4.4 序列无约束优化方法 — 概述

§4.4　序列无约束方法

考虑一般约束优化问题(4.1.1), 即:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \geq 0, \ \ \forall i \in I = \{1, 2, \ldots, p\}, \\ & c_i(x) = 0, \ \ \forall i \in E = \{p+1, \ldots, m\}, \end{aligned} \tag{4.4.1}$$

其中$x \in \mathbb{R}^n$, 函数$f, c_i \ \ (i \in \{1, \ldots, m\})$具有连续可微性. 对于此约束优化问题，一类重要的求解方法是用一系列的无约束优化问题来逼近原约束优化问题，通过一系列无约束优化问题的最优解，逼近原约束优化问题的最优解. 直到迭代点列收敛到原约束优化问题的最优解为止. 这种求解约束优化问题的方法称为**序列无约束极小化技术** (简称SUMT法). 函数$m(x, \mu)$通常称为“罚”函数，对应的方法称为罚函数法. 本节先介绍三种类型的罚函数法：外罚函数法，内罚函数法以及混合罚函数法. 然后，介绍利用Lagrange函数对罚函数法进行改进，得到另外一种方法即为乘子法.

4.4 序列无约束优化方法 — 概述

外罚函数法的不足之处主要体现在以下两点：

(I) 对于给定罚因子σ的值，其对应无约束优化问题的最优解$x(\sigma)$往往不满足可行性条件, 即说明得到原约束优化问题的近似最优解$x(\sigma)$往往是不可行解，这对某些实际问题是不能接受的.

(II) 根据收敛性定理可知，σ_k值取得越大，得到的解$x(\sigma_k)$越接近于原约束优化问题的最优解x^*, 但是，σ_k值越大，造成惩罚函数$P(x, \sigma_k)$的Hesse阵越趋向于病态矩阵, 增大求解困难程度.

§4.4.2 内罚函数法

外罚函数法在每步迭代所得到的迭代点均在原约束优化问题可行域$\mathscr{F}$的外部，但在某些实际问题中是不想得到这样的解. 为了使保证每次的迭代点总在可行域的内部，于是，本小节介绍另外一种方法, 即为内罚函数法. 内罚函数法是通过在约束区域的边界上设置一个障碍，确保将所有的迭代点都限制在可行域的内部，故内罚函数法又称为障碍函数法. 其主要思想是: 当迭代点在可行域的内部时，惩罚函数中的惩罚项取得非常小的值，不起到惩罚作用；而当迭代点靠近可行域的边界时，惩罚函数值徒然增大，以示惩罚，以至于当迭代点充分靠近可行域的边界时，惩罚函数值趋于无穷大，这样可以将问题的最优解“挡”在可行域的内部.

4.4 序列无约束优化方法 — 内罚函数法

由于内罚函数法使所有的迭代点始终在可行域内部，因而，这种方法适用于只含有不等式约束的优化问题：

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \geq 0, \ \ \forall i \in I = \{1, 2, \ldots, m\}, \end{aligned} \tag{4.4.8}$$

其中，$f, c_i \ (i \in \{1, \ldots, m\})$为连续函数, 并记该问题的可行域为:

$$\mathscr{F} := \{x \in \mathbb{R}^n \mid c_i(x) \geq 0, \ \ \forall i \in I\},$$

其内部为:

$$int\mathscr{F} = \{x \in \mathbb{R}^n \mid c_i(x) > 0, \ \ \forall i \in \{1, \ldots, m\}\}.$$

4.4 序列无约束优化方法 — 内罚函数法

对于不等式约束优化问题(4.4.8), 假设$int\mathscr{F} \neq \emptyset$. 构造障碍函数

$$B(x, r) = f(x) + rb(x),$$

其中参数r为充分小的正数, 也称之为罚因子，$rb(x)$仍称为惩罚项. 为了使迭代点保持在可行域的内部$int\mathscr{F}$中，函数$b(x)$的选取往往具有下列性质:

- 函数$b(\cdot)$关于变量x是连续函数;
- 对于$x \in int\mathscr{F}$，有$b(x) \geq 0$;
- x趋近于可行域$\mathscr{F}$的边界时，$b(x) \to \infty$.

4.4 序列无约束优化方法 — 内罚函数法

满足上述条件常用的函数$b(x)$主要有以下两种:

$$b(x) = \sum_{i=1}^{m} \frac{1}{c_i(x)} \quad 或 \quad b(x) = -\sum_{i=1}^{m} ln\ c_i(x),$$

则对应的障碍函数分别称为倒数型障碍函数

$$B(x,r) = f(x) + r\sum_{i=1}^{m} \frac{1}{c_i(x)},$$

和对数型障碍函数

$$B(x,r) = f(x) - r\sum_{i=1}^{m} ln\ c_i(x).$$

4.4 序列无约束优化方法 — 内罚函数法

由障碍函数$B(x,r)$的定义可知，r值越小，对应优化问题

$$\begin{aligned} \min \quad & B(x,r) = f(x) + rb(x), \\ \text{s.t.} \quad & x \in int\mathscr{F} \end{aligned} \tag{4.4.9}$$

的最优解越接近于原约束优化问题(4.4.8)的最优解. 但类似于外罚函数法的情况，若r值越小，则对优化问题(4.4.9)的求解带来的困难越大. 此处仍采用序列无约束极小化方法(SUMT)取一个严格单调递减趋于0的罚因子数列$\{r_k\}$. 对每个罚因子r_k，从可行域$\mathscr{F}$的内部$int\mathscr{F}$出发，求解优化问题

$$\begin{aligned} \min \quad & B(x,r_k) = f(x) + r_k b(x) \\ \text{s.t.} \quad & x \in int\mathscr{F}. \end{aligned}$$

当$r_k \to 0$时，上述优化问题的最优解会一直逼近于原约束优化问题(4.4.8)的最优解x^*. 由此，内罚函数法又称为**SUMT内点法(或序列无约束极小化技术内点法)**.

4.4 序列无约束优化方法 — 内罚函数法

注. (i) 从形式上看，问题(4.4.9)仍然是一个约束优化问题，但是在可行域$\mathscr{F}$的边界上，惩罚函数$B(x,r)$值趋于无穷大，因而，只要从可行域$\mathscr{F}$的一个内点开始迭代，并适当控制一维搜索的步长，就可以使迭代点不会越出可行域，也就是说，从计算的观点来看，问题(4.4.9)可当做为无约束优化问题

$$\min B(x,r) = f(x) + rb(x)$$

来处理.

(ii) 考虑到原约束优化问题(4.4.8)的最优值有可能在可行域$\mathscr{F}$的边界上取得，所以在极小化惩罚函数的过程中，为了使惩罚函数$B(x,r)$的极小值点能更好地接近可行域$\mathscr{F}$ 的边界，需要不断地减小罚因子r的值，来减小惩罚项$rb(x)$的影响.

4.4 序列无约束优化方法 — 内罚函数法

例 4.4.3 用内罚函数法求解约束优化问题

$$
\begin{aligned}
\min \quad & f(x) = \frac{1}{12}(x_1 + 1)^3 + x_2 \\
\text{s.t.} \quad & x_1 - 1 \geq 0, \\
& x_2 \geq 0.
\end{aligned}
$$

解 首先引入障碍函数

$$
B(x, r_k) = \frac{1}{12}(x_1 + 1)^3 + x_2 + r_k\left(\frac{1}{x_1 - 1} + \frac{1}{x_2}\right).
$$

采用解析方法求解下面的无约束优化问题:

$$
\min \quad B(x, r_k). \tag{4.4.10}
$$

4.4 序列无约束优化方法 — 内罚函数法

根据惩罚函数$B(x, r_k)$的表达式，则

$$\begin{cases} \dfrac{\partial B}{\partial x_1} = \dfrac{1}{4}(x_1+1)^2 - \dfrac{r_k}{(x_1-1)^2}, \\ \dfrac{\partial B}{\partial x_2} = 1 - \dfrac{r_k}{x_2^2}. \end{cases}$$

由最优解存在的一阶必要条件，令

$$\frac{\partial B}{\partial x_1} = 0 \quad 且 \quad \frac{\partial B}{\partial x_2} = 0,$$

解之得$x^k = \left(\sqrt{1 + 2\sqrt{r_k}}, \sqrt{r_k} \right)^\top$.

当$r_k \to 0$时，得$x^* = \lim_{k\to\infty} x^k = (1,0)^\top$, 则点$x^*$是原约束优化问题的最优解. 从$r_k \to 0$的过程来看，无约束优化问题(4.4.10)最优解的点列$\{x^k\}$从原约束优化问题可行域的内部

$$int\mathscr{F} = \{x = (x_1, x_2)^\top \mid x_1 - 1 > 0, x_2 > 0\}$$

向可行域$\mathscr{F}$边界上的最优解$x^* = (1,0)^\top$无限逼近. □

4.4 序列无约束优化方法 — 内罚函数法

算法 4.4.2 *(内罚函数法)* 选取初始可行内点$x^0 \in int(\mathscr{F})$, 给定允许误差$\varepsilon > 0$, 初始参数罚因子$r_1 > 0$, 以及缩小系数$\beta \in (0, 1)$, 置$k = 1$.

步1 以x^{k-1}为初始点, 求解无约束优化问题

$$\min \quad B(x, r_k) = f(x) + r_k b(x) \tag{4.4.11}$$

得最优解为$x^k := x(r_k)$.

步2 若$r_k b(x^k) < \varepsilon$, 则算法终止, 点x^k即为原约束最优化问题*(4.4.8)*的近似最优解; 否则, 置$r_{k+1} = \beta r_k$ 以及$k := k + 1$, 转步*1*.

4.4 序列无约束优化方法 — 内罚函数法

定理 4.4.3 对于不等式约束优化问题*(4.4.8)*，设$\mathscr{F}$内部非空且存在最优解, 罚因子数列$\{r_k\}$是严格单调递减趋于0的正数列，且对每个r_k, 对应的优化问题*(4.4.11)*存在最优解x^k，则由算法*4.4.2*产生迭代点列$\{x^k\}$的任意聚点$\bar{x}$都是原问题*(4.4.8)*的最优解.

证明 先证罚函数数列$\{B(x^k, r_k)\}$ 单调递减且有下界. 设x^*为原不等式约束优化(4.4.8) 的最优解, $x^k, x^{k+1} \in \mathscr{F}$分别是罚因子$r_k$和$r_{k+1}$对应无约束优化问题(4.4.11) 的最优解且$r_{k+1} < r_k$, 于是

$$
\begin{aligned}
B(x^{k+1}, r_{k+1}) &\leq B(x^k, r_{k+1}) = f(x^k) + r_{k+1}b(x^k) \\
&\leq f(x^k) + r_k b(x^k) = B(x^k, r_k),
\end{aligned}
$$

且

$$B(x^k, r_k) = f(x^k) + r_k b(x^k) \geq f(x^k) \geq f(x^*). \tag{4.4.12}$$

这表明数列$\{B(x^k, r_k)\}$是单调递减且有下界的数列. 因此，数列$\{B(x^k, r_k)\}$存在极限, 记为$\bar{B}$. 另外, 由(4.4.12)式可知: $\bar{B} \geq f(x^*)$.

4.4 序列无约束优化方法 — 内罚函数法

接下来，假如我们能证得$\bar{B} = f(x^*)$, 则由(4.4.12)式可知

$$f(x^*) = \bar{B} = \lim_{k\to\infty} B(x^k, r_k) \geq \lim_{k\to\infty} f(x^k) \geq f(x^*).$$

结合函数f的连续性以及$\bar{x}$为点列$\{x^k\}$的聚点，不失一般性，不妨设$\lim_{k\to\infty} x^k = \bar{x}$, 于是有

$$f(\bar{x}) = \lim_{k\to\infty} f(x^k) = f(x^*),$$

即$\bar{x}$为原约束优化问题(4.4.8)的最优解.

最后，用反证法证$\bar{B} = f(x^*)$. 设$\bar{B} > f(x^*)$, 则存在$\varepsilon > 0$使得

$$\bar{B} > f(x^*) + 2\varepsilon.$$

4.4 序列无约束优化方法 — 内罚函数法

由函数f的连续性可知，必存在可行点$\hat{x} \in \mathscr{F}$使得

$$f(\hat{x}) < f(x^*) + \varepsilon.$$

进而

$$B(x^k, r_k) \quad \leq \quad B(\hat{x}, r_k) = f(\hat{x}) + r_k b(\hat{x}) < f(x^*) + \varepsilon + r_k b(\hat{x}).$$

由于当$k \to \infty$时，$r_k \to 0$，因而$r_k b(\hat{x}) < \varepsilon$，所以

$$\lim_{k\to\infty} B(x^k, r_k) \leq f(x^*) + 2\varepsilon,$$

即$\bar{B} \leq f(x^*) + 2\varepsilon$. 此式与假设$\bar{B} > f(x^*)$相矛盾，所以$\bar{B} = f(x^*)$. 定理得证. $\square$

4.4 序列无约束优化方法 — 内罚函数法

本小节曾指出，内罚函数法仅限于求解不等式约束优化问题，且初始点的选取必须在可行域的内部. 而外罚函数法可适用于求解一般约束优化问题, 初始点可以是可行域外部的点，但是得到的近似最优解往往是不可行的. 鉴于上述情况，对于求解一般约束优化问题，可以把外罚函数法和内罚函数法联合起来，形成所谓的**混合罚函数法**，其基本思想为：对于得到的近似最优解$x^{k-1} \in \mathbb{R}^n(k \in \{1,2,\ldots\})$, 在点$x^{k-1}$处, 对于满足严格不等式成立的那些不等式约束，采用内罚函数法来构造惩罚函数$B(x,r)$; 对剩余的不等式约束和等式约束，我们采用外罚函数法来构造惩罚函数$P(x,\sigma)$, 其中罚因子σ 和r也可以用同一个参数表示，不妨用r_k表示，即混合罚函数为：

$$F(x,r_k) = f(x) + \frac{1}{r_k}\bar{p}(x) + r_k b(x),$$

其中

$$\bar{p}(x) = \sum_{i\in I_1}(\min\{0,c_i(x)\})^2 + \sum_{i=1}^{p} c_i^2(x) \quad 且 \quad b(x) = -\sum_{i\in I_2} ln\ c_i(x),$$

$$I_1 = \{i \mid c_i(x^{k-1}) \le 0,\ \ i \in I\},\ \ I_2 = \{i \mid c_i(x^{k-1}) > 0,\ \ i \in I\},\ \ I = I_1 \cup I_2.$$

$\{r_k\}$单调递减且$\lim_{k\to\infty} r_k = 0$，其迭代步骤与内罚函数法相同.

4.4 序列无约束优化方法 — 内罚函数法

上面介绍的两种方法，外罚函数法与内罚函数法均采用序列无约束极小化技术，它们的共同优点是：

- 方法简单，易编制程序，使用方便.
- 对目标函数和约束函数的要求不高，能够求解导数不存在的问题，适用范围广.

因此，这类算法受到工程技术人员的欢迎，但是上述罚函数法也存在一些固有的缺陷：

- 为求解一个约束优化问题，需要求解一系列的无约束优化问题，工作量必然增大.
- 在迭代过程中，当$\sigma \to \infty$(或$r \to 0$)时，惩罚函数Hesse 阵的条件数会速然增大，Hesse 阵的性态也就变得越来越病态，这给无约束优化问题的求解带来很大的困难.

为了克服上述的缺点，Hestenes 和Powell于1969年分别提出了乘子法.

4.4 序列无约束优化方法 — 内罚函数法作业

4.19 用内点法求解下列约束优化问题：

(1) $\max \quad f(x) = x_1 x_2$

$\text{s.t.} \quad c(x) = -2x_1 + x_2 + 3 \geq 0.$

(2) $\min \quad f(x) = x_1^2 + x_2^2$

$\text{s.t.} \quad c_1(x) = 2x_1 + x_2 - 2 \leq 0,$

$c_2(x) = -x_1 + 1 \leq 0.$