Optimality conditions in the unconstrained case

Find extrema of a function defined on E.

Find

$$\inf_{x \in E} f(x),$$

infimum of f is the correct notation as long as we don't know if a minimum exists

with

$$f : E \longrightarrow \mathbb{R}$$

E normed finite dimension vector space $= \mathbb{R}^n$

Necessary optimality conditions

$\exists \varepsilon$ s.t. $\forall y$ $\|x^* - y\| \leq \varepsilon$ $f(x^*) \leq f(y)$

Let $x^\star$ local minimum of a function $f : \mathbb{R}^n \longrightarrow \mathbb{R}$.

1. First order optimality condition : if f is differentiable on an open neighborhood V of $x^\star$, then $\nabla f(x^\star) = 0$
2. Second order optimality condition : moreover if f is twice differentiable on V , then H$f(x^\star)$ is positive semi definite only and f is locally convex in $x^\star$.

Scalar case: $f : \mathbb{R} \rightarrow \mathbb{R}$

1) if $f(x^*)$ local minimum $f'(x^*) = 0$

2) $f''(x^*) \geq 0$ f locally convex in x^*

$f''(x^*) \geq 0$ in a neighborhood of x^*

suppose $\boxed{f(x^*) \le f(x)}$ for $\|x - x^*\| \le \varepsilon$

$\lambda \ge 0$ $\quad f(x^* + \lambda(x - x^*)) = f(x^*) + \langle \nabla f(x^*), \lambda(x - x^*)\rangle + o(\|\lambda(x - x^*)\|)$

$\underbrace{f(x^* + \lambda(x - x^*)) - f(x^*)}_{\ge 0} = \lambda \langle \nabla f(x^*), x - x^*\rangle + o(|\lambda|)$

for λ small : divide by λ $\quad 0 \le \langle \nabla f(x^*), x - x^*\rangle + \frac{o(|\lambda|)}{\lambda}$

$\lambda \le 0 \quad \lambda \langle \nabla f(x^*), x - x^*\rangle + o(|\lambda|) \ge 0$

divide by λ : $\quad \langle \nabla f(x^*), x - x^*\rangle + \frac{o(|\lambda|)}{\lambda} \le 0 \quad \lambda \to 0$

$\langle \nabla f(x^*), x - x^*\rangle \le 0$

$\langle \nabla f(x^*), x - x^*\rangle = 0 \quad \forall x \Rightarrow \nabla f(x^*) = 0$

2) $f(x^* + \lambda(x - x^*)) = f(x^*) + \frac{1}{2}\langle Hf(x^*)\lambda(x - x^*), \lambda(x - x^*)\rangle + o(\|\lambda(x - x^*)\|)$

$= f(x^*) + \boxed{\frac{1}{2}\lambda^2 \langle Hf(x^*)(x - x^*)\rangle + o(\lambda^2)}$

λ small enough $\quad f(x^* + \lambda(x - x^*)) - f(x^*) \ge 0$

divide by λ^2 $\quad \langle Hf(x^*), x - x^*\rangle \ge 0 \ \forall x$

Examples

$f'(x) = 4x^3 \quad f'(0) = 0$

$f''(x) = 12x^2 \quad f''(0) = 0$

Counter example : $f(x) = x^4$. global minimum $x^* = 0$

Counter example : $f(x) = x^3$. $f'(0) = 0 \quad f''(0) = 0$

the reciprocal of the necessary condition doesn't hold.

Sufficient optimality conditions

$f(x^* + \lambda(x - x^*)) = f(x^*) + \frac{1}{2}\lambda^2 \langle Hf(x^*)(x-x^*), x-x^* \rangle + o(\lambda^2)$

$\underbrace{f(x^* + \lambda(x-x^*)) - f(x^*)}_{\geq 0} = \frac{1}{2}\lambda^2 \Big(\underbrace{\underbrace{\langle Hf(x^*)(x-x^*), x-x^* \rangle}_{>0} + \frac{o(\lambda^2)}{\lambda^2}}_{\geq 0} \Big)$ for $\lambda < \lambda_0$

$\Rightarrow x^*$ is a local minimum

If f is twice differentiable in $x^\star$, and if $\nabla f(x^\star) = 0$ and moreover

- either $Hf(x^\star)$ is positive definite $\leftarrow f(x) = x^2$
- either f is twice differentiable in a neighborhood V of $x^\star$ and $Hf(x)$ is positive semi definite on V $\leftarrow f(x) = x^4$

then $x^\star$ is a strict (isolated) minimizer of f on V.

$f(x^* + \lambda(x-x^*)) - f(x^*) = \frac{1}{2}\lambda^2 \langle Hf(x^* + \lambda\tau(x-x^*))(x-x^*), x-x^* \rangle$

with $\tau \in]0,1[$

for λ small enough $x^* + \lambda(x - x^*) \in V$

$f(x^* + \lambda(x - x^*)) - f(x^*) \geq 0$ $\quad x^*$ local minimum

Uniqueness condition in the convex case

$f(x^*) \leq f(x)$ on V neighborhood of x^*, suppose it is not global

false

suppose that $\exists\, y \in C$ such that $\boxed{f(y) < f(x^*)}$

$$f(x^* + \lambda(y - x^*)) = f(\lambda y + (1-\lambda)x^*) \leq \lambda f(y) + (1-\lambda) f(x^*) \qquad \lambda \in [0,1]$$

for λ small enough $f(x^* + \lambda(y - x^*)) \geq f(x^*)$

$$f(x^*) \leq \lambda f(y) + (1-\lambda) f(x^*) \iff \boxed{\lambda f(x^*) \leq \lambda f(y)}$$

contradiction

(i) If f is convex on a convex subset $C \in \mathbb{R}^n$, any local minimum of f on C is global.

(ii) If f is strictly convex it has at most one global minimum.

x^* and y^* global minimum and $x^* \neq y^*$

$$f\left(\tfrac{1}{2}x^* + \tfrac{1}{2}y^*\right) < \tfrac{1}{2} f(x^*) + \tfrac{1}{2} f(y^*) = f(x^*)$$

contradiction with $f(x^*) \leq f(z)\ \forall z$

Necessary and sufficient optimality condition in the convex case

$\Rightarrow$ x^* global minimum, local $\Rightarrow \nabla f(x^*) = 0$

$\Leftarrow$ $\nabla f(x^*) = 0$ and f convex $\Rightarrow Hf(x) \in S_+^n$ on $\mathbb{R}^n$

x^* is local minimum $\Rightarrow$ global minimum

If f is convex on $\mathbb{R}^n$ and C^1, $x^\star \in \mathbb{R}^n$ realizes a global minimum of f if and only if $\nabla f(x^\star) = 0$.