

Theory of Optimization

Jayadev Naram

Contents

1	Unconstrained Minimization	2
----------	-----------------------------------	----------

1 Unconstrained Minimization

Let \mathbb{R}^n denote the n -dimensional Euclidean real vector space with the inner product defined for any $x, y \in \mathbb{R}^n$ as $\langle x, y \rangle = x^T y = \sum_{i=1}^n x_i y_i$, where $x = [x_1, \dots, x_n]^T$ and $y = [y_1, \dots, y_n]^T$ are the coordinates of x and y respectively. Let the norm and the metric on \mathbb{R}^n be defined as $\|x\| = \sqrt{\langle x, x \rangle}$ and $d(x, y) = \|x - y\|$, respectively.

Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a real-valued function on \mathbb{R}^n . We consider the following optimization problem

$$\min_{x \in \mathbb{R}^n} f(x). \quad (1)$$

The problem (1) is called as the unconstrained minimization problem as there are no constraints on x .

Definition 1.1. A **local minimum** of f in the problem (1) is a vector $x^* \in \mathbb{R}^n$ for which there exists $\varepsilon > 0$ such that for all $x \in \mathbb{R}^n$ we have

$$f(x^*) \leq f(x), \quad (2)$$

when $\|x - x^*\| \leq \varepsilon$. A **global minimum** of f in the problem (1) is a vector $x^* \in \mathbb{R}^n$ such that for all $x \in \mathbb{R}^n$ we have

$$f(x^*) \leq f(x). \quad (3)$$

The global or local minimum x^* is said to be **strict** if the corresponding inequality given above is strict for $x \neq x^*$.

Proposition 1.2 (Necessary Optimality Conditions). *Let x^* be an unconstrained local minimum of $f : \mathbb{R}^n \rightarrow \mathbb{R}$, and assume that f is continuously differentiable in an open set U containing x^* . Then*

$$\nabla f(x^*) = 0. \quad (\text{First Order Necessary Condition})$$

If in addition f is twice continuously differentiable within U , then

$$\nabla^2 f(x^*) \succeq 0. \quad (\text{Second Order Necessary Condition})$$

Proof. Fix some $d \in \mathbb{R}^n$. Then, using chain rule to differentiate the function $g(\alpha) = f(x^* + \alpha d)$, we have

$$0 \leq \lim_{\alpha \downarrow 0} \frac{f(x^* + \alpha d) - f(x^*)}{\alpha} = \frac{dg(0)}{d\alpha} = \nabla f(x^*)^T d,$$

where the inequality follows from the assumption that x^* is a local minimum and $\alpha \downarrow 0$ indicates the right-hand limit, i.e., $\alpha > 0$ and $\alpha \rightarrow 0$. Since d is arbitrary, the same inequality holds with d replaced by $-d$. Therefore, $\nabla f(x^*)^T d = 0$ for all $d \in \mathbb{R}^n$, which shows that $\nabla f(x^*) = 0$.

Assume that f is twice continuously differentiable, and let d be any vector in \mathbb{R}^n . For all $\alpha \in \mathbb{R}$, the second order Taylor expansion yields

$$f(x^* + \alpha d) - f(x^*) = \alpha \nabla f(x^*)^T d + \frac{\alpha^2}{2} d^T \nabla^2 f(x^*) d + o(\alpha^2).$$

Using the condition $\nabla f(x^*) = 0$ and the local optimality of x^* , we see that there is a sufficiently small $\varepsilon > 0$ such that for all $\alpha \in (0, \varepsilon)$,

$$0 \leq \frac{f(x^* + \alpha d) - f(x^*)}{\alpha^2} = \frac{1}{2} d^T \nabla^2 f(x^*) d + \frac{o(\alpha^2)}{\alpha^2}.$$

Taking the limit $\alpha \rightarrow 0$ and using the fact that $\lim_{\alpha \rightarrow 0} o(\alpha^2)/\alpha^2 = 0$, we obtain $d^T \nabla^2 f(x^*) d \geq 0$, showing that $\nabla^2 f(x^*)$ is positive semidefinite. \square