第13章　惩罚函数法

SUMT(Sequential unconstrained minimization technique)
序列无约束最小化技术

$$\text{罚函数法}\begin{cases}\text{外点法：迭代点在可行域外部移动}\\ \text{内点法：迭代点在可行域内部移动}\end{cases}$$

外点法

$$
(A)\quad \begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \geq 0\ i=1,2,\cdots,m \\ \qquad\quad h_j(x)=0\ \ j=1,2,\cdots,l \end{cases}
$$

其中$f(x), g_i(x)(i=1,2,\cdots,m), h_j(x)(j=1,2,\cdots,l)$

在E^n上连续。

$$
S=\{x \mid g_i(x) \geq 0(i=1,2,\cdots,m), h_j(x)=0(j=1,2,\cdots,l)\}
$$

引入罚项

$$p(x)=\sum_{i=1}^{m}\varphi[g_i(x)]+\sum_{j=1}^{l}\psi[h_j(x)]$$

其中$\varphi(y),\psi(y)$是连续函数，且满足

$$\begin{cases}\varphi(y)=0 & \text{当} y\geq 0\\ \varphi(y)>0 & \text{当} y<0\end{cases}$$

$$\begin{cases}\psi(y)=0 & \text{当} y=0\\ \psi(y)>0 & \text{当} y\neq 0\end{cases}$$

函数φ和ψ的典型取法：

$$\varphi[g_i(x)]=[\max\{0,-g_i(x)\}]^{\alpha}\qquad \psi[h_j(x)]=|h_j(x)|^{\beta}$$

其中$\alpha\geq 1,\beta\geq 1$均为给定常数，通常取$\alpha=\beta=2$。

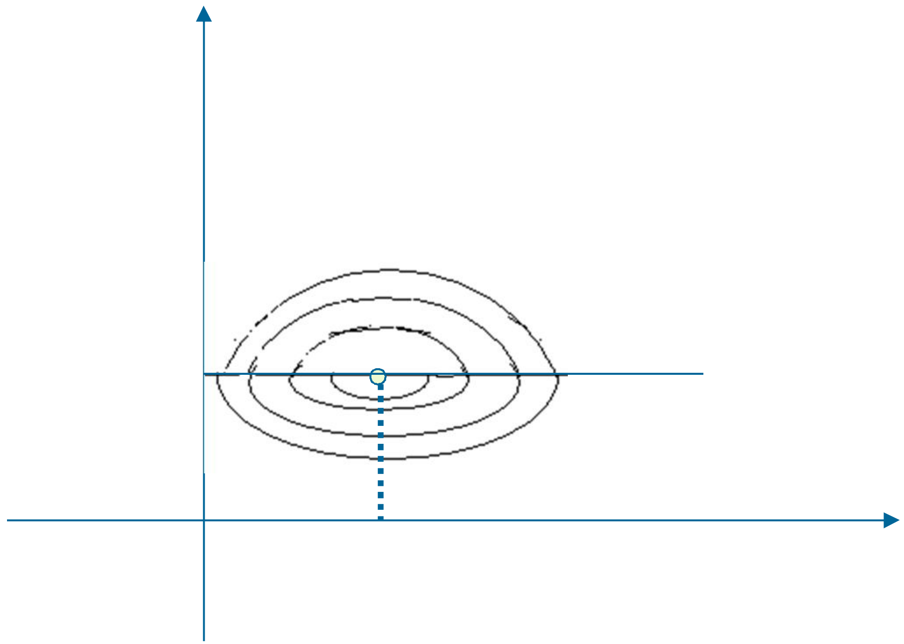
$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ i = 1,2,\cdots,m \\ \qquad\quad h_j(x) = 0 \ j = 1,2,\cdots,l \end{cases}$$

$$\Longleftrightarrow \qquad \min \ F(x,\sigma) = f(x) + \sigma p(x)$$

$$\min \ F(x,\sigma) = f(x) + \sigma\left[\sum_{i=1}^{m} \varphi[g_i(x)] + \sum_{j=1}^{l} \psi[h_j(x)]\right]$$

其中σ是很大的正数。

$$\min\ f(x) = (x_1 - 1)^2 + x_2^2$$
$$s.t.\quad g(x) = x_2 - 1 \geq 0$$



解：定义罚函数

$$F(x,\sigma) = (x_1 - 1)^2 + x_2^2 + \sigma\left[\max\{0, -(x_2 - 1)\}\right]^2$$
$$= \begin{cases} (x_1 - 1)^2 + x_2^2 & x_2 \geq 1 \\ (x_1 - 1)^2 + x_2^2 + \sigma(x_2 - 1)^2 & x_2 < 1 \end{cases}$$

$$\frac{\partial F}{\partial x_1} = 2(x_1 - 1) \quad \frac{\partial F}{\partial x_2} = \begin{cases} 2x_2 & x_2 \geq 1 \\ 2x_2 + 2\sigma(x_2 - 1) & x_2 < 1 \end{cases}$$

令 $\frac{\partial F}{\partial x_1} = 0 \quad \frac{\partial F}{\partial x_2} = 0$ 得

$$\overline{x}_\sigma = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{\sigma}{1+\sigma} \end{pmatrix} \to \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad (\sigma \to +\infty)$$

$$\min\ x_1 + x_2$$

$$s.t.\quad x_1 - x_2^2 = 0$$

定义罚函数

$$F(x,\sigma) = x_1 + x_2 + \sigma(x_1 - x_2^2)^2$$

$$\nabla F(x,\sigma) = \begin{pmatrix} 1 + 2\sigma(x_1 - x_2^2) \\ 1 + 2\sigma(x_1 - x_2^2)(-2x_2) \end{pmatrix}$$

令$\nabla F(x,\sigma) = 0$，得

$$\bar{x}_\sigma = \left(\frac{1}{4} - \frac{1}{\sigma}, -\frac{1}{2}\right)^T \to \left(\frac{1}{4}, -\frac{1}{2}\right)^T (\sigma \to \infty)$$

$$\min f(x) = x_1 + x_2$$

$$s.t. \quad g_1(x) = -x_1^2 + x_2 \geq 0$$

$$g_2(x) = x_1 \geq 0$$

解：定义罚函数

$$F(x,\sigma) = x_1 + x_2 + \sigma\left[\max\left\{0, x_1^2 - x_2\right\}\right]^2 + \sigma\left[\max\left\{0, -x_1\right\}\right]^2$$

$$\frac{\partial F}{\partial x_1} = 1 + 2\sigma\left[2\max\left\{0, x_1^2 - x_2\right\}x_1 + \max\left\{0, -x_1\right\}(-1)\right]$$

$$\frac{\partial F}{\partial x_2} = 1 + 2\sigma\max\left\{0, x_1^2 - x_2\right\}(-1)$$

令 $\dfrac{\partial F}{\partial x_1}=0 \quad \dfrac{\partial F}{\partial x_2}=0$ 得

$$\begin{cases} 1+4\sigma x_1(x_1^2-x_2)+2\sigma x_1=0 \\ 1-2\sigma(x_1^2-x_2)=0 \end{cases}$$

$$\Rightarrow \quad x_1=-\frac{1}{2+2\sigma} \quad x_2=\frac{1}{(2+2\sigma)^2}-\frac{1}{2\sigma}$$

当$\sigma \to +\infty$时，有$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \to \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

步骤：

1. 给定初始点$x^{(0)}$，初始罚因子$\sigma_1 > 0(\sigma_1 = 1)$，放大系数$c > 1$，允许误差$\varepsilon > 0$，置$k = 1$。

2. 以$x^{(k-1)}$为初始点，求解无约束问题

$$\min f(x) + \sigma_k p(x)$$

设其极小点为$x^{(k)}$。

3. 若$\sigma_k p(x^{(k)}) < \varepsilon$，则停止计算，得到点$x^{(k)}$；否则，令$\sigma_{k+1} = c\sigma_k$，置$k := k+1$，返回2。

例:用外点法求解

$$\min\ (x-1)^2$$

$$s.t.\quad x-2\geq 0$$

解：取$x^{(0)}=0,\sigma_1=1$，令

$$p(x)=\left[\max\{0,-x+2\}\right]^2$$ 则

$$p(x)=\begin{cases}0 & x\geq 2\\ (-x+2)^2 & x<2\end{cases}$$

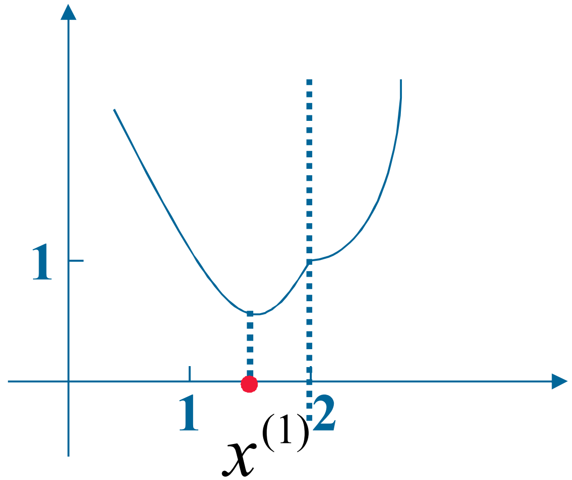
第一次迭代

求解无约束最优化问题：

$$\min F(x,\sigma_1)=(x-1)^2+1\times p(x)$$

其中$F(x,\sigma_1)=\begin{cases}(x-1)^2 & x\geq 2\\ (x-1)^2+1\times(-x+2)^2 & x<2\end{cases}$

解得： $x^{(1)}=\dfrac{3}{2}$



令$\sigma_2=10\times\sigma_1=10$

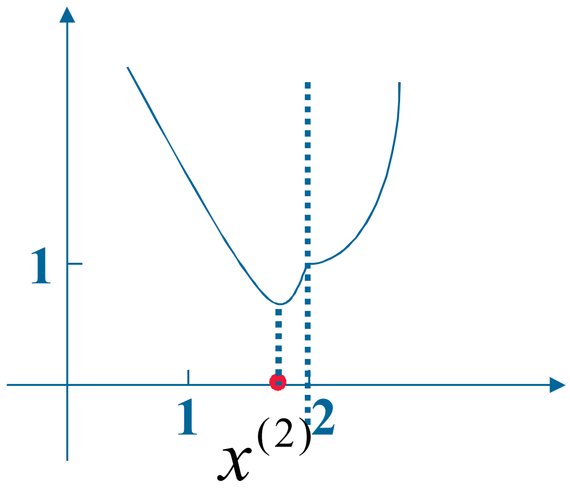
第二次迭代

求解无约束最优化问题：

$$\min F(x,\sigma_2)=(x-1)^2+10\times p(x)$$

$$\text{其中}F(x,\sigma_2)=\begin{cases}(x-1)^2 & x\ge 2\\(x-1)^2+10\times(-x+2)^2 & x<2\end{cases}$$

解得： $x^{(2)}=\dfrac{21}{11}$



$$\text{令}\sigma_3=10\times\sigma_2=100$$

第三次迭代

求解无约束最优化问题：

$$\min F(x,\sigma_3) = (x-1)^2 + 100 \times p(x)$$

其中$F(x,\sigma_3) = \begin{cases} (x-1)^2 & x \geq 2 \\ (x-1)^2 + 100 \times (-x+2)^2 & x < 2 \end{cases}$

解得： $x^{(3)} = \dfrac{201}{101}$

$$x^* = 2$$

以此类推，得序列：

$$\frac{3}{2}, \frac{21}{11}, \frac{201}{101}, \frac{2001}{1001}, \cdots$$

引理1 对于由外点法所产生的序列$\{x^{(k)}\}$，总有

(1) $F(x^{(k+1)},\sigma_{k+1}) \geq F(x^{(k)},\sigma_k)$

(2) $p(x^{(k+1)}) \leq p(x^{(k)})$

(3) $f(x^{(k+1)}) \geq f(x^{(k)})$

证明:(1)由$F(x,\sigma)=f(x)+\sigma p(x)$和$\sigma_{k+1}>\sigma_k$知

$$F(x^{(k+1)},\sigma_{k+1})=f(x^{(k+1)})+\sigma_{k+1}p(x^{(k+1)})$$

$$\geq f(x^{(k+1)})+\sigma_k p(x^{(k+1)})=F(x^{(k+1)},\sigma_k)$$

$\because x^{(k)}$是$F(x,\sigma_k)$的极小点，$\therefore$ 对$\forall x$，有$F(x,\sigma_k)\geq F(x^{(k)},\sigma_k)$

$$\Rightarrow F(x^{(k+1)},\sigma_k)\geq F(x^{(k)},\sigma_k)$$

$$\Rightarrow F(x^{(k+1)},\sigma_{k+1})\geq F(x^{(k)},\sigma_k)$$

(2) $x^{(k)}$和$x^{(k+1)}$分别使$F(x,\sigma_k),F(x,\sigma_{k+1})$取极小

$$\therefore f(x^{(k+1)})+\sigma_k p(x^{(k+1)})\geq f(x^{(k)})+\sigma_k p(x^{(k)}) \quad (*)$$

$$f(x^{(k)})+\sigma_{k+1} p(x^{(k)})\geq f(x^{(k+1)})+\sigma_{k+1} p(x^{(k+1)})$$

$$\sigma_k p(x^{(k+1)})+\sigma_{k+1} p(x^{(k)})\geq \sigma_k p(x^{(k)})+\sigma_{k+1} p(x^{(k+1)})$$

$$\Rightarrow \left(\sigma_{k+1}-\sigma_k\right)p(x^{(k)})\geq \left(\sigma_{k+1}-\sigma_k\right)p(x^{(k+1)})$$

$$\Rightarrow p(x^{(k)})\geq p(x^{(k+1)})$$

$$\boxed{\begin{array}{l}(2)\ p(x^{(k+1)})\leq p(x^{(k)})\\(3)\ f(x^{(k+1)})\geq f(x^{(k)})\end{array}}$$

(3) 由(*),得

$$f(x^{(k+1)})-f(x^{(k)})\geq \sigma_k\left(p(x^{(k)})-p(x^{(k+1)})\right)$$

$$\geq 0$$

引理2 设x^*是问题(A)的一个最优解,则对$\forall k$,有

$$f(x^*) \geq F(x^{(k)},\sigma_k) \geq f(x^{(k)}).$$

证明:因为x^*是问题(A)的最优解,所以有

$$p(x^*) = 0$$

$$\begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \geq 0 \\ \qquad\quad h_j(x) = 0 \end{cases}$$

$\because x^{(k)}$是$F(x,\sigma_k)$的极小点

$$\therefore f(x^*) = F(x^*,\sigma_k) \geq F(x^{(k)},\sigma_k)$$

又$\because \sigma_k p(x^{(k)}) \geq 0$

$$\therefore F(x^{(k)},\sigma_k) = f(x^{(k)}) + \sigma_k p(x^{(k)}) \geq f(x^{(k)})$$

定理 设$\left\{x^{(k)}\right\}$是由外点法产生的一个序列,则$\left\{x^{(k)}\right\}$的任何收敛子序列的极限都是问题(A)的最优解.设x^*是问题(A)的最优解，若$f(x), g_i(x), h_j(x)$具有连续的一阶偏导数，且在x^*处

$$\left\{\nabla g_i\left(x^*\right), i \in I(x^*), \nabla h_j\left(x^*\right), j=1,2,\cdots,l\right\}$$

线性无关，则

$$\lim_{k\to\infty} 2\sigma_k \max\left\{0, -g_i\left(x^{(k)}\right)\right\} = w_i, \quad i \in I(x^*)$$

$$\lim_{k\to\infty} 2\sigma_k h_j\left(x^{(k)}\right) = v_i, \quad j=1,2,\cdots,l.$$

证明:设$\left\{x^{(k_j)}\right\}$是$\left\{x^{(k)}\right\}$的任一收敛子序列,且有极限$\bar{x}$

$\because f(x)$在E^n上连续,$\therefore \lim\limits_{k_j\to\infty} f\left(x^{(k_j)}\right)=f(\bar{x})$

令$f^*=f(x^*)$为问题(A)的最优值

由引理1,2,知$\left\{F\left(x^{(k_j)},\sigma_{k_j}\right)\right\}$是非减的,且以$f^*$为上界的数列

$$\begin{cases}\min f(x)\\ s.t. \quad g_i(x)\geq 0\\ \qquad\quad h_j(x)=0\end{cases}$$

$\therefore \lim\limits_{k_j\to\infty} F\left(x^{(k_j)},\sigma_{k_j}\right)=F^*\leq f^*$

$\because F\left(x^{(k_j)},\sigma_{k_j}\right)=f\left(x^{(k_j)}\right)+\sigma_{k_j}p\left(x^{(k_j)}\right)$

$\therefore \lim\limits_{k_j\to\infty} \sigma_{k_j}p\left(x^{(k_j)}\right)=F^*-f(\bar{x})$

$\because\ p\left(x^{(k_j)}\right) \geq 0, \sigma_{k_j} \to \infty$

$\therefore \lim\limits_{k_j \to \infty} p\left(x^{(k_j)}\right) = 0$

又因为$p(x)$为连续函数

$\therefore \lim\limits_{k_j \to \infty} p\left(x^{(k_j)}\right) = p(\bar{x}) \quad \Rightarrow \bar{x}$是可行解

由引理2知, $f\left(x^{(k_j)}\right) \leq f^*$

$\Rightarrow f(\bar{x}) \leq f^* \quad \Rightarrow f(\bar{x}) = f^*$

$\therefore \bar{x}$为最优解.

因为$x^{(k)}$是$F(x,\sigma_k)$的最优解，所以有

$$\nabla_x F\left(x^{(k)},\sigma_k\right)=0$$

$$\Rightarrow \nabla f\left(x^{(k)}\right)-2\sigma_k\sum_{i=1}^{m}\max\left\{0,-g_i\left(x^{(k)}\right)\right\}\nabla g_i\left(x^{(k)}\right)$$

$$+2\sigma_k\sum_{j=1}^{l}h_j\left(x^{(k)}\right)\nabla h_j\left(x^{(k)}\right)=0.$$

当$i\notin I(x^*)$时，有$g_i(x^*)>0$,则存在K,当$k>K$时，

有$g_i\left(x^{(k)}\right)>0$，因此

$$\lim_{k\to\infty}\sigma_k\max\left\{0,-g_i\left(x^{(k)}\right)\right\}=0=w_i.$$

$$\boxed{F(x,\sigma)=f(x)+\sigma\left[\sum_{i=1}^{m}\varphi\left[g_i(x)\right]+\sum_{j=1}^{l}\psi\left[h_j(x)\right]\right]}$$

$$\Rightarrow \nabla f\left(x^{(k)}\right)-2\sigma_k\sum_{i\in I(x^*)}\max\left\{0,-g_i\left(x^{(k)}\right)\right\}\nabla g_i\left(x^{(k)}\right)$$

$$+2\sigma_k\sum_{j=1}^{l}h_j\left(x^{(k)}\right)\nabla h_j\left(x^{(k)}\right)=0.$$

由于x^*是问题(A)的最优解，所以有

$$\nabla f\left(x^*\right)-\sum_{i\in I(x^*)}w_i\nabla g_i\left(x^*\right)+\sum_{j=1}^{l}v_j\nabla h_j\left(x^*\right)=0$$

$$\therefore \quad 2\sigma_k\max\left\{0,-g_i\left(x^{(k)}\right)\right\}\to w_i,\quad i\in I\left(x^*\right)$$

$$2\sigma_k h_j\left(x^{(k)}\right)\to v_j,\quad j=1,2,\cdots,l.$$

外点法的一个重要特点:

函数$F(x,\sigma)$是在整个空间E^n内进行优化,初始点可任意选择,且外点法也可用于非凸规划的最优化.

缺点:

1. 惩罚项$\sigma p(x)$的二阶偏导数一般不存在;

2. 外点法的中间结果不是可行解,不能作为近似解;

3. 当点$x^{(k)}$接近最优解时,罚因子σ_k很大.可能使罚函数性质变坏,使搜索产生极大困难.

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

$$s.t.\quad x_1 + 1 = 0$$

其罚函数为

$$F(x,\sigma) = x_1^2 + x_2^2 + \sigma(x_1+1)^2.$$

$Hesse$矩阵为

$$\nabla_x^2 F(x,\sigma) = \begin{pmatrix} 2+2\sigma & 0 \\ 0 & 2 \end{pmatrix}$$

$$条件数=\frac{最大特征值}{最小特征值} = \frac{2+2\sigma}{2} \to \infty\ (\sigma \to \infty)$$

$$\begin{cases} \min -x^4 \\ s.t. \quad x = 0 \end{cases} \qquad x^* = 0$$

若取$p(x)=x^2$，则$F(x,\sigma)=-x^4+\sigma x^2$，没有最小点.

若取$p(x)=x^8$，则$F(x,\sigma)=-x^4+\sigma x^8$，有极小点

$$x_\sigma = \left(\frac{1}{2\sigma}\right)^{\frac{1}{4}} \to 0$$

内点法

基本思想： 迭代总是从内点出发,并保持在可行域内部进行搜索.

$$\begin{cases} \min \ f(x) \\ s.t. \quad g_i(x) \geq 0 \ \ i = 1,2,\cdots,m \end{cases}$$

其中$f(x), g_i(x)(i=1,2,\cdots,m)$是连续函数。

$$S = \{x \mid g_i(x) \geq 0, i = 1,2,\cdots,m, x \in E^n\}$$

$$\operatorname{int} S = \{x \mid g_i(x) > 0, i = 1, 2, \cdots, m, x \in E^n\}$$

障碍函数

障碍因子

$$G(x,r) = f(x) + rB(x)$$

其中r是很小的正数，$B(x)$定义在可行域内部，它满足两个条件：

(1) $B(x)$是连续函数；

(2) 当点x趋向可行域边界时，$B(x) \to +\infty$。

两种最重要的形式：

倒数障碍函数

对数障碍函数

$$B(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)} \qquad B(x) = -\sum_{i=1}^{m} \ln g_i(x)$$

两种障碍函数的比较

$$\min -x$$

$$s.t. \quad -x-1 \geq 0$$

取 $B(x)=\dfrac{1}{-x-1}$，则内罚函数为

$$G(x,r)=-x-\frac{r}{x+1}$$

令 $\dfrac{\partial G}{\partial x}=-1+\dfrac{r}{(x+1)^2}=0$

得到 $\bar{x}=\bar{x}(r)=-1-\sqrt{r}$

当$r \to 0$时，有$\bar{x}(r) \to -1 = x^*$.

两种障碍函数的比较

$$\min -x$$

$$s.t.\quad -x-1\geq 0$$

取 $B(x)=-\ln\left(-x-1\right)$,则内罚函数为

$$G(x,r)=-x-r\ln\left(-x-1\right)$$

令 $\dfrac{\partial G}{\partial x}=-1-\dfrac{r}{x+1}=0$

得到 $\bar{x}=\bar{x}(r)=-1-r$

当$r\to 0$时，有$\bar{x}(r)\to -1=x^*$.

例:考虑约束优化问题

$$\min \frac{x}{2}$$

$$s.t.\quad x \geq 1$$

该问题的对数障碍函数为

$$G(x,r)=\frac{x}{2}-r\ln(x-1)$$

$G(x,r)$的最小点为:

$$x_r = 1+2r \to 1 (r \to 0)$$

$$G(x_r,r)=\frac{1}{2}+r-r\ln 2r \to \frac{1}{2}(r \to 0)$$

$$\begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \geq 0 \quad i = 1,2,\cdots,m \end{cases}$$

$$\Rightarrow \begin{cases} \min G(x,r) = f(x) + rB(x) \\ s.t. \quad x \in \operatorname{int} S \end{cases}$$

$$\begin{cases} \min G(x,r_k) = f(x) + r_k B(x) \\ s.t. \quad x \in \operatorname{int} S \end{cases}$$

其中$\{r_k\}$为严格单调减且趋于0的障碍因子数列。

步骤:

1. 给定初始点$x^{(0)} \in \operatorname{int} S$，允许误差$\varepsilon > 0$，初始参数$r_1$, 缩小系数$\beta \in (0, 1)$, 置$k = 1$。

2. 以$x^{(k-1)}$为初始点，求解下列问题

$$\min \ f(x) + r_k B(x)$$
$$s.t. \quad x \in \operatorname{int} S$$

设其极小点为$x^{(k)}$。

3. 若$r_k B(x^{(k)}) < \varepsilon$，则停止计算，得到点$x^{(k)}$；否则，令$r_{k+1} = \beta r_k$，置$k := k + 1$，返回2。

例:用内点法求解下列问题

$$\min\ x_1 + x_2$$

$$s.t.\quad -x_1^2 + x_2 \ge 0$$

$$x_1 \ge 0$$

解：定义障碍函数

$$G(x,r) = x_1 + x_2 - r_k \ln\left(-x_1^2 + x_2\right) - r_k \ln x_1$$

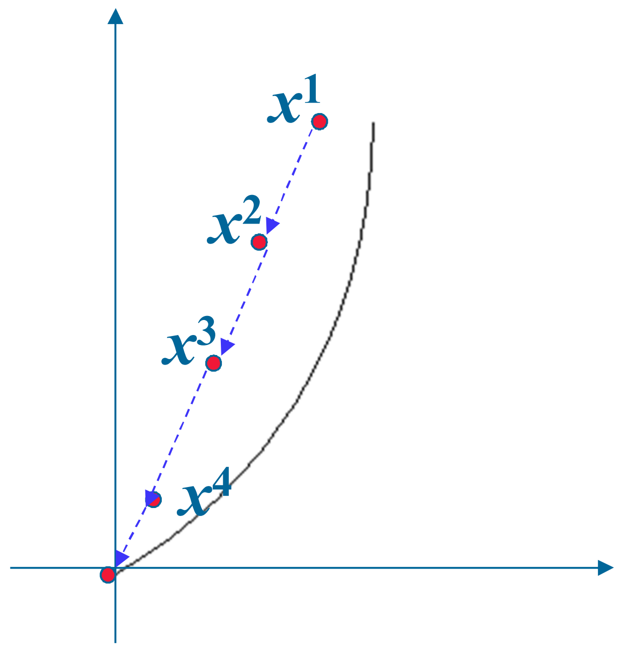
令 $\dfrac{\partial G}{\partial x_1} = 1 - \dfrac{-2x_1 r_k}{-x_1^2 + x_2} - \dfrac{r_k}{x_1} = 0,\ \dfrac{\partial G}{\partial x_2} = 1 - \dfrac{r_k}{-x_1^2 + x_2} = 0$

$\Rightarrow x_1 = \dfrac{1}{4}\left(-1 + \sqrt{1+8r_k}\right), x_2 = \dfrac{3r_k}{2} - \dfrac{1}{8}\left(-1 + \sqrt{1+8r_k}\right)$

当$r_k \to 0$时，$x^* = (0,0)^T$

$$x_1 = \frac{1}{4}\left(-1+\sqrt{1+8r_k}\right), \quad x_2 = \frac{3r_k}{2} - \frac{1}{8}\left(-1+\sqrt{1+8r_k}\right)$$

	r_k	$x_1(r_k)$	$x_2(r_k)$
1	1	0.5	1.25
2	0.5	0.309	0.595
3	0.25	0.183	0.283
4	0.1	0.085	0.107
5	0.0001	0	0



引理1 对于由内点法所产生的序列$\{x^{(k)}\}$，总有

(1) $G(x^{(k+1)}, r_{k+1}) \le G(x^{(k)}, r_k)$

(2) $B(x^{(k+1)}) \ge B(x^{(k)})$

(3) $f(x^{(k+1)}) \le f(x^{(k)})$

证明:(1)由$G(x,r) = f(x) + rB(x)$和$r_{k+1} < r_k$知

$$G(x^{(k)}, r_k) = f(x^{(k)}) + r_k B(x^{(k)})$$

$$\ge f(x^{(k)}) + r_{k+1} B(x^{(k)}) = G(x^{(k)}, r_{k+1})$$

$\because x^{(k+1)}$是$G(x, r_{k+1})$的极小点,$\therefore$ 对$\forall x$,

有$G(x, r_{k+1}) \ge G(x^{(k+1)}, r_{k+1}) \Rightarrow G(x^{(k)}, r_{k+1}) \ge G(x^{(k+1)}, r_{k+1})$

$\Rightarrow G(x^{(k+1)}, r_{k+1}) \le G(x^{(k)}, r_k)$

(2) $x^{(k)}$和$x^{(k+1)}$分别使$G(x,r_k),G(x,r_{k+1})$取极小

$$\therefore f(x^{(k+1)})+r_kB(x^{(k+1)})\geq f(x^{(k)})+r_kB(x^{(k)})$$

$$f(x^{(k)})+r_{k+1}B(x^{(k)})\geq f(x^{(k+1)})+r_{k+1}B(x^{(k+1)}) \quad (*)$$

$$r_kB(x^{(k+1)})+r_{k+1}B(x^{(k)})\geq r_kB(x^{(k)})+r_{k+1}B(x^{(k+1)})$$

$$\Rightarrow \left(r_k-r_{k+1}\right)B(x^{(k)})\leq \left(r_k-r_{k+1}\right)B(x^{(k+1)})$$

$$\Rightarrow B(x^{(k)})\leq B(x^{(k+1)})$$

$$\boxed{\begin{array}{l}(2)\ B(x^{(k+1)})\geq B(x^{(k)})\\(3)\ f(x^{(k+1)})\leq f(x^{(k)})\end{array}}$$

(3) 由(*),得

$$f(x^{(k+1)})-f(x^{(k)})\leq r_{k+1}\left(B(x^{(k)})-B(x^{(k+1)})\right)$$

$$\leq 0$$

引理2 设x^*是问题(A)的一个最优解,则对$\forall k$,有

$$f(x^*) \le f(x^{(k)}) \le G(x^{(k)}, r_k).$$

证明:由$G(x)$的定义及

$$r_k B(x^{(k)}) \ge 0$$

$$\therefore G(x^{(k)}, r_k) = f(x^{(k)}) + r_k B(x^{(k)})$$

$$\ge f(x^{(k)}) \ge f(x^*)$$

$$\boxed{\begin{cases} \min\ f(x) \\ s.t.\quad g_i(x) \ge 0 \quad i = 1, 2, \cdots, m \end{cases}}$$

定理 设$\{x^{(k)}\}$是由内点法产生的一个序列,则$\{x^{(k)}\}$的任何收敛子序列的极限都是原问题的最优解.

证明: 设$\{x^{(k_j)}\}$是$\{x^{(k)}\}$的任一收敛子序列,且有极限$\bar{x}$

$\because f(x)$在E^n上连续,$\therefore \lim\limits_{k_j \to \infty} f\left(x^{(k_j)}\right) = f(\bar{x})$

令$f^* = f(x^*)$为原问题的最优解

由引理1,2,知$\left\{G\left(x^{(k_j)}, r_{k_j}\right)\right\}$是非增的,且以$f^*$为下界的数列

$\therefore \lim\limits_{k_j \to \infty} G\left(x^{(k_j)}, r_{k_j}\right) = G^* \ge f^*$

由f的连续性知，对$\forall \varepsilon > 0, \exists \hat{x} \in \operatorname{int} S$,当$\hat{x}$与$x^*$充分接近时

有 $f(\hat{x}) - f(x^*) < \dfrac{\varepsilon}{2}$.

由G的定义知，对$\forall x \in \operatorname{int} S$，有

$$G\left(x^{k_j}, r_{k_j}\right) \le f(x) + r_{k_j} B(x)$$

特别的，$G\left(x^{k_j}, r_{k_j}\right) \le f(\hat{x}) + r_{k_j} B(\hat{x})$

$$\Rightarrow G\left(x^{k_j}, r_{k_j}\right) - f(x^*) \le f(\hat{x}) - f(x^*) + r_{k_j} B(\hat{x})$$

$\because \lim\limits_{k_j \to \infty} r_{k_j} = 0, \therefore$ 对同一个ε，$\exists K$, 当$k > K$时，有

$$r_{k_j} B(\hat{x}) < \frac{\varepsilon}{2} \Rightarrow G\left(x^{k_j}, r_{k_j}\right) - f(x^*) < \varepsilon$$

$\therefore G^* = f(x^*)$

由引理2, $\lim\limits_{k_j \to \infty} f(x^{k_j}) = f^*$.

$\because f$连续, $\therefore f(\bar{x}) = \lim\limits_{k_j \to \infty} f(x^{k_j}) = f(x^*)$

求初始内点的迭代步骤

1. 任取$x^{(0)} \in E^n, r_0 > 0$(如取$r_0 = 1$),置$k := 0$。

2. 令 $$S_k = \{i \mid g_i(x^{(k)}) \le 0, 1 \le i \le m\},$$
$$T_k = \{i \mid g_i(x^{(k)}) > 0, 1 \le i \le m\}.$$

3. 若$S_k = \varnothing$，停止计算；否则，转4。

4. 构造函数

$$\widetilde{P}(x, r_k) = -\sum_{i \in S_k} g_i(x) + r_k \sum_{i \in T_k} \frac{1}{g_i(x)} \quad (r_k > 0)$$

记 $\widetilde{R}_k = \{x \mid g_i(x) > 0 \quad i \in T_k\}$

5. 以$x^{(k)}$为初始点，在$\widetilde{R}_k$域内，求障碍函数$\widetilde{P}(x,r_k)$的极小点：

$$\min\ \widetilde{P}(x,r_k)$$
$$s.t.\quad x\in\widetilde{R}_k$$

得$x^{(k+1)}$，转6。

6. 令$0<r_{k+1}<r_k\left(\text{如取}r_{k+1}=\dfrac{1}{10}r_k\right)$，置$k:=k+1$，转2。

内点法优点

迭代总在可行域内进行，每一个中间结果都是可行解，可以作为近似解。

内点法缺点

选取初始可行点较困难，且只适用于含不等式约束的非线性规划问题。

乘子法

改进罚函数法的设想

$$(1)\quad \begin{cases} \min\ f(x) \\ s.t.\quad c(x)=0 \end{cases}$$

罚函数问题为： $\min F(x,\sigma)=f(x)+\frac{\sigma}{2}c^2(x) \quad (2)$

问题1: 设x^*是(1)的最优解，x^*是否是(2)的（局部）最优解?

问题2: 是否存在函数$\phi(x,\sigma)$，使x^*恰好是无约束问题

$$\min\ \phi(x,\sigma)$$

全局或局部最优解?

定理（二阶充分条件）：设f，$g_i(i=1,\cdots,m)$和$h_j(j=1,\cdots,l)$是二次连续可微函数，$\bar{x}$为可行解，若存在$\bar{w},\bar{v}$，使$(\bar{x},\bar{w},\bar{v})$为$(LM)$的解且矩阵$\nabla^2 L_x(\bar{x},\bar{w},\bar{v})$在子空间$G$上是正定的，则$\bar{x}$是严格局部极小点。

其中
$$G=\left\{d\neq 0\left|\begin{array}{l}\nabla g_i(\bar{x})^T d=0, i\in I\text{且}\bar{w}_i>0\\ \nabla g_i(\bar{x})^T d\geq 0, i\in I\text{且}\bar{w}_i=0\\ \nabla h_j(\bar{x})^T d=0, j=1,2,\cdots,l\end{array}\right.\right\}.$$

$$\min\ x_1^2 - 3x_2 - x_2^2$$
$$s.t.\quad x_2 = 0$$
$$\text{最优解} x^* = (0,0)^T$$

Lagrange函数为

$$L(x,v) = x_1^2 - 3x_2 - x_2^2 - vx_2 = x_1^2 - (v+3)x_2 - x_2^2$$

$$\nabla L(x,v) = \begin{pmatrix} 2x_1 \\ -(v+3) - 2x_2 \end{pmatrix}, \quad \nabla^2 L(x,v) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

Lagrange函数不存在极小点。

引理1 设W是$n\times n$阶矩阵，a为n阶向量，若对一切d满足$d\neq 0$，$a^T d=0$，均有$d^T W d>0$，则存在$\sigma^*>0$，使当$\sigma\geq\sigma^*$时，矩阵$W+\sigma aa^T$正定.

证明：考虑集合

$$K=\{d \mid \|d\|=1\},$$

$$K'=\left\{d \mid d^T W d\leq 0, d\in K\right\}.$$

对任意$z\neq 0, d=\dfrac{z}{\|z\|}\in K$，若$d\in K\setminus K'$，则$d^T W d>0$，

$$\therefore \forall\sigma>0, \text{有} d^T\left(W+\sigma aa^T\right)d\geq d^T W d>0.$$

引理1 设W是$n\times n$阶矩阵，a为n阶向量，若对一切d满足$d\neq 0$，$a^T d=0$，均有$d^T Wd>0$，则存在$\sigma^*>0$，使当$\sigma\geq\sigma^*$时，矩阵$W+\sigma aa^T$正定.

证明：若$d\in K'$,由于K'是有界闭集，则函数

$d^T Wd$与$\left(a^T d\right)^2$在K'取到极小值$d^{(1)T}Wd^{(1)}$与$\left(a^T d^{(2)}\right)^2$，

并且$a^T d^{(2)}\neq 0$.取

$$\sigma^*>\frac{-d^{(1)T}Wd^{(1)}}{\left(a^T d^{(2)}\right)^2}\geq 0$$

当$\sigma\geq\sigma^*$时，有

$$\begin{aligned} d^T\left(W+\sigma aa^T\right)d &= d^T Wd+\sigma\left(a^T d\right)^2 \\ &\geq d^{(1)T}Wd^{(1)}+\sigma\left(a^T d^{(2)}\right)^2>0. \end{aligned}$$

定理：设x^*, v^*满足问题(1)的二阶充分条件，则存在$\sigma^* > 0$，使当$\sigma \geq \sigma^*$时，x^*是无约束问题

$$\min \phi(x,\sigma) = f(x) - v^* c(x) + \frac{\sigma}{2} c^2(x)$$

$$\begin{cases} \min & f(x) \\ s.t. & c(x) = 0 \end{cases}$$

的严格局部最优解。反之，若$x^{(k)}$是$\phi(x,\sigma_k)$的极小点，并且$c(x^{(k)}) = 0$，则$x^{(k)}$是问题(1)的局部最优解.

证明：设x^*, v^*满足问题(1)的二阶充分条件, 则

$$\nabla_x \phi(x^*,\sigma) = \nabla f(x^*) - v^* \nabla c(x^*) + \sigma c(x^*) \nabla c(x^*) = 0$$

$$\nabla_x^2 \phi(x^*,\sigma) = \nabla^2 f(x^*) - v^* \nabla^2 c(x^*) + \sigma \nabla c(x^*) \nabla c(x^*)^T$$

$$= \nabla_x^2 L(x^*, v^*) + \sigma \nabla c(x^*) \nabla c(x^*)^T$$

取$W = \nabla_x^2 L(x^*, v^*), a = \nabla c(x^*)$,由引理1,存在$\sigma^*$,当$\sigma \geq \sigma^*$时,

$\nabla_x^2 \phi(x^*,\sigma)$正定，所以$x^*$是无约束问题的严格局部最优解。

定理：设x^*, v^*满足问题(1)的二阶充分条件，则存在$\sigma^* > 0$，使当$\sigma \ge \sigma^*$时，x^*是无约束问题

$$\min \phi(x,\sigma) = f(x) - v^* c(x) + \frac{\sigma}{2} c^2(x)$$

$$\begin{cases} \min \ f(x) \\ s.t. \quad c(x) = 0 \end{cases}$$

的严格局部最优解。反之，若$x^{(k)}$是$\phi(x,\sigma_k)$的极小点，并且$c(x^{(k)}) = 0$，则$x^{(k)}$是问题(1)的局部最优解.

证明：若$x^{(k)}$是$\phi(x,\sigma_k)$的极小点，且$c(x^{(k)}) = 0$

则对任意的x,有

$$\phi(x,\sigma_k) \ge \phi(x^{(k)},\sigma_k)$$

$$= f(x^{(k)}) - v^* c(x^{(k)}) + \frac{\sigma_k}{2} c^2(x^{(k)}) = f(x^{(k)})$$

特别地，当x是问题(1)的可行点时，有$f(x) \ge f(x^{(k)})$,

所以$x^{(k)}$是问题(1)的局部极小点。

$$\min\ x_1^2 - 3x_2 - x_2^2 \qquad \text{最优解} x^* = (0,0)^T$$

$$s.t.\quad x_2 = 0 \qquad\qquad v^* = -3$$

定义函数

$$\phi(x,\sigma) = \left(x_1^2 - 3x_2 - x_2^2\right) + 3x_2 + \frac{\sigma}{2}x_2^2$$

取$\sigma^* = 2+\varepsilon$，$\varepsilon > 0$, 当$\sigma \geq \sigma^*$时

$$\nabla\phi(x^*,\sigma) = \begin{pmatrix} 2x_1 \\ -3-2x_2+3+\sigma x_2 \end{pmatrix}\Bigg|_{x^*} = 0$$

$$\nabla^2\phi = \begin{pmatrix} 2 & 0 \\ 0 & -2+\sigma \end{pmatrix} \quad \Rightarrow x^* \text{是} \phi(x,\sigma) \text{的极小点.}$$

乘子法的基本思想

$$\min\ f(x)$$

$$s.t.\quad h_j(x)=0,\ j=1,2,\cdots,l$$

其中$f,h_j\left(j=1,2,\cdots,l\right)$是二次连续可微函数

定义增广Lagrange函数（乘子罚函数）：

$$\phi(x,v,\sigma)=f(x)-\sum_{j=1}^{l}v_jh_j(x)+\frac{\sigma}{2}\sum_{j=1}^{l}h_j^2(x)$$

$$=f(x)-v^Th(x)+\frac{\sigma}{2}h(x)^Th(x)$$

其中 $\sigma>0, v=\left(v_1,\cdots,v_l\right)^T,$

$$h(x)=\left(h_1(x),\cdots,h_l(x)\right)^T.$$

乘子迭代公式：

设在第k次迭代中，$Lagrange$乘子向量的估计为$v^{(k)}$，罚因子取σ,得到$\phi(x, v^{(k)}, \sigma)$的极小点$x^{(k)}$.这时有

$$\nabla_x \phi(x^{(k)}, v^{(k)}, \sigma) = \nabla f(x^{(k)}) - \sum_{j=1} (v_j^{(k)} - \sigma h_j(x^{(k)})) \nabla h_j(x^{(k)})$$

$$= 0.$$

修正乘子v的公式：

$$v_j^{(k+1)} = v_j^{(k)} - \sigma h_j(x^{(k)}),$$

$$j = 1, 2, \cdots, l.$$

等式约束问题乘子法计算步骤

1. 给定初始点$x^{(0)}$，乘子向量初始估计$v^{(1)}$，参数σ，允许误差$\varepsilon>0$，常数$\alpha>1$，$\beta\in(0,1)$，置$k:=1$.

2. 以$x^{(k-1)}$为初始点，解无约束问题

$$\min\ \phi(x,v^{(k)},\sigma).$$

3. 若$\left\|h(x^{(k)})\right\|<\varepsilon$，则停止计算，得到点$x^{(k)}$；否则转4。

4. 若$\dfrac{\left\|h(x^{(k)})\right\|}{\left\|h(x^{(k-1)})\right\|}\ge\beta$，则置$\sigma:=\alpha\sigma$，转5；否则进行5。

5. 计算$v_j^{(k+1)}=v_j^{(k)}-\sigma h_j(x^{(k)}),\ j=1,2,\cdots,l$，置$k:=k+1$，转2。

$$\min\ f(x)=\frac{1}{2}\left(x_1^2+\frac{1}{3}x_2^2\right)$$

$$s.t.\quad h(x)=x_1+x_2-1=0$$

$$x^*=\left(\frac{1}{4},\frac{3}{4}\right)^T$$

$$v^*=\frac{1}{4}$$

增广Lagrange函数为

$$\phi(x,v,\sigma)=\frac{1}{2}\left(x_1^2+\frac{1}{3}x_2^2\right)-v(x_1+x_2-1)+\frac{\sigma}{2}(x_1+x_2-1)^2$$

$$\frac{\partial\phi}{\partial x_1}=x_1-v+\sigma(x_1+x_2-1)=0$$

$$\frac{\partial\phi}{\partial x_2}=\frac{1}{3}x_2-v+\sigma(x_1+x_2-1)=0$$

$$\Rightarrow\quad \bar{x}(v,\sigma)=\left(\frac{\sigma+v}{1+4\sigma},\frac{3(\sigma+v)}{1+4\sigma}\right).$$

$$\bar{x}(v,\sigma)=\left(\frac{\sigma+v}{1+4\sigma},\frac{3(\sigma+v)}{1+4\sigma}\right) \quad (1)$$

由迭代公式$v_j^{(k+1)}=v_j^{(k)}-\sigma h_j(x^{(k)}), j=1,2,\cdots,l,$

得$v^{(k+1)}=v^{(k)}-\sigma\left(\frac{\sigma+v^{(k)}}{1+4\sigma}+\frac{3(\sigma+v^{(k)})}{1+4\sigma}-1\right)$

$$=v^{(k)}-\sigma\frac{4v^{(k)}-1}{1+4\sigma}=\frac{v^{(k)}+\sigma}{1+4\sigma}$$

当$\sigma>0$时，$\left\{v^{(k)}\right\}$收敛，且σ越大收敛越快，如取σ=10

令$v^{(k+1)}\to\bar{v}$，$v^{(k)}\to\bar{v}$，得$\bar{v}=\frac{1}{4}=v^{*}$

在(1)中取$\sigma=10,\bar{v}=\frac{1}{4}$, 得最优解$x^*=\left(\frac{1}{4},\frac{3}{4}\right)$.

乘子法得到的点列为：

$$x^{(k)} = \left(\frac{\sigma_k + v^{(k)}}{1+4\sigma_k}, \frac{3(\sigma_k + v^{(k)})}{1+4\sigma_k} \right)^T$$

迭代到第8步得：$(0.25, 0.75)$

罚函数法得到的点列为：

$$x^{(k)} = \left(\frac{\sigma_k}{1+4\sigma_k}, \frac{3\sigma_k}{1+4\sigma_k} \right)^T$$

迭代到第20步得：$(0.25, 0.75)$

不等式约束问题的乘子法

$$\min\ f(x)$$

$$s.t.\quad g_j(x)\ge 0,\ j=1,2,\cdots,m$$

引入变量y_j，把不等式约束问题化为等式约束问题

$$\begin{cases}\min\ f(x)\\ s.t.\quad g_j(x)-y_j^2=0,\ j=1,2,\cdots,m\end{cases}$$

增广Lagrange函数为

$$\Phi(x,y,w,\sigma)=f(x)-\sum_{j=1}^{m}w_j(g_j(x)-y_j^2)$$

$$+\frac{\sigma}{2}\sum_{j=1}^{m}(g_j(x)-y_j^2)^2$$

$$\min f(x)$$

$$s.t. \quad g_j(x) \geq 0, j = 1, 2, \cdots, m$$

$$\Longleftrightarrow \quad \min \Phi(x, y, w, \sigma)$$

分两步进行 (1) 求解问题

$$\min_{y} \Phi(x, y, w, \sigma)$$

得到最优解$\overline{y} = \overline{y}(x, w, \sigma)$。

(2) 将$\overline{y} = \overline{y}(x, w, \sigma)$代入$\Phi(x, y, w, \sigma)$中，求解问题

$$\min_{x} \phi(x, w, \sigma) = \Phi(x, y(x, w, \sigma), w, \sigma)$$

得到最优解$\overline{x} = \overline{x}(w, \sigma)$。

求$\bar{y} = y(x, v, \sigma)$的解析表达式

$$\Phi = f(x) - \sum_{j=1}^{m} w_j (g_j(x) - y_j^2) + \frac{\sigma}{2} \sum_{j=1}^{m} (g_j(x) - y_j^2)^2$$

$$= f(x) + \sum_{j=1}^{m} \left\{ \frac{\sigma}{2} \left[y_j^2 - \frac{1}{\sigma} (\sigma g_j(x) - w_j) \right]^2 - \frac{w_j^2}{2\sigma} \right\}$$

为使Φ关于y_j取极小，y_j的取值为

若$\sigma g_j(x) - w_j \geq 0$，则$y_j^2 = \frac{1}{\sigma}(\sigma g_j(x) - w_j)$;

若$\sigma g_j(x) - w_j < 0$，则$y_j = 0$.

$$y_j^2 = \frac{1}{\sigma} \max\left\{0, \sigma g_j(x) - w_j\right\}.$$

$$\min\ f(x)$$
$$s.t.\quad g_j(x) \geq 0,\ j = 1, 2, \cdots, m$$

增广Lagrange函数为

$$\phi(x, w, \sigma) = f(x)$$
$$+\frac{1}{2\sigma}\sum_{j=1}^{m}\left\{\left[\max\left(0,\ w_j - \sigma g_j(x)\right)\right]^2 - w_j^2\right\}.$$

$$\min \ f(x)$$

$$s.t. \quad g_j(x) \geq 0, j = 1, 2, \cdots, m$$

$$h_j(x) = 0, j = 1, 2, \cdots, l$$

增广Lagrange函数为

$$\phi(x, w, v, \sigma) = f(x)$$

$$+\frac{1}{2\sigma}\sum_{j=1}^{m}\left\{\left[\max\left(0, w_j - \sigma g_j(x)\right)\right]^2 - w_j^2\right\}$$

$$-\sum_{j=1}^{l} v_j h_j(x) + \frac{\sigma}{2}\sum_{j=1}^{l} h_j^2(x).$$

乘子迭代公式为

$$\begin{cases} w_j^{(k+1)} = \max(0, w_j^{(k)} - \sigma g_j(x^{(k)})), j = 1, 2, \cdots, m \\ v_j^{(k+1)} = v_j^{(k)} - \sigma h_j(x^{(k)}), \quad j = 1, 2, \cdots, l \end{cases}$$

$$\min\ x_1^2 + x_2^2$$
$$s.t.\quad x_1 + x_2 \geq 2$$

增广Lagrange函数为

$$\phi = x_1^2 + x_2^2 + \frac{1}{2\sigma}\left\{\left[\max\left(0, w - \sigma\left(x_1 + x_2 - 2\right)\right)\right]^2 - w^2\right\}$$

$$= \begin{cases} x_1^2 + x_2^2 - \dfrac{w^2}{2\sigma} & w - \sigma\left(x_1 + x_2 - 2\right) < 0 \\ x_1^2 + x_2^2 + \dfrac{1}{2\sigma}\left\{\left[w - \sigma\left(x_1 + x_2 - 2\right)\right]^2 - w^2\right\} & \text{否则} \end{cases}$$

当$w-\sigma\left(x_1+x_2-2\right)<0$时，令

$$\frac{\partial\phi}{\partial x_1}=2x_1=0,\quad \frac{\partial\phi}{\partial x_2}=2x_2=0,$$得$\bar{x}=(0,0)^T$.

当σ充分大时，$\bar{x}$不满足$w-\sigma\left(x_1+x_2-2\right)<0$,

即$\bar{x}$不是$\phi\left(x,w,\sigma\right)$的极小点.

当$w-\sigma\left(x_1+x_2-2\right)\geq 0$时，令

$$\frac{\partial\phi}{\partial x_1}=2x_1-\left[w-\sigma\left(x_1+x_2-2\right)\right]=0,$$

$$\frac{\partial\phi}{\partial x_2}=2x_2-\left[w-\sigma\left(x_1+x_2-2\right)\right]=0,$$得

$$\overline{x}=\left(\frac{2\sigma+w}{2\sigma+2},\frac{2\sigma+w}{2\sigma+2}\right)^T$$

当σ充分大时，$\overline{x}$满足$w-\sigma\left(x_1+x_2-2\right)\geq 0$.

代入迭代公式

$$w^{(k+1)} = \max\left(0, w^{(k)} - \sigma\left(x_1 + x_2 - 2\right)\right)$$

$$= \max\left(0, \frac{2\sigma + w^{(k)}}{\sigma + 1}\right)$$

若给定$w^{(1)} > 0, \sigma > 0$(如$\sigma = 10$)，则有

$$w^{(k+1)} = \frac{2\sigma + w^{(k)}}{\sigma + 1} = \frac{w^{(k)}}{\sigma + 1} + \frac{2\sigma}{\sigma + 1}$$

令$w^{(k)} \to w^*$，得$w^* = 2.$

$$\therefore \bar{x} = \left(\frac{2\sigma + w^*}{2\sigma + 2}, \frac{2\sigma + w^*}{2\sigma + 2}\right)^T = (1,1)^T.$$

序列线性规划法

基本思想：将非线性规划线性化，通过解线性规划问题来求解原问题的近似解。

近似规划法

$$
(NP)\quad \begin{cases} \min\ f(x) \\ s.t.\quad g_j(x) \ge 0 \quad j=1,\cdots,m \\ \qquad\quad h_j(x)=0 \quad j=1,\cdots,l \end{cases}
$$

其中$x\in R^n, f, g_i, h_j$均存在一阶连续偏导数。

基本思想：

将(NP)中的目标函数$f(x)$和约束函数$g_i(x)$,$h_j(x)$线性化，并对变量的取值范围加以限制，从而得到线性近似规划，用单纯形方法求解此线性规划问题，把其最优解作为(NP)的解的近似。

设$x^{(k)}$是原问题的可行解，将$f(x), g_i(x)$,$h_j(x)$在$x^{(k)}$点$Taylor$展开。

$$\begin{cases} \min\ f(x^{(k)}) + \nabla f(x^{(k)})^T (x - x^{(k)}) \\ s.t.\quad g_i(x^{(k)}) + \nabla g_i(x^{(k)})^T (x - x^{(k)}) \ge 0, i = 1, \cdots m \\ \qquad\quad h_j(x^{(k)}) + \nabla h_j(x^{(k)})^T (x - x^{(k)}) = 0, j = 1, \cdots l \end{cases}$$

$$
\begin{cases}
\min\ f(x^{(k)})+\nabla f(x^{(k)})^T(x-x^{(k)}) \\
s.t.\quad g_i(x^{(k)})+\nabla g_i(x^{(k)})^T(x-x^{(k)})\geq 0, i=1,\cdots m \\
\qquad\quad h_j(x^{(k)})+\nabla h_j(x^{(k)})^T(x-x^{(k)})=0, j=1,\cdots l \\
\qquad\quad |x_j-x_j^{(k)}|\leq \delta_j^{(k)},\quad j=1,2,\cdots,n
\end{cases}
$$

步骤

1. 给定初始可行解$x^{(1)}$，$\delta_j^{(1)}, j=1,2,\cdots,n$，缩小误差$\beta\in(0,1)$，允许误差$\varepsilon_1,\varepsilon_2$，置$k:=1$。

2. 求解线性规划问题：

$$
\begin{cases}
\min\ f(x^{(k)})+\nabla f(x^{(k)})^T(x-x^{(k)}) \\
s.t.\quad g_i(x^{(k)})+\nabla g_i(x^{(k)})^T(x-x^{(k)})\ge 0, i=1,\cdots m \\
\qquad\quad h_j(x^{(k)})+\nabla h_j(x^{(k)})^T(x-x^{(k)})=0, j=1,\cdots l \\
\qquad\quad |x_j-x_j^{(k)}|\le\delta_j^{(k)},\quad j=1,2,\cdots,n
\end{cases}
$$

得最优解$\bar{x}$.

3. 若$\bar{x}$是(NP)的可行解，则令$x^{(k+1)}=\bar{x}$，转4；否则，置$\delta_j^{(k)}:=\beta\delta_j^{(k)}, j=1,2,\cdots,n$，返回2。

4. 若$|f(x^{(k+1)})-f(x^{(k)})|<\varepsilon_1$，且$\left\|x^{(k+1)}-x^{(k)}\right\|<\varepsilon_2$或$|\delta_j^{(k)}|<\varepsilon_2, j=1,2,\cdots,n$，则点$x^{(k+1)}$为近似最优解；否则，令$\delta_j^{(k+1)}:=\delta_j^{(k)}$，$j=1,2,\cdots,n$，置$k:=k+1$，返回2。

割平面法

$$(NP)\quad \begin{cases} \min\ f(x) = cx \\ s.t. \quad g_i(x) \geq 0,\ i = 1,2,\cdots,m \end{cases}$$

其中$c_{1\times n}, g_i(x)$为凹函数，可行域为S。

$$(NP)\quad \begin{cases} \min\ f(x) \\ s.t. \quad g_i(x) \geq 0,\ i = 1,2,\cdots,m \end{cases}$$

其中$\boldsymbol{f(x)}$是凸函数

$$\Longleftrightarrow \quad \begin{cases} \min\ z \\ s.t. \quad z - f(x) \geq 0 \\ \qquad\quad g_i(x) \geq 0,\ i = 1,2,\cdots,m \end{cases}$$

基本思想：用多面集取代可行域，并在多面集上极小化目标函数。

给定一组线性约束，确定一个多面集S_1，使得

$S_1 \supset S$. 求解线性规划问题：

$$\begin{cases} \min\ f(x) = cx \\ s.t. \quad x \in S_1 \end{cases}$$

$$\boxed{\begin{cases} \min\ f(x) = cx \\ s.t. \quad g_i(x) \geq 0, i = 1, 2, \cdots, m \end{cases}}$$

设问题的最优解为$\bar{x}$，设

$$g_r(\bar{x}) = \min\{g_i(\bar{x}) \mid i = 1, 2, \cdots, m\}$$

将$g_r(x)$线性化：

$$g_r(\bar{x}) + \nabla g_r(\bar{x})^T (x - \bar{x}) \geq 0$$

令$S_2 = S_1 \cap \{x \mid g_r(\bar{x}) + \nabla g_r(\bar{x})^T (x - \bar{x}) \geq 0\}$。

步骤

1. 给定初始多面集S_1，使得$S_1 \supset S$，并且使cx在S_1上有界，给定允许误差$\varepsilon > 0$，置$k := 1$。

2. 求$\begin{cases} \min\ cx \\ s.t. \quad x \in S_k \end{cases}$，设其最优解为$x^{(k)}$.

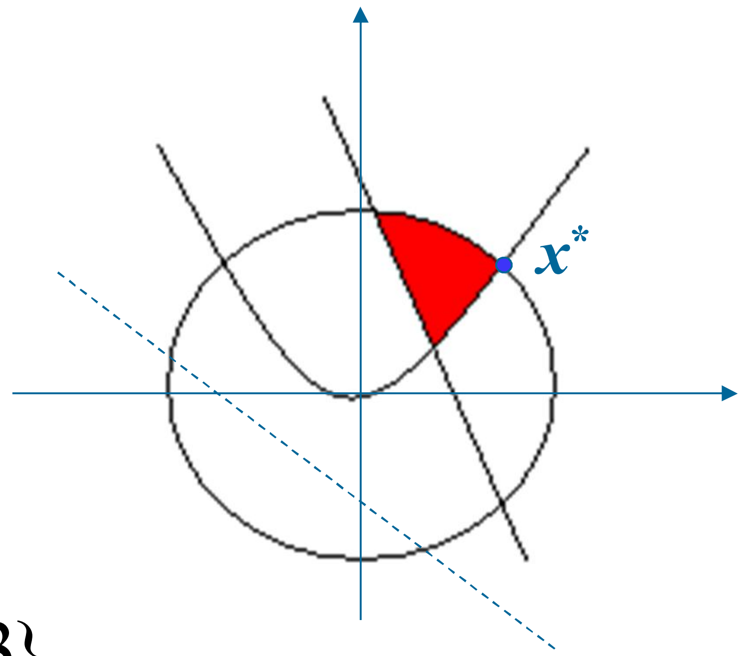
3. 若$g_i(x^{(k)}) \geq -\varepsilon, i = 1, 2, \cdots, m$，则停止计算，得到近似解$x^{(k)}$；否则，选择下标$r$，使得

$$g_r(x^{(k)}) = \min_i g_i(x^{(k)})$$

令$S_{k+1} = \left\{ x \in S_k \mid g_r(x^{(k)}) + \nabla g_r(x^{(k)})^T (x - x^{(k)}) \geq 0 \right\}$

置$k := k + 1$，返回2。

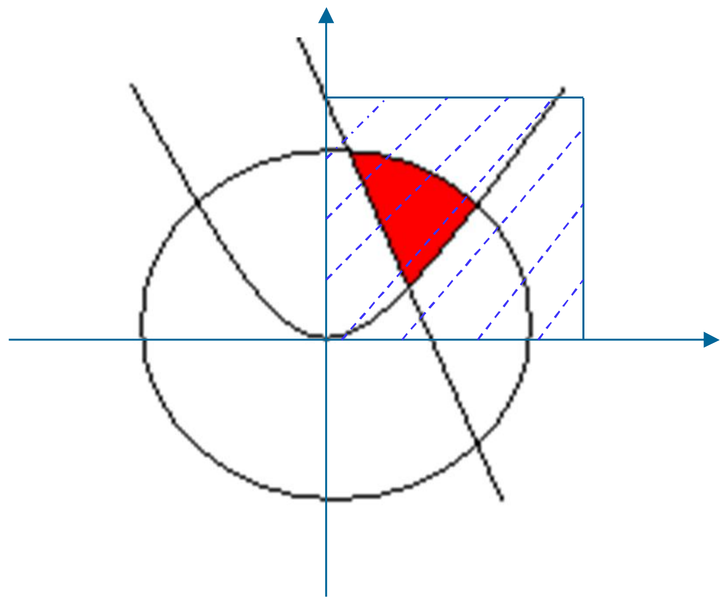
$$\begin{cases} \min\ f(x) = -4x_1 - x_2 \\ s.t. \quad g_1(x) = 8 - x_1^2 - x_2^2 \geq 0 \\ \qquad\ \ g_2(x) = 2x_2 - x_1^2 \geq 0 \\ \qquad\ \ g_3(x) = 3x_1 + x_2 - 3 \geq 0 \end{cases}$$



取$S_1 = x \mid 0 \leq x_1 \leq 3, 0 \leq x_2 \leq 3\}$

解线性规划问题：

$$\begin{cases} \min -4x_1 - x_2 \\ s.t. \quad x_1 \leq 3 \\ \qquad\ \ x_2 \leq 3 \\ \qquad\ \ x_1, x_2 \geq 0 \end{cases}$$



得最优解 $x^{(1)} = (x_1, x_2)^T = (3, 3)^T$

由于$g_1(x^{(1)}) = -10,\ g_2(x^{(1)}) = -3,\ g_3(x^{(1)}) = 9$

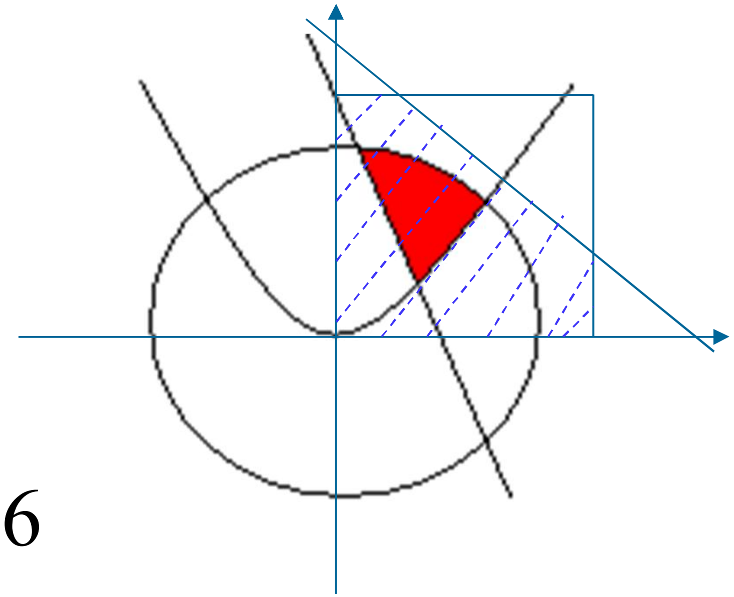
所以$r = 1$。令

$$g_1(x) \approx 26 - 6x_1 - 6x_2$$

$$S_2 = \{x \in S_1 \mid 26 - 6x_1 - 6x_2 \geq 0\}.$$

解线性规划问题：

$$\begin{cases} \min -4x_1 - x_2 \\ s.t. \quad x_1 \leq 3 \\ \qquad\quad x_2 \leq 3 \\ \qquad\quad 6x_1 + 6x_2 \leq 26 \\ \qquad\quad x_1, x_2 \geq 0 \end{cases}$$



得最优解 $x^{(2)}=\left(x_1, x_2\right)^T=\left(3, \dfrac{4}{3}\right)^T$

由于$g_1(x^{(2)})=-\dfrac{25}{9}, g_2(x^{(2)})=-\dfrac{19}{3}, g_3(x^{(3)})=\dfrac{22}{3}$

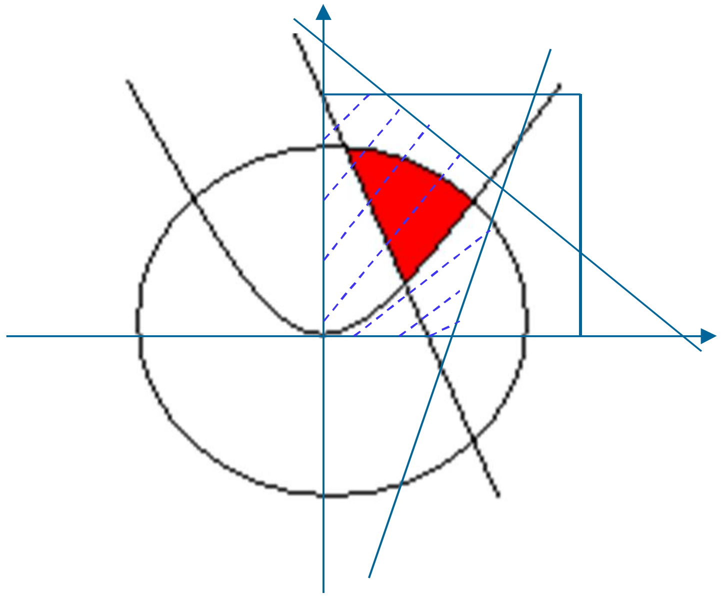
所以$r=2$。令

$$g_2(x)\approx 8-6x_1+2x_2$$

$$S_3=\left\{x\in S_2 \mid 8-6x_1+2x_2\geq 0\right\}.$$

解线性规划问题：

$$\begin{cases} \min -4x_1 - x_2 \\ s.t. \quad x_1 \le 3 \\ \qquad x_2 \le 3 \\ \qquad 6x_1 + 6x_2 \le 26 \\ \qquad 6x_1 - 2x_2 \le 8 \\ \qquad x_1, x_2 \ge 0 \end{cases}$$



得最优解 $x^{(3)} = (x_1, x_2)^T = \left(\dfrac{25}{12}, \dfrac{27}{12}\right)^T$