Deterministic Optimization

Unconstrained Optimization: Derivative based

Shabbir Ahmed
Anderson-Interface Chair and Professor
School of Industrial and Systems Engineering

Optimality Conditions

Optimality Conditions

Learning objective:

- Identify first and second order optimality conditions

Unconstrained Optimization: Derivative Based

$$(P): \quad \min f(x) \;\; \text{s.t. } x \in \mathbb{R}^n$$

where $f: \mathbb{R}^n \to \mathbb{R}$ is continuous and twice differentiable.

- Lesson 1: Optimality Conditions
- Lesson 2: Gradient Descent
- Lesson 3: Newton's Method

Unconstrained Optimization: Derivative Based

$$(P): \quad \min f(x) \quad \text{s.t. } x \in \mathbb{R}^n$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is continuous and twice differentiable.

- Lesson 1: Optimality Conditions
- Lesson 2: Gradient Descent
- Lesson 3: Newton's Method

First Order Optimality Conditions

What are the conditions under which x^* is a local optima?

From Taylor's expansion:

$$\begin{aligned} f(x^* + \Delta x) &= f(x^*) + \frac{\nabla f(x^*)^\top \Delta x}{1!} \\ &\quad + \frac{\Delta x^\top \nabla^2 f(x^*) \Delta x}{2!} + \cdots \end{aligned}$$

First Order Optimality Conditions

If x^* is a local minima, then $f(x^* + \Delta x) \geq f(x^*)$ for Δx sufficiently small, i.e.

$$\frac{\nabla f(x^*)^\top \Delta x}{1!} \geq 0$$

since lower order term dominates.

But Δx can be positive or negative. Thus we must have

$$\nabla f(x^*) = 0.$$

The above argument also holds for local maxima, i.e. if x^* is a local maxima or minima, then $\nabla f(x^*) = 0$.

A point where the gradient vanishes is called a stationary point.

Second Order Optimality Conditions

Consider the higher order terms in Taylor's expansion

$$\begin{aligned} f(x^* + \Delta x) &= f(x^*) + \frac{\nabla f(x^*)^\top \Delta x}{1!} \\ &\quad + \frac{\Delta x^\top \nabla^2 f(x^*) \Delta x}{2!} + \cdots \end{aligned}$$

If x^* is a local min, we have $\nabla f(x^*)^\top \Delta x = 0$. Then for $f(x^* + \Delta x) \geq f(x^*)$ to hold for all Δx we must have

$$\Delta x^\top \nabla^2 f(x^*) \Delta x \geq 0 \text{ for all } \Delta x$$

That is, $\nabla^2 f(x^*)$ (the Hessian) is positive semi-definite.

First Order Optimality Conditions

The two conditions are only necessary, not sufficient

E.g. Consider $f(x) = x^3$. At the point $x = 0$, $f'(0) = 0$ and $f''(0) = 0$. Both conditions are satisfied, but $x = 0$ is neither a local min or max.

Sufficient (but not necessary) conditions for x^* to be a local minima: $\nabla f(x^*) = 0$ and $\nabla^2 f(x^*)$ is positive-definite.

Summary

- If a solution is a local optimal solution of an unconstrained problem, then the gradient vanishes at that point
- Also the Hessian is positive semidefinite
- The conditions are necessary but not sufficient