Continuous Optimization

Chapter 3: Constrained Optimization

1 Definitions

In this chapter we will consider constrained optimization problems with the following shape

$$\begin{aligned} \min \ & f(x) \\ \text{s.t.} \ & x \in C \end{aligned} \tag{1}$$

Definition 1.1 (Convex Set). *A set C is said to be convex if given $x_1, x_2 \in C$ and $\lambda \in [0,1]$, then $\lambda x_1 + (1-\lambda)x_2 \in C$.*

Definition 1.2 (Convex Function). *A function $f : C \to \mathbb{R}$ defined on a convex set C is said to be convex if given $x_1, x_2 \in C$ and $\lambda \in [0,1]$, then*

$$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2).$$

Definition 1.3 (Strictly Convex Function). *A function $f : C \to \mathbb{R}$ defined on a convex set C is said to be strictly convex if given $x_1, x_2 \in C$ and $\lambda \in [0,1]$, then*

$$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2).$$

A function is called concave if $-f$ is convex and strictly concave if $-f$ is strictly convex.

2 Characterizations of Convex Functions

Theorem 2.1 (Gradient characterization of convex functions). *Let $f \in \mathrm{C}^1(C)$, where C is convex. Then f is convex over C if and only if*

$$f(x) + \nabla f(x)^T (y - x) \leq f(y) \quad \forall x, y \in C. \tag{2}$$

Proof. Exercise. □

Proposition 2.1 (Sufficiency of stationarity under convexity). *Let $f \in \mathrm{C}^1(C)$, where $C \subseteq \mathbb{R}^n$ is convex. Suppose that $\nabla f(x^*) = 0$ for some $x^* \in C$. Then x^* is a global minimizer of f over C.*

Proof. Let $z \in C$. Plugging $x = x^*$ and $y = z$ in Theorem 2.1 we obtain that

$$f(z) \geq f(x^*) + \nabla f(x^*)^T (z - x^*),$$

which implies that $f(z) \geq f(x^*)$ because $\nabla f(x^*) = 0$. □

We note that Proposition 2.1 establishes only the sufficiency of the stationarity condition $\nabla f(x^*) = 0$ for guaranteeing that x^* is a global optimal solution. When $C$is not the entire space, this condition is not necessary, in fact it might be that the points for which $\nabla f(\cdot) = 0$ are not in C. On the other hand, when $C = \mathbb{R}^n$ and f is convex, $\nabla f(x^*) = 0$ is both sufficient and necessary condition for x^* to be a global minimum. We can now establish the conditions under which a twice continuously differentiable function f is convex.

Theorem 2.2 (Second order characterization of convexity). *Let $f \in \mathrm{C}^2(C)$, where $C \subseteq \mathbb{R}^n$ is convex and open. Thus, we have that f is convex iff $\nabla^2 f(x) \succcurlyeq 0 \quad \forall x \in C$.*

Proof. Suppose that $\nabla^2 f(x) \succcurlyeq 0$ for all $x \in C$. We will prove (2) which is enough to establish convexity. Let $x, y \in C$, then by the Mean Value Theorem[2] we get that there exists $z \in [x, y]$ (and hence $z \in C$) for which

$$f(y) = f(x) + \nabla f(x)^T (y - x) + \frac{1}{2}(y - x)^T \nabla^2 f(z)(y - x). \tag{3}$$

Since $\nabla^2 f(z) \succcurlyeq 0$, it follows that $(y - x)^T \nabla^2 f(z)(y - x) \geq 0$, which implies (2). □

3 Stationarity

Definition 3.1 (Stationary points of convex constrained problems)**.** *Let $f \in \mathrm{C}^1(C)$, where C is closed and convex. Then x^* is a stationary point of* (1) *if $\nabla f(x^*)(x - x^*) \geq 0\ \forall x \in C$.*

In words, this means that there are no feasible descent directions of f at x^*. This suggests that stationarity is in fact a necessary condition for a local minimum of (1).

Theorem 3.1 (Stationarity as necessary optimality condition)**.** *Let $f \in \mathrm{C}^1(C)$, where C is closed and convex and let x^* be a local minimum of* (1)*. Then x^* is a stationary point of* (1)*.*

Proof. Let x^* be a local minimum of f and assume by contradiction that is not a stationary point of (1). Then there exists $x \in C$ such that $\nabla f(x^*)(x - x^*) < 0$. Therefore, $f'(x, d) < 0$, where $d = x - x^*$. Hence, by Lemma 1.1 of Chapter 2, there exists $\epsilon \in (0, 1)$ such that $f(x^* + td) < f(x^*)\ \forall t \in (0, \epsilon)$. Since C is convex, we have that $x + td = (1 - t)x^* + tx \in C$, leading to the conclusion that x^* is not a local optimum of (1), which is a contradiction. □

Theorem 3.2 (Stationarity as necessary optimality condition)**.** *Let $f \in \mathrm{C}^1(C)$, where C is closed and convex and f is also convex. Let x^* be a local minimum of* (1)*. Then x^* is a stationary point of* (1) *iff x^* is a optimal solution of* (1)*.*

Proof. The necessity of the stationarity condition follows from Theorem 3.1. To prove the sufficiency, assume that x^* is a stationary point of (1) and let $x \in C$. Then

$$s$$

□

References