CSE 592 Project: Gradient-Free Stochastic Unconstrained Convex Optimizations
Progress Report

Xuan Li, Xiaofei Sun, Shaohui Liu

April 16, 2018

1 Overview

In this project, we are going to implement three different methods for gradient-free stochastic convex optimizations [1][2][3].

The problem we want to solve is stated as follows:

$$\min_{x \in \mathbb{R}^d} \mathbb{E}_\xi F(x, \xi) \tag{1}$$

where $f(x) = \mathbb{E}_\xi F(x, \xi)$ is a convex function.

We can think f as the true function, F as the noisy measure of that funtion. But we cannot access $f(x)$, we can only get noisy function values $F(x, \xi)$.

2 Assumptions

These methods assume that the true gradient ∇f is L-Lipschitz. The Lipschitz constant appears as a parameter in step sizes.

Besides, the oracle assumptions are slightly different. There are two different assumptions:

One-point Feedbacks. In this assumption, given a realization of ξ, we can measure once. That is, different measures must have different realization of ξ.

Two-point Feedbacks. In this assumption, given a realization of ξ, we can measure on a pair of points. That is, given (x, y), we can get two measures $F(x, \xi), F(y, \xi)$ with the same realization of ξ.

3 Basic Ideas

The three methods use a common technique: gradient estimation, but with different schemes. With estimated gradients, we can use first-order methods to update.

The updating scheme are slightly different, two papers use gradient descend, and one paper uses mirror descend.

The proofs of bounds are based on the schemes to estimate gradients and the schemes to update.

4 Plan

In this project, we mainly focus on the implementations of these methods, and compare them on different oracles.

The main steps of our project are as follows:

1. Each one in our group chooses a paper to read, figure out all assumptions, the method of gradient estimation, and the method to update.

2. Each one implement one paper using the same code structures: similar APIs, similar styles, etc.

3. Experiment to compare these different algorithms. The basic idea is to compare the convergence rates.

References

[1] CHEN, R., AND WILD, S. M. Randomized derivative-free optimization of noisy convex functions.

[2] DVURECHENSKY, P., GASNIKOV, A., AND GORBUNOV, E. An accelerated method for derivative-free smooth stochastic convex optimization, 2018.

[3] GHADIMI, S., AND LAN, G. Stochastic first- and zeroth-order methods for nonconvex stochastic programming. *SIAM Journal on Optimization 23* (2013), 2341–2368.