AM8211 Non-Linear Optimization Project Proposal

Sep. 20, 2023

Mariam Walaa (500-744-640)

Project Title:

Subdifferential Calculus for Non-Differentiable Convex Functions

Background:

When working with non-linear programming problems, it is possible to encounter objective functions that are not necessarily possible to differentiate, which is commonly done to objective functions in non-linear optimization. One example is any objective function that uses the absolute value function.

The theory of convex analysis gives a characterization of continuously differentiable convex functions, but not non-differentiable convex functions. The theory of subdifferential calculus introduces a new characterization of non-differentiable convex functions, which generalizes the concept of gradient to all kinds of convex functions.

An important result that stems from the theory of subdifferentials is the subdifferential optimality condition, which is a necessary and sufficient condition for optimality in convex optimization problems involving non-differentiable convex functions.

There are numerous applications of the theory of subdifferentials, one of the most prominent being the L1 norm which is typically used in L1 regularization in machine learning and data analysis. Another example in machine learning is the hinge loss of support vector machines (SVMs).

This report will aim to summarize the topic of subdifferential calculus for non-differential convex functions by providing the main ideas, theorems, and applications of subdifferential calculus.

Reference #1: Nonlinear Optimization by Andrzej Ruszczyński (2006)

Ruszczynski introduces the notion of a subgradient with theorems and examples, and transferable rules of calculus such as the chain rule. Section 2.5 Subdifferential Calculus (pg. 57 to 74) from Part I: Theory (Chapter 2 – Elements of Convex Analysis in Nonlinear Optimization by Andrzej Ruszczyński will be referenced.

Reference #2: Nonlinear Programming by Dimitri P. Bertsekas (2018)

P. Bertsekas discusses subgradients with respect to two optimization algorithms: the subgradient method and the cutting plane method. The former method uses a single subgradient at each iteration, while the cutting plane method uses all previously calculated subgradients at each iteration. Chapter 7.5 Sub-gradient Optimization Methods (pg. 709 to 744) from Nonlinear Programming by Dimitri P. Bertsekas will be referenced.

Additional references, not mentioned here, may be included in the report.