6-3：罚函数方法与内点障碍法

罚函数方法

罚函数方法是一类重要的约束优化求解方法，通常通过解一系列**无约束最优化问题**来获取原约束优化问题最优解。

罚函数法是一类不可行点的方法，对初始点的选取没有可行性的限制。对不可行的迭代点施加惩罚，并随着迭代的进展增大惩罚量，增加迭代点可行性。

本节主要学习如下两种经典罚函数方法：

- 二次罚函数法
- 障碍罚函数法

二次罚函数定义

考虑等式约束问题：

$$\begin{aligned} \min_x \quad & f(\boldsymbol{x}) \\ \text{s.t.} \quad & c_i(\boldsymbol{x}) = 0,\ i \in \mathcal{E}, \end{aligned} \tag{6.3.1}$$

罚函数由目标函数加上惩罚项组成, 定义为:

$$\mathrm{P_E}(\boldsymbol{x}, \sigma) \stackrel{\text{def}}{=} f(\boldsymbol{x}) + \frac{1}{2}\sigma \sum_{i\in\mathcal{E}} c_i^2(\boldsymbol{x}), \tag{6.3.2}$$

这里$\sigma > 0$是罚参数，当σ趋于无穷时，违反约束的惩罚项剧烈地增大。该罚函数是针对非可行点，所以该函数称为**外点罚函数**，亦称为**二次罚函数方法**。

例6.3.1 考虑等式约束优化问题

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

二次罚函数为：

$$\mathrm{P_E}(\boldsymbol{x}, \sigma) = x_1^2 + x_2^2 + \frac{1}{2}\sigma(x_1 - x_2 + 1)^2.$$

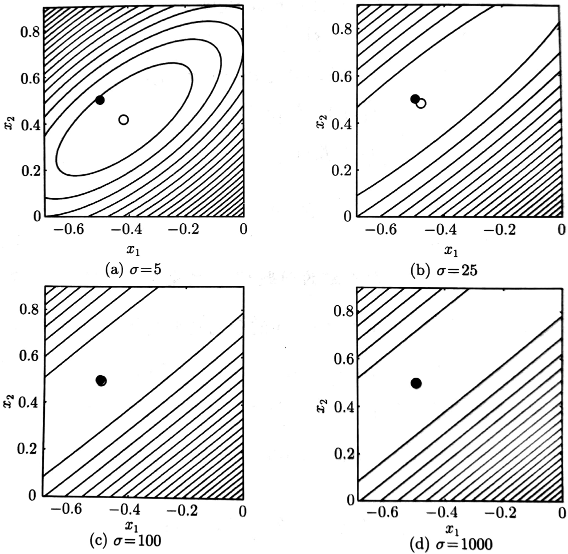
原问题极小点为：

$$\left(-\frac{1}{2}, \frac{1}{2}\right)$$

二次罚函数极小点为：

$$\frac{\sigma}{2(1+\sigma)}(-1, 1)^T$$

例题分析



0.8
0.6
0.4
0.2
0
x_2
-0.6
-0.4
-0.2
0
x_1
(a) $\sigma=5$
(b) $\sigma=25$
(c) $\sigma=100$
(d) $\sigma=1000$

二次罚函数方法步骤

基本思想: 确定一个逐渐递增参数序列$\{\sigma_k\}$时，用无约束优化技术来求解$\mathrm{P_E}(\boldsymbol{x},\sigma_k)$的极小点$\boldsymbol{x}(\sigma_k)$, 可作为第$k+1$次迭代优化$\mathrm{P_E}(\boldsymbol{x},\sigma_{k+1})$的高质量初始点。当$\sigma_k\downarrow\infty$时，罚函数$\mathrm{P_E}(\boldsymbol{x},\sigma_k)$的极小点$\boldsymbol{x}(\sigma_k)$就是原问题的极小点。

算法6.3.1（二次罚函数方法）

步1 给定$\sigma_1>0,\varepsilon_1>0,\varepsilon>0$, 初始点$\boldsymbol{x}_0$，$k=0$;

步2 以$\boldsymbol{x}_{k-1}$为初始点，求$\boldsymbol{x}(\sigma_k)=\arg\min \mathrm{P_E}(\boldsymbol{x},\sigma_k)$。求解该问题的停止准则$\|\nabla \mathrm{P_E}(\boldsymbol{x}(\sigma_k),\sigma_k)\|\leq\varepsilon_1$;

步3 当$\|\boldsymbol{c}(\boldsymbol{x}(\sigma_k))\|\leq\varepsilon$时，终止迭代并输出$\boldsymbol{x}_k$;

步4 令$\boldsymbol{x}_k=\boldsymbol{x}(\sigma_k)$, 选$\sigma_{k+1}>\sigma_k$, $k:=k+1$, 转步2。

注意： 罚参数序列$\{\sigma_k\}$要合适地选择, 例如，$\sigma_{k+1}=10\sigma_k$。

定理6.3.1.

假定$f(\boldsymbol{x})$在可行域上有下界，并且**算法**6.3.1步2中求得$\mathrm{P_E}(\boldsymbol{x},\sigma_k)$的全局极小点$\boldsymbol{x}(\sigma_k)$. 若序列$\{\sigma_k\}$满足$\sigma_{k+1}\geq\sigma_k$，则：

$$\mathrm{P_E}(\boldsymbol{x}_k,\sigma_k)\leq\mathrm{P_E}(\boldsymbol{x}_k,\sigma_{k+1}) \tag{6.3.4a}$$

$$\boldsymbol{c}_k^T\boldsymbol{c}_k\geq\boldsymbol{c}_{k+1}^T\boldsymbol{c}_{k+1} \tag{6.3.4b}$$

$$f_k\leq f_{k+1} \tag{6.3.4c}$$

且当$\sigma_k\to\infty$时有：$\boldsymbol{c}_k\to 0$, $\{\boldsymbol{x}_k\}$的任何一个聚点$\boldsymbol{x}^*$为问题(6.3.1)的全局最优解。

证明：因为$\boldsymbol{x}_k=\boldsymbol{x}(\sigma_k)$为$\mathrm{P_E}(\boldsymbol{x},\sigma_k)$的全局最优解以及$\sigma_{k+1}\geq\sigma_k$, 有：

$$\mathrm{P_E}(\boldsymbol{x}_k,\sigma_k)\leq\mathrm{P_E}(\boldsymbol{x}_{k+1},\sigma_k)\leq\mathrm{P_E}(\boldsymbol{x}_{k+1},\sigma_{k+1})\leq\mathrm{P_E}(\boldsymbol{x}_k,\sigma_{k+1}) \tag{a}$$

证明续：从以上不等式的两端可以得到(6.3.4a)。另外，该不等式外层与内层分别相减可得：

$$(\sigma_{k+1}-\sigma_k)\boldsymbol{c}_k^T\boldsymbol{c}_k \geq (\sigma_{k+1}-\sigma_k)\boldsymbol{c}_{k+1}^T\boldsymbol{c}_{k+1}$$

由此可以得到(6.3.4b)。

由(a)中的第一个不等式与(6.3.4b)得：

$$0 \leq \frac{1}{2}\sigma_k(\boldsymbol{c}_k^T\boldsymbol{c}_k - \boldsymbol{c}_{k+1}^T\boldsymbol{c}_{k+1}) \leq f_{k+1}-f_k$$

即得到(6.3.4c)。

因为$\boldsymbol{x}_k$是$\mathrm{P_E}(\boldsymbol{x},\sigma_k)$的全局极小点，有：

$$\mathrm{P_E}(\boldsymbol{x}_k,\sigma_k) \leq \mathrm{P_E}(\boldsymbol{x},\sigma_k), \boldsymbol{x}\in\mathcal{D} \tag{b}$$

定理6.3.1证明

证明续：由$f(\boldsymbol{x})$在可行域上有下界及$f(\boldsymbol{x}) = \mathrm{P_E}(\boldsymbol{x}, \sigma_k)(\boldsymbol{x} \in \mathcal{D})$知，$\mathrm{P_E}(\boldsymbol{x}, \sigma$
可行域上也有下界，从而：

$$\mathrm{P_E}(\boldsymbol{x}_k, \sigma_k) \leq \inf_{\boldsymbol{x} \in \mathcal{D}} \mathrm{P_E}(\boldsymbol{x}, \sigma_k)$$

即：

$$f(\boldsymbol{x}_k) + \frac{1}{2}\sigma_k \boldsymbol{c}_k^T \boldsymbol{c}_k \leq \inf_{\boldsymbol{x} \in \mathcal{D}} \mathrm{P_E}(\boldsymbol{x}, \sigma_k). \qquad \text{(c)}$$

令$\sigma_k \to \infty$, 由(6.3.4c)知$\{f_k\}$非减，必有:

$$\boldsymbol{c}_k \to 0.$$

设$\boldsymbol{x}_k \to \boldsymbol{x}^*$，$\boldsymbol{x}^*$为可行点，则：

$$f(\boldsymbol{x}^*) \geq \inf_{\boldsymbol{x} \in \mathcal{D}} f(\boldsymbol{x}) \qquad \text{(d)}$$

定理6.3.1证明

证明续：因为$\mathrm{P_E}(\boldsymbol{x},\sigma_k)=f(\boldsymbol{x})(\boldsymbol{x}\in\mathcal{D})$，由(c)式知：

$$f(\boldsymbol{x}_k)\le\inf_{\boldsymbol{x}\in\mathcal{D}}f(\boldsymbol{x})$$

令$k\to\infty$，得到：

$$f(\boldsymbol{x}^*)\le\inf_{\boldsymbol{x}\in\mathcal{D}}f(\boldsymbol{x})\tag{e}$$

综合(d)与(e)式知，$\boldsymbol{x}^*$是约束优化问题(6.3.1)的全局最优解。 ■

上面关于外罚函数算法的收敛定理要求罚函数的全局极小点，对很多问题而言很难做到。

罚函数方法收敛性

定理6.3.2.

在**算法**6.3.1中，若在步2有$\|\nabla \mathrm{P}_{\mathrm{E}}(\boldsymbol{x}_k,\sigma_k)\| \le \varepsilon_k$，而：

$$\lim_{k\to\infty} \varepsilon_k = 0, \sigma_k \to \infty,$$

对$\{\boldsymbol{x}_k\}$的任何极限点$\boldsymbol{x}^*$，$\{\nabla c_i(\boldsymbol{x}^*), i\in\mathcal{E}\}$线性无关，则$\boldsymbol{x}^*$是等式约束最优化问题(6.3.1)的KKT点，且:

$$\lim_{k\to\infty} (-\sigma_k c_i(\boldsymbol{x}_k)) = \lambda_i^*, i \in \mathcal{E} \tag{6.3.5}$$

其中$\boldsymbol{\lambda}^*$是$\boldsymbol{x}^*$的相应Lagrange乘子。

定理6.3.2证明

证明：先证明$\boldsymbol{x}^*$是问题(6.3.1)的可行点。因为：

$$\nabla_{\boldsymbol{x}} \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k) = \boldsymbol{g}(\boldsymbol{x}_k) + \sum_{i \in \mathcal{E}} \sigma_k c_i(\boldsymbol{x}_k) \nabla c_i(\boldsymbol{x}_k), \tag{a}$$

由**算法**6.3.1中步2求子问题$\mathrm{P_E}(\boldsymbol{x}, \sigma_k)$的算法的终止准则有：

$$\left\| \boldsymbol{g}(\boldsymbol{x}_k) + \sum_{i \in \mathcal{E}} \sigma_k c_i(\boldsymbol{x}_k) \nabla c_i(\boldsymbol{x}_k) \right\| \leq \varepsilon_k$$

上式可转化为：

$$\left\| \sum_{i \in \mathcal{E}} c_i(\boldsymbol{x}_k) \nabla c_i(\boldsymbol{x}_k) \right\| \leq \frac{1}{\sigma_k} \left[\varepsilon_k + \|\boldsymbol{g}(\boldsymbol{x}_k)\| \right]$$

定理6.3.2证明

证明续：令$k \to \infty$, 有$\sigma_k \to \infty$, 则：

$$\sum_{i \in \mathcal{E}} c_i(\boldsymbol{x}^*) \nabla c_i(\boldsymbol{x}^*) = 0$$

因为$\{\nabla c_i(\boldsymbol{x}^*), i \in \mathcal{E}\}$线性无关, 所以：

$$c_i(\boldsymbol{x}^*) = 0, i \in \mathcal{E}$$

故$\boldsymbol{x}^*$为问题(6.3.1)的可行点。

下证$\boldsymbol{x}^*$是问题(6.3.1)的KKT点. 记：

$$A(\boldsymbol{x}) = [\nabla c_1(\boldsymbol{x}), \cdots, \nabla c_m(\boldsymbol{x})], \quad \boldsymbol{\lambda}_k = -\sigma_k \boldsymbol{c}(\boldsymbol{x}_k)$$

定理6.3.2证明

证明续：(a)式可写成：

$$A_k \boldsymbol{\lambda}_k = \boldsymbol{g}(\boldsymbol{x}_k) - \nabla_{\boldsymbol{x}} \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k) \tag{b}$$

对足够大的k，A_k列满秩(约束函数极限点处梯度向量线性无关)，因而$A_k^T A_k$非奇异。(b)式两边左乘A_k^T得：

$$\boldsymbol{\lambda}_k = (A_k^T A_k)^{-1} A_k^T (\boldsymbol{g}(\boldsymbol{x}_k) - \nabla_{\boldsymbol{x}} \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k))$$

因而$\|\nabla_{\boldsymbol{x}} \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k)\| \le \varepsilon_k$, 而$\lim_{k\to\infty} \varepsilon_k = 0$, 故：

$$\lim_{k\to\infty} \boldsymbol{\lambda}_k = ({A^*}^T A^*)^{-1} {A^*}^T \boldsymbol{g}^* \triangleq \boldsymbol{\lambda}^*$$

令$k \to \infty$，由(b)式可得：$\boldsymbol{g}^* - A^* \boldsymbol{\lambda}^* = 0$, 所以,$\boldsymbol{x}^*$是KKT点，$\boldsymbol{\lambda}^*$为对应的Lagrange乘子。 ■

罚函数方法数值困难性

当$\sigma_k \to \infty$时，**算法**6.3.1中步2对无约束优化问题求解变得越来越困难。这是因为随着σ_k的增大，Hesse矩阵$\nabla^2 \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k)$越来越恶化。具体来说，Hesse矩阵：

$$\nabla^2 \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k) = \nabla^2_{\boldsymbol{x}} L(\boldsymbol{x}_k, \boldsymbol{\lambda_k}) + \sigma_k A_k A_k^T.$$

矩阵：$\sigma_k A_k A_k^T$的秩为m, 这里假定$m < n$。随着$\sigma_k \to \infty$, $\nabla^2 \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k)$有$n$特征值趋于无穷，其余特征值有界。

对于前面**例**6.3.1中：

$$\nabla^2 \mathrm{P_E}(\boldsymbol{x}_k, \sigma_k) = \begin{bmatrix} 2+\sigma & -\sigma \\ -\sigma & 2+\sigma \end{bmatrix}$$

当$\sigma = 5, 25, 100, 1000$，该矩阵对应的条件数分别为：$6, 26, 101, 1001$。

不等式约束二次罚函数

考虑不等式约束最优化问题：

$$\min \quad f(\boldsymbol{x}) \tag{6.3.6a}$$

$$\text{s.t.} \quad c_i(\boldsymbol{x}) \geq 0, i \in \mathcal{I} \tag{6.3.6b}$$

其外点罚函数定义为：

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

例6.3.2 考虑不等式约束优化问题:

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

该问题的最优解为：$\boldsymbol{x}^* = \left(-\frac{1}{2}, \frac{1}{2}\right)$.

例题分析

二次罚函数为:

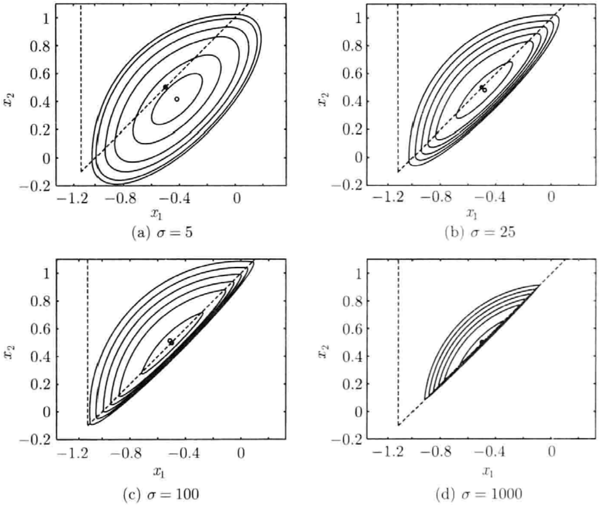
$$
\begin{aligned}
\mathrm{P_I}(\boldsymbol{x},\sigma) &= x_1^2+x_2^2+\frac{\sigma}{2}[\min\{-x_1+x_2-1,0\}]^2\\
&= \begin{cases} x_1^2+x_2^2, & x_1-x_2+1\le 0\\ x_1^2+x_2^2+\frac{\sigma}{2}(-x_1+x_2-1)^2, & x_1-x_2+1>0 \end{cases}
\end{aligned}
$$

所以:

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

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

当$x_1-x_2+1>0$时，外点罚函数的极小点为$\frac{\sigma}{2(1+\sigma)}(-1,1)^T$。



1
0.8
0.6
0.4
0.2
0
−0.2
−1.2
−0.8
−0.4
0
x_1
x_2
(a) $\sigma = 5$
(b) $\sigma = 25$
(c) $\sigma = 100$
(d) $\sigma = 1000$

外点罚函数方法

对于一般约束优化问题，既有等式约束也有不等式约束，外点罚函数定义为：

$$\mathrm{P_{EI}}(\boldsymbol{x},\sigma)=f(\boldsymbol{x})+\frac{1}{2}\sigma\left[\sum_{i\in\mathcal{E}}c_i^2(\boldsymbol{x})+\sum_{i\in\mathcal{I}}\tilde{c}_i^2(\boldsymbol{x})\right]$$

其中：

$$\tilde{c}_i(\boldsymbol{x})=\min\{c_i(\boldsymbol{x}),0\},i\in\mathcal{I}$$

一般约束优化问题的外点罚函数方法和等式约束最优化问题类似，在收敛性定理中，若外点罚函数为$\mathrm{P_{EI}}(\boldsymbol{x},\sigma)$，定理结论仍然成立。

障碍罚函数法

内点障碍函数法通过在目标函数上引入一个关于约束的障碍项，当迭代点由可行域的内部接近可行域的边界时，障碍项将趋于无穷大来迫使迭代点返回可行域的内部，从而保持迭代点的严格可行性。

障碍函数法与外点罚函数方法都是把约束优化问题转化为无约束优化问题来求解的方法。不同之处在于：

- 外点罚函数方法中无约束优化问题的最优解序列由可行域的外部逼近约束优化问题的最优解。
- 障碍函数方法中的无约束优化问题的最优解序列由可行域的内部逼近约束优化问题的最优解，是一类保持严格可行性的方法。因此，这类方法适用于不等式约束的最优化问题。

障碍函数

考虑不等式约束最优化问题(6.3.6), 主要有如下两种障碍函数：

- **倒数障碍函数:**

$$\mathrm{B_I}(\boldsymbol{x};\mu) = f(\boldsymbol{x}) + \mu \sum_{i\in\mathcal{I}} \frac{1}{c_i(\boldsymbol{x})} \tag{6.3.7}$$

  其中$\mu > 0$为障碍因子。

- **对数障碍函数:**

$$\mathrm{B_L}(\boldsymbol{x};\mu) = f(\boldsymbol{x}) - \mu \sum_{i\in\mathcal{I}} \ln c_i(\boldsymbol{x}) \tag{6.3.8}$$

  其中$\ln(\cdot)$为自然对数。

障碍函数

定义严格可行内点区域为:

$$\mathrm{int}(\mathcal{D}) \stackrel{\text{def}}{=} \{\boldsymbol{x} \in \mathbb{R}^n |\ c_i(\boldsymbol{x}) > 0, \forall i \in \mathcal{I}\}$$

并假设$\mathrm{int}(\mathcal{D})$是非空的。

障碍函数具有如下性质：

- 障碍函数亦称为内点罚函数，该函数的极小点为严格可行点，及满足$c_i(\boldsymbol{x}), (i \in \mathcal{I})$, 在外部的值或者无定义或者是无穷的。
- 当障碍函数的极小点序列从可行域内部接近可行域的边界时，至少某一个约束接近于起作用，此时障碍项会无限增大，以防止迭代点跃出可行域;
- 因为约束优化问题最优解可能位于边界上，允许$\mu \to 0$使得障碍函数极小点能够接近可行域的边界，来减少障碍项的数值。

障碍函数方法步骤

算法6.3.2（障碍函数方法）

步1 给定初始内点$\boldsymbol{x}_0$, $\mu_1 > 0$，$\varepsilon_1 > 0, \epsilon > 0$，$k := 1$.

步2 以$\boldsymbol{x}_{k-1}$为初始点，求:$\boldsymbol{x}(\mu_k) = \arg\min B(\boldsymbol{x}, \mu_k)$ 其迭代停止准则为：$\|\nabla_{\boldsymbol{x}} B(\boldsymbol{x}(\mu_k), \mu_k)\| \leq \varepsilon_1$.

步3 若收敛准则$\mu_k \sum_{i\in\mathcal{I}} c_i(\boldsymbol{x}(\mu_k))^{-1} \leq \epsilon$ 满足，迭代停止；否则，$\boldsymbol{x}_k = \boldsymbol{x}(\mu_k)$, 选$\mu_{k+1} < \mu_k$, $k := k+1$, 转步2.

以上方法中：

- 障碍函数$B(\boldsymbol{x}, \mu_k)$可以选用倒数障碍函数或者对数障碍函数二者之中任何一个。
- 障碍参数μ_{k+1}可采用规则: 若求$B(\boldsymbol{x}, \mu_k)$的极小点不太困难，以及初始点比较可靠，则可显著减小μ_{k+1}，例如取$\mu_{k+1} = 0.1\mu_k$. 否则，取$\mu_{k+1} = 0.7\mu_k$.

障碍函数方法停机准则

可以证明：若$f(\boldsymbol{x})$在可行域$\mathcal{D}$上有下界，对$\epsilon > 0$, **算法**6.3.2或者有限终止，或对倒数障碍函数及对数障碍函数分别有：

$$\lim_{k\to\infty} \mu_k \sum_{i\in\mathcal{I}} c_i(\boldsymbol{x}_k)^{-1} = 0$$

$$\lim_{k\to\infty} \mu_k \sum_{i\in\mathcal{I}} \ln c_i(\boldsymbol{x}_k) = 0$$

且

$$\lim_{k\to\infty} f(\boldsymbol{x}_k) = \inf_{i\in\mathrm{int}(\mathcal{D})} f(\boldsymbol{x}).$$

算法产生的序列$\{\boldsymbol{x}_k\}$的聚点是不等式约束优化问题的最优解，所以**算法**6.3.2步3中的终止准则以倒数障碍函数为例的终止准则。

例题分析

例6.3.3 考虑非线性约束最优化问题:

$$\begin{aligned} \min \quad & f(x) = x_1 - 2x_2, \\ \text{s.t.} \quad & 1 + x_1 - x_2^2 \geq 0, x_2 \geq 0. \end{aligned}$$

解：该问题的最优解为$\boldsymbol{x}^* = (0, 1)^T$, 对数障碍函数为

$$\mathrm{B_L}(\boldsymbol{x}; \mu) = x_1 - 2x_2 - \mu[\log(1 + x_1 - x_2^2) + \log(x_2)].$$

为确定$\mathrm{B_L}(\boldsymbol{x}; \mu)$的无约束最优解，考察一阶最优性条件:

$$\begin{cases} 1 - \frac{\mu}{1+x_1-x_2^2} = 0, \\ -2 + \frac{2\mu x_2}{1+x_1-x_2^2} - \frac{\mu}{x_2} = 0. \end{cases}$$

例题分析

解续：求解这个方程组，得

$$x_1(\mu)=\frac{\sqrt{1+2\mu}+3\mu-1}{2},\quad x_2(\mu)=\frac{\sqrt{1+2\mu}+1}{2}.$$

关于障碍参数$\mu\to 0$取极限，得

$$\lim_{\mu\to 0}x_1(\mu)=\lim_{\mu\to 0}\frac{\sqrt{1+2\mu}+3\mu-1}{2}=0,$$
$$\lim_{\mu\to 0}x_2(\mu)=\lim_{\mu\to 0}\frac{\sqrt{1+2\mu}+1}{2}=1.$$

即有：$\lim_{\mu\to 0}\boldsymbol{x}(\mu)=\boldsymbol{x}^*$.

注意： 上述例子中$\boldsymbol{x}(\mu)$关于μ连续可微，因而$\boldsymbol{x}(\mu)$在问题的可行域内部定义了一条以μ 为参数的光滑曲线(通常称为障碍轨迹Barrier Trajectory), 这条曲线在$\mu\to 0$时的极限点即为问题的最优解。

对数障碍函数数值困难

对数障碍函数的Hesse矩阵：

$$\begin{aligned}\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathrm{B_L}(\boldsymbol{x}_k;\mu) &= \nabla^2 f(\boldsymbol{x}_k) - \sum_{i\in\mathcal{I}} \frac{\mu_k}{c_i(\boldsymbol{x}_k)} \nabla^2 c_i(\boldsymbol{x}_k) \\ &\quad + \sum_{i\in\mathcal{I}} \frac{\mu_k}{c_i^2(\boldsymbol{x}_k)} \nabla c_i(\boldsymbol{x}_k) \nabla c_i(\boldsymbol{x}_k)^T.\end{aligned} \tag{6.3.9}$$

令

$$\lambda_i^{(k)} = \frac{\mu_k}{c_i(\boldsymbol{x}_k)}, \quad i \in \mathcal{I},$$

并将其代入(6.3.9)式，利用Lagrange函数，有:

$$\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathrm{B_L}(\boldsymbol{x}_k;\mu_k) = \nabla^2_{\boldsymbol{x}\boldsymbol{x}} L(\boldsymbol{x}_k, \lambda^{(k)}) + \sum_{i\in\mathcal{I}} \frac{1}{\mu_k} (\lambda_i^{(k)})^2 \nabla c_i(\boldsymbol{x}_k) \nabla c_i(\boldsymbol{x}_k)^T. \tag{6.3.10}$$

当$\mu_k \to 0$时，$\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathrm{B_L}(\boldsymbol{x}_k;\mu_k)$的条件数与$\nabla^2_{\boldsymbol{x}\boldsymbol{x}} \mathrm{P_E}(\boldsymbol{x}_k;\sigma_k)$情形相同。

障碍罚函数等高线

例6.3.4 考虑等式约束优化问题

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

该问题的最优解为：

$$\boldsymbol{x}^* = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

对数障碍函数为:

$$\mathrm{B_L}(\boldsymbol{x}, \mu) = x_1^2 + x_2^2 + \mu \log(-x_1 + x_2 - 1)$$

障碍罚函数等高线