Performance of New Line Search Methods with Three-Term Hybrid Descent Approach for Unconstrained Optimization Problems

Okundalaye Oluwaseun Olumide
Department of Mathematical Sciences,
Faculty of Science,
E-mail: oluwaseun.okundalaye@aaua.edu.ng
Adekunle Ajasin University, Akungba–Akoko (AAUA),
P. M. Box 01, Akungba-Akoko, Ondo State, Nigeria.

ABSTRACT

In this paper, I demonstrates the performance and efficiency of new line search methods with three-term hybrid descent method for the solution of unconstrained optimization problems. The techniques advanced the sustainable range of step-length to a broader level than the previous ones and give a suitable initial step-length at each step of the iterations. The global convergence rate of the new line with three-term hybrid descent method search is carried studied. Some numerical results through performance profile shows that among the new search method modified Wolfe line search method in CPU time and iterations is best in practical computation.
Keywords: Quasi-Newton method, search direction, step-length, global convergence, performance profile.

1. Introduction

Considering an unconstrained optimization problem of the form

$$\min_{x \in R^n} f(x). \tag{1.1}$$

where R^n is the n-dimensional Euclidean space and $f : R^n \to R$ is continuous differentiable.
The solution of (1.1) require using an iterative methods with a starting initial point x_0 to obtain a points of sequence $\{x_k\}$, $k = 1, 2, 3...n$, and show progressive approximations to the required solution using the iterative technique below

$$x_{k+1} = x_k + \alpha_k d_k \qquad k = 0, 1, ...n. \tag{1.2}$$

Where x_k is the current iterate, d_k is the search direction, and α_k the step-length. Let (x^*) be a minimizer of (1.1) and thus be a stationary point that satisfies $g(x^*) = 0$. We denote $f(x_k)$ by f_k, f_{x^*} by f^*, and $\nabla f(x_k)$ by g_x respectively. The line search method required two step at each iterations; the first step is to find a search direction d_k and the second step is to select the step-length α_k along the search direction. On the other hand, the d_k is typically needed to satisfy the descent condition $g_k^T d_k < 0$ that guarantees d_k is a descent direction of $f(x)$ at x_k. It has been proved that search direction perform an essential role in line search techniques and that the step-length methods mainly guarantee global convergence.
The following condition holds in order to obtain the global convergence of the line search methods.

$$-\frac{g_k^T d_k}{||g_k||.||d_k||} \geq c. \tag{1.3}$$

where $c \in (0,1]$ is a constant. The condition (1.3) is called angle property. Several methods of selecting d_k and α_k yield different convergence properties. Also, for the step-length, the sequence of iterates x_k defined by (1.2) globally converges with some rate of convergence.
The step-length can be determine either by using; exact line search or inexact line search. For exact line search, α_k is obtained by using the formula below

$$\alpha_k = \arg\min_{\alpha>0}(f(x_x + \alpha_k d_k)). \tag{1.4}$$

However, (1.4) is complex and often problematic to find in practical computation. Therefore, the inexact line search has been introduced by previous researchers; (Armijo, 1966), (Wolfe, 1969), and (Goldstein, 1965) to overcome this challenge. Recently, (Yuan, Wei, & Yang, 2019) proposed the global convergence of some conjugate gradient method with inexact line search, (Berahas, Cao, & Scheinberg, 2021) analyse the global convergence of some line search methods, (Hosseini Dehmiry, 2020) show the global convergence of quasi family under conjugate technique, (Yuan, Sheng, Wang, Hu, & Li, 2018) the global convergence under quasi family, (Masmali, Salleh, & Alhawarat, 2021) proposed the global convergence properties on a large scale problem, and (Wang, Yin, & Zeng, 2019) show the global convergence of non-convex optimization problem.
This work focused on new inexact line search rule called modified line search rules that advanced the scope of appropriate step-length and give a good initial step-length at each iteration.
Modified Armijo rule.
Set scalar $l_k > 0, \beta \in (1,0)$, $\sigma \in (0, \frac{1}{2})$, and set $S_k = -\frac{g_k^T d_k}{l_k ||d_k||^2}$. Let α_k be the largest α in $\{s_k, \beta s_k, \beta^2 s_k ....\}$ such that

$$f_k - f(x_k + \alpha d_k) \geq \sigma\alpha ||d_k|| w_k(\alpha), \tag{1.5}$$

Modified Goldstein rule.
A fixed scalar $\sigma = (0, \frac{1}{2})$ is selected and α_k is chosen to satisfy

$$-(1-\sigma)\alpha_k g_k^T d_k \geq f_k - f(x_k + \alpha_k d_k) \geq \sigma\alpha_k ||d_k|| w_k(\alpha_k), \tag{1.6}$$

Modified Wolfe rule.
The step α_k is chosen to satisfy

$$f_k - f(x_k + \alpha_k d_k) \geq \sigma\alpha_k ||d_k|| w_k(\alpha_k),$$

and

$$g(x_k + \alpha_k d_k)^T d_k \geq y g_k^T d_k. \tag{1.7}$$

where σ and y are some scalars with $\sigma \in (0, \frac{1}{2})$ and $y \in (0,1)$ for $k = 0, 1, 2, \ldots n$. Then, the sequence of $\{x_k\}_{k=0}^{\infty}$ converges to the optimal point x^* which minimizes $f(x)$. Hence, modified Armijo, Goldsten, and Wolfe line search methods are used in this research associated with three-term hybrid descent search direction.
This paper is organized as follows; In section (2), I illustrate and discussed extensively the importance of search direction in iterative method. The new three-term hybrid method and its convergence analysis are discussed in section (3). Numerical results and discussion are given in section (4). The paper ends with a short conclusion in section (5).

2. The Search Direction

In an iterative method of solving an unconstrained optimization problem, search direction is most important and essential which includes; conjugate gradient method, Newton method, and quasi-Newton method. Many of these techniques used in solving unconstrained optimization problems relies solely on the results of search direction d_k. The conjugate gradient (CG) approach contain a class of unconstrained optimization algorithms with a properties of low memory, easy computation and global strong convergence, making them efficient for solving large-scale problems in the form of $min_{x\in\mathbb{R}^n} f(x)$ with the differentiable non-linear function $f : R^n \to R$. The CG method is also essential in finding the minimum value of functions of an unconstrained optimization problems. (Stiefel, 1952) starts a CG method of solving a linear system of equations with a symmetric positive definite coefficient matrix for minimizing a strictly convex quadratic function. After, (Fletcher & Reeves, 1964) used the CG method to solve unconstrained optimization problems. Recently, CG methods is becoming more popular iterative methods to solve large-scale unconstrained optimization problems, since they do not required the storage of matrices (Hanke, 2017; Meurant, 2020; Alhawarat, Alhamzi, Masmali, & Salleh, 2021; Livieris, Tampakas, & Pintelas, 2018; Waziri, Ahmed, & Sabi'u, 2020; Yuan, Wang, & Sheng, 2020; Hassan, Abdullah, & Jabbar, 2019). The conjugate gradient search direction approach is defined by

$$d_k = \begin{cases} -g_k & k = 0, \\ -g_k + \beta_k d_{k-1} & k \geq 1. \end{cases} \tag{2.1}$$

where $g_k = \nabla f(x_k)$ and β_k is known as the CG coefficient. We have several ways of calculating β_k and some well-known formulae are;

$$\begin{aligned} \beta_k^{FR} &= \frac{g_k^T g_k}{||g_{k-1}||^2}, \\ \beta_k^{PR} &= \frac{g_k^T (g_k - g_{k-1})}{||g_{k-1}||^2}, \\ \beta_k^{HS} &= \frac{g_k^T (g_k - g_{k-1})}{(g_k - g_{k-1})^T d_{k-1}}, \\ \beta_k^{BAN} &= \frac{-g_k^T (g_k - g_{k-1})}{g_{k-1}^T (g_k - g_{k-1})}, \\ \beta_k^{HZ} &= \left(y_k - \frac{(2d_k(\|y_k\|)^2)}{(d_k y_k)} \right)^T \left(\frac{g_{k+1}}{(d_k^T y_k)} \right). \end{aligned}$$

where g_k and g_{k-1} are gradients of $f(x)$ at the points x_k and x_{k-1} and $y_k = g_k - g_{k-1}$ respectively. While $||.||$ is a norm of vectors and d_{k-1} is a direction for the previous iteration. The above corresponding coefficients are known as (Fletcher and Reeves, 1964), (Polak and Ribiere, 1969) and (Hestenes and Stiefel, 1952) and (Zhang, Zhou, & Li, 2007; Narushima, Yabe, & Ford, 2011; Andrei, 2013; Liu & Li, 2014; Dong, Liu, & He, 2015; Moyi & Leong, 2016). Recently, (Kobayashi, Narushima, & Yabe, 2017; Gao & He, 2018; Bojari & Eslahchi, 2020; Baluch, Salleh, & Alhawarat, 2018; ABDULLAH & JAMEEL, 2019) proposed three-term CG methods which satisfy the sufficient descent condition.

$$g_k^T d_k \leq -\bar{c}||g_k||^2 \qquad for \quad all \quad k = 0, 1, 2...n. \tag{2.2}$$

and a positive constant $\bar{c}$, independently of line searches. They proposed the modified FR method defined by

$$d_k = -\bar{\theta_k} g_k + \beta^{FR} d_{k-1}$$

Where $\bar{\theta_k} = \frac{d_{k-1}^T y_{k-1}}{||g_{k-1}||^2}$. Since this search direction satisfies $g_k^T d_k < -||g_k||^2$ for all k, it can be written by the three-term form:

$$d_k = -g_k + \beta^{FR} d_{k-1} - \theta_k^1 g_k, \tag{2.3}$$

where $\theta_k^1 = \frac{g_k^T d_{k-1}}{||g_{k-1}||^2}$. They also proposed the modified PR methods and the modified HS method, which are respectively given by

$$d_k = -g_k + \beta^{PR} d_{k-1} - \theta_k^{(2)} y_{k-1}, \tag{2.4}$$

$$d_k = -g_k + \beta^{HS} d_{k-1} - \theta_k^{(3)} y_{k-1}, \tag{2.5}$$

Where $\theta_k^{(2)} = \frac{g_k^T d_{k-1}}{||g_{k-1}||^2}$ and $\theta_k^{(3)} = \frac{g_k^T d_{k-1}}{d_{k-1}^T y_{k-1}}$. Cheng gave another modification of PR method:

$$d_k = -g_k + \beta_k^{PR}(I - \frac{g_k g_k^T}{g_k^T g_k}) d_{k-1} = -g_k + \beta_k^{PR} d_{k-1} - \beta_k^{PR} \frac{g_k^T d_{k-1}}{g_k^T g_k} g_k. \tag{2.6}$$

They obtained their global convergence properties under appropriate line searches. The recent modification can be seen (Masmali et al., 2021; Liu, Feng, & Zou, 2018; Liu & Du, 2019; Abubakar, Kumam, Ibrahim, Chaipunya, & Rano, 2021; Bojari & Eslahchi, 2020). We observe that these approach always satisfy $g_k^T d_k = -||g_k||^2 < 0$ for all k, which indicate the sufficient descent condition with $\bar{c} = 1$.
In quasi-Newton family, the search direction is the solution of linear system

$$d_k = -H_k g_k. \tag{2.7}$$

where H_k is an approximation of Hessian. Initial matrix H_0 is selected by the identity matrix, which thereafter updates by an update formula. There are a few update formulae that are widely used like Davidon-Fletcher-Powell(DFP), BFGS, and Broyden family formula. This study employs a BFGS formula in a classical algorithm and the new hybrid method. The update formula for BFGS is

$$H_{k+1} = H_k - \frac{H_k s_k s_k^T H_k}{s_k^T H_k s_k} + \frac{y_k y_k^T}{s_k^T y_k}, \tag{2.8}$$

with $s_k = x_k - x_{k+1}$ and $y_k = g_k - g_{k-1}$. The approximation that the Hessian must fulfil is

$$H_{k+1} s_k = y_k, \tag{2.9}$$

This condition is essential to hold for the updated matrix H_{k+1}. Note that it is only feasible to fulfil the equations if

$$s_k^T y_k > 0. \tag{2.10}$$

which is called the curvature condition.

3. The Proposed Three-Term Method

The modification on three-term approach have been proposed by many researchers; One of the research is by Ludwig (Ludwig, 2007) which is a hybrid between quasi-Newton methods with Gauss-Siedel method to solve the system of linear equation. Then, Luo et.al (Luo, Tang, & Zhou, 2008) suggested the new hybrid method which can solve the system of non-linear equations by combining quasi-Newton method with chaos optimization. Besides, Han and Newman (Han & Neumann, 2003) combine the Quasi-Newton methods and Cauchy descent method to solve unconstrained optimization problems and recognized as quasi-Newton-SD method. Also Ibrahim et.al(Ibrahim, Mamat, & Leong, 2014) proposed BFGS-CG method which is between quasi-Newton and conjugate gradient method and come out with this search direction.

$$d_k = \begin{cases} -H_k g_k & k = 0, \\ -H_k g_k + \eta(-g_k + \beta_k d_{k-1}) & k \geq 1. \end{cases}$$

where $\eta > 0$ $\qquad \beta_k = \frac{g_k^T g_k - 1}{g_k^T d_k - 1}$.

Hence, the modification on Quasi-Newton by previous researchers spawned the new idea on hybrid; the classical method to yield the new hybrid method. Hence, a new hybrid search direction which combines the concept of search direction of quasi-Newton and conjugate gradient method is created. It yields a new search direction of hybrid method which is known as Three-term BFGS-CG method. Search direction for Three-term BFGS-CG method

$$d_k = \begin{cases} -H_k g_k & k = 0, \\ -H_k g_k + \eta(-g_k + \beta_k d_{k-1} - \beta_k \frac{g_k^T d_{k-1}}{g_k^T g_k} g_k) & k \geq 1. \end{cases} \tag{3.1}$$

where $\eta > 0$ $\qquad \beta_k = \frac{g_k^T g_k - 1}{g_k^T d_k - 1}$

Hence, the complete algorithms for BFGS, (CG-HS, CG-PR, CG-FR), and Three-Term BFGS-CG method will be arranged in Algorithm(16), Algorithm(17) and (18) respectively.

Algorithm(1) Modified Armijo line search for Three-term BFGS-CG.

Step 1. Given a starting point x_0 and $H_o = I_n$, choose values for s, β, and σ. Set $k = 1$
Step 2. Terminate if $||g(x_{k+1})|| < 10^{-6}$ or $k \leq 1000$
Step 3. Calculate the search direction by (3.1).
Step 4. Calculate the step length α_k by (1.5).
Step 5. Compute the difference between $s_k = x_k - x_{k-1}$ and $y_k = g_k - g_{k-1}$
Step 6. Update H_{k+1} by (12) to obtain H_k
Step 7. Set $k = k + 1$ and go to step 2.

Algorithm(2) Modified Goldstein line search for Three-term BFGS-CG method .

Step 1. Giving a starting initial point x_0 and choose values for s, β, and σ. Set $k = 1$
Step 2. Terminate if $||g(x_{k+1})|| < 10^{-6}$ or $k \leq 1000$.
Step 3. Calculate the search direction by (3.1)
Step 4. Calculate the step size α_k by (1.6)
Step 5 Compute the difference $s_k = x_k - x_{k-1}$ and $y_k = g_k - g_{k-1}$
Step 6. Update H_{k+1} by (2.8) to obtain H_k.
Step 7.Set $k = k + 1$ and go to step 2.

Algorithm(3) Modified Wolfe line search for Three-term BFGS-CG method.

Step 1. Given a starting point x_0 and $H_o = I_n$, choose values for s,β, and σ. Set $k = 1$.
Step 2. Terminate if $||g(x_{k+1})|| < 10^{-6}$ or $k \leq 1000$.
Step 3. Calculate the search direction by(3.1)
Step 4. Calculate the step length α_k by (1.7).
Step 5. Compute the difference between $s_k = x_k - x_{k-1}$ and $y_k = g_k - g_{k-1}$.
Step 6. Update H_{k+1} by (2.8) to obtain H_k.
Step 7. Set $k = k + 1$ and go to step 2.

Based on Algorithm (1), (2), and (3), we assume that every search direction d_k satisfied the descent condition $g_k^T d_k < 0$.

Hence, we need to make a few assumption based on the objective function

Assumption 3.1

H1: The objective function f is twice continuously differentiable.

H2: The level set L is convex. Moreover, positive constants c_1 and c_2 exist, satisfying

$$c_1||z||^2 \leq z^T F(x) z \leq c_2||z||^2. \tag{3.2}$$

for all $z \in R^n$ and $x \in L$ where $f(x)$ is the Hessian matrix of f.

H3: The Hessian matrix is Lipschitz continuous at the point x^* that is, there exist the positive constant c_3 satisfying

$$||g(x) - g(x^*)|| \leq c_3||x - x^*||. \tag{3.3}$$

for all x in a neighbourhood of x^*

Theorem (3.2). Let $\{B_k\}$ be generated by BFGS formal (2.8), where B_k is symmetric and positive definite, and $y_k^T s_k > 0$ for k. Furthermore, assume that $\{s_k\}$ and $\{y_k\}$ are such that

$$\frac{||(y_k - G_*)s_k||}{||s_k||} \leq \epsilon_k. \tag{3.4}$$

for some symmetric and definite matrix $G(x^*)$ and for some sequence ϵ_k with the property. $\sum_{k=1}^{\infty} \epsilon_k < \infty$. Then

$$lim_{k\to\infty} \frac{||(B_k - G_* d_k)||}{||d_k||} = 0. \tag{3.5}$$

and the sequence $||\{B_k\}||$, $||\{B_k^{-1}\}||$ are bounded.

Theorem (3.3). Global convergence.

Suppose that Assumption (3.1) and Theorem(3.2) hold. Then

$$\lim_{k\to\infty} ||g_k||^2 = 0. \tag{3.6}$$

Proof.

from the condition $g_k^T d_k < 0$, we see that

$$g_k^T d_k = -g_k^T B_k^{-1} g_k + \eta g_k^T (-g_k + \beta_k d_{k-1} - \beta_k \frac{g_k^T d_{k-1}}{g_k^T g_k} g_k), \tag{3.7}$$

$$= -g_k^T B_k^{-1} g_k + \eta(-g_k^T g_k + \frac{g_k^T g_{k-1}}{g_k^T d_{k-1}} g_k^T d_{k-1} - \frac{g_k^T g_{k-1}}{g_k^T d_{k-1}} \frac{g_k^T d_{k-1}}{g_k^T g_k} g_k^T g_k), \tag{3.8}$$

$$= -g_k^T B_k^{-1} g_k + \eta(-g_k^T g_k + g_k^T g_{k-1} - g_k^T g_{k-1}), \tag{3.9}$$

then

$$g_k^T d_k = -g_k^T B_k^{-1} g_k + \eta(-||g_k||^2), \tag{3.10}$$

$$\leq -\lambda_k ||g_k||^2 + \eta(-||g_k||^2), \tag{3.11}$$

$$g_k^T d_k \leq c_1 ||g_k||^2, \tag{3.12}$$

where $c_1 = -(\lambda_k + \eta)$ which is bounded away from zero. Hence, from the Armijo line search condition, we have that .

$$f_k - f_{k+1} \leq \sigma \alpha_k g_k^T d_k, \tag{3.13}$$

$$\leq \sigma \alpha_k c_1 ||g_k||^2, \tag{3.14}$$

holds for all k. Since f_k is decreasing and the sequence $\{f_k\}$ is bounded below by H2, we have that

$$\lim_{k\to\infty} (f_k - f_{k+1}) = 0, \tag{3.15}$$

Hence, this (3.14) and (3.15) imply

$$\lim_{k\to\infty} ||g_k||^2 = 0. \tag{3.16}$$

Table 1: Unconstrained optimization test problems

Test Problems	n-dimension	Sources
Powell badly scaled	2	More et al.
Beale	2	More et al.
Biggs Exp	6 6	More et al.
Chebyquad	4 6	More et al.
Colville polynomial	4	Michalewicz
Variably dimensioned	4, 8	More et al.
Freudenstein and Roth	2	More et al.
Goldstein price polynomial	2	Michalewicz
Himmelblau	2	Andrei
Penalty	1 2 4	More et al.
Extended Powell singular	4, 8	More et al.
Extended Rosenbrock	2, 10, 100, 200, 500, 1000	Andrei
Arwhead	10,50,100,500,1000	Andrei
PSC 1	2	More et al.
Six-hump camel back polynomial	2	Michalewicz
Extended Cliff	2, 4, 10, 100, 200, 500, 1000	Andrei
Extended Hiebert	2, 4, 10, 100, 200, 500, 1000	Andrei
Extended EP1	2,4,10	Michalewicz
Raydan	1 2, 4	Andrei
Raydan	2 2, 4	Andrei
Diagonal	3 2	Andrei
Cube	2, 10, 100, 200	More et al.

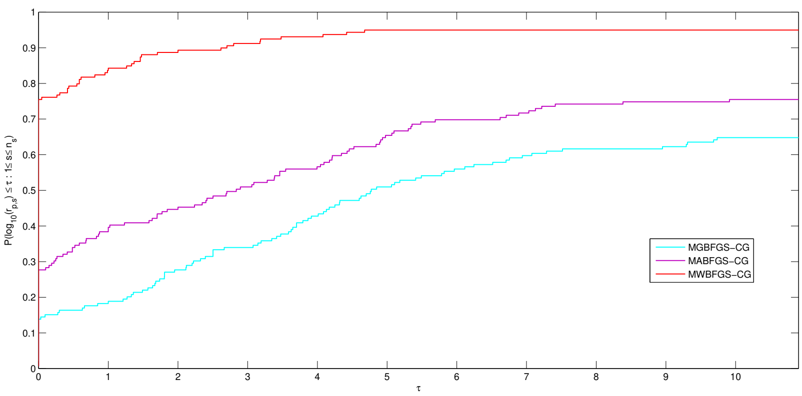

Figure 1: Performance Profile in a log_{10} scale based on iteration

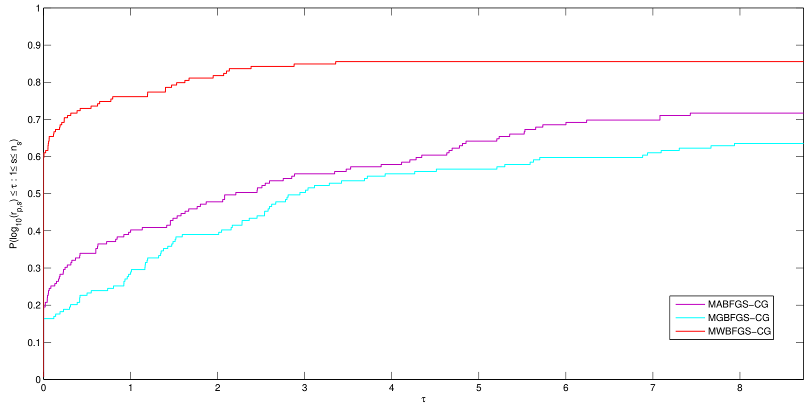

Figure 2: Performance Profile in a log_{10} scale based on CPU time

4. Numerical Results and Discussion

In this section, l used a set of some selected unconstrained optimization problems from the CUTEr suite to analyse the performance and efficiency of several new line search methods with three-term hybrid descent method. Each of the test problems were tested with dimensions varying from 2 to 1000. For each of the test problems, the initial point x_0 take further away from the minimum point. In doing so, it leads to test the global convergence properties and the robustness of the method. For the modified Armijo line search, modified Goldstein line search, and modified Wolfe line search l used $\sigma = (o, \frac{1}{2})$, stopping criteria $||g_k|| \leq 10^{-6}$, and the number of iterations exceeds a limit of 10,000. In general $p(\tau)$ is the fraction of problems with performance ratio τ; thus, a solver with high values of $p(\tau)$ is preferable. Performance profile were drawn for the above methods for performance and efficiency. The performance profile seeks to find how well the solvers perform relative to the other solvers on a set of problems. The implementations, numerical tests was performed on Matlab 2021a software. Table 1 show the unconstrained optimization test problems used to test the efficiency of the proposed line search methods. Performance profiles of methods are illustrated in Figures 1 and 2. From Figures 1 and 2, three-term hybrid modified Wolfe line search approach has the best performance since it can solve (97%) of the test problems compared with the three-term hybrid modified Armijo line of search(77%), and three-term hybrid modified Goldstein line of search(62%).

5. Conclusion

In summary, l presented performance and efficiency of a several line search methods with three-term BFGS-CG descent search for solving unconstrained optimization problems that guaranteed sufficient descent condition. I deduced that modified Wolfe line search technique perform best on performance profile as shown on the Figures. l also concluded that forming an hybrid method out of the existing methods is more efficient especially when the strength of the component are the target of the hybridization. Researchers and authors can implement modified Wolfe line search for large scale unconstrained optimization problem.

References

ABDULLAH, Z. M., & JAMEEL, M. S. (2019). Three-term conjugate gradient (ttcg) methods. *Journal of Multidisciplinary Modeling and Optimization*, *2*(1), 16–26.

Abubakar, A. B., Kumam, P., Ibrahim, A. H., Chaipunya, P., & Rano, S. A. (2021). New hybrid three-term spectral-conjugate gradient method for finding solutions of nonlinear monotone operator equations with applications. *Mathematics and Computers in Simulation*.

Alhawarat, A., Alhamzi, G., Masmali, I., & Salleh, Z. (2021). A descent four-term conjugate gradient method with global convergence properties for large-scale unconstrained optimisation problems. *Mathematical Problems in Engineering*, *2021*.

Andrei, N. (2013). A simple three-term conjugate gradient algorithm for unconstrained optimization. *Journal of Computational and Applied Mathematics*, *241*, 19–29.

Armijo, L. (1966). Minimization of functions having lipschitz continuous first partial derivatives. *Pacific Journal of mathematics*, *16*(1), 1–3.

Baluch, B., Salleh, Z., & Alhawarat, A. (2018). A new modified three-term hestenes–stiefel conjugate gradient method with sufficient descent property and its global convergence. *Journal of Optimization*, *2018*.

Berahas, A. S., Cao, L., & Scheinberg, K. (2021). Global convergence rate analysis of a generic line search algorithm with noise. *SIAM Journal on Optimization*, *31*(2), 1489–1518.

Bojari, S., & Eslahchi, M. (2020). Two families of scaled three-term conjugate gradient methods with sufficient descent property for nonconvex optimization. *Numerical Algorithms*, *83*(3), 901–933.

Dong, X. L., Liu, H. W., & He, Y. B. (2015). New version of the three-term conjugate gradient method based on spectral scaling conjugacy condition that generates descent search direction. *Applied Mathematics and Computation*, *269*, 606–617.

Fletcher, R., & Reeves, C. M. (1964). Function minimization by conjugate gradients. *The computer journal*, *7*(2), 149–154.

Gao, P., & He, C. (2018). An efficient three-term conjugate gradient method for nonlinear monotone equations with convex constraints. *Calcolo*, *55*(4), 1–17.

Goldstein, A. A. (1965). On steepest descent. *Journal of the Society for Industrial and Applied Mathematics, Series A: Control*, *3*(1), 147–151.

Han, L., & Neumann, M. (2003). Combining quasi-newton and steepest descent directions. *International Journal of Applied Mathematics*, *12*, 167–171.

Hanke, M. (2017). *Conjugate gradient type methods for ill-posed problems*. Chapman and Hall/CRC.

Hassan, B. A., Abdullah, Z. M., & Jabbar, H. N. (2019). A descent extension of the dai-yuan conjugate gradient technique. *Indonesian Journal of Electrical Engineering and Computer Science*, *16*(2), 661–668.

Hosseini Dehmiry, A. (2020). The global convergence of the bfgs method under a modified yuan-wei-lu line search technique. *Numerical Algorithms*, *84*(2), 781–793.

Ibrahim, M. A. H., Mamat, M., & Leong, W. J. (2014). The hybrid bfgs-cg method in solving unconstrained optimization problems. In *Abstract and applied analysis* (Vol. 2014).

Kobayashi, H., Narushima, Y., & Yabe, H. (2017). Descent three-term conjugate gradient methods based on secant conditions for unconstrained optimization. *Optimization Methods and Software*, *32*(6), 1313–1329.

Liu, J., & Du, S. (2019). Modified three-term conjugate gradient method and its applications. *Mathematical Problems in Engineering*, *2019*.

Liu, J., Feng, Y., & Zou, L. (2018). Some three-term conjugate gradient methods with the inexact line search condition. *calcolo*, *55*(2), 1–16.

Liu, J., & Li, S. (2014). New three-term conjugate gradient method with guaranteed global convergence. *International Journal of Computer Mathematics*, *91*(8), 1744–1754.

Livieris, I. E., Tampakas, V., & Pintelas, P. (2018). A descent hybrid conjugate gradient method based on the memoryless bfgs update. *Numerical Algorithms*, *79*(4), 1169–1185.

Ludwig, A. (2007). The gauss–seidel–quasi-newton method: A hybrid algorithm for solving dynamic economic models. *Journal of Economic Dynamics and Control*, *31*(5), 1610–1632.

Luo, Y.-Z., Tang, G.-J., & Zhou, L.-N. (2008). Hybrid approach for solving systems of nonlinear equations using chaos optimization and quasi-newton method. *Applied Soft Computing*, *8*(2), 1068–1073.

Masmali, I. A., Salleh, Z., & Alhawarat, A. (2021). A decent three term conjugate gradient method with global convergence properties for large scale unconstrained optimization problems. *AIMS Mathematics*, *6*(10), 10742–10764.

Meurant, G. (2020). On prescribing the convergence behavior of the conjugate gradient algorithm. *Numerical Algorithms*, *84*(4), 1353–1380.

Moyi, A. U., & Leong, W. J. (2016). A sufficient descent three-term conjugate gradient method via symmetric rank-one update for large-scale optimization. *Optimization*, *65*(1), 121–143.

Narushima, Y., Yabe, H., & Ford, J. A. (2011). A three-term conjugate gradient method with sufficient descent property for unconstrained optimization. *SIAM Journal on Optimization*, *21*(1), 212–230.

Stiefel, E. (1952). Methods of conjugate gradients for solving linear systems. *J. Res. Nat. Bur. Standards*, *49*, 409–435.

Wang, Y., Yin, W., & Zeng, J. (2019). Global convergence of admm in nonconvex nonsmooth optimization. *Journal of Scientific Computing*, *78*(1), 29–63.

Waziri, M., Ahmed, K., & Sabi'u, J. (2020). A dai–liao conjugate gradient method via modified secant equation for system of nonlinear equations. *Arabian Journal of Mathematics*, *9*(2), 443–457.

Wolfe, P. (1969). Convergence conditions for ascent methods. *SIAM review*, *11*(2), 226–235.

Yuan, G., Sheng, Z., Wang, B., Hu, W., & Li, C. (2018). The global convergence of a modified bfgs method for nonconvex functions. *Journal of Computational and Applied Mathematics*, *327*, 274–294.

Yuan, G., Wang, X., & Sheng, Z. (2020). Family weak conjugate gradient algorithms and their convergence analysis for nonconvex functions. *Numerical Algorithms*, *84*(3), 935–956.

Yuan, G., Wei, Z., & Yang, Y. (2019). The global convergence of the polak–ribière–polyak conjugate gradient algorithm under inexact line search for nonconvex functions. *Journal of Computational and Applied Mathematics*, *362*, 262–275.

Zhang, L., Zhou, W., & Li, D. (2007). Some descent three-term conjugate gradient methods and their global convergence. *Optimisation Methods and Software*, *22*(4), 697–711.