International Journal of Computer Applications (0975 – 8887)
Volume 178 – No. 51, September 2019

An Implementation of Conjugate Gradient Methods for Estimating Polynomial Models

Osman O. O. Yousif
Department of Mathematics,
Faculty of Mathematical and Computer Sciences,
University of Gezira, Sudan

Adam Hussein
Department of Statistics and population studies,
Faculty of Economics and Social Sciences,
Red Sea University, Sudan

Abdelrhman Abashar
Department of Basic Science,
Faculty of Engineering,
Red Sea University, Sudan

ABSTRACT
Conjugate gradient (CG) methods are one of the most widely used methods for solving nonlinear unconstrained optimization problems, especially of large scale. That is, due to their simplicity and low memory requirement. To analyze the convergence properties of a CG method, it implemented into two line searches; exact and inexact. In this paper, given some data, some CG methods will be used to find a polynomial function that fitting the data. To show the efficiency, a comparison between CG methods and least square method will be done.

Keywords
Conjugate gradient methods; unconstrained optimization; least square; data fitting.

1. INTRODUCTION
To solve the unconstrained optimization problem

$$min\, f(x)\,,\quad x \in R^n, \tag{1.1}$$

where $f: R^n \to R$ is continuously differentiable and it's gradient at a point x_k is denoted by g_k or $\nabla f(x_k)$, the CG methods use the iterative formula

$$x_{k+1} = x_k + \alpha_k d_k, \tag{1.2}$$

where α_k is step size which can be obtained by the exact line search

$$f(x_k + \alpha_k d_k) = \min_{\alpha \geq 0} f(x_k + \alpha d_k) \tag{1.3}$$

which simply means the orthogonality condition

$$g_k^T d_{k-1} = 0,$$

is hold. Also, the step size can be computed by inexact line search methods such as the strong Wolfe

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

$$|g(x_k + \alpha_k d_k)^T d_k| \leq \sigma |g_k^T d_k|,$$

where $0 < \delta < \sigma < 1$. In (1.2), x_k is the current iterate and d_k is the search direction that defined by

$$d_k = \begin{cases} -g_k\,, & \text{if } k = 0 \\ -g_k + \beta_k d_{k-1}, & \text{if } k \geq 1, \end{cases} \tag{1.4}$$

where β_k is a parameter that whose different forms determine different conjugate gradient methods. Well known formulas for β_k are the Hestenes-Stiefel (HS) [6], Fletcher-Reeves (FR) [10], Polak-Ribière-Polyak (PRP) [2, 3], conjugate descent (CD) [9], Liu-Storey (LS) [11], and Dai-Yuan (DY) [12] formulas. For more formulas for the coefficient β_k see [4, 5, 9, 13, 14, 15, and 16].

In this paper, a new algorithm that uses CG methods to fit some data by a polynomial of degree n will be presented in the next section. In Section 3, a numerical experiment that shows the efficiency of the new algorithm will be done. A conclusion will be presented in Section 4.

2. THE NEW ALGORITHM
It is well known that the method of least squares is a mathematical procedure for finding the best fitting-curve to a given set of points by minimizing the sum of the squares [1]. It can be used to fit the data $(x_i, y_i),\ i = 1,2, \dots, n$ using a polynomial of degree n, that is

$$y = a_0 + a_1 x + \cdots a_n x^n.$$

When the data $(x_i, y_i),\ i = 1,2, \dots, n$ substituted in in the polynomial above, errors $\in_i$ occurs, that is,

$$y_i = a_0 + a_1 x_i + \cdots a_n x_i^n + \in_i\,, \quad i = 1,2, \dots, n.$$

Therefore, the sum of the square of errors is given by

$$E(a) = \sum_{i=1}^{n} (y_i - a_0 + a_1 x_i + \cdots a_n x_i^n)^2$$

$$= \sum_{i=1}^{n} (y_i - a^T x_i)^2 \tag{2.1}$$

where

$$a^T = [a_0 \quad a_1 \quad \dots \quad a_n] \text{ and } x_i = [1 \quad x_i \quad \dots \quad x_i^n]^T,$$

$$i = 1, 2, \dots, n.$$

Equation (2.1) is a continuously differentiable function in the variable a and its gradient is given by

$$\nabla E(a) = \left[\frac{\partial E}{\partial a_0} \quad \frac{\partial E}{\partial a_1} \quad \dots \quad \frac{\partial E}{\partial a_n}\right]^T \tag{2.2}$$

where

$$\frac{\partial E}{\partial a_j} = \sum_{i=1}^{n} 2x_i^j (y_i - a^T x_i)\,, \qquad j = 0, 1, 2, \dots, n.$$

To find approximation values of the coefficients a, the least square minimizes the sum of the square of errors, that is solve

$$min\, E(a) \tag{2.3}$$

The problem (2.3) is a nonlinear unconstrained optimization in the variable a which can be solved using CG methods.

The following algorithm uses the CG techniques to solve problem (2.3) provides the line search is exact.

Algorithm 2.1: An algorithm for solving problem (2.3)

Step 1: Choose $x_0 \in R^n, \varepsilon \geq 0$.

Step 2: Compute the gradient (2.2) at x_0

Step 3: Set $\mathrm{k} = 0$ and $d_{\mathrm{k}} = -g_{\mathrm{k}}$

Step 4: if $\|g_0\| \leq \varepsilon$, then stop.

Step 5: Compute α_{k} by exact line search (1.3).

Step 6: Set $a_{k+1} = a_k + \alpha_k d_k$

Step 7: Compute the gradient (2.2) at a_{k+1}

Step 8: if $\|g_{k+1}\| < \varepsilon$, then stop.

Step 9: Compute the parameter β_k and generates d_k by (1.4).

Step 10: Set *k=k+1* go to Step 5.

3. NUMERICAL EXPERIMENT

In this section, to show the efficiency of Algorithm 2.1, it applied on some data obtained from [7]. This data is listed in Table 1.

Table 1: Unemployment rate in Terengganu from 2000 to 2015

No.	Years	Unemployment Rate
1	2000	3.0
2	2001	2.7
3	2002	3.2
4	2003	2.9
5	2004	3.2
6	2005	3.1
7	2006	3.6
8	2007	2.6
9	2008	3.4
10	2009	3.8
11	2010	3.7
12	2011	3.2
13	2012	3.0
14	2013	3.4
15	2014	4.2
16	2015	4.0

The data in Table 1 can be plotted using MATLAB in Fig. 1.

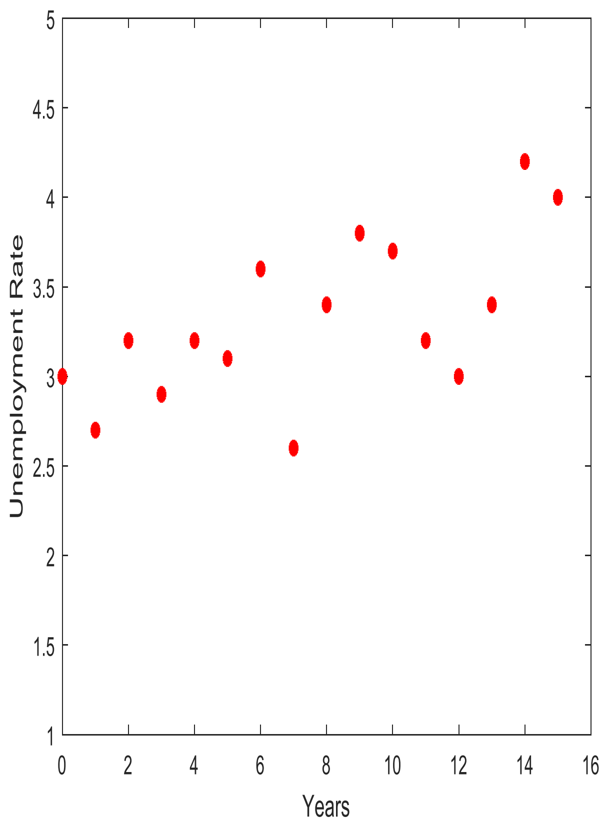

Fig. 1: Plotting of unemployment versus years

Using the least square calculations, it can be found that the linear, quadratic, and cubic polynomials that fitting the data in Table 1 are

$y = 2.8493 + 0.0618x,$

$$a = [2.8493 \quad 0.0618]^T, \tag{3.1}$$

$y = 2.9547 + .0166x + 0.0030x^2,$

$$a = [2.9547 \quad 0.0166 \quad 0.0030]^T \tag{3.2}$$

$y = 2.8233 + 0.1424x + -0.0186x^2 + 0.001x^3,$

$$a = [2.8233 \quad 0.1424 \quad -0.0186 \quad 0.001]^T \tag{3.3}$$

respectively.

To solve (2.3), a MATLAB coded program for Algorithm 2.1, with different values of the parameter β_k, hence different CG methods, namely, FR, PRP, and the new method [8] whose parameter given by

$$\beta_k^{OR} = \begin{cases} \dfrac{g_k^T g_k - |\, g_k^T d_{k-1}|}{d_{k-1}^T d_{k-1}}, & \text{if } g_k^T g_k \geq |\, g_k^T d_{k-1}| \\ 0, & \text{otherwise,} \end{cases}$$

and with different initial points, has been run with stopping criterion $\|g_k\| < 10^{-6}$. The numerical results are in Table 2.

Table 2: Implementation of FR, PRP, and OR methods for solving problem (2.3)

Poly.	Initial point	FR	PRP	OR
Linear	(0, 0)	(2.8493,0.0618)	(2.8493,0.0618)	(2.8493,0.0618)
	(10,10)	(2.8493,0.0618)	(2.8493,0.0618)	(2.8493,0.0618)
	(50,50)	(2.8493,0.0618)	(2.8493,0.0618)	(2.8493,0.0618)
	(-100,-100)	(2.8493,0.0618)	(2.8493,0.0618)	(2.8493,0.0618)
Quad.	(0,0,0)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)
	(20,20,20)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)
	(-10,-10,-10)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)
	(-40,-40,-40)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)	(2.9547,0.0166,0.0030)
Cubic	(0,0,0,0)	(2.8233,0.1424,-.0186,0.001)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-.0186,0.001)
	(5,5,5,5)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-.0186,0.001)
	(-1,-1,-1,-1)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-.0186,0.001)
	(1,10,1,10)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-0.0186,0.0010)	(2.8233,0.1424,-.0186,0.001)

Obviously, from Table 2, FR, PRP, and OR solve problem (2.3) and the results coincide with (3.1), (3.2), and (3.3). Therefore, they give a good fitting for the data in Table1 as in Figure 2.

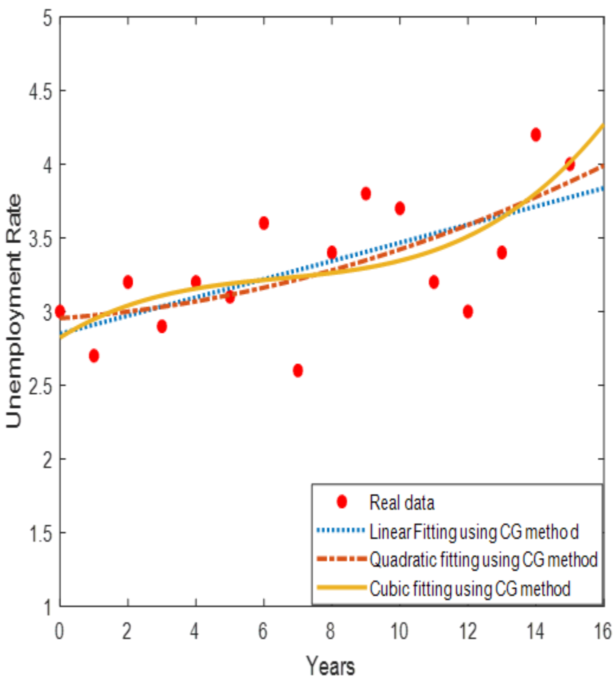

Figure 2: Curve fitting using CG methods

4. CONCLUSION

In this paper, based on some data, a conjugate gradient algorithm with exact line search has been applied to construct linear, quadratic, and cubic polynomials that fitting the data. To show the efficiency of conjugate gradient methods, a numerical experiment has been done. It has been shown that FR, PRP, and the new OR methods give the same results when they used for fitting data by polynomials.

5. REFERENCES

[1] Abdelwab Kharab and Ronald B. Guenther (2012), An introduction to numerical methods A matlab approach, CRC press.

[2] B. T. Polyak, The conjugate gradient method in extremem problems, *USSR Comp. Math. Math. Phys*., 9 (1969), pp. 94-112.

[3] E. Polak, Optimization: Algorithms and Consistent Approximations, vol. 124 of Applied Mathematical Sciences, *Springer*, New York, NY, USA, 1997.

[4] G. Yuan and X. Lu, "A modified PRP conjugate gradient method", *Annals of Operations Research*, vol. 166, pp. 73-90, 2009.

[5] L. Zhang, An improved Wei-Yao-Liu nonlinear conjugate gradient method for optimization computation, *Appl. Math. Comput*. 215 (2009), pp. 2269-2274.

[6] M. R. Hestenes and E. L. Stiefel, Methods of conjugate gradients for solving linear systems, *J. Research Nat. Bur. Standards*. 49 (1952), pp. 409-436.

[7] Nur Syarafina, M., Mustafa, M., Rivaie, M., Nur Hamizah, A., Norhaslinda Z., Syazni S., Estimating the unemployment rate using least square and conjugate gradient methods. International Journal of Engineering and Technology, 7(2.15), (2018) 94- 97.

[8] Osman, O.,Mamat, M., Abdelrhaman, A., Rivaie, M. (2014).The global convergence properties of a conjugate gradient methods. American Institute of Physics Conference Proceeding, 1602 (286), 286-295.

[9] R. Fletcher, *Practical Method of Optimization*, Vol. 1, Unconstrained Optimization, John Wiley & Sons, New York, 1987.

[10] R. Fletcher and C. Reeves, Function minimization by conjugate gradients, Comput. J., 7 (1964), pp. 149-154.

[11] Y. Liu, C. Storey, Efficient generalized conjugate gradient algorithms. Part 1: Theory, *J. Optimiz. Theory Appl*. 69 (1992), pp. 129-137.

[12] Y. H. Dai, Y. Yuan, A nonlinear conjugate gradient method with a strong global convergence property, *SIAM J. Optim*. 10 (2000), pp. 177-182.

[13] W. W. Hager and H. Zhang, "A new conjugate gradient method with guaranteed descent and an efficient line search," *SIAM Journal on Optimization*, vol. 16, no. 1, pp. 170-192, 2005.

[14] Y. Dai, "A nonmonotone conjugate gradient algorithm for unconstrained optimization," *Journal of Systems Science and Complexity*, vol. 15, no. 2, pp. 139-145, 2002.

[15] Z. Wei, G. Li, and L. Qi, "New nonlinear conjugate gradient formulas for large-scale unconstrained optimization problems," *Applied Mathematics and Computation*, vol. 179, no. 2, pp. 407-430, 2006.

[16] Z. Wei, S. Yao, and L. Liu, "The convergence properties of some new conjugate gradient methods," *Applied Mathematics and Computation*, vol. 183, no. 2, pp. 1341-1350, 2006.