约束极值问题的算法

一、惩罚函数法（SUMT）

1. 问题：$\min\ f(x)$

$$s.t.\ g_i(x) \geq 0 \quad i = 1, \cdots, \quad m$$

$$\Leftrightarrow \min\ f(x)$$

$$s.t.\ \ x \in D$$

$D = \{x \mid g(x) \geq 0\}$：可行点集或可行解集

2、算法思想：

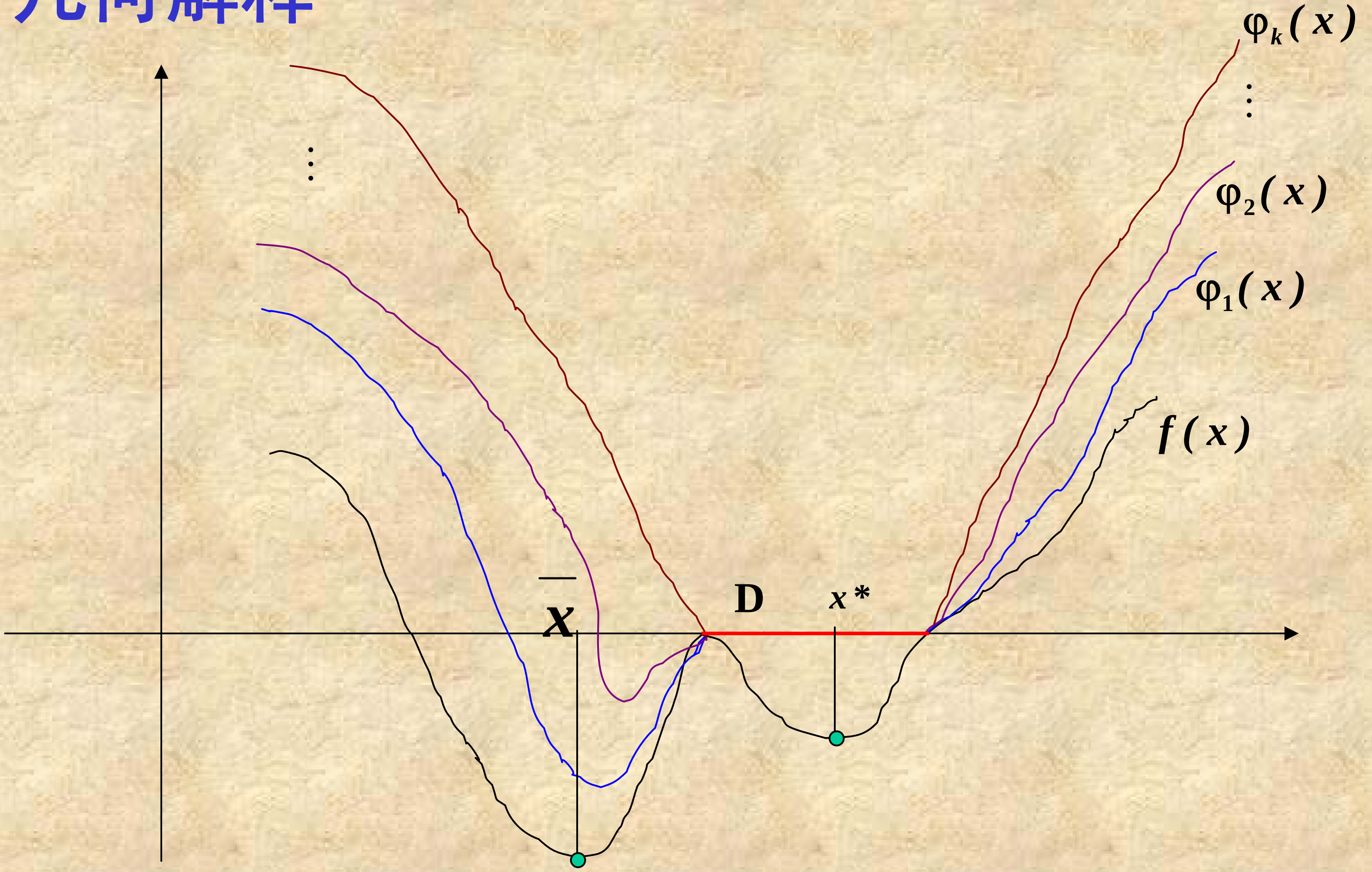
将有约束优化问题转化为一系列无约束优化问题进行求解。（**Sequential Unconstrained Minimization Technique - SUMT**）

3、算法类型：

- 外点法（外惩法）
- 内点法（内惩法）

4、外点法（外部惩罚函数法）

（1）几何解释



（2）算法分析

$$\min_{x\in D} f(x) \Rightarrow \min_{x\in R^n} \varphi_1(x) \Rightarrow \min_{x\in R^n} \varphi_2(x) \Rightarrow \cdots \Rightarrow \min_{x\in R^n} \varphi_k(x)$$

如何构造 $\varphi_k(x)$?

$\varphi_k(x)$满足： $\varphi_k(x) = f(x) \quad x \in D$

$\varphi_k(x) > f(x) \quad x \notin D$且$\varphi_k(x)\uparrow(k\uparrow)$

则可取： $\varphi_k(x) = f(x) + \lambda_k p(x),$

其中 $p(x)$ 满足

(1) $p(x) = 0 \quad x \in D$;

(2) $p(x) > 0 \quad x \notin D$。

其中 $\lambda_1 < \lambda_2 < \cdots < \lambda_k \to +\infty$

这样一来，我们只需构　造 $p(x)$，事实上，

因为　$g_j(x)\ge 0$

$$\Leftrightarrow \min\{g_j(x),0\}=0 \qquad \Leftrightarrow \min^2\{g_j(x),0\}=0$$

所以可设

$$p(x)=\sum_{j=1}^{m}(\min\{g_j(x),0\})^2=\sum_{j=1}^{m}\min^2\{g_j(x),0\}$$

显然 $p(x)$ 满足恰前面的条件（1）和（2）。

结论 1: 如果 $g_j(x)$（$j=1,2,\cdots,m$）连续，那么 $p(x)$ 也连续。

事实上，只须注意:

$$\min\{f_1(x),f_2(x)\}=\frac{f_1(x)+f_2(x)-|f_1(x)-f_2(x)|}{2}$$

结论 2：如果 (1) 问题（B）：$\min\limits_{x\in R^n}\varphi_k(x)$ 有最优解 $x^*(\lambda_k)$；

(2) $x^*(\lambda_k)\in D$，即$g_j(x^*(\lambda_k))\geq 0$（$j=1,2,\cdots,m$）。

则 $x^*(\lambda_k)$ 也是问题（A）：$\min\limits_{x\in D} f(x)$ 的最优解。

证明：因为 $x^*(\lambda_k)$ 是（B）：$\min\limits_{x\in R^n}\varphi_k(x)$ 的最优解。

所以 $\varphi_k(x^*(\lambda_k))\leq\varphi_k(x),\quad \forall\ x\in R^n$。

又 $x^*(\lambda_k)\in D$，即$g_j(x^*(\lambda_k))\geq 0$（$j=1,2,\cdots,m$），

所以 $p(x^*(\lambda_k))=0$。

所以 $\forall x\in D$，有 $f(x^*(\lambda_k))= f(x^*(\lambda_k))+\lambda_k p(x^*(\lambda_k))$

$$=\varphi_k(x^*(\lambda_k))$$

$$\leq\varphi_k(x)$$

$$= f(x) + \lambda_k p(x)$$

$$= f(x)$$

所以 $x^*(\lambda_k)$是问题 $\min\limits_{x\in D} f(x)$ 的最优解。

（3）算法步骤（外点法）：

***step*1.** 给定初始点 $x^{(0)}$，初始惩罚因子 $\lambda_1 > 0$（可取 $\lambda_1 = 1$），精度$\varepsilon > 0, k := 0$。

***step*2.** 以 $x^{(k)}$ 为初始点，求解无约束优化问题

$$\min_{x \in R^n} \varphi_k(x) = f(x) + \lambda_k p(x) = f(x) + \lambda_k \sum_{j=1}^{m} \min{}^2(g_j(x), 0)$$

得到极小点为 $x^*(\lambda_k)$，记为 $x^{(k+1)}$.

***step* 3:** 如果 $x^*(\lambda_k) \in D$，即 $g_j(x^*(\lambda_k)) \geq \varepsilon$（$j = 1, 2, \cdots, m$），

则 $x^*(\lambda_k)$ 就是问题（A）：$\min_{x \in D} f(x)$ 的最优解，*stop*;

否则转 ***step* 4.**

***step* 4:** 给定 $\lambda_{k+1} > \lambda_k$（可取 $\lambda_{k+1} = \alpha \lambda_k$ 这里 $\alpha > 1$ 为惩罚因子的放大系数），$k := k + 1$，转 ***step* 2.**

（4）应注意的问题

(a) 在$step2$中，可用无约束优化问题的算法求解

$$\min_{x\in R^n} \varphi_k(x) = f(x) + \lambda_k p(x)$$

(b) 在实际计算中，判断 $x^*(\lambda_k) \in D$ 用 $g_j(x^*(\lambda_k)) \geq \varepsilon$

（$j = 1,2,\cdots,m$）或 $\lambda_k p(x) < \varepsilon$.

(c) 在算法中，一开始λ_k不宜取的过大。否则会造成无约束优化问题 $\min_{x\in R^n} \varphi_k(x)$求解困难（$\lambda_k$越大，$\varphi_k(x)$的$Hesse$矩阵的条件数越坏）。

(d) 通常称:

$\varphi_k(x)$: 增广函数，$p(x)$: 惩罚函数

λ_k: 惩罚因子，$\lambda_k p(x)$: 惩罚项

例. 试用外点法（外部惩罚 函数法）求解如下优化 问题

$$\min f(x) = (x-1)^2$$
$$s.t. \;\; x - 2 \geq 0$$

解: 构造增广函数 $\varphi_k(x)$ 如下:

$$\varphi_k(x) = (x-1)^2 + \lambda_k \min{}^2\{x-2, 0\}$$

$$= \begin{cases} (x-1)^2 & if \;\; x \geq 2 \\ (x-1)^2 + \lambda_k (x-2)^2 & if \;\; x < 2 \end{cases}$$

$$\frac{d\varphi_k(x)}{dx}=\begin{cases}2(x-1) & \text{if } x\ge 2\\ 2(x-1)+2\lambda_k(x-2) & \text{if } x<2\end{cases}$$

由 $\dfrac{d\varphi_k(x)}{dx}=0$ 可得:

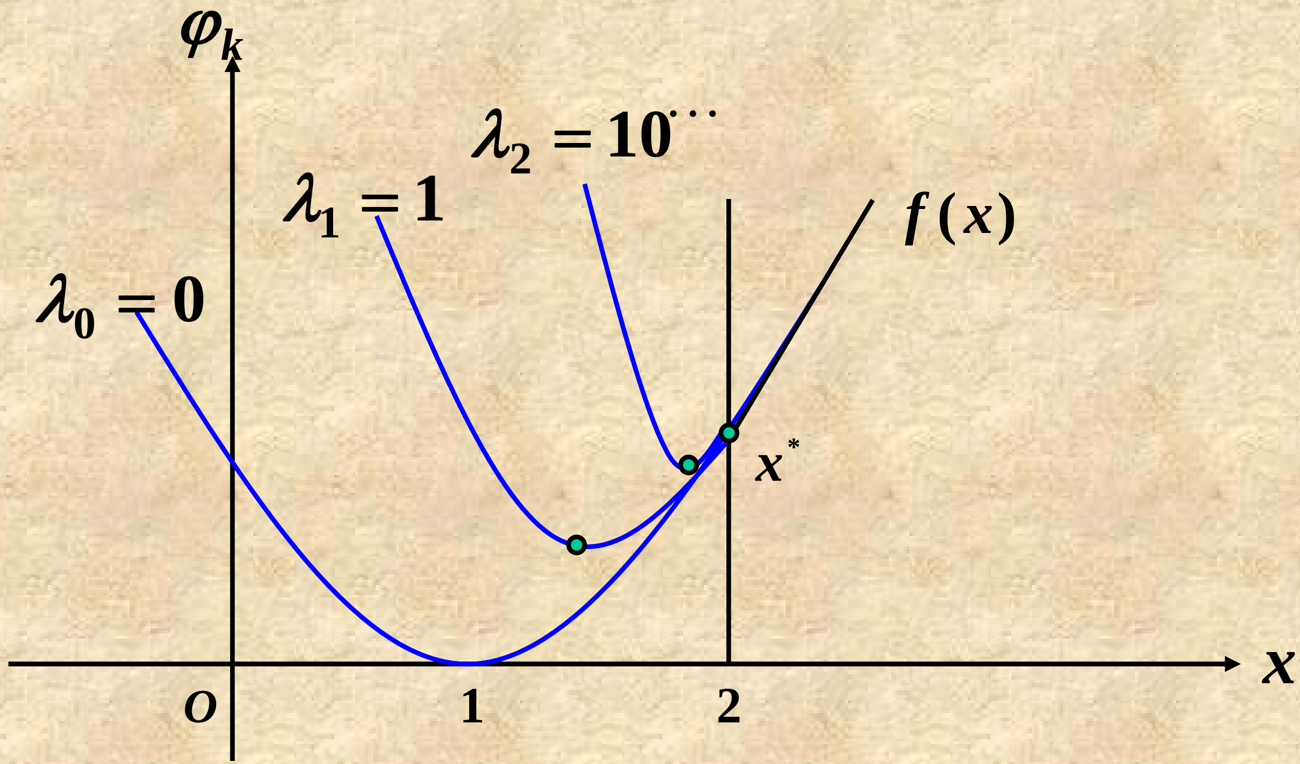
$$2(x-1)+2\lambda_k(x-2)=0$$

所以 $x^k=x^*(\lambda_k)=\dfrac{1+2\lambda_k}{1+\lambda_k}\notin D$

这就是对于固定的λ_k，问题$\min\limits_{x\in R^n}\varphi_k(x)$的最优解。

$$\lim_{k\to\infty}x^{(k)}=\lim_{k\to\infty}x^*(\lambda_k)=\lim_{k\to\infty}\frac{1+2\lambda_k}{1+\lambda_k}=2=x^*,$$

x^*就是所求原问题的最优 解。



φ_k
$\lambda_2 = 10 \cdots$
$\lambda_1 = 1$
$\lambda_0 = 0$
$f(x)$
x^*
x
O
1
2

（5）一般模型的外点法

- $$\min \ f(x)$$
  $$s.t. \begin{cases} g_i(x) \geq 0 & i = 1, \cdots, \ m \\ h_j(x) = 0 & j = 1, \cdots, \ p \end{cases}$$

- 构造如下的惩罚函数

$$p(x) = \sum_{j=1}^{m} (\min\{g_j(x), 0\})^2 + \sum_{k=1}^{l} (h_k(x))^2$$

$$= \sum_{j=1}^{m} \min{}^2\{g_j(x), 0\} + \sum_{k=1}^{l} h_k^2(x)$$

- **算法步骤相同**

例 2　写出用外点法求解下列非线性规划问题时的增广函数

$$\min \quad f(x) = 3x_1 + x_3 x_1^3 - 2x_2^2$$

$$s.t. \quad \begin{cases} 2x_1 x_3 - x_2^2 + 2x_3 < 6 \\ x_2 + x_3 = 2x_1^2 - 1 \\ x_1 - x_3^2 + 4x_2 > 0 \end{cases}$$

解　令　$g_1(x) = 6 + x_2^2 - 2x_1 x_3 - 2x_3$,

$g_2(x) = x_1 - x_3^2 + 4x_2$, $h(x) = x_2 + x_3 - 2x_1^2 + 1$,

则增广函数为

$$F(x,\sigma) = 3x_1 + x_3 x_1^3 - 2x_2^2 + \sigma \min{}^2\{6 + x_2^2 - 2x_1 x_3 - 2x_3 , 0\}$$

$$+ \sigma \min{}^2\{x_1 - x_3^2 + 4x_2 , 0\}$$

$$+ \sigma (x_2 + x_3 - 2x_1^2 + 1)^2 .$$

例 2　用外点法求解非线性规划问题

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

$$s.t. \quad \begin{cases} g_1(x) = x_1 \geq 1 \\ g_2(x) = x_2 \geq 1 \end{cases}$$

解　构造增广函数

$$F(x,\sigma) = x_1^2 + x_2^2 + \sigma[\min{}^2\{x_1 - 1, 0\} + \min{}^2\{x_2 - 1, 0\}]$$

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

利用必要条件求解 $\min\limits_{x\in R^2} \quad F(x,\sigma)$.

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

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

令 $\dfrac{\partial F}{\partial x_1} = 0, \dfrac{\partial F}{\partial x_2} = 0$.

$$解得: x_1 = \begin{cases} \dfrac{\sigma}{1+\sigma}, & x_1 < 1, x_2 \ge 1 \\ \dfrac{\sigma}{1+\sigma}, & \boxed{x_1 < 1, x_2 < 1} \end{cases} \qquad x_2 = \begin{cases} \dfrac{\sigma}{1+\sigma}, & x_1 \ge 1, x_2 < 1 \\ \dfrac{\sigma}{1+\sigma}, & \boxed{x_1 < 1, x_2 < 1} \end{cases}$$

综上可得当$x_1 < 1,\ x_2 < 1$ 时有最优解 $x^*(\sigma) = (\dfrac{\sigma}{1+\sigma}, \dfrac{\sigma}{1+\sigma})^T$.

当 $\sigma \to +\infty$ 时，有 $x^*(\sigma) \to x^* = (1,1)^T$.

x^* 即为所求非线性规划问题的最优解。

(6) 算法收敛性

结论1 若点列$\{x^{(k)}\}$是由外点法产生的，则 有

$$\varphi_{k+1}(x^{(k+1)}) \geq \varphi_k(x^{(k)})$$

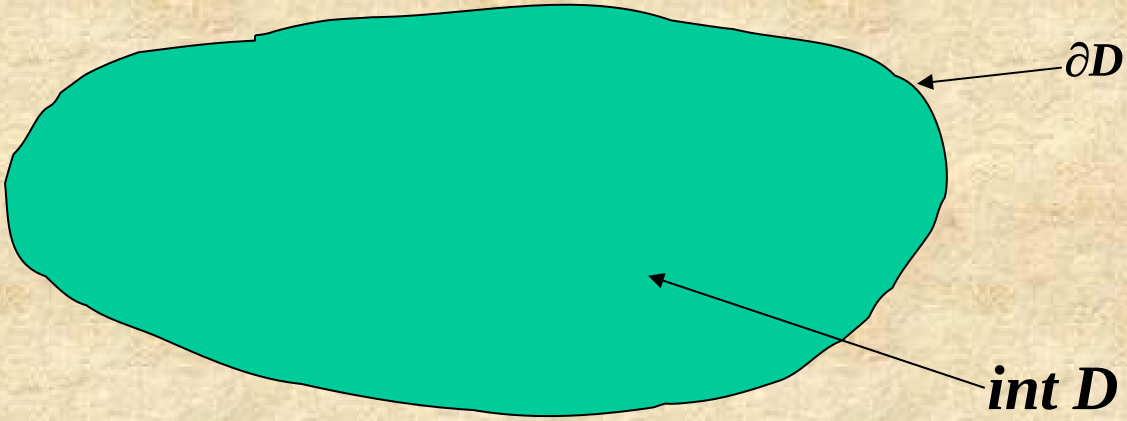
$$p(x^{(k+1)}) \leq p(x^{(k)})$$

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

结论2 设 $f(x)$、$g_j(x)(j=1,\cdots, m)$和$h_k(x)(k=1,\cdots, l)$ 都是 R^n 上的连续函数, 则由外点法产生的点列$\{x^{(k)}\}$的任意极限点一定是所 求问题的极小点。

5、内点法（障碍函数法）

（1）集合结构



$$D=\partial D\cup int\ D=\text{边界点集}\cup\text{内点集}$$

$$\partial D=\{x\,|\,\text{至少存在一个 } j,\ \text{使得}\quad g_j(x)=0\};$$

$$\text{int}\ D=\{x\,|\,g_j(x)<0\ ,\forall\ j\}\text{。}$$

（2）算法思想

内点法（障碍函数法）的迭代点是在可行域点集内部移动的，对接近可行域边界上的点施加越来越大的惩罚，对可行域边界上的点施加无限大的惩罚，这好比边界是一道障碍物，阻碍迭代点穿越边界。

内点法要求可行点集的内点集合非空，否则算法无法运行。这样一来内点法只对不等式约束的优化问题才可能有效。

(3) 算法分析

考虑如下优化问题:

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

转化为无约束优化问题:

$$\min_{x \in R^n} \psi_k(x) = f(x) + \mu_k q(x)$$

构造函数 $\psi_k(x) = f(x) + \mu_k q(x)$,

其中 $q(x)$ 满足:

(1) $q(x) > 0 \qquad if\ \ x \in \operatorname{int} D$

(2) $q(x) \uparrow +\infty \quad if\ \ x \to \partial D$

$$\mu_1 > \mu_2 > \cdots > \mu_k \downarrow 0 .$$

例如：$q(x)=\sum_{j=1}^{m}\frac{1}{g_j(x)}$ 或 $q(x)=\sum_{j=1}^{m}\frac{1}{g_j^2(x)}$ 等；

$$q(x)=\sum_{j=1}^{m}\ln\frac{1}{g_j(x)} \quad \text{或} \quad q(x)=-\sum_{j=1}^{m}\ln g_j(x) \text{ 等。}$$

通常称：

$\psi_k(x)$: 增广函数，$q(x)$: 障碍函数，

μ_k: 惩罚因子，$\mu_k q(x)$: 惩罚项。

（4）算法步骤（内点法）：

$step1.$ 给定初始点 $x^{(0)} \in \operatorname{int} D$，初始惩罚因子 $\mu_1 > 0$,（可取 $\mu_1 = 1$), 精度 $\varepsilon > 0, k := 0$。

$step\,2.$ 以 $x^{(k)}$ 为初始点，求解无约束 优化问题

$$\min_{x \in R^n} \psi_k(x) = f(x) + \mu_k q(x) = f(x) + \mu_k \sum_{j=1}^{m} \frac{1}{g_j(x)}$$

得到其最优解 $x^*(\mu_k)$，记为 $x^{(k+1)}$。

$step\,3.$ 如果 $\mu_k q(x^{(k+1)}) \le \varepsilon$, 则 $x^{(k+1)}$ 就是问题（A）: $\min\limits_{x \in D} f(x)$ 的最优解，$stop$;否则转 $step\,4$。

$step\,4.$ 给定 $\mu_{k+1} < \mu_k$（可取 $\mu_{k+1} = \beta\mu_k$ 这里 $\beta < 1$为惩罚因子的缩小系数），$k := k + 1$,转 $step\,2$。

例 3　写出用内点法求解时下列非线性规划问题的增广函数

$$\min \quad f(x) = x_1 \sin x_2 + 2x_2^2 - 2x_2 x_3$$

$$s.t. \quad \begin{cases} 2x_1^2 + x_2 - 4x_3 < 2 \\ x_2 + 2x_3 \geq 3x_1^2 - 1 \\ x_1 \geq 1 \end{cases}$$

解　令　$g_1(x) = 2 + 4x_3 - 2x_1^2 - x_2$，$g_2(x) = x_2 + 2x_3 - 3x_1^2 + 1,$

$g_3(x) = x_1 - 1,$

则增广函数为

$$\psi(x) = x_1 \sin x_2 + 2x_2^2 - 2x_2 x_3 + \frac{u}{2 + 4x_3 - 2x_1^2 - x_2} + \frac{u}{x_2 + 2x_3 - 3x_1^2 + 1} + \frac{u}{x_1 - 1}$$

或者

$$\psi(x) = x_1 \sin x_2 + 2x_2^2 - 2x_2 x_3 - u\ln(2 + 4x_3 - 2x_1^2 - x_2)$$

$$- u\ln(x_2 + 2x_3 - 3x_1^2 + 1) - u\ln(x_1 - 1)$$

例4 试用内点法求解如下优化问题

$$\min f(x)=\frac{1}{3}x^3$$

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

解：构造增广函数 $\psi_k(x)$ 如下：

$$\psi_k(x)=\frac{1}{3}x^3+\mu_k\frac{1}{x-1}$$

由 $\dfrac{d\psi_k(x)}{dx}=x^2-\dfrac{\mu_k}{(x-1)^2}=0$ 可得：

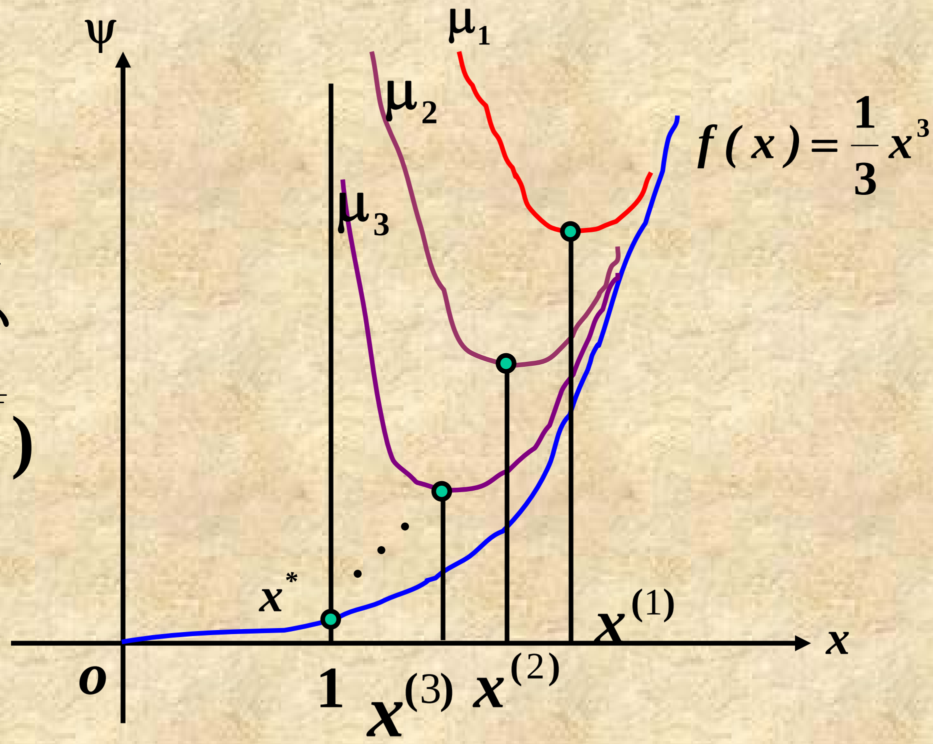
$$x(x-1)=\pm\sqrt{\mu_k}$$

因为 $x\geq 1$ 所以 $x(x-1)=\sqrt{\mu_k}$。

因此 $x=\frac{1}{2}(1\pm\sqrt{1+4\sqrt{\mu_k}})$

因为负根不满足条件，所以

$$x^{(k+1)}=\frac{1}{2}(1+\sqrt{1+4\sqrt{\mu_k}})$$



由此可以得到：

$$x^*=\lim_{k\to\infty}x^{(k+1)}=\lim_{k\to\infty}\frac{1}{2}(1+\sqrt{1+4\sqrt{\mu_k}})=1$$

例 5　用内点法求解下列问题

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

$$s.t. \quad \begin{cases} x_2 - x_1^2 \geq 1 \\ x_2 \leq 2 \end{cases}$$

解　构造增广函数

$$G(x,r) = x_1^2 + 2x_2 - r\ln(x_2 - x_1^2 - 1) - r\ln(2 - x_2)$$

利用必要条件求解 $\min\limits_{x\in \text{int}\, D} \quad G(x,r)$.

令 $$\frac{\partial G}{\partial x_1} = 2x_1 + \frac{2x_1 r}{(x_2 - x_1^2 - 1)} = 0,$$

因为 $x_2 - x_1^2 > 1,$　　解得 $x_1 = 0$.

令 $$\frac{\partial G}{\partial x_2}=2-\frac{r}{(x_2-x_1^2-1)}+\frac{r}{2-x_2}=0,$$

即 $$2-\frac{r}{(x_2-1)}+\frac{r}{2-x_2}=0,$$

解得 $$x_2=\frac{3+r\pm\sqrt{r^2+1}}{2}$$

因为 $x_2=\dfrac{3+r+\sqrt{r^2+1}}{2}>\dfrac{3+r+1}{2}>2,$ 舍去此解。

所以 $$x_2=\frac{3+r-\sqrt{r^2+1}}{2}.$$

当 $r\to 0$ 时，点 $(0,\dfrac{3+r-\sqrt{r^2+1}}{2})^T\to(0,1)^T$ 。

所以原问题最优解为 $x^*=(0,1)^T$ 。

(5) 算法收敛性

结论1 若点列$\{x^{(k)}\}$是由内点法产生的，则 有

$$\psi_{k+1}(x^{(k+1)}) \le \psi_k(x^{(k)}).$$

结论 2 设 $f(x)$、$g_j(x)$($j=1,\cdots,m$)都是 R^n 上的连续函数，则由内点法产生的点列 $\{x^{(k)}\}$的任意极限点一定是所求问题 的极小点。