EFFICIENCY OF SUBGRADIENT METHOD IN SOLVING NONSMOOTTH OPTIMIZATION PROBLEMS

NUR AZIRA BINTI ABDULLAH

A dissertation submitted in partial fulfilment of the requirements for the award of the degree of Master of Science (Mathematics)

Faculty of Science
Universiti Teknologi Malaysia

JANUARY 2014

In memory of my beloved father, you will always remain in our hearts

To my beloved mother and sister

Puan Adnan binti Ismail and Nur Azwani binti Abdullah

For your love, endless support and encouragement.

ACKNOWLEDGEMENT

Assalamualaikum warahmatullahi wabarakatuh…

I would like to thank all those who have contributed in so many ways to the completion of this report. Firstly I acknowledge my great supervisor, PM DR. Rohanin binti Ahmad who gave me the opportunity to do this wonderful project on the topic 'Efficiency of Subgradient Method in Solving Nonsmooth Optimization Problems'. I am extremely grateful for her excellent guidance, supervision, suggestions, and efforts throughout the duration of this study

Secondly, I am most indebted to my mother and sister who helped me, encouraged me and supported me each step of the way in finishing this study. I would like to acknowledge their valuable time and advice as well as their intellectual support and encouragement throughout my academic journey.

I would like to express my sincere appreciation to my friends who help me a lot in this study. Without their help, advice and contributions, the study would not have been completed. Lastly, I would like to thank again all who helped me.

ABSTRACT

Nonsmooth optimization is one of the hardest type of problems to solve in optimization area. This is because of the non-differentiability of the function itself. Subgradient method is generally known as the method to solve nonsmooth optimization problems, where it was first discovered by Shor [17]. The main purpose of this study is to analyze the efficiency of subgradient method by implementing it on these nonsmooth problems. This study considers two different types of problems, the first of which is Shor's piecewise quadratic function type of problem and the other is L1-regularized form type of problems. The objectives of this study are to apply the subgradient method on nonsmooth optimization problems and to develop matlab code for the subgradient method and to compare the performance of the method using various step sizes and matrix dimensions. In order to achieve the result, we will use matlab software. At the end of this study we can conclude that subgradient method can solve nonsmooth optimization problems and the rate of convergence varies based on the step size used.

ABSTRAK

Pengoptimuman tak licin adalah salah satu jenis masalah yang paling sukar untuk diselesaikan di dalam bidang pengoptimuman. Ini adalah kerana ketidakbolehan pembezaan oleh fungsi itu sendiri. Kaedah subkecerunan diketahui umum sebagai kaedah untuk menyelesaikan masalah pengoptimuman tak licin, yang pertama kali ditemui oleh Shor [17]. Tujuan utama kajian ini adalah untuk menganalisis keberkesanan kaedah subkecerunan dengan mengaplikasikannya kepada masalah-masalah tak licin ini. Kajian ini mempertimbangkan dua jenis masalah, di mana yang pertama adalah masalah Shor fungsi kuadratik jenis cebis demi cebis, dan yang kedua ialah masalah berbentuk nalar-L1. Objektif kajian ini adalah untuk mengaplikasikan kaedah subkecerunan pada masalah pengoptimuman yang tak licin dan untuk membina kod matlab bagi kaedah subkecerunan serta untuk membandingkan prestasi kaedah ini dengan menggunakan saiz langkah dan dimensi matrik yang berbeza. Dalam usaha untuk mencapai keputusan, kita akan menggunakan perisian matlab. Pada akhir kajian ini, kita boleh membuat kesimpulan bahawa kaedah subkecerunan ini boleh menyelesaikan masalah pengoptimuman tak licin dan kadar penumpuan berbeza berdasarkan saiz langkah yang digunakan.

TABLE OF CONTENTS

LIST OF FIGURE

LIST OF SYMBOLS

g	-	Subgradient of function.
K	-	Lipshitz constant.
α_k	-	Step size for function.
Z	-	Upper bound for all subgradients.
S	-	Upper bound on distance of the initial point to the optimal set.

CHAPTER 1

INTRODUCTION

1.1 Introduction

Optimization is one of the areas of applied mathematics and it has been widely used in many areas such as control theory, optimal control, engineering, and economics. Optimization is about finding the largest or smallest values which can be attained by functions of several real variables subject to constraints [1]. Example of optimization problems are:

a) An architect wants to design a house of 460 million square feet and will be divided one third of it to build a playground for his children. How can he do this so as to minimize the cost of all the materials?
b) How to maximize the interest you get on your money once you invest it?

Optimization can be categorized into many different problems, such as linear, nonlinear, discrete, continuous, stochastic, smooth and nonsmooth optimization problems. Here we will explain briefly about the stated types of optimization problems above. Linear optimization which also known as linear programming (LP) involve linear objective function, subject to the linear equality and inequality constraints [2]. Most of the practical problems found in operational research are recognized as LP problems. It was first developed by the Leonid Kantorovich in 1939 for use during World War II. Whereas nonlinear optimization is the opposite of the LP problem where some of the objective functions and constraints are nonlinear. Examples of nonlinear optimization are problems which involve quadratic and cubic functions in the objective function.

In continuous optimization, the variables in the problems usually are from continuous range of values, usually real numbers [3]. This characteristic differentiates continuous optimization from discrete optimization in which the variables may be binary (limited to the values 0 and 1) and only integer allowable. Other than that, stochastic optimization is the process to find the best (maximum or minimum) optimal value of mathematical or statistical function which involves randomness [4].

A function is smooth if it is differentiable and the derivatives are continuous and more specifically, this is first-order smoothness. Second-order smoothness means that second derivatives exist and are continuous, and so forth, while infinite smoothness refers to continuous derivatives of all orders [5]. As opposed to the smooth optimization, nonsmooth optimization (NSO) refers to the more general problem of minimizing functions that are typically not differentiable at their minimizers. [6]. Some of the methods that are well known to solve nonsmooth optimization include subgradient method, cutting planes method, bundle method, trust region method and proximal point method [7]. In this study, we will be focusing on how to solve the NSO problems.

1.2 Background of the Study

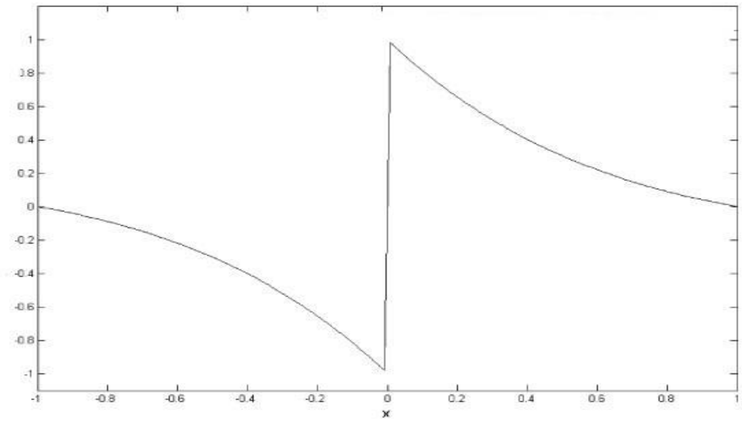

Figure 1.1: Nonlinear nonsmooth function $|x|e^{-|x|}$.

Nonsmooth, optimization (NSO) looks at problems where the functions involved are not continuously differentiable [8]. Figure 1.1 is one example of the nonsmooth function involved in NSO. The graph shows that, there are some regions of the function which is not differentiable as the graph has a corner or kink at $x=0$ and thus cannot be solved by classical nonlinear programming algorithm because the derivative information cannot be used to find the direction in which the function is increasing or decreasing. Although the function is not differentiable at the corner, it is still continuous at that point. Because of the very little information we get from the function, there will be some modification in the original method in order to derive a method that can solve such optimization problem involving nonsmooth functions. Most modification is at the derivative part of the algorithm where the direction can be estimated.

This study concentrates on nonsmooth optimization where we only consider subgradient method to solve the optimization problems. Subgradient method is a direct generalization of the steepest descent method, which generates a sequence by using $-g^k$ as a direction [9]. Steepest descent is the basic algorithm to solve nonlinear optimization and it is one of the methods that have gradient-type algorithm.

Since nonsmooth optimization involves nondifferentiable function, thus steepest descent cannot solve this type of problem because it requires information from the derivative of the function. As indicated by Wolfe [10] when the function is nondifferentiable, if one uses the steepest descent method with line search to solve the problem, it is possible to generate a sequence which most likely will converge to a non stationary point. One of the differences between the subgradient method and the steepest descent method is in the rule for finding the step size.

An extension of the subgradient method is the conjugate subgradient method which was developed by Wolfe [10] and it is similar to the well-established conjugate gradient method for solving quadratic optimization problems. However Wolfe's method remained practically unused and a method has been extended from it, which is the Bundle method. Bundle methods are at the moment the most efficient and promising methods for nonsmooth optimization [11]. However, in this study we will focus on the subgradient method only.

1.3 Statement of the Problem

This study proposes to solve the optimization problems involving nonsmooth functions using the subgradient method.

1.4 Objectives of the Study

The objectives of this study are:

1. To apply the subgradient method on NSO problem.
2. To develop matlab code for the subgradient method.
3. To compare the performance of the method using various step sizes and matrix dimensions.

1.5 Scope of the Study

We will limit our study to unconstrained Nonsmooth Optimization (NSO) problems where the function is convex and locally Lipschitz continuous, but not continuously differentiable. To solve such problem we use subgradient method and the problems is unconstrained type of problems.

1.6 Significance of the Study

This study contributes knowledge to better understand and solve NSO problem by using subgradient methods. Besides that, mathematicians can further investigate the modification of the present methods to have better method to solve the more difficult problems in the optimization areas.

1.7 Outline of the Study

This dissertation consists of five chapters. The arrangement of the content of this study is as follows.

Chapter 1: In this chapter we briefly discuss the general introduction of this study which consists of introduction, background of the study, statement of the problem, objectives of the study, scope of the study, and significance of the study.

Chapter 2: In Chapter 2, we will discuss the literature review on the topics selected. Most of the topics involve are nonsmooth optimization and subgradient method.

Chapter 3: In this chapter, we discuss and describe the method that we will use throughout this study in order to solve the nonsmooth problems.

Chapter 4: Here we implementing the method that we have discussed on the selected nonsmooth problems. We discuss in detail the nonsmooth problems, and the results obtained.

Chapter 5: This chapter is the summary and conclusion of the result of this study. Some recommendations will be given for future work.

1.8 Summary

In this chapter we discuss the background of the study, statement of the problem, the objectives of the study, scope of the study and the significance of the study. Chapter 1 also provides outline of this dissertation for the reader to overview the flow of this study. In next chapter we will discuss the literature review involve in this study, such as the nonsmooth optimization and history of subgradient method.

REFERENCES

1. Kiwiel, K.C. Methods of Descent for Nondifferentiable Optimization. *Volume 1133 of Lecture Notes in Mathematics*. Pennsylvania State University. 1985.
2. Luenberger, D. G. and Ye, Y. Linear and Nonlinear Programming. *International Series in Operations Research & Management Science, Springer*. Volume 116, 2008.
3. Wright, S. J. *Continuous Optimization (Nonlinear And Linear Programming)*. Computer Sciences Department, University Of Wisconsin, Madison, Wisconsin, Usa.
4. Fouskakis, D. and Draper, D. *Stochastic Optimization: A Review, International Statistical Review*. The Netherlands. 2002. 315-349.
5. Rockafellar, T. Nonsmooth Optimization. *Mathematical Programming: State Of The Art.* 1994. 248–258.
6. Erfanian, H. R., Skandari, M. H. N., and Kamyad, A. V. A Novel Approach For Solving Nonsmooth Optimization Problems With Application To Nonsmooth Equation. *Journal of Mathematics*. Volume 2013.
7. Ruszcynki, A. *Nonlinear Optimization*. Princeton University Press. 2011.
8. Goffin, J. L. Subgradient Optimization In Nonsmooth Optimization (Including The Soviet Revolution). *Mathematics Subject Classification*. 2010.
9. Sun, W. and Yuan, Y. X. *Optimization Theory And Methods: Nonlinear Programming*. 2006.
10. Wolfe, P. A Method of Conjugate Subgradients for Minimizing Nondifferentiable Functions. *Math. Prog. Study 3*. 1975. 145-173
11. Mäkelä, M. Survey of Bundle Methods for Nonsmooth Optimization, *Optimization Methods and Software* . 2002.

12. Clarke, F. *Nonsmooth Analysis in Systems and Control Theory*. Institute Universitaire De France Et Université De Lyon. 2008.

13 Karmitsa, N. *Nonsmooth Optimization In 30 Minutes*. Department of Mathematics and Statistics, University Of Turku, Finland. 2013.

14. Sagastizábal, C. and Solodov, M. An Infeasible Bundle Method for Nonsmooth Convex Constrained Optimization without a Penalty Function or a Filter. *SIAM Journal on Optimization*. Volume 16, Issue 1. 2005. 146 - 169

15. Meza, J. C. *Steepest Descent*. Lawrence Berkeley National Laboratory: Lawrence Berkeley National Laboratory. LBNL Paper LBNL-3395E. 2010.

16. Shor, N. Z. Nondifferentiable Optimization And Polynomial Problems. *Nonconvex Optimization And Its Applications*. Volume 24, Springer. 1998.

17. Shor, N. Z. Minimization Methods For Non-Differentiable Functions. *Springer Series In Computational Mathematics*. 1985.

18. Held, M. and Karp, R. M. The Traveling-Salesman Problem And Minimum Spanning Trees: Part II. *Mathematical Programming 1*. 1971. 6–25.

19. Boyd, S., Xiao, L., and Mutapcic, A. *Subgradient Methods*. Stanford University, Autumn. 2003.

20. S. Boyd and Mutapcic, A. *Subgradient Methods*. Stanford University. 2006-07.

21. Camerini, P.M., Fratta, L., and Maffioli, F. On Improving Relaxation Methods By Modified Gradient Techniques. *Math. Prog. Study 3*. 1975. 26-34

22. Bagirov, A. M., Karasözen, B., and Sezer, M. Discrete Gradient Method: Derivative-Free Method for Nonsmooth Optimization, *Journal of Optimization Theory and Applications*. Volume 137, Issue 2. 2008. 317-334

23. Nedi´c, A. and Ozdaglar, A. Subgradient Methods in Network Resource Allocation: Rate Analysis. *In 42nd Annual Conference on Information Sciences and Systems*. CISS, Princeton. 2008. 1189-1194

24. Nedi´c, A. and Bertsekas, D. P. Incremental Subgradient Methods for Nondifferentiable Optimization. *SIAM Journal on Optimization*. Volume 12 Issue 1. 2001. 109-138

25. Ratliff, N., Bagnell, A., and Zinkevich,M. (Online) subgradient methods for structured prediction. 2007.

26. Dixon, L., Spedicato, E., and Szego, G. *Nonlinear Optimization, Theory And Algorithms*, Springer. Bikhauser, Boston. 1980.

27. Bonnans, J. F. *Numerical Optimization: Theoretical And Practical Aspects: With 26 Figures*, 2003.

28. Lemarechal, C. and Mifflin, R. *Nonsmooth optimization*, Proceedings of a IIASA Workshop. 1997.

29. Luksan, L. and Vlcek, J. *Test Problems For Nonsmooth Unconstrained and Linearly Constrained Optimization*, Institute of Computer Science , Academy of Science of the Czech Republic. 2000.