UNIVERSITI PUTRA MALAYSIA

# *SCALED THREE-TERM CONJUGATE GRADIENT METHOD VIA DAVIDON-FLETCHER-POWELL UPDATE FOR UNCONSTRAINED OPTIMIZATION*

ARZUKA IBRAHIM

IPM 2015 18

# SCALED THREE-TERM CONJUGATE GRADIENT METHOD VIA DAVIDON-FLETCHER-POWELL UPDATE FOR UNCONSTRAINED OPTIMIZATION

By

ARZUKA IBRAHIM

Thesis Submitted to the School of Graduate Studies, Universiti Putra Malaysia, in Fulfilment of the Requirements for the Degree of Master of Science

May, 2015

**DEDICATIONS**

*To my Dad Late Alhaji Arzuka Jega*

Abstract of thesis presented to the Senate of Universiti Putra Malaysia in fulfilment of the requirement for the degree of Master of Science

# SCALED THREE-TERM CONJUGATE GRADIENT METHOD VIA DAVIDON-FLETCHER-POWELL UPDATE FOR UNCONSTRAINED OPTIMIZATION

By

**ARZUKA IBRAHIM**

**May 2015**

**Chair: Associate Professor Mohd Rizam Abu Bakar, PhD**
**Faculty: Institute for Mathematical Research**

This thesis focus on the development of Scaled Three-Term Conjugate Gradient Method via the Davidon-Fletcher-Powell (DFF) quasi-Newton update for unconstrained optimization. The DFP method possess the merits of Newton's method and steepest descent method while overcoming their disadvantages. Over the years the DFP update has been neglected as a result of lacking the self correcting property for bad Hessian approximation. In this thesis, we proposed a Scaled Three-Term Conjugate Gradient Method by utilizing the DFP update for the inverse Hessian approximation via memoryless quasi Newton's method which satisfies both the sufficient descent and the conjugacy conditions. The basic philosophy is to restart the DFP update with a multiple of identity matrix in every iteration. An acceleration scheme is incorporated in the proposed method to enhance reduction in function value. Numerical results from an implementation of the proposed method on some standard unconstrained optimization problem shows that the proposed method is promising and exhibits superior numerical performance in comparison with other well-known conjugate gradient methods.

Abstrak tesis yang dikemukakan kepada Senat Universiti Putra Malaysia sebagai memenuhi keperluan untuk ijazah Master Sains

**KAEDAH KECERUNAN KONJUGAT BERSKALA TIGA SEBUTAN MELALUI KEMASKINI DAVIDON-FLETCHER-POWELL UNTUK PENGOPTIMUNAN TAK BERKEKANGAN**

Oleh

**ARZUKA IBRAHIM**

**Mei 2015**

**Pengerusi: Profesor Madya Mohd Rizam Abu Bakar, PhD**
**Fakulti: Institut Penyelidikan Matematik**

Tumpuan tesis ini adalah terhadap pembentukan kaedah kecerunan konjugat berskala tiga sebutan melalui kemaskini kuasi-Newton Davidon-Fletcher-Powell (DFP) untuk pengoptimuman tak berkekangan. Kaedah Davidon-Fletcher-Powell memiliki merit kaedah Newton dan kaedah penurunan tercuram di samping berupaya mengatasi kelemahan kaedah-kaedah ini. Sejak beberapa tahun kebelakangan ini kemaskini DFP diabaikan kerana kelemahan sifat pembetulan sendiri untuk penghamiran Hessian yang lemah. Dalam tesis ini, kami mencadangkan kaedah kecerunan konjugat berskala tiga sebutan dengan mengounakan kemaskini DFP untuk penghampiran Hessian songsang melalui kaedah kuasi-Newton tanpa memori yang memenuhi syarat penurunan kecerunan yang mencukupi dan syarat konjugasi. Falsafah asas adalah untuk mula semula kemaskini DFP dengan beberapa matrik identiti dalam setiap lelaran. Skim pecutan digabungkan ke dalam kaedah yang dicadangkan untuk meng galakkan pengurangan nilai fungsi. Keputusan berangka dari pelaksanaan kaedah yang dicadang ke atas beberapa masalah pengoptimuman tak berkekangan piawai, menunjukkan bahawa kaedah ini berpotensi dan mempamerkan prestasi berangka unggul berbanding dengan kaedah-kaedah kecerunan konjugat lain yang terkenal.

# ACKNOWLEDGEMENTS

All praise is due to Allah who in His infinite mercy grant's me life, good health, ability, hope, assistance and determination to put this thesis together. May His peace and blessings continue to shower on the best creature, the seal of prophet-hood, Muhammad Sallallahu Alaihim Wasallam and on his descendants

I would like to express my appreciation to my supervisor, Associate Prof. Dr Mohd Rizam Abu Bakar, for his excellent supervision, guidance, and words of encouragement. You have been a tremendous mentor during the undertaking of this research.

I will also like to thanks the member of my committee Associate Prof. Dr Leong Wah June and Dr Norfifah Bacok for their valuable contribution, observation and assistance up to the completion of this thesis.

My sincere appreciation also goes to the entire staffs and students of Institute for Mathematical Research for their assistance and hospitality while in Malaysia.

A special thanks to my family, especially my mother for her prayers, love and sacrifices. I would like to express my deepest appreciation to Bauchi state university, Gadau for their support through out the entire process. And I would also like to thanks Dr. Aliyu Usman Moyi, Dr. Muhd Waziri Yusuf , Umar Ali Umar and my loved ones who have supported me in one way or the other throughout the entire process

I certify that a Thesis Examination Committee has met on **20/05/2015** to conduct the final examination of Arzuka Ibrahim on his thesis entitled Scaled Three-Term Conjugate Gradient Method via Davidon-Fletcher-Powell update for unconstrained optimization in accordance with the Universities and University Colleges Act 1971 and the Constitution of the Universiti Putra Malaysia [P.U.(A) 106] 15 March 1998. The Committee recommends that the student be awarded the Master of science.

Members of the Thesis Examination Committee were as follows:

**Zarina Bibi Ibrahim,Phd**
Associate Professor
Faculty of Science
Universiti Putra Malaysia
(Chairperson)

**Mansor Monsi, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Faculty of Science
Universiti Putra Malaysia
(Internal Examiner)

**Name of External Examiner, Ph.D.**
Title (e.g. Professor/Associate Professor/Ir) – Omit if not relevant
Name of Department and/or Faculty
Name of Organisation (University/Institute)
Country
(External Examiner)

**NORITAH OMAR, PhD**
Professor and Deputy Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

This thesis was submitted to the Senate of Universiti Putra Malaysia and has been accepted as fulfilment of the requirement for the degree of Master of Science. The members of the Supervisory Committee were as follows:

**Mohd Rizam Abu Bakar, PhD**
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Chairperson)

**Wah June Leong, PhD**
Associate Professor
Institute for Mathematical Research
Universiti Putra Malaysia
(Member)

**Norfifah Bachok @ Lati, PhD**
Senior Lecturer
Faculty of Science
Universiti Putra Malaysia
(Member)

**BUJANG BIN KIM HUAT, PhD**
Professor and Dean
School of Graduate Studies
Universiti Putra Malaysia

Date:

**Declaration by Graduate Student**

I hereby confirm that:

- this thesis is my original work;
- quotations, illustrations and citations have been duly referenced;
- this thesis has not been submitted previously or concurrently for any other degree at any other institutions;
- intellectual property from the thesis and copyright of thesis are fully-owned by Universiti Putra Malaysia, as according to the Universiti Putra Malaysia (Research) Rules 2012;
- written permission must be obtained from supervisor and the office of Deputy Vice-Chancellor (Research and Innovation) before thesis is published (in the form of written, printed or in electronic form) including books, journals, modules, proceedings, popular writings, seminar papers, manuscripts, posters, reports, lecture notes, learning modules or any other materials as stated in the Universiti Putra Malaysia (Research) Rules 2012;
- there is no plagiarism or data falsification/fabrication in the thesis, and scholarly integrity is upheld as according to the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) and the Universiti Putra Malaysia (Research) Rules 2012. The thesis has undergone plagiarism detection software.

Signature: ______________________ Date: ______________________

Name and Matric No.: ________________________________________

**Declaration by Members of Supervisory Committee**

This is to confirm that:
- the research conducted and the writing of this thesis was under the supervision;
- supervision responsibilities as stated in the Universiti Putra Malaysia (Graduate Studies) Rules 2003 (Revision 2012-2013) are adhered to.

Signature: ____________________
Name of Chairman of Supervisory Committee: ____________________

Signature: ____________________
Name of Member of Supervisory Committee: ____________________

Signature: ____________________
Name of Member of Supervisory Committee: ____________________

# TABLE OF CONTENTS

## LIST OF TABLES

# LIST OF FIGURES

## LIST OF ABBREVIATIONS

| | |
|---|---|
| $f$ | Objective function |
| $\alpha, \alpha_k$ | Steplength (at $k$th iteration) |
| $\Re^n$ | n-dimensional space |
| $x$ | Vector |
| $x_{k+1}, x_k$ | Current and previous point |
| $\nabla f(x_k)$ | Gradient of the $f$ at the point $x_k$ |
| $\nabla^2 f(x_k)$ | Second derivative of $f$ at the point $x_k$ Hessian |
| $x^*$ | Minimizer of |
| $(\|.\|_2)\, x$ | Is the norm of a vector or matrix |
| $x$ | Variable of an optimization problem |
| $I$ | Identity matrix |
| $tr$ | Trace operator |
| NI | Number of iterations |
| NF | Number of function evalution |
| SD | Steepest decent method |
| DFP | Davidon-Fletcher-Powell |
| SR1 | Symmetric Rank One |
| BFGS | Broyden-Fletcher-Goldfarb-Shanno |
| STCG | Scaled three-term Conjugate Method |
| TTHS | Three-term Hestenes-Stiefel |
| TTPRP | Three-term Polaik-Reeve-Polaik |
| TTDFP | Three-term Davidon-Fletcher-Powell |
| MTTPRP | Modified Three-term Polaik-Reeve-Polaik |

# CHAPTER 1

# INTRODUCTION

## 1.1 Introduction

Numerous problems in the real-life situation possess several solutions, Ancona-Navarrete and Tawn (2000) in some cases infinite number of solution may be possible Antoniou and Lu (2007). Optimization is the process of locating the best out of the available solutions, it occurs in various fields of application such as engineering, sciences, social sciences and administration. It trace back to early century and become more independent in the late $1940's$ when G.B Dantzing suggested the famous simplex method. With the advent of digital computers, this area received great attention. Optimization methods are employed in order to obtain the most or the least in a given situation, these methods are iterative rules design to determine the optimal solution of a given optimization problem. In this thesis, we focus on solving nonlinear large-scale unconstrained optimization problem of the form

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

where $f : \Re^n \to \Re$ is a continuously differentiable function and n is the dimension of $x_k$ which is assume to be large. Several iterative schemes were proposed to generate a sequence of approximation to the solution of (1.1). The basic idea behind optimization method is to either minimize or maximize the problem depending upon it nature. Given an initial point $x_0 \in \Re^n$, generate a sequences $\{x_k\}$ of approximation using some iterative scheme such that the iterate $x_k$ moves toward the neighborhood of the optimal value and readily converges when a given stopping criterion is satisfied. In most situation $\|g_k\| < \epsilon$ where $\epsilon$ is pre-determined tolerance, that is when the gradient approaches zero, the approximations $x_k$ of the iterate converges to the optimal solution of the given problem. The sequences of approximations to the solution of (1.1) are generated by the following iterative scheme

$$x_{k+1} = x_k + \alpha_k d_k, \qquad k = 0, 1, 2, 3, ..., \tag{1.2}$$

where $x_{k+1}$ and $x_k$ are the current and previous iterate, $\alpha_k > 0$ is the steplength determine by a line search strategy and $d_k$ is the search direction. Different $d_k$ will result to a different iterative scheme, the search direction are mainly classified into two classes. In the first class we have

$$d_k = -H_k g_k. \tag{1.3}$$

If $(H_k)$ in (1.3) is an identity matrix then (1.3) corresponds to steepest descent direction and the iterative scheme (1.2) become steepest descent method which is one of the earliest method for solving (1.1). If $H_k = \nabla^2 f(x_k)$ in (1.3), then (1.3) corresponds to Newton's direction and the iterative scheme (1.2) becomes the Newton's method. Moreover if $(H_k)$ in (1.3) is an approximation of $\nabla^2 f(x_k)$ then (1.3) corresponds to quasi-Newton direction and the iterative scheme (1.2) becomes quasi-Newton methods which were developed based on the shortcoming of the Newton method, they require the gradient of the objective function to generate and update the inverse Hessian approximate matrix at every step such that the secant condition

$$H_k y_k = s_k, \tag{1.4}$$

is satisfied, with $y_k = g_{k+1} - g_k$ and $s_k = x_{k+1} - x_k$. In the second class, the search direction is given by

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

where the parameter $\beta_k$ is a scalar known as the conjugate gradient parameter and $g_k$ is the gradient of $f$. In this case, the iterative scheme for solving (1.1) corresponds to conjugate gradient method, it is one of the famous iterative scheme in this aspect due to its simplicity, low memory requirement and a good convergences property.

## 1.2 Basic Definitions and Theorems

We begin with the definitions and Theorems that are needed in this thesis which are in Sun and Yuan (2006)

**Definition 1.1** The search direction $d_k$ is said to be a descent direction of the objective function $f$, if $g_k^T d_k < 0$. Similarly, the search direction $d_k$ is said to be sufficient descent direction of the objective function $f$ if

$$g_k^T d_k \leq -c\|g_k\|^2, \tag{1.5}$$

where $c > 0$.

**Definition 1.2** A point $x^* \in \Re^n$ is said to be a stationary point (or a critical point) point of the objective function if $\nabla f(x^*) = 0$.

**Definition 1.3** A point $x^* \in \Re^n$ is said to be a local minimizer of the objective function $f$, if $f(x^*) \leq f(x_k)$ for all $x_k \in \Re^n$.

**Definition 1.4** A point $x^* \in \Re$ is said to be global minimizer of the objective function $f$, if $f(x^*) \leq f(x_k)$ for all $x_k \in \Re^n$ and $x_k \neq x^*$.

**Definition 1.5** A point $x^* \in \Re^n$ is said to be a strict minimizer of the objective function $f$, if $f(x^*) < f(x_k)$ for all $x_k \in \Re^n$.

**Lemma 1.1** Let U and V be matrices belonging to $\Re^{n\times m}$ for some $m \in [1, n]$ and suppose that $L(\Re^n)$ defined the linear space of all matrices of order n. The rank-k update of matrix that is nonsingular $A \in L(\Re^n)$ of the form

$$\hat{A} = A + UV^T$$

is nonsingular if and only if $\kappa = I + V^T A^{-1} U \neq \mathbf{0}$ specifically if $\kappa \neq \mathbf{0}$, then, the inverse of the matrix A is given by

$$\hat{A}^{-1} = A^{-1} - A^{-1}U(I + V^T A^{-1} U)^{-1} V^T A^{-1}, \tag{1.6}$$

which is known as Sharman-Morrison formular (see Nocedal and Wright (2006)).

**Theorem 1.1** Let $P \in \Re^n$ be a non empty open convex set and let $f : P \in \Re^n \rightarrow \Re$ be a differentiable function, then $f$ is convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^T (y - x)$$

for all $x, y \in P$. Similarly, f is strictly convex on P if and only if

$$f(y) > f(x) + \nabla f(x)^T (y - x)$$

for all $x, y \in P$. Further more f is strongly or uniformly convex if and only if

$$f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{1}{2} c \|y - x\|^2$$

for all $x, y \in P$ where $c > 0$ is a constant.

Proof omitted see (Sun and Yuan (2006))

## 1.3 Optimality Conditions

The following Theorem (1.2-1.4) are due to (Sun and Yuan (2006))

**Theorem 1.2** *(First Order Necessary Condition)* Let $f : U \in \Re^n \rightarrow \Re$ be continuously differentiable on an open set U, if $x^* \in \Re^n$ is a local minimizer then $\nabla f(x^*) = 0$.

**Theorem 1.3** *(Second Order Necessary Condition)* Let $f : U \in \Re^n \rightarrow R$ be continuously differentiable on an open set U, if $x^* \in U$ is a local minimizer then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive -semi definite.

**Theorem 1.4** *(Second Order Sufficient Condition)* Let $U \in \Re^n \rightarrow \Re$ be continuously differentiable on an open set U, if $x^* \in U$ is a strict local minimizer then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite.

## 1.4 Rate of Convergence

The rate of convergence refers to the speed in which the sequence generated by an algorithm tends to its limit point (Nocedal and Wright (2006). The following definition (1.6-1.4) are from (Nocedal and Wright (2006)

**Definition 1.6** Let the sequence $\{x_k\}$ in $\Re^n$ converge to a finite limit $x^*$.

- The sequence is said to be $q$-linear convergence if there exist a scalar $r \in (0, 1)$ such that
$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} \leq r$$
for sufficiently large k.

- The convergence is said also to be $q$-superlinear if
$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|} = 0$$

- The convergence is said to be $q$-quadratic if there exist a scalar $N$ such that
$$\frac{\|x_{k+1} - x_*\|}{\|x_k - x_*\|^2} \leq N$$
for a very large $k$.

**Definition 1.7** Let the sequence $\{x_k\}$ in $\Re^n$ converge to a finite limit $x^*$. The sequence convergence is said to be $r$-(linear, superlinearly quadratically) if there exist a sequence converging $q$-(linear, superlinearly quadratically) to zero.

### 1.5 Motivation

Among the quasi-Newton methods is the DFP method which serve as one of the clever optimization method for determining the inverse Hessian approximation, its encompasses the merits of Newton's method and steepest descent method while avoiding their disadvantages Goldfarb (1969). It been neglected over the years as the result of

1. Computation and storage of the Hessian or inverse Hessian approximation in every iteration

2. Its inability to adjust within a fewer step whenever a less informative inverse Hessian matrix is approximated (Nocedal and Wright (2006)).

On the other hand, the conjugate gradient methods are fascinating iterative scheme with low memory requirement. For non-linear problems most conjugate gradient method cannot guarantee the sufficient descent condition
$$g_k^T d_k \leq -c\,\|g_k\|$$
which facilitate it convergence (Gilbert and Nocedal (1992)).

## 1.6 Objectives of the Thesis

The main objectives are

1. To develop some three-term conjugate gradient method via DFP update which would satisfies both the sufficient and the conjugacy conditions

2. To establish the global convergence properties of the proposed methods

3. To present some numerical results by comparing the proposed method with other existing conjugate gradient methods.

4. To apply the proposed method on real-life optimization problems.

## 1.7 Structure of the Thesis

The remaining part of this thesis is structure as: An overview of nonlinear unconstrained optimization methods is present in chapter 2. Chapter 3 deals with the derivation of the proposed method STCG, the convergence properties, numerical result and discussion. In chapter 4 we present the real-life application of the proposed method on statistical problems. Chapter 5 present the implementation of the proposed method with nonmonotone line search and chapter 6 gives conclusion and future research.

## REFERENCES


Al-Baali, M. 1985. Descent property and global convergence of the Fletcher—Reeves method with inexact line search. *Journal of Numerical Analysis* 5 (1): 121–124.

Al-Bayati, A. Y. and Altae, W. 2010. A new three-term non-linear conjugate gradient method for unconstrained optimization. *Canadian Journal on Science & Engineering Mathematics* 1 (5): 108–124.

Amini, K., Ahookhosh, M. and Nosratipour, H. 2014. An inexact line search approach using modified nonmonotone strategy for unconstrained optimization. *Numerical Algorithms* 66 (1): 49–78.

Ancona-Navarrete, M. A. and Tawn, J. A. 2000. A comparison of methods for estimating the extremal index. *Extremes* 3 (1): 5–38.

Andrei, N. 2006. An acceleration of gradient descent algorithm with backtracking for unconstrained optimization. *Numerical Algorithms* 42 (1): 63–73.

Andrei, N. 2007. Scaled conjugate gradient algorithms for unconstrained optimization. *Computational Optimization and Applications* 38 (3): 401–416.

Andrei, N. 2008a. Another hybrid conjugate gradient algorithm for unconstrained optimization. *Numerical Algorithms* 47 (2): 143–156.

Andrei, N. 2008b. An unconstrained optimization test functions collection. *Advance. Modelling. Optimimization* 10 (1): 147–161.

Andrei, N. 2009. Acceleration of conjugate gradient algorithms for unconstrained optimization. *Applied Mathematics and Computation* 213 (2): 361–369.

Andrei, N. 2011. A modified Polak-Ribière-Polyak conjugate gradient algorithm for unconstrained optimization. *Optimization* 60 (12): 1457–1471.

Andrei, N. 2013. On three-term conjugate gradient algorithms for unconstrained optimization. *Applied Mathematics and Computation* 219 (11): 6316–6327.

Antoniou, A. and Lu, W.-S. 2007. *Practical optimization: algorithms and engineering applications*. Springer.

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

Babaie-Kafaki, S. 2013. A modified scaled memoryless BFGS preconditioned conjugate gradient method for unconstrained optimization. *4OR* 11 (4): 361–374.

Bard, Y. 1968. On a numerical instability of Davidon-like methods. *Mathematics of Computation* 22 (103): 665–666.

Barzilai, J. and Borwein, J. M. 1988. Two-point step size gradient methods. *Journal of Numerical Analysis* 8 (1): 141–148.

Beale, E. M. L. 1972. A derivation of conjugate gradients. *Numerical methods for Nonlinear Optimization* 39–43.

Birgin, E. G. and Martínez, J. M. 2001. A spectral conjugate gradient method for unconstrained optimization. *Applied Mathematics and Optimization* 43 (2): 117–128.

Broyden, C. 1967. Quasi-Newton methods and their application to function minimisation. *Mathematics of Computation* 368–381.

Broyden, C. G. 1970. The convergence of a class of double-rank minimization algorithms 2. The new algorithm. *Journal of Applied Mathematics* 6 (3): 222–231.

Byrd, R. H. and Nocedal, J. 1989. A tool for the analysis of quasi-Newton methods with application to unconstrained minimization. *Journal on Numerical Analysis* 26 (3): 727–739.

Cauchy, A. 1847. Méthode générale pour la résolution des systemes d'équations simultanées. *Comp. Rend. Sci. Paris* 25 (1847): 536–538.

Conn, A. R., Gould, N. I. and Toint, P. L. 1991. Convergence of quasi-Newton matrices generated by the symmetric rank one update. *Mathematical Programming* 50 (1-3): 177–195.

Contreras, M. and Tapia, R. 1993. Sizing the BFGS and DFP updates: numerical study. *Journal of Optimization Theory and Applications* 78 (1): 93–108.

Dai, Y. and Yuan, Y. 2001. A three-parameter family of nonlinear conjugate gradient methods. *Mathematics of Computation* 70 (235): 1155–1167.

Dai, Y. and Yuan, Y. 2003. A class of globally convergent conjugate gradient methods. *Science in China Series A: Mathematics* 46 (2): 251–261.

Dai, Y. H. 1997. *Analyses of conjugate gradient methods*. , vol. 116. Springer.

Dai, Y. H. and Liao, L. Z. 2001. New conjugacy conditions and related nonlinear conjugate gradient methods. *Applied Mathematics and Optimization* 43 (1): 87–101.

Dai, Y. H. and Yuan, Y. 1999. A nonlinear conjugate gradient method with a strong global convergence property. *Journal on Optimization* 10 (1): 177–182.

Davidon, W. C. 1991. Variable metric method for minimization. *Journal on Optimization* 1 (1): 1–17.

Dennis, J. and More, J. J. 1977. Quasi-Newton methods, motivation and theory. *Review* 19 (1): 46—-89.

Dolan, E. D. and More, J. J. 2002. Benchmarking optimization software with performance profiles. *Mathematical Programming* 91 (2): 201–213.

Fletcher, R. 1970. A new approach to variable metric algorithms. *The Computer Journal* 13 (3): 317–322.

Fletcher, R. 1987. *Practical Methods of Optimization*. Wiley.

Fletcher, R. 2013. *Practical methods of optimization*. John Wiley & Sons.

Fletcher, R. and Powell, M. J. 1963. A rapidly convergent descent method for minimization. *The Computer Journal* 6 (2): 163–168.

Fletcher, R. and Reeves, C. M. 1964. Function minimization by conjugate gradients. *The Computer Journal* 7 (2): 149–154.

Gilbert, J. C. and Nocedal, J. 1992. Global convergence properties of conjugate gradient methods for optimization. *Journal on Optimization* 2 (1): 21–42.

Gill, P. and Murray, W. 1972. Quasi-Newton methods for unconstrained optimization. *Journal of Applied Mathematics* 9 (1): 91–108.

Gill, P. E., Golub, G. H., Murray, W. and Saunders, M. A. 1974. Methods for modifying matrix factorizations. *Mathematics of Computation* 28 (126): 505–535.

Goldfarb, D. 1969. Extension of Davidon's variable metric method to maximization under linear inequality and equality constraints. *Journal on Applied Mathematics* 17 (4): 739–764.

Goldfarb, D. 1970. A family of variable-metric methods derived by variational means. *Mathematics of Computation* 24 (109): 23–26.

Goldfeld, S. M., Quandt, R. E. and Trotter, H. F. 1966. Maximization by quadratic hill-climbing. *Journal of the Econometric Society* 541–551.

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

Grippo, L., Lampariello, F. and Lucidi, S. 1986. A nonmonotone line search technique for Newton's method. *Journal on Numerical Analysis* 23 (4): 707–716.

Hager, W. W. and Zhang, H. 2005. A new conjugate gradient method with guaranteed descent and an efficient line search. *Journal on Optimization* 16 (1): 170–192.

Hager, W. W. and Zhang, H. 2006. A survey of nonlinear conjugate gradient methods. *Pacific Journal of Optimization* 2 (1): 35–58.

Hassan, M. A., Leong, W. J. and Farid, M. 2009. A new gradient method via quasi-Cauchy relation which guarantees descent. *Journal of Computational and Applied Mathematics* 230 (1): 300–305.

Hestenes, M. R. 1956. The conjugate gradient method for solving linear systems. In *Proceeding. Symposium. Application. Math VI, American Mathematical Society*, 83—-102.

Khalfan, H. F., Byrd, R. H. and Schnabel, R. B. 1993. A theoretical and experimental study of the symmetric rank-one update. *Journal on Optimization* 3 (1): 1–24.

Kutner, M. H., Nachtsheim, C. and Neter, J. 2004. *Applied linear regression models*. McGraw-Hill/Irwin.

Leong, W. J. and Hassan, M. A. 2009. A restarting approach for the symmetric rank one update for unconstrained optimization. *Computational Optimization and Applications* 42 (3): 327–334.

Leong, W. J. and San Goh, B. 2013. Convergence and Stability of Line Search Methods for Unconstrained Optimization. *Acta Applicandae Mathematicae* 127 (1): 155–167.

Li, G., Tang, C. and Wei, Z. 2007. New conjugacy condition and related new conjugate gradient methods for unconstrained optimization. *Journal of Computational and Applied Mathematics* 202 (2): 523–539.

Liao, A. 1997. Modifying the BFGS method. *Operations Research Letters* 20 (4): 171–177.

Liu, J. and Feng, Y. 2011. Global convergence of a modified LS nonlinear conjugate gradient method. *Procedia Engineering* 15: 4357–4361.

Liu, Y. and Storey, C. 1991. Efficient generalized conjugate gradient algorithms, part 1: theory. *Journal of Optimization Theory and Applications* 69 (1): 129–137.

Livieris, I. E. and Pintelas, P. 2013. A new class of spectral conjugate gradient methods based on a modified secant equation for unconstrained optimization. *Journal of Computational and Applied Mathematics* 239: 396–405.

Luenberger, D. G. and Ye, Y. 2008. *Linear and nonlinear programming*. , vol. 116. Springer.

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

Nazareth, J. L. 2001. *Conjugate-gradient Methods*. Springer.

Nazareth, L. 1977. A conjugate direction algorithm without line searches. *Journal of Optimization Theory and Applications* 23 (3): 373–387.

Nocedal, J. 1996. Conjugate gradient methods and nonlinear optimization. *Linear and Nonlinear Conjugate Gradient Related Methods* 9–23.

Nocedal, J. and Wright, S. 2006. *Numerical optimization, series in operations research and financial engineering*. Springer, New York.

Oren, S. S. and Luenberger, D. G. 1974. Self-Scaling Variable Metric (SSVM) Algorithms: Part I: Criteria and Sufficient Conditions for Scaling a Class of Algorithms. *Management Science* 20 (5): 845–862.

Pearson, J. D. 1969. Variable metric methods of minimisation. *The Computer Journal* 12 (2): 171–178.

Perry, J. M. 1977. A class of conjugate gradient algorithms with a two step variable metric memory. *Center for Mathematical Studies in Economies and Management Science* .

Polak, E. and Ribiere, G. 1969. Note sur la convergence de méthodes de directions conjuguées. *Mathematical Modelling and Numerical Analysis-Modélisation Mathématique et Analyse Numérique* 3 (R1): 35–43.

Powell, M. J. D. 1977a. Restart procedures for the conjugate gradient method. *Mathematical Programming* 12 (1): 241–254.

Powell, M. J. D. 1977b. Restart procedures for the conjugate gradient method. *Mathematical Programming* 12 (1): 241–254.

Powell, M. J. D. 1984. *Nonconvex minimization calculations and the conjugate gradient method*. Springer.

Raydan, M. 1997. The Barzilai and Borwein gradient method for the large scale unconstrained minimization problem. *Journal on Optimization* 7 (1): 26–33.

Shanno, D. F. 1970. Conditioning of quasi-Newton methods for function minimization. *Mathematics of Computation* 24 (111): 647–656.

Shanno, D. F. 1978. Conjugate gradient methods with inexact searches. *Mathematics of Operations Research* 3 (3): 244–256.

Shi, Z.-J. and Shen, J. 2006. Convergence of nonmonotone line search method. *Journal of Computational and Applied Mathematics* 193 (2): 397–412.

Sun, W. and Yuan, Y.-X. 2006. *Optimization theory and methods: nonlinear programming*. , vol. 1. Springer Science & Business Media.

Touati-Ahmed, D. and Storey, C. 1990. Efficient hybrid conjugate gradient techniques. *Journal of Optimization Theory and Applications* 64 (2): 379–397.

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

Wolkowicz, H. 1994. Measures for symmetric rank-one updates. *Mathematics of Operations Research* 19 (4): 815–830.

Xiao, Y., Song, H. and Wang, Z. 2012. A modified conjugate gradient algorithm with cyclic Barzilai–Borwein steplength for unconstrained optimization. *Journal of Computational and Applied Mathematics* 236 (13): 3101–3110.

Yu, G., Guan, L. and Chen, W. 2008. Spectral conjugate gradient methods with sufficient descent property for large-scale unconstrained optimization. *Optimization Methods and Software* 23 (2): 275–293.

Yuan, G. 2009. Modified nonlinear conjugate gradient methods with sufficient descent property for large-scale optimization problems. *Optimization Letters* 3 (1): 11–21.

Yuan, G. and Wei, Z. 2009. New line search methods for unconstrained optimization. *Journal of the Korean Statistical Society* 38 (1): 29–39.

Yuan, Y. X. 1993. Analysis on the conjugate gradient method. *Dynamical Systems* 2 (1): 19–29.

Yuhong, D. and Yaxiang, Y. 1996. Convergence Properties of the Conjugate Descent Method. *Advances in Mathematics* 06.

Yuhong, D. A. I. 2002. A nonmonotone conjugate gradient algorithm for unconstrained optimization. *Journal of Systems Science and Complexity* 15 (2): 139–145.

Zhang, J., Xiao, Y. and Wei, Z. 2009. Nonlinear conjugate gradient methods with sufficient descent condition for large-scale unconstrained optimization. *Mathematical Problems in Engineering* 2009.

Zhang, J. and Xu, C. 2001. Properties and numerical performance of quasi-Newton methods with modified quasi-Newton equations. *Journal of Computational and Applied Mathematics* 137 (2): 269–278.

Zhang, J. Z., Deng, N. Y. and Chen, L. H. 1999. New quasi-Newton equation and related methods for unconstrained optimization. *Journal of Optimization Theory and Applications* 102 (1): 147–167.

Zhang, L., Zhou, W. and Li, D. 2006a. Global convergence of a modified Fletcher–Reeves conjugate gradient method with Armijo-type line search. *Numerische Mathematik* 104 (4): 561–572.

Zhang, L., Zhou, W. and Li, D. 2007a. Global convergence of the DY conjugate gradient method with Armijo line search for unconstrained optimization problems. *Optimisation Methods and Software* 22 (3): 511–517.

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

Zhang, L., Zhou, W. and Li, D.-H. 2006b. A descent modified Polak–Ribière–Polyak conjugate gradient method and its global convergence. *Journal of Numerical Analysis* 26 (4): 629–640.

Zoutendijk, G. 1970. Nonlinear programming, computational methods. *Integer and Nonlinear Programming* 143 (1): 37–86.