The application of new conjugate gradient methods in estimating data

Syazni Shoid1*, Norrlaili Shapiee1, Norhaslinda Zull1, Nur Hamizah Abdul Ghani1, Nur Syarafina Mohamed1, Mohd Rivaie2, Mustafa Mamat1

1Faculty of Informatics and Computing, Universiti Sultan Zainal Abidin (Unisza), Terengganu, Malaysia
2Department of Computer Sciences and Mathematics, Universiti Teknologi MARA (UiTM) Terengganu, Malaysia
*Corresponding author E-mail: syaznishoid@yahoo.com

Abstract

Many researchers are intended to improve the conjugate gradient (CG) methods as well as their applications in real life. Besides, CG become more interesting and useful in many disciplines and has important role for solving large-scale optimization problems. In this paper, three types of new CG coefficients are presented with application in estimating data. Numerical experiments show that the proposed methods have succeeded in solving problems under strong Wolfe Powell line search conditions.

Keywords: Conjugate Gradient Coefficient, Inexact Line Search, Least Squares, Regression, Strong Wolfe Powell

1. Introduction

The conjugate gradient (CG) method become attentive because this method does not require matrix storage and their iteration cost is low compared to other methods. Consider the CG method as a solution to the numerical problem for the following constrained problem,

$$\min f(x) \text{ subject to } x \in X,$$  \hspace{1cm} (1)

where $f(x)$ is an objective function, the $x \in \mathbb{R}^n$ is a decision variable and $X \in \mathbb{R}^n$ is a constraint set. If $x = \mathbb{R}^n$, then optimization problem (1) can be express as an unconstrained optimization problem,

$$\min_{x \in \mathbb{R}^n} f(x).$$  \hspace{1cm} (2)

The CG problems is solved iteratively, $x_0 \in \mathbb{R}^n$ is the starting point, using a recurring formula,

$$x_{k+1} = x_k + \alpha_k d_k \hspace{1cm} k = 0, 1, 2, ..., $$  \hspace{1cm} (3)

where $x_k$ is the current iteration point and the $\alpha_k > 0$ is a step-size, obtained through some line search methods.

Among the line searches used in practice is an inexact line searches. One of the most common and popular of inexact line searches is Wolfe line search\(^{11}\). This line search introduced two conditions as follow,

$$f(x_k + \alpha_k d_k) \leq f(x_k) + \mu \alpha_k g_k^T d_k \\\ \\ \\ (4)$$

$$g_{k+1}^T d_k \geq \sigma g_k^T d_k \\\ \ \ \ (5)$$

where $0 < \mu < \sigma < 1$. By replacing (5) with condition (6), then this is called as strong Wolfe line search,

$$g_{k+1}^T d_k \leq -\sigma g_k^T d_k. \\ \ \ \ (6)$$

The CG methods also determines the search direction $d_k$ with

$$d_k = \begin{cases} -g_k & \text{if} \ k = 0 \\ -g_k + \beta_k d_{k-1} & \text{if} \ k \geq 1 \end{cases} \hspace{1cm} (7)$$

where $g_k = \nabla f(x_k)$, and $\beta_k \in R$ is a coefficients which determines the difference between the CG method. The examples of $\beta_k$ suggested by early researchers are Fletcher and Reeves (FR) method\(^1\), Polak Ribiere and Rolyak (PR) method\(^2\), Hestenes and Stiefel (HS) method\(^3\), Rivaie, Mustafa, Ismail and (RAMI) method\(^4\), Syazni, Rivaie, Mustafa and Ismail (SRMI) method\(^5\), Norrlaili, Rivaie, Mustafa and Ismail (NRMI) method\(^6\), Rivaie, Abashar, Mustafa and Ismail (RAMI) method\(^7\). Here are the formulas,

$$\beta_{FR} = \frac{\|g_k\|^2}{\|g_{k-1}\|^2} \quad \hspace{1cm} (8)$$

$$\beta_{PR} = \frac{g_k^T (g_k - g_{k-1})}{\|g_{k-1}\|^2} \quad \hspace{1cm} (9)$$

$$\beta_{HS} = \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}} \quad \hspace{1cm} (10)$$

$$\beta_{RMMIL} = \frac{g_k^T (g_k - g_{k-1})}{\|d_{k-1}\|^2} \quad \hspace{1cm} (11)$$

$$\beta_{SRMI} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k - d_{k-1})} \quad \hspace{1cm} (12)$$

$$\beta_{NRMI} = \frac{g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k - d_{k-1})} \quad \hspace{1cm} (13)$$

Copyright © 2018 Authors. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
\[ \beta_k^\text{RAMI} = \frac{g_{k+1}^T (g_{k+1} - g_k)}{d_{k+1}^T (d_{k+1} - g_k)} \]  
where \( g_{k+1} = \nabla f(x_{k+1}) \) and \( ||.|| \) is the Euclidian norm of vectors.

Regression is used to determine the relationship between two or more variables and use this relationship to make estimation. The linear regression model is defined by

\[
y = a_0 + a_1 x\]

where \( a_0 \) and \( a_1 \) are the regression parameters. The value of parameters is calculated by method of least squares that minimized the problem to solve the above regression model

\[
\min Q = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i)]^2
\]

Then, the least square problems is transform to an unconstrained optimization problem as

\[
\min f(a) = \sum_{i=1}^{n} [y_i - a(1, x_i)^T]^2
\]

The quadratic regression model is defined by

\[
y = a_0 + a_1 x + a_2 x^2 \]

where \( a_0, a_1 \) and \( a_2 \) are the regression parameters. Similar to the problem above, the least squares method can be used to solve the problem as

\[
\min Q = \sum_{i=1}^{n} [y_i - (a_0 + a_1 x_i + a_2 x_i^2)]^2
\]

Similar transformation of least squares problem into an unconstrained optimization problem yields the following as

\[
\min f(a) = \sum_{i=1}^{n} [y_i - a(1, x_i, x_i^2)^T]^2
\]

The structure of this paper is organized as follows. In the section 2, new types of CG coefficients are presented. In section 3, we presented the data analysis, while in Section 4, some numerical outcome and discussion corresponding to these new \( \beta_k \) is given. Lastly, our conclusion is presented in Section 5.

2. New types of conjugate gradient coefficients

We used three types of \( \beta_k \) formulas named as \( \beta_k^{\text{SYRM}} \), \( \beta_k^{\text{SRM}} \) and \( \beta_k^{\text{NRMI+}} \). The formulas are shown below.

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

The SYRM is classified as a modified CG method. The original idea is from RAMI in (14) where a new formula for numerator has been proposed and the original formula for denominator as the RAMI formula has been retained.

\[
\beta_k^{\text{SRM}} = \begin{cases} 
\frac{g_{k+1}^T (g_k - g_{k-1})}{d_{k-1}^T (d_{k-1} - g_k)}, & g_{k+1}^T (g_k - g_{k-1}) > 0 \\
\|g_k\|^2 / d_{k-1}^T (d_{k-1} - g_k), & \text{else}
\end{cases}
\]

\[
\beta_k^{\text{NRMI+}} = \max \left\{ 0, \frac{g_{k+1}^T (g_k - g_{k-1})}{d_{k-1}^T (g_k - d_{k-1})} \right\}
\]

The SRM is classified as hybrid CG method. For this method, we combine two methods: HS and SYRM method with certain conditions.

The last method is the NRMI+ method which is derived from NRMI in (13). We extend the original method by assuming it as a nonnegative value.

The following algorithm is a common algorithm of CG method used in this study.

**ALGORITHM 1:** The basic of Conjugate Gradient algorithm

**Step 1:** Given an initial point \( x_0 \) and set \( k = 0 \).

**Step 2:** Computing conjugate gradient coefficients, \( \beta_k \) based on (21) until (23).

**Step 3:** Computing search direction

\[
d_k = \begin{cases} 
-g_k, & k = 0 \\
-g_k + \beta_k d_{k-1}, & k \geq 1
\end{cases}
\]

If \( g_k = 0 \), terminate the execution of the algorithm.

**Step 4:** Computing step size \( \alpha_k \) by inexact line search rule

Solve, (4) and (6).

**Step 5:** Updating new point

\[
x_{k+1} = x_k + \alpha_k d_k
\]

**Step 6:** Convergent test and stopping criteria.

If \( f(x_{k+1}) < f(x_k) \) and \( \|d_k\| < \varepsilon \), then terminate.

Otherwise go to Step 1 with \( k = k + 1 \).

3. Data analysis

Table 1 below shows the actual data obtained from Traffic Statistic Branch, Bukit Aman PDRM. The data represent the number of road death from 2004 to 2014.

<table>
<thead>
<tr>
<th>Year</th>
<th>Number Of Road Death</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2004</td>
</tr>
<tr>
<td>2</td>
<td>2005</td>
</tr>
<tr>
<td>3</td>
<td>2006</td>
</tr>
<tr>
<td>4</td>
<td>2007</td>
</tr>
<tr>
<td>5</td>
<td>2008</td>
</tr>
<tr>
<td>6</td>
<td>2009</td>
</tr>
<tr>
<td>7</td>
<td>2010</td>
</tr>
<tr>
<td>8</td>
<td>2011</td>
</tr>
<tr>
<td>9</td>
<td>2012</td>
</tr>
<tr>
<td>10</td>
<td>2013</td>
</tr>
<tr>
<td>11</td>
<td>2014</td>
</tr>
</tbody>
</table>

However, for the purpose of computation, only data from 2004 to 2013 are considered while data for 2014 are reserved for error calculation. In this experiment, the function used to solve the problem above is considered linear and quadratic cases and solved by SYRM, SRM and NRMI methods. Furthermore, for the purpose of comparison, sum squares method is also applied to solve the problem. Then, the results are analyzed by finding the relative error of the CG and sum squares methods using the formula shown below.

\[
\text{Relative error} = \frac{\text{exact value} - \text{approximate value}}{\text{exact value}}
\]

Calculation of relative error is important to determine the best method to estimate the number of road death for 2014.
4. Results and discussion

The data is implemented in SYRM, SRM, NRMI+ and sum square methods. All calculations involve the use of Microsoft Excel and Matlab 12 subroutine program. The results of linear and quadratic regression model for each method are shown below in Table 2 and Table 3 respectively.

<table>
<thead>
<tr>
<th>Method</th>
<th>Result</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYRM</td>
<td>4.5246666666666665-0.165575757575757809x</td>
</tr>
<tr>
<td>SRM</td>
<td>4.5246666666666661-0.165575757575757497x</td>
</tr>
<tr>
<td>NRMI+</td>
<td>4.5246666666666661-0.165575757575757497x</td>
</tr>
<tr>
<td>Least Squares</td>
<td>4.5246666666666661-0.165575757575757497x</td>
</tr>
</tbody>
</table>

As we can see, all methods can solve the problems for linear and quadratic cases except SYRM, where it failed to solve quadratic case. The relative error for each method is compared by estimating the actual data and the approximation function. Table 4 and table 5, show the results of the relative error of each method for linear and quadratic cases respectively.

<table>
<thead>
<tr>
<th>Method</th>
<th>RESULT</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYRM</td>
<td>FAIL</td>
</tr>
<tr>
<td>SRM</td>
<td>4.659666663129-0.23075766305x+0.00613636734e2x</td>
</tr>
<tr>
<td>NRMI+</td>
<td>4.659666663121-0.23075766325x+0.00613636736e2x</td>
</tr>
<tr>
<td>Least Squares</td>
<td>4.659666666667-0.2307576756e+0.00613636636e2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYRM</td>
<td>0.01072100313498</td>
</tr>
<tr>
<td>SRM</td>
<td>0.01072100313479</td>
</tr>
<tr>
<td>NRMI+</td>
<td>0.01072100313479</td>
</tr>
<tr>
<td>Least Squares</td>
<td>0.01072100313480</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>RELATIVE ERROR</th>
</tr>
</thead>
<tbody>
<tr>
<td>SYRM</td>
<td>FAIL</td>
</tr>
<tr>
<td>SRM</td>
<td>0.01467085292636</td>
</tr>
<tr>
<td>NRMI+</td>
<td>0.01467085292636</td>
</tr>
<tr>
<td>Least Squares</td>
<td>0.01467085403949</td>
</tr>
</tbody>
</table>

As we can see, all methods can solve the problems for linear and quadratic cases except SYRM, where it failed to solve quadratic case. The relative error for each method is compared by estimating the actual data and the approximation function. Table 4 and table 5, show the results of the relative error of each method for linear and quadratic cases respectively.

Based on the above results, we can say that the linear case gives the smallest relative error value compared to the quadratic case. In the case of linear, we can see the SRM and NRMI+ method is comparable which gives the smallest relative error value compared to SYRM and least squares methods. Thus, SRM and NRMI+ method are the best method to estimate the number of road deaths in 2014. As a result when x=11, then y=2.70.

5. Conclusion

The data is applied to the CG methods which are SYRM, SRM, NRMI+ method and sum squares methods. Then, it is found that all methods can solve the problem for a linear case. Instead, for quadratic case only the SYRM method fails to solve the problem. As a conclusion, the number of road death for 2014 is estimated as 2.70.

Acknowledgement

We would like to thank The Ministry of Higher Education of Malaysia for their sponsorship MyPhD KPM.

References