第26-27讲 增广目标函数法

约束优化问题

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \geq 0,\ i \in \mathcal{I} = \{1, \ldots, m_1\}, \\
& h_j(x) = 0,\ j \in \mathcal{E} = \{m_1 + 1, \ldots, m\}.
\end{aligned}
$$

$$g_{\mathcal{I}}(x) = (g_1(x), \ldots, g_{m_1}(x))^T, \quad h_{\mathcal{E}}(x) = (h_{m_1+1}(x), \ldots, h_m(x))^T.$$

$$
\Longleftrightarrow \qquad
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & g_{\mathcal{I}}(x) \geq 0, \\
& h_{\mathcal{E}}(x) = 0.
\end{aligned}
\qquad (12.1)
$$

可行域为

$$D = \{x \mid g_{\mathcal{I}}(x) \geq 0,\ h_{\mathcal{E}}(x) = 0\}.$$

罚函数法

外点罚函数法

基本思想: 构造辅助函数 $F_\mu : R^n \to R$, 其中参数 $\mu > 0$.

函数 F_μ 在可行域 D 的 内部的取值与问题 (12.1) 的目标函数 f 的取值相等, 对可行域外部点的目标函数值加以惩罚, $\min F_\mu(x)$ 的解是(12.1) 的近似解.

当 $\mu \to \infty$ 时, 问题 $\min F_\mu(x)$ 的解 $x(\mu)$ 趋于约束问题 (12.1) 的解.

例

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

令 $S(x) = \min{}^2\{g(x), 0\}.$

则 $S(x) = 0 \iff x \in D.$

令 $P_\mu(x) = \frac{1}{2}\mu S(x) = \frac{1}{2}\mu \min{}^2\{g(x), 0\}.$

构造辅助函数 F_μ

$$F_\mu(x) = f(x) + P_\mu(x) = f(x) + \frac{1}{2}\mu \min{}^2\{g(x), 0\}.$$

则

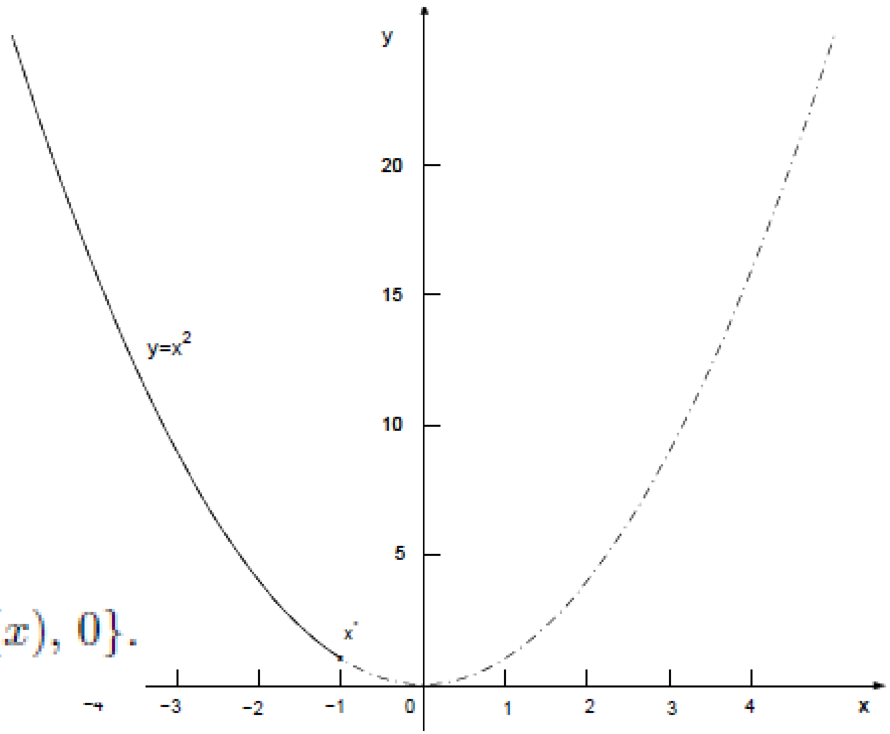

图 12.1： 问题 (12.2) 的可行域和最优解.

$D = (-\infty, -1]$ 和 $x^* = -1.$

无约束问题 $\min F_\mu(x)$ 的解为 $x(\mu) = -\dfrac{\mu}{2+\mu}$.

当 $\mu \to +\infty$ 时， $x(\mu) \to -1 = x^*$.

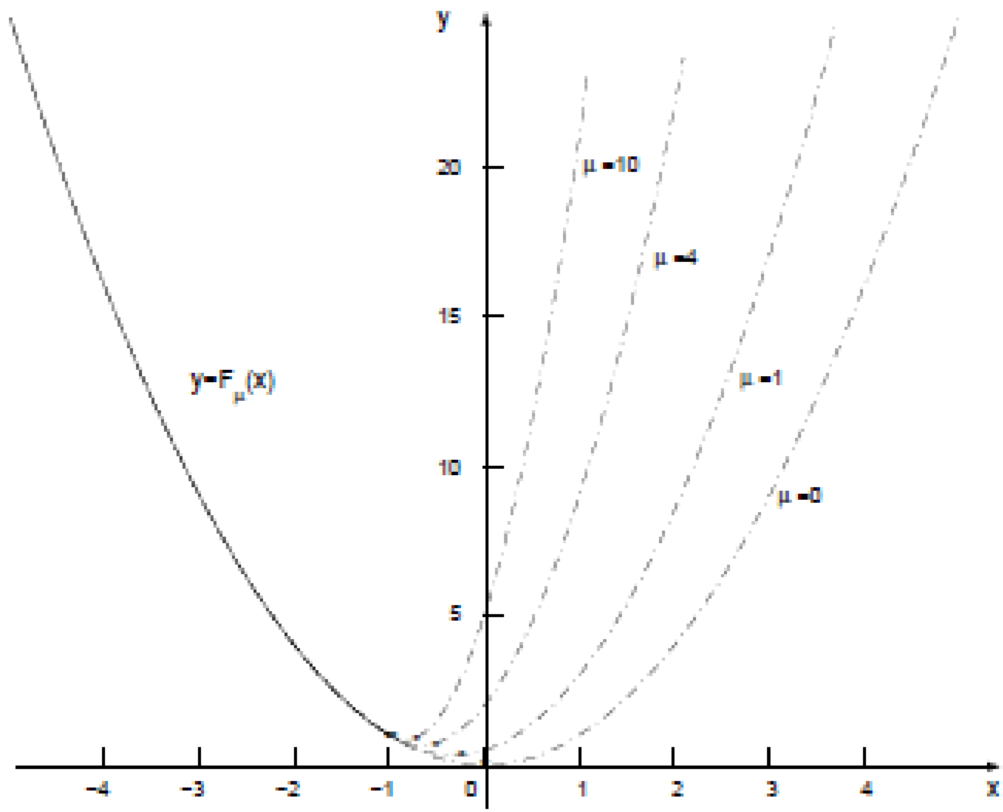

图 12.2： 增广目标函数 $F_\mu(x)$ 及其极小值点.

对一般的约束问题 (12.1), 我们引入如下函数:

$$S(x) = \|h_{\mathcal{E}}(x)\|^2 + \|\min\{g_{\mathcal{I}}(x), 0\}\|^2,$$

$$P_\mu(x) = \frac{1}{2}\mu S(x), \qquad F_\mu(x) = f(x) + P_\mu(x).$$

显然有 $F_\mu(x) = f(x), \forall x \in D$. 而且，当 $\mu > 0$ 很大时，对任何 $x \notin D$, 有 $F_\mu(x) \gg f(x)$.

算法 12.1 *(外点罚函数法)*

步 *0* 取初始点 $x^{(0)} \in R^n$. 罚参数序列 $\{\mu_k\}$, 精度 $\epsilon > 0$. 令 $k := 0$.

步 *1* 构造增广目标函数

$$F_{\mu_k}(x) = f(x) + P_{\mu_k}(x) = f(x) + \frac{1}{2}\mu_k\Big(\|h_{\mathcal{E}}(x)\|^2 + \|\min\{g_{\mathcal{I}}(x), 0\}\|^2\Big).$$

步 *2* 求解无约束问题

$$\min F_{\mu_k}(x), \quad x \in R^n$$

得解 $x^{(k)}$.

步 *3* 若 $P_{\mu_k}(x^{(k)}) \le \epsilon$ 或

$$-\min\{g_i(x^{(k)}) \,\big|\, i \in \mathcal{I}\} \le \epsilon, \quad \max\{|h_j(x^{(k)})| \,\big|\, j \in \mathcal{E}\} \le \epsilon,$$

则得解 $x^{(k)}$, 否则, 令 $k := k + 1$. 转步 *1*.

定理 12.1.1 设函数 f, g_i, h_j, $i \in \mathcal{I}$, $j \in \mathcal{E}$ 连续且约束问题 (12.1) 的解存在. 设 $\{x^{(k)}\}$ 由算法 *12.1* 产生, 其中罚参数序列 $\{\mu_k\}$ 单调递增且趋于 $+\infty$. 则 $\{x^{(k)}\}$ 的任何极限点都是问题 (12.1) 的解.

例 12.1.1 用外点罚函数法求解约束问题

$$\begin{aligned} \min \quad & f(x) = \frac{1}{2}x_1^2 + \frac{1}{6}x_2^2 \\ \text{s.t.} \quad & h(x) = x_1 + x_2 - 1 = 0. \end{aligned}$$

取罚参数序列为 $\mu_k = 2^{k-1}$.

解 构造增广目标函数 F_μ 如下:

$$F_\mu(x) = f(x) + \frac{1}{2}\mu h^2(x) = \frac{1}{2}x_1^2 + \frac{1}{6}x_2^2 + \frac{1}{2}\mu(x_1 + x_2 - 1)^2.$$

由计算得 $\nabla F_\mu(x) = (x_1 + \mu(x_1 + x_2 - 1),\ \frac{1}{3}x_2 + \mu(x_1 + x_2 - 1))^T.$

F_μ 是二次凸函数

$$\min F_\mu(x), \quad x \in R^n \quad \Longleftrightarrow \quad \nabla F_\mu(x) = 0$$

$$\Longleftrightarrow \quad \begin{cases} x_1 + \mu(x_1 + x_2 - 1) & = & 0, \\ \frac{1}{3}x_2 + \mu(x_1 + x_2 - 1) & = & 0 \end{cases}$$

$$\Longrightarrow \quad x_\mu = \left(\frac{\mu}{1+4\mu}, \frac{3\mu}{1+4\mu}\right)^T. \quad \Longrightarrow \quad x^{(k)} = \left(\frac{2^{k-1}}{1+2^{k+1}}, \frac{3 \times 2^{k-1}}{1+2^{k+1}}\right)^T.$$

$$\Longrightarrow \quad x^{(k)} \to x^* = (1/4,\ 3/4)^T.$$

结论说明

约束优化问题有最优解，不能保证罚函数有最优解.

A、罚因子取值太小, 目标函数的下降度不能抵消对不可行点的惩罚度.

最优解(1,0)

例

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

$$\text{s.t. } x_1 = 1$$

$$P(\boldsymbol{x},\pi) = -x_1^2 + x_2^2 + \pi(x_1 - 1)^2$$

在 $\pi < 1$ 时没有最优解。

目标函数与惩罚项同阶，罚因子太小，目标函数主导罚函数的值

B、目标函数阶数高于惩罚项阶数，目标函数的下降速度远高于惩罚项的下降速度

最优解(1;0)

例 $$\begin{aligned} &\min -x_1^5 \\ &\text{s.t. } x_1^2 + x_2^2 = 1 \end{aligned}$$

$P(\boldsymbol{x}, \pi) = -x_1^5 + \pi(x_1^2 + x_2^2 - 1)^2$ 对任意 $\pi > 0$ 均无最优解。

目标函数阶数高于惩罚项阶数，目标函数主导罚函数

算法缺陷

随罚因子增大，子问题病态加剧，从而导致算法稳定性差.

原问题 $\min\{f(\boldsymbol{x}) \mid c_i(\boldsymbol{x}) = 0, i \in \mathcal{E}\}$

外点罚函数 $P(\boldsymbol{x},\pi) = f(\boldsymbol{x}) + \pi \sum_{i\in\mathcal{E}} c_i^2(\boldsymbol{x})$

Hesse阵 $\nabla_{\boldsymbol{x}\boldsymbol{x}} P(\boldsymbol{x},\pi) = \nabla^2 f(\boldsymbol{x}) + \sum_{i\in\mathcal{E}} 2\pi c_i(\boldsymbol{x}) \nabla^2 c_i(\boldsymbol{x}) + 2\pi \boldsymbol{N}(\boldsymbol{x}) \boldsymbol{N}(\boldsymbol{x})^{\mathrm{T}}$

$\lambda_i = -2\pi c_i(x)$

$$\nabla_{xx} P(x,\pi) \approx \nabla_{xx} L(x,\lambda) + 2\pi N(x) N(x)^T$$

矩阵$\boldsymbol{N}(\boldsymbol{x})\boldsymbol{N}(\boldsymbol{x})^T$奇异

罚函数条件数 $\boldsymbol{\kappa}\left(\boldsymbol{\nabla}_{xx}\boldsymbol{P}(\boldsymbol{x},\boldsymbol{\pi})\right) \rightarrow \infty, \boldsymbol{\pi} \rightarrow \infty$ 子问题病态！

外点罚函数病态实例

例：考虑约束优化问题

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

分析：外点罚函数

$$\mathrm{F}(x,\pi) = x_1^2 + x_2^2 + \pi(x_1+1)^2$$

Hesse矩阵

$$\nabla_{xx}\mathrm{F}(x,\pi) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 2\pi & 0 \\ 0 & 0 \end{pmatrix}$$

Hesse矩阵条件数

$$\kappa(\nabla_{xx}\mathrm{F}(x,\pi)) = \frac{2+2\pi}{2} \to \infty, \quad (\pi \to \infty)$$

内点罚函数法

考察约束问题

$$\begin{array}{ll} \min & f(x) \\ \text{s.t.} & g_i(x) \geq 0, \quad i \in \mathcal{I}. \end{array}$$

内点罚函数法的基本思想是：构造辅助（光滑）函数 F_μ，该函数在严格可行域 D_0 以外的取值为无穷大，而且，当点 x 从 D_0 趋于 D 的边界时，函数值趋于无穷大．这样，无约束问题 $\min F_\mu(x)$ 的解一定在 D_0 内．我们希望当 μ 趋于无穷大时，无约束问题 $\min F_\mu(x)$ 的解 $x(\mu)$ 趋于约束问题 (12.6) 的解．辅助函数 F_μ 在 D 的边界筑起了一道很高的墙，把无约束问题 $\min F_\mu(x)$ 的解挡在了可行域 D 的内部．因此，内点罚函数法也称为障碍函数法．

$$F_\mu(x) = f(x) - \mu^{-1} \sum_{i \in \mathcal{I}} \log g_i(x).$$

称函数 $S(x) = -\sum_{i \in \mathcal{I}} \log g_i(x)$ 为对数障碍函数．

例

$$\begin{array}{ll} \min & f(x) = x^2 \\ \text{s.t.} & g(x) = -x - 1 \geq 0. \end{array}$$

$$F_\mu(x) = x^2 - \mu^{-1} \log[-(x+1)],$$

相应的无约束问题 $\min F_\mu(x), \quad x \in R$ 的解为

$$x(\mu) = -\frac{1+\sqrt{1+2\mu^{-1}}}{2} \longrightarrow -1 = x^*,\ \mu \to \infty.$$

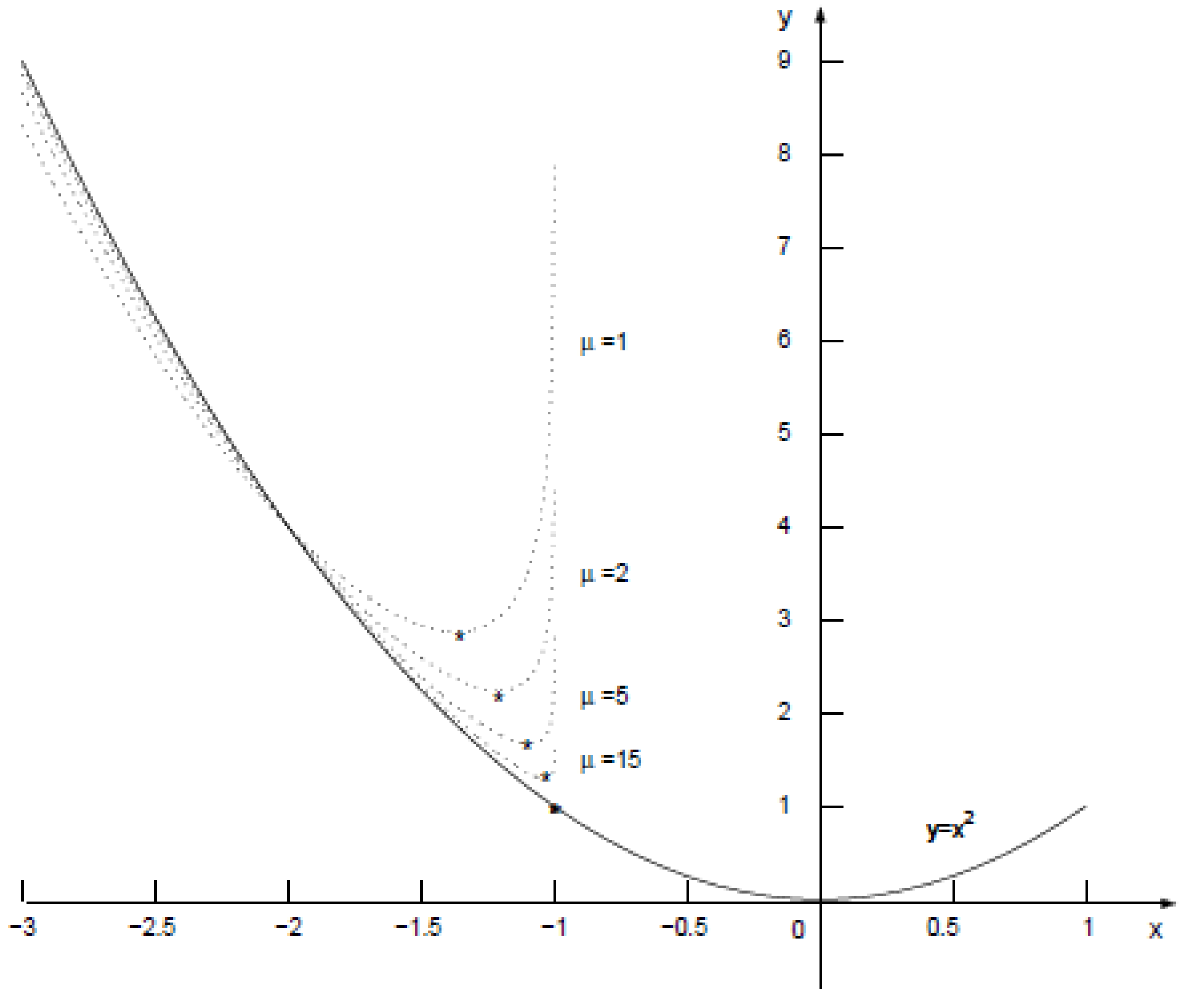

图 12.3： 辅助目标函数 $F_\mu(x)$ 及其极小值点.

算法 12.2 *(内点罚函数法)*

步 *0* 取初始点 $x^{(0)} \in D_0$. 罚参数序列 $\{\mu_k\}$, 精度 $\epsilon > 0$. 令 $k := 0$.

步 *1* 构造辅助函数

$$F_{\mu_k}(x) = f(x) + \mu_k^{-1} S(x) = f(x) - \mu_k^{-1} \sum_{i \in \mathcal{I}} \log g_i(x).$$

步 *2* 以 $x^{(k-1)}$ 作为初始点, 求解无约束问题 $\min F_{\mu_k}(x)$, $x \in R^n$ 得解 $x^{(k)}$.

步 *3* 若

$$-\mu_k^{-1} \sum_{i \in \mathcal{I}} \log g_i(x) \leq \epsilon,$$

则得解 $x^{(k)}$. 否则, 令 $k := k + 1$. 转步 *1*.

定理 12.1.2 设函数 $f, g_i, i \in \mathcal{I}$ 连续可微, D_0 非空且其闭包为 D. 再设问题 (12.6) 有解, $\{x^{(k)}\}$ 由算法 *12.2* 产生, 其中罚参数序列 $\{\mu_k\}$ 单调递增且趋于 $+\infty$. 则 $\{x^{(k)}\}$ 的任何极限点都是问题 (12.6) 的解.

内点罚函数方法的收敛性

对数障碍函数的最优值点列的聚点为原问题的K-T点

一阶最优性条件：

$$\nabla_x \mathrm{F}(x_\pi, \pi) = \nabla f(x_\pi) - \sum_{i \in I} \frac{1}{\pi c_i(x_\pi)} \nabla c_i(x_\pi) = 0$$

$$\lambda_i(\pi) := \frac{1}{\pi c_i(x_\pi)}, i \in I$$

$$\nabla_x L(x_\pi, \lambda(\pi)) = \nabla f(x_\pi) - \sum_{i \in I} \lambda_i(\pi) \nabla c_i(x_\pi) = 0$$

KKT条件第一式

对任意$\boldsymbol{i} \in \boldsymbol{I}$，$\boldsymbol{\lambda_i(\pi) c_i(x_\pi) = \frac{1}{\pi} \to 0, \quad \pi \to \infty}$

互补松弛条件

罚函数法存在的问题:

$x^{(k)}$ 是问题 (12.8) 的解 x^* 的近似， μ_k 必须充分大.

当 μ_k 很大时，函数 F_{μ_k} 的 Hessian 阵会出现病态.

为了克服由于罚参数过大而引起的数值计算上的困难，我们构造函数 $L_\mu(x)$，使得当罚参数 μ 相对较小时，无约束问题 $\min L_\mu(x)$ 的解 $x(\mu)$ 也是约束问题 (12.8) 的解的一个很好的近似.

Hestenes和Powell(1969)将Lagrange函数与外点罚函数相结合，建立了等式约束优化问题的增广Lagrange罚函数(乘子罚函数)。

Rockafeller (1973)将其推广到不等式约束优化问题, 建立了约束优化问题的乘子罚函数。

乘子法

等式约束问题的乘子法

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_i(x) = 0,\ i \in \mathcal{E} = \{m_1+1, \dots, m\} \end{aligned}$$

增广 Lagrange 函数

$$L_\mu(x,\lambda) = L(x,\lambda) + \frac{1}{2}\mu S(x) = f(x) - \sum_{i\in\mathcal{E}} \lambda_i h_i(x) + \frac{1}{2}\mu \sum_{i\in\mathcal{E}} h_i^2(x).$$

迭代格式

给定μ_{k}和$\lambda^{(\mathrm{k})}$,

$$x^{(k)} = \operatorname{argmin} \mathrm{L}_{\mu_{\mathrm{k}}}(x, \lambda^{(k)})$$

$$\lambda_i^{(k+1)} = \lambda_i^{(k)} - \mu_k h_i(x^{(k)}), i \in E$$

定理 12.2.1 设 $x^{(k)}$ 是无约束问题

$$\min L_{\mu_k}(x, \lambda^{(k)}), \qquad x \in R^n$$

的解. 则 $x^{(k)}$ 也是约束问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & h_i(x) = h_i(x^{(k)}),\ i \in \mathcal{E} \end{aligned}$$

的解.

定理 12.2.2 设 x^* 是问题 (12.8) 的一个局部最优解且 *LICQ* 在 x^* 处成立, 即 $\nabla h_j(x^*)$, $j \in \mathcal{E}$ 线性无关. 再设在 x^* 处二阶充分条件成立. 则存在 $\bar{\mu} > 0$, 使得对所有 $\mu \geq \bar{\mu}$, x^* 是无约束问题

$$\min L_\mu(x, \lambda^*) = f(x) - \sum_{i \in \mathcal{E}} \lambda_i^* h_i(x) + \frac{1}{2}\mu \sum_{i \in \mathcal{E}} h_i^2(x) \tag{12.15}$$

的严格局部最优解, 其中 λ^* 为解 x^* 处的 *Lagrange* 乘子.

增广拉格朗日法本质上是: 近似牛顿法

结论意义：无需罚因子趋于无穷,由乘子罚函数可得原问题最优解.

缺陷：无法保证Lagrange乘子的取值！

解决途径：逐步调整

对给定的$\boldsymbol{\lambda}^k, \pi_k$, 计算乘子罚函数$P(\boldsymbol{x}, \boldsymbol{\lambda}^k, \pi_k)$值点得$\boldsymbol{x}_k$

$$\nabla f(\boldsymbol{x}_k) - \sum_{i\in\mathcal{E}} \lambda_i^k \nabla c_i(\boldsymbol{x}_k) + 2\pi_k \sum_{i\in\mathcal{E}} c_i(\boldsymbol{x}_k)\nabla c_i(\boldsymbol{x}_k) = \mathbf{0}$$

$$\Downarrow$$

$$\nabla f(\boldsymbol{x}_k) = \sum_{i\in\mathcal{E}} \left(\lambda_i^k - 2\pi_k c_i(\boldsymbol{x}_k)\right)\nabla c_i(\boldsymbol{x}_k) \xrightarrow{\text{KKT条件}} \lambda_i^{k+1} = \lambda_i^k - 2\pi_k c_i(\boldsymbol{x}_k)$$

若$\boldsymbol{x}_k$ 为乘子罚函数最优解,且为可行点,则为最优解和K-T点,算法终止

终止条件：$\|\boldsymbol{c}(\boldsymbol{x}_k)\|_\infty \leqslant \varepsilon$

初始步

步1、取 $\pi_1 > 0$, $\boldsymbol{\lambda}^1 = \mathbf{0}$, $\boldsymbol{x}_0 \in \mathbb{R}^n$, $\gamma > 1$, $\varepsilon > 0$. 令 $k = 1$.

迭代步

步2、以 $\boldsymbol{x}_{k-1}$ 为初始点， 求 $P(\boldsymbol{x}, \boldsymbol{\lambda}^k, \pi_k)$ 的最小值点得 $\boldsymbol{x}_k$.

步3、若 $\max\{|c_i(\boldsymbol{x}_k)| \mid i \in \mathcal{E}\} \leqslant \varepsilon$, 算法终止. 否则， 转下一步.

校正步

罚因子

乘 子

步4、若 $\|\boldsymbol{c}(\boldsymbol{x}_k)\|_\infty \geqslant \|\boldsymbol{c}(\boldsymbol{x}_{k-1})\|_\infty$, 令 $\pi_{k+1} = \gamma\pi_k$, $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k$,

置 $k = k + 1$, 转步2. 否则， 转步5.

步5、若 $\pi_k > \pi_{k-1}$ 或 $\|\boldsymbol{c}(\boldsymbol{x}_k)\|_\infty \leqslant \frac{1}{4}\|\boldsymbol{c}(\boldsymbol{x}_{k-1})\|_\infty$, 令 $\pi_{k+1} = \pi_k$,

调整Lagrange乘子： $\lambda_i^{k+1} = \lambda_i^k - 2\pi_k c_i(\boldsymbol{x}_k)$

置 $k = k + 1$, 转步2； 否则， 令 $\boldsymbol{\lambda}^{k+1} = \boldsymbol{\lambda}^k$, $\pi_{k+1} = \gamma\pi_k$,

置 $k = k + 1$, 转步2.

不等式约束优化

$$\min\ f(x)\quad \text{s.t. } g_i(x)\ge 0,\ i\in I$$

$$\min\ f(x)\quad \text{s.t. } g_i(x)-z_i^{\,2}=0,\ i\in I$$

$$L_\mu(x,z,\lambda)=f(x)-\sum_{i\in I}\lambda_i(g_i(x)-z_i^2)+\frac{\mu}{2}\sum_{i\in I}(g_i(x)-z_i^2)^2$$

$$\min_x\min_{s\ge 0} L_\mu(x,s,\lambda)=f(x)-\sum_{i\in I}\lambda_i(g_i(x)-s_i)+\frac{\mu}{2}\sum_{i\in I}(g_i(x)-s_i)^2$$

$$\min_{s\ge 0} L_\mu(x,s,\lambda)=f(x)-\sum_{i\in I}\lambda_i(g_i(x)-s_i)+\frac{\mu}{2}\sum_{i\in I}(g_i(x)-s_i)^2$$

$$s_i=\max\{0,g_i(x)-\frac{1}{\mu}\lambda_i\}$$

不等式约束问题

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \geq 0,\ i \in \mathcal{I} = \{1, 2, \ldots, m_1\}, \end{aligned}$$

$$\begin{aligned} \bar{g}_i(x, z(x)) &= g_i(x) - z_i^2(x) = g_i(x) - \mu^{-1} \max\{0, \mu g_i(x) - \lambda_i\} \\ &= \mu^{-1}\Big(\min\{\mu g_i(x),\ \lambda_i\}\Big) = \mu^{-1}\Big(\min\{\mu g_i(x) - \lambda_i,\ 0\} + \lambda_i\Big). \end{aligned} \tag{12.23}$$

增广 Lagrange 函数为

$$\begin{aligned} L_\mu(x, \lambda) &= f(x) - \sum_{i\in\mathcal{I}} \lambda_i \bar{g}_i(x, z(x)) + \frac{1}{2}\mu \sum_{i\in\mathcal{I}} \bar{g}_i^2(x, z(x)) \\ &= f(x) - \mu^{-1} \sum_{i\in\mathcal{I}} \lambda_i \Big(\min\{\mu g_i(x) - \lambda_i,\ 0\} + \lambda_i\Big) + \frac{1}{2}\mu^{-1} \sum_{i\in\mathcal{I}} \Big(\min\{\mu g_i(x) - \lambda_i,\ 0\} + \lambda_i\Big)^2 \\ &= f(x) + \frac{1}{2}\mu^{-1} \sum_{i\in\mathcal{I}} \Big(\min{}^2\{\mu g_i(x) - \lambda_i,\ 0\} - \lambda_i^2\Big). \end{aligned}$$

迭代格式

给定μ_k和$\lambda^{(k)}$,

$$x^{(k)} = \underset{x}{\operatorname{argmin}}\ L_{\mu_k}(x, \lambda^{(k)})$$

$$\lambda^{(k+1)} = \underset{\lambda \geq 0}{\arg\max}\, L_{\mu_k}(x^{(k)}, \lambda),$$

$$\max_{\lambda\ge 0} L_\mu(x^k,\lambda) \Leftrightarrow \min_{\lambda\ge 0} \ -L_\mu(x^k,\lambda)$$

$$\boxed{L_\mu(x,\lambda)= f(x)-\sum_{i\in I}\lambda_i g_i(x)+\frac{\mu}{2}\min{}^2\{g_i(x),0\}}$$

$$\nabla_\lambda[-L_\mu(x^k,\lambda)]=g(x^k)$$

$$\boxed{\begin{aligned}\min_{\lambda\ge 0} \ -L_\mu(x^k,\lambda) &\Leftrightarrow g(x^k)^T(\lambda-\lambda^*)\ge 0,\forall\lambda\ge 0\\ &\Leftrightarrow \lambda^*=\mathrm{P}_{\lambda\ge 0}(\lambda^*-\rho g(x^k))\\ &= \max\{\lambda^*-\rho g(x^k),0\}\end{aligned}}$$

特别地，取$\rho=\mu_k$，有

$$\lambda^{(k+1)}=\max\{\lambda^k-\mu_k g(x^k),0\}$$

一般约束问题

$$
\begin{aligned}
\min \quad & f(x) \\
\text{s.t.} \quad & g_i(x) \geq 0, \ i \in \mathcal{I}, \\
& h_j(x) = 0, \ j \in \mathcal{E}
\end{aligned}
$$

增广 Lagrange 函数为

$$
L_\mu(x, \lambda) = f(x) + \frac{1}{2}\mu^{-1} \sum_{i \in \mathcal{I}} \left(\min{}^2\{\mu g_i(x) - \lambda_i, 0\} - \lambda_i^2 \right) - \sum_{j \in \mathcal{E}} \lambda_j h_j(x) + \frac{1}{2}\mu \sum_{j \in \mathcal{E}} h_j^2(x).
$$

相应的乘子迭代格式为：

$$
\lambda_j^+ = \begin{cases} \lambda_j - \mu h_j(x), & j \in \mathcal{E}, \\ \max\{\lambda_i - \mu g_i(x), 0\}, & i \in \mathcal{I}, \end{cases} \tag{12.28}
$$

其中， x, μ, λ 表示当前迭代点的值， λ^+ 表示下一次迭代的乘子向量.

算法 12.4（乘子法）

步 *0* 取初始点 $x^{(0)} \in R^n$, 初始乘子向量 $\lambda^{(0)}$. 给定罚参数序列 $\{\mu_k\}$, 精度 $\epsilon > 0$. 令 $k := 0$.

步 *1* 由 (12.27) 构造增广 *Lagrange* 函数 $L_\mu(x, \lambda)$.

步 *2* 以 $x^{(k-1)}$ 作为初始点 *(*$k=0$ 时，初始点任意*)*，求解无约束问题

$$\min L_{\mu_k}(x, \lambda^{(k)}), \quad x \in R^n$$

得解 $x^{(k)}$.

步 *3*（终止准则）若

$$\|h_{\mathcal{E}}(x^{(k)})\| + \|\min\{g_{\mathcal{I}}(x^{(k)}),\ \mu_k^{-1}\lambda_{\mathcal{I}}^{(k)}\}\| \le \epsilon,$$

则得解 $x^{(k)}$.

步 *4*（进行乘子迭代）：由 (12.28) 确定 $\lambda^{(k+1)}$. 令 $k := k+1$. 转步 *1*.

例 12.2.2 用算法 *12.4* 求解下面的约束问题:

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

取 $\mu_k = 4,\ \lambda_0 = 0.$

解 问题的增广 Lagrange 函数为

$$\begin{aligned} L_\mu(x,\lambda) &= x_1^2 + x_2^2 + \frac{1}{2}\mu^{-1}\left[\min{}^2\{\mu(x_1-1)-\lambda,\ 0\} - \lambda^2\right] \\ &= \begin{cases} x_1^2 + x_2^2 + \frac{1}{2}\mu(x_1-1)^2 - \lambda(x_1-1), & \text{若 } x_1 \leq 1 + \mu^{-1}\lambda, \\ x_1^2 + x_2^2 - \frac{1}{2}\mu^{-1}\lambda^2, & \text{若 } x_1 > 1 + \mu^{-1}\lambda. \end{cases} \end{aligned}$$

$$\Longrightarrow \quad \frac{\partial L_\mu(x,\lambda)}{\partial x_1} = \begin{cases} 2x_1 + \mu(x_1-1) - \lambda, & \text{若 } x_1 \leq 1 + \mu^{-1}\lambda, \\ 2x_1, & \text{若 } x_1 > 1 + \mu^{-1}\lambda. \end{cases}$$

$$\frac{\partial L_\mu(x,\lambda)}{\partial x_2} = 2x_2.$$

由 $\nabla_x L_\mu(x,\lambda)=0$ 得无约束问题 $\min\limits_x L_\mu(x,\lambda_k)$, $x\in R^n$ 的极小值点

$$x^{(k)}=\Big(\frac{\lambda_k+\mu_k}{2+\mu_k},\ 0\Big)^T=(\frac{\lambda_k+4}{6},\ 0)^T.$$

乘子满足:

$$\lambda_{k+1}=\max\{\lambda_k-4(x_1^{(k)}-1),\ 0\}=\max\{\frac{\lambda_k+4}{3},\ 0\}=\frac{\lambda_k+4}{3}.$$

因此, $\lambda_k\to 2,\ x^{(k)}\to(1,0)^T,\ k\to\infty.$

可以看出，用乘子法产生的点列收敛于问题的解. 而且，罚参数 μ_k 不必趋于 $+\infty$.

作业

- 10(1) 内点和外点罚函数法分别求解
- 20(ii)
- 16(思考题，不做为作业)