Optimal Theory and Method

最优化理论与方法

程春杰

杭州电子科技大学 自动化学院

科技馆512

Email: cjzhai@hdu.edu.cn

目录

- 无约束问题的极值条件（考）
- 约束极值的最优性条件（**K-T**条件）（考）

目录

- 无约束问题的极值条件（考）
- 约束极值的最优性条件（**K-T**条件）（考）

■ 无约束问题的极值条件

- 什么是无约束极值问题
- 无约束极值的必要条件
- 无约束极值的二阶充分条件
- 无约束极值的充要条件

■ 无约束问题的极值条件

➢无约束极值问题

考虑非线性规划问题

$$\min f(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbf{R}^n,$$

这个问题是求$f(x)$在 n维欧氏空间中的极小点，称为无约束极值问题

■ 无约束问题的极值条件

➢无约束极值的必要条件

泰勒展开式

设 $f(x)$ 在点 x_0 的邻域 $U(x_0)$ 内有连续的 $n+1$ 阶导数，则 $\forall x \in U(x_0)$，有

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+R_n(x),$$

其中 $R_n(x)=\dfrac{1}{n!}\displaystyle\int_{x_0}^{x} f^{(n+1)}(x)(x-t)^n\,dt$

对 $R_n(x)=\dfrac{1}{n!}\displaystyle\int_{x_0}^{x} f^{(n+1)}(x)(x-t)^n\,dt$ 右边用积分中值定理，有

$$R_n(x)=\frac{1}{n!}\int_{x_0}^{x} f^{(n+1)}(x)(x-t)^n\,dt=\frac{1}{n!}f^{(n+1)}(\xi)\int_{x_0}^{x}(x-t)^n\,dt$$

$$=\frac{1}{(n+1)!}f^{(n+1)}(\xi)(x-x_0)^{n+1}=(x-x_0)^n\alpha(x_0,x-x_0)$$

■ 无约束问题的极值条件

➢无约束极值的必要条件

$$f(x) = f(\overline{x}) + \Delta f(\overline{x})^T(x-\overline{x}) + \frac{1}{2}(x-\overline{x})^T\Delta^2 f(\overline{x})(x-\overline{x}) + \cdots$$

一阶泰勒展开式

$$f(x) = f(\overline{x}) + \Delta f(\overline{x})^T(x-\overline{x}) + R_2(x)$$

$$R_2(x) = \frac{1}{2}(x-\overline{x})^T\Delta^2 f(\xi)(x-\overline{x})$$

$$= \|x-\overline{x}\|\alpha(\overline{x}, x-\overline{x})$$

二阶泰勒展开式

$$f(x) = f(\overline{x}) + \Delta f(\overline{x})^T(x-\overline{x}) + \frac{1}{2}(x-\overline{x})^T\Delta^2 f(\overline{x})(x-\overline{x}) + R_3(x)$$

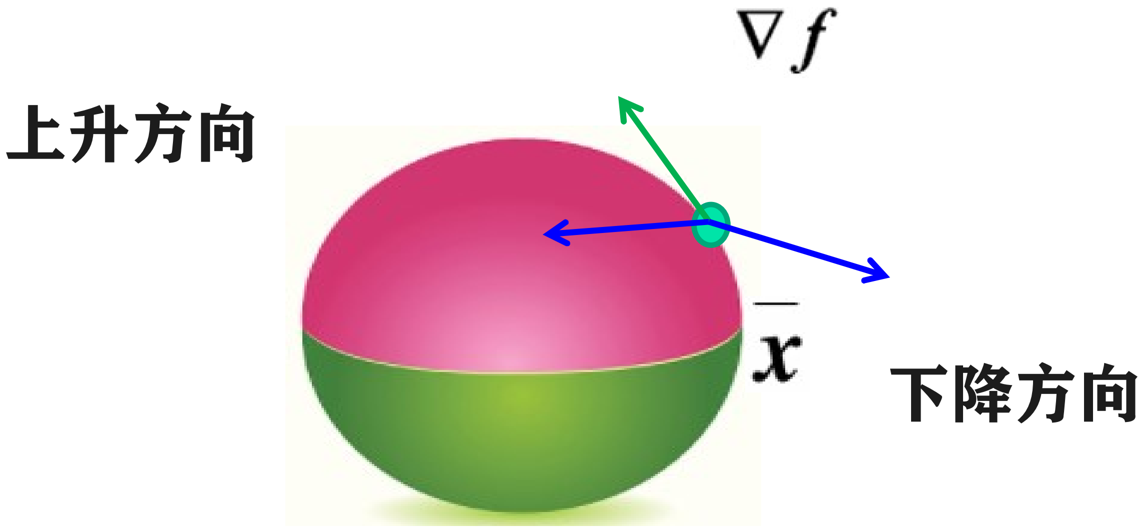
$$R_3(x) = \|x-\overline{x}\|^2\alpha(\overline{x}, x-\overline{x})$$

■ 无约束问题的极值条件

➢无约束极值的必要条件

定理 7.1.1

设函数 $f(\boldsymbol{x})$在点 $\bar{\boldsymbol{x}}$ 可微，如果存在方向 $\boldsymbol{d}$，使$\nabla f(\bar{\boldsymbol{x}})^{\mathrm{T}}\boldsymbol{d}<0$，则存在数 $\delta>0$，使得对每个 $\lambda\in(0,\delta)$，有 $f(\bar{\boldsymbol{x}}+\lambda\boldsymbol{d})<f(\bar{\boldsymbol{x}})$.



■ 无约束问题的极值条件

➢无约束极值的必要条件

证明 根据函数可微性的定义，有

$$f(\bar{x}+\lambda d)=f(\bar{x})+\lambda\nabla f(\bar{x})^T d+\|\lambda d\|\alpha(\bar{x},\lambda d)$$

$$=f(\bar{x})+\lambda\underbrace{\left[\underbrace{\nabla f(\bar{x})^T d}_{<0}+\frac{|\lambda|}{\lambda}\|d\|\underbrace{\alpha(\bar{x},\lambda d)}_{\lambda\to 0\quad 0}\right]}_{<0}$$

存在数δ >0，使得对每个λ ∈ (0， δ)

■ 无约束问题的极值条件

➢无约束极值的必要条件

定理 7.1.2

设函数 $f(x)$在点 $\bar{x}$ 可微，若 $\bar{x}$ 是局部极小点，

则梯度 $\Delta f(\bar{x})=\mathbf{0}$

证明 用反证法．设$\Delta f(\bar{x})\neq\mathbf{0}$，令方向 $\boldsymbol{d}=-\Delta f(\bar{x})$，则有

$$\Delta f(\bar{x})^{\mathrm{T}}\boldsymbol{d}=-\Delta f(\bar{x})^{\mathrm{T}}\Delta f(\bar{x})=-\|\Delta f(\bar{x})\|^{2}<0$$

必存在 $\delta>0$，使得当 $\lambda\in(0,\delta)$时，成立

$$f(\bar{x}+\lambda\boldsymbol{d})<f(\bar{x}),$$

这与 $\bar{x}$ 是局部极小点矛盾．

■ 无约束问题的极值条件

➢无约束极值的必要条件

定理 7.1.3

设函数 $f(\boldsymbol{x})$在点 $\bar{\boldsymbol{x}}$ 处二次可微，若 $\bar{\boldsymbol{x}}$ 是局部极小点，则梯度$\Delta f(\bar{\boldsymbol{x}})=\mathbf{0}$，并且 Hesse 矩阵$\Delta^2 f(\bar{\boldsymbol{x}})$半正定．

证明**Hessian**矩阵半正定

设**d**是任意一个**n**维向量，由于**f(x)**在$\bar{\mathbf{x}}$处二次可微，且$\nabla f(\bar{\mathbf{x}})=0$，因此有

$$f(\bar{x}+\lambda d)=f(\bar{x})+\frac{1}{2}\lambda^2 d^T\nabla^2 f(\bar{x})d+\lambda^2\|d\|^2\alpha(\bar{x},\lambda d)$$

$$\frac{f(\bar{x}+\lambda d)-f(\bar{x})}{\lambda^2}=\frac{1}{2}d^T\nabla^2 f(\bar{x})d+\|d\|^2\alpha(\bar{x},\lambda d)$$

$\downarrow$ $\qquad\qquad\qquad\qquad$ $\downarrow$ $\qquad\qquad\qquad\qquad$ $\downarrow$

$\geqslant 0$,极小点 $\longrightarrow$ $\geqslant 0$ $\qquad\qquad\qquad$ 0

■ 无约束问题的极值条件

➢无约束极值的二阶充分条件

定理 7.1.4

设函数 $f(\boldsymbol{x})$在点 $\bar{\boldsymbol{x}}$ 处二次可微，若梯度$\Delta f(\bar{\boldsymbol{x}})=\mathbf{0}$，且 Hesse 矩阵 $\Delta^2 f(\bar{\boldsymbol{x}})$正定，则 $\bar{\boldsymbol{x}}$ 是局部极小点．

证明：反正法，假如$\bar{x}$不是极小点，则必有下降方向d≠0，**存在**θ，**当$0<\lambda<\theta$时满足**

$$\boldsymbol{f}(\bar{x}+\lambda \boldsymbol{d})<\boldsymbol{f}(\bar{\boldsymbol{x}})$$

$$\boldsymbol{f}(\bar{x}+\lambda \boldsymbol{d})=\boldsymbol{f}(\bar{\boldsymbol{x}})+\lambda\Delta \boldsymbol{f}(\bar{\boldsymbol{x}})^T\boldsymbol{d}+\frac{1}{2}\lambda^2\boldsymbol{d}^T\Delta^2\boldsymbol{f}(\bar{\boldsymbol{x}})\boldsymbol{d}+\boldsymbol{R}_3(\boldsymbol{x})$$

$$\underbrace{\frac{\boldsymbol{f}(\bar{x}+\lambda \boldsymbol{d})-\boldsymbol{f}(\bar{\boldsymbol{x}})}{\lambda^2}}_{<0}=\frac{1}{2}\underbrace{\boldsymbol{d}^T\Delta^2\boldsymbol{f}(\bar{\boldsymbol{x}})\boldsymbol{d}}_{>0}+\|\boldsymbol{d}\|^2\underbrace{\boldsymbol{\alpha}(\bar{\boldsymbol{x}},\lambda\boldsymbol{d})}_{\text{趋近于}0}$$

$f(\boldsymbol{x}^{(2)}) \geqslant f(\boldsymbol{x}^{(1)}) + \Delta f(\boldsymbol{x}^{(1)})^{\mathrm{T}}(\boldsymbol{x}^{(2)} - \boldsymbol{x}^{(1)})$,

■ 无约束问题的极值条件

■ 无约束极值的充要条件

定理 7.1.5

设 $f(\boldsymbol{x})$是定义在$\mathbf{R}^n$上的可微凸函数，$\bar{\boldsymbol{x}} \in \mathbf{R}^n$，则 $\bar{\boldsymbol{x}}$ 为全极小点的充分必要条件是梯度$\Delta f(\bar{\boldsymbol{x}}) = \mathbf{0}$

证明

必要性是显然的，若 $\bar{\boldsymbol{x}}$ 是全局极小点，自然是局部极小点

$$\Delta f(\bar{\boldsymbol{x}}) = \mathbf{0}$$

设$\Delta f(\bar{\boldsymbol{x}}) = \mathbf{0}$，则对任意的 $\boldsymbol{x} \in \mathbf{R}^n$，有$\Delta f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) = 0$，

由于 $f(\boldsymbol{x})$是可微凸函数，根据定理 1.4.14，

$$f(\boldsymbol{x}) \geqslant f(\bar{\boldsymbol{x}}) + \Delta f(\bar{\boldsymbol{x}})^{\mathrm{T}}(\boldsymbol{x} - \bar{\boldsymbol{x}}) = f(\bar{\boldsymbol{x}}),$$

无约束问题的极值条件

无约束极值的充要条件

例7.1.2 利用极值条件解下列问题：

$$\min \quad f(x) \triangleq (x_1^2-1)^2+x_1^2+x_2^2-2x_1$$

先求驻点：$\dfrac{\partial f}{\partial x_1}=4x_1^3-2x_1-2$

令$\nabla f(x)=0$

$$\dfrac{\partial f}{\partial x_2}=2x_2$$

$\Rightarrow \quad =0$

驻点 $\bar{x}=(\bar{x}_1, \bar{x}_2)^T=(1, 0)^T$

无约束问题的极值条件

无约束极值的充要条件

再利用极值条件判断$\bar{x}$是否为极小点。 由于目标函数的Hessian矩阵

$$\nabla^2 f(x)=\begin{bmatrix}12x_1^2-2 & 0\\ 0 & 2\end{bmatrix}$$

$$\rightarrow \quad \nabla^2 f(\bar{x})=\begin{bmatrix}10 & 0\\ 0 & 2\end{bmatrix} \quad \text{正定}$$

$\rightarrow$ $\bar{x}$是局部极小点

■ 作业

P243、P244：

4、5