A New Projection Technique with Gradient Property to Solve Optimization Problems

Hussain Ali Mueen 1 and Mushtak A.K. Shiker 2

1, 2 Department of Mathematics, College of Education for Pure Sciences, University of Babylon, Hilla-Iraq.
E-mail: 1 hussain.moeen.pure234@student.uobabylon.edu.iq, 2 mmtmnmhh@yahoo.com, pure.mushtaq@uobabylon.edu.iq

Abstract. In this study, a new gradient projection technique has been proposed that consists of three boundaries with achieving the unadulterated descent feature. In this technique, we worked on combining the conjugate gradient algorithm with projection techniques to obtain a new algorithm for solving a wide range of unconstrained optimization problems. We have established global convergence with some hypotheses, and it has become clear to us through our results that the new formula is good and promised.

Keywords: Optimization Problem. Projection Method, Conjugate Gradient Approach.

1. Introduction

Numerous applications of the conjugate gradient projection technique exist, including mathematical programming and machine learning [1, 2]. It is a technique for solving unconstrained optimization problems that is both efficient and effective:

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

where $F: \mathbb{R}^n \to \mathbb{R}$ is a continuously differentiable real-valued function. The iterative gradient projection technique is an effective tool for solving problems (1) [3, 4]. To solve this problem using recurrence, the iterative sequence $x_k$ is generated from the initial point $x_0 \in \mathbb{R}^n$.

$$x_{k+1} = x_k + \alpha_k d_k, \quad k \in \mathbb{N},$$  \hspace{1cm} (2)

and $d_k$ is a search path with $\alpha_k > 0$, obtained via a line search

$$d_k = \begin{cases} -F_k & \text{if } k = 0 \\ -F_k + \beta_k d_{k-1} & \text{if } k \geq 1 \end{cases}$$  \hspace{1cm} (3)

$F_k = \nabla f(x_k)^T$ is an essential parameter. $\beta_k$ is a parameter [5, 6]. Therefore, famous instances many of $\beta_k$ are introduced by Fletcher-Reeves (FR), Polak- Ribere- Polyak (PRP) and others. As a result, Fletcher-Reeves (FR) suggested a well-known formula for update parameters $\beta_k$ [7], as follows:
\[ \beta_k^{FR} = \frac{F_k^T y_{k+1}}{d_k^T y_k} \]

Where \( y_{k-1} = F_k - F_{k-1} \). The techniques used when \( f \) is a solid convex quadratic function are equivalent to an exact line scan. When \( f \) is a non-quadratic functional, each type of action, as well as the convergence of the corresponding algorithms can be very different. The authors published many papers in various field of science such as transportation problems [8- 11], line search methods [12-15], optimization [16- 22], and reliability [23- 28].

2. The New Algorithm

We define the framework of the proposed approach in this work and show that it satisfies a nice descent property. For large- scale unconstrained optimization problems, we used the projection technique with conjugate gradient direction (CGD). The appropriate hyperplane \( H_k \) was constructed by Solodov and Svaiter [29], which strictly separates \( x_k \) from the solution set of problem:

\[ H_k = \{ x \in \mathbb{R}^n \mid F(z_k)^T (x - z_k) = 0 \}. \]

The other iteration point \( x_{k+1} \) is constructed by projecting \( x_k \) onto \( H_k \), which is \( x_{k+1} \) is defined by Solodov and Svaiter advice:

\[ x_{k+1} = x_k - \frac{F(z_k)^T (x_k - z_k)}{\|F(z_k)\|^2} F(z_k), \quad (4) \]

Our method's latest search path can be calculated by

\[ d_{k+1} = \begin{cases} -F_k + 2\beta_k^0 y_{k-1} + \zeta_k d_{k-1}, & k \geq 1 \\ -F_k, & k = 0 \end{cases} \]

Our method's latest search path can be calculated by

\[ v_k = x_k - x_{k-1}, \quad y_{k-1} = F_k - F_{k-1}, \quad \zeta_k = \frac{y_{k-1}}{\|y_{k-1}\|} \frac{v_k F_k + F_k^T y_{k-1}}{\|y_{k-1}\| d_{k-1}} \]

\[ \beta_k^u = \frac{d_{k-1}^T F_k + F_k^T d_{k-1}}{\|y_{k-1}\| d_{k-1}}. \quad (6) \]

An appealing property of the \( NO \) algorithm is that \( F_k^T d_k = -\|F_k\|^2 \) and it is independent of any line search technique, as shown in (4), (5), and (6). On the other hand, we have \( d_k \) defined by our method almost holds the descent state, i.e. (3), (5), and (6)

\[ (d_k, F_k) = - \|F_k\|^2 - \frac{d_{k-1}^T F_k + F_k^T d_{k-1}}{\|y_{k-1}\| d_{k-1}} F_k^T y_{k-1} + \frac{y_{k-1}^T F_k + F_k^T y_{k-1}}{\|y_{k-1}\| d_{k-1}} F_k^T d_{k-1} \]

\[ = - \|F_k\|^2 - \frac{d_{k-1}^T F_k y_{k-1} + \|F_k\|^2 d_{k-1}}{\|y_{k-1}\| d_{k-1}} F_k^T y_{k-1} + \frac{y_{k-1}^T F_k^T d_{k-1} + y_{k-1}^T F_k^T d_{k-1}}{\|y_{k-1}\| d_{k-1}} \]

\[ = - \|F_k\|^2. \]

As a result, the path \( d_k \) defined by our method typically meets the appropriate descent condition, i.e.

\[ F_k^T d_k \leq -c \|F_k\|^2, \quad \forall k \geq 0 \quad (7). \]

where \( c = 1 \).
The following are the steps of a new algorithm for the suggested approach.

3. Conjugate Gradient Projection Algorithm (NO).

1. The starting point \( x_0 \in \mathbb{R}^n \) is given, as well as parameters. \( \gamma, \tau > 0 \) and \( 0 < \rho, \sigma < 1 \)

   Set \( d_0 = -F_{x_0} \), \( k = 0 \);

   If \( \|F(x_k)\| \leq \varepsilon \) stop

2. Let \( \alpha_k = \max \{ \rho^i \alpha_{k-i}, i = 0,1,2, \ldots \} \), generated by a Wolfe condition

   \[
   f(x_k + \alpha_k d_k) \leq f(x_k) + \rho \alpha_k F_x^T d_k, \\
   F(x_k + \alpha_k d_k)^T d_k \leq \sigma F_x^T d_k. \tag{8}
   \]

3. Set \( z_k = x_k + \alpha_k d_k \) and calculate \( x_{k+1} \) by

   \[
   x_{k+1} = P_{\Omega} [x_k - \varphi F(z_k)],
   \]

   where \( \varphi = \frac{F(x_k)^T (x_k - x_{k-1})}{\|F(x_k)\|^2} \).

4. If \( \|F(x_k)\| \leq \varepsilon \) stop. Otherwise calculate \( d_{k+1} \) by (6).

5. Put \( k = k + 1 \), and return to (2)

Remark (2.1): According to the upper discussed, the search direction \( d_k \) holds (7) is a sufficient descent direction at iteration \( x_k \) in our approach. Furthermore, by (7)

\[
\|d_k\| \geq \|F_k\|. \quad k \in \mathbb{N} \tag{9}
\]

4. Global Convergence of New Method

Assumptions (3.1)

**A1:** The \( f(x) \) is bounded below and continuous on the level set \( \mathbb{R}^n \), as well as differentiable in the level set's neighborhood \( \mathbb{N} \),

\[
\sigma(x_0) = \{ x \in \mathbb{R}^n \mid f(x) \leq f(x_0) \}
\]

**A2:** The mapping \( F(x) \) is LC, i.e. \( \exists L > 0 \) s.t.

\[
\|F(x) - F(y)\| \leq L \|x - y\|, \quad x, y \in \mathbb{N} \tag{10}
\]

Assumptions (3.1) imply that \( \exists U > 0 \) such that

\[
\|F(x)\| \leq U \quad \forall x \in \sigma. \tag{11}
\]

To evaluate the global convergent, we work to determine the selection of \( \alpha_k \) for all \( d_k \) generated by the new technique holds (7). In the next lemma, we show a lower bound for \( \alpha_k \) using the Wolfe line search.

**Lemma (3.1):** Assume that algorithm (4) determines the sequence of search directions \( \{d_k\} \) and \( \{F_k\} \), which means that

\[
\alpha_k \geq \frac{\|F_k\|^2}{\|d_k\|^2}. \tag{12}
\]

**proof:** We have (8) by the specified condition.

\[
(\sigma - 1) F_x^T d_k \leq (F_k - F_{k-1})^T d_k \leq \|F_k - F_{k-1}\| \|d_k\| \leq \alpha_k L \|d_k\|^2,
\]
Then
\[ \alpha_k \geq \frac{\|F_k\|^2}{\|d_k\|^2}. \]

Where \( \alpha \geq \frac{1-\sigma}{\gamma} \), step 2 is represented by the Cauchy-Schwarz inequality, and step 3 is represented by the LC. Since \( d_k \) holds, \( \sigma > 1 \), and \( 12 \) is satisfied. The proof is finished.

We used the Zoutendijk condition [30] with the Wolfe condition, to access the global convergent of conjugate gradient projection process. The following lemma is so important to show that the line search of proposed algorithms holds the Zoutendijk condition.

**Lemma (3.2):** Assume that assumptions (10) are met, and the sequence of search directions \( d_k \) and \( F_k \) are determined using the suggested methodology, and the step length \( \alpha_k \) is measured using the Wolfe line search, then
\[ \sum_{k=0}^{\infty} \frac{\|F_k\|^4}{\|d_k\|^2} < +\infty. \quad (13) \]

**Proof** From (7) for any \( k \) we get:
\[ f(x_k) - f(x_{k+1}) \geq -\rho \alpha_k F_k^T d_k \geq \rho \alpha_k \|F_k\|^2 \geq \frac{\rho(1 - \sigma)\|F_k\|^4}{L\|d_k\|^2}. \]

which is responsible for the second and third inequalities (7) and (12). As a result of Assumptions (10)/A1, we get (13). The proof is finished. □

As a result, we know that the iteration of (CGM) can fail, implying that\( \|F_k\| \geq \mu. \forall k \geq 0 \), only if \( \|d_k\| \rightarrow \infty \) is sufficiently quick. To put it another way, the sequence \( \{\|F_k\|\} \) can only be bounded away from zero if \( \sum_{k=0}^{\infty} \frac{1}{\|d_k\|^2} < +\infty \).

The global convergence of the proposed solution is now investigated in the following theorem.

**Theorem (3.3):** Assume that the sequence \( F_k \) and \( d_k \) is generated by our new the suggested algorithm, then
\[ \lim_{k \to \infty} \inf \|F_k\| = 0 \quad (14) \]

**Proof:** Assume that (14) is not satisfy, then
\[ \|F_k\| \geq M. \forall k \geq 0, \quad (15) \]

where \( M > 0 \).

From (5), (6) and (7) we have
\[ \|d_k\| \leq \|F_k\| + \beta\|\gamma_{k-1}\| + \|\gamma_{k-1}\|d_{k-1}\| \]
\[ \|F_k\| \leq \|F_k\| + \frac{\|F_{k-1}\|\|F_k\|\|d_{k-1}\|}{\|y_{k-1}\|\|d_{k-1}\|} + \frac{\|F_{k-1}\|\|F_k\|\|y_{k-1}\|}{\|y_{k-1}\|\|d_{k-1}\|} \leq \|F_k\| + 2\frac{\|F_{k}\|\|y_{k-1}\|}{\|d_{k-1}\|} \leq \|F_k\| + \frac{2||F_k||\|y_{k-1}\|}{\|d_{k-1}\|} \]

\[ \|d_k\| \leq \|F_k\| + \frac{\|F_{k-1}\|\|F_k\|\|d_{k-1}\|}{\|y_{k-1}\|\|d_{k-1}\|} + \frac{\|F_{k-1}\|\|F_k\|\|y_{k-1}\|}{\|y_{k-1}\|\|d_{k-1}\|} \]
\[ \leq \|F_k\| + \frac{2\|F'_k\|}{\|F_{k-1}\|} [\|x_k - x_{k-1}\| + \|F_k - F_{k-1}\|] \]
\[ = \|F_k\| + \frac{2U(\|x_k - x_{k-1}\| + L)(x_k - x_{k-1})\|}{\|F_{k-1}\|} = U + \frac{2U(1 + L)}{M} \leq \theta \]

The second, third, and fourth inequalities, respectively, are caused by Cauchy-Schwarz inequality (9), (10), and Trigonometric inequality. (11), (15), and Assumption (10) are the last three inequalities.

In summary, our algorithm determines the sequence \(d_k\), which has a common upper bound such that
\[ \|d_k\| \leq \Gamma, \forall k \geq 0. \] (16)

And \(\Gamma = \theta\). Know by using (15) and (6), it’s show that
\[ \sum_{k=0}^{\infty} \frac{\|F_k\|^4}{\|d_k\|^2} \geq \sum_{k=0}^{\infty} \frac{M^4}{\Gamma^2} = +\infty. \]

This contradicts with (13), so, (14) holds. The proof is finished. □

5. Numerical Results

The suggested algorithm (NO) is compared with three other algorithms, NHS: Which produced by Mahdi M M and Shiker M A K (2020) [31]. HHT: This algorithm created by Mahdi M M and Shiker M A K (2020) [32]. HGL: Is produced by Dreeb N K, et al. (2021) [33].

To compare the performance of above algorithms in terms of time spent (CPU-Time), iterations number \((N_i)\), and functions Evaluating number \((N_f)\) that were achieved by finding the solution, all operating processes were suggested using MATLAB R 2018, running the computer with 4GH, CPU 2.30 – Windows 10 operating system. When \(\|F_k\| \leq 10^{-6}\) and the total number of iterations reach 20000, we used a criterion to stop the executions. The new algorithm's parameters were set as follows: \(\tau = 10^{-4}, \sigma = 0.35\) and \(\gamma = 0.1\). The problem that we took is
\[ f(x) = (x_2 - x_1)^2 + (1 - x_1)^2 \]

Table 4.1: Numerical results (CPU time)

| P. Dim. | S.I | NO | NHS | HHT | HGL |
|---------|-----|----|-----|-----|-----|
| 20000   | \(x_0\) | 1.321 | 2.812 | 3.640 | 2.250 |
| 20000   | \(x_1\) | 0.953 | 2.109 | 2.953 | 1.810 |
| 20000   | \(x_2\) | 0.734 | 1.546 | 1.875 | 1.200 |
| 20000   | \(x_3\) | 0.603 | 1.453 | 1.765 | 1.562 |
| 20000   | \(x_4\) | 1.721 | 0.828 | 1.093 | 0.250 |
| 20000   | \(x_5\) | 0.531 | 1 | 1.156 | 0.875 |
| 20000   | \(x_6\) | 2.453 | 1.062 | 1.500 | 0.750 |
| 20000   | \(x_7\) | 1.078 | 2.062 | 1.375 | 0.750 |
6 Table 4.2: Numerical results \((N_i, N_f)\)

The tables (4.1) and (4.2) indicate that the (CPU) time, the iterations number \((N_i)\) and the functions evaluation number \((N_f)\) of the new algorithm is less and better than the results of the other three algorithms, so, the new algorithm is effective and promising.

| P. | Dim. | S.P | NO | HNS | HHT | HGL |
|----|------|-----|----|-----|-----|-----|
| P  |      |     | 33 | 71  | 139 | 280 |
| 20000 |    |     | 31 | 74  | 159 | 310 |
| 20000 |    |     | 32 | 68  | 147 | 276 |
| 20000 |    |     | 37 | 68  | 137 | 276 |
| 20000 |    |     | 54 | 279 | 89  | 180 |
| 20000 |    |     | 38 | 58  | 176 | 234 |
| 20000 |    |     | 57 | 320 | 163 | 268 |
| 20000 |    |     | 47 | 430 | 133 | 268 |

6. Conclusion

In this work, we suggested a new projection technique to solve unconstrained optimization problems. The proposed method achieves global convergence with standard Wolfe line search. All of the results of the numerical experiment indicate that the new algorithm is better and more efficient than the other three algorithms that comparing with.

7. References

[1] Qi H D, Han J Y and Liu G H 1999 A modified Hestenes-Stiefel conjugate gradient algorithm, *Chinese Annals of Mathematics, Series A*, 17: 3, p. 177–184.
[2] Wasi H A and Shiker M A K 2020 A new hybrid CGM for unconstrained optimization problems, *J. Phys.: Conf. Ser.* 1664 012077.
[3] Shiker M A K and Sahib Z 2018 A modified trust-region method for solving unconstrained optimization, *Journal of Engineering and Applied Sciences*, 13: 22, p. 9667–9671.
[4] Mahdi M M et al. 2021 Solving systems of nonlinear monotone equations by using a new projection approach, *J. Phys.: Conf. Ser.* 1804 012107.
[5] Mahdi M M and Shiker M A K 2020 A New Class of Three-Term Double Projection Approach for Solving Nonlinear Monotone Equations, *J. Phys.: Conf. Ser.* 1664 012147.
[6] Mahdi M M and Shiker M A K 2020 Three terms of derivative free projection technique for solving nonlinear monotone equations, *J. Phys.: Conf. Ser.* 1591 012032.
[7] Shiker, M. A. K. and Z. Sahib, (2018), A modified technique for solving unconstrained optimization, *J. Eng. Applied Sci*. 13, p. 9667–9671.
[8] Zabiba M S M, Al-Dallal H A H, Hashim, Hashim K H, Mahdi M M and Shiker M A K 2021 A new technique to solve the maximization of the transportation problems. “in press”, *ICCEPS* - April.
[9] Hussein H A and Shiker M A K 2020 Two New Effective Methods to Find the Optimal Solution for the Assignment Problems, *Journal of Advanced Research in Dynamical and Control Systems*, 12: 7, p. 49–54.
[10] Hussein H A and Shiker M A K 2020 A modification to Vogel’s approximation method to solve transportation problems, *J. Phys.: Conf. Ser.* 1591 012029.
[11] Hussein H A, Shiker M A K and Zabiba M S M 2020 A new revised efficient of VAM to find the initial solution for the transportation problem, *J. Phys.: Conf. Ser.* 1591 012032.
[12] Hashim K H and Shiker M A K 2021 Using a new line search method with gradient direction to solve nonlinear systems of equations, *J. Phys.: Conf. Ser.* 1804 012106.
[13] Hashim K H, et al. 2021 Solving the Nonlinear Monotone Equations by Using a New Line Search Technique, *J. Phys.: Conf. Ser.* 1818 012099.
[14] Hashim L H, et al. 2021 An application comparison of two negative binomial models on rainfall count data, J. Phys.: Conf. Ser. 1818 012100.
[15] Hashim L H, et al. 2021 An application comparison of two Poisson models on zero count data, J. Phys.: Conf. Ser. 1818 012165.
[16] Dwail H H et al. 2021 A new modified TR algorithm with adaptive radius to solve a nonlinear systems of equations, J. Phys.: Conf. Ser. 1804 012108.
[17] Dwail H H and Shiker M A K 2020 Reducing the time that TRM requires to solve systems of nonlinear equations, IOP Conf. Ser.: Mater. Sci. Eng. 928 042043.
[18] Dwail H H and Shiker M A K 2020 Using a trust region method with nonmonotone technique to solve unrestrictes optimization problem, J. Phys.: Conf. Ser. 1664 012128.
[19] Dwail H H and Shiker M A K 2021 Using trust region method with BFGS technique for solving nonlinear systems of equations, J. Phys.: Conf. Ser. 1818 012022.
[20] Wasi H A and Shiker M A K 2021 Nonlinear conjugate gradient method with modified Armijo condition to solve unconstrained optimization, J. Phys.: Conf. Ser. 1818 012021.
[21] Wasi H A and Shiker M A K 2020 Using of generalized baye’s theorem to evaluate the reliability of aircraft systems, Journal of Engineering and Applied Sciences, (Special Issue13), 10797–10801.
[22] Abdulla g and Hassan Z A H 2020 Using of particle swarm optimization (PSO) to addressed reliability allocation of complex network, J. Phys.: Conf. Ser. 1664 012125.
[23] Abdulla g and Hassan Z A H 2020 Using of Genetic Algorithm to Evaluate Reliability Allocation and Optimization of Complex Network, IOP Conf. Ser.: Mater. Sci. Eng. 928 0420333.
[24] Solodov, M. V. and Svaiter, B. F. (1999), "A globally convergent inexact Newton method for systems of monotone equations", in: M. Fukushima, L. Qi (Eds.), Reformulation: Nonsmooth, Piecewise Smooth, Semismooth and Smoothing Methods, Kluwer Academic Publishers, p. 355-369.
[25] Zoutendijk G 1970 Nonlinear programming, computational methods. Integer and nonlinear programming, 37, p 86–96.
[26] Mahdi M M and Shiker M A K 2020 Three-term of new conjugate gradient projection approach under Wolfe condition to solve unconstrained optimization problems, Journal of Advanced Research in Dynamical and Control Systems. 12: 7, p 788–795.
[27] Mahdi M M and Shiker M A K 2020 A new projection technique for developing a Liu-Storey method to solve nonlinear systems of monotone equations, J. Phys.: Conf. Ser. 1591 012030.
[28] Dreeb N K, et al. 2021, Using a new projection approach to find the optimal solution for nonlinear systems of monotone equation, J. Phys.: Conf. Ser. 1818 012101.