南昌大学
NANCHANG UNIVERSITY

大数据优化：理论、算法及其应用

——来源于《最优化计算方法》（高教社）

★★★ 主讲人：彭振华 ★★★

联系方式：zhenhuapeng@ncu.edu.cn
zhenhuapeng@whu.edu.cn
15870605317（微信同号）

研究兴趣：1. 非凸非光滑优化算法与理论
2. 智能决策
3. 智能计算与机器学习

数学与计算机学院

2026年

课件资源：https://zhenhuapeng.github.io//coursematerials/

目录

南昌大学
NANCHANG UNIVERSITY

无约束优化理论

第 三 部分

存在性

可微情形

不可微情形

解的存在性

$$\min_{x} f(x) \quad s.t. \ x \in \chi.$$

- 经典 Weierstrass 极值定理：下半连续 （lower semi-continuous）的函数 f 在紧集 C 上一定存在最小值点.
- 下半连续函数：若对任意x, 有 $\liminf_{y\to x} f(y) \geq f(x)$

在许多实际问题中, 定义域可能不是紧的, 目标函数也不一定下半连续, 因此需要将此定理推广来保证最优化问题解的存在性.

- 推广的Weierstrass 定理：若函数f 适当且闭, 且以下条件中任意一个成立:

I. $\mathrm{dom}\, f = \{x \in X : f(x) < +\infty\}$ 是有界的;

II. 存在一个常数γ 使得下水平集$C_\gamma = \{x \in X : f(x) \leq \gamma\}$ 是非空且有界的;

III. f 是强制的, 即对于任一满足$\|x_k\| \to +\infty$ 的点列$\{x_k\} \subset X$, 都有$\lim_{k\to\infty} f(x_k) = +\infty$,

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

解的唯一性

➢ **强拟凸函数：** 给定凸集X和函数f. 若任取$x \neq y$和$\lambda \in (0, 1)$, 有

$$f(\lambda x + (1-\lambda)y) < max\{f(x), f(y)\}$$

□ **任意强凸函数均为强拟凸的, 但凸函数并不一定是强拟凸的.**

□ **唯一性：** 设X是R^n的一个非空, 紧且凸的子集, 如果f是适当, 闭且强拟凸函数, 那么存在唯一的x^*满足

$$f(x^*) < f(x), \ \forall x \in X \backslash \{x\}.$$

对于一般的凸函数, 其最优解可能不唯一

$$\min_x \ f(x)$$

➢ 情形一：函数 f 是连续可微的

如何判断 x^* 是否是函数 f 的一个局部极小解或者全局极小解?

□ (一阶必要条件) 假设 f 在全空间 R^n 可微. 如果 x^* 是一个局部极小点, 那么

$\nabla f(x^*) = 0$. (若函数 f 是凸的，则 $\nabla f(x^*) = 0 \to x^*$ 是 f 的全局极小点)

Proof. 根据 x^* 的最优性, 在上式中进行整理分别对 t 取点 0 处的左, 右极限可知

$$\lim_{t\to 0+} f(x^* + tv) - f(x^*) / t = v^T \nabla f(x^*) \geq 0,$$

$$\lim_{t\to 0-} f(x^* + tv) - f(x^*) / t = v^T \nabla f(x^*) \leq 0,$$

即对任意的 v 有 $v^T \nabla f(x^*) = 0$, 由 v 的任意性知 $\nabla f(x^*) = 0$.

现在能否判断x^*是否是函数 f 的一个局部极小解?

- (**二阶最优性条件**)

必要条件: 若x^* 是 f 的一个局部极小点, 则 $\nabla f(x^*) = 0, \nabla^2 f(x^*) \succeq 0$.

充分条件: 若 $\nabla f(x^*) = 0, \nabla^2 f(x^*) \succ 0$, 则$x^*$ 是 f 的一个局部极小点.

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

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

上式 $= 1/2\lambda_- + o(1)$

充分性: 上式 $\geq 1/2\lambda_{min} + o(1)$

最优性理论

- **线性最小二乘**

$$(VI)\ \min_{\theta} \|X\theta - y\|^2$$

- **非线性最小二乘**

$$\min_{\theta}\ f(\theta) = \sum_{i=1}^{N} r_i^2(\theta), r_i(\theta) = \left(a_i^T\theta\right)^2 - b_i^2$$

- **神经网络模型**

$$\min_{w_1, w_2} \frac{1}{N} \sum_{i=1}^{N} \ell\left(\sigma_2\left(w_2^T \sigma_1\left(w_2^T x_i\right)\right), y_i\right)$$

最优性理论

- import numpy as np
- import gurobipy as gp
- from gurobipy import GRB

- m = gp.Model("my_model")
- x = m.addVars(range(3),vtype=GRB.CONTINUOUS, name="x")
- m.setObjective(x[2] ** 2 + x[1] ** 2 - x[1] * x[2] - x[2], GRB.MINIMIZE)
- m.addGenConstrSin(x[0], x[2])
- m.Params.NonConvex=2
- m.optimize()
- for i in x.keys():
-     print('x[%d] = %f' % (i,x[i].x))

- def Gradientob(x):
-     return np.array([2*np.sin(x[0])*np.cos(x[0])-x[1]*np.cos(x[0])-np.cos(x[0]),2*x[1]-np.sin(x[0])])
- def Hessianob(x):
-     return np.array([[2*np.cos(2*x[0])+x[1]*np.sin(x[0])+np.sin(x[0]),-np.cos(x[0])],[-np.cos(x[0]),2]])
- print(Gradientob([x[0].x,x[1].x]))
- A = Hessianob([x[0].x,x[1].x])
- eigenvalues, eigenvectors = np.linalg.eig(A)
- print("特征值：", eigenvalues)

```
Optimal solution found (tolerance 1.00e-04)
Best objective -3.333333333333e-01, best bound -3.333333333333e-01, gap 0.0000%
x[0] = 101.260693
x[1] = 0.333333
x[2] = 0.666667
[ 1.31684663e-11 -8.83371154e-12]
特征值:  [0.68775004 2.42336108]
```

$$\min_x \ f(x)$$

➢ 情形二：函数 f 是不可微的

如何判断 x^* 是否是函数 f 的一个局部极小解或者全局极小解?

- **(一阶充要条件)** 假设 f 是适当且凸的函数.. 如果 x^* 是一个全局极小点, 当且仅当 $0 \in \partial f(x^*)$.

Proof. $f(y) \geq f(x^*) + 0^T (y - x^*), \ \forall y.$

- 假设 f 是可微的, h 是凸函数(可能非光滑). 如果 x^* 是 $f + h$ 的一个局部极小点, 当且仅当 $-\nabla f(x^*) \in \partial h(x^*)$.

1. 写出Lasso问题和神经网络模型的一阶充要条件？

2. 对于非凸优化，其最优性理论如何？

课件资源：https://zhenhuapeng.github.io//coursematerials/

南昌大学
NANCHANG UNIVERSITY

感　谢　观　看

大数据优化：理论、算法及其应用

——来源于《最优化计算方法》（高教社）

★★★　主讲人：彭振华　★★★

联系方式：zhenhuapeng@ncu.edu.cn
zhenhuapeng@whu.edu.cn
15870605317（微信同号）

研究兴趣：1. 非凸非光滑优化算法与理论
2. 智能决策
3. 智能计算与机器学习

数学与计算机学院