Introduction to Optimization

Lecture 03: Optimality conditions. Examples. Convex functions.

university of groningen

First order optimality condition

Theorem (Fermat's Rule)

Let $f : A \subset \mathbb{R}^N \to \mathbb{R}$ and let $\emptyset \neq C \subset A$ be convex. If $\hat{x} \in C$ is such that $f(\hat{x}) \leq f(y)$ for all $y \in C$, and if f is differentiable at $\hat{x}$, then

$$\nabla f(\hat{x}) \cdot (y - \hat{x}) \geq 0$$

for all $y \in C$. If, moreover, $\hat{x} \in \mathrm{int}(C)$, then $\nabla f(\hat{x}) = 0$.

First order optimality condition

Theorem (Fermat's Rule)

Let $f : A \subset \mathbb{R}^N \to \mathbb{R}$ and let $\emptyset \neq C \subset A$ be convex. If $\hat{x} \in C$ is such that $f(\hat{x}) \leq f(y)$ for all $y \in C$, and if f is differentiable at $\hat{x}$, then

$$\nabla f(\hat{x}) \cdot (y - \hat{x}) \geq 0$$

for all $y \in C$. If, moreover, $\hat{x} \in \operatorname{int}(C)$, then $\nabla f(\hat{x}) = 0$.

Question

What if C is affine?

Example

Compute the maximum value of the expression

$$\sum_{i=1}^{N} \alpha_i \ln(x_i)$$

subject to the constraint that

$$\sum_{i=1}^{N} x_i = b,$$

where $\alpha_1, \ldots, \alpha_N, b > 0$.

Second order conditions

Theorem (Second order optimality conditions)

Let $f : A \subset \mathbb{R}^N \to \mathbb{R}$ be twice differentiable at $\hat{x} \in \mathrm{int}(A)$.

i) *If $\hat{x}$ is a local minimizer of f, then $\nabla f(\hat{x}) = 0$ and $\nabla^2 f(\hat{x})$ is positive semidefinite ($\nabla^2 f(\hat{x})d \cdot d \geq 0$ for all $d \in \mathbb{R}^N$).* ✍

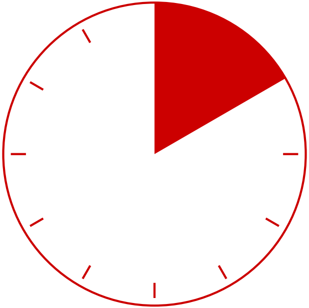
ii) *If $\nabla f(\hat{x}) = 0$ and $\nabla^2 f(\hat{x})$ is positive definite ($\nabla^2 f(\hat{x})d \cdot d > 0$ for all $d \neq 0$), then $\hat{x}$ is a strict local minimizer of f.*

Lemma (Taylor's Approximation)

Let $f : A \subset \mathbb{R}^N \to \mathbb{R}$ be of class C^2, and let $x \in A$. For each $d \in \mathbb{R}^N$,

$$\lim_{t\to 0} \frac{1}{t^2} \left| f(x+td) - f(x) - t\nabla f(x) \cdot d - \frac{t^2}{2}\nabla^2 f(x)d \cdot d \right| = 0.$$

Break



Convex functions

A function $f : D \subset \mathbb{R}^N \to \mathbb{R}$ is convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in D$ and all $\lambda \in (0,1)$ (or $[0,1]$, if you prefer).

Convex functions

A function $f : D \subset \mathbb{R}^N \to \mathbb{R}$ is convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in D$ and all $\lambda \in (0,1)$ (or $[0,1]$, if you prefer).

Proposition

The function $f(x) = \frac{1}{2}\|Ax - b\|^2$ is convex.

Convex functions

A function $f : D \subset \mathbb{R}^N \to \mathbb{R}$ is convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda) f(y)$$

for all $x, y \in D$ and all $\lambda \in (0,1)$ (or $[0,1]$, if you prefer).

Proposition

The function $f(x) = \frac{1}{2}\|Ax - b\|^2$ is convex.

Proposition

Local minimizers of convex functions are global minimizers.

Characterizations of differentiable convex functions

Proposition

Let $f : D \subset \mathbb{R}^N \to \mathbb{R}$ be differentiable. The following are equivalent:

1. *f is convex;*
2. *for all $x, y \in D$, $f(y) \geq f(x) + \nabla f(x) \cdot (y - x)$;*
3. *for all $x, y \in D$, $\big(\nabla f(y) - \nabla f(x)\big) \cdot (y - x)$.*

If f is twice differentiable, the three statements above are equivalent to

4. *for all $x \in D$, $\nabla^2 f(x)$ is positive semidefinite.*

Strict and strong convexity

A function $f : D \subset \mathbb{R}^N \to \mathbb{R}$ is strictly convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) < \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in D$ and all $\lambda \in (0,1)$

Strict and strong convexity

A function $f : D \subset \mathbb{R}^N \to \mathbb{R}$ is strictly convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) < \lambda f(x) + (1-\lambda)f(y)$$

for all $x, y \in D$ and all $\lambda \in (0,1)$, and it is strongly convex if D is convex and

$$f\big(\lambda x + (1-\lambda)y\big) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\alpha}{2}\lambda(1-\lambda)\|x-y\|^2$$

for all $x, y \in D$ and all $\lambda \in (0,1)$.

Exercises

1. Find examples of: a function that is convex, but not strictly convex; and a function that is strictly convex, but not strongly convex.

2. When is the function $f(x) = \frac{1}{2}\|Ax - y\|^2$ strictly/strongly convex?

3. Prove that every strictly convex function $f : \mathbb{R}^N \to \mathbb{R}$ has at most one minimizer, and every strongly convex function $f : \mathbb{R}^N \to \mathbb{R}$ has exactly one minimizer.

4. Can you obtain characterizations of strict and strong convexity of f in terms of properties of ∇f and $\nabla^2 f$?