(5,5), (8,8), (-11,-11), (-15,-15)
Leon	2	(4,4), (-4,-4), (6,6), (-10,-10)
Extended Penalty	2,4,10,50	(2,..,2), (-2,..,-2), (5,..,5), (-5,..,-5)
Ext. quadratic penalty QP2	10,100	(2,..,2), (6,..,6), (8,..,8), (-10,..,-10)
Ext. quadratic penalty QP1	10,100	(5,..,5), (-5,..,-5), (8,..,8), (-8,..,-8)
Power	2,4,50,100	(5,..,5), (5,..,5), (100,..,100), (100,..,-100)
Extended Himmelblau	1000	(2,..,2), (-2,..,-2), (25,..,25), (-25,..,-25)
Quadratic QF1	10,100,1000,10000	(2,….,2), (6,..,6), (8,..,8), (10,..,10)
Rosenbrock	2,4,10,100,1000,10000	(5,..,5), (13,..,13), (20,..,20), (40,..,40)
White and Holst	2,4,10,100,1000,10000	(2,..,2), (5,..,5), (9,..,9), (-9,..,-9)
Shallow	2,4,10,100,1000,10000	(100,100),(200,200),(400,400),(500,500)

5. Conclusion

A proposed spectral PRP conjugate gradient method satisfies a suffi-cient descent conditions and converges globally. The method overcomes the shortcomings of PRP method and the exhibited numerical results apparently shows that the method performed excellently compared with the famous PRP and FR methods in terms of CPU time and number of iterations respectively.

Acknowledgement

The authors thank the referees for their contributions and also grateful to the Malaysian government and Universiti Sultan Zainal Abidin for funding this research under the Fundamental Research Grant Scheme (FRGS/1/2017/STG06/Unisza/01/1).

References

[1] M. Raydan (1997). The Barzilai and J.M. Borwein gradient methods for the large scale unconstrained minimization in extreme problems, SIAM. J. Optim., 7(1), 26-33.
[2] E. Dolan, J.J. More (2002). Benchmarking optimization software with performance profile, Math. Prog., 91, 201-213.
[3] E.G. Birgin and J.M. Martinez (2011). A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Optim., 43(2), 117-128.
[4] X. Wu (2015). A new spectral Polak- Ribière -Polak conjugate gradient method, ScienceAsia, 41, 345-349.
[5] G. Zoutendijk (1970). Nonlinear programming, computational methods, in J Abadie (Ed.), Integer and Nonlinear Programming, Amsterdam: North-Holland, pp. 37-86.

[6] J. C. Gilbert, J. Nocedal (1992). Global convergence properties of conjugate gradient methods for optimization, SIAM. J. Optim., 2, 21-42.
[7] C. Hu, Z. Wan (2013). An extended spectral conjugate gradient method for unconstrained optimization problems, British Journal of Math. and Computer Science, 3, 86-98.
[8] H. Huang, Z. Wei, Y. Shengwei (2007). The proof of the sufficient decent condition of the Wei-Yao-Liu conjugate gradient method under the strong Wolfe-Powell line search, Applied Mathematics and Computation., 189, 1241–1245.
[9] M.J.D. Powell (1984). Non-convex minimization calculations and the conjugate gradient method. Lecture Notes in Mathematics, 1066, 122-241.
[10] M.J.D. Powell (1977). Restart procedures for the conjugate gradient method, Math. Program., 12, 241-254.
[11] J. Barzilai, J.M. Borwein (1988). Two point stepsize gradient methods, IMA J Numer Anal., 8, 141-148.
[12] X. Du, J. Liu (2011). Global convergence of a spectral HS conjugate gradient method, Procedia Engineering, 15, 1487-1492.
[13] N. Zull, M. Rivaie, M. Mamat, Z Salleh, Z. Amani (2015). Global convergence of a Spectral conjugate gradient by using strong Wolfe line search, Appl. Math. Sci., 63, 3105-3117.
[14] W.W. Hager, H. Zhang (2006). A survey of nonlinear conjugate gradient methods, Pacific Journal of Optimization, 2(1), 35-58.
[15] N. Andrei (2008). An unconstrained optimization test functions collection, Adv. Modell. Optim., 10, 147-161.
[16] A. Y. Usman, M. Mamat, M. Rivaie, A. M. Mohamad and B. Y. Rabi'u (2018). Secant free condition of a spectral WYL and its global convergence properties. Far East Journal of Mathematical Science, 103(12), 1889-1902.
[17] A. Y. Usman, M. Mamat, M. Rivaie, A. M. Mohamad and J. Sabi'u (2018). A recent modification on Dai-Liao conjugate gradient method for solving symmetric nonlinear equations. Far East Journal of Mathematical Science, 103(12), 1961-1974.
[18] N. Z. Abidin, M. Mamat, B. Dangerfield, J. H. Zulkepli, M. A. Baten and A. Wibowo (2014). Combating obesity through healthy eating behavior: A call for system dynamics optimization. Plos One, 9(12), 1-17.
[19] M. Mamat, Y. Rokhayati, N. M. M. Noor and I. Mohd (2011). Optimizing human diet problem with puzzy price using fuzzy linear programming approach. Pakistan Journal of Nutrition, 10(6), 594-598.
[20] A. Abashar, M. Mamat, M. Rivaie and I. Mohd (2014). Global convergence properties of a new class of conjugate gradient method for unconstrained optimization. Applied Mathematical Sciences Issue, 65-68, 3307-3319.
[21] A. Abashar, M. Mamat, M. Rivaie, I. Mohd and O. Omer (2014). The proof of sufficient descent condition for a new type of conjugate gradient methods. AIP Conference Proceedings, 1602, 296-303.