Model Function Based Conditional Gradient Method with Armijo-like Line Search

Peter Ochs

Mathematical Optimization Group
Saarland University

— 13.06.2019 —

joint work: Yura Malitsky

Classic Conditional Gradient Method

Constrained Smooth Optimization Problem:

$$\min_{x \in C} \; f(x)$$

- $C \subset \mathbb{R}^N$ compact and convex constraint set

Conditional Gradient Method: Update step:

$$y^{(k)} \in \operatorname*{argmin}_{y \in C} \; \langle \nabla f(x^{(k)}), y \rangle$$

$$x^{(k+1)} = \gamma_k y^{(k)} + (1 - \gamma_k) x^{(k)}$$

Convergence mainly relies on:

- step size $\gamma_k \in [0, 1]$ (*we consider Armijo line search*)
- **Descent Lemma** (implies curvature condition)

Generalizing the Descent Lemma

Descent Lemma:

$$|f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle | \leq \frac{L}{2} \|x - \bar{x}\|^2$$

provides a measure for the **linearization error**

$\rightsquigarrow$ quadratic growth

- f smooth non-convex
- L is the Lipschitz constant of ∇f

Generalizing the Descent Lemma

Generalization of the Descent Lemma:

$$|f(x) - f(\bar{x}) - \langle \nabla f(\bar{x}), x - \bar{x} \rangle| \leq \omega(\|x - \bar{x}\|)$$

provides a measure for the **linearization error**

$\rightsquigarrow$ growth given by ω

- f smooth non-convex
- $\omega\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a growth function

Generalizing the Descent Lemma

Generalization of the Descent Lemma:

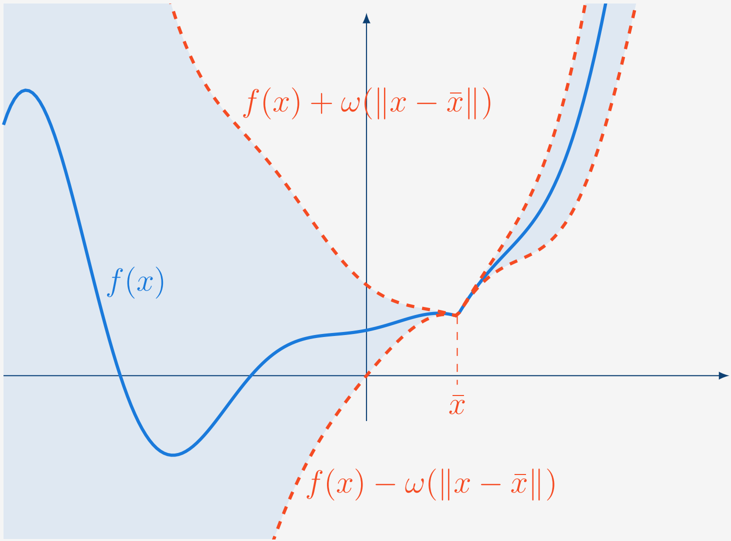
$$|f(x) - f_{\bar{x}}(x)| \leq \omega(\|x - \bar{x}\|)$$

provides a measure for the **approximation error**

$\rightsquigarrow$ growth given by ω

- f non-smooth non-convex
- $\omega\colon \mathbb{R}_+ \to \mathbb{R}_+$ is a growth function

Model Assumption $|f(x) - f_{\bar{x}}(x)| \le \omega(\|x - \bar{x}\|)$



Model Function based Conditional Gradient Method:

$$y^{(k)} \in \operatorname*{argmin}_{y \in C} f_{x^{(k)}}(y)$$
$$x^{(k+1)} = \gamma_k y^{(k)} + (1 - \gamma_k) x^{(k)}$$

Examples for Model Assumption: $|f(x) - f_{\bar{x}}(x)| \leq \omega(\|x - \bar{x}\|)$

- **additive composite problem:**

$$\min_{x \in C} \{ f(x) = \underbrace{g(x)}_{\text{non-smooth}} + \underbrace{h(x)}_{\text{smooth}} \}$$

- **model function:** $f_{\bar{x}}(x) = g(x) + h(\bar{x}) + \langle \nabla h(\bar{x}), x - \bar{x} \rangle$
- **oracle:** $\operatorname*{argmin}_{y \in C} g(y) + \left\langle \nabla h(x^{(k)}), y \right\rangle$

Examples

Examples for Model Assumption: $|f(x) - f_{\bar{x}}(x)| \leq \omega(\|x - \bar{x}\|)$

- **hybrid Proximal–Conditional Gradient**, example:

$$\min_{\substack{x_1 \in C_1 \\ x_2 \in C_2}} \{ f(x_1, x_2) = \underbrace{g(x_1)}_{\text{non-smooth}} + \underbrace{h(x_2)}_{\text{smooth}} \}$$

- $f_{\bar{x}}(x_1, x_2) = h(\bar{x}_2) + \langle \nabla h(\bar{x}_2), x_2 - \bar{x}_2 \rangle + g(x_1) + \frac{1}{2\lambda}\|x_1 - \bar{x}_1\|^2$

- **oracle:** $\begin{cases} \underset{y_1 \in C_1}{\operatorname{argmin}} \quad g(y_1) + \frac{1}{2\lambda}\|y_1 - x_1^{(k)}\|^2 \\ \underset{y_2 \in C_2}{\operatorname{argmin}} \quad \left\langle \nabla h(x_2^{(k)}), y_2 \right\rangle \end{cases}$

Examples

- **composite problem**
- **second order Conditional Gradient**

Design model functions for your problem!