Preface

This book is based on the author's lecture "Modern Optimization Methods" given to graduate students at the Beijing Institute of Technology since 2015. It aimed at presenting complete and systematic theories of numerical optimization and their latest applications in different areas, especially in machine learning, statistics, and computer science.

This book aims to introduce the basic definitions and theory of numerical optimization, including optimality conditions for unconstrained and constrained optimization, as well as algorithms for unconstrained and constrained problems. Moreover, it also includes the nonsmooth Newton's method, which plays an important role in large-scale numerical optimization. Finally, based on the author's research experiences, several latest applications for optimization are introduced, including optimization algorithms for hypergraph matching, support vector machine, and bilevel optimization approach for hyperparameter selection in machine learning.

The structure of book is organized into three parts. In the first part (chapters 1–6), after the introduction in chapter 1, we start with the fundamentals of optimization, followed by the typical methods for unconstrained optimization problems. Such methods are classified as line search methods and trust region methods. In particular, we include semismooth Newton's method as an independent chapter to show how we can apply it to solve the data-based problem, *i.e.*, support vector machine (SVM) in machine learning. In the second part (chapters 7 and 8), we introduce the theories and typical methods for a constrained optimization problem. Moreover, in order to show how the methods are applied in problems arising from the application, we demonstrate the quadratic penalty method by applying it to the hypergraph matching, and show the interesting exact recovery property based on the author's research experience. We also show how the augmented Lagrange method is used in the L1 regularized SVM. The final part (chapter 9) is to show how optimization can be used in hyperparameter selection in machine learning.

It involves bilevel optimization and mathematical program with equilibrium constraints (MPEC), which are the two branches of optimization.

We sincerely wish that the reader will learn and understand the main ideas and the essence of the basic optimality theories and the numerical algorithms. We also wish that the reader can deeply understand the ideas of how to implement fast algorithms and further undertake the related research after having learned references from this book.

We also want to take this opportunity to thank all the people who were concerned about us, including my parents, husband, kids, and my teachers, colleagues, and collaborators.

Qingna LI
March, 2023