Application of spectral conjugate gradient methods for solving unconstrained optimization problems

Sulaiman Mohammed Ibrahim\textsuperscript{a}, Usman Abbas Yakubu\textsuperscript{b}, Mustafa Mamat\textsuperscript{a}

\textsuperscript{a} Faculty of Informatics and Computing, University Sultan Zainal Abidin, Terengganu, Malaysia
\textsuperscript{b} Department of Mathematics, Yusuf Maitama Sale University, Kano, Nigeria
sulaimanib@unisza.edu.my, usman.abbas84@yahoo.com, must@unisza.edu.my

\begin{abstract}
Conjugate gradient (CG) methods are among the most efficient numerical methods for solving unconstrained optimization problems. This is due to their simplicity 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 established under exact line search. Numerical results has shown that the proposed methods are efficient and promising.
\end{abstract}

\section{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.

\begin{equation}
\min f(x), \quad x \in \mathbb{R}^n
\end{equation}

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

\begin{equation}
x_{k+1} = x_k + y_k d_k, \quad k = 0,1,2,3,4,\ldots
\end{equation}

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

\begin{equation}
\gamma_k = \arg \min_{\gamma \geq 0} f(x_k + \gamma d_k)
\end{equation}

also, $d_k$ is the classical search direction given as

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

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

\begin{equation}
\beta_k^{PRP} = \frac{g_k^T(g_k - g_{k-1})}{\|g_k - g_{k-1}\|^2}
\end{equation}

\begin{equation}
\beta_k^{WYL} = \frac{g_k^T(g_k - \|g_k\| g_{k-1})}{\|g_k - g_{k-1}\|^2} \leq \frac{2\|g_k\|^2}{\|g_{k-1}\|^2}
\end{equation}

where $g_k$ and $g_{k-1}$ are gradient vectors at points $x_k$, $x_{k-1}$ respectively, and $\|\cdot\|$ 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
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 used 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 $j = 1, 2, 3, ..., n$. The analysis is used to obtain the relative error. The model is defined by

$$y = l(x_j + e), \quad \text{for } x_j = x_1, x_2, ..., x_n$$

and generalized as follows

$$y = u_0 + u_1x_1 + u_2x_2 + ... + u_nx_n + e$$

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

$$\min E(u) = \sum_{j=1}^{n} (y_j - u_0 + y_1x_1j + y_2x_2j + ... + y_nx_nj)^2$$

where $y_j$ is the estimated data of $j$th response and $x_1j, x_2j, ..., x_nj$ are data evaluation of the response variables [16]. The formula for predicting data in regression analysis is derived from calculating the relative error. However, the error is obtained by comparing the approximate value and actual value as described below

Relative error $= \frac{|\text{Exact Value} - \text{Approximate Value}|}{\text{Exact Value}}$ (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 $d_k = -\varphi_k g_k + \beta_k s_{k-1}$. where

$$s_{k-1} = y_k - d_{k-1}$$

Motivated by the procedure of [5], we proposed the following search direction

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

$$d_k = \begin{cases} \frac{g_k}{\varphi_k}, & \text{if } k = 0 \\ \frac{1}{\varphi_k} g_k + \beta_k^{WYL} d_{k-1}, & \text{if } k \geq 1 \end{cases}$$

From (11) and (12), $d_k = -\varphi_k g_k + \beta_k^{PRP} d_{k-1}$, $d_k = -\varphi_k g_k$. Also, $d_k = -g_k$, then substituting equation (5) we have,

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

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

$$\varphi_k = \left(1 - \frac{g_k^T d_{k-1}}{g_k^T d_{k-1}}\right)^{-1}$$

Recall that the orthogonality of gradients $g_k^T d_{k-1} = 0$ and thus, $\varphi_k$ and $\varphi_k$ 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 $\beta_k$ by (5) and (6)
Step 3: Compute $d_k$ by (11) and (12). If $\|g_k\| = 0$, then stop.
Step 4: Compute $y_k$ by (3).
Step 5: Update the new point by the recurrence expression (2).
Step 6: If $f(x_{k+1}) < f(x_k)$ and $\|g_k\| < \epsilon$ 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.

$$g_k^T d_k \leq -C\|g_k\|^2 \quad \text{for } k \geq 0 \text{ and } C > 0$$

Theorem 1.1 Suppose a CG method with search direction (11), (12) and $\beta_k^{PRP}$, $\beta_k^{WYL}$ given by equation (5), (6), then condition (15) holds for
all \( k \geq 0 \).

**Proof.** With \( \beta_k^{\text{PRP}} \), we proceed by induction, since \( g_0^Td_0 = -\|g_0\|^2 \), 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 \( g_k^{T}d_{k+1} \) and substitute parameter (13) gives

\[
g_k^{T}d_{k+1} = -\left(1 \right)g_k^{T}d_{k-1} + \beta_k^{\text{PRP}} g_k^{T}d_k
\]

It is known from the conjugacy conditions \( g_k^T d_k = 0 \). Hence for constant \( C = 1 \) condition (15) is true for \( k + 1 \). \( \blacksquare \)

**Proof.** With \( \beta_k^{\text{WYL}} \), 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 \( g_k^{T}d_{k+1} \) and substitute (14) gives

\[
g_k^{T}d_{k+1} = -\left(1 \right)g_k^{T}d_{k-1} + \beta_k^{\text{WYL}} g_k^{T}d_k
\]

Therefore, from the conjugacy conditions \( g_k^T d_k = 0 \). Hence for constant \( C = 1 \) condition (15) hold for \( k + 1 \). \( \blacksquare \)

The following assumptions are needed for the convergence analysis of the CG method.

**Assumptions 1.**

\( i \) A level set \( \Omega = \{ x \in \mathbb{R}^n \mid f(x) \leq f(x_0) \} \) is bounded, the function \( f \) is continuously differentiable in a neighborhood \( N \) of the level set \( \Omega \) 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), \( \gamma_k \) computed using (3). Then Zoutendijk condition (16) holds.

\[
\sum_{k=0}^{\infty} \frac{(g_k^T d_k)^2}{\|d_k\|^2} < \infty
\]  \hspace{1cm} (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, \( \gamma_k \) determined by equation (3) and \( \beta_k \) in equation (5) and (6). Then

\[
\lim_{k \to \infty} \|g_k\| = 0 \hspace{1cm} (17)
\]

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

\[
(d_{k+1}^T + \varphi_k g_k^T d_{k+1})^2 = (\beta_k^{\text{PRP}} d_k)^2
\]

\[
\|d_{k+1}\|^2 = (\beta_k^{\text{PRP}} d_k)^2(\|d_k\|^2 - 2\varphi_k g_k^T d_{k+1})
\]

Substituting (5) into (18) and recall that \( g_k^T d_k = -\|g_k\|^4, \) rewrite equation (18) as

\[
\|d_{k+1}\|^2 = \frac{\|g_k\|^4}{\|d_k\|^4}(\|d_k\|^2 - 2\varphi_k g_k^T d_{k+1})
\]

Multiply both sides of equation (19) by \( \frac{\|g_k\|^2}{\|d_k\|^2} \), we get

\[
\frac{\|d_{k+1}\|^2}{\|d_k\|^2} = \frac{\|g_k\|^4}{\|d_k\|^4}(2\varphi_k g_k^T d_{k+1} + \frac{\|g_k\|^2}{\|d_k\|^2} )
\]

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 \( g_k^T d_k = 0 \) we have,

\[
\frac{\|d_{k+1}\|^2}{\|d_k\|^2} \leq \frac{\|g_k\|^4}{\|d_k\|^4} \hspace{1cm} (21)
\]

Thus, from the Lemma 1.1 above. It implies that Theorem 1.2 does not hold true, then \( \lim_{k \to \infty} \frac{\|g_{k+1}\|^2}{\|d_{k+1}\|^2} = \infty \) and from equation (21) this is true \( \infty \leq \frac{\|g_k\|^4}{\|d_k\|^4} \). So Theorem 1.2 is true for a sufficient large \( k \). \( \blacksquare \)

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

\[
(d_{k+1}^T + \frac{1}{\varphi_k} g_k^T d_{k+1})^2 = (\beta_k^{\text{WYL}} d_k)^2
\]

\[
\|d_{k+1}\|^2 = (\beta_k^{\text{WYL}} d_k)^2\|d_k\|^2 - \frac{2}{\varphi_k} g_k^T d_{k+1}^T d_{k+1}
\]

\[
- \frac{1}{\varphi_k^2} \|g_k\|^4
\]

Substituting equation (6) into (22) and recall that \( g_k^T d_k = -\|g_k\|^4, \) rewrite (22) as

\[
\|d_{k+1}\|^2 = \frac{4\|g_{k+1}\|^4}{\|g_k\|^4}(\|d_k\|^2 - \frac{2C}{\varphi_k} \|g_k\|^2)
\]

\[
- \frac{1}{\varphi_k^2} \|g_k\|^4
\]

\[
\|d_{k+1}\|^2 = \frac{4\|g_{k+1}\|^4}{\|g_k\|^4}(\|d_k\|^2 - \frac{2C}{\varphi_k} \|g_k\|^2)
\]

\[
- \frac{1}{\varphi_k^2} \|g_k\|^4
\]

\[
\|d_{k+1}\|^2 = \frac{4\|g_{k+1}\|^4}{\|g_k\|^4}(\|d_k\|^2 - \frac{2C}{\varphi_k} \|g_k\|^2)
\]

\[
- \frac{1}{\varphi_k^2} \|g_k\|^4
\]
Multiply both sides of (23) by \( \frac{\| g_{k+1} \|^2}{\| d_{k+1} \|^2} \), we get

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

(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,

\[
\frac{\| d_{k+1} \|^2 \| g_{k+1} \|^2}{\| d_{k+1} \|^2} \leq \frac{\| g_{k+1} \|^4}{\| d_{k+1} \|^2} \]

Thus, from the Lemma 1.1 above. It implies that

Theorem 1.2 does not hold true, then \( \lim_{k \to \infty} \frac{\| g_{k+1} \|^2}{\| d_{k+1} \|^2} = \infty \) and from equation (25) this is true \( \infty \leq \frac{\| g_{k+1} \|^2}{\| d_{k+1} \|^2} \). 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

\[
f(x) = -0.05690476 x^2 + 0.68404762 x + 0.13285714\]

Thus, the function \( f(x) \) is used 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

\[
\min_{x \in \mathbb{R}^2} \sum_{j=1}^{n} (u_0 + u_1 x_j + u_2 x_j^2 - y_j)^2 \quad (26)
\]

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

<table>
<thead>
<tr>
<th>Number of Data (x)</th>
<th>Years</th>
<th>Rate of Drug Abuse (Y)%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2009</td>
<td>0.78</td>
</tr>
<tr>
<td>2</td>
<td>2010</td>
<td>1.35</td>
</tr>
<tr>
<td>3</td>
<td>2011</td>
<td>1.59</td>
</tr>
<tr>
<td>4</td>
<td>2012</td>
<td>1.88</td>
</tr>
<tr>
<td>5</td>
<td>2013</td>
<td>1.95</td>
</tr>
<tr>
<td>6</td>
<td>2014</td>
<td>2.46</td>
</tr>
<tr>
<td>7</td>
<td>2015</td>
<td>2.26</td>
</tr>
<tr>
<td>8</td>
<td>2016</td>
<td>1.81</td>
</tr>
<tr>
<td>9</td>
<td>2017</td>
<td>1.83</td>
</tr>
</tbody>
</table>

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 \( y_j \) have parabolic relations with the regression function defined by (26) and the regression parameters \( u_0, u_1 \text{ and } u_2 \).

\[
\min_{x \in \mathbb{R}^2} \sum_{j=1}^{n} E_j^2 = \sum_{j=1}^{n} ((u_0 + u_1 x_j + u_2 x_j^2) - y_j)^2 \quad (27)
\]

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

\[
f(u_0, u_1, u_2) = (8u_0 + 36u_1 + 204u_2 - 140.08)^2 \quad (28)
\]

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.
To avoid computing the values of \( u_0, u_1, u_2 \) using matrix inverse, we employ the Spectral PRP, Spectral WYL, SHS and MCG 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 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 2. Numerical Results for SPRP, SWYL, MCG and SHS Methods based on CPU Time.

<table>
<thead>
<tr>
<th>Initial value</th>
<th>SPRP</th>
<th>SWYL</th>
<th>MCG</th>
<th>SHS</th>
</tr>
</thead>
<tbody>
<tr>
<td>(-5, -5)</td>
<td>41.3119</td>
<td>49.5704</td>
<td>14.6685</td>
<td>0.00066</td>
</tr>
<tr>
<td>(-1, 0, -1)</td>
<td>41.3443</td>
<td>44.2581</td>
<td>4.19763</td>
<td>0.13285714</td>
</tr>
<tr>
<td>(11, 11, 11)</td>
<td>97.317</td>
<td>103.905</td>
<td>5.58880</td>
<td>0.13285714</td>
</tr>
<tr>
<td>(-2, -2, -2)</td>
<td>41.3556</td>
<td>50.5668</td>
<td>4.55623</td>
<td>0.13285714</td>
</tr>
</tbody>
</table>

**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-Ribiere-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™ i5-3317U (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 \( P(x_0) \) and the solver that reached the solution point foremost are regarded as the best performing CG methods for unconstrained optimization problems [25, 27].

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

<table>
<thead>
<tr>
<th>Functions</th>
<th>Dimensions</th>
<th>Initial Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trecanni</td>
<td>2</td>
<td>(3,3), (7,7), (11,11), (15,15)</td>
</tr>
<tr>
<td>Zelt</td>
<td>2</td>
<td>(10,10), (25,25), (100,100), (10,10)</td>
</tr>
<tr>
<td>Leon</td>
<td>2</td>
<td>(-100,-100), (4,4), (10,10), (-10,-10)</td>
</tr>
<tr>
<td>Quartic</td>
<td>4</td>
<td>(-3,-3), (5,-5), (15,15), (20,20)</td>
</tr>
<tr>
<td>Wood</td>
<td>4</td>
<td>(3,3), (3,-3), (14,14), (-14,-14)</td>
</tr>
<tr>
<td>Hager</td>
<td>4</td>
<td>(2,2), (10,10), (-10,-10)</td>
</tr>
<tr>
<td>Fletcher</td>
<td>100</td>
<td>(13,13), (25,25), (40,40), (49,49)</td>
</tr>
<tr>
<td>Raydan</td>
<td>100</td>
<td>(2,2), (5,5), (9,9), (-9,-9)</td>
</tr>
<tr>
<td>Gen. Quartic</td>
<td>1000,10000</td>
<td>(3,3), (5,5), (15,15), (-20,-20)</td>
</tr>
<tr>
<td>Freud &amp; Roth</td>
<td>4,10,100,500</td>
<td>(2,2), (5,5), (7,7), (-21,-21)</td>
</tr>
<tr>
<td>White and Holst</td>
<td>10,100,1000</td>
<td>(4,4), (4,-4), (9,9), (-9,9)</td>
</tr>
<tr>
<td>Shallow</td>
<td>100,1000,10000</td>
<td>(100,100), (200,200), (300,300), (400,400)</td>
</tr>
<tr>
<td>Rosenbrock</td>
<td>2,4,10,100,10000,100000,500000,100000</td>
<td>(13,13), (25,25), (40,40), (49,49)</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Functions</th>
<th>Dimensions</th>
<th>Initial Points</th>
</tr>
</thead>
<tbody>
<tr>
<td>Trecanni</td>
<td>2</td>
<td>(5,5), (8,8), (-11,-11), (-15,-15)</td>
</tr>
<tr>
<td>Leon</td>
<td>2</td>
<td>(4,4), (-4,4), (6,6), (-10,-10)</td>
</tr>
<tr>
<td>Extended Penalty</td>
<td>2,4,10,50</td>
<td>(2,2), (-2,-2), (5,5), (-5,-5)</td>
</tr>
<tr>
<td>Power</td>
<td>2,4,50,100</td>
<td>(5,5), (-5,5), (100,100), (-100,-100)</td>
</tr>
<tr>
<td>Quadratic QF1</td>
<td>10,100,100,10000</td>
<td>(5,5), (-5,5), (100,100), (-100,-100)</td>
</tr>
<tr>
<td>Ext. Quadratic</td>
<td>10,100</td>
<td>(5,5), (-5,5), (8,8), (-8,8)</td>
</tr>
<tr>
<td>Penalty QP1</td>
<td>10,100</td>
<td>(2,2), (6,6), (8,8), (10,10)</td>
</tr>
<tr>
<td>Ext. Quadratic</td>
<td>10,100</td>
<td>(2,2), (-2,-2), (25,25), (-25,-25)</td>
</tr>
<tr>
<td>Penalty QP2</td>
<td>10000</td>
<td>(7,7), (11,11), (13,13), (25,25)</td>
</tr>
<tr>
<td>Freud &amp; Roth</td>
<td>2,4,10,100,1000,10000</td>
<td>(2,2), (5,5), (9,9), (-9,-9)</td>
</tr>
<tr>
<td>White and Holst</td>
<td>2,4,10,100,1000,10000</td>
<td>(100,100), (200,200), (400,400)</td>
</tr>
<tr>
<td>Shallow</td>
<td>2,4,10,100,1000,10000</td>
<td>(5,5), (13,13), (20,20), (40,40)</td>
</tr>
</tbody>
</table>

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.
The efficiency of each method is measured 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 estimates 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 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/STG06/Unisza/01/1).

References

[16] Batu, T., Dasgupta, S., Kumar, R., & Rubinfeld, R.
Application of spectral conjugate gradient methods for solving unconstrained optimization problems


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 specializing 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

An International Journal of Optimization and Control: Theories & Applications (http://ijocta.balikesir.edu.tr)

This work is licensed under a Creative Commons Attribution 4.0 International License. The authors retain ownership of the copyright for their article, but they allow anyone to download, reuse, reprint, modify, distribute, and/or copy articles in IJOCTA, so long as the original authors and source are credited. To see the complete license contents, please visit http://creativecommons.org/licenses/by/4.0/.