COMP 414/514:
Optimization – Algorithms, Complexity and Approximations

Lecture 4

Overview

- In the last lecture, we:
  - Introduced some notions on convex optimization
  - Studied how such global assumptions affect the performance of gradient descent and what we can say about its convergence rate

- In this lecture, we will:
  - Solely focus on an important variant in convex optimization, the Frank–Wolfe algorithm

Thus far:

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & x \in \mathcal{C} \end{aligned}$$

Thus far:

$$\min_{x} \quad f(x)$$

$$\text{s.t.} \quad \cancel{x \in \mathcal{C}}$$

– Either the problem is unconstrained (for general objectives)

Thus far:

$$\min_x \quad f(x)$$
$$\text{s.t.} \quad x \in \mathcal{C}$$

– Either the problem is unconstrained (for general objectives)

– Or the problem is constrained (but the analysis assumes convex objective)

Frank–Wolfe or the Conditional Gradient Method

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & x \in \mathcal{C} \end{aligned}$$

Frank–Wolfe or the Conditional Gradient Method

$$\begin{aligned} \min_{x} \quad & f(x) \\ \text{s.t.} \quad & x \in \mathcal{C} \end{aligned}$$

- Designed for convex optimization (originally)

Frank–Wolfe or the Conditional Gradient Method

$$\min_{x} \quad f(x)$$
$$\text{s.t.} \quad x \in \mathcal{C}$$

- Designed for convex optimization (originally)
- Its purpose is to handle constraints in a more efficient way (while remaining convex)

Frank–Wolfe or the Conditional Gradient Method

$$\min_{x} \quad f(x)$$
$$\text{s.t.} \quad x \in \mathcal{C}$$

- Designed for convex optimization (originally)
- Its purpose is to handle constraints in a more efficient way (while remaining convex)
- We will see that, compared to convex projected gradient descent, we can achieve practical acceleration, without losing theoretical guarantees

Whiteboard

Conclusion

- We have introduced the notion of convexity
- We studied some of the merits of convex optimization

Conclusion

- We have introduced the notion of convexity
- We studied some of the merits of convex optimization

Next lecture

- We will consider an important variant for convex optimization for large-scale computing: Frank-Wolfe (conditional gradient) algorithm