

# 第四章 最优性理论

修贤超

<https://xianchaoxiu.github.io>

- 4.1 最优化问题解的存在性
- 4.2 无约束可微问题的最优性理论
- 4.3 无约束不可微问题的最优性理论
- 4.4 对偶理论
- 4.5 一般约束优化问题的最优性理论
- 4.6 带约束凸优化问题的最优性理论

# 最优化问题解的存在性

## ■ 考虑优化问题

$$\begin{array}{ll}\min_{x \in \mathbb{R}^n} & f(x) \\ \text{s.t.} & x \in \mathcal{X}\end{array}$$

- 首先分析最优解的存在性
- 然后考虑如何求出其最优解

- (Weierstrass 定理) 紧集上的连续函数一定存在最大 (最小) 值
- 而在许多实际问题中, 定义域可能不是紧的, 目标函数也不一定连续

# 推广的 Weierstrass 定理

- 若函数  $f : \mathcal{X} \rightarrow (-\infty, +\infty]$  适当且闭, 且以下条件中任意一个成立

- $\text{dom } f = \{x \in \mathcal{X} : f(x) < +\infty\}$  是有界的
- 存在一个常数  $\bar{\gamma}$  使得下水平集

$$C_{\bar{\gamma}} = \{x \in \mathcal{X} \mid f(x) \leq \bar{\gamma}\}$$

是非空且有界的

- $f$  是强制的, 即对于任一满足  $\|x^k\| \rightarrow +\infty$  的点列  $\{x^k\} \subset \mathcal{X}$ , 都有

$$\lim_{k \rightarrow \infty} f(x^k) = +\infty$$

则函数  $f$  的最小值点集  $\{x \in \mathcal{X} \mid f(x) \leq f(y), \forall y \in \mathcal{X}\}$  非空且紧

- 三个条件在本质上都是**保证  $f(x)$  的最小值不能在无穷远处取到**

# 例子

- 当定义域不是有界闭集时, 对于强制函数

$$f(x) = x^2, x \in \mathcal{X} = \mathbb{R}$$

其全局最优解一定存在

- 对于适当且闭的函数

$$f(x) = e^{-x}, x \in \mathcal{X} = \mathbb{R}$$

不满足三个条件中任意一个, 因此不能断言其全局极小值点存在

- 4.1 最优化问题解的存在性
- 4.2 无约束可微问题的最优性理论
- 4.3 无约束不可微问题的最优性理论
- 4.4 对偶理论
- 4.5 一般约束优化问题的最优性理论
- 4.6 带约束凸优化问题的最优性理论

# 无约束可微问题的最优性理论

- 无约束可微优化问题通常表示为

$$\min_{x \in \mathbb{R}^n} f(x) \quad (1)$$

- 对于可微函数  $f$  和点  $x \in \mathbb{R}^n$ , 如果存在向量  $d$  满足

$$\nabla f(x)^\top d < 0$$

那么称  $d$  为  $f$  在点  $x$  处的一个下降方向

- 一阶最优性条件是利用梯度 (一阶) 信息来判断给定点的最优性
- 在局部最优点处不能有下降方向

# 一阶必要条件

- 假设  $f$  在全空间  $\mathbb{R}^n$  可微. 如果  $x^*$  是 (1) 的一个局部极小点, 那么

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

**证明** 任取  $v \in \mathbb{R}^n$ , 考虑  $f$  在点  $x = x^*$  处的泰勒展开

$$\frac{f(x^* + tv) - f(x^*)}{t} = v^\top \nabla f(x^*) + o(1)$$

根据  $x^*$  的最优性, 分别对  $t$  取点 0 处的左、右极限可知

$$\begin{aligned}\lim_{t \rightarrow 0^+} \frac{f(x^* + tv) - f(x^*)}{t} &= v^\top \nabla f(x^*) \geq 0 \\ \lim_{t \rightarrow 0^-} \frac{f(x^* + tv) - f(x^*)}{t} &= v^\top \nabla f(x^*) \leq 0\end{aligned}$$

- 称满足  $\nabla f(x) = 0$  的点  $x$  为  $f$  的**稳定点 (或驻点、临界点)**



## 二阶最优性条件

- 对于  $f(x) = x^3$ , 满足  $f'(x) = 0$  的点为  $x^* = 0$ , 但其不是局部最优解
- 假设  $f$  在点  $x$  的一个开邻域内是二阶连续可微的, 考虑

$$f(x + d) = f(x) + \nabla f(x)^\top d + \frac{1}{2}d^\top \nabla^2 f(x)d + o(\|d\|^2)$$

则以下最优性条件成立

- (二阶必要条件) 若  $x^*$  是  $f$  的一个局部极小点, 则

$$\nabla f(x^*) = 0, \nabla^2 f(x^*) \succeq 0$$

- (二阶充分条件) 若满足

$$\nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$$

则  $x^*$  是  $f$  的一个局部极小点

- **必要性** 若  $\nabla^2 f(x^*)$  有负的特征值  $\lambda_- < 0$ , 设  $\nabla^2 f(x^*)d = \lambda_- d$ , 则

$$\frac{f(x^* + d) - f(x^*)}{\|d\|^2} = \frac{1}{2} \frac{d^\top}{\|d\|} \nabla^2 f(x^*) \frac{d}{\|d\|} + o(1) = \frac{1}{2} \lambda_- + o(1)$$

当  $\|d\|$  充分小时,  $f(x^* + d) < f(x^*)$ , 这和点  $x^*$  的最优性矛盾

- **充分性** 由  $\nabla f(x^*) = 0$  时的二阶展开

$$\frac{f(x^* + d) - f(x^*)}{\|d\|^2} = \frac{d^\top \nabla^2 f(x^*) d + o(\|d\|^2)}{\|d\|^2} \geq \frac{1}{2} \lambda_{\min} + o(1)$$

当  $\|d\|$  充分小时有  $f(x^* + d) \geq f(x^*)$ , 即二阶充分条件成立

# 实例：实数情形的相位恢复

## ■ 考虑

$$\min_{x \in \mathbb{R}^n} f(x) = \sum_{i=1}^m r_i^2(x)$$

其中  $r_i(x) = (a_i^\top x)^2 - b_i^2, i = 1, 2, \dots, m$

## ■ 计算梯度和的海瑟矩阵

$$\nabla f(x) = 2 \sum_{i=1}^m r_i(x) \nabla r_i(x) = 4 \sum_{i=1}^m ((a_i^\top x)^2 - b_i^2) (a_i^\top x) a_i$$

$$\nabla^2 f(x) = \sum_{i=1}^m (12(a_i^\top x)^2 - 4b_i^2) a_i a_i^\top$$

- 4.1 最优化问题解的存在性
- 4.2 无约束可微问题的最优性理论
- 4.3 无约束不可微问题的最优性理论
- 4.4 对偶理论
- 4.5 一般约束优化问题的最优性理论
- 4.6 带约束凸优化问题的最优性理论

# 无约束不可微问题的最优性理论

## ■ 考虑不可微优化问题

$$\min_{x \in \mathbb{R}^n} f(x) \quad (2)$$

■ 假设  $f$  是适当且凸的函数, 则  $x^*$  为 (2) 的全局极小点当且仅当  $0 \in \partial f(x^*)$

□ **必要性** 因  $x^*$  为全局极小点, 有

$$f(y) \geq f(x^*) = f(x^*) + 0^\top (y - x^*), \quad \forall y \in \mathbb{R}^n$$

$$\implies 0 \in \partial f(x^*)$$

□ **充分性** 如果  $0 \in \partial f(x^*)$ , 那么根据次梯度的定义

$$f(y) \geq f(x^*) + 0^\top (y - x^*) = f(x^*), \quad \forall y \in \mathbb{R}^n$$

$$\implies x^* \text{ 为一个全局极小点}$$

# 复合优化问题的一阶必要条件

- 考虑一般复合优化问题

$$\min_{x \in \mathbb{R}^n} \quad \psi(x) = f(x) + h(x) \quad (3)$$

其中  $f$  为光滑函数 (可能非凸),  $h$  为凸函数 (可能非光滑)

- **定理 4.5** 令  $x^*$  为复合优化问题 (3) 的一个局部极小点, 那么

$$-\nabla f(x^*) \in \partial h(x^*)$$

- 由于目标函数可能是整体非凸的, 因此一般没有一阶充分条件

## 实例: $\ell_1$ 范数优化问题

### ■ 考虑

$$\min_{x \in \mathbb{R}^n} \quad \psi(x) = f(x) + \mu \|x\|_1$$

### ■ $\|x\|_1$ 不是可微的, 但可以计算其次微分

$$\partial_i \|x\|_1 = \begin{cases} \{1\}, & x_i > 0 \\ [-1, 1], & x_i = 0 \\ \{-1\}, & x_i < 0 \end{cases}$$

### ■ 若 $x^*$ 是局部最优解, 则 $-\nabla f(x^*) \in \mu \partial \|x^*\|_1$ , 即

$$\nabla_i f(x^*) = \begin{cases} -\mu, & x_i^* > 0 \\ a \in [-\mu, \mu], & x_i^* = 0 \\ \mu, & x_i^* < 0 \end{cases}$$

## ■ 无约束优化问题及其最优性条件

问题	一阶条件	二阶条件
可微问题	$\nabla f(x^*) = 0$ (必要)	必要/充分
凸问题	$0 \in \partial f(x^*)$ (充要)	—
复合优化问题	$-\nabla f(x^*) \in \partial h(x^*)$ (必要)	—
非凸非光滑	$0 \in \partial f(x^*)$ (必要)	—



*Q&A*

*Thank you!*

感谢您的聆听和反馈