Overview

R.D.C. Monteiro[1]

[1]School of Industrial and SystemsEngineering
Georgia Institute of Technology

January 10, 2024

Outline

Composite optimization (CO) problem is

$$\phi_* = min\{\phi(x) := (f+h)(x) : x \in \mathbb{R}^n\}$$

where $h \in \overline{\mathrm{Conv}}\,(\mathbb{R}^n)$ and f is possibly convex or nonconvex (e.g., weakly convex), and either smooth or nonsmooth

First-order methods assume that any subproblem

$$\min_x \ \lambda h(x) + \frac{1}{2}\|x-a\|^2$$

where $a \in \mathbb{R}^n$ and $\lambda > 0$ is relatively easy to solve

Special case (Set Optimization):

Let $h(\cdot) = \delta_X(\cdot)$ where $\delta_X(\cdot)$ is the indicator function of a nonempty closed convex set X defined as

$$\delta_X(x) := \begin{cases} 0 & \text{if } x \in X \\ +\infty & \text{otherwise} \end{cases}$$

CO with such $h(\cdot)$ reduces to the set optimization

$$\min\{f(x) : x \in X\}$$

Solution of the above subproblem is then given by

$$Proj_X(a) = \operatorname{argmin}\{\|x - a\| : x \in X\}$$

Convex smooth CO

Assumption:

- f is convex and L-smooth on $\mathrm{dom}\, h$

Typical methods include:

- composite gradient (CG)
- accelerated composite gradient (ACG)

Convex nonsmooth CO

Assumption:

- f is convex on $\operatorname{dom} h$
- $\partial f(x) \neq \emptyset$ for all $x \in \operatorname{dom} h$

Typical methods include:

- subgradient methods
- proximal bundle methods

Nonconvex smooth CO

Assumption:

- f is L-smooth on $\operatorname{dom} h$

Typical methods include:

- CG methods
- ACG methods
- inexact proximal methods

Stochastics CO

CO is the same as before but have access only to a stochastic subgradient, i.e., to the following oracle

Oracle: Given $x \in \operatorname{dom} h$, it returns $F(x;\xi) \in \mathbb{R}$ and $s(x;\xi) \in \mathbb{R}^n$ such that

$$\mathbb{E}_\xi[F(x;\xi)] = f(x), \quad \mathbb{E}_\xi[s(x;\xi)] \in \partial f(x)$$

If $F(\cdot;\xi)$ is convex, we can take

$$s(x;\xi) \in \partial F(x;\xi) =: [\partial F(\cdot;\xi)](x)$$

Many of the previous methods generalize to this setting

Linearly Constrained CO

We will also study first-order methods for linearly constrained CO

$$\min\{(f+h)(x) : Ax = b\}$$

where A is a linear map

Moreover, we also consider first-order methods for solving block structured COs, namely,

$$\min(f+h)(x)$$

where $f \in \overline{\mathrm{Conv}}\left(\mathbb{R}^n\right)$ and for any $x = (x_1, \ldots, x_B)$

$$h(x) := \sum_{i=1}^{B} h_i(x_i) \quad h_i \in \overline{\mathrm{Conv}}\left(\mathbb{R}^{n_i}\right)$$