An International Journal of Optimization and Control: Theories & Applications
ISSN: 2146-0957 eISSN: 2146-5703
Vol.10, No.2, pp.198-205 (2020)
https://doi.org/10.11121/ijocta.01.2020.00859

RESEARCH ARTICLE

Application of spectral conjugate gradient methods for solving unconstrained optimization problems

Sulaiman Mohammed Ibrahim[a*], Usman Abbas Yakubu[b], Mustafa Mamat[a]

[a] *Faculty of Informatics and Computing, University Sultan Zainal Abidin, Terengganu, Malaysia*
[b] *Department of Mathematics, Yusuf Maitama Sule University, Kano, Nigeria*
sulaimanib@unisza.edu.my, usman.abbas84@yahoo.com, must@unisza.edu.my

ARTICLE INFO

Article history:
Received: 3 September 2019
Accepted: 8 March 2020
Available Online: 7 June 2020

Keywords:
Sufficient descent property
Exact line search
Regression analysis
Spectral CG

AMS Classification 2010:
65K1023, 90C56

ABSTRACT

Conjugate gradient (CG) methods are among the most efficient numerical methods for solving unconstrained optimization problems. This is due to their simplicty and less computational cost in solving large-scale nonlinear problems. In this paper, we proposed some spectral CG methods using the classical CG search direction. The proposed methods are applied to real-life problems in regression analysis. Their convergence proof was estabilised under exact line search. Numerical results has shown that the proposed methods are efficient and promising.

1. Introduction

The spectral CG methods are among the most efficient variant of CG methods designed to solve large-scale problems. The methods possess the global convergent properties in addition to the sufficient descent condition. Moreso, the spectral CG methods are less expensive and requires less storage location. Some outstanding features of the spectral CG method are their simplicity in algebraic processes and development of computer codes [1]. Spectral CG method is formulated by combining the CG search direction and a scalar spectral parameter to form a new search direction. Birgin and Martinez [2], introduced a spectral CG method using standard secant equation [3].

Consider the following minimization problem.

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

where $f: R^n \to R$ is continuous and differentiable, g_k is the gradient of $f(x)$ and the vector $x_0 \in R^n$ is known as the initial point. The CG method are iterative scheme of the form

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

where the vector x_k is the current iterate, x_{k+1} is the new iteration point, and $\gamma_k > 0$ is the step-dimension obtained by the line search method defined as

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

also, d_k is the classical search direction given as

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

$g_k = \nabla f(x)$, is the gradient and the parameter $\beta_k \in R$ is the CG coefficient that characterizes different CG methods. Some known CG coefficients are the Polak-Ribiére-Polyak (PRP) and Wei-Yao-Liu (WYL) methods with formulas as follows.

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

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

where g_k and g_{k-1} are gradient vectors at points x_k, x_{k-1} respectively, and $\|.\|$ represent the Euclidian norm. The PRP method is regarded as the best CG method due to its rapid convergence. However, its convergence analysis for nonlinear function is

*Corresponding author

uncertain [12]. For further references on the CG and spectral CG methods, please refer to [4-14, 20, 26].

In this paper, the spectral PRP and spectral WYL CG methods are presented without the secant equation. Their performance is verified using the least square and trend line methods in regression analysis. The regression analysis is an important tool for the analysis of statistical data utilized in the field of economics, engineering, sciences and many more [15]. The analysis is use for forecasting and to comprehend the relation between dependent and independent variables in real life applications. The dependent variable is denoted by y and independent is denoted by x_j for $j = 1,2,3 \ldots \ldots, n, n > 0$, and e is an integer constant in the error term. The model is defined by

$$y = l(x_j + e), \qquad for\ x_j = x_1, x_2 \ldots . x_n \tag{7}$$

and generalized as follows

$$y = u_0 + u_1x_1 + u_2x_2 + \ldots \ldots . u_nx_n + e \tag{8}$$

where $u_0, u_1, u_2, \ldots \ldots \ldots, u_n$ are the parameters for the regression analysis. The values of the parameters are estimate by using the nonlinear least square method defined by

$$\min E(u) = \sum_{j=1}^{n} (y_i - u_0 + y_1x_{j1} + y_2x_{j2} + .. y_nx_{jn})^2 \tag{9}$$

where y_i is the estimated data of j^{th} response and $x_{j1}, x_{j2}, \ldots \ldots, x_{jn}$ are n data evaluation of the response variables [16]. The formula for predicting data in regression analysis is derive from calculating the relative error. However, the error is obtained by comparing the approximate value and actual value as described below

$$\text{Relative error} = \left| \frac{\text{Exact Value} - \text{Approximate Value}}{\text{Exact Value}} \right| \tag{10}$$

The least square determines the best approximation models by comparing the total least square errors. The error is defined as

$$E_j = (u_0 + u_1x) - y_j$$

The strategy of fitting the best line through the data would minimize the sum of the residual error squares for the data available. This problem is similar to the minimization problem in unconstrained optimization [17]. Thus, we employ the spectral PRP and WYL CG parameter to obtain the solution of the given unconstrained optimization problem.

2. Derivation of spectral CG methods

Spectral CG method was introduce by [2] with direction defined as $d_k = -\varphi_k\, \mathrm{g}_k + \beta_k s_{k-1}$, where $s_{k-1} = \gamma_{k-1}d_{k-1}$ and φ_k is a spectral scalar parameter. Motivated by the procedure of [5], we proposed the following search direction

$$d_k = \begin{cases} -\mathrm{g}_k, & if\ k = 0 \\ -\varphi_k \mathrm{g}_k + \beta_k^{PRP} d_{k-1}, & if\ k \geq 1 \end{cases} \tag{11}$$

$$d_k = \begin{cases} -\mathrm{g}_k, & if\ k = 0 \\ -\frac{1}{\emptyset_k} \mathrm{g}_k + \beta_k^{WYL} d_{k-1}, & if\ k \geq 1 \end{cases} \tag{12}$$

From(11), $d_k = -\varphi_k \mathrm{g}_k + \beta_k^{PRP} d_{k-1} \rightarrow d_k - \beta_k^{PRP} d_{k-1} = -\varphi_k \mathrm{g}_k$. Also, $d_k = -\mathrm{g}_k$, then substituting equation (5) we have,

$$\varphi_k = 1 - \frac{\mathrm{g}_k^T d_{k-1}}{\mathrm{g}_{k-1}^T d_{k-1}} \tag{13}$$

From equation (12), $d_k = -\frac{1}{\emptyset_k} \mathrm{g}_k + \beta_k^{WYL} d_{k-1}$ which is rewritten as $d_k - \beta_k^{WYL} d_{k-1} = -\frac{1}{\emptyset_k} \mathrm{g}_k$. This implies $\frac{1}{\emptyset_k} = \frac{d_k}{-\mathrm{g}_k} + \frac{\beta_k^{WYL} d_{k-1}}{\mathrm{g}_k}$. Substituting (6) in the equation, we have

$$\emptyset_k = \left(1 - \frac{2\mathrm{g}_k^T d_{k-1}}{\mathrm{g}_{k-1}^T d_{k-1}}\right)^{-1} \tag{14}$$

Recall that the orthogonality of gradients $\mathrm{g}_k^T \mathrm{g}_{k-1} = 0$ and thus, φ_k and $\emptyset_k$ are the new spectral parameters computed by exact line search procedure.

Algorithm 1.1 (Spectral CG method)

Step 1: Given a starting point $x_0 \in R^n$ set $k = 0$
Step 2: Compute β_k by (5) and (6)
Step 3: Compute d_k by (11) and (12). If $\|\mathrm{g}_k\| = 0$, then stop.
Step 4: Compute γ_k by (3).
Step 5: Update the new point by the recurrence expression (2).
Step 6: If $f(x_{k+1}) < f(x_k)$ and $\|\mathrm{g}_k\| < \varepsilon$ then stop, otherwise go to step 1 with $k = k + 1$.

3. The global convergence analysis of spectral CG methods

The Sufficient descent condition ensures that global convergence of iterative procedures or algorithm is achieved. Therefore, all CG methods must satisfy the following.

$$\mathrm{g}_k^T d_k \leq -C\|\mathrm{g}_k\|^2 \quad for\ k \geq 0\ and\ C > 0 \tag{15}$$

Theorem 1.1 *Suppose a CG method with search direction* (11) , (12) *and* β_k^{PRP}, β_k^{WYL} *given by equation* (5), (6), *then condition* (15) *holds for*

all $k \geq 0$.

Proof. With β_k^{PRP}, we proceed by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (15) satisfied as $k = 0$. Now we assume it is true for $k \geq 0$. Also, the inequality (15) as well hold.

From the search direction (11) multiply both sides by g_{k+1}^T and substitute parameter (13) 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 (15) is true for $k + 1$. ■

Proof. With β_k^{WYL}, also by induction, since $g_0^T d_0 = -\|g_0\|^2$, the condition (15) satisfied as $k = 0$. Now we assume it is true for $k \geq 0$.
Also, the inequality (15) hold true, from the search direction (12) multiply both sides of the equation by g_{k+1}^T and substitute (14) gives

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

Therefore, from the conjugacy conditions $g_{k+1}^T d_k = 0$. Hence for constant $C = 1$ condition (15) hold for $k + 1$. ■

The following assumptions are needed for the convergence analysis of the 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 neighborhood* N *of the level set* Ω *and* x_0 *is a starting point.*
(ii) $g(x)$ *is 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 Assumption 1.1 hold and consider any recurrence expression* (2) *with search direction* (11) and (12), γ_k *computed using* (3). *Then Zoutendijk condition (16) holds.*

$$\sum_{k=0}^{\infty} \frac{(g_K^T d_k)^2}{\|d_k\|^2} < \infty \tag{16}$$

Proof: The proof of this Lemma is given in [18].

Theorem 1.2 *Suppose Assumptions 1.1 hold, for any CG sequence* $\{x_k\}$, $\{d_k\}$ *be given as spectral PRP, spectral WYL CG methods,* γ_k *determined by equation* (3) *and* β_k *in equation* (5) *and* (6). *Then*

$$\lim_{k\to\infty} \|g_k\| = 0 \tag{17}$$

Proof. With β_k^{PRP}, from the search direction (11), square both sides of equation,

$$\begin{aligned}(d_{k+1} + \varphi_k g_{k+1})^2 &= (\beta_k^{PRP} d_k)^2 \\ \|d_{k+1}\|^2 &= (\beta_k^{PRP})^2\|d_k\|^2 - 2\varphi_k g_{k+1}^T d_{k+1} \\ &\quad - \varphi_k^2\|g_{k+1}\|^2 \end{aligned} \tag{18}$$

Substituting (5) into (18) and recall that $g_{k+1}^T d_{k+1} = -C\|g_{k+1}\|^2$, rewrite equation (18) as

$$\begin{aligned}\|d_{k+1}\|^2 &= \frac{\|g_{k+1}\|^4}{\|g_k\|^4}\|d_k\|^2 \\ &\quad -\|g_{k+1}\|^2(\varphi_k^2 - 2C\varphi_k) \end{aligned} \tag{19}$$

Multiply both sides of equation (19) by $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$, we get

$$\begin{aligned}\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} &= \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\Bigg((2C\varphi_k - \varphi_k^2) \\ &\quad + \frac{\|g_k\|^4}{\|g_{k-1}\|^4}\|d_k\|^2\Bigg) \end{aligned} \tag{20}$$

From the theorem 1.1 the value of the constant $C = 1$ therefore, substituting equation (13) in (20) and note that from the conjugacy conditions $g_{k+1}^T d_k = 0$ we have,

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} \leq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \tag{21}$$

Thus, from the Lemma 1.1 above. It implies that Theorem 1.2 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 (21) this is true $\infty \leq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So Theorem 1.2 is true for a sufficient large k. ■

Proof. With β_k^{WYL}, from the search direction equation (12), square both sides we have,

$$\begin{aligned}\left(d_{k+1} + \frac{1}{\emptyset_k} g_{k+1}\right)^2 &= (\beta_k^{WYL} d_k)^2 \\ \|d_{k+1}\|^2 &= (\beta_k^{WYL})^2\|d_k\|^2 - \frac{2}{\emptyset_k} g_{k+1}^T d_{k+1} \\ &\quad - \frac{1}{\emptyset^2}\|g_{k+1}\|^2 \end{aligned} \tag{22}$$

Substituting equation (6) into (22) and recall that $g_{k+1}^T d_{k+1} = -C\|g_{k+1}\|^2$, rewrite (22) as

$$\begin{aligned}\|d_{k+1}\|^2 &= \frac{4\|g_{k+1}\|^4}{\|g_k\|^4}\|d_k\|^2 + \frac{2C}{\emptyset_k}\|g_{k+1}\|^2 \\ &\quad - \frac{1}{\emptyset_k^2}\|g_{k+1}\|^2 \end{aligned} \tag{23}$$

Multiply both sides of (23) by $\frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2}$, we get

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}\left(\left(\frac{2C}{\emptyset_k} - \frac{1}{\emptyset_k^2}\right) + \frac{4\|g_{k+1}\|^2}{\|g_k\|^4}\|d_k\|^2\right) \quad (24)$$

From the theorem 1.1 the value of the constant $C = 1$ therefore, substituting (14) in (24) and note that from the conjugacy conditions $g_{k+1}^T d_k = 0$ we equally have,

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

$$\frac{\|d_{k+1}\|^2\|g_{k+1}\|^2}{\|d_{k+1}\|^2} \leq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2} \quad (25)$$

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 (25) this is true $\infty \leq \frac{\|g_{k+1}\|^4}{\|d_{k+1}\|^2}$. So, Theorem 1.2 is true for a sufficient large k. ■

4. Description of the real life application

In this section, the detailed description of the real-life problem considered are in Table 1. These problems were obtained from [19]. The approximate function for the nonlinear least square method is formed as follows

$$f(x) = -0.05690476x^2 + 0.68404762x + 0.13285714$$

Thus, the function $f(x)$ is use to approximate the value of y based on value of x, that is, the rate of drug abuse within the city from year 2009 to 2016. The least square method can easily be transformed into unconstrained minimization problems as follows

$$\min_{x\in R^n} f(x) = \sum_{j=1}^{n}\left(\left(u_0 + u_1x_j + u_2x_j^2\right) - y_j\right)^2 \quad (26)$$

The data set in Table 1 shows the rate of drug abuse among the youth with aged 18 to 25 in Kano city, Nigeria for the years 2009-2017. The statistical data was obtained yearly by the National Drug Law enforcement agency (NDLEA), Kano. From the Table 1, the x-variable represent the year of the operation while the y-variable represent the rate of drug abuse among the youth in the city. For the data fitting, only the data from 2009 to 2016 is been considered. The data for the year 2017 is reserved for the error analysis.

Table 1. Rate of Drug Abuse in Kano City for the Year 2009 to 2017 in Percentage

Number of Data (x)	Years	Rate of Drug Abuse (y)%
1	2009	0.78
2	2010	1.35
3	2011	1.59
4	2012	1.88
5	2013	1.95
6	2014	2.46
7	2015	2.26
8	2016	1.81
9	2017	1.83

Let the number of data x_j be the number of years and the value y_j be the rate of drug abuse in percentages. Then, the data from 2009 to 2016 are utilized to formulate the nonlinear quadratic model for the least square method and the corresponding test function of unconstrained optimization problem. From the above problem, observation reveals that the data x_j and the value of y_j have parabolic relations with the regression function defined by (26) and the regression parameters u_0, u_1 and u_2.

$$\min_{x\in R^2} \sum_{j=1}^{n} E_j^2 = \sum_{j=1}^{n}((u_0 + u_1x + u_2x^2) - y_j)^2 \quad (27)$$

Similar transformation of the above least squares problem using the data from Table 1 for nonlinear quadratic unconstrained minimization model is

$$f(u_0, u_1, u_2) = (8u_0 + 36u_1 + 204u_2 - 14.08)^2 \quad (28)$$

Equation (28) is similar to equation (27). Therefore, expanding (28) we have

$$f(u_0, u_1, u_2) = 64u_0^2 + 1296u_1^2 + 41616u_2^2 + 576u_0u_1 + 3264u_0u_2 + 14688u_1u_2 - 225.28u_0 - 1013.76u_1 - 5744.64u_2 + 198.2464 \quad (29)$$

However, the data for 2017 is excluded from the unconstrained optimization function so that it could be used to compute the relative errors of the predicted data. Therefore, the proposed spectral PRP and WYL CG methods are applied to solve the test function using exact line search technique. Table 2 and Table 3 shows the test results for the spectral PRP, spectral HS, spectral WYL and MSCG methods for some selected initial point.

Table 2. Numerical Results for SPRP, SWYL, MSCG and SHS Methods based on CPU Time.

Initial value	CPU Time			
	SPRP	SWYL	MSCG	SHS
(-5,-5,-5)	41.3119	49.5704	14.6685	0.00063
(-1,0,-1)	41.3443	44.2581	4.19763	0
(11,11,11)	97.317	103.905	5.58880	0
(-2,-2,-2)	41.3556	50.5668	4.55623	0

Table 3. Numerical Results for SPRP, SWYL, MSCG and SHS Methods based on Number of Iterations.

Initial value	Number of Iteration			
	SPRP	SWYL	MSCG	SHS
(-5,-5,-5)	10000	10000	2	0
(-1,0,-1)	10000	1000	3	NaN
(11,11,11)	1000	1000	2	NaN
(-2,-2,-2)	1000	1000	2	NaN

To avoid computing the values of u_0, u_1, u_2 using matrix inverse, we employ the Spectral PRP, Spectral WYL, SHS and MSCG using four initial points as presented in the Table 2 and Table 3. The iteration is terminated if the number of iterations exceed 10000 or if the method fails to solve a test problem and denoted the point of failure as 'NaN'. The approximation functions of the spectral CG methods is given in Table 4.

Table 4. Approximation Functions for Different Initial Point

Initial values	Methods	Approximate Function
(-5, -5, -5)	SPRP	$y = 0.5243x^2 + 31.9303x - 5$
	SWYL	$y = 0.5243x^2 + 31.9303x - 5$
	MSCG	NaN
	SHS	NaN
(-1, 0,-1)	SPRP	$y = 0.2142975x^2 + 6.6407718x - 1$
	SWYL	$y = 0.2142975x^2 + 6.6407718x - 1$
	MSCG	NaN
	SHS	NaN
(11, 11,11)	SPRP	$y = -0.7754x^2 - 69.2073x + 11$
	SWYL	$y = -0.7754x^2 - 69.2073x + 11$
	MSCG	NaN
	SHS	NaN
(-2, -2,-2)	SPRP	$y = 0.27574716x^2 + 12.96824x - 2$
	SWYL	$y = 0.27574716x^2 + 12.96824x - 2$
	MSCG	NaN
	SHS	NaN

4.1. Trend line method

The rate of drug abuse in Kano city, Nigeria is estimated using the least square method and the proposed spectral CG methods. The tread line is plotted based on the original data from Table 1 using Microsoft Excel software. The equation for the trend line is in the form of nonlinear quadratic equation. Based on the actual data, the index of drug abuse denoted by y is represented in the y-axis. The x-axis represent the year and denoted by x.

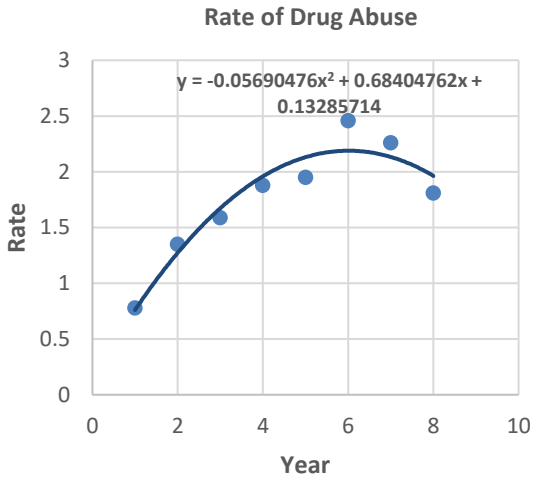

Figure 1. Nonlinear Quadratic Trend Line for Rate of Drug Abuse in Kano City

The functions of trend line and least square methods are compared with approximation functions obtained from the Spectral CG methods presented in Table 4.

5. Numerical result

Algorithm 1.1 have been tested on some benchmark problems and its performance are compared with RSPRP method [10], Wei-Yao-Liu (WYL) method [23], and Polak-Ribierre-Polyak (PRP) method [24] respectively. The comparisons are based on CPU time and number of iterations. The stopping criteria used is $\|g_k\| < \varepsilon$ where $\varepsilon = 10^{-6}$ as suggested by Hillstrom [21].

The set of standard test functions are considered from [1] and utilised with four different initial values. The codes are written on *MatlabR2015 subroutine* programming and run on an Intel® Core™ *i*5-3317U (1.7GHz with 4 GB (RAM)).

Table 5 and 6 presents the list of standard test problems with dimensions and initial points used to test the efficiency of the proposed spectral CG methods. The numerical performance of the proposed algorithms is presented in Figures 1.2 - 1.5 based on a number of iterations and CPU time. The maximum value of the percentage of probability $P_s(t)$ and the solver that reached the solution point foremost are regarded as the best performing CG methods for unconstrained optimization problems [25, 27].

Table 5. Standard Test Problems functions for Figure 2 and Figure 3

Functions	Dimensions	Initial Points
Trecanni	2	(3,3), (7,7), (11,11), (15,15)
Zettl	2	(10,10), (25,25), (100,100), (-100,-100)
Leon	2	(4,4), (-4,-4), (10,10), (-10,-10)
Quartic	4	(-3,-3), (5,-5), (15,15), (-20,-20)
Wood	4	(3,3), (-3,-3),(14,14),(-14,-14)
Hager	4	(2, 2), (10,10), (-10,-10),(15,15)
Fletcher	100	(13,13), (25,25),(40,40),(49,49)
Raydan	100	(2,2), (6,6), (8,8), (10,10)
Gen. Quartic	1000,10000, 50000,100000	(3,3), (5,5), (15,15), (-20,-20)
Freud. & Roth	4,10,100,500, 1000,10000, 50000,100000	(2,2), (5,5), (7,7), (-21,-21)
White and Holst	10,100,1000	(4,4), (-4,-4), (9,9), (-9,-9)
Shallow	100,1000, 10000	(100,100), (200,200), (300,300), (400,400)
Rosenbrock	2,4,10,100,1000, 10000,50000, 100000	(13,13), (25,25), (40,40), (49,49)

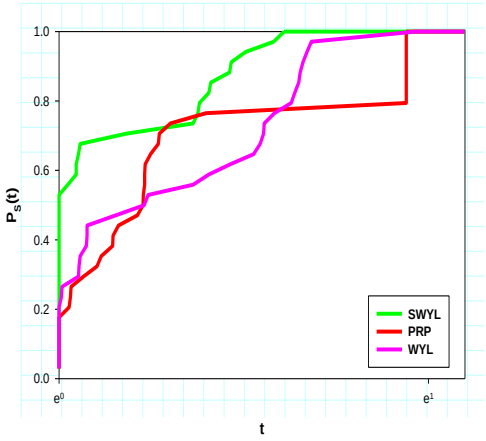

Figure 2. Performance outline based on the number of iterations

From the results obtained, the SPRP and SWYL CG methods are able to solve the standard benchmark problems as compared to the existing methods used in the analysis. Similarly, the data for 2017 are estimated using the nonlinear unconstrained optimization model in Table 4. and the relative error for each model using equation (10) is presented in Table 7.

Table 6. Standard Test Problems functions for Figure 4 and Figure 5

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)
Power	2,4,50,100	(5,5), (-5,-5), (100,100),(-100,-100)
Quadratic QF1	10,100,1000, 10000	(5,5), (-5,-5), (100,100),(-100,-100)
Ext. Quadratic Penalty QP1	10,100	(5,5), (-5,-5), (8,8), (-8,-8)
Ext. Quadratic Penalty QP2	10,100	(2,2), (6,6), (8,8), (10,10)
Himmelblau	10000	(2,2), (-2,-2), (25,25), (-25,-25)
Freud. & Roth	2,4,10,100,1000, 10000	(7,7), (11,11), (13,13), (25,25)
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)
Rosenbrock	2,4,10,100,1000, 10000	(5,5), (13,13), (20,20), (40,40)

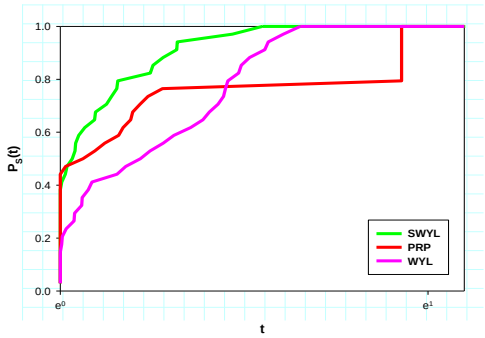

Figure 3. Performance outline based on CPU time

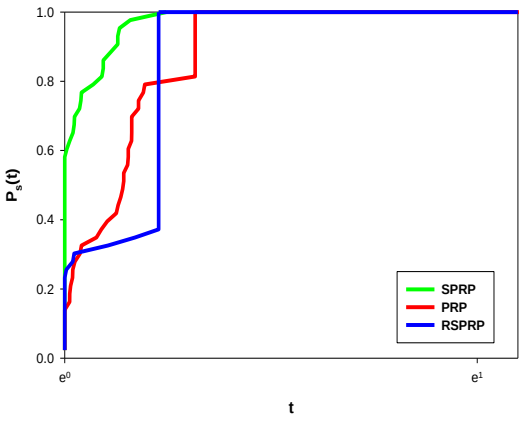

Figure 4. Performance outline based on the number of iterations

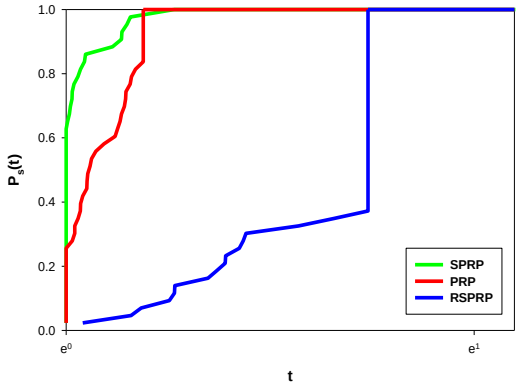

Figure 5. Performance outline based on CPU time

Table 7. Estimation Point and Relative Errors for 2017 Data

Models	Estimation Point	Relative Error
SPRP	1.130602014195	0.3821846916967
SWYL	1.130602014195	0.3821846916967
MSCG	NaN	NaN
SHS	NaN	NaN
Least Square	0.1686095216	0.907863649398907
Trend line	0.1686095216	0.907863649398907

The efficiency of each method is measure by equation (9). All the computations are carried out using Microsoft Excel 2016 and MATLAB 2015a *subroutine* programme. The model with the smallest relative error is considered the best model that estimate the rate of drug abuse in Kano city for the year 2017.

6. Conclusion

This paper focuses on the application of the spectral CG methods for unconstrained optimization. The proposed methods are compared with the classical WYL, PRP, least square and Trend line methods. The sum of relative error for the proposed spectral CG methods are computed based on four categories of initial values and three set of real numbers for nonlinear quadratic model. From the Table 7, the average relative error for the predicted data against the actual data 1.83 are calculated. The relative error for the data generated from nonlinear quadratic models of spectral PRP and spectral WYL methods are smaller compared to the least square and trend line models, which is around 0.3821846916967. The smallest relative error signifies the success of the spectral CG methods.

Acknowledgments

The authors would like to thank the anonymous reviewers for their valuable comments. This work was funded by Malaysian government under the grant number (FRGS/1/2017/STGO6/Unisza/01/1).

References

[1] Andrei, N. (2008). An unconstrained optimization test functions collection, *Advanced Modeling and Optimization*, 10, 147-161.

[2] Barzilai, J. & Borwein, J.M. (1988). Two-point step size gradient methods, *IMA J Numer Anal.* 8, 141-148.

[3] Birgin, E.G. & Martinez, J. M. (2011). A spectral conjugate gradient method for unconstrained optimization, *Applied Mathematics and Optimization*, 43 (2), 117-128.

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

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

[6] Yakubu, U.A. Mamat, M. Mohamad, A.M. Rivaie, M. & 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.

[7] Yakubu, U.A. Mamat, M. Mohamad, A.M. Rivaie, M. & 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.

[8] Zull, N. Rivaie, M. Mamat, M. Salleh, Z. & Amani, Z. (2015). Global convergence of a Spectral conjugate gradient by using strong Wolfe line search, *Applied Mathematical Sciences*, 9(63), 3105-3117.

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

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

[11] Sulaiman, I. M., Sukono, Sudradjat, S., and Mamat, M. (2019). New class of hybrid conjugate gradient coefficients with guaranteed descent and efficient line search. *IOP Conference Series: Materials Science and Engineering*, 621(012021), 1-7.

[12] *Mathematical Sciences* Issue 65-68, 3307-3319.

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

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

[15] Raydan, M. (1997). The Barzilai and J.M. Borwein gradient methods for the large scale unconstrained minimization in extreme problems, *SIAM Journal on Optimization*, 7 (1), 26-33.

[16] Batu, T., Dasgupta, S., Kumar, R., & Rubinfeld, R.

(2005). The complexity of approximating the entropy. *SIAM Journal on Computing*, 35(1): 132–150.

[17] Motulsky, H., & Christopoulos, A. (2004). *Fitting models to biological data using linear and nonlinear regression.* Oxford University Press, New York.

[18] Shirin, S., Mahmudul. H., & Laek, S.A. (2015). Age-Structured Population Projection of Bangladesh by Using a Partial Differential Model with Quadratic Polynomial Curve Fitting. *Journal of Applied Sciences*, 5: 542-551.

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

[20] Yakubu, U.A. (2019). *Enhancing spectral conjugate gradient method for solving unconstrained optimization problems*. PhD Thesis. University Sultan Zainal Abidin Kuala Terengganu, Malaysia.

[21] Yakubu, U.A. Mamat, M. Mohamad, A.M. Puspa, L.G. & 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.

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

[23] Dolan, E. & More, J.J. (2002). Benchmarking optimization software with performance profile, *Math. Prog*. Vol. 91, 201-213.

[24] Z. Wei, S. Yao, and L. Liu. (2006). The convergence properties of some new conjugate gradient methods, *Applied Mathematics and Computation*, Vol. 183(2), pp. 1341–1350.

[25] E. Polak & G. Ribiere. (1969). Note sur la convergence de directions conjug´ees, *Rev. Francaise Informat Recherche Opertionelle, 3e ann´ee*, Vol 16, pp. 35–43.

[26] Sulaiman, I. M. Mamat, M. Abashar, A. Rivaie M. (2015). The global convergence properties of an improved conjugate gradient method *Applied Mathematical Science*, 9(38), pp. 1857-1868.

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

[28] Sulaiman, I. M., Mamat, M., Abashar, A., Zabidin, S. (2015): A Modified Nonlinear Conjugate Gradient Method for Unconstrained Optimization, Applied Mathematical Sciences 9(54), pp. 2671-2682.

[29] Aini, N. Rivaie, M. Mamat, M. and Sulaiman, I. M. (2019). A Hybrid of Quasi-Newton Method with CG Method for Unconstrained Optimization. J. Phys.: Conf. Ser. 1366(012079), 1-10.

Ibrahim Sulaiman Mohammed is currently a post-*doctoral researcher at Faculty of informatics and computing, Universiti Sultan Zainal Abidin (UniSZA), Malaysia from 2019 till date. He obtained his PhD from UniSZA in 2018 specializing in the field of fuzzy systems. He has published research papers in various international journals and attended international conferences. His research interest includes Numerical research, Fuzzy nonlinear systems, unconstrained optimization.*

https://orcid.org/0000-0001-5246-6636

***Usman Abbas Yakubu** is a lecturer at Yusuf Maitama Sule University Kano, Nigeria. He obtained his PhD from Universiti Sultan Zainal Abidin (UniSZA), in 2019 specializing in the field of Computational Mathematics/ Unconstrained optimization. He has published research papers in various international journals and attended international conferences.*

https://orcid.org/0000-0002-6718-0523

***Mustafa Mamat** is currently a Professor of Computational and Applied Mathematics at Universiti Sultan Zainal Abidin, Malaysia since 2013. He obtained his PhD from UMT in 2007 specialization in optimization field. To date, he has successfully supervised more than 70 postgraduate students and published more than 260 research papers in various international journals. His research interest includes unconstrained optimization. Currently, he is the Editor in Chief for Malaysian Journal of Computing and Applied Mathematics and an editor for Indonesian Journal of Science and Technology.*

https://orcid.org/0000-0001-5358-3797