PART 02

内点法

惩罚函数法基本思想

根据约束函数的特点，构造“惩罚项”加到目标函数中去，构成“惩罚函数”，将约束优化问题化为一系列的无约束优化问题来求解。

根据惩罚项的不同形式，惩罚函数法可以分为：
- 内罚函数法
- 外罚函数法
- 混合罚函数法

内点法——迭代点在可行域之内

内点法基本思想：通过在目标函数上引入一个关于约束的惩罚项，当迭代点有可行域的内部接近可行域的边界时，罚项的值将趋近于无穷大，以迫使迭代点回到可行域的内部，从而保持迭代点的严格可行性。

内点法基本概念

基本原理：

内点法的特点是将构造的新的无约束目标函数-惩罚函数定义在可行域内，并在可行域内求惩罚函数的极值点，即求解无约束问题时的探索点总是在可行域内部，这样，在求解内点惩罚函数的序列无约束优化问题的过程中，所求得的系列无约束优化问题的解总是可行解，从而在可行域内部逐步逼近原约束优化问题的最优解。

内点法是求解不等式约束最优化问题的一种十分有效方法，但不能处理等式约束。因为构造的内点惩罚函数是定义在可行域内的函数，而等式约束优化问题不存在可行域空间，因此，内点法不能用来求解等式约束优化问题。

对于目标函数为：

$$\begin{cases} \min f(X) \\ s.t. \quad g_i(X) \geq 0 \quad i = 1, \cdots, m \end{cases}$$

的最优化问题，利用内点法进行求解时，构造惩罚函数的一般表达式为：

$$\varphi(X, r^{(k)}) = f(X) - r^{(k)} \sum_{i=1}^{m} \frac{1}{g_i(X)}$$

或者 $\varphi(X, r^{(k)}) = f(X) + r^{(k)} \left| \sum_{i}^{m} ln|g_i(X)| \right| = f(X) - r^{(k)} \sum_{i=1}^{m} ln[-g_i(X)]$

内点法基本概念

而对于$f(X)$受约束于$g_i(X) \geq 0 \ \ (i=1,\cdots,m)$的最优化问题，其惩罚函数的一般形式为：

$$\varphi\left(X,\ r^{(k)}\right) = f(X) \boxed{- r^{(k)} \sum_{i=1}^{m} \frac{1}{g_i(X)}}$$

或

$$\varphi\left(X,\ r^{(k)}\right) = f(X) \boxed{- r^{(k)} \sum_{i=1}^{m} ln[g_i(X)]}$$

惩罚项（障碍项）

式中，$r^{(k)}$——惩罚因子，是递减的正数序列，即

$r^{(0)} > r^{(1)} > r^{(2)} > \ldots > r^{(k)} > r^{(k+1)} > \ldots > 0 \qquad \lim_{k\to\infty} r^{(k)} = 0$

通常取$r^{(k)}$=1.0，0.1，0.01，0.001，…。

内点法基本概念

注：

当迭代点在可行域内部时，有$g_i(X) \geq 0 \ (i = 1, \cdots, m)$，而$r^{(k)} > 0$，则惩罚项恒为正值，当设计点由可行域内部向约束边界移动时，惩罚项的值要急剧增大并趋向无穷大，于是惩罚函数的值也急剧增大直至无穷大，起到惩罚作用，使其在迭代过程中始终不会触及约束边界。

懵了是吧？看个例题，看个例题......

例：
$$\min f(X)=x$$
$$s.t.\ \ g(X)=1-x\leq 0$$

这是一个一维优化问题，很显然，$x^* = 1 \quad f(x^*) = 1$，那利用内点法怎么算？

内点法基本概念

建立一个内罚函数：

$$\varphi\left(X,\ r^{(k)}\right)=x-r^{(k)}\frac{1}{1-x}$$

分析：

x在可行域内，即$g(X)=1-x<0$，惩罚项>0；当x趋向于约束边界时，$g(X)=1-x\to 0$，惩罚项$\to\infty$. 因为求的是最小值，所以x不可能越出来约束边界，保证迭代过程始终在可行域内进行。

当$r^{(k)}\to 0$时，惩罚项$\to 0$，$x^*(r^{(k)})\to x^*$，得到原问题最优解。

设x为内点，按无约束优化问题的极值条件：

$$\frac{\partial\varphi\left(X,\ r^{(k)}\right)}{\partial x}=1-r^{(k)}\frac{1}{(1-x)^2}$$

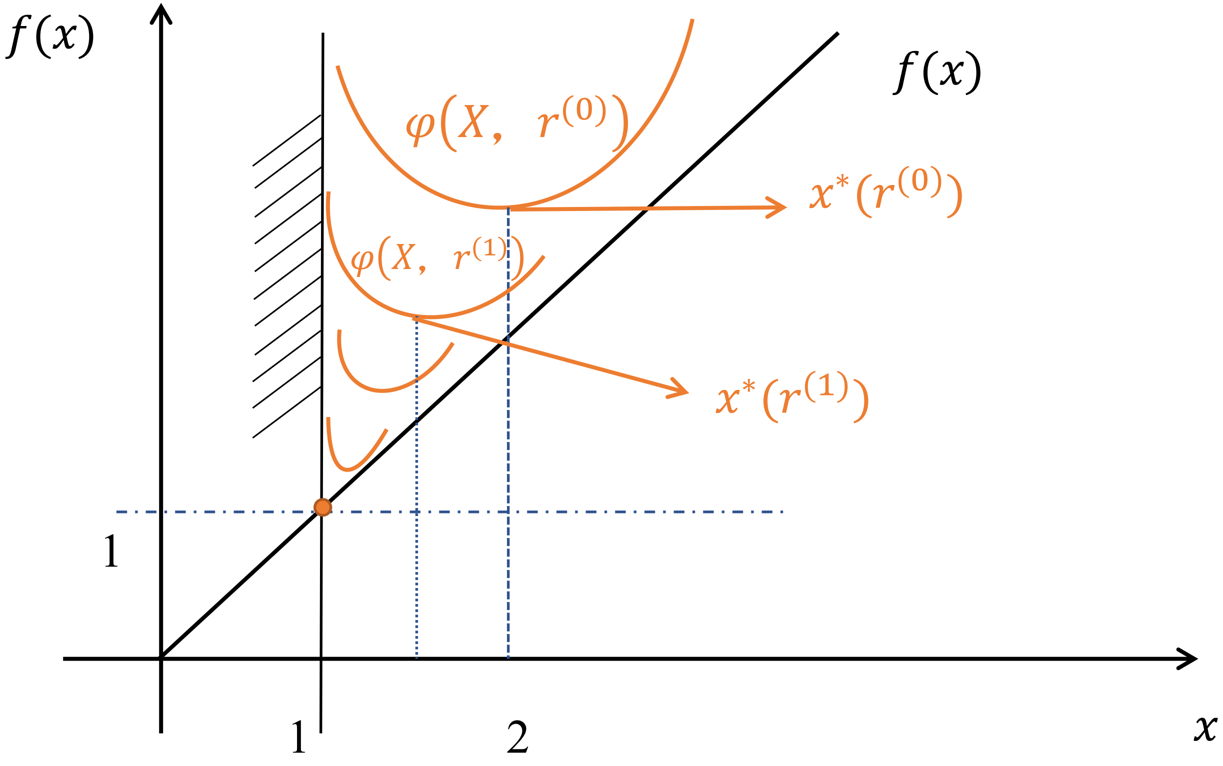
令$1-r^{(k)}\frac{1}{(1-x)^2}=0$，得到$x^*(r^{(k)})=1+\sqrt{r^{(k)}}$，再让

$$r=1,\ 0.1,\ 0.01,\ 0.001\ldots\ldots$$

直至趋向于零，得到以下表格：

内点法基本概念

$r^{(k)}$	1	0.1	0.01	0.001	……	0
$x^*(r^{(k)})$	2	1.316	1.1	1.032	……	1
$\varphi(X,\ r^{(k)})$	3	1.632	1.2	1.036	……	1



利用内点法求解问题的计算步骤

(1)取初始惩罚因子$r^{(0)} > 0$，允许误差$\varepsilon > 0$；

(2)在可行域D内取初始点$X^{(0)}$，令$k = 1$；

(3)构造惩罚函数$\varphi\left(X,\ r^{(k)}\right)$，从$X^{(k-1)}$点出发用无约束优化方法求解惩罚函数$\varphi\left(X,\ r^{(k)}\right)$的极值点$X^*(r^{(k)})$；

(4)检查迭代终止准则：如果满足

$$\left\|X^*(r^{(k)}) - X^*(r^{(k-1)})\right\| \le \varepsilon_1 = 10^{-5} - 10^{-7}$$

或

$$\left\|\frac{\varphi\left(X^*,\ r^{(k)}\right) - \varphi\left(X^*,\ r^{(k-1)}\right)}{\varphi\left(X^*,\ r^{(k-1)}\right)}\right\| \le \varepsilon_2 = 10^{-3} - 10^{-4}$$

则停止迭代计算，并以$X^*(r^{(k)})$为原目标函数$f(X)$的约束最优解，否则转入下一步

(5)取$r^{(k+1)} = Cr^{(k)}$，$X^{(0)} = X^*(r^{(k)})$，$k = k + 1$，转入步骤(3).

递减系数$C = 0.1$~0.5，常取0.1，也可以取0.02.

利用内点法求解问题的计算步骤

第四步中，根据情况，终止准则还可以有如下形式：

$$\left\|f(X^{(k)})\ -f(X^{(k-1)})\ \right\| \le \varepsilon$$

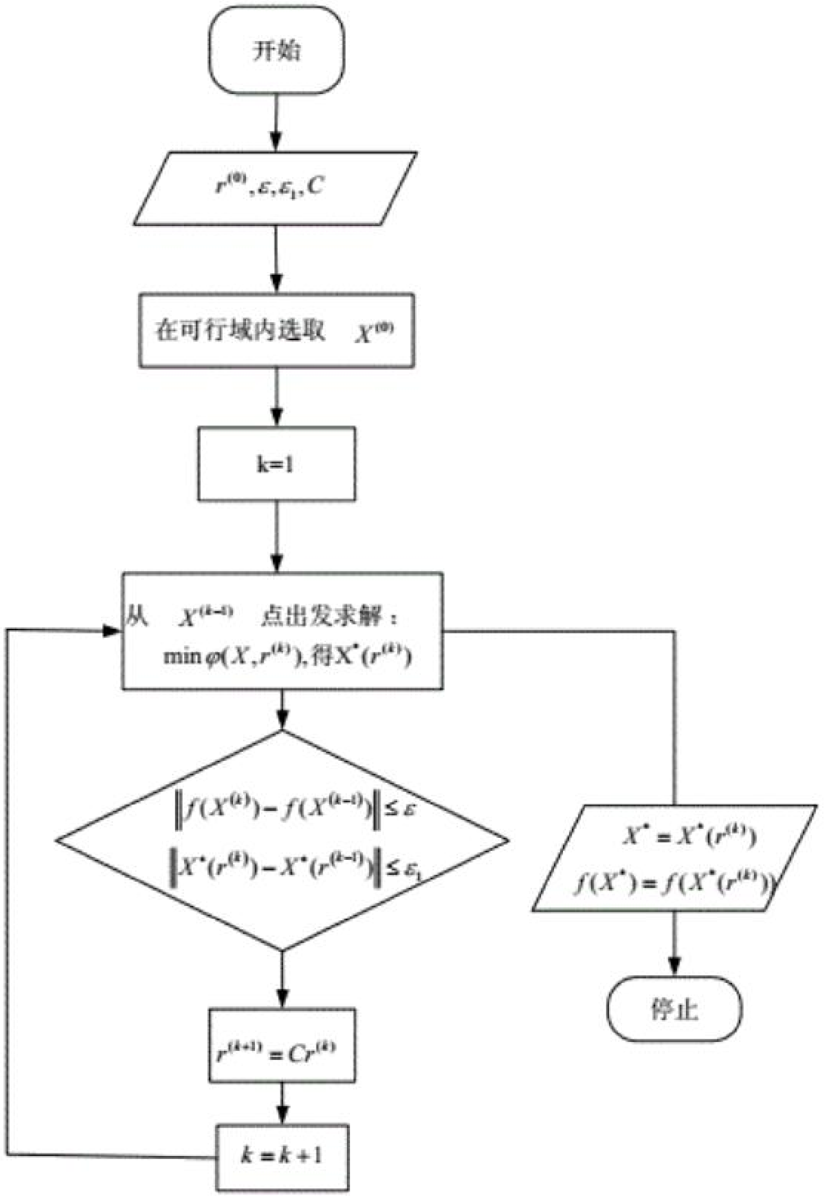
或

$$\left\|r^{(k)}\sum_{i=1}^{m}\frac{1}{g_i(X)}\right\| \le \varepsilon$$

或

$$\left\|r^{(k)}\sum_{i}^{m} ln|g_i(X)|\right\| \le \varepsilon$$

利用内点法求解问题的程序框图



收敛性

内点罚函数收敛性定理：设一个非等式约束优化问题中，可行域内部 $int\ S$ 非空，且存在最优解，又设对每一个 r_k， $G(X, r_k)$ 在 $int\ S$ 内存在极小点，并且在内点罚函数法产生的全局极小点序列 $\{X^{(k)}\}$ 存在子序列收敛到 $\overline{X}$,则 $\overline{X}$ 是问题的全局最优解。

证明：

①首先证明 $\{G(X^{(k)}, r_k)\}$ 是单调递减有下界的序列

设 $X^{(k)}$， $X^{(k+1)} \in int\ S$ 分别是 $G(X, r_k)$， $G(X, r_{k+1})$ 的全局最小点，由于 $r_{k+1} < r_k$，因此有

$$G(X^{(k+1)}, r_{k+1}) = f(X^{(k+1)}) + r_{k+1}B(X^{(k+1)}) \le f(X^{(k)}) + r_{k+1}B(X^{(k+1)})$$
$$\le f(X^{(k)}) + r_kB(X^{(k)}) = G(X^{(k)}, r_k)$$

收敛性

设X^*是原问题的全局最优解，由于$X^{(k)}$是可行点，因此有

$$f(X^{(k)}) \geq f(X^*)$$

又

$$G\left(X^{(k)}, r_k\right) \geq f(X^{(k)})$$

故有

$$G\left(X^{(k)}, r_k\right) \geq f(X^*)$$

综上，$\left\{G\left(X^{(k)}, r_k\right)\right\}$是单调递减有下界的序列，设其极限为$\widehat{G}$。

②证明$\widehat{G} = f(X^*)$

利用反证法，设$\widehat{G} > f(X^*)$,由于$f(X)$是连续函数，取一常数$\frac{1}{2}\left[\widehat{G} - f(X^*)\right]$，存在正数$\delta$，当$\|X - X^*\| \leq \delta$且$X \in int\ S$时，有

$$f(X) - f(X^*) \leq \frac{1}{2}\left[\widehat{G} - f(X^*)\right]$$

即

$$f(X) \le \frac{1}{2}\left[\widehat{G} + f(X^*)\right]$$

在X^*的邻域$\|X - X^*\| \le \delta$内取一点$\overline{X}$， $\overline{X} \in int\, S$， 由于$r_k \to 0$， 故存在K，当$k > \mathrm{K}$时， 有

$$r_k B(\overline{X}) < \frac{1}{4}\left[\widehat{G} - f(X^*)\right]$$

这样， 当$k > \mathrm{K}$时， 有

$$\begin{aligned} G\left(X^{(k)}, r_k\right) &= f\left(X^{(k)}\right) + r_k B\left(X^{(k)}\right) \le f\left(\widehat{X}\right) + r_k B\left(\widehat{X}\right) \\ &\le \frac{1}{2}\left[\widehat{G} + f(X^*)\right] + \frac{1}{4}\left[\widehat{G} - f(X^*)\right] = \widehat{G} - \frac{1}{4}\left[\widehat{G} - f(X^*)\right] \end{aligned}$$

收敛性

上式与$G\left(X^{(k)}, r_k\right) \to \widehat{G}$矛盾，因此必有$\widehat{G} = f(X^*)$

③证明$\overline{X}$，$\overline{X} \in int\ S$为全局最优点：

设$\left\{X^{(k_j)}\right\}$是$\left\{X^{(k)}\right\}$的收敛子序列，且$\lim\limits_{k_j \to \infty} X^{(k_j)} = \overline{X}$，

由于$X^{(k_j)}$是可行域内的点，故有

$$g_i\left(X^{(k_j)}\right) > 0, i = 1, \dots, m$$

又$g_i(X)$为连续函数，有

$$\lim_{k_j \to \infty} g_i\left(X^{(k_j)}\right) = g_i(\overline{X}) \geq 0, i = 1, \dots, m$$

$\overline{X}$为可行点，因此有$f(X^*) \leq f(\overline{X})$，则有：

$$\lim_{k_j \to \infty} \left\{f\left(X^{(k_j)}\right) - f(X^*)\right\} = f(\overline{X}) > 0$$

收敛性

这样，$k_j \to \infty$时，

$$G(X^{(k_j)}, r_{k_j}) - f(X^*) = f(X^{(k_j)}) - f(X^*) + r_{k_j} B(X^{(k_j)})$$
$$\geq f(X^{(k_j)}) - f(X^*)$$

不趋近于0，因此与$\lim\limits_{k \to \infty} G\left(X^{(k)}, r_k\right) = \widehat{G} = f(X^*)$相矛盾，故必有$f(X^*) = f(\overline{X})$，

$\overline{X}$是原问题的全局最优解。

内点法与外点法的比较

1.外点法的初始点可以任意取，内点法的初始点必须取在可行域内。

2.外点法对等式约束也适用，内点法对等式约束不适用，因为此时没有内点存在。

3.外点法只有迭代到最后才能得到可行解；而内点法每一步得到的点都是可行解，这在实际问题中很方便，随时停止迭代，都可以得到运问题的一个近似最优解。

4.外点法和内点法对非凸规划均适用。

5.内点法与外点法收敛速度的快慢，与罚因子的选取有关，而罚因子只能根据经验来选取。

例题

用内点法求解下列问题

$$\begin{cases} \min \dfrac{1}{12}(x_1+1)^3 + x_2 \\ s.t. \qquad x_1 - 1 \geq 0, \\ \qquad\qquad x_2 \geq 0. \end{cases}$$

例题

解：定义罚函数

$$\varphi\left(X, r^{(k)}\right)=\frac{1}{12}(x_1+1)^3+x_2+r^{(k)}\left(\frac{1}{x_1-1}+\frac{1}{x_2}\right)$$

令

$$\frac{\partial\varphi}{\partial x_1}=\frac{1}{4}(x_1+1)^2-\frac{r^{(k)}}{(x_1-1)^2}=0,$$

$$\frac{\partial\varphi}{\partial x_2}=1-\frac{r^{(k)}}{{x_2}^2}=0,$$

解得$\overline{X}_{r^{(k)}}=(x_1,x_2)=\left(\sqrt{1+\sqrt{r^{(k)}}},\sqrt{r^{(k)}}\right).$

当$r^{(k)}$取1，0.1，0.01，0.001，……，0时，可得最优解

例题

根据内点法的求值步骤可得下表：

r_k	1	0.1	0.01	······	$r_k \to 0$
$x_1^{(k)}$	1.4142	1.1473	1.0488	······	1
$x_2^{(k)}$	1	0.3163	0.1	······	0
$\varphi(X，r^{(k)})$	5.5868	1.8206	1.1216	······	0

即当$r_k \to 0$，得到最优解$X = (1,0)^T$， $f(X) = \frac{2}{3}$

练习

用内点法求解下列问题

$$\begin{cases} \min f(X) = {x_1}^2 + {x_2}^2 \\ \quad s.t. \qquad 1 - x_1 \leq 0 \end{cases}$$

练习

解：构造内点罚函数为：

$$\varphi\left(X,\ r^{(k)}\right) = {x_1}^2 + {x_2}^2 - r_k \ln[-(1 - x_1)]$$

利用极值条件求解，令

$$\frac{\partial \varphi}{\partial x_1} = 2x_1 - \frac{r^{(k)}}{(x_1 - 1)} = 0$$

$$\frac{\partial \varphi}{\partial x_2} = 2x_2 = 0$$

联立上式，解得$x_1 = \frac{1 \pm \sqrt{1+2r^{(k)}}}{2}, x_2 = 0$

由于约束条件限制，可得无约束极值点为

$\overline{X}_{r^{(k)}} = (\frac{1+\sqrt{1+2r^{(k)}}}{2}, 0).$

当$r^{(k)}$取1，0.1，0.01，0.001，……，0时，可得最优解

例题

根据内点法的求值步骤可得下表：

r_k	1	0.1	0.01	……	$r_k \to 0$
$x_1^{(k)}$	1.4	1.0477	1.0049	……	1
$x_2^{(k)}$	0	0	0	……	0
$\varphi(X，r^{(k)})$	0.9162	0.3043	0.0532	……	0

即当$r_k \to 0$，得到最优解$X=(1,0)^T$，$f(X)=1$