A new hybrid CGM for unconstrained optimization problems

H A Wasi and Mushtak A.K. Shiker
1 Mathematics Department, College of Education for Pure Sciences, University of Babylon

E-mail: 1mathman1972@gmail.com, 2mmtmnh@yahoo.com

2 Corresponding Author

Abstract: In this work, a hybrid CGM (conjugate gradient method) has been suggested to solve the unconstrained optimization problems by combining a (Polak–Ribiére–Polyak) method with (Fletcher-Reeves) method. The suggested method has the sufficient descent property under the suggestion of a suitable line search and appropriate conditions. The global convergence is constructing for this method. The numerical results display that this method is better than the other method comparing with.

Keywords: Conjugate gradient method, hybrid method, descent property and global convergence.

1. Introduction

There are many methods to solve large scale unconstrained optimization problems in the form

$$\min f(x)_{x \in \mathbb{R}^n},$$

Where $f: \mathbb{R} \rightarrow \mathbb{R}^n$ is differentiable and smooth. The Conjugate gradient method is one of the important methods to solve (1.1) by using the iterative form

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

Where $x_k$ is the $k_{th}$ iterative point and $d_k$ is the search direction and $\alpha_k > 0$ is the step length. $d_k$ is determined by:

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

(1.2)

(1.3)

Numerical implementation of the Fletcher-Reeves method is basically affected by jamming, i.e., the method can be taking a lot of short steps without making advanced to the minimum. To obtain good computational results and to keep of the good feature of strong global convergence, we are trying to
use the hybridization principle to the Fletcher Reeves method with smoothing techniques. Also, the simple use of smoothing technicality can help stopped of ill-conditioning [3-6]. CG method suggestion by Polak, Ribiére and Polyak (PRP) In general it is believed to be one of the most effective methods [7]. The CG parameter of the PRP method is given by

$$\beta_k^{PRP} = \frac{F_k^T y_k}{\|F_k\|^2},$$  

(1.5)

where $y_k = F_k - F_{k-1}$. As we know that $F_k$ is the gradient of $f(x_k)$ and $F_{k-1}$ is the gradient of $f(x_{k-1})$. Much research has used hybridization methods to obtain satisfactory results with a high accuracy and speed to solve Unconstrained Optimization Problems. There are other methods that came from making some improvements to the mentioned methods. The authors introduced many techniques to solve various optimization and reliability problems [see 8-18], but in this work, we will combined the $\beta_k^{PRP}$ (Polak–Ribiére–Polyak) method and (Fletcher-Reeves) method $\beta_k^{FR}$ to obtain more accurate results, was used the Strong Wolfe-Powell Line Search, our method has the sufficient descent property, and it’s satisfy the sufficient descent condition under Appropriate condition by restricting some parameters. The sufficient descent condition:

$$\frac{\|F_k\|^2}{\|\nabla f_k\|^2} \leq - c \|F_k\|^2,$$

(1.6)

where $c > 0$.

The numerical results were compared with other algorithms, and with regard to global convergence, it was proven by imposing two of the assumptions applied in this field.

2. The New Hybrid Algorithm

In this section we suggest a new hybrid algorithm by mix (PRP) method and (FR) method as follows: we use $\beta_k^{Hybrid}$ as follows:

$$\beta_k^{Hybrid} = \zeta_1 \beta_k^{PRP} + \zeta_2 \beta_k^{FR},$$

(2.1)

where $\zeta_1, \zeta_2 > 0$, and $\beta_k^{FR}, \beta_k^{PRP}$ they are computed by (1.4), (1.5) respectively.

We will use the following search direction $d_k$

$$d_k = \begin{cases} -F_k & \text{if } k = 0 \\ -F_k + \beta_k^{Hybrid} \eta_k \lambda y_k & \text{if } k \geq 1 \\ \end{cases},$$

(2.2)

where $\lambda > 0$, and $\eta_k = \|w_k\|^2 F_k^T y_k / \|F_k\|^2$, and $w_k = x_{k+1} - x_k$, $y_k = F_{k+1} - F_k$.

With regard to the line search used in the proposed algorithm, it was as follows

$$|f(x_k + \alpha_k d_k)^T d_k| \leq \sigma |\alpha_k| |d_k|^2,$$

(2.3)

where $\sigma > 0$.

From the above, the proposed algorithm will be as the following:

2.1 The hybrid Algorithm

Step 1. Select an primary point $x_0 \in \mathbb{R}^n$, $\varepsilon \in (0,1), \zeta_1 > 0, \zeta_2 > 0, \lambda > 0, \alpha > 0, d_0 = -F_0 = -\nabla f(x_0), k := 0$.

Step 2. if $\|F_{k-1}\| \leq \varepsilon$, then stop, otherwise go to the next step.
Step 3. Compute $\alpha_k = \frac{F(x_k)^T d_k}{d_k \cdot \nabla F(x_k) \cdot d_k}$.

Step 4. $x_{k+1} = x_k + \alpha_k d_k$, if $\|F_k\| \leq \varepsilon$, then stop.

Step 5. Compute the search direction $d_k$ by (2.2), where $\beta_k^{Hybrid}$ calculated by (2.1).

Step 6. Set $k := k + 1$, go to step 3.

3. Global Convergence Property

The following assumptions are used to prove the global convergence of the suggest conjugate gradient method.

3.1: Assumption

(i) Where $x_0$ is given as initial point then $f(x)$ it’s have a lower bound in the set $\Omega = \{x \in \mathbb{R}^n \ni f(x) \leq f(x_0)\}$.

(ii) $f$ is differentiable, and its gradient $g$ is Lipschitz continuous, namely, there exists a constants $L > 0$ such that: $\|g[x] - g[y]\| \leq L\|x - y\|, \forall x, y \in \Omega_o$, where $\Omega_o$ be any convex set that contains $\Omega$.

3.1 Lemma. Assume the Assumption 3.1 holds, and suppose that the sequences $\{F_k\}, \{d_k\}$ are generated by Algorithm 2.1. Let the condition (1.6) hold, then

$$\sum_{k=0}^{\infty} \frac{(F_k^T d_k)^2}{\|F_k\|^2} < +\infty \quad (3.1)$$

Proof: The proof can be seen in [19] $\square$

3.2 Theorem. Suppose that Assumption 3.1 holds and the two sequences $\{F_k\}, \{d_k\}$ are generated by Algorithm 2.1, assume the parameter $2\gamma_1 + \gamma_2 \leq 1$ then

$$\lim_{k \to \infty} \inf \|F_k\| = 0 \quad (3.2)$$

Proof: by (1.6), (2.3) and (3.1) we get

$$\sum_{k=0}^{\infty} \frac{(F_k^T d_k)^4}{\|F_k\|^4} < +\infty \quad (3.4)$$

Suppose we have $S_k$ equal to

$$\frac{\|d_k\|^2}{\|F_k\|^4} \quad (3.5)$$

Now (3.1) becomes

$$\sum_{k=0}^{\infty} \frac{1}{S_k} < +\infty \quad (3.6)$$

We use contradiction concepts to prove this theorem, Assume it is not true,
Then there exist a positive number $\mu > 0$ such that:

$$\|F_k\| \geq \mu \quad \forall k \geq 0 \quad (3.7)$$

From (2.1) we have

$$\beta^H_k = \zeta_1 \beta^P_k + \zeta_2 \beta^F_k \geq \zeta_2 \frac{\|F_k\|^2}{\|F_{k-1}\|^2} \quad (3.8)$$

Squaring both sides of (2.2), we obtain

$$\|d_k\|^2 = (F_k + \beta^H_k \cdot \theta_k \cdot \lambda)^2 \quad (3.9)$$

Dividing both sides by $\|F_k\|^4$ and applied (3.5), (3.8) and the parameter $2\zeta_1 + \zeta_2 \leq 1$ we get

$$S_k \leq S_{k-1} + \frac{1}{\|F_k\|^2} (1 + \frac{2|F_k \theta_k y_k|}{\|F_k\|^2}) \quad (3.10)$$

From the strong Wolfe-Powell line search and (1.6) we get

$$S_k \leq S_{k-1} + \frac{1}{\|F_k\|^2} \left(1 + \frac{2\sigma|F_{k-1} y_{k-1} \theta_{k-1}|}{\|F_{k-1}\|^2}\right) \leq S_{k-1} + \frac{[1 + 2\sigma(2 - c)]}{\|F_k\|^2} \quad (3.11)$$

Using the fact $S_0 = 1/\|F_0\|^2$ we get

$$S_k \leq \left[1 + 2\sigma(2 - c)\right] \sum_{k=0}^\infty \frac{1}{\|F_k\|^2} \quad (3.12)$$

By applying (3.7) we get

$$S_k \leq \left[1 + 2\sigma(2 - c)\right] \sum_{k=0}^\infty \frac{k+1}{\mu^2} \quad (3.13)$$

That's mean

$$\sum_{k=0}^\infty \frac{1}{S_k} = +\infty \quad (3.14)$$

This contradiction with (3.6) then (3.2) is true and the proof is complete. □

4. Numerical Results

In this part we introduce the numerical results of the proposed algorithm and the comparison with some algorithms that have addressed the same problem within the unconstrained optimization Problems proposed by Q Li et al. [20], B. A. Hassan et al. [21] and H. Liu et al. [22], which is encoded with numerical results tables as follows QD, BA and HX respectively. The parameters selected as follows: $\rho = 0.8$, $\sigma = 0.4$, $\epsilon = 10^{-9}$, $\lambda = 0.8$. 
\( \zeta_1 = 0.5, \zeta_2 = 0.3 \), And the stop condition \( \|F_{k-1}\| \leq 10^{-9} \). all algorithms are perform through MATLAB R2014 and run on PC with 2.5 GHz CPU processor and 12 GB RAM and Windows XP operation system. The results are shown in the following tables.

**Table 1:** Functions evaluations (f eval) and iterations (iter).

| problem | Dim | f eval | Hybrid | QD | BA | HX | Iter |
|---------|-----|--------|--------|----|----|----|------|
| P1      | 500000 | 90 | 163 | 112 | 681 | 18 | 22 | 20 | 170 |
|         | 500000 | 90 | 163 | 112 | 681 | 18 | 22 | 20 | 170 |
|         | 500000 | 84 | 153 | 77 | 675 | 16 | 19 | 16 | 168 |
|         | 500000 | 84 | 153 | 77 | 675 | 16 | 19 | 16 | 168 |
| P2      | 500000 | 90 | 163 | 112 | 681 | 18 | 22 | 20 | 170 |
|         | 500000 | 58 | 12 | 102 | 660 | 14 | 2 | 20 | 165 |
|         | 500000 | 49 | 270 | 88 | 655 | 11 | 31 | 16 | 163 |
| P3      | 500000 | 76 | 597 | 77 | 20693 | 18 | 78 | 5 | 545 |
|         | 500000 | 132 | 947 | 77 | 27781 | 32 | 114 | 5 | 734 |
|         | 500000 | 83 | 279 | 77 | 19115 | 20 | 44 | 5 | 502 |
|         | 500000 | 104 | 389 | 77 | 29216 | 25 | 65 | 5 | 764 |
| P4      | 500000 | 261 | 261 | 32 | 880 | 61 | 61 | 3 | 293 |
|         | 500000 | 268 | 268 | 32 | 1579 | 64 | 64 | 3 | 525 |
|         | 500000 | 265 | 265 | 32 | 217 | 62 | 62 | 3 | 72 |
|         | 500000 | 266 | 266 | 32 | 1264 | 62 | 62 | 3 | 421 |
| P5      | 500000 | 91 | 10606 | 121 | 506 | 19 | 585 | 6 | 127 |
|         | 500000 | 92 | 19868 | 101 | 526 | 19 | 1014 | 6 | 132 |
|         | 500000 | 82 | 1978 | 103 | 450 | 17 | 142 | 6 | 113 |

**Table 2:** CPU-Time (in seconds)

| problem | Dim | CPU-Time |
|---------|-----|----------|
|         |     | Hybrid | QD | BA | HX |
| P1      | 500000 | 0.453125 | 1.03125 | 0.5 | 3.25 |
|         | 500000 | 0.421875 | 0.6875 | 0.484375 | 3.1875 |
|         | 500000 | 0.3125 | 0.578125 | 0.328125 | 3.015625 |
|         | 500000 | 0.40625 | 0.609375 | 0.28125 | 3.203125 |
| P2      | 500000 | 0.421875 | 0.65625 | 0.51625 | 3.3125 |
|         | 500000 | 0.296875 | 0.03125 | 0.390625 | 3.453125 |
|         | 500000 | 0.40625 | 0.5625 | 0.28125 | 3.171875 |
|         | 500000 | 0.21875 | 1.171875 | 0.328125 | 3.15625 |
From above tables generally our algorithm it showed better results compared with the other algorithms. Note that the comparison was made with the following terms:

* Number of iterations.
* Number of functions evaluations.
* CPU time.

5. Conclusions
In this paper a hybrid CGM has been suggested to solve the unconstrained optimization problems by combining a (Polak–Ribiére–Polyak) method and (Fletcher-Reeves) method. The global convergence has been proven for the proposed algorithm. Numerical results showed that hybrid method have competitive edge to other three conjugate gradient methods and it's promising and effective tool for solving optimization problems.

6. References

[1] Yu G Guan L and Chen W 2008. “Spectral conjugate gradient methods with sufficient descent property for large-scale unconstrained optimization”, Optimization Methods & Software, V 23, p. 275–293.

[2] Abbo K K 2008. “New CG Method for Large-Scale Unconstrained Optimization Based on Nazareth theorem”, Iraqi Journal of Statistical Sciences, V 13, p. 53–65.

[3] Pang D and Du S 2016. “The smoothing Fletcher-Reeves conjugate gradient method for solving finite minimax problems”, Science Asia, V 42, p. 40–45.

[4] Wasi H A and Shiker M A K 2020. "A new conjugate gradient method for solving large scale systems of monotone equations", International Journal of Advanced Science and Technology, V 29 (4), p.2303- 2314.

[5] Shiker M A K and Sahib Z 2018."a modified trust-region method for solving unconstrained optimization". Journal of Engineering and Applied Sciences, V 13(22), p. 9667-9671. https://doi.org/10.3923/jeasci.2018.9667.9671

[6] Dwail H H Mahdi M M Wasi H A Hashim K H Dreeb N K Hussein A H and Shiker M A K 2020. "A new modified TR algorithm with adaptive radius to solve a nonlinear systems of equations", “in press”, accepted paper for publication in IOP Science, 1st International Virtual Conference of the University of Babylon (IVCUB)- Iraq.
[7] Kafaki S B and Ghanbari R 2014. "A descent extension of the Polak–Ribiére–Polyak conjugate gradient method", Computers and Mathematics with Applications, V 68, p. 2005–2011.

[8] Dwail H H and Shiker M A K 2020. "Using a new trust region algorithm with nonmonotone adaptive radius for solving nonlinear systems of equations", International Journal of Advanced Science and Technology, V 29(4), p 2351- 2360.

[9] Hussein H A and Shiker M A K 2020. "A modification to Vogel’s approximation method to Solve transportation problems", “in press”, accepted paper for publication in IOP Science, 5th International Scientific Conference for Iraqi Al-Khwarizmi Society, Iraq.

[10] Hussein H A and Shiker M A K 2020. "Two New Effective Methods to Find the Optimal Solution for the Assignment Problems", “in press”, accepted paper for publication in Journal of Advanced Research in Dynamical and Control Systems.

[11] Hassan Z A H H and Shiker M A K 2018. "Using of generalized baye’s theorem to evaluate the reliability of aircraft systems", Journal of Engineering and Applied Sciences, (Special Issue13), p. 10797-10801. https://doi.org/10.36478/jeasci.2018.10797.10801.

[12] 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", “in press”, accepted paper for publication in IOP Science, 5th International Scientific Conference for Iraqi Al-Khwarizmi Society, Iraq.

[13] Mahdi M M and Shiker M A K 2020. "Three terms of derivative free projection technique for solving nonlinear monotone equations”, “in press”, accepted paper for publication in IOP Science, 5th International Scientific Conference for Iraqi Al-Khwarizmi Society, Iraq.

[14] Mahdi M M and Shiker M A K 2020. "Solving systems of nonlinear monotone equations by using a new projection approach”, “in press”, accepted paper for publication in IOP Science, 1st International Virtual Conference of the University of Babylon (IVCUB)- Iraq.

[15] 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”, “in press”, accepted paper for publication in Journal of Advanced Research in Dynamical and Control Systems.

[16] Shiker M A K and Amini K 2018. "A new projection-based algorithm for solving a large scale nonlinear system of monotone equations", Croatian operational research review, V 9(1), p. 63-73. https://doi.org/10.17535/crorr.2018.0006.

[17] Hassan Z A H and Mutar E K 2017. "Geometry of reliability models of electrical system used inside spacecraft”, Second Al-Sadiq International Conference on Multidisciplinary in IT and Communication Science and Applications (AIC-MITCSA), p 301-306.

[18] Hussein H A and 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", “in press”, accepted paper for
publication in *IOP Science*, 5th International Scientific Conference for Iraqi Al-Khwarizmi Society, Iraq.

[19] Dai Y H and Yan Y 2000. "Nonlinear Conjugate Gradient Methods", *Shanghai Science and Technology Publisher*, Shanghai, 2000.

[20] Li Q and Huili D 2011. "A class of derivative-free methods for large-scale nonlinear monotone equations", *IMA Journal of Numerical Analysis*, V 31, p 1625–1635.

[21] Hassan B A Wais O M T and Mahmood A A 2019. "A Class of Descent Conjugate Gradient Methods for Solving Optimization Problems", *Applied Mathematical Sciences*, V 13, p 559–567.

[22] Liu H Wang H Qian X and Rao F 2015. "A conjugate gradient method with sufficient descent property", *Springer*, V 70, p 269–286.