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:

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

其中$x \in \mathbb{R}^n$, 函数$f, c_i \quad (i \in \{1, \ldots, m\})$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 — 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 — 乘子法

下面以外罚函数法为例，分析罚因子σ无限增大的原因

对于等式约束优化问题：

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

设x^*为该问题的最优解，则$c_i(x^*) = 0$ $(\forall i \in E)$. 引入惩罚函数

$$P(x, \sigma) = f(x) + \sigma \sum_{i=1}^{p} c_i^2(x),$$

则惩罚函数$P(x, \sigma)$在点x^*处的梯度为：

$$\nabla P(x^*, \sigma) = \nabla f(x^*) + 2\sigma \sum_{i=1}^{p} c_i(x^*) \nabla c_i(x^*) = \nabla f(x^*).$$

而对应于等式约束优化问题的Lagrange函数

$$L(x, \lambda) = f(x) - \sum_{i=1}^{p} \lambda_i c_i(x),$$

4.4 序列无约束优化方法 — 乘子法

根据无约束优化问题的最优性条件, 则存在相应的Lagrange乘子向量$\lambda^* = (\lambda_1^*, \ldots, \lambda_p^*)^\top$使得

$$\nabla_x L(x^*, \lambda^*) = \nabla f(x^*) - \sum_{i=1}^{p} \lambda_i^* \nabla c_i(x^*) = 0,$$

即$\nabla f(x^*) = \sum_{i=1}^{p} \lambda_i^* \nabla c_i(x^*)$. 因而一般来说$\nabla f(x^*) \neq 0$, 即$\nabla P(x^*, \sigma) \neq 0$. 这说明$x^*$ 不可能是无约束优化问题$\min P(x, \sigma)$的稳定点. 所以，对于一个有限数值σ, 不能期望仅仅求解一个无约束优化问题$\min P(x, \sigma)$ 来获得原约束优化问题的最优解x^*, 只能期望

$$\lim_{\sigma \to \infty} \nabla P(x^*, \sigma) = 0.$$

通过上述的讨论，根据Lagrange函数和罚函数的特点，我们把这两类函数结合起来，构造出新的目标函数，使得在罚因子σ取值为适当大的情况下，借助于Lagrange乘子能逐步达到原约束优化问题的最优解，这也就是乘子法的基本思想.

4.4 序列无约束优化方法 — 乘子法

一、等式约束优化问题的乘子法

对于等式约束优化问题(4.4.2), 即

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) = 0, \quad \forall i \in E = \{1, 2, \ldots, p\}. \end{aligned} \tag{4.4.13}$$

设$\lambda = (\lambda_1, \ldots, \lambda_p)^\top \in \mathbb{R}^p$为相应的Lagrange乘子向量，则对应问题(4.4.13)的Lagrange函数为：

$$L(x, \lambda) = f(x) - \sum_{i=1}^{p} \lambda_i c_i(x) = f(x) - \lambda^\top c(x)$$

其中$c(x) = (c_1(x), \ldots, c_p(x))^\top$. 对Lagrange函数引入惩罚项，得

$$\begin{aligned} M(x, \lambda, \sigma) &= f(x) - \sum_{i=1}^{p} \lambda_i c_i(x) + \frac{\sigma}{2} \sum_{i=1}^{p} c_i^2(x) \\ &= L(x, \lambda) + \frac{\sigma}{2} c(x)^\top c(x), \end{aligned}$$

称函数$M(\cdot, \cdot, \cdot)$为问题(4.4.13)的增广Lagrange函数.

4.4 序列无约束优化方法 — 乘子法

对于等式约束优化问题(4.4.13)，设x^*是其最优解. 由最优性条件知：存在乘子向量λ^*，使得点对(x^*, λ^*) 为Lagrange函数的稳定点，即$\nabla_x L(x^*, \lambda^*) = 0$. 而附加的惩罚项$\frac{\sigma}{2}c(x)^\top c(x)$在点$x^*$处的梯度为0，由此

$$\nabla_x M(x^*, \lambda^*, \sigma^*) = \nabla_x L(x^*, \lambda^*) + \sigma \sum_{i=1}^{p} c_i(x) \nabla c_i(x^*) = 0,$$

这说明x^*是增广Lagrange函数$M(x, \lambda^*, \sigma)$的稳定点. 另外，也可证得：当σ取适当大的数值时，x^*将是函数$M(x, \lambda^*, \sigma)$的极小值点. 因而，对问题(4.4.13)的求解可等价转化为对某个λ^* 值, 求解增广Lagrange函数$M(x, \lambda^*, \sigma)$的极小值点.

4.4 序列无约束优化方法 — 乘子法

引理 4.4.2 设矩阵$A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$, 则下列命题等价:

(i) 对于满足$B^\top x = 0$的任意向量$x \neq 0$, 都有$x^\top Ax > 0$;

(ii) 存在$\delta^* > 0$, 使得对任意的$\delta \geq \delta^*$和任意的$0 \neq x \in \mathbb{R}^n$, 都有

$$x^\top (A + \delta BB^\top)x > 0. \tag{4.4.14}$$

定理 4.4.4 对于等式约束优化问题*(4.4.13)*, 设$x^* \in \mathbb{R}^n$和$\lambda^* \in \mathbb{R}^p$满足约束优化问题的二阶充分条件, 则存在一个数$\sigma^*$, 使得对所有的$\sigma \geq \sigma^*$, x^*都是无约束优化问题

$$\min_{x\in\mathbb{R}^n} M(x, \lambda^*, \sigma) \tag{4.4.15}$$

的严格局部最优解;

反之, 若$c(\bar{x}) = 0$, 并且对某个向量$\bar{\lambda}$, $\bar{x}$是问题*(4.4.15)*的局部最优解, 则$\bar{x}$是原约束优化问题*(4.4.13)*的局部最优解.

4.4 序列无约束优化方法 — 乘子法

证明　为证x^*是(4.4.15)的严格局部最优解，根据无约束优化问题的二阶充分条件，只需证：对任意的$\sigma \geq \sigma^*$, $\nabla_x M(x^*, \lambda^*, \sigma) = 0$且对满足$\nabla c(x)^\top z = 0$的非零向量$z \neq 0$，都有

$$z^\top \nabla_{xx}^2 M(x^*, \lambda^*, \sigma) z > 0,$$

其中$\nabla c(x)^\top = (\nabla c_1(x), \nabla c_2(x), \ldots, \nabla c_p(x))^\top$. 由$M(x, \lambda, \sigma)$的定义，则

$$\nabla_x M(x, \lambda, \sigma) = \nabla_x L(x, \lambda) + \sigma \nabla c(x) c(x).$$

所以，

$$\nabla_x M(x^*, \lambda^*, \sigma) = \nabla_x L(x^*, \lambda^*) + \sigma \nabla c(x^*) c(x^*) = \nabla_x L(x^*, \lambda^*) = 0.$$

再由$\nabla_x M(x, \lambda, \sigma)$的表达式，则

$$\nabla_{xx}^2 M(x, \lambda, \sigma) = \nabla_{xx}^2 L(x, \lambda) + \sigma \nabla^2 c(x) c(x) + \sigma \nabla c(x) \nabla c(x)^\top.$$

4.4 序列无约束优化方法 — 乘子法

进而有

$$\nabla_{xx}^2 M(x^*, \lambda^*, \sigma) = \nabla_{xx}^2 L(x^*, \lambda^*) + \sigma \nabla c(x^*) \nabla c(x^*)^\top.$$

根据二阶充分条件，对任意满足$\nabla c(x)^\top z = 0$的非零向量z，都有

$$z^\top \nabla_{xx}^2 L(x^*, \lambda^*) z > 0,$$

结合引理4.4.2, 于是存在$\sigma^* > 0$, 使得当$\sigma \geq \sigma^*$ 且$z \neq 0$时，都有

$$z^\top \nabla_{xx}^2 M(x^*, \lambda^*, \sigma) z > 0.$$

再由二阶充分条件，所以点x^*是无约束优化问题(4.4.15)的严格局部最优解.

反之，由于$\bar{x}$是无约束优化问题(4.4.15) 的局部最优解且$c(\bar{x}) = 0$, 因而对任意的$\hat{x} \in \mathscr{F} \cap B(\bar{x})$(其中$\mathscr{F}$为问题(4.4.13)的可行域，$B(\bar{x})$是$\bar{x}$的某邻域), 都有

$$M(\bar{x}, \bar{\lambda}, \sigma) \leq M(\hat{x}, \bar{\lambda}, \sigma).$$

由$c(\bar{x}) = c(\hat{x}) = 0$, 故$M(\bar{x}, \bar{\lambda}, \sigma) = f(\bar{x})$且$M(\hat{x}, \bar{\lambda}, \sigma) = f(\hat{x})$. 所以$f(\bar{x}) \leq f(\hat{x})$, 即$\bar{x}$为原约束优化问题(4.4.13)的局部最优解. □

4.4 序列无约束优化方法 — 乘子法

例 4.4.4 考虑约束优化问题

$$\begin{aligned} \min \quad & f(x) = x_1^2 - 3x_2 - x_2^2 \\ \text{s.t.} \quad & x_2 = 0. \end{aligned}$$

解 易知该问题的最优解为$x^* = (0,0)^\top$, 且相应的Lagrange乘子为$\lambda^* = -3$. 此外，该问题的增广Lagrange函数为：

$$\begin{aligned} M(x,\lambda,\sigma) &= x_1^2 - 3x_2 - x_2^2 - \lambda x_2 + \frac{\sigma}{2}x_2^2 \\ &= x_1^2 - (\lambda+3)x_2 + \frac{\sigma-2}{2}x_2^2. \end{aligned}$$

若取$\lambda = \lambda^* = -3$, 当$\sigma \geq \sigma^* = 2$时，原约束优化问题的最优解$x^* = (0,0)^\top$ 是函数$M(x,\lambda^*,\sigma)$的极小值点. 反之，对于无约束优化问题

$$\min_{x\in\mathbb{R}^2} M(x,\lambda,\sigma) = x_1^2 - (\lambda+3)x_2 + \frac{\sigma-2}{2}x_2^2.$$

4.4 序列无约束优化方法 — 乘子法

因为

$$\begin{aligned}\frac{\partial M}{\partial x_1} &= 2x_1;\\ \frac{\partial M}{\partial x_2} &= (\sigma-2)x_2-(\lambda+3).\end{aligned}$$

令$\dfrac{\partial M}{\partial x_1}=0$ 且$\dfrac{\partial M}{\partial x_2}=0$, 解之得$\bar{x}=(0,\dfrac{\lambda+3}{\sigma-2})^{\top}$. 如果取$\bar{\lambda}=-3$，$\sigma\geq 2$可知, $c(\bar{x})=\bar{x}_2=0$ 且$\bar{x}=(0,0)^{\top}$是问题

$$\min_{x\in\mathbb{R}^2} M(x,\bar{\lambda},\sigma) \qquad (\sigma\geq 2)$$

的局部最优解, 由定理4.4.4，则$\bar{x}=(0,0)^{\top}$为原约束优化问题的最优解. □

4.4 序列无约束优化方法 — 乘子法

注. 通过上面定理4.4.4 和例4.4.4 可以看出：对于等式约束优化问题，乘子法并不需要罚因子σ的值趋于无穷大，只要求选取罚因子σ的值为适当大的一个正数，即可保证无约束优化问题$\min_{x\in\mathbb{R}^n} M(x,\lambda^*,\sigma)$的最优解为原约束优化问题的最优解. 但是，乘子向量λ^*事先未知，如何确定乘子向量λ^*? 接下来将讨论此问题.

在例4.4.4中，发现乘子向量λ^*的取值正好是Lagrange函数在最优解处所对应的Lagrange乘子向量. 事实上，对于更一般约束优化问题也是如此. 但是，在实际问题中，我们需要通过迭代法来求解Lagrange乘子向量λ^*，即求得迭代点列$\{\lambda^k\}$满足$\lambda^k \to \lambda^*$. 下面讨论λ^*的求法.

4.4 序列无约束优化方法 — 乘子法

对于已知的点列$\{\lambda^k\}$, 求解对应无约束优化问题

$$\min M(x, \lambda^k, \sigma),$$

设其最优解为x^k, 则由一阶最优性条件可知,

$$\nabla_x M(x^k, \lambda^k, \sigma) = \nabla f(x^k) - \nabla c(x^k)(\lambda^k - \sigma c(x^k)) = 0. \qquad (4.4.16)$$

由于我们希望$x^k \to x^*, \lambda^k \to \lambda^*$, 且在$(x^*, \lambda^*)$ 处，有

$$\nabla f(x^*) - \nabla c(x^*)\lambda^* = 0. \qquad (4.4.17)$$

结合(4.4.16)和(4.4.17)式，可用下列迭代公式来修正乘子向量λ^k,

$$\lambda^{k+1} = \lambda^k - \sigma c(x^k) \qquad (4.4.18)$$

或$\lambda_i^{k+1} = \lambda_i^k - \sigma c_i(x^k)$, $\forall i \in E = \{1, \ldots, l\}$. 再由(4.4.18)式可知，若序列$\{\lambda^k\}$收敛，则$c(x^k) \to 0$. 因而，当$x^k \to x^*$时有$c(x^*) = 0$，即$x^*$为原约束优化问题的可行解. 在(4.4.16)中，令$k \to \infty$，得(4.4.17)成立，即x^*为原约束优化问题的K-T点.

4.4 序列无约束优化方法 — 乘子法

注. 通过上面推导可知，若x^k是问题$\min_{x\in\mathbb{R}^n} M(x,\lambda^k,\sigma)$ 的最优解，且(4.4.16) 式和(4.4.17) 式成立，则有$c(x^k)=0$. 反之，如果最优解x^k满足$c(x^k)=0$, 也可证得: x^k满足(4.4.16)式和(4.4.17)式成立. 由此可根据约束函数c在迭代点x^k处的信息，来作为乘子法的终止规则，即当$\|c(x^k)\|\le\varepsilon$时，算法停止迭代. 同时, 在迭代过程中，如果乘子序列$\{\lambda^k\}$ 不收敛或者收敛太慢，则可增大参数σ的值，再进行迭代. 收敛速度的快慢一般用$\|c(x^k)\|/\|c(x^{k-1})\|$ 来度量.

综上所述，下面给出求解等式约束优化问题乘子法的迭代算法，由于该算法最初是由Powell和Hestenes分别提出的，故又称为PH算法.

4.4 序列无约束优化方法 — 乘子法

算法 4.4.3 (等式约束优化问题的乘子法)　选取初始数据. 给定初始点x^0, 初始乘子向量λ^1, 初始罚因子$\sigma_1 > 0$ 以及放大系数$c > 1$, 允许误差$\varepsilon > 0$和常数$\theta \in (0, 1)$, 置$k = 1$.

步1 以x^{k-1}为初始点，求解无约束优化问题

$$\min M(x, \lambda^k, \sigma_k) = f(x) - (\lambda^k)^\top c(x) + \frac{\sigma_k}{2} c(x)^\top c(x),$$

得最优解为x^k.

步2 若$\|c(x^k)\| < \varepsilon$, 则算法终止, 点x^k为所求优化问题的近似最优解；否则，转步3.

步3 判断收敛速度快慢. 若$\dfrac{\|c(x^k)\|}{\|c(x^{k-1})\|} < \theta$，转步4; 否则，置$\sigma_k := c\sigma_k$, 转步4.

步4 置

$$\lambda^{k+1} = \lambda^k - \sigma_k c(x^k)$$

及$k := k + 1$, 转步1.

4.4 序列无约束优化方法 — 乘子法

例 4.4.5 用乘子法求解等式约束优化问题

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}x_1^2 + \frac{1}{6}x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 = 1. \end{aligned}$$

解 构造增广Lagrange函数为：

$$M(x_1, x_2, \lambda^k, \sigma_k) = \frac{1}{2}x_1^2 + \frac{1}{6}x_2^2 - \lambda^k(x_1 + x_2 - 1) + \frac{\sigma_k}{2}(x_1 + x_2 - 1)^2.$$

则

$$\begin{aligned} \frac{\partial M}{\partial x_1} &= x_1 - \lambda^k + \sigma_k(x_1 + x_2 - 1); \\ \frac{\partial M}{\partial x_2} &= \frac{1}{3}x_2 - \lambda^k + \sigma_k(x_1 + x_2 - 1). \end{aligned}$$

令 $\frac{\partial M}{\partial x_1} = 0$ 且 $\frac{\partial M}{\partial x_2} = 0$, 用解析方法求解无约束优化问题

$$\min_{x \in \mathbb{R}^2} M(x, \lambda^k, \sigma_k),$$

得最优解为:

$$x^k = (x_1^k, x_2^k)^\top = \left(\frac{\sigma_k + \lambda^k}{1 + 4\sigma_k}, \frac{3(\sigma_k + \lambda^k)}{1 + 4\sigma_k}\right)^\top, \qquad \forall k \in \{1, 2, \ldots\}.$$

把x_1, x_2的值代入乘子迭代公式(4.4.18)中，得

$$\lambda^{k+1} = \lambda^k - \sigma_k(x_1^k + x_2^k - 1),$$

即

$$\lambda^{k+1} = \frac{1}{1 + 4\sigma_k}\lambda^k + \frac{\sigma_k}{1 + 4\sigma_k}.$$

显然，当$\sigma_k > 0$时，数列$\{\lambda^k\}$都收敛，且σ_k值越大收敛速度越快. 不妨设$\sigma_k = \sigma_*$, $\forall k \geq N$, 其中N是一个充分大的正整数且σ_*是一个固定的正实数. 若令$k \to \infty$, $\lambda^k \to \lambda^*$，两边取极限，得$\lambda^* = \dfrac{1}{4}$. 进而，在x_1^k和x_2^k的表达式中，并令$k \to \infty$取极限，可得原问题的最优解为

$$x^* = (x_1^*, x_2^*)^\top = \left(\frac{1}{4}, \frac{3}{4}\right)^\top.$$

二、不等式约束优化问题的乘子法

考虑不等式约束优化问题(4.4.3), 即:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \geq 0, \quad \forall i \in I = \{1, 2, \ldots, m\}, \end{aligned} \tag{4.4.19}$$

为了利用等式约束优化问题所得到的结论，需引进辅助变量z_i ($i \in \{1, \ldots, m\}$)，使得上述不等式约束优化问题可等价转化为等式约束优化问题：

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) - z_i^2 = 0, \quad \forall i \in \{1, 2, \ldots, m\}. \end{aligned} \tag{4.4.20}$$

此时，问题(4.4.20)所对应的增广Lagrange函数为：

$$\hat{M}(x, z, \lambda, \sigma) = f(x) - \sum_{i=1}^{m} \lambda_i \left(c_i(x) - z_i^2 \right) + \frac{\sigma}{2} \sum_{i=1}^{m} \left(c_i(x) - z_i^2 \right)^2 \tag{4.4.21}$$

4.4 序列无约束优化方法 — 乘子法

从而把问题(4.4.19)转化为无约束优化问题

$$\min_{x\in\mathbb{R}^n, z\in\mathbb{R}^m} \hat{M}(x,z,\lambda,\sigma). \tag{4.4.22}$$

为了把上述无约束优化问题(4.4.22)化为只含有变量x求极小的问题，我们先考虑函数$\hat{M}$关于变量z的极小化问题：

$$\min_{z\in\mathbb{R}^m} \hat{M}(x,z,\lambda,\sigma).$$

由此，先求解出z^*，再代入优化问题(4.4.22)中，可得只含有变量x的无约束优化问题. 为求出z^*，根据无约束优化问题的一阶必要条件，令$\nabla_z \hat{M}(x,z,\lambda,\sigma)=0$, 整理得

$$z_i\left(\lambda_i - \sigma(c_i(x) - z_i^2)\right) = 0, \quad \forall i \in \{1,2,\ldots,m\}.$$

4.4 序列无约束优化方法 — 乘子法

(I) 当$\sigma c_i(x) - \lambda_i \geq 0$时，则$z_i^2 = \dfrac{1}{\sigma}(\sigma c_i(x) - \lambda_i)$ 或$z_i = 0$;

(II) 当$\sigma c_i(x) - \lambda_i < 0$时，则$z_i = 0$, 从而有$z_i^2 = 0$.

综合上述(I)和(II)的分析, 于是有

$$z_i^2 = \frac{1}{\sigma}\max\{0, \sigma c_i(x) - \lambda_i\}, \quad \forall i \in \{1, 2, \ldots, m\}. \tag{4.4.23}$$

将(4.4.23)式代入(4.4.21)式中, 可得原不等式约束优化问题(4.4.19)的增广Lagrange函数为:

$$\hat{M}(x, \lambda, \sigma) = f(x) + \frac{1}{2\sigma}\sum_{i=1}^{m}\left\{[\max\{0, \lambda_i - \sigma c_i(x)\}]^2 - \lambda_i^2\right\}.$$

4.4 序列无约束优化方法 — 乘子法

根据等式约束优化问题乘子法的乘子校正迭代公式，可以推出不等式约束优化问题乘子法的乘子校正迭代公式为:

$$\begin{aligned}
\lambda_i^{k+1} &= \lambda_i^k - \sigma[c_i(x^k) - z_i^2] \\
&= \lambda_i^k - \sigma c_i(x^k) + \max\{0, \sigma c_i(x^k) - \lambda_i^k\} \\
&= \max\{0, \lambda_i^k - \sigma c_i(x^k)\}, \qquad (\forall i \in \{1, 2, \ldots, m\}).
\end{aligned}$$

从而，求解原不等式约束优化问题(4.4.19)可转化为求解一系列无约束优化问题

$$\min_{x \in \mathbb{R}^n} M(x, \lambda^k, \sigma_k).$$

此外，终止规则为:

$$\left(\sum_{i=1}^{m} \left[c_i(x^k) - z_i^2 \right]^2 \right)^{\frac{1}{2}} = \left(\sum_{i=1}^{m} \left[\min\left\{ c_i(x^k), \frac{\lambda_i^k}{\sigma} \right\} \right]^2 \right)^{\frac{1}{2}} < \varepsilon.$$

三、一般约束优化问题的乘子法

对于一般约束优化问题(4.1.1), 即:

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & c_i(x) \geq 0, \quad \forall i \in I = \{1, 2, \ldots, p\}, \\ & c_i(x) = 0, \quad \forall i \in E = \{p+1, \ldots, m\}, \end{aligned} \tag{4.4.24}$$

根据上述对等式约束优化问题和不等式约束优化问题两种情况的讨论，可定义一般约束优化问题(4.4.24)相应的增广Lagrange函数为:

$$\begin{aligned} M(x, \lambda, \sigma) \quad = \quad & f(x) - \sum_{i \in E} \lambda_i c_i(x) + \frac{\sigma}{2} \sum_{i \in E} c_i^2(x) \\ & + \frac{1}{2\sigma} \sum_{i \in I} \left\{ [\max\{0, \lambda_i - \sigma c_i(x)\}]^2 - \lambda_i^2 \right\}. \end{aligned}$$

4.4 序列无约束优化方法 — 乘子法

对应的乘子校正迭代公式为：

$$
\begin{aligned}
\lambda_i^{k+1} &= \lambda_i^k - \sigma c_i(x^k), \quad \forall i \in E, \\
\lambda_i^{k+1} &= \max\{0, \lambda_i^k - \sigma c_i(x^k)\}, \quad \forall i \in I.
\end{aligned}
$$

因而，求解一般约束优化问题(4.4.24)可转化为求解一系列无约束优化问题

$$
\min_{x\in\mathbb{R}^n} M(x, \lambda^k, \sigma_k),
$$

且具体的计算步骤与等式和不等式约束情况类似. 此外，终止规则为:

$$
\phi_k := \left\{ \sum_{i\in E} c_i^2(x) + \sum_{i\in I} \left[\min\{c_i(x^k), \frac{\lambda_i^k}{\sigma}\} \right]^2 \right\}^{\frac{1}{2}} < \varepsilon. \qquad (4.4.25)
$$

4.4 序列无约束优化方法 — 乘子法

算法 4.4.4 *(一般约束优化问题的乘子法⟶PHP算法)* 选取初始数据. 给定初始点x^0, 初始乘子向量λ^1, 初始罚因子$\sigma_1 > 0$ 以及放大系数$c > 1$, 允许误差$\varepsilon > 0$和常数$\theta \in (0, 1)$, 置$k = 1$.

步1 以x^{k-1}为初始点, 求解无约束优化问题

$$\min M(x, \lambda^k, \sigma_k),$$

得最优解为x^k.

步2 按照*(4.4.25)*式计算ϕ_k, 若$\phi_k < \varepsilon$, 则算法终止, 点x^k为所求优化问题的近似最优解; 否则, 转步*3*.

步3 当$\phi_k/\phi_{k-1} < \theta$时, 转步*4*; 否则, 置$\sigma_k := c\sigma_k$, 转步*4*.

步4 根据乘子向量的校正迭代公式, 校正乘子向量λ_k,

$$\begin{aligned}
\lambda_i^{k+1} &= \lambda_i^k - \sigma_k c_i(x^k), \quad \forall i \in E, \\
\lambda_i^{k+1} &= \max\{0, \lambda_i^k - \sigma_k c_i(x^k)\}, \quad \forall i \in I.
\end{aligned}$$

置$k := k + 1$, 转步*1*.

4.4 序列无约束优化方法 — 乘子法

例 4.4.6 用乘子法求解不等式约束优化问题

$$\begin{aligned} \min \quad & f(x) = x_1^2 + 2x_2^2 \\ \text{s.t.} \quad & x_1 + x_2 \geq 1. \end{aligned}$$

解 构造增广Lagrange函数为：

$$\begin{aligned} M(x_1, x_2, \lambda, \sigma) &= x_1^2 + 2x_2^2 + \frac{1}{2\sigma}\left\{[\max\{0, \lambda - \sigma(x_1 + x_2 - 1)\}]^2 - \lambda^2\right\} \\ &= \begin{cases} x_1^2 + 2x_2^2 + \dfrac{1}{2\sigma}\{[\lambda - \sigma(x_1 + x_2 - 1)]^2 - \lambda^2\}, & x_1 + x_2 - 1 \leq \dfrac{\lambda}{\sigma}, \\ x_1^2 + 2x_2^2 - \dfrac{\lambda^2}{2\sigma}, & x_1 + x_2 - 1 > \dfrac{\lambda}{\sigma}. \end{cases} \end{aligned}$$

4.4 序列无约束优化方法 — 乘子法

当$x_1 + x_2 - 1 > \dfrac{\lambda}{\sigma}$时，令

$$
\begin{aligned}
\frac{\partial M}{\partial x_1} &= 2x_1 = 0, \\
\frac{\partial M}{\partial x_2} &= 4x_2 = 0,
\end{aligned}
$$

得$\bar{x} = (x_1, x_2)^\top = (0, 0)^\top$. $\bar{x}$不满足可行性条件$x_1 + x_2 \geq 1$, 说明$\bar{x}$不是原问题的最优解.

当$x_1 + x_2 - 1 \leq \dfrac{\lambda}{\sigma}$时，令

$$
\begin{aligned}
\frac{\partial M}{\partial x_1} &= 2x_1 - [\lambda - \sigma(x_1 + x_2 - 1)] = 0, \\
\frac{\partial M}{\partial x_2} &= 4x_2 - [\lambda - \sigma(x_1 + x_2 - 1)] = 0,
\end{aligned}
$$

得

$$
x_1 = \frac{2(\lambda + \sigma)}{4 + 3\sigma}, \quad x_2 = \frac{\lambda + \sigma}{4 + 3\sigma}, \tag{4.4.26}
$$

且$\tilde{x} = (x_1, x_2)^\top$满足条件$x_1 + x_2 - 1 \leq \dfrac{\lambda}{\sigma}$.

4.4 序列无约束优化方法 — 乘子法

将λ替换成λ^k代入乘子的修正公式中，修正λ^k, 得

$$\lambda^{k+1} = \max\{0, \lambda^k - \sigma_k(x_1^k + x_2^k - 1)\} = \max\left\{0, \frac{4(\lambda^k + \sigma_k)}{4 + 3\sigma_k}\right\}.$$

不妨取$\lambda^1 > 0$, 由于所有的$\sigma_k > 0$, 所以由上式可得

$$\lambda^{k+1} = \frac{4(\lambda^k + \sigma_k)}{4 + 3\sigma_k} = \frac{4}{4 + 3\sigma_k}\lambda^k + \frac{4\sigma_k}{4 + 3\sigma_k}. \qquad (4.4.27)$$

因此, 序列$\{\lambda^k\}$是收敛的, 不妨设其收敛到λ_*. 在(4.4.27)中, 令$k \to \infty$且$\lim_{k\to\infty} \sigma_k = \sigma_*$, 则容易得到: $\lambda_* = \frac{4}{3}$. 进一步, 在(4.4.26)中, 令$k \to \infty$, 可得$x^* = (\frac{2}{3}, \frac{1}{3})^\top$. 因此, x^*为原约束优化问题的最优解. □

4.4 序列无约束优化方法 — 乘子法作业

4.25 用乘子法求解下列约束优化问题：

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

$\text{s.t.} \quad c(x) = x_1 + 2x_2 = 4.$

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

$\text{s.t.} \quad c(x) = x_1 + x_2 \geq 1.$