外点法

第九组

目录

基本概念

1. 定义：外点法是对违反约束的点在目标函数中加入相应的惩罚，可行点不予惩罚，这种方法的迭代点在可行域的外部移动。

2. 基本格式： $F(x,\sigma)=f(x)+\sigma P(x)$

其中 $f(x)$ 是目标函数，$F(x,\sigma)$ 为惩罚函数，$P(x)$ 为惩罚项，σ 为惩罚因子。

惩罚项 $P(x)$ 应满足以下三个条件：

(1) $P(x)$ 连续

(2) $P(x)=0, \forall X \in D$

(3) $P(x)>0, \forall X \notin D$

3. 原理：从非可行解出发逐渐移动到可行区域的方法

4. 基本思想：约束优化问题→无约束优化问题

基本概念

惩罚函数也叫乘子法，求解带约束的非线性规划问题时，常用KKT条件列出满足条件的方程组，解方程组后即可得到最值点，当满足KKT条件的方程组是一个非线性方程组，利用计算机求解很难给出通用算法。本篇介绍的惩罚函数可以将一个带约束非线性问题转化为无约束的非线性规划，而无约束线性规划可以用梯度法等实现求解，利用惩罚函数更方便我们制成计算机算法，在现代计算机算法中，凡涉及到求解最值，都会大量的运用惩罚函数或者借鉴惩罚函数思想。

惩罚函数可以分为外点法和内点法，其中外点法更通用，可解决约束为等式和不等式混合的情形，外点法对初始点也没有要求，可以任意取定义域内任意一点，外点法是真正的将问题转化为无约束问题；而内点法初始点必须为可行区内一点，在约束比较复杂时，这个选择内点法的初始点是有难度的，并且内点法只能解决约束为不等式情形。

分类

1. 仅含有等式约束
2. 仅含有不等式约束
3. 既含有不等式又含有等式的约束问题

第一种情况：仅含有等式约束

$$\min f(x) \quad (1)$$
$$s.t.\ h_j(x)=0\ \ j=1,2,\dots,m$$

构造辅助函数

$$F_1(X,\sigma)=f(X)+\sigma\sum_{i=1}^{m}[h_j(x)]^2 \quad (2)$$

令 $P(X)=\sum_{i=1}^{m}[h_j(x)]^2$ ，

p(x)刻画可行性的破化程度；σ表示惩罚的力度。

σ为非常大的正数，称为罚因子/罚系数/罚参数（可以理解为权重）。

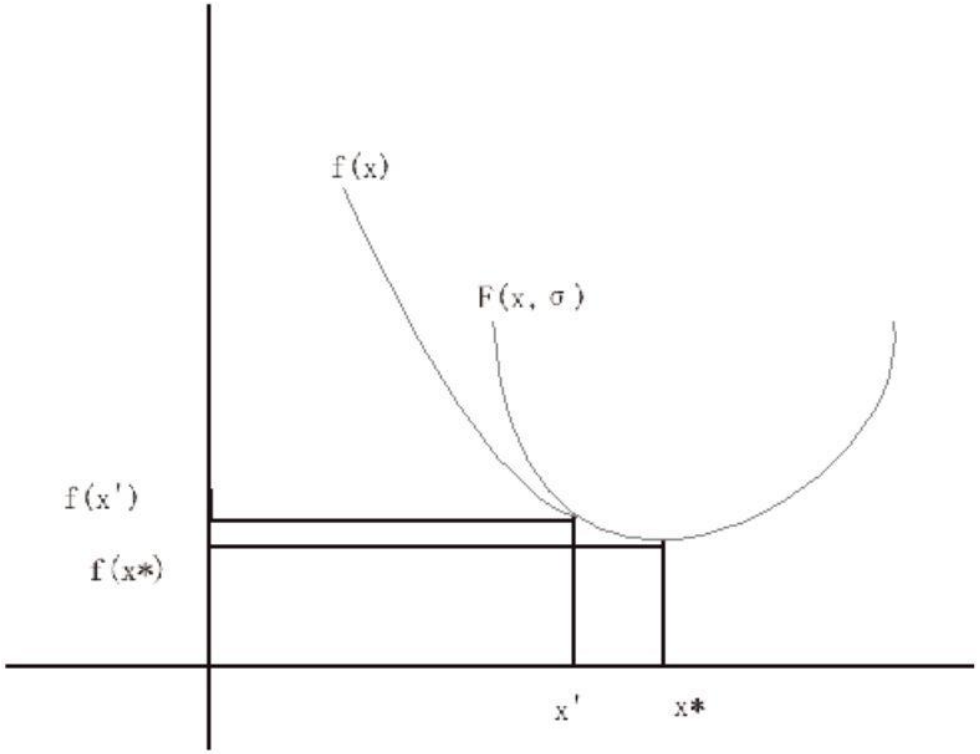
$F_1(X,\sigma)$为罚函数，则原问题转化为了求解无约束的问题

$$\min F_1(X,\sigma) \quad (3)$$

显然，（3）的最优解使得$h_j(x)$接近零，因为如果不这样的话，$\sigma\sum_{i=1}^{m}[h_j(x)]^2$这一项将是很大的正数，现行点必不是极小点。

由此可见，求解问题（3）就能够得到问题（1）的近似解。

辅助函数的构造思路：
（1）在可行点辅助函数等于原来的目标函数值；
（2）在不可行点，辅助函数等于原来的目标函数值加上一个很大的正数。



x’点是原问题最小值点，即x’也是可行点，满足原问题所有等式条件，F(x, σ)与f(x)在该点函数值相等F(x’，σ)=f(x’)，x’点是目标最值点。通过观察可以发现，虽然F(x, σ)是f(x)的上界函数，但最值F(x*)是小于f(x’)的，这是由于F(x*)是F(x, σ)是在无约束情形下的最小值，F(x, σ)参数x可取值范围比原带约束的问题大的多（原问题x必须满足所有等式），极点x*是函数F(x, σ)全局最小值点,外点惩罚函数通过每轮增大参数σ，逐渐让x*逼近x’,通过不断迭代计算，最终的x*便是近似的目标函数极值点x’,这样一来，就实现了利用求解无约束最值得到另一个有约束问题最值。

第二种情况：仅含有不等式约束

$$\min f(X)$$
$$s.t.\ g_i(X) \geq 0, i = 1,2,\ldots,l \qquad (4)$$

构造辅助函数

$$F_2(X,\sigma) = f(X) + \sigma \sum_{i=1}^{l} [\max\{0, -g_i(X)\}]^2 \qquad (5)$$

其中， σ是非常大的正数。 $F_2(X,\sigma)$为罚函数，则原问题转化为了求解下列无约束的问题。

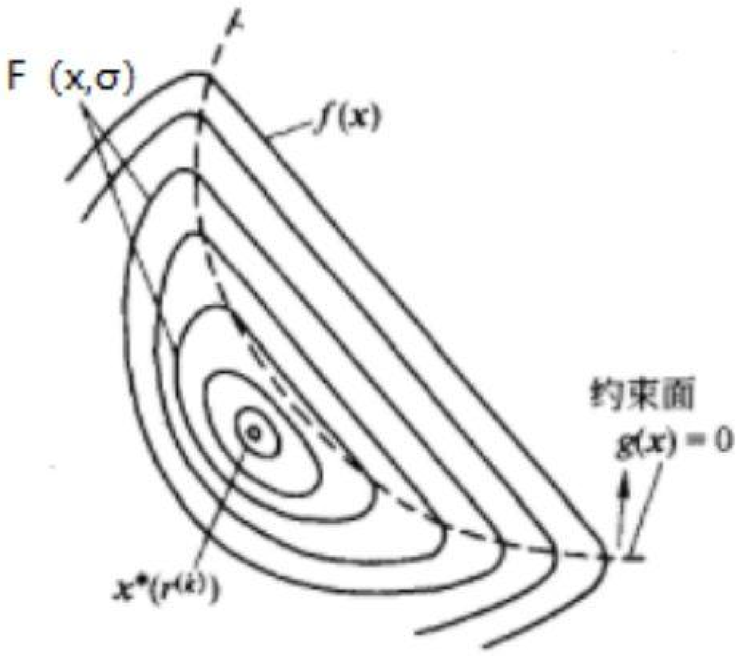
当X为可行点时， $\max\{0, -g_i(X)\} = 0$
当X是不可行点时， $\max\{0, -g_i(X)\} = -g_i(X)$
将问题（4）转化为了无约束问题：

$$\min F_2(X,\sigma) \qquad (6)$$

通过（6）求得（4）的近似解

注意，如果$g_i(X) \leq 0$，则令$F_2(X,\sigma) = f(X) + \sigma \sum_{i=1}^{l} [\min\{0, -g_i(X)\}]^2$

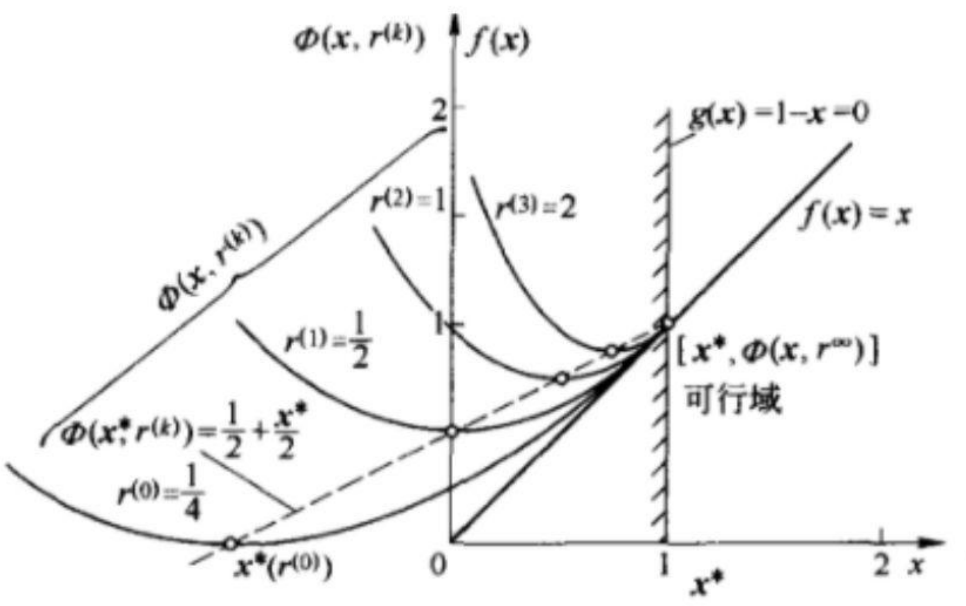
当惩罚函数参数x在可行区内时,惩罚项不等式惩罚项=0，此时F(x, σ)=f(x)；当x在可行区外时F(x, σ)>f(x)，与约束为等式时一样，F(x, σ)的最值F(x*)在初始阶段是小于f(x’)的，可以用等值线来解释这种特性：



g(x)>=0为不等式约束，F(x, σ)函数极值点x*初始时在可行区外，点x*往g(x)=0方向移动时，F(x*, σ)函数值逐渐增大。在可行区内，F(x, σ)与f(x)等值线是一样的，代表两个函数在可行区内函数值相同，约束问题极值点x’一定是在可行区内，即在上图虚线上以及虚线右侧的范围内；在虚线左侧时F(x, σ)极值F(x*)小于f(x’)。惩罚函数解决不等式约束时，是通过调整参数σ，让x*逐渐接近可行区，从而得到约束极值点x’。

通过上面两种情形的分析，可以发现外点法是从可行区外慢慢接近边界，在接近的过程中计算每个阶段极值点，一旦到达可行区范围内，该极值点即为原带约束非线性规划的极值点，接下来还有最后一个问题，为什么增大参数值σ可以让极值点x*接近可行区，以公式（2）为例，做适当变换可得：

$$\sum_{i=1}^{l}[\max\{0,-g_i(X)\}]^2=\frac{F(x,\sigma)-f(x)}{\sigma}$$



σ增大时，惩罚项将逐渐等于0，代表F(x, σ)等值线整体逐渐向可行区移动。这里也可以看出，σ之所以叫惩罚因子是对惩罚函数的极值点不在可行区时，对惩罚函数进行相应的惩罚，因此，调整惩罚因子将移动惩罚函数极值点，这个动态的过程可参考下图，惩罚函数Φ(x, rk)不断增大rk后，Φ(x, rk)的极值点x*不断接近目标函数f(x)极值点。

第三种情况，既含有不等式又含有等式的约束问题

$$\min f(X) \quad X \in R^n$$

$$s.t. \begin{cases} g_i(X) \geq 0, i = 1,2,\dots,l \\ h_j(X) = 0,\ j = 1,2,\dots,m \end{cases}$$

构造辅助函数

$$F_3(X,\sigma) = f(X) + \sigma P(X)$$

其中， $P(X)$具有以下形式

$$P(X) = \sum_{i=1}^{m} \phi(g_i(X)) + \sum_{j=1}^{l} \varphi(h_j(X))$$

$P(X)$为连续函数，典型取法如下

$$\phi\big(g_i(X)\big) = [\max\{0, -g_i(X)\}]^{\alpha}$$

$$\varphi\big(h_j(X)\big) = |h_j(X)|^{\beta}$$

其中，$\alpha \geq 1$，$\beta \geq 1$，均为给定的常数。通常取$\alpha = \beta = 2$

这样，把原问题转化为无约束最优化问题

$$\min F_3(X,\sigma)$$

罚因子的选择

实际计算中，罚因子的选择十分重要。

1、如果σ过大且最优解在可行域边界，进化算法将很快被推到可行域以内，不能返回正确解；

2、在搜索开始的时候，一个较大的罚因子将会阻碍非可行域的搜索，即如果可行域是几个非连通区域，则进化算法将仅移动到其中一个区域，很难搜索到其他区域；

3、如果σ过小，则花费较多时间在非可行域的搜索上，计算效率很差。

最小罚因子规则概念是很简单的，但是实现起来却是非常的困难。对于一个确定的进化算法，很多问题的可行域和非可行域的边界是未知的，因此很难确定它的精确位置。

罚因子的选择

因此，一般的策略是选取一个无穷大的严格递增的数列$\{\sigma_k\}$，从某个σ_1开始，对每个k，求解

$$\min f(X) + \sigma P(X)$$

从而得到一个极小点的序列$\{\bar{x}_{\sigma_k}\}$，在适当的条件下，这个序列将收敛于最优解。（该方法称为序列无约束极小化方法，即SUMT法）

算法步骤

Step1:选定初始点$X^{(0)}$，初始罚因子$\sigma_1 > 0$，取$(\sigma_1 = 1)$，罚因子的放大系数$C(C > 1)(C = 10)$，允许误差$\varepsilon > 0$，置$k = 1$

Step2:以$X^{(k-1)}$为初始点，求解无约束最优化问题

$$\min F\left(X, \sigma\right) = f(X) + \sigma P(X)$$

得到解$X^{(k)}$

Step3:若$\sigma_k P\left(X^{(k)}\right) < \varepsilon$，则$X^{(k)}$就是最优解，停止迭代；否则，转Step4

Step4:置$\sigma_{k+1} = C\sigma_k$；转步骤二

例题一（仅包含不等式约束）

用外点函数法求解如下优化问题：

$$\min {x_1}^2 + 2{x_2}^2$$
$$s.t.\ -x_1 - x_2 + 1 \le 0$$

解：问题只有不等式约束，对应的罚函数为：

$$p(x) = (\min\{x_1 + x_2 - 1,0\})^2$$

$$F(x, M_k) = {x_1}^2 + 2{x_2}^2 + M_k(\min\{x_1 + x_2 - 1,0\})^2$$
$$= \begin{cases} {x_1}^2 + 2{x_2}^2 & 若 x_1 + x_2 \ge 1 \\ {x_1}^2 + 2{x_2}^2 + M_k(-x_1 - x_2 + 1)^2 & 若 x_1 + x_2 < 1 \end{cases}$$

用解析法求$F(x,M_k)$的驻点，令

$$0=\frac{\partial F(x,M_k)}{\partial x_1}=\begin{cases}2x_1 & 若x_1+x_2\geq 1\\ 2x_1+2M_k(x_1+x_2-1) & 若x_1+x_2<1\end{cases}$$

$$0=\frac{\partial F(x,M_k)}{\partial x_2}=\begin{cases}4x_2 & 若x_1+x_2\geq 1\\ 4x_2+2M_k(x_1+x_2-1) & 若x_1+x_2<1\end{cases}$$

当$x_1+x_2\geq 1$时， $x_1=0=x_2$，舍去该点；

当$x_1+x_2<1$时，由$\frac{\partial F(x,M_k)}{\partial x_1}=\frac{\partial F(x,M_k)}{\partial x_2}=0$，得到

$$\begin{cases}2x_1+2M_k(x_1+x_2-1)=0\\ 4x_2+2M_k(x_1+x_2-1)=0\end{cases}$$

所以， $x_1=\frac{2M_k}{2+3M_k}$，$x_1=\frac{M_k}{2+3M_k}$，令$M_k\to+\infty$，得到

$$x^*=(\frac{2}{3},\frac{1}{3})^T$$

例题二（同时包含等式约束和不等式约束）

用罚函数法（外点法）解下列问题

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

解：定义罚函数

$$F(X,\sigma) = {x_1}^2 + {x_2}^2 + \sigma\{\max\{0, -x_1 + \frac{1}{2}\}^2 + (x_1 + x_2 - 1)^2\}$$

$$F(X,\sigma)=\begin{cases} x_1^2+x_2^2+\sigma(x_1+x_2-1) & x_1\geq\frac{1}{2} \\ x_1^2+x_2^2+\sigma(x_1-\frac{1}{2})^2+\sigma(x_1+x_2-1)^2 & x_1<\frac{1}{2} \end{cases}$$

当$x_1\geq\frac{1}{2}$

$$\frac{\partial F}{\partial x_1}=2x_1+2\sigma(x_1+x_2-1)$$

$$\frac{\partial F}{\partial x_2}=2x_2+2\sigma(x_1+x_2-1)$$

令$\frac{\partial F}{\partial x_1}=\frac{\partial F}{\partial x_2}=0$

则$x_1=x_2=\frac{\sigma}{1+2\sigma}$

$\sigma\to\propto$ 时，$x_1=x_2=\frac{1}{2}$

当$x_1 < \frac{1}{2}$时

$$\frac{\partial F}{\partial x_1} = 2x_1 + 2\sigma(x_1 + x_2 - 1) + 2\sigma(x_1 - \frac{1}{2})$$

$$\frac{\partial F}{\partial x_2} = 2x_2 + 2\sigma(x_1 + x_2 - 1)$$

令$\frac{\partial F}{\partial x_1} = \frac{\partial F}{\partial x_2} = 0$

$$x_1 = \frac{\sigma(3+\sigma)}{2(1+3\sigma+\sigma^2)}, x_2 = \frac{\sigma(2+\sigma)}{2(1+3\sigma+\sigma^2)}$$

当$\sigma \to \infty$时，$x_1 = x_2 = \frac{1}{2}$

综上，$\sigma \to \infty$时，$X = (\frac{1}{2}, \frac{1}{2})^T$

THANKS FOR YOUR WATCHING

your content is entered here, or by copying your text, select paste in this box and choose to retain only text. your content is typed here, or by copying your text, select paste in this box.your content is entered here, or by copying your text, select paste in this box and choose to retain only text. your content is typed here, or by copying your text, select paste in this box.