1 Optimality for Smooth Functions

Basic Necessary and Sufficient Conditions

Theorem 1.1. *Suppose f is continuously differentiable on domain neighborhood $\mathcal{N}(x^*, r)$. If x^* is a local minimizer of f, then $\nabla f(x^*) = 0$.*

Proof. Let f be continuously differentiable let x^* a local minimizer of f on $\mathcal{N}(x^*, r)$ for some $r > 0$. Suppose, by way of contradiction, that $\nabla f(x^*) \neq 0$.

Let $p = -\nabla f(x^*)$. Then, $p^\mathsf{T} \nabla f(x^*) = -\|\nabla f(x^*)\|^2 < 0$. Because $\nabla f(x)$ is continuous on $\mathcal{N}$, there exists $T > 0$ such that $p^\mathsf{T} \nabla f(x^* + tp) < 0$ for all $t \in (0, T]$. Now, let $s \in (0, T]$. Then, by Taylor's Theorem

$$\begin{aligned} f(x^* + sp) &= f(x^*) + sp^\mathsf{T} \nabla f(x^* + tsp), \quad t \in (0, 1) \\ &= f(x^*) + sp^\mathsf{T} \nabla f(x^* + tp), \quad t \in (0, s) \\ &< f(x^*) \end{aligned}$$

Thus, $f(x^* + sp) < f(x^*)$ for all $s \in (0, T]$, a contradiction. Therefore $\nabla f(x^*) = 0$ if x^* is a local minimizer. $\square$

Theorem 1.2. *Suppose f is twice continuously differentiable on domain neighborhood $\mathcal{N}(x^*, r)$. If x^* is a local minimizer of f, then $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive semi-definite.*

Proof. Let f be twice continuously differentiable on $\mathcal{N}(x^*, r)$ for some $r > 0$, and x^* a local minimizer of f. From Theorem 1.1, $\nabla f(x^*) = 0$. Now, assume, by way of contradiction, that $\nabla^2 f(x^*)$ is not positive semi-definite.

Let p be an eigenvector of $\nabla^2 f$ with a negative eigenvalue. Then $p^\mathsf{T} \nabla f(x^*) p < 0$ ($\nabla f(x^*)$. Therefore, there exists $T > 0$ such that $p^\mathsf{T} \nabla f(x^* + tp) p < 0$ for all $t \in (0, T]$. Let $s \in (0, t]$ then by Taylor's Theorem:

$$\begin{aligned} f(x^* + sp) &= f(x^*) + sp^\mathsf{T} \nabla f(x^*) + \tfrac{1}{2} s^2 p^\mathsf{T} \nabla^2 (x^* + tp) p, \quad t \in (0, s) \\ &= f(x^*) + \tfrac{1}{2} s^2 p^\mathsf{T} \nabla^2 (x^* + tp) p, \quad t \in (0, s) \\ &< f(x^*) \end{aligned}$$

Thus, $f(x^* + sp) < f(x^*)$ for all $s \in (0, T]$, a contradiction. Therefore, $\nabla^2 f(x^*)$ is positive semi-definite if x^* is a local minimizer of f. $\square$

Theorem 1.3. *Suppose f is twice continuously differentiable on domain neighborhood $\mathcal{N}(x^*, r)$. If $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive definite, then x^* is a strict local minimizer.*

Proof. Let f be twice continuously differentiable on $\mathcal{N}(x^*, r)$, $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ positive definite. Then, there exists $\mathcal{N}(x^*, s)$ with $s < r$ and arbitrary p satisfying $x^* + p \in \mathcal{N}(x^*, s)$ and $\nabla^2 f(x^* + p)$ positive definite. Then,

$$f(x^* + p) = f(x^*) + \tfrac{1}{2} p^\mathsf{T} \nabla^2 f(x^* + tp) p, \quad \text{for some } t \in (0, 1).$$

Therfore, $f(x^* + p) > f(x^*)$ on $\mathcal{N}(x^*, s)$. Thus, x^* is a strict local minimizer of f. □

Additional Optimality Conditions

Theorem 1.4. *Suppose f is convex. Then any local minimizer of f is a global minimizer of f.*

Proof. Let f be convex and x^* a local minimizer of f on $\mathcal{N}(x^*, r)$. Suppose, by way of contradiction, that x^* is not a global minimizer of f. Then, there exists y such that $f(y) < f(x^*)$. Let $w = x^* + \alpha(y - x^*)$ with $\alpha \in (0, 1)$ small enough so that $w \in \mathcal{N}(x^*, r)$. We have

$$\begin{aligned} f(w) &= f(\alpha(y - x^*) + x^*) \\ &= f(\alpha y + (1 - \alpha)x^*) \\ &\leq \alpha f(y) + (1 - \alpha) f(x^*) \\ &< \alpha f(x^*) + (1 - \alpha) f(x^*) \\ &= f(x^*) \end{aligned}$$

Now, $f(w) < f(x^*)$ violates the fact that x^* is a local minimizer on $\mathcal{N}(x^*, r)$, a contradiction. Thus, $f(y) \geq f(x^*)$ for all $y \in \mathbb{R}^n$ and x^* is a global minimizer of f. □

Theorem 1.5. *Suppose f is convex and continuously differentiable, then x^* is a global minimizer of f if and only if $\nabla f(x^*) = 0$.*

Proof. Let f be convex and continuously differentiable. Then, we have $f(y) \geq f(x^*) + (y - x^*)^\intercal \nabla f(x^*)$ for all $y \in \mathbb{R}^n$.
($\Rightarrow$) Suppose x^* is a global minimizer of f. Then, $f(y) \geq f(x^*)$ for all $y \in \mathbb{R}^n$. Thus, $(y - x^*)\nabla f(x^*) \geq 0$ for all $y \in \mathbb{R}^n$. If direction vector $(y_1 - x^*)$ results in $(y_1 - x^*)\nabla f(x^*) > 0$ then the negative of that direction $y_2 = 2x^* - y_1$ would result in $(y_2 - x^*)\nabla f(x^*) < 0$. Thus, it must follow that $\nabla f(x^*) = 0$.
($\Leftarrow$) Suppose $\nabla f(x^*) = 0$. Then, $f(y) \geq f(x^*)$ for all $y \in \mathbb{R}^n$. That is, x^* is a global minimizer of f. □

Lemma 1.6. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is coercive if and only if every sublevel set of f is bounded.*

Theorem 1.7. *If $f : \mathbb{R}^n \to \mathbb{R}$ is either convex or continuous, and there exists at least one non-empty bounded sublevel set of f, then a global minimizer of f exists.*

The essence of a proof is that if there exists a bounded non-empty level set Ω of f, then any global minimizer of f must be in Ω. Now, if f is convex on Ω then it is also continuous on Ω. Then, it follows that f is bounded below and continuous and thus attains its infimum. That is, a global minimizer must exist.