罚函数法

翟卫欣　副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院

目录

目录

罚函数法的含义

- 罚函数法是一种求解约束问题的特殊方法，其核心思想是将约束问题的约束条件转化为无约束问题的目标函数的一部分，从而可将约束问题转化为无约束问题进行求解。

目录

外点罚函数方法

以最小化问题为例，通常约束函数有如下特点：在该点，如果约束均满足，则约束函数为0，如果有的约束不满足，则该项为正，对其进行“惩罚”。原问题(P)：

$$min\, f(\boldsymbol{x})$$
$$s.t. c_i(\boldsymbol{x}) = 0, i \in \boldsymbol{E}$$

对于等式约束优化问题，常定义罚函数为：

$$P_E(\boldsymbol{x}, \sigma) = f(\boldsymbol{x}) + \frac{1}{2}\sigma \sum_{i \in \boldsymbol{E}} c_i^2(\boldsymbol{x})$$

等式右端的第二项称为惩罚项，$\sigma > 0$称为罚因子，P_E的P表示penalty，E表示Equality。

因为该函数的惩罚项是针对非可行点的，所以该函数被称为外点罚函数。惩罚项的形式可以有多种。本式由于其惩罚项的形式被称为二次罚函数。

课堂上对方法的收敛性不作证明。

目录

外点罚函数方法-等式约束

$$P_E(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\frac{1}{2}\sigma\sum_{i\in\boldsymbol{E}}c_i^2(\boldsymbol{x})$$

- 对非可行点而言，当σ变大时，惩罚项在罚函数中的比重加大，我们对函数P_E求极小值，相当于迫使其极小点向可行域靠近；
- 在可行域中，$c_i(\boldsymbol{x})=0,(i\in\boldsymbol{E})$，$P_E(\boldsymbol{x},\sigma)$的极小点与约束最优化问题的最优解相同。

外点罚函数方法-等式约束

利用外点罚函数来求解约束最优化问题的算法如下：

- 给定$\sigma_1 > 0, \epsilon_1 > 0, \epsilon > 0, k = 1, x_0$
- 以$\boldsymbol{x}_{k-1}$为初始点，求代数解$\boldsymbol{x}(\sigma_k) = argmin\ P_E(\boldsymbol{x}, \sigma_k)$；如果无法求代数解，也可以求解该无约束最优化问题的数值解，例如可以选择$||\nabla P_E(\boldsymbol{x}(\sigma_k), \sigma_k)|| \leqslant \epsilon_1$作为终止条件
- 当$||c(\boldsymbol{x}(\sigma_k))|| \leqslant \epsilon$时，迭代停止
- $\boldsymbol{x}_k = \boldsymbol{x}(\sigma_k)$，选取$\sigma_{k+1} = \sigma_k, k = k + 1$，转到步骤2

算法步骤2中，我们假定$P_E(\boldsymbol{x}, \sigma_k)$的局部极小点$\boldsymbol{x}(\sigma_k)$是存在的。

如何选取递增序列$\{\sigma_k\}$的问题直接影响到算法的有效性。如果增长过快，可能会影响到无约束子问题的求解，因为每一步无约束最优化问题的解是下一步无约束最优化问题的初始点。

而如果该序列增长缓慢，无约束最优化问题固然可以更好求解，但迭代速度必然受到影响。常设该序列为$\{10^k\}$。

外点罚函数方法-等式约束举例

$$min\ x_1^2 + x_2^2$$
$$s.t. x_1 - x_2 + 1 = 0$$

- 罚函数为$P_I(\boldsymbol{x}, \sigma) = x_1^2 + x_2^2 + \frac{\sigma}{2}(x_1 - x_2 + 1)^2$
- 罚函数的极小点为$\frac{\sigma}{2+2\sigma}(-1, 1)^T$

中国农业大学 China Agricultural University

外点罚函数方法-等式约束举例

- 图中绘制出了$P_E(\boldsymbol{x}, \sigma)$对于不同$\sigma$的等高线，$\sigma$取值分别为5，25，100，1000。由该图可以看出，随着σ的增大，$P_E(\boldsymbol{x}, \sigma)$的极小值趋于原问题最优解$\boldsymbol{x}^* = (-\frac{1}{2}, \frac{1}{2})$。

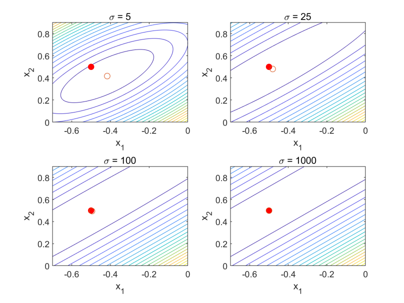

图 1: 变化趋势

中国农业大学 China Agricultural University

目录

外点罚函数方法-不等式约束

对于不等式约束问题：

$$\min f(\boldsymbol{x})$$
$$s.t. c_i(\boldsymbol{x}) \geqslant 0, i \in \boldsymbol{I}$$

常定义罚函数为：

$$P_I(\boldsymbol{x},\sigma) = f(\boldsymbol{x}) + \frac{1}{2}\sigma \sum_{i\in\boldsymbol{I}} (min\{c_i(\boldsymbol{x}),0\})^2$$

等式右端的第二项称为惩罚项，$\sigma > 0$称为罚因子，P_I的P表示penalty，I表示Inequality。

常定义$\tilde{c}_i(\boldsymbol{x}) = min\{c_i(\boldsymbol{x}),0\}$，所以罚函数可以改写为：

$$P_I(\boldsymbol{x},\sigma) = f(\boldsymbol{x}) + \frac{1}{2}\sigma \sum_{i\in\boldsymbol{I}} \tilde{c}_i(\boldsymbol{x})^2$$

外点罚函数方法-不等式约束举例

$$
\begin{aligned}
& min\ x_1^2 + x_2^2 \\
& s.t. x_1 - x_2 + 1 \leqslant 0
\end{aligned}
$$

罚函数为

$$
\begin{aligned}
P_E(\boldsymbol{x}, \sigma) &= x_1^2 + x_2^2 + \frac{\sigma}{2}(min\{-x_1 + x_2 - 1, 0\})^2 \\
&= \begin{cases} x_1^2 + x_2^2, & x_1 - x_2 + 1 \leqslant 0 \\ x_1^2 + x_2^2 + \dfrac{1}{2}\sigma(-x_1 + x_2 - 1)^2, & x_1 - x_2 + 1 > 0 \end{cases}
\end{aligned}
$$

所以

$$
\begin{aligned}
\frac{\partial P_I(\boldsymbol{x}, \sigma)}{\partial x_1} &= \ = \begin{cases} 2x_1, & x_1 - x_2 + 1 \leqslant 0 \\ 2x_1 - \sigma(-x_1 + x_2 - 1), & x_1 - x_2 + 1 > 0 \end{cases} \\
\frac{\partial P_I(\boldsymbol{x}, \sigma)}{\partial x_2} &= \ = \begin{cases} 2x_2, & x_1 - x_2 + 1 \leqslant 0 \\ 2x_2 + \sigma(-x_1 + x_2 - 1), & x_1 - x_2 + 1 > 0 \end{cases}
\end{aligned}
$$

外点罚函数方法-不等式约束举例

- 对于等高线$P_I(\boldsymbol{x},\sigma)=c$，当$c>\frac{1}{2}$时，等高线由两部分曲线组成：可行域内，等高线是一族圆的一部分；可行域外，等高线是一族椭圆的一部分。当$c\leqslant\frac{1}{2}$时，等高线是一族椭圆。图中绘制出了$P_I(\boldsymbol{x},\sigma)$对于$\sigma=5,25,100,1000$的等高线。随着$\sigma$的增大，$P_I(\boldsymbol{x},\sigma)$的极小值趋于原问题最优解$\boldsymbol{x}^*=(-\frac{1}{2},\frac{1}{2})$。

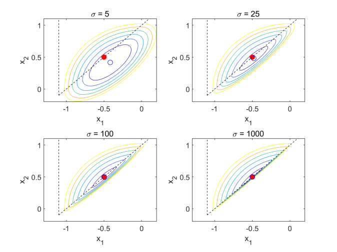

图 2: 变化趋势

目录

外点罚函数方法-一般最优化问题

- 对于一般约束最优化问题，将等式约束和不等式约束的惩罚项均考虑在内，我们得到如下的外点罚函数的形式：

$$P_{EI}(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\frac{1}{2}\sigma[\sum_{i\in\boldsymbol{E}}c_i(\boldsymbol{x})^2+\sum_{i\in\boldsymbol{I}}(min\{c_i(\boldsymbol{x}),0\})^2]$$

目录

内点罚函数方法

- 内点罚函数方法（又称障碍函数方法）与外点罚函数方法都是把约束最优化问题化为无约束最优化问题来求解的方法。
- 它们的不同之处在于：外点罚函数方法中无约束最优化问题的最优解序列由可行域外部逼近约束最优化问题的最优解，而内点罚函数方法中无约束最优化问题的最优解序列由可行域内部逼近约束最优化问题的最优解。
- 内点罚函数方法适宜于解决不等式约束最优化问题。

内点罚函数方法（障碍函数方法）

- 考虑不等式约束优化问题。该问题的倒数障碍函数定义为：
  $$B_I(\boldsymbol{x},\mu) = f(\boldsymbol{x}) + \mu \sum_{i\in \boldsymbol{I}} c_i(\boldsymbol{x})^{-1}$$
- 这里$\mu > 0$是惩罚因子，B_I的B代表Barrier，其下标I代表inverse；对数障碍函数定义为：
  $$B_L(\boldsymbol{x},\mu) = f(\boldsymbol{x}) - \mu \sum_{i\in \boldsymbol{I}} ln c_i(\boldsymbol{x})$$
- 这里$\mu > 0$是惩罚因子，B_I的B代表Barrier，其下标L代表ln。

内点罚函数方法

内点罚函数的特点如下：

- 内点罚函数（记为$B(\boldsymbol{x}, \mu)$）的极小点为严格极小点，即满足$c_i(\boldsymbol{x}) > 0 (i \in \boldsymbol{I})$的点；
- 当内点罚函数的极小点序列从可行域内部接近可行域的边界时，至少某一约束接近于起作用，此时占该项会无限增大，以防止迭代点越出可行域；
- 因为约束最优化问题的最优解可能落在边界上，为了使内点罚函数的极小点序列能够接近可行域的边界，允许$\mu \to 0$，以减少障碍项的数值。

内点罚函数方法

内点罚函数方法的算法：

- 给定初始内点$\boldsymbol{x}_0, \mu_1, \epsilon_1 > 0, \epsilon > 0, k = 1$；
- 以$\boldsymbol{x}_{k-1}$为初始点，求$\boldsymbol{x}(\mu_k) = argminB(\boldsymbol{x}, \mu_k)$，其迭代当$||\nabla_{\boldsymbol{x}} B(\boldsymbol{x}(\mu_k), \mu_k)|| \leqslant \epsilon_1$时停止；
- 当$\mu_k \sum_{i \in \boldsymbol{I}} c_i(\boldsymbol{x}(\mu_k))^{-1} \leqslant \mu$时，迭代停止；否则$\boldsymbol{x}_k = \boldsymbol{x}(\mu_k)$，选$\mu_{k+1} < \mu_k, k = k + 1$，转步骤2。

内点罚函数方法举例

$$min\ x_1^2+x_2^2$$

$$s.t. x_1-x_2+1\leqslant 0$$

构造内点罚函数$B(\boldsymbol{x},\mu)=x_1^2+x_2^2-\mu ln(-x_1+x_2-1)$

取μ分别为$0.1, 0.05, 0.01, 0.001$，绘制四幅等高线图。

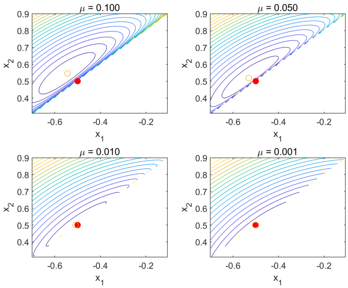

图 3: 变化趋势

Thank you for your attention!

翟卫欣　副教授

zhaiweixin@cau.edu.cn

中国农业大学信息与电气工程学院