LECTURE NOTE 06
Linhui Xie
December 1, 2022

6 Basics for optimization

6.1 Introduction

- Optimization problem:

$$\begin{array}{rl}\text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega\end{array}$$

  - Ω : constraint set or feasible set.
  - $f(\boldsymbol{x})$ is called the objective function or cost function;
  - $\Omega = \mathbb{R}^n$: **unconstrained** minimization.
  - $\Omega \subset \mathbb{R}^n$ explicitly given: set **constrained** minimization.
  - $\Omega = \{\boldsymbol{x} \in \mathbb{R}^n : g(\boldsymbol{x}) = 0, h(\boldsymbol{x}) \leq 0\}$: functional **constrained** minimization.
  - Solution to the problem: a minimizer, $\boldsymbol{x}^*$.

- Global minimizer: $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}$.

- Local minimizer: there exists $\varepsilon > 0$ such that $f(\boldsymbol{x}^*) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega$ and $\|\boldsymbol{x} - \boldsymbol{x}^*\| < \varepsilon$.

- **Strict** global minimizer: $f(\boldsymbol{x}^*) < f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}$.

- **Strict** local minimizer: there exists $\varepsilon > 0$ such that $f(\boldsymbol{x}^*) < f(\boldsymbol{x})$ for all $\boldsymbol{x} \in \Omega \setminus \{\boldsymbol{x}^*\}$ and $\|\boldsymbol{x} - \boldsymbol{x}^*\| < \varepsilon$.

- $\mathcal{THEOREM}$ 4.2 Theorem of Weierstrass: If f is continuous and Ω is *closed* and *bounded*, then a global minimizer exists.

6.2 Conditions: Necessary and Sufficient

- Two types of conditions: necessary and sufficient, that characterize minimizers.

- *Necessary* condition $\Longleftarrow$: If $\boldsymbol{x}^*$ is a minimizer, then $\boldsymbol{x}^*$ satisfies this particular condition.

- *Sufficient* condition $\Longrightarrow$: If $\boldsymbol{x}^*$ satisfies this particular condition, then $\boldsymbol{x}^*$ is a minimizer.

- A *necessary* condition limits the set of candidates for minimizers.

- A *sufficient* condition guarantees that a point is a minimizer.

- In this course, consider conditions that are based on **gradients** and **Hessians** and apply these conditions to local minimizers.

6.3 Reminder methods of proof

1. By direct method,
$$S_1 \Longrightarrow S_2.$$

2. By contraposition,
$$\begin{array}{ccc} S_1 & \Longrightarrow & S_2, \\ & \Updownarrow & \\ (\text{ NOT } S_2) & \Longrightarrow & (\text{ NOT } S_1). \end{array}$$

3. By contradiction,
$$\begin{array}{ccc} S_1 & \Longrightarrow & S_2, \\ & \Updownarrow & \\ \text{NOT } \Big(S_1, & \text{AND} & (\text{ NOT } S_2)\Big). \end{array}$$

6.4 Conditions for local minimizers

- An **unconstrained** problem, with assumption $f \in \mathcal{C}^1$, a real valued function on $\mathbb{R}^n$:

$$\begin{gathered} \text{minimize } f(\boldsymbol{x}) \\ \text{subject to } \boldsymbol{x} \in \mathbb{R}^n. \end{gathered}$$

♠ *THEOREM* If $\boldsymbol{x}^*$ is a local minimizer, then

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

- An **constrained** problem, with assumption $f \in \mathcal{C}^1$, a real valued function on Ω:

$$\begin{gathered} \text{minimize } f(\boldsymbol{x}) \\ \text{subject to } \boldsymbol{x} \in \Omega. \end{gathered}$$

♠ *COROLLARY* 6.1 If $\boldsymbol{x}^*$ is a local minimizer and an interior point of Ω, then

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

$$\boxed{\text{interior point } \boldsymbol{x}^* \text{ is a local minimizer,} \quad \xrightarrow{\text{FONC}} \quad \nabla f(\boldsymbol{x}^*) = \mathbf{0}.}$$

- Proof of theorem (by contraposition): suppose $\nabla f(\boldsymbol{x}^*) \neq \mathbf{0}$.
  - Since $-\nabla f(\boldsymbol{x}^*)$ points in the direction of decreasing f, there will be some points close to $\boldsymbol{x}^*$ that have smaller f value.
  - Specifically, consider $\boldsymbol{x}_\alpha = \boldsymbol{x}^* - \alpha \nabla f(\boldsymbol{x}^*), \alpha > 0$.
  - From an equation derived by Taylor's formula,

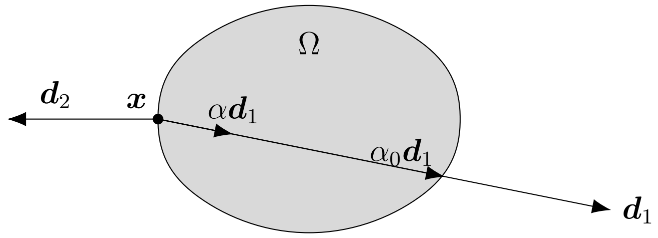
$$f(\boldsymbol{x}_\alpha) = f(\boldsymbol{x}^*) - \alpha \|\nabla f(\boldsymbol{x}^*)\|^2 + o(\alpha).$$

$$f(\boldsymbol{x}_\alpha) < f(\boldsymbol{x}^*).$$

$$\begin{array}{ccc} \neg S_2 & & \neg S_1 \\ \nabla f(\boldsymbol{x}^*) \neq \mathbf{0}, & \Longrightarrow & \boldsymbol{x}^* \text{ is definitely not a local minimizer.} \\ & \Updownarrow & \\ S_1 & & S_2 \\ \boldsymbol{x}^* \text{ is a local minimizer,} & \Longrightarrow & \nabla f(\boldsymbol{x}^*) = \mathbf{0}. \end{array}$$

6.5 Feasible directions

- We only consider a nonzero vector $\boldsymbol{d} \in \mathbb{R}^n$.
- $\boldsymbol{d}$ is a feasible direction at $\boldsymbol{x} \in \Omega$ if there exists $\alpha_0 > 0$ such that $\boldsymbol{x} + \alpha \boldsymbol{d} \in \Omega$ for all $\alpha \in [0, \alpha_0]$.
- At an *interior* point, all directions are feasible.
- At a *boundary* point, only some directions "point interior points of " the set are feasible.



- $\boldsymbol{d}_1$ is a feasible direction, whereas $\boldsymbol{d}_2$ is not a feasible direction.

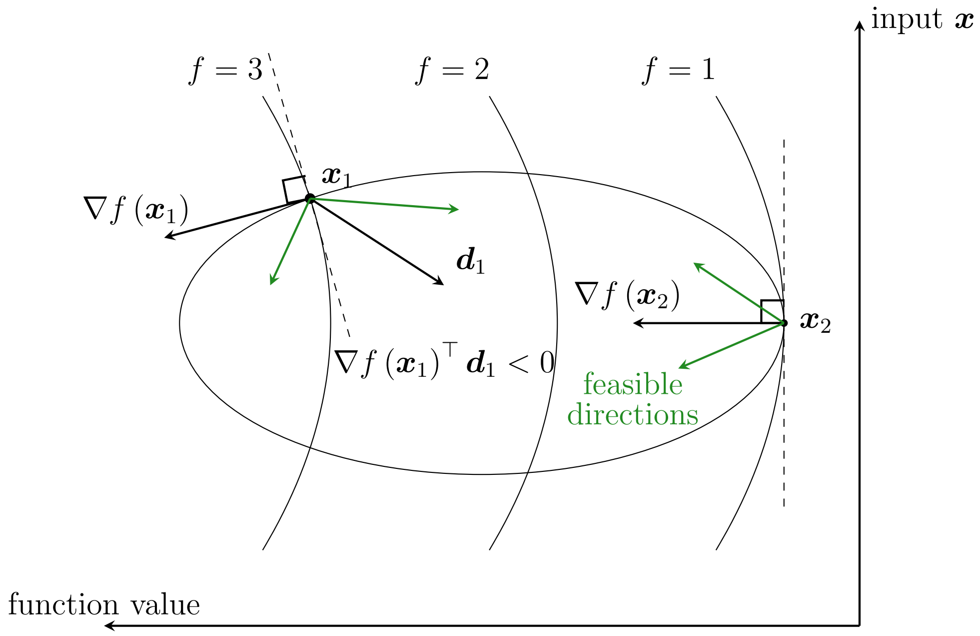
6.6 First order necessary conditions

- $\mathcal{THEOREM}$ 6.1 (FONC) If $\boldsymbol{x}^*$ is a local minimizer, then

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq \boldsymbol{0}$$

  for all feasible directions $\boldsymbol{d}$.

$$\boxed{\boldsymbol{x}^* \text{ is a local minimizer,} \quad \xRightarrow{\text{FONC}} \quad \boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0, \text{ for all feasible } \boldsymbol{d}.}$$



- $\boldsymbol{x}_1$ does not satisfy FONC, but $\boldsymbol{x}_2$ satisfies FONC.
- If $\boldsymbol{x}^*$ is a local minimizer, then the *directional derivative* of f in any feasible direction must be ≥ 0 (since the function must be increasing in that direction).
- General case: $\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq \boldsymbol{0}$ for all feasible directions $\boldsymbol{d}$.

- Example: consider the constrained problem

$$\begin{array}{ll} \text{minimize} & f(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in \Omega \end{array}$$

where $f \in \mathcal{C}^2$, $f : \mathbb{R}^2 \to \mathbb{R}, \Omega = \left\{ \boldsymbol{x} = [x_1, x_2]^\top : x_1 \geq 1 \right\}, \boldsymbol{x}^* = [1, 2]^\top$, and gradient $\nabla f(\boldsymbol{x}^*) = [1, 1]^\top$. Determine the type of minimizer for the given point $\boldsymbol{x}^*$.

Short answer,

- For $\boldsymbol{x}^* = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$, exists a *feasible direction* $\boldsymbol{d} = \begin{bmatrix} 1 \\ -5 \end{bmatrix}$,

$$\exists\ \alpha_0 = 1 > 0, \text{ s.t. } \boldsymbol{x}^* + \alpha \boldsymbol{d} = \begin{bmatrix} 1 + \alpha \\ 2 - 5\alpha \end{bmatrix} \in \Omega, \text{ for all } \alpha \in [0, \alpha_0].$$

- However,

$$\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) = \begin{bmatrix} 1 & -5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = -4 < 0$$

violates the FONC.

$$\underbrace{\boldsymbol{x}^* \text{ is a local minimizer,}}^{S_1} \quad \xRightarrow{\text{FONC}} \quad \overbrace{\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) \geq 0, \text{ for all feasible } \boldsymbol{d}.}^{S_2}$$

- Based on truth table and by contraposition,

$$S_1 \Rightarrow S_2, \quad \Longleftrightarrow \quad \neg S_2 \Rightarrow \neg S_1.$$

$$\overbrace{\boldsymbol{d}^\top \nabla f(\boldsymbol{x}^*) < 0, \text{ for some } \boldsymbol{d},}^{\neg S_2} \quad \Longrightarrow \quad \overbrace{\boldsymbol{x}^* \text{ is definitely not a local minimizer.}}^{\neg S_1}$$

[Ref]: Edwin K.P. Chong, Stanislaw H. Żak, "PART II UNCONSTRAINED OPTIMIZATION" in "An introduction to optimization", 4th Edition, John Wiley and Sons, Inc. 2013.