SPRINGER BRIEFS IN OPTIMIZATION

Xin-She Yang
Xing-Shi He

Mathematical Foundations of Nature-Inspired Algorithms

Springer

SpringerBriefs in Optimization

SpringerBriefs in Optimization showcases algorithmic and theoretical techniques, case studies, and applications within the broad-based field of optimization. Manuscripts related to the ever-growing applications of optimization in applied mathematics, engineering, medicine, economics, and other applied sciences are encouraged.

More information about this series at http://www.springer.com/series/8918

Xin-She Yang • Xing-Shi He

Mathematical Foundations of Nature-Inspired Algorithms

Springer

Xin-She Yang
School of Science and Technology
Middlesex University
London, UK

Xing-Shi He
College of Science
Xi'an Polytechnic University
Xi'an, China

ISSN 2190-8354 ISSN 2191-575X (electronic)
SpringerBriefs in Optimization
ISBN 978-3-030-16935-0 ISBN 978-3-030-16936-7 (eBook)
https://doi.org/10.1007/978-3-030-16936-7

Mathematics Subject Classification: 80M50, 68Q25, 68W20, 90C26, 90C31, 46N10, 97N40, 60J10, 37N40

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Optimization is everywhere, from engineering design and business planning to data mining and industrial applications. However, many optimization problems are very challenging to solve, especially highly nonlinear problems, combinatorial problems and large-scale problems. Traditional optimization struggles to cope such tough problems. In recent years, nature-inspired algorithms have become popular for solving tough problems concerning optimization, data mining and machine learning. Despite their wide applications, their mathematical analysis is weak, and, for most new algorithms, there is no analysis at all. Even if there exist some fragmental theoretical studies, it lacks a unified theoretical framework for analysing such algorithms.

This book will attempt to provide a systematic approach to rigorous mathematical analysis of these algorithms based on solid mathematical foundations, and the analyses will be from different angles and perspectives. There are many good textbooks on the detailed introduction of traditional optimization techniques and new algorithms. In contrast, the introduction of algorithms in this book will be relatively brief and will introduce the most relevant algorithms for the purpose of analysing and understanding nature-inspired algorithms. Thus, the introduction of algorithms includes traditional algorithms such as gradient-based methods and new algorithms such as swarm intelligence-based algorithms. The emphasis will be placed on the latest algorithms and their analyses from a wide spectrum of theories and frameworks. Thus, readers can gain greater insight into the main characteristics of different algorithms and understand how and why they work for solving optimization problems in real-world settings. More specifically, in-depth mathematical analyses will be carried out from different perspectives, including complexity theory, fixed-point theory, dynamical system, self-organization, Bayesian framework, Markov chain framework, filter theory, statistical learning and statistical measures. Such mathematical analyses will provide some insight into relevant algorithms and thus allow readers to see the advantages and disadvantages of different algorithms. As a result, it provides further insight and understanding on the performance of each class of algorithms and thus the proper choice of relevant algorithms for a given class of optimization problems to be solved.

It is worth pointing out that there are still many open problems, despite the efforts by all the researchers in this active research area. Our aims here are twofold: to summarize and to highlight. One main intention is to summarize all the major approaches for analysing algorithms from at least ten different angles and perspectives. The other intention is to highlight the key issues and open problems so as to encourage and inspire more research in this area. We hope that more understanding and insight will be gained by new studies and research in the near future.

Therefore, this brief and timely book can serve as a textbook for many optimization courses at both advanced undergraduate and graduate levels. It can also serve as a timely reference for lecturers, graduates and researchers in optimization, operations research, artificial intelligence, data mining, machine learning, computer science and management sciences.

London, UK
Xi'an, China
January 2019

Xin-She Yang
Xing-Shi He

Contents

About the Authors

Xin-She Yang is a reader/professor at the School of Science and Technology, Middlesex University London. He is also an elected bye-fellow and supervisor at Cambridge University and a guest chief scientist and distinguished professor at Xi'an Polytechnic University, China. He worked at the National Physical Laboratory as a senior research scientist after obtaining his DPhil in Applied Mathematics at Oxford University. With more 250 publications and more than 30 books, his research has been cited more than 36,000 times (according to Google Scholar) with an h-index of 71. He has been on the prestigious lists of Highly Cited Researchers (2016, 2017 and 2018), according to Web of Science/Clarivate Analytics. He is the chair of IEEE CIS Task Force on Business Intelligence and Knowledge Management. He has given many invited keynote talks at international conferences such as ICCS2015 (Iceland), SIBGRAPI2015 (Brazil), IEEE OIPE2016 (Italy), BDIOT2017 (UK), ICIST2018(UK), CPMMI2019 (Serbia), ICAAI2018 (Spain) and LION2019 (Greece). He has also given tutorials at international conferences such as FedCSIS2011 (Poland), IAMG2011 (Austria), ECTA2015 (Portugal), MOD2017 (Italy) and EANN2018 (UK).

Xing-Shi He is a professor in the College of Science at Xi'an Polytechnic University. He graduated with an MSc in Mathematics from Shaanxi Normal University. He has held many professional positions including the deputy dean of the College of Science at Xi'an Polytechnic University. He was a recipient of Shaanxi Province Distinguished Teaching Achievement Award, Hong Kong SangMa Teaching Award and a few National and Provincial Awards for Scientific Achievements. He has been on the scientific committee of Shaanxi Society for Industrial and Applied Mathematics and Chinese Statistical Association and has been advisor of Shaanxi Mathematical Modelling Committee. He has published more than 130 research papers and authored/edited 8 books. He had more than 15 research projects and has organized special sessions in many international conferences, including the IEEE SSCI2014 (Orlando, USA). His research interests include algorithms, mathematical modelling, nature-inspired computation, statistics and probability and computational intelligence.

Chapter 1
Introduction to Optimization

Optimization is part of many university courses because of its importance in many disciplines and applications such as engineering design, business planning, computer science, data mining, machine learning, artificial intelligence and industries. The techniques and algorithms for optimization are diverse, ranging from the traditional gradient-based algorithms to contemporary swarm intelligence based algorithms. This chapter introduces the fundamentals of optimization and some of the traditional optimization techniques.

1.1 Introduction

Optimization is everywhere, though it can mean different things from different perspectives. From basic calculus, optimization can simply be to find the maximum or minimum of a function $f(x)$ such as $f(x) = x^4 + 2x^2 + 1$ in the real domain $x \in \mathbb{R}$. In this case, we can either use a gradient-based method or simply spot the solution due to the fact that x^4 and x^2 are always non-negative; the minimum values of x^4 and x^2 are zero; thus, $f(x) = x^4+2x^2+1$ has a minimum $f_{\min} = 1$ at $x_* = 0$. This can be confirmed easily by taking the first derivative of $f(x)$ with respect to x, we have

$$f'(x) = 4x^3 + 4x = 4x(x^2 + 1) = 0, \tag{1.1}$$

which has only one real solution $x = 0$ because $x^2 + 1$ cannot be zero for any real number x. The condition $f'(x) = 0$ seems to be sufficient to determine the optimal solution in this case. In fact, this function is convex with only one minimum in the whole real domain.

However, things become more complicated when $f(x)$ is highly nonlinear with multiple optima. For example, if we try to find the maximum value of $f(x) =$

X.-S. Yang, X.-S. He, *Mathematical Foundations of Nature-Inspired Algorithms*,
SpringerBriefs in Optimization, https://doi.org/10.1007/978-3-030-16936-7_1

$\text{sinc}(x) = \sin(x)/x$ in the real domain, we can naively use

$$f'(x) = \Big[\frac{\sin(x)}{x}\Big]' = \frac{x\cos(x) - \sin(x)}{x^2} = 0, \tag{1.2}$$

which has an infinite number of solutions for $x \neq 0$. There is no simple formula for these solutions; thus, a numerical method has to be used to calculate these solutions. In addition, even with all the efforts to find these solutions, care has to be taken because the actual global maximum $f_{\max} = 1$ occurs at $x_* = 0$. However, this solution can only be found by taking the limit $x \to 0$, and it is not part of the solutions from the above condition of $f'(x) = 0$. This highlights the potential difficulty for nonlinear problems with multiple optima or multi-modality.

Furthermore, not all functions are smooth. For example, if we try to use $f'(x) = 0$ to find the minimum of

$$f(x) = |x|\mathrm{e}^{-\sin(x^2)}, \tag{1.3}$$

we will realize that $f(x)$ is not differentiable at $x = 0$, though the global minimum $f_{\min} = 0$ occurs at $x_* = 0$. In this case, optimization techniques that require the calculation of derivatives will not work.

Problems become more challenging in higher-dimensional spaces. For example, the nonlinear function

$$f(\boldsymbol{x}) = \Big\{\Big[\sum_{i=1}^{n} \sin^2(x_i)\Big] - \exp\Big(-\sum_{i=1}^{n} x_i^2\Big)\Big\} \cdot \exp\Big[-\sum_{i=1}^{n} \sin^2\sqrt{|x_i|}\Big], \tag{1.4}$$

where $-10 \leq x_i \leq 10$ (for $i = 1, 2, \ldots, n$), has the global minimum $f_{\min} = -1$ at $\boldsymbol{x}_* = (0, 0, \ldots, 0)$, but this function is not differentiable at $\boldsymbol{x}_*$.

Therefore, optimization techniques have to be diverse to use gradient information when appropriate, and not to use it when it is not defined or not easily calculated. Though the above nonlinear optimization problems can be challenging to solve, constraints on the search domain and certain independent variables can make the search domain much more complicated, which can consequently make the problem even harder to solve. In addition, sometime, we have to optimize several objective functions instead of just one function, which will in turn make a problem more challenging to solve.

1.2 Essence of an Algorithm

An algorithm is a computational, iterative procedure. For example, Newton's method for finding the roots of a polynomial $p(x) = 0$ can be written as

$$x_{t+1} = x_t - \frac{p(x_t)}{p'(x_t)}, \tag{1.5}$$