Malaysian Journal of Computing and Applied Mathematics 2019, Vol 2(1): 34-41

(Online)

MALAYSIAN JOURNAL OF COMPUTING AND APPLIED MATHEMATICS

Scalar Parameter of a Spectral PRP Conjugate Gradient Method for Unconstrained optimization

Usman A. Yakubu[a], Abdul Iguda[b], Abba V. Mandara[c] and Salisu Murtala[d]

[a] Department of Mathematics, Northwest University, Kano, Nigeria
[b] Department of Mathematical Sciences, Bayero University Kano, Nigeria
[c] Department of Mathematics, University of Maiduguri, Borno, Nigeria
[d] Department of Mathematics, Federal University Dutse, Jigawa, Nigeria

*Corresponding author: usman.abbas84@yahoo.com

Received: 01/03/2019, Accepted: 15/05/2019

Abstract

In recent times, conjugate gradient method (CG) have been broadly used to solve nonlinear unconstrained minimization problems as a result of fewer storage locations and computational expensive in dealing with large-scale problems. In this work, we present a spectral PRP CG method which derived from the CG search direction without secant condition and utilized some of the benchmark test problem functions with several variables to prove its global convergence properties and satisfies sufficient descent condition, the results are validated by exact line search techniques.

Keywords: Sufficient descent property; exact line search; spectral CG; global convergence.

Introduction

The CG method nowadays is the standard method for solving large-scale unconstrained optimization problems, one of the most important characteristics is to solve a large number of problems within the shortest period of time and less number of iterations. CG methods simply require a small storage location and are less computational expensive since they do not use the Hessian matrix or its approximation as normally used in other minimization methods. The CG method has rapid global convergent properties. The peculiarity of this method is paramount due to its simplicity in both algebraic processes and the development of computer codes. Therefore, the method is effective and capable in solving large-scale unconstrained minimization problems (Andre, 2008). In recent times, Birgin and Martinez (2011) introduced a spectral CG method and they computed their spectral parameter using standard secant equation as initially used by (Barzilai and Borwein, 1988). Spectral CG method combines CG search direction and scalar spectral parameter to construct a new search direction, see (Yakubu et al., (2018a, b, c); Abba et al., (2018); Zull et al., (2015); Hu (2013); Wu (2015); Abashar, (2014); Zhang, (2006); Du et al. (2011); Andrei (2008); and Raydan et al. (1997)), for more details.

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

where $f: R^n \to R$ is continuous as well as differentiable, g_k is a gradient vector of f and initial point $x_0 \in R^n$ are solved iteratively using recurrence expression below

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

the vector x_k stand for a current iteration and x_{k+1} is a new iteration, $\gamma_k > 0$ represent a step size obtained by the line search method given as

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

also d_k is a search direction

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

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

$$\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_{k-1}\|^2} \tag{5}$$

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. PRP method is the best CG method amongst others but in general, the convergence analysis for a nonlinear function is uncertain Zhang, (2006).
Nevertheless, Nocedal, (1992) established a global convergence result of the PRP method by limiting a scalar β_k to be non-negative that is ${\beta_k}^{PRP} = max\{0, {\beta_k}^{PRP}\}$. In this work, a new spectral PRP (SpPRP) CG method is presented without using the secant equation and verified its performance with recent spectral PRP (RSPRP) developed by Wu, (2015) and PRP CG methods.

Details of SpPRP CG Methods

Spectral CG method is suggested initially by Barzilai and Borwein, (1988) the direction is produced as $d_k = -\varphi_k\, g_k + \beta_k s_{k-1}$, where $s_{k-1} = \gamma_{k-1} d_{k-1}$ and φ_k is a spectral scalar parameter. The recent articles by Yakubu et al., (2018a) and Zull et al., (2015) motivate the researcher's on this piece of work to determine the SpPRP CG method, where the search direction is defined as

$$d_k = \begin{cases} -g_k, & if\ k = 0 \\ -\varphi_k g_k + \beta_k^{PRP} d_{k-1}, & if\ k \ \geq 1 \end{cases} \tag{6}$$

From the above search direction (6), $d_k = -\varphi_k g_k + \beta_k^{PRP} d_{k-1} \ \to \ d_k - \beta_k^{PRP} d_{k-1} = -\varphi_k g_k$ using the fact that $d_k = -g_k$ also from equation (6) and substituting equation (5) we have,

$$\varphi_k = 1 - \frac{g_k^T d_{k-1}}{g_{k-1}^T d_{k-1}} \tag{7}$$

Recall that the orthogonality of gradients $g_k^T g_{k-1} = 0$, thus φ_k is a new spectral parameter computed by exact line search technique.

Algorithm 1 (SpPRP CG Method)
Step 1: Given a starting point $x_0 \in R^n$ set $k = 0$
Step 2: Compute β_k as given in formula (5) above
Step 3: Compute d_k given as in (6). If $\|g_k\| = 0$, then stop.
Step 4: Compute γ_k given in equation (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 $k = k + 1$.

Global Convergence Analysis

A. Sufficient Descent Condition

Sufficient descent condition ensures that global convergence of iterative procedures or algorithm is achieved. Therefore, the following inequality must hold true.

$$g_k^T d_k \leq -C\|g_k\|^2 \qquad for\ k \geq 0\ and\ C > 0 \tag{8}$$

Theorem 1.1. *Suppose a CG method with search direction* (6) *and* β_k^{PRP} *given by* (5)*, the condition* (8) *holds* $\forall\ k \geq 0$.

Proof. We proceed by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (8) satisfied as $k = 0$. Now we assume it is true for $k \geq 0$. Also, the inequality (8) as well hold, from the search direction (6) multiply both sides by g_{k+1}^T and substitute (7) it will gives

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

It is known from the conjugacy conditions $g_{k+1}^T d_k = 0$. Hence for constant $C = 1$ condition (8) is true for $k + 1$. ■

B. Global Convergence Properties

The assumptions below applied to the objective function for the analysis of a global convergence properties of a general CG method.

Assumptions 1.1. (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 neighbourhood 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 1.1*. Suppose the assumptions 1.1 holds and consider any recurrence expression* (2) *and search direction* (6)*,* γ_k *is found in equation* (3)*. Then Zoutendijk condition below holds.*

$$\sum_{k=0}^{\infty} \frac{(g_K^T d_k)^2}{\|d_k\|^2} < \infty \tag{9}$$

Proof of this Lemma is given in Zoutendijk, (1970).

Theorem 1.2. *Suppose the assumptions 1.1 holds, for any* $\{x_k\}$, $\{d_k\}$ *be given as SpPRP CG method,* γ_k *are determined by equation* (3) *and* β_k *in equation* (5)*. Then*

$$\lim_{k\to\infty} \|g_k\| = 0 \tag{10}$$

Proof. From the search direction equation (6), square both sides we have,

$$(d_{k+1} + \varphi_k g_{k+1})^2 = \left(\beta_k^{PRP} d_k\right)^2$$
$$\|d_{k+1}\|^2 = \left(\beta_k^{PRP}\right)^2 \|d_k\|^2 - 2\varphi_k g_{k+1}^T d_{k+1} - \varphi_k^2 \|g_{k+1}\|^2 \tag{11}$$

Substituting equation (5) into (11), recall that $g_{k+1}^T d_{k+1} = -C\|g_{k+1}\|^2$ and rewrite (11) as

$$\|d_{k+1}\|^2 = \frac{\|g_{k+1}\|^4}{\|g_k\|^4} \|d_k\|^2 - \|g_{k+1}\|^2 \left(\varphi_k^2 - 2C\varphi_k\right) \tag{12}$$

Multiply both sides of equation (12) by $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$, we get

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \left(\left(2C\varphi_k - \varphi_k^2\right) + \frac{\|g_k\|^4}{\|g_{k-1}\|^4} \|d_k\|^2 \right) \tag{13}$$

Substituting (7) in (13) and note that from the conjugacy conditions $g_{k+1}^T d_k = 0$ we have,

$$\frac{\|d_{k+1}\|^2 \|g_{k+1}\|^2}{\|d_{k+1}\|^2} \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \tag{14}$$

Thus, from the Lemma 1.1 above. It implies that Theorem 1.2 does not hold true, then $\lim_{k\to\infty} \frac{\left(g_{k+1}^T d_{k+1}\right)^2}{\|d_{k+1}\|^2} = \infty$ and from equation (14) this is true $\infty \le \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So Theorem 1.2 is true for a sufficient large k. ■

Numerical Results

In this section, we test the Algorithm 1 above and compared its performance with the recent spectral PRP (RSPRP) developed by Wu, (2015) and classical PRP CG methods. The comparisons are arranged on CPU time and number of iterations in each test function. The stopping criteria used for both methods is $\varepsilon = 10^{-6}$ as suggested by Hillstrom, (1977) and $\|g_k\| < \varepsilon$. A standard test problem functions in Yakubu et al., (2018a) and Andrei (2008) were utilised and established a problems in Table 1 below with four different initial values using

MatlabR2015 subroutine programming by Intel® Core™ *i*5-3317U (1.7GHz with 4 GB (RAM)). In this work, failure is represented due to (i) memory requirement (ii) number of iterations exceed 1000 (iii) CPU time running in seconds reaches 1000 (iv) Line search method failed to find the step size or step length γ_k. The results shown in a Fig. 1 and Fig. 2 below are performance framework introduced by Dolan and More, (2002). Performances are measured based on CPU time running in seconds and number of iterations respectively. Therefore, the highest value of a percentage probability $P_s(t)$ indicated in the Fig. 1 and Fig. 2 below will be regarded as the best performing method and so also the method that reached the top foremost is consider most sophisticated upon all the CG methods.

Table1. Standard 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	10000	(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)
Freud. & Roth	2,4,10,100,1000,10000	(7,7), (11,11), (13,13), (25,25)

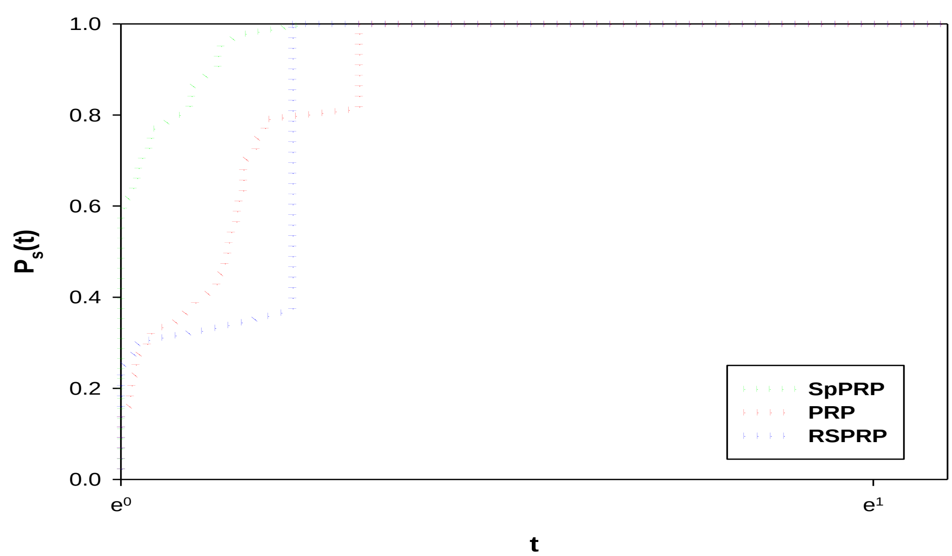

Figure 1. Performance framework based on the number of iterations

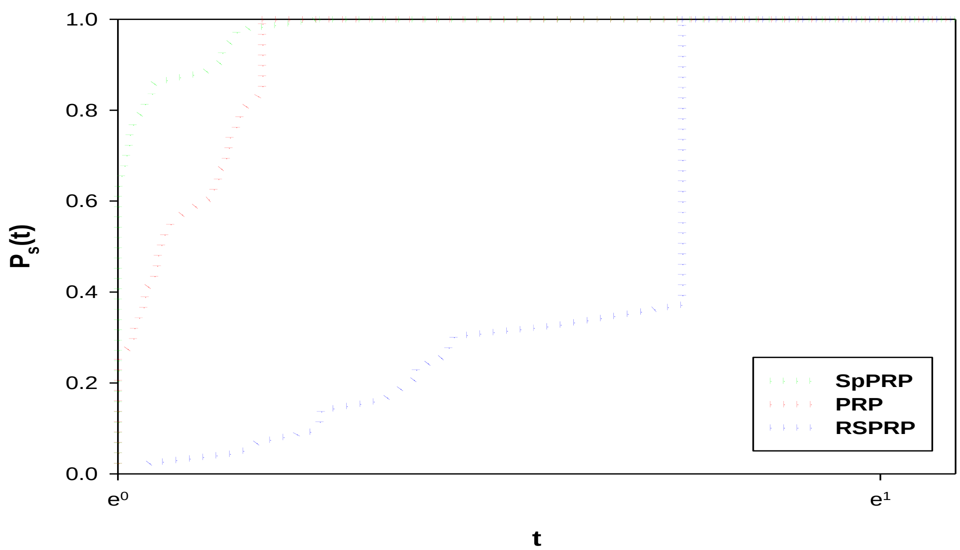

Figure 2. Performance framework based on CPU time.

It is certain that the spectral PRP (SpPRP) CG method has performed wonderfully since it solved the entire test problem functions used in table 1 above, the method also has complete potentials when compared with RSPRP and PRP CG methods. Fig.1 and Fig.2 shows that SpPRP, PRP and RSPRP CG methods successfully reached the solution point but nevertheless the SpPRP method has the highest probability of being the fastest optimum solver on approximately $P_s(t)$ = 60% and 65% respectively of the standard test problem functions used in this paper. Denoting the ‘‘successful’’ in the Table 2 and Table 3 meaning that SpPRP CG method has less number of iterations and less CPU time compared to RSPRP and PRP CG methods while denoting the ‘‘failed’’ meaning that SpPRP method produces more iterations and consumes more CPU time as compared to RSPRP and PRP CG methods. If the SpPRP CG method provides the same number of iterations or CPU time with RSPRP and PRP CG methods then it's equivalent. For both Table 2 and Table 3, the results are presented in percentages as follows. The highest value of percentage indicated that the method will be victorious over the other methods.

Table 2. Percentage Performance of SpPRP against RSPRP, PRP CG Methods on Number of Iterations

Method		Comparison of the Methods (%)	
		RSPRP	PRP
SpPRP	Successful	67.44	79.67
	Equivalent	16.28	9.30
	Failed	16.28	11.03

Therefore, SpPRP CG method in Table 2 above solves entirely the test problem functions with the highest percentage of success and its corresponding equivalent of 83.72% compared to RSPRP CG method, 88.97% compared with renowned PRP CG method in terms of a number of iterations.

Table 3. Percentage Performance of SpPRP against RSPRP, PRP CG Methods on CPU time

Method		Comparison of the Methods (%)	
		RSPRP	PRP
SpPRP	Successful	100	72.09
	Equivalent	0.0	0.0
	Failed	0.0	27.91

Consequently, the percentage of success and its corresponding equivalent of SpPRP method based on CPU time is 100% as compared with the RSPRP CG method and 72.09% compared to the famous PRP CG method. Hence, the overall average percentage of success for both a number of iterations and CPU time of SpPRP method is 79.08% and 86.20% for the average success and its corresponding equivalent. In the end, the new method has absolute potentials as compared with RSPRP and PRP CG methods. Certainly, the efficiency of the SpPRP CG method is highly applauded.

Conclusion

The proposed spectral PRP (SpPRP) CG method converges globally. The method overcomes the shortcomings of the recent spectral PRP and classical PRP CG method and the numerical results apparently show that the method performed admirably in terms of CPU time running in seconds and number of iterations reserved to achieve a sufficient decrease in the objective function.

Conflicts of Interest

There is no any conflict of interest regarding the publication of this article.

References

Raydan, M. (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.

Dolan, E. and More, J.J. (2002). Benchmarking optimization software with performance profile, Math. Prog., 91, 201-213.

Birgin, E.G. and Martinez, J. M. (2011). A spectral conjugate gradient method for unconstrained optimization, Appl. Math. Optim., 43(2), 117-128.

X. Wu. (2015). A new spectral Polak- Ribière -Polak conjugate gradient method, ScienceAsia, 41, 345-349.

Zoutendijk, G. “Nonlinear programming, computational methods, in J Abadie (ED)”, Integer and Nonlinear Programming, North-Holland, Amsterdam, (1970), 37-86.

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

Hu, C. and Wan, Z. (2013). An Extended Spectral Conjugate Gradient Method for unconstrained optimization problems, British Journal of Mathematics & Computer Science, 3, 86-98.

Yakubu, U.A. Mamat, M. Mohamad, A.M. Sukono, A.M. and Rivaie, M. (2018a). Secant free condition of a spectral PRP conjugate gradient method. International Journal of Engineering & Technology, 7(3.28), 325 – 328.

Yakubu, U.A. Mamat, M. Mohamad, A.M. Sukono, A.M. and Rivaie, M. (2018b). Modification on spectral conjugate gradient method for unconstrained optimization. International Journal of Engineering & Technology, 7(3.28), 307 – 311.

Yakubu, U.A. Mamat, M. Mohamad, A.M. Puspa, L.G. and Rivaie, M. (2018c).Secant free condition of a spectral Hestenes-Stiefel conjugate gradient method and its sufficient descent properties. International Journal of Engineering & Technology, 7(3.28), 312 – 315.

Powell, M.J.D. (1977). Restart procedures for the conjugate gradient method, Math. Program., 12, 241-254.

Barzilai, J. and Borwein, J.M. (1988). Two point step size gradient methods, IMA J Numer Anal., 8, 141-148.

Du, X. and Liu, J. (2011). Global convergence of a spectral HS conjugate gradient method, Procedia Engineering, 15, 1487 – 1492.

Zull, N. Rivaie, M. Mamat, M. Salleh, Z. Amani, Z. (2015). Global convergence of a Spectral conjugate gradient by using strong Wolfe line search, Appl. Math. Sci., 63, 3105-3117.

Andrei, N. (2008). An unconstrained optimization test functions collection, Adv. Modell. Optim., 10, 147-161.

Yakubu, U.A. Mamat, M. Mohamad, A.M. Rivaie, M. and B.Y. Rabi'u. (2018d). Secant free condition of a spectral WYL and its global convergence properties. Far East Journal of Mathematical science, 12, 1889 – 1902.

Yakubu, U.A. Mamat, M. Mohamad, A.M. Rivaie, M. and J. Sabi'u. (2018e). A recent modification on Dai-Liao conjugate gradient method for solving symmetric nonlinear equations. Far East Journal of Mathematical science, 12, 1961 – 1974.

Hager, W.W., and H. Zhang. (2006). A survey of nonlinear conjugate gradient methods, Pacific Journal of Optimization, 2(1), 35-58.

Yakubu, U.A., and U. Yuksel. (2015f). Necessary and sufficient conditions for first-order differential operators to be associated with disturbed Dirac operator in quaternionic analysis. Adv. appl. Clifford alg., 25(1), 1-12.

Abashar, A., 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, 65-68, 3307-3319.

Hillstrom, K.E. (1977). A simulation test approach to the evaluation of the nonlinear optimization algorithm. *Journal ACM Trans.* Mathematics Software, 3(4), 305-315.

Abba,V.M., M. Mamat, M.Y. Waziri, A.M. Mohamad, and U.A. Yakubu. (2018). A new conjugate gradient coefficient with exact line search for unconstrained optimization. Far East Journal of Mathematical science, 105(2), 193 – 206.

Umar, O.A., M. Mamat, M.Y. Waziri, I.M. Sulaiman. (2018). Solving dual fuzzy nonlinear equation via a modification on Shamanskii steps. Malaysian Journal of Computing and Applied Mathematics, 1(2), 1– 9.