First and Second Order Necessary Conditions for Optimality

20/01/25

1 First Order Necessary Condition (FONC)

1.1 Derivation and Introduction

In optimization, we seek to minimize a function $f(x)$ over a feasible set S. Let us consider a point $\bar{x} \in S$ and a feasible direction $\bar{d}$ at $\bar{x}$. A feasible direction $\bar{d}$ indicates a direction in which we can move from $\bar{x}$ and remain within the feasible set, at least for a small step. To understand the behavior of the function near $\bar{x}$ along this feasible direction, we define a univariate function $\phi(\alpha)$ for $\alpha \geq 0$ as:

$$\phi(\alpha) = f(\bar{x} + \alpha\bar{d})$$

where α represents the step size in the direction $\bar{d}$. We observe that:

$$\phi(0) = f(\bar{x})$$

Assuming f is differentiable, we can find the derivative of $\phi(\alpha)$ with respect to α using the chain rule:

$$\phi'(\alpha) = \frac{d}{d\alpha} f(\bar{x} + \alpha\bar{d}) = \nabla f(\bar{x} + \alpha\bar{d})^T \bar{d}$$

Evaluating at $\alpha = 0$:

$$\phi'(0) = \nabla f(\bar{x})^T \bar{d}$$

This $\phi'(0)$ represents the directional derivative of f at $\bar{x}$ in the direction $\bar{d}$.

Now, consider the first order Taylor approximation of $\phi(\alpha)$ around $\alpha = 0$:

$$\phi(\alpha) \approx \phi(0) + \alpha\phi'(0) = f(\bar{x}) + \alpha\nabla f(\bar{x})^T \bar{d} \quad (1) \tag{1}$$

If $\bar{x}$ is a local minimizer of f over S, it means that for any feasible direction $\bar{d}$, moving a small distance from $\bar{x}$ along $\bar{d}$ should not decrease the function value. More formally, for sufficiently small $\alpha > 0$ along any feasible direction $\bar{d}$, we must have $f(\bar{x} + \alpha\bar{d}) \geq f(\bar{x})$. This implies that there exists some $\alpha_0 \in \mathbb{R}^+$ such that,

$$f(\bar{x} + \alpha\bar{d}) \geq f(\bar{x}) \quad \forall \alpha \in (0, \alpha_0]$$

From the Taylor approximation (1), for small $\alpha > 0$, for $f(\bar{x} + \alpha\bar{d}) \geq f(\bar{x})$ to hold, the term $\alpha\nabla f(\bar{x})^T\bar{d}$ must be non-negative (or zero). Since $\alpha > 0$, this leads to the condition:

$$\nabla f(\bar{x})^T\bar{d} \geq 0 \quad (2) \tag{2}$$

Intuitively, at a local minimizer, the gradient must be "uphill" or "flat" in any feasible direction.

1.2 First Order Necessary Condition (FONC) - General Case

Theorem 1 (First Order Necessary Condition (FONC))**.** Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be continuously differentiable on S. If $\bar{x}^*$ is a local minimizer of f over S, then for any feasible direction $\bar{d}$ at $\bar{x}^*$, we have:

$$\nabla f(\bar{x}^*)^T\bar{d} \geq 0$$

1.3 First Order Necessary Condition (FONC) - Interior Case

Theorem 2 (First Order Necessary Condition (FONC) - Interior Case)**.** Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be continuously differentiable on S (i.e., $f \in C^1(S)$). If $\bar{x}^*$ is a local minimizer of f over S and $\bar{x}^*$ is an interior point of S, then:

$$\nabla f(\bar{x}^*) = \mathbf{0}$$

Proof 1 (Proof of FONC - Interior Case)**.** Let $\bar{x}^*$ be a local minimizer of f over S and an interior point of S. Because $\bar{x}^*$ is an interior point, any direction $\bar{d} \in \mathbb{R}^n$ is a feasible direction at $\bar{x}^*$. From the general FONC, for any feasible direction $\bar{d}$, we must have:

$$\nabla f(\bar{x}^*)^T\bar{d} \geq 0$$

Since every direction is feasible at an interior point, the opposite direction $-\bar{d}$ is also a feasible direction. Applying the FONC to the direction $-\bar{d}$, we get:

$$\nabla f(\bar{x}^*)^T(-\bar{d}) \geq 0$$

which simplifies to:

$$-\nabla f(\bar{x}^*)^T\bar{d} \geq 0 \implies \nabla f(\bar{x}^*)^T\bar{d} \leq 0$$

Combining the inequalities $\nabla f(\bar{x}^*)^T\bar{d} \geq 0$ and $\nabla f(\bar{x}^*)^T\bar{d} \leq 0$, we conclude that:

$$\nabla f(\bar{x}^*)^T\bar{d} = 0$$

This equality must hold for all directions $\bar{d} \in \mathbb{R}^n$. The only vector that has a zero dot product with every vector in $\mathbb{R}^n$ is the zero vector itself. Therefore, it must be that:

$$\nabla f(\bar{x}^*) = \mathbf{0}$$

Remark 1. The **First Order Necessary Condition (FONC)** is a necessary condition for local minimality, but it is not sufficient. A point $\bar{x}^*$ satisfying FONC is called a stationary point. Such a point could be a local minimum, a local maximum, or a saddle point. Therefore, verifying FONC is a crucial first step in finding local minimizers, but further conditions are needed to guarantee minimality.

2 Second Order Necessary Condition (SONC)

To refine our optimality conditions and gain more insight, we now consider the second order Taylor expansion. This leads to the Second Order Necessary Condition (SONC), which provides a stronger condition for local minimizers.

2.1 Second Order Necessary Condition (SONC) - General Case

Theorem 3 (Second Order Necessary Condition (SONC))**.** Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be twice continuously differentiable on S (i.e., $f \in C^2(S)$). Let $\bar{x}^*$ be a local minimizer of f over S. Let $\bar{d}$ be a feasible direction at $\bar{x}^*$. If the first-order condition is "flat" in the direction $\bar{d}$, i.e., $\nabla f(\bar{x}^*)^T \bar{d} = 0$, then the second-order condition must be non-negative in this direction:

$$\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} \geq 0$$

where $\nabla^2 f(\bar{x}^*)$ is the Hessian matrix of f at $\bar{x}^*$.

Proof 2 (Proof of SONC - General Case (by Contradiction))**.** We will prove this by contradiction. Assume that $\bar{x}^*$ is a local minimizer, $\bar{d}$ is a feasible direction at $\bar{x}^*$, $\nabla f(\bar{x}^*)^T \bar{d} = 0$, but contrary to the theorem, $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} < 0$.

Consider again the function $\phi(\alpha) = f(\bar{x}^* + \alpha \bar{d})$. We use the second order Taylor approximation of $\phi(\alpha)$ around $\alpha = 0$:

$$\phi(\alpha) = \phi(0) + \alpha \phi'(0) + \frac{\alpha^2}{2} \phi''(0) + o(\alpha^2)$$

We have the following values at $\alpha = 0$:

- $\phi(0) = f(\bar{x}^*)$
- $\phi'(0) = \nabla f(\bar{x}^*)^T \bar{d} = 0$ (by the condition of the theorem)
- $\phi''(0) = \bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} < 0$ (by our contradiction assumption)

Substituting these values into the Taylor expansion:

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

Since $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} < 0$, for sufficiently small $\alpha > 0$, the quadratic term $\frac{\alpha^2}{2} \bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d}$ will be negative and will dominate the higher order term $o(\alpha^2)$. Thus, the sum $\frac{\alpha^2}{2} \bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} + o(\alpha^2)$ will also be negative for small enough $\alpha > 0$. This implies that for such α:

$$f(\bar{x}^* + \alpha \bar{d}) < f(\bar{x}^*)$$

This contradicts the assumption that $\bar{x}^*$ is a local minimizer, because we have found a feasible direction $\bar{d}$ along which the function values decrease for points arbitrarily close to $\bar{x}^*$. Therefore, our initial assumption that $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} < 0$ must be false. Hence, we must have $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} \geq 0$.

2.2 Second Order Necessary Condition (SONC) - Interior Case

Theorem 4 (Second Order Necessary Condition (SONC) - Interior Case)**.** Let $S \subseteq \mathbb{R}^n$ and $f : S \to \mathbb{R}$ be twice continuously differentiable on S (i.e., $f \in C^2(S)$). Let $\bar{x}^*$ be a local minimizer of f over S and $\bar{x}^*$ be an interior point of S. Then, the following conditions must hold:

1. $\nabla f(\bar{x}^*) = \mathbf{0}$ (Stationarity)
2. $\nabla^2 f(\bar{x}^*)$ is positive semi-definite, i.e., $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} \geq 0 \quad \forall$ directions $\bar{d} \in \mathbb{R}^n$.

This means that for any direction $\bar{d}$, the quadratic form associated with the Hessian matrix at $\bar{x}^*$ must be non-negative. Equivalently, all eigenvalues of $\nabla^2 f(\bar{x}^*)$ must be non-negative.

Proof 3 (Proof of SONC - Interior Case)**.** Let $\bar{x}^* \in \text{int}(S)$ be a local minimizer. From the FONC for interior points, we know that the first condition, $\nabla f(\bar{x}^*) = \mathbf{0}$, must hold.

Now, for any direction $\bar{d} \in \mathbb{R}^n$, since $\bar{x}^*$ is an interior point, $\bar{d}$ is a feasible direction. As $\nabla f(\bar{x}^*)^T \bar{d} = 0$ (because $\nabla f(\bar{x}^*) = \mathbf{0}$), we can apply the SONC for the general case. By SONC, we must have:

$$\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} \geq 0 \quad \forall \bar{d} \in \mathbb{R}^n$$

This precisely means that the Hessian matrix $\nabla^2 f(\bar{x}^*)$ is positive semi-definite.

Alternatively, consider the function $\phi(\alpha) = f(\bar{x}^* + \alpha \bar{d})$. We know from FONC that $\nabla f(\bar{x}^*) = \mathbf{0}$, so $\phi'(0) = 0$. Assume for contradiction that $\nabla^2 f(\bar{x}^*)$ is not positive semi-definite. Then there exists a direction $\bar{d}$ such that $\bar{d}^T \nabla^2 f(\bar{x}^*) \bar{d} < 0$, i.e., $\phi''(0) < 0$. Since $f \in C^2(S)$, $\phi \in C^2$ and ϕ'' is continuous. As $\phi''(0) < 0$, there exists $\delta > 0$ such that $\phi''(\alpha) < 0$ for all $\alpha \in (0, \delta)$.

Using the second order Taylor series with remainder (truncated form) for some $\zeta \in (0, \delta)$:

$$\phi(\delta) = \phi(0) + \delta \phi'(0) + \frac{\delta^2}{2} \phi''(\zeta)$$

Since $\phi'(0) = 0$ and $\phi''(\zeta) < 0$ for $\zeta \in (0, \delta)$, we have:

$$\phi(\delta) = \phi(0) + \frac{\delta^2}{2}\phi''(\zeta) < \phi(0)$$

Thus, $f(\bar{x}^* + \delta\bar{d}) < f(\bar{x}^*)$, which contradicts the local minimality of $\bar{x}^*$. Therefore, we must have $\nabla^2 f(\bar{x}^*)$ be positive semi-definite.