

第 7 章：不等式约束

(1): 障碍方法

曹语

课程主页: <https://yucaoyc.github.io/math3806>

问题背景

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) \\ \text{subject to} & f_i(\mathbf{x}) \leq 0, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{array} \quad (\text{新增部分})$$

假设:

- ▶ $f_0, f_1, \dots, f_m : \mathbb{R}^n \rightarrow \mathbb{R}$ 是二次可微的凸函数;
- ▶ $\mathbf{A} \in \mathbb{R}^{p \times n}$, $\text{rank } \mathbf{A} = p$;
- ▶ Slater 条件成立 (或者直接考虑强对偶成立的情况)。

目录

1. 障碍方法
2. 初始阶段
3. 总结

障碍方法

思路：

- ▶ 等式约束问题 $\rightarrow$ 无约束问题
- ▶ 不等式约束问题 $\rightarrow$ 等式约束问题

障碍方法

思路:

- ▶ 等式约束问题 $\rightarrow$ 无约束问题
- ▶ 不等式约束问题 $\rightarrow$ 等式约束问题

构造想法:

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) + \sum_{i=1}^m l_-(f_i(\mathbf{x})) \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} \end{array}$$

其中

$$l_-(u) = \begin{cases} 0 & u \leq 0 \\ \infty & u > 0 \end{cases}$$

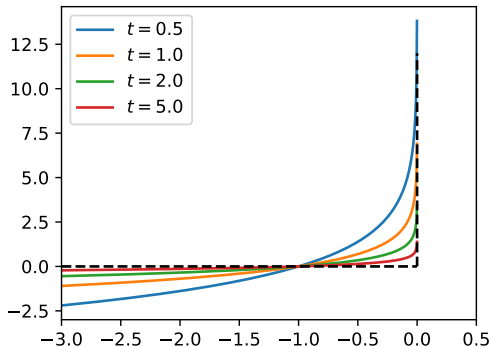
近似问题

为了应用梯度、牛顿法等，我们需要一个二阶可微的目标函数，因此我们可以考虑使用近似的函数 $\hat{l}_-(u) \approx l_-(u)$

$$\hat{l}_-(u) = -\frac{1}{t} \log(-u),$$

定义域 = $(-\infty, 0)$

当 t 很大时，该近似比较精确。



因此，问题可以变成

$$\begin{array}{ll} \text{minimize} & f_0(\mathbf{x}) + \sum_{i=1}^m -\frac{1}{t} \log(-f_i(\mathbf{x})) \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} \end{array}$$

上述问题可以等价于

$$\begin{array}{ll} \text{minimize} & tf_0(\mathbf{x}) + \phi(\mathbf{x}) \\ \text{subject to} & \mathbf{Ax} = \mathbf{b} \end{array} \quad (1)$$

函数 $\phi(\mathbf{x}) = -\sum_{i=1}^m \log(-f_i(\mathbf{x}))$ 被称为对数障碍函数或者对数障碍；该问题的最优解 $\mathbf{x}^*(t)$ 被称为中心点。

练习：该问题的 KKT 条件是什么？

KKT 条件是:

$$\begin{aligned}\mathbf{A}\mathbf{x}^*(t) &= \mathbf{b}, \quad f_i(\mathbf{x}^*(t)) < 0, \quad i = 1, \dots, m, \\ t\nabla f_0(\mathbf{x}^*(t)) + \sum_{i=1}^m \frac{1}{-f_i(\mathbf{x}^*(t))} \nabla f_i(\mathbf{x}^*(t)) + \mathbf{A}^\top \hat{\boldsymbol{\nu}} &= \mathbf{0}\end{aligned}$$

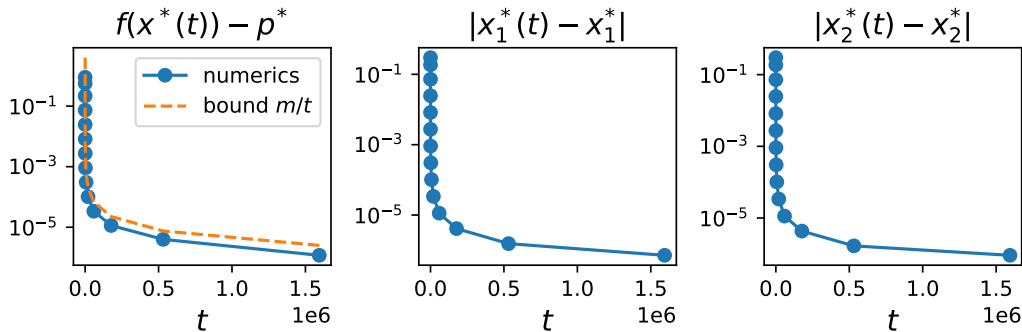
如果我们把

$$\tilde{\lambda}_i(t) = -\frac{1}{tf_i(\mathbf{x}^*(t))}, \quad i = 1, \dots, m, \quad \tilde{\boldsymbol{\nu}}(t) = \frac{\hat{\boldsymbol{\nu}}}{t}$$

该方程和原始带不等式约束的优化问题的 KKT 条件很类似。因此

$$\begin{aligned}g(\tilde{\boldsymbol{\lambda}}, \tilde{\boldsymbol{\nu}}) &= f_0(\mathbf{x}^*(t)) + \sum_{i=1}^m \tilde{\lambda}_i(t) f_i(\mathbf{x}^*(t)) + \tilde{\boldsymbol{\nu}}(t)^\top (\mathbf{A}\mathbf{x}^*(t) - \mathbf{b}) \\ &= f_0(\mathbf{x}^*(t)) - \frac{m}{t} \quad \implies \quad f_0(\mathbf{x}^*(t)) - p^* \leq \frac{m}{t}\end{aligned}$$

对于作业 1 里的线性规划问题: $m = 4$



此处验证了, 当 $t \rightarrow \infty$, 该近似问题的最优解收敛到真解, 且最优值的误差小于等于 m/t

因此，对于误差阈值 ϵ ，令 $t = m/\epsilon$ ，并求解上述的问题得到 $\mathbf{x}^*(t)$ ，我们可以保证函数值的误差小于 ϵ

即通过对数障碍函数来近似原问题的误差是可控的

挑战：对于很小的误差阈值 ϵ ，直接设定值很大的 t 对于求解原始优化问题有困难。

思路：对于困难的问题，分多步走

障碍方法/路径跟踪方法

在上世纪 60 年代，由 Anthony V. Fiacco, Garth P. McCormick 提出：

Algorithm 1: 障碍方法/路径跟踪方法

Input: 给定严格可行的 $\mathbf{x}$ ，阈值 ϵ ， t ， $\mu > 1$

```
1 while True do
2   通过  $\mathbf{x}$  为初始值，求解优化问题(1)得到  $\mathbf{x}^*(t)$ 
3   改进： $\mathbf{x} \leftarrow \mathbf{x}^*(t)$ 
4   停止准则：如果  $m/t \leq \epsilon$ ，则退出
5   令  $t \leftarrow t\mu$ ,
6 end
```

一般而言，我们可以取 μ 在 $3 \sim 100$ ；这里的符号 t 不是回溯搜索的 t

问题： μ 取得过大或者过小会发生什么？

牛顿步径

练习：如果使用牛顿法来求解问题(1)，牛顿步径如何求解呢？

牛顿步径

练习：如果使用牛顿法来求解问题(1)，牛顿步径如何求解呢？

答案：对于 $\tilde{f}(\mathbf{x}) = tf_0(\mathbf{x}) + \phi(\mathbf{x})$

$$\begin{bmatrix} \nabla^2 \tilde{f}(\mathbf{x}) & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_{\text{nt}} \\ \boldsymbol{\omega} \end{bmatrix} = \begin{bmatrix} -\nabla \tilde{f}(\mathbf{x}) \\ \mathbf{b} - \mathbf{A}\mathbf{x} \end{bmatrix}$$

变成了

$$\begin{bmatrix} t\nabla^2 f_0(\mathbf{x}) + \nabla^2 \phi(\mathbf{x}) & \mathbf{A}^\top \\ \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \Delta \mathbf{x}_{\text{nt}} \\ \boldsymbol{\omega} \end{bmatrix} = \begin{bmatrix} -t\nabla f_0(\mathbf{x}) - \nabla \phi(\mathbf{x}) \\ \mathbf{b} - \mathbf{A}\mathbf{x} \end{bmatrix}$$

因此，我们需要求解 Hessian 矩阵 $\nabla^2 \phi$ ，结果如下

$$\nabla^2 \phi(\mathbf{x}) = \sum_{i=1}^m \frac{1}{f_i(\mathbf{x})^2} \nabla f_i(\mathbf{x}) \nabla f_i(\mathbf{x})^\top + \sum_{i=1}^m \frac{1}{-f_i(\mathbf{x})} \nabla^2 f_i(\mathbf{x})$$

目录

1. 障碍方法
2. 初始阶段
3. 总结

构造 1

为了运行障碍法，我们需要找到某个 $\mathbf{x}$ 满足

$$f_i(\mathbf{x}) < 0, \quad i = 1, 2, \dots, m, \quad \mathbf{Ax} = \mathbf{b}$$

为了找到这样的初值，我们可以求解如下的优化问题：

$$\begin{aligned} \bar{p}^* = \text{minimize} \quad & s \\ \text{subject to} \quad & f_i(\mathbf{x}) \leq s, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \end{aligned}$$

- ☺ 若 $\bar{p}^* < 0$ ，我们即得到原优化问题的一个严格可行解。
- ☹ 若 $\bar{p}^* > 0$ ，我们可知原优化问题无可行解。
- ☹ 若 $\bar{p}^* = 0$ ，且能达到，则具有可行解（但不是严格可行解）。
- ☹ 若 $\bar{p}^* = 0$ ，且不能达到，则无可行解。

该问题其实依旧是不等式约束的凸优化问题，但是相比于原问题初始点很容易满足不等式约束。

构造 2

若我们仅需要找到可行解，我们也可以考虑如下：

$$\begin{array}{ll}\text{minimize} & \mathbf{1}^\top \mathbf{s} \\ \text{subject to} & f_i(\mathbf{x}) \leq s_i, \quad i = 1, \dots, m \\ & \mathbf{Ax} = \mathbf{b} \\ & \mathbf{s} \succeq \mathbf{0}\end{array}$$

若该问题的最优值达到 0，则原问题有可行解。

目录

1. 障碍方法
2. 初始阶段
3. 总结

总结

主要需要掌握的知识：

- ▶ 将不等式约束的问题变成一个近似的等式约束问题，即公式(1)
- ▶ 能够复述障碍法的流程

阅读作业 & 参考资料：

- ▶ 课本第 11.1 - 11.4 章