First Order Nonsmooth Optimization: Algorithm Design, Analysis, Convergence, and Applications

[*]Author Name 1 [†] and Author Name 2

December 11, 2024

Abstract

Abstract abstract abstract.

Contents

1 Introduction **2**

2 Preliminaries **3**

2.1 Fundamentals in convex analysis 4

2.2 Fundamentals in non-convex analysis 4

3 Unifying variants of Nesterov's accelerated methods **4**

4 Method of spectral momentum **4**

5 Catalyst accelerations and future works **4**

[*]Department of Mathematics, I.K. Barber Faculty of Science, The University of British Columbia, Kelowna, BC Canada V1V 1V7. E-mail: `alto@mail.ubc.ca`.

[†]Department of Mathematics, I.K. Barber Faculty of Science, The University of British Columbia, Kelowna, BC Canada V1V 1V7. E-mail: `shawn.wang@ubc.ca`.

6 Performance estimation problems **4**

7 Methods of inexact proximal point **4**

8 Nestrov's acceleration in the non-convex case **4**

9 Using PostGreSQL and big data analytic method for species classification on Sentinel-2 Satelite remote sensing imagery **4**

1 Introduction

Let $\mathbb{R}^n$ be the ambient space. We consider

$$\min_{x\in\mathbb{R}^n} \{F(x) : f(x) + g(x)\}\,. \tag{1.1}$$

Unless specified, assume f is a smooth function with L-Lipschitz gradient operator. Unless specified, assume g is a convex function. In the literatures this type of problems are referred to as additive composite problems.

Our ongoing research concerns accelerated proximal gradient type method for solving (1). In the expository writing by Walkington [2], a variant for of accelerated gradient method for strongly convex function f is discussed. We had two lingering questions after reading it.

(i) Do there exist a unified description for the convergence for both variants of the algorithm.

(ii) Is it possible to attain faster convergence rate without knowledge about the strong convexity of function f.

The good news is we have definitive answers for both questions through our own efforts of research. Section 3, 4 are our ongoing research which present the answers to the questions.

In Section 3, we proposed the method of "Relaxed Weak Accelerated Proximal Gradient (R-WAPG)" as the foundation to describe several variants of Accelerated proximal gradient method in the literatures. The convergence theories of R-WAPG allows us to model convergence of accelerated proximal gradient method where the momentum sequence doesn't strictly follow the conditions presented in the literatures. The descriptive power of R-WAPG allows convergence analysis for all the variants using one single theorem.

In Section 4 we propose a practical algorithm that exploits a specific term in the proof of R-WAPG to achieve faster convergence for solving (1) without knowing parameter L, μ in prior. Results of numerical experiments are presented.

2 Preliminaries

Beck's book [1].

2.1 Fundamentals in convex analysis

2.2 Fundamentals in non-convex analysis

3 Unifying variants of Nesterov's accelerated methods

4 Method of spectral momentum

5 Catalyst accelerations and future works

6 Performance estimation problems

7 Methods of inexact proximal point

8 Nestrov's acceleration in the non-convex case

9 Using PostGreSQL and big data analytic method for species classification on Sentinel-2 Satelite remote sensing imagery

References

[1] A. Beck, *First-order Methods in Optimization*, MOS-SIAM Series in Optimization, SIAM, israel, 2017.

[2] W. Noel, *Nesterov's method for convex optimization*, SIAM Review, 65, pp. 539–562.