运筹学与优化方法

晁国清
计算机科学与技术学院

课程回顾

- 对偶范数

对偶范数$\|x\|_*$定义为

$$\|x\|_* = \max_{\|z\|\le 1} z^T x$$

含义：对于一个范数小于等于1的向量z, z与x的内积最大值就是x的对偶范数

- 共轭函数

- 给定一个函数$f: R^n \to R$, 定义它的共轭函数$f^*: R^n \to R$,

$$f^*(y) = \max_x y^T x - f(x)$$

- 对偶锥

- 对于一个锥$K \subseteq \mathbb{R}^n$ $(x \in K, t \ge 0 \Rightarrow tx \in K)$, 其对偶锥定义为：

$$K^* = \{y: y^T x \ge 0 \ \ for\ all\ x \in K\}$$

- 对偶技巧

- 推导无约束问题对偶的一个常见技巧是增加一个哑变量

(dummy variable)和等式约束来转换原问题

不同的选择会产生不同的对偶问题

课程内容

- 制约函数法
- 外点法(惩罚函数法)
- 内点法(障碍函数法)
- 乘子法

制约函数法

基本思想：通过构造制约函数，把约束优化问题转化为一系列无约束优化问题，进而用无约束优化方法求解。

该方法也被称为序列无约束最优化方法SUMT(Sequential Unconstrained Minimization Technique)

常用的制约函数法分为两类：惩罚函数法(外点法)和障碍函数法(内点法)

惩罚函数法

1. 惩罚函数概念：

$$
(fghD)\begin{cases} \min f(x) & f: \mathbb{R}^n \to \mathbb{R} \\ s.t.\, g(x) \le 0 & g: \mathbb{R}^n \to \mathbb{R}^m \\ h(x) = 0 & h: \mathbb{R}^n \to \mathbb{R}^l \\ x \in D & D \subset \mathbb{R}^n \end{cases}
$$

构造惩罚函数：

$$
\alpha(x) = \sum_{i=1}^{m} \phi\big(g_i(x)\big) + \sum_{j=1}^{l} \varphi(h_j(x))
$$

目的：使满足约束的x有$\alpha(x) = 0$

不满足约束的x有$\alpha(x) > 0$

惩罚函数法

其中：$\phi(t) = \begin{cases} > 0, & if\ \ t > 0 \\ = 0, & if\ t \leq 0 \end{cases}$

$$\varphi(t) = \begin{cases} > 0, & if\ t \neq 0 \\ = 0, & if\ t = 0 \end{cases}$$

当$\mu > 0$, 可构造

$\mu\alpha(x) - -$惩罚项$\begin{cases} = 0, & 可行 \\ \to \infty & 不可行 \end{cases}$

$f(x) + \mu\alpha(x) - -$辅助函数

$\phi(t), \varphi(t)$的典型取法：

$\phi(t) = [\max\{0, t\}]^p$　　$\varphi(t) = |t|^p$　　　p为正整数。

当$p = 2$时，称2次罚函数。（常用：因2次是最低次的光滑函数）

惩罚函数法

举例：$\begin{cases} \min & x \\ s.t. & -x+2 \le 0 \end{cases}$

取二次惩罚函数：$\alpha(x) = [\max\{0, -x+2\}]^2 = \begin{cases} (x-2)^2, & x < 2 \\ 0, & x \ge 2 \end{cases}$

如图6-10所示

当$\mu \to \infty$时，$\min f(x) + \mu\alpha(x) \to f(x^*) = x^* = 2$

解析解：构造辅助函数

$$g(x,\mu) = f(x) + \mu\alpha(x)$$
$$= \begin{cases} x + \mu(x-2)^2 = \mu x^2 - (4\mu - 1)x + 4\mu, & x < 2 \\ x, & x \ge 2 \end{cases}$$

$$\left.\begin{array}{l} \text{当}x < 2\text{时，}g(x,\mu)\text{的驻点}\bar{x} = \dfrac{4\mu - 1}{2\mu} \xrightarrow{\mu \to \infty} 2 \\ \text{当}x \ge 2\text{时，}g(x,\mu)\text{的最小值点}\bar{x} = 2 \end{array}\right\} \text{故}x^* = 2 - -opt$$

惩罚函数法

注意：

1. 求解一般约束问题时，为方便起见，常把一部分易于处理的约束(如上下界约束、线性约束等)归入集合D, 从而使得到的辅助问题易于求解。特别地，当$D=\mathbb{R}^n$时，辅助问题就变成了无约束问题；

2. 惩罚函数可从任一点$x^{(1)}$开始，一般情况下$x^{(1)}\notin S, S=\{x|x\in D, g(x)\le 0, h(x)=0\}$, 随着$\mu$的增大，$x^{(k)}$逐步接近$S$，因此，惩罚函数法也称外点法。

图示

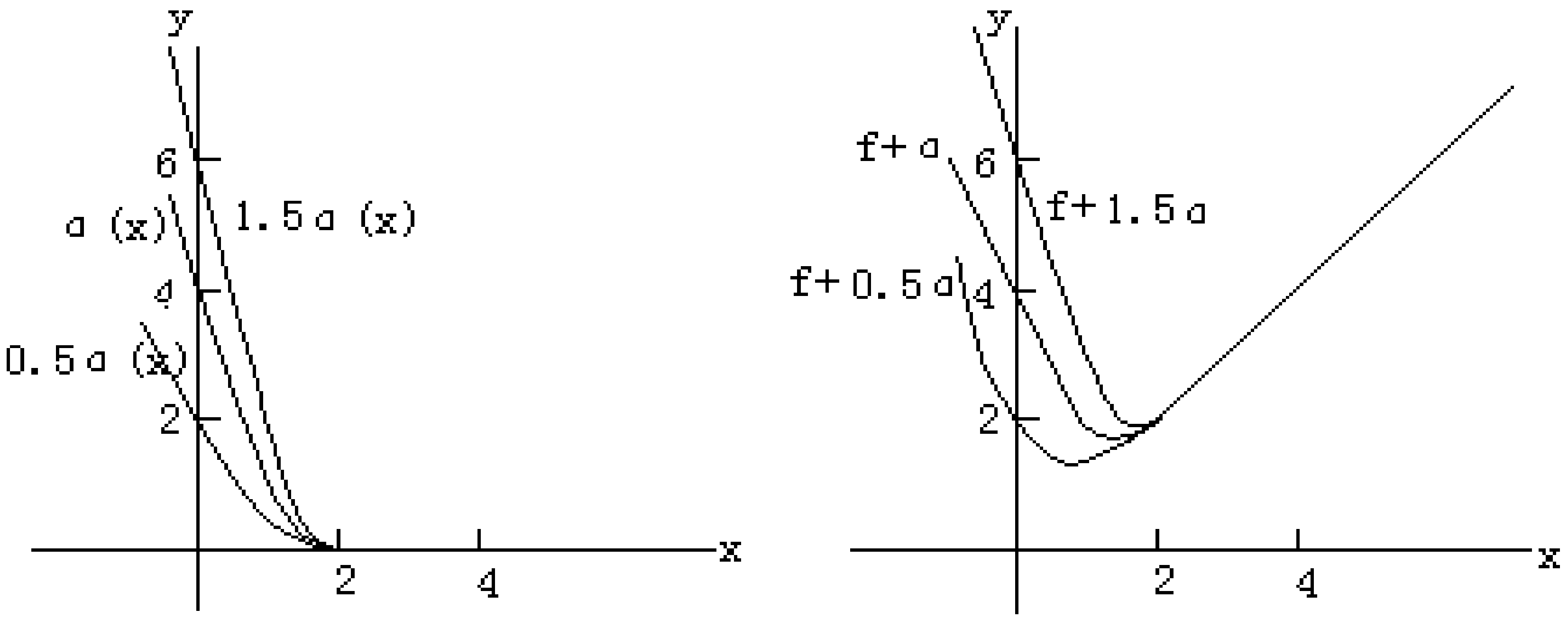


图 6 － 10

惩罚函数法

2. 惩罚函数法

定义：$\theta(\mu)=\inf\{f(x)+\mu\alpha(x)\}$　　下确界

记问题$(fghD)$的可行集为$S=\{x|x\in D,g(x)\le 0,h(x)=0\}$

惩罚函数法产生的辅助问题为：$\begin{cases}\max & \theta(\mu)\\ s.t. & \mu\ge 0\end{cases}$

原问题与辅助问题关系相当于一种对偶，具有弱对偶性质并且各相应函数有单调性

引理：设f,g,h连续，$\alpha(x)$为惩罚函数，连续，再设$\forall\mu\ge 0,\exists x_\mu\in D$，使$\theta(\mu)=f(x_\mu)+\mu\alpha(x_\mu)$，则：

1. $\inf\{f(x)|x\in S\}\ge\sup\{\theta(\mu)|\mu\ge 0\}$
2. $f(x_\mu),\theta(\mu)$是关于$\mu\ge 0$的单调非降函数
3. $\alpha(x_\mu)$是关于$\mu\ge 0$的单调非增函数

惩罚函数法

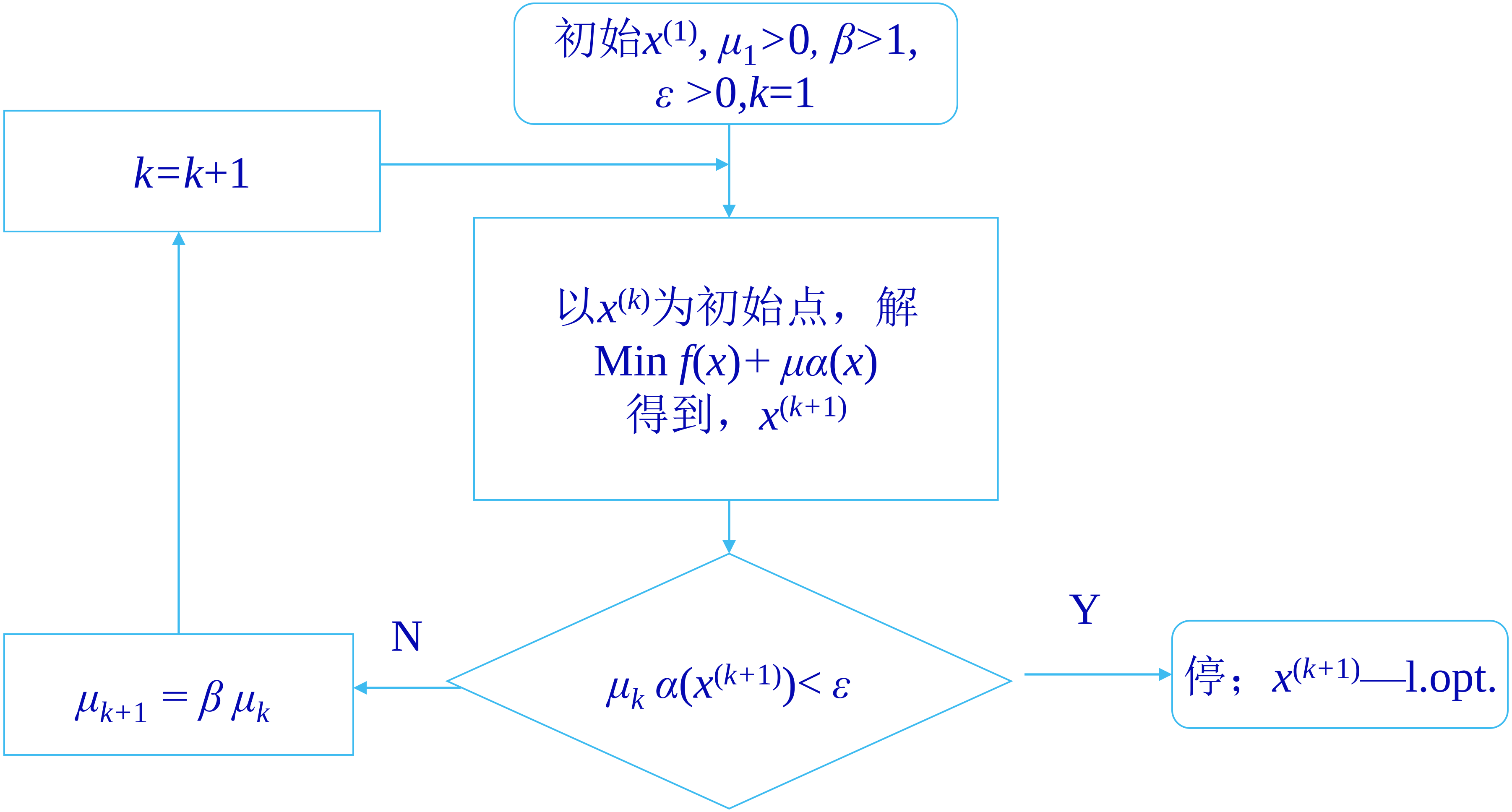
定理: $(fghD), S=\{\mathrm{x}|\mathrm{g}(\mathrm{x})\leq 0, h(x)=0\}\neq\emptyset$, 在引理假设下，设∃紧子集$D_0$，使$\{\mathrm{x}_\mu|\mu\geq 0\}\subset D_0$, 则

1. $\inf\{f(x)|x\in S\}=\sup\{\theta(\mu)|\mu\geq 0\}=\lim_{\mu\to\infty}\theta(\mu)$
2. $\{x_\mu\}$的任意极限点x^*是（$fghD$）的最优解
3. $x^*{\sim}opt$ 且$\mu\alpha(x_\mu)\xrightarrow{\mu\to\infty}0$

推论：在定理条件下，若∃$\mu\geq 0$使$\alpha\left(x_\mu\right)=0$, 则$x_\mu-$
$-\,opt$.

惩罚函数法

算法



初始$x^{(1)}$, μ_1>0, β>1,
ε >0,k=1
k=k+1
以$x^{(k)}$为初始点，解
Min $f(x)$+ $\mu\alpha(x)$
得到，$x^{(k+1)}$
$\mu_k\ \alpha(x^{(k+1)})< \varepsilon$
N
Y
$\mu_{k+1} = \beta\ \mu_k$
停； $x^{(k+1)}$—l.opt.

障碍函数法（闸函数法或内点法）

障碍函数法

$$（fgD）\begin{cases} \min f(x) & f: \mathbb{R}^n \to \mathbb{R} \\ s.t.\ g(x) \le 0 & g: \mathbb{R}^n \to \mathbb{R}^m \\ x \in D & D \subset \mathbb{R}^n \end{cases}$$

记可行集$S = \{x|g(x) \le 0\}, S_0 = \{x|g(x) < 0\} \neq \emptyset$

基本思想：

从S_0中的一个点（内点）出发，在目标函数中加入惩罚项，使迭代保持在S_0内

构造障碍函数(Barrier Function):$B(x) = \sum_{i=1}^{m} \phi(g_i(x))$，使

$$B(X)\begin{cases} 0 \ge, & x \in S_0 \\ \to \infty, & x \to \partial S_0(\text{边界}) \end{cases}$$

为方便应有$B(x)$连续。

障碍函数法

典型取法：

$\phi(t) = -\frac{1}{t}$或$\phi(t) = |\ln(-t)|$

惩罚项：

$$\mu B(x) = \begin{cases} \to 0 & x \in S_0 \\ +\infty & x \in \partial S_0 \end{cases}$$

由于当$x \in S_0$时，$B(x) > 0$且$B(x) \xrightarrow{x \to \partial S_0} \infty$

故需要随着$x \to \partial S_0, \mu \to 0^+$

辅助问题：

$$\begin{cases} \min & f(x) + \mu B(x) \\ s.t. & x \in D \end{cases}$$

障碍函数法

举例：

$$\begin{cases}\min & x \\ s.t. & -x+2\le 0\end{cases}$$

障碍函数$B(x)=\frac{1}{x-2}, x>2$

求解$\begin{cases}\min & g(x,\mu)=x+\frac{\mu}{x-2} \\ s.t. & x>2\end{cases}$

目标函数关于x是凸的，求驻点：$x_\mu = 2+\sqrt{\mu}$

$$x_\mu = 2+\sqrt{\mu} \xrightarrow{\mu\to 0^+} 2 = x^*$$

$$g(x_\mu,\mu) \xrightarrow{\mu\to 0^+} 2 - -opt$$

障碍函数法

定义：$\theta(\mu)=\inf\{f(x)+\mu B(x)|x\in S_0\}$

有类似于罚函数法的理论结果：

定理：$(fg), f, g$连续，$S_0\neq\emptyset$, 最优解$x^*\in S_0$

则

1. $\min\{f(x)|x\in S\}=\sup\{\theta(\mu)|\mu>0\}=\lim_{\mu\to 0^+}\theta(\mu)$

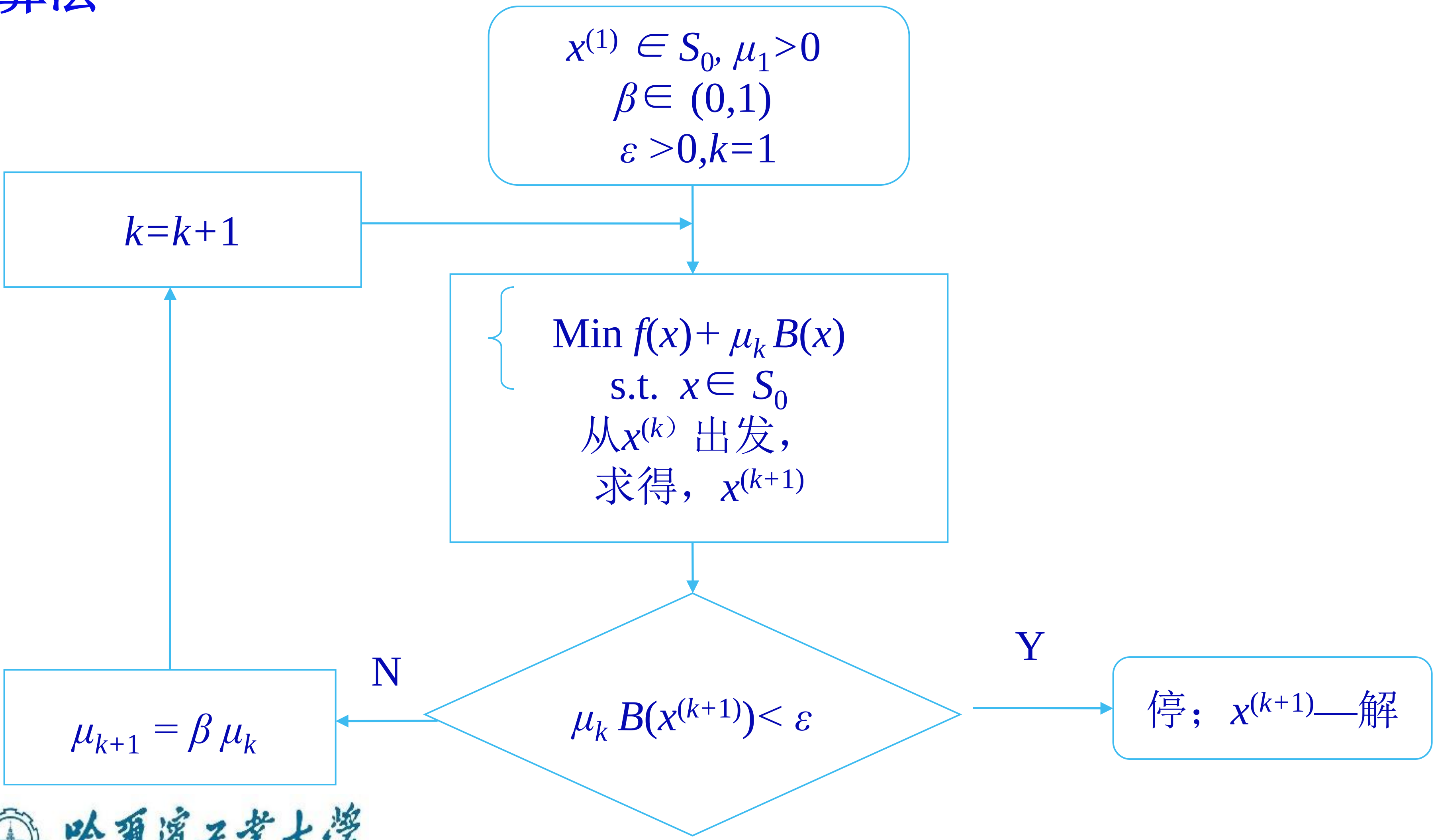
2. 若$\forall\mu>0, \exists x_\mu\in S_0$, 使$\theta(\mu)=f\left(x_\mu\right)+\mu B(x_\mu)$

那么，$\{x_\mu\}$的极限点是(fg)的opt.且

$\lim_{\mu\to 0^+}\mu B\left(x_\mu\right)=0$

障碍函数法

算法


$x^{(1)} \in S_0, \mu_1 > 0$
$\beta \in (0,1)$
$\varepsilon > 0, k=1$
$k=k+1$
$\begin{cases} \text{Min } f(x) + \mu_k B(x) \\ \text{s.t. } x \in S_0 \end{cases}$
从$x^{(k)}$出发，
求得，$x^{(k+1)}$
$\mu_k B(x^{(k+1)}) < \varepsilon$
N
Y
$\mu_{k+1} = \beta \mu_k$
停；$x^{(k+1)}$——解

障碍函数法

求初始内点：

1. $\forall x^{(1)} \in D, k = 1$, 转步骤2；

2. 令$J_k = \{j | g_j(x^{(k)}) < 0\}$, 若$J_k = \{1,2,\cdots,m\}$, 停止，并取$x^{(k)}$为初始内点，否则转步骤3；

3. 用障碍函数法求解

$$\begin{cases} \min g_i(x) \\ s.t.\, g_j(x) < 0, j \in J_k \\ x \in D \end{cases}$$

其中，i使$g_i(x^{(k)}) = \max\{g_l(x^{(k)}) | l \notin J_k\}$。以$x^{(k)}$为初始点，得到$x^{(k+1)}$。若$g_i(x^{(k+1)}) \geq 0$, 则停止，说明$S_0 = \emptyset$; 否则$k = k + 1$，转步骤2；

惩罚函数法和障碍函数法的优缺点

惩罚函数法和障碍函数法的缺点：

1. 当惩罚函数法(障碍函数法)的$\mu \to \infty(\mu \to 0^+)$时，惩罚项趋于$+\infty \circ 0(0 \circ +\infty)$形式，导致计算困难；

2. 计算一系列无约束问题，计算量大

3. 惩罚函数法可用于只含等式约束、只含不等式约束以及既含有等式约束又含有不等式约束的问题；障碍函数法只用于只含不等式约束的问题

4. 惩罚函数法中间步骤的点不是近似解，障碍函数法中间步骤的点都是可行解，可作为近似解

乘子法

$$(fhD)\begin{cases}\min & f(x) \quad f:\mathbb{R}^n\to\mathbb{R}\\ s.t. & h(x)=0 \quad h:\mathbb{R}^n\to\mathbb{R}^l\\ & x\in D \quad D\subset\mathbb{R}^n\text{是一个集合}\end{cases}$$

用增广$Lagrange$函数代替$f(x)$:

乘子罚函数：

$$\phi(x,v,\mu)=f(x)+\sum_{i=1}^{l} v_i h_i(x)+\sum_{i=1}^{l}\mu_i h_i^2(x)$$

其中：$v\in\mathbb{R}^l$为乘子，$\mu\in\mathbb{R}^l$为罚因子。

求解$\begin{cases}\min & \phi(x,v^{(k)},\mu^{(k)})\\ s.t. & x\in D\end{cases}$得到$x^{(k+1)},k=0,1,\cdots$

若$h\left(x^{(k+1)}\right)=0$得到解$x^{(k+1)}$及乘子$v^{(k)}$;否则，调整$v^{(k)}$及$\mu^{(k)}$

参考文献

- 运筹学与最优化方法 第6章 P160-171