КРАТКИЕ СООБЩЕНИЯ

MSC 46N10, 65K10, 90C06 DOI: 10.14529/mmp220309

A MODIFICATION OF DAI-YUAN'S CONJUGATE GRADIENT ALGORITHM FOR SOLVING UNCONSTRAINED OPTIMIZATION

***Y. Najm Huda*[1], *I. Ahmed Huda*[2]**
[1]University of Duhok, Duhok, Kurdistan Region, Iraq
[2]University of Mosul, Mosul, Iraq
E-mail: huda.najm@uod.ac, dr.hudaea@uomosul.edu.iq

The spectral conjugate gradient method is an essential generalization of the conjugate gradient method, and it is also one of the effective numerical methods to solve large scale unconstrained optimization problems. We propose a new spectral Dai-Yuan (SDY) conjugate gradient method to solve nonlinear unconstrained optimization problems. The proposed method's global convergence was achieved under appropriate conditions, performing numerical testing on 65 benchmark tests to determine the effectiveness of the proposed method in comparison to other methods like the AMDYN algorithm and some other existing ones like Dai-Yuan method.

Keywords: unconstrained optimization; conjugate gradient method; spectral conjugate gradient; sufficient descent; global convergence.

Introduction

The conjugate gradient method (CGM) is one of the most widely used approaches for solving nonlinear optimization problems. We consider the following unconstrained optimization problem:

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

where $f(x):\ R^n \to R$ is continuously differentiable and bounded below when solving equation (1), the iteration method is employed and is written as:

$$x_{r+1} = x_r + \alpha_r d_r, \tag{2}$$

where α_r is a line search-based optimal move [1]. One of the most inexact line searches is the Wolfe conditions to achieve convergence in the conjugate gradient methods, α_r is guaranteed to be present as:

$$f(x_{r+1}) - f(x_r) \le \delta\alpha_r \nabla f(x_r)^T d_r, \tag{3a}$$

$$\left|{g_{r+1}}^T d_r\right| \le -\sigma {g_r}^T d_r, \tag{3b}$$

where $0 < \delta < \sigma < 1$. More information may be found in [2]. The directions d_r are calculated using the following rule:

$$d_{r+1} = \left\{ \begin{array}{ccc} -g_{r+1} & ; & for \quad r=0 \\ -g_{r+1} + \beta_r d_r & ; & for \quad r \ge 1 \end{array} \right\}, \tag{4}$$

where $\beta_r \in R$ is a conjugacy coefficient. There are many Conjugate Gradient techniques based on different β_r process choices, for example [3–8]. The most significant distinction between the spectral gradient approach and the gradient conjugate method is calculating the search path. The spectral gradient method's search path is as follows:

$$d_{r+1} = -\theta_{r+1} g_{r+1} + \beta_r s_r, \tag{5}$$

where $s_r = \alpha_r d_r$ and θ_{r+1} is the spectral parameter. The spectral conjugate gradient is a powerful tool for solving large-scale unconstrained optimization problems. The spectral conjugate gradient algorithms are presented in many papers, for more details see [9–20].

1. Derivation of New Parameter

An essential idea is to equate the vector of the conjugate gradient method to the vector of quasi-Newton method such that $d_{r+1}^{QN} = d_{r+1}^{CG}$, then

$$-\rho_r H_{r+1} g_{r+1} = -\theta_{r+1} g_{r+1} + \beta_r^{DY} s_r, \tag{6}$$

where the matrix H_{r+1} is positive definite and symmetric, and we use the scale from [21] defined as follows:

$$\rho_r = \frac{y_r^T y_r}{s_r^T y_r}. \tag{7}$$

We use the DY conjugate gradient algorithm because it achieves convergence according to the Wolfe criterion and has the same property as FR [4] because it is similar in the numerator and the same property as HS [3], which is similar in the denominator. Therefore,

$$-\frac{y_r^T y_r}{s_r^T y_r} H_{r+1} g_{r+1} = -\theta_{r+1} g_{r+1} + \frac{\|g_{r+1}\|^2}{s_r^T y_r} s_r, \tag{8}$$

$$-\frac{y_r^T y_r}{s_r^T y_r} y_r^T H_{r+1} g_{r+1} = -\theta_{r+1} y_r^T g_{r+1} + \frac{\|g_{r+1}\|^2}{s_r^T y_r} y_r^T s_r,$$

since $H_{r+1} y_r = s_r \;\rightarrow\; y_r^T H_{r+1} = s_r^T$,

$$-\frac{y_r^T y_r}{s_r^T y_r} s_r^T g_{r+1} = -\theta_{r+1} y_r^T g_{r+1} + \|g_{r+1}\|^2.$$

Then the new spectral is known in the following form:

$$\theta_{r+1}^{SDY} = \frac{\|g_{r+1}\|^2}{y_r^T g_{r+1}} + \left(\frac{y_r^T y_r}{s_r^T y_r}\right) \frac{s_r^T\, g_{r+1}}{y_r^T g_{r+1}}, \tag{9}$$

then the proposed search direction is defined as

$$d_{r+1} = -\left[\frac{\|g_{r+1}\|^2}{y_r^T g_{r+1}} + \left(\frac{y_r^T y_r}{s_r^T y_r}\right) \frac{s_r^T\, g_{r+1}}{y_r^T g_{r+1}}\right] g_{r+1} + \frac{\|g_{r+1}\|^2}{s_r^T y_r} s_r. \tag{10}$$

Lemma 1. *Suppose that the sequences generate (2) and (10) computed by line search (3a) and (3b), then the sufficient descent condition holds.*

Proof. Define the spectral gradient direction as follows:

$$d_{r+1} = -\theta_{r+1}^{SDY} g_{r+1} + \beta_r^{DY} s_r,$$

where θ_{r+1}^{SDY} is defined in (9)

$$d_{r+1} = -\left[\frac{\|g_{r+1}\|^2}{y_r^T g_{r+1}} + \left(\frac{y_r^T y_r}{s_r^T y_r}\right)\frac{s_r^T\ \mathrm{g_{r+1}}}{y_r^T g_{r+1}}\right]\ \mathrm{g_{r+1}} + \frac{\|g_{r+1}\|^2}{s_r^T y_r} s_r. \tag{11}$$

Multiply equation (11) by $(\frac{g_{r+1}}{\|g_{r+1}\|^2})$. We get

$$\frac{d_{r+1}^T\ g_{r+1}}{\|g_{r+1}\|^2} = -\left[\frac{\|g_{r+1}\|^2}{y_r^T g_{r+1}} + \left(\frac{y_r^T y_r}{s_r^T y_r}\right)\frac{s_r^T\ g_{r+1}}{y_r^T g_{r+1}}\right]\frac{\|g_{r+1}\|^2}{\|g_{r+1}\|^2} + \frac{\|g_{r+1}\|^2}{s_r^T y_r}\frac{s_r^T\ g_{r+1}}{\|g_{r+1}\|^2}.$$

Since $y_r^T g_{r+1} \le \|y_r\|\,\|g_{r+1}\|$ and $s_r^T g_{r+1} \le s_r^T y_r$, then

$$\frac{d_{r+1}^T\ g_{r+1}}{\|g_{r+1}\|^2} \le -\frac{\|g_{r+1}\|}{\|y_r\|} - \frac{\|y_r\|}{\|g_{r+1}\|} + \frac{s_r^T\ g_{r+1}}{s_r^T y_r}.$$

From the strong Wolfe condition, $\mathrm{s_r^T g_{r+1}} \le -\sigma\ \mathrm{s_r^T g_r}$, and since we get $\mathrm{s_r^T g_r} \le \frac{-\mathrm{s_r^T y_r}}{\sigma+1}$

$$\frac{d_{r+1}^T\ \mathrm{g_{r+1}}}{\|g_{r+1}\|^2}\ \le -\frac{\|g_{r+1}\|}{\|y_r\|} - \frac{\|y_r\|}{\|g_{r+1}\|} - \sigma\frac{\left(\frac{-s_r^T y_r}{\sigma+1}\right)}{s_r^T y_r} \le -\frac{\|y_r\|}{\|g_{r+1}\|} + \frac{\sigma}{\sigma+1}.$$

Since $\mathrm{y_r = g_{r+1} - g_r}$ we get:

$$\frac{d_{r+1}^T\ \mathrm{g_{r+1}}}{\|g_{r+1}\|^2} \le -\frac{\|\mathrm{g_{r+1} - g_r}\|}{\|g_{r+1}\|} + \frac{\sigma}{\sigma+1} \le -\frac{\|g_{r+1}\|}{\|g_{r+1}\|} + \frac{\sigma}{\sigma+1}.$$

Let $\psi = \left(1 - \frac{\sigma}{\sigma+1}\right)$ then we get:

$$d_{r+1}^T\ \mathrm{g_{r+1}}\ \le -\psi\|g_{r+1}\|^2.$$

□

2. Convergence Analysis of New Method

Assumption I. The following are some fundamental assumptions in this paper. (a) The level set Ω is bounded. (b) The function is convex and continuously differentiable, its gradient $g(x)$ is Lipchitz continuous, i.e., there exists a constant $\overline{L} > 0$ such that $\|g(x) - g(y)| \le \overline{L}\|x - y\|$ for all $x, y \in \Omega$.

Lemma 2. *Suppose that Assumption **I** holds then for any CG-method, the direction* d_{r+1} *is a descent, and the step size* α_r *is achieved by (3a, 3b) if*

$$\sum_{r\ge 1}\frac{1}{\|d_{r+1}\|^2} = \infty.$$

Then

$$\lim_{r\to\infty}(\inf\|g_r\| = 0)\ . \tag{12}$$

Lemma 3. *Suppose that Assumption **I** holds, where* d_{r+1} *is defined by* (10), *then the method satisfies equation* (12).

Proof. Since β^{DY} is positive [7], then $0 < \beta < 1$,

$$\therefore \left|\beta^{DY}\right| \le \psi_1, \psi_1 > 0, \tag{13}$$

$$\left|\theta_{r+1}^{SDY}\right| = \left|\frac{\|g_{r+1}\|^2}{y_r^T g_{r+1}} + \left(\frac{\|y_r\|^2}{s_r^T y_r}\right)\frac{s_r^T\ \mathrm{g_{r+1}}}{y_r^T g_{r+1}}\right|,$$

$$\left|\theta_{r+1}^{SDY}\right| \le \left|\frac{\|g_{r+1}\|^2}{\|g_{r+1}\|^2 - \left|g_{r+1}^T g_k\right|}\right| + \left|\left(\frac{\|y_r\|^2}{s_r^T y_r}\right)\frac{-\sigma s_r^T\ g_r}{\|g_{r+1}\|^2 - \left|g_{r+1}^T g_r\right|}\right|,$$

from Powell restart $|g_{r+1}^T g_r| > 0,2\,\|g_{r+1}\|^2$,

$$\left|\theta_{r+1}^{SDY}\right| \le \frac{\|g_{r+1}\|^2}{0,8\ \|g_{r+1}\|^2} + \left|\frac{\|y_r\|^2}{s_r^T y_r}\frac{\sigma\left(\frac{s_r^T\ y_r}{\sigma+1}\right)}{0,8\ \|g_{r+1}\|^2}\right|,$$

$$\left|\theta_{r+1}^{SDY}\right| \le \psi_2, \quad \psi_2 > 0. \tag{14}$$

From equation (5) and using (13) and (14)

$$\|d_{r+1}\|^2 \le \psi_2 \|g_{r+1}\|^2 + \psi_1 \|s_r\| = \psi_3,$$

$$\sum_{k\ge1}^{\infty}\frac{1}{\|\ d_{r+1}\|} \ge \frac{1}{\psi_3}\sum_{r\ge1}^{\infty} 1 = \infty. \tag{15}$$

□

3. Results and Discussion

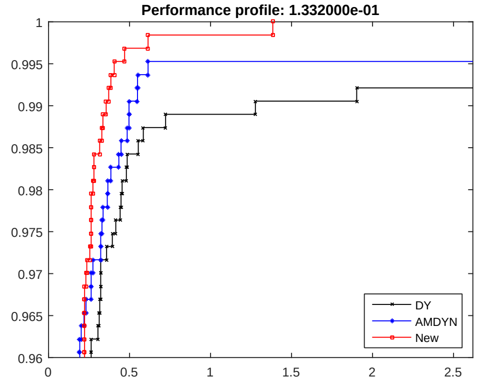
In this section, the new algorithm specified by equation (10) is compared to the classical DY algorithm [7] as well as AMDYN algorithm [11], which define θ_{r+1} as follows:

$$\theta_{r+1} = \frac{1}{y_r^T g_{r+1}}\left[\|g_{r+1}\|^2 - \frac{\|g_{r+1}\|^2 (s_r^T g_{r+1})}{y_r^T s_r} + s_r^T g_{r+1}\right].$$

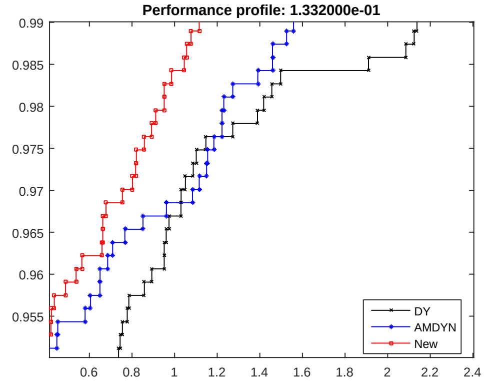
We chosen 65 unconstrained optimizations in the range $[n = 1000, 2000, ..., 10,000]$, broadly and based on generalized [22]. All algorithms used Wolfe condition $\sigma = 0,9, \delta = 0,0001$. The codes are adopted with double precision and using the Fortran language. All of these codes are authored by Andrei [11,23]. We applied the performance profile of Dolan and More [24] to demonstrate the algorithm's efficiency, where the upward curve indicates that the new algorithm (SDY) is better than the classic DY. Also, we compare the results with AMDYN based on the number of iterations (NOI), number of functions and gradient ratings (NOF & g) and CPU time (time). Representation of practical results is presented in Figure.

Conclusion

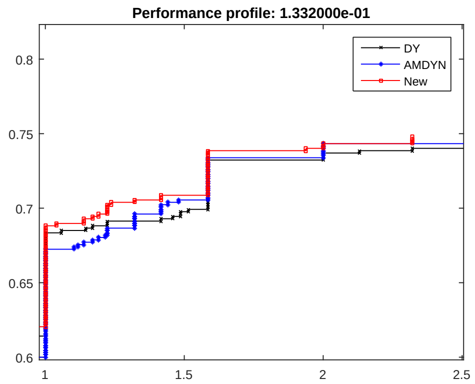
The aim of this work is to propose new computing schemes for the spectral parameter θ_{r+1}^{SDY}. This technique is such that the search direction always possesses a sufficient descent property. As a result, the presented new spectral conjugate gradient possesses global convergence under a robust Wolfe search. Using a collection of 65 optimization tests to solve problems, and numerically comparing the conduct of this new algorithm to that of some conjugated gradient algorithms DY and AMDYN, we show that the proposed algorithm outperforms previous conjugate gradient algorithms due to computational performance.



(a) NOI



(b) NOF & g



(c) Time of CPU

Performance profile

References

1. Alhawarat A., Salleh Z. Modification of Nonlinear Conjugate Gradient Method with Weak Wolfe-Powell Line Search. *Abstract and Applied Analysis*, 2017, vol. 2017, no. 2, pp. 1–6. DOI: 10.1155/2017/7238134

2. Nocedal J., Yuan Ya-Xiang. Analysis of a Self-Scaling Quasi-Newton Method. *Mathematical Programming*, 1993, vol. 61, no. 1, pp. 19–37. DOI: 10.1007/BF01582136

3. Hestenes M.R., Stiefel E. Methods of Conjugate Gradients for Solving Linear Systems. *Journal of Research of the National Bureau of Standards*, vol. 49, no. 6, pp. 409–435. DOI: 10.6028/jres.049.044

4. Fletcher R., Reeve C.M., Function Minimization by Conjugate Gradients. *The Computer Journal*, 1964, vol. 7, no. 2, pp. 149–154. DOI: 10.1093/comjnl/7.2.149

5. Polak E., Ribiere G. Note sur la convergence de méthodes de directions conjuguées. *Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique*, 1969, vol. 3, no. 1, pp. 35–43. (in French)

6. Polyak B.T. The Conjugate Gradient Method in Extremal Problems. *Computational Mathematics and Mathematical Physics*, 1969, vol. 9, no. 4, pp. 94–112. DOI: 10.1016/0041-5553(69)90035-4

7. Dai Yu-Hong, Yuan Yaxiang. A Nonlinear Conjugate Gradient Method with a Strong Global Convergence Property. *SIAM Journal on Optimization*, 1999, vol. 10, no. 1, pp. 177–182.

8. Dai Yu-Hong, Li-Zhi Liao. New Conjugacy Conditions and Related Nonlinear Conjugate Gradient Methods. *Applied Mathematics and Optimization*, 2001, vol. 43, no. 1, pp. 87–101. DOI: 10.1007/s002450010019
9. Hassan B., Abdullah Z., Jabbar H. A Descent Extension of the Dai-Yuan Conjugate Gradient Technique. *Indonesian Journal of Electrical Engineering and Computer Science*, 2019, vol. 16, no. 2, pp. 661–668. DOI: 10.11591/ijeecs.v16.i2.pp661-668
10. Jie Guo, Zhong Wan. A Modified Spectral PRP Conjugate Gradient Projection Method for Solving Large-Scale Monotone Equations and its Application in Compressed Sensing. *Mathematical Problems in Engineering*, 2019, vol. 2019, pp. 23–27. DOI: 10.1155/2019/5261830
11. Neculai A. New Accelerated Conjugate Gradient Algorithms as a Modification of Dai–Yuan's Computational Scheme for Unconstrained Optimization. *Journal of Computational and Applied Mathematics*, 2010, vol. 234, no. 12, pp. 3397–3410. DOI: 10.1016/j.cam.2010.05.002
12. Eman H., Rana A.Z., Abbas A.Y. New Investigation for the Liu-Story Scaled Conjugate Gradient Method for Nonlinear Optimization. *Journal of Mathematics*, 2020, vol. 2020, article ID: 3615208, 10 p. DOI: 10.1155/2020/3615208
13. Gaohang Yu, Lutai Guan, Wufan Chen. Spectral Conjugate Gradient Methods with Sufficient Descent Property for Large-Scale Unconstrained Optimization. *Optimization Methods and Software*, 2008, vol. 23, no. 2, pp. 275–293. DOI: 10.1080/10556780701661344
14. Ibrahim Sulaiman Mohammed, Yakubu Usman Abbas, Mamat Mustafa. Application of Spectral Conjugate Gradient Methods for Solving Unconstrained Optimization Problems. *An International Journal of Optimization and Control: Theories and Applications*, 2020, vol. 10, no. 2, pp. 198–205.
15. Wang Li, Cao Mingyuan, Xing Funa, Yang Yueting. The New Spectral Conjugate Gradient Method for Large-Scale Unconstrained Optimisation. *Journal of Inequalities and Applications*, 2020, vol. 2020, no. 1, pp. 1–11. DOI: 10.1186/s13660-020-02375-z
16. Jian Jinbao, Yang Lin, Jiang Xianzhen, Liu Pengjie, Liu Meixing. A Spectral Conjugate Gradient Method with Descent Property. *Mathematics*, 2020, vol. 8, no. 2, article ID: 280, 13 p. DOI: 10.3390/math8020280
17. Danhausa A.A., Odekunle R.M., Onanaye A.S. A Modified Spectral Conjugate Gradient Method for Solving Unconstrained Minimization Problems. *Journal of the Nigerian Mathematical Society*, 2020, vol. 39, no. 2, pp. 223–237.
18. Al-Arbo A., Rana Al-Kawaz. A Fast Spectral Conjugate Gradient Method for Solving Nonlinear Optimization Problems. *Indonesian Journal of Electrical Engineering and Computer Science*, 2021, vol. 21, no. 1, pp. 429–439. DOI: 10.11591/ijeecs.v21.i1.pp429-439
19. Hassan Basim, Jabbar Hawraz. A New Spectral on the Gradient Methods for Unconstrained Optimization Minimization Problem. *Journal of Zankoy Sulaimani*, 2020, vol. 22, no. 2, pp. 217–224. DOI: 10.17656/jzs.10822
20. Liu J.K., Feng Y.M., Zou L.M. A Spectral Conjugate Gradient Method for Solving Large-Scale Unconstrained Optimization. *Computers and Mathematics with Applications*, 2019, vol. 77, no. 3, pp. 731–739.
21. Al-Bayati A.Y. A New Family of Self-Scaling Variable Metric Algorithms for Unconstrained Optimization. *Journal of Education and Science*, 1991, vol. 12, pp. 25–54.
22. Neculai A. An Unconstrained Optimization Test Functions Collection. *Advanced Modeling and Optimization*, 2008, vol. 10, no. 1, pp. 147–161.
23. Neculai A. New Accelerated Conjugate Gradient Algorithms for Unconstrained Optimization. *Technical Report*, 2008, vol. 34, pp. 319–330.
24. Dolan E.D., More J.J. Benchmarking Optimization Software with Performance Profiles. *Mathematical Programming*, 2002, vol. 91, no. 2, pp. 201–213. DOI: 10.1007/s101070100263

Received November 25, 2021

УДК 519.6+517.972 **DOI: 10.14529/mmp220309**

МОДИФИКАЦИЯ АЛГОРИТМА СОПРЯЖЕННЫХ ГРАДИЕНТОВ ДАЙ-ЮАНЯ ДЛЯ РЕШЕНИЯ БЕЗУСЛОВНОЙ ОПТИМИЗАЦИИ

***Ю. Наджм Худа*[1], *И. Ахмед Худа*[2]**
[1]Университет Духока, г. Духок, Ирак
[2]Университет Мосула, г. Мосул, Ирак

Метод спектральных сопряженных градиентов является существенным обобщением метода сопряженных градиентов, а также одним из эффективных численных методов для решения крупномасштабных задач безусловной оптимизации. Мы предложили новый спектральный метод сопряженных градиентов Дай–Юаня для решения нелинейных задач безусловной оптимизации. Глобальная сходимость предложенного метода была достигнута при соответствующих условиях, проведены численные эксперименты на 65 эталонных тестах, показывающие эффективность предложенного метода по сравнению с другими методами, такими как алгоритм AMDYN и некоторыми другими существующими методами, такими как метод Дай–Юаня.

Ключевые слова: неограниченная оптимизация; метод сопряженных градиентов; спектральный сопряженный градиент; достаточный спуск; глобальная конвергенция.

Наджм Ю. Худа, кафедра математики, колледж наук, университет Духока (г. Духок, Ирак), huda.najm@uod.ac.

Ахмед И. Худа, PhD, кафедра операционных исследований и интеллектуальных техник, колледж компьютерных наук и математики, Университет Мосула (г. Мосул, Ирак), dr.hudaea@uomosul.edu.iq.

Поступила в редакцию 25 ноября 2021 г.