Optimal Theory and Method

最优化理论与方法

翟春杰

杭州电子科技大学 自动化学院

科技馆512

Email: cjzhai@hdu.edu.cn

约束非线性优化方法

■ 直接法

- ➢ 在满足不等式约束的可行设计区域内直接求出问题的约束最优解
- ➢ **可行方向法**

■ 间接法

- ➢ 将约束优化问题转化为一系列无约束优化问题来解的一种方法
- ➢ 惩罚函数法

■ 外点法　　**■ 内点法**

目录

- 外点法（考）
- 内点法（考）

目录

- 外点法（考）
- 内点法（考）

■ 外点法

➢外点法的基本思想

➢外点法的惩罚函数构造

➢外点法的特点

■ 外点法

➢基本思想

考虑约束问题

$$\min\ f(x)$$

$$\text{s. t}\quad g_i(x) \geq 0,\quad i=1,2,\cdots,m$$

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

其中$f(x)$, $g_i(\mathrm{x})$, $h_j(\mathrm{x})$是E^n上的连续函数

由目标函数和约束函数组成辅助函数，把原来的约束问题转化为极小化辅助函数的无约束问题

- 在可行点辅助函数值等于原来的目标函数值
- 在不可行点，辅助函数值等于原来的目标函数值加上一个很大的正数

■ 外点法

➢惩罚函数的构造

对于等式约束问题

$$\min\ f(x)$$
$$\text{s. t}\ h_j(\mathbf{x}) = 0\ \ j=1,2,\cdots,l \qquad (13.1.2)$$

可定义**辅助函数**

$$F_1(x,\sigma) = f(x) + \sigma \sum_{j=1}^{l} h_j^2(\mathbf{x}) \qquad (13.1.3)$$

参数σ是**很大的正数**

转化为无约束问题 $\min F_1(x,\sigma)$ (13.1.4)

- ➢ (13.1.4)的最优解必使得$h_j(\mathbf{x})$ **接近零**;
- ➢ 问题（13.1.4）的解即是问题（13.1.2）的**近似解**

■ 外点法

➢惩罚函数的构造

对于不等式约束问题

$$\min\ f(x)$$
$$\text{s. t}\ g_i(x) \ge 0\ \ i=1,2,\cdots,m \qquad (13.1.2)$$

定义函数

$$F_2(x,\sigma) = f(x) + \sigma \sum_{i=1}^{m} [max\{0, -g_i(x)\}]^2$$

σ是很大的正数

$$max\{0, -g_i(x)\} = \begin{cases} 0, & \text{当}x\text{为可行点时} \\ -g_i(x), & \text{当}x\text{不是可行点时} \end{cases}$$

转化为无约束问题 $\min\ F_2(x,\sigma)$

→ **求得原约束问题的近似解**

■ 外点法

➢惩罚函数的构造

对于一般情形的约束优化问题

$$\min\ f(x)$$

$$\text{s.t}\ \ g_i(x) \geq 0,\quad i=1,2,\cdots,m$$

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

定义函数

$$F(x,\sigma)= f(x)+\sigma P(x)$$

其中P(x)具有下列形式

$$P(x)=\sum_{i=1}^{m}\phi(g_i(x)) + \sum_{j=1}^{l}\varphi(h_j(x))$$

典型取法

$$\phi = [\max\{0,-g_i(x)\}]^\alpha$$

$$\varphi = \left|h_j(x)\right|^\beta$$

通常取$\alpha = \beta = 2$

■ 外点法

➢惩罚函数的构造

这样，把一般约束优化问题转化为无约束优化问题

$$\min \quad F(x,\sigma)=f(\mathrm{x})+\sigma P(\mathrm{x})$$

其中σ是很大的正数， P(x)是连续函数

$\sigma P(\mathrm{x})$称为**惩罚项**

σ**称为惩罚因子**

$F(x,\sigma)$称为**罚函数**

求解无约束优化问题能够得到一般约束优化问题的近似解，而且σ越大，近似解越好

■ 外点法

➢惩罚函数的构造

用外点法求解规划问题

$$\begin{aligned} \min \quad & x \\ \text{s.t} \quad & x-2\geq 0 \end{aligned}$$

解: **惩罚项:**

$$P(x)=[max\{0,-g_i(x)\}]^2=\begin{cases} 0, \text{当} x\geq 2 \\ (x-2)^2, \ \text{当} x<2 \end{cases}$$

罚函数:

$$F(x,\sigma)=x+\sigma P(x)$$

■ 外点法

➢惩罚函数的构造

求解下列无约束问题，求得原约束问题的近似解

$$\min \quad x+\sigma P(x)$$

解析方法解无约束问题

$$\frac{dF}{dx}=\begin{cases} 1, \text{当} x \geq 2 \\ 1+2\sigma(x-2), \text{当} x < 2 \end{cases}$$

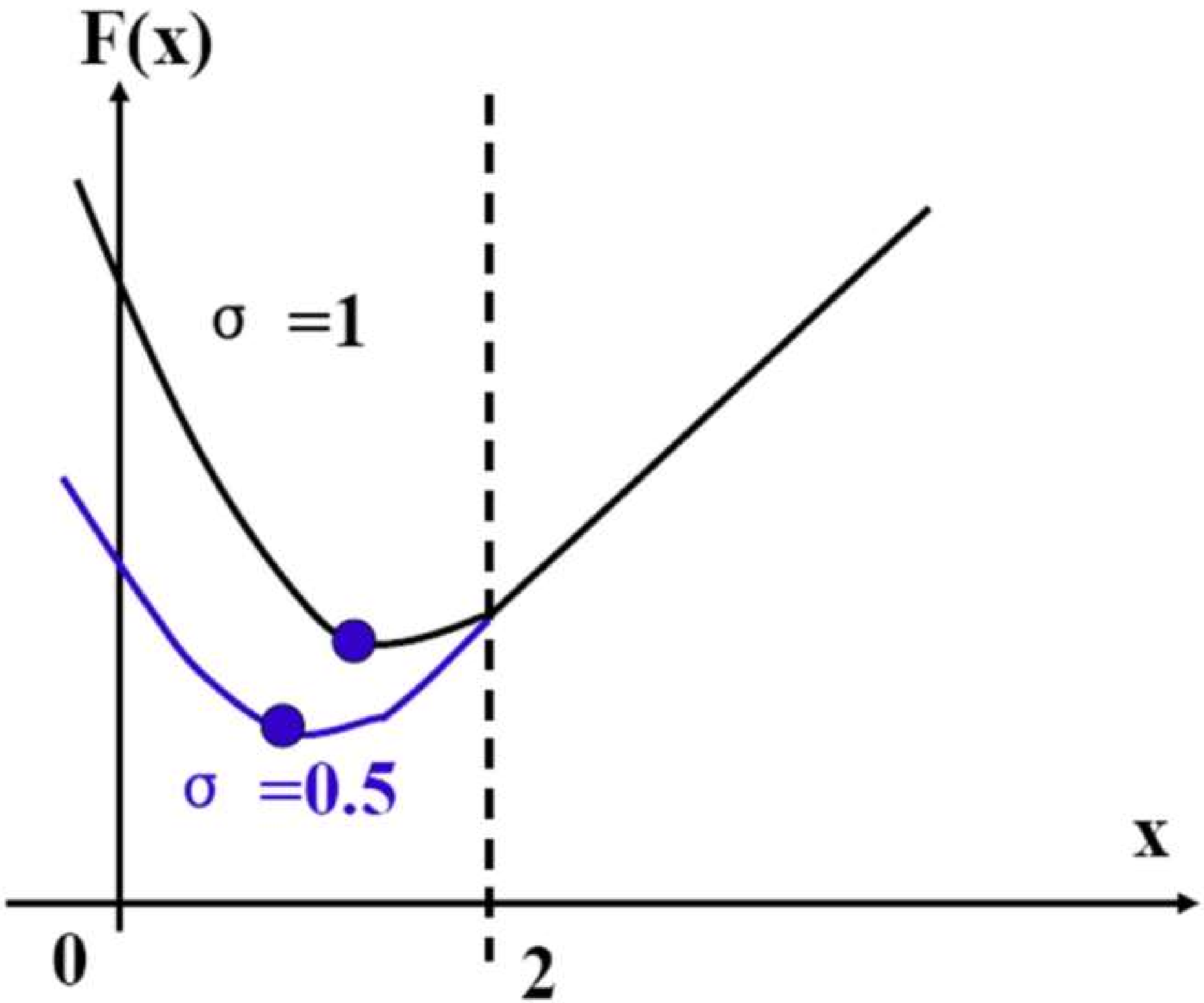
令$\frac{dF}{dx}=0$， $\bar{x}_\sigma = 2 - \frac{1}{2\sigma}$

显然，σ越大，$\bar{x}_\sigma$越接近（13.1.11）的最优解

$\sigma \to +\infty$时，$\bar{x}_\sigma \to \bar{x}=2$

■ 外点法

➢惩罚函数的构造



$\overline{x}=2$

■ 外点法

➢惩罚函数的构造

用外点法求解规划问题

$$\min \quad f(x) \triangleq (x_1-1)^2+x_2^2 \qquad (13.1.11)$$

$$\text{s.t} \quad g(x)=x_2-1\geq 0$$

解:

罚函数：

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

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

■ 外点法

➢惩罚函数的构造

解析方法

$$\frac{\mathrm{dF}}{\mathrm{d}x_1} = 2(x_1 - 1) = 0$$

$$\frac{\mathrm{dF}}{\mathrm{d}x_2} = \begin{cases} 2x_2, & 当x_2 \geq 1 \\ 2x_2 + 2\sigma(x_2 - 1), & 当x_2 < 1 \end{cases} = 0$$

$$x_2 = \frac{\sigma}{1+\sigma} \qquad \bar{x} = \begin{bmatrix} 1 \\ \frac{\sigma}{1+\sigma} \end{bmatrix}$$

令$\sigma \to +\infty$时，$\bar{x}_\sigma \to \bar{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

■ 外点法

➢特点

➢ 无约束问题的最优解$\bar{x}$，往往不满足原来的约束条件，它是从可行域外部趋向$\bar{x}$的

➢ $\mathrm{F}(x,\sigma)$也称为外罚函数，相应的最优化方法也称为外点法或外罚函数法

➢ 罚因子σ的选择十分重要

如果σ过大，增加计算上的困难

如果σ太小，罚函数的极小点远离约束问题的最优解

目录

- 外点法（考）
- 内点法（考）

■ 内点法

➢ 内点法的基本思想

➢ 内点法的障碍函数构造

➢ 内点法的特点

■ 内点法

➢基本思想

内点法在迭代中总是**从内点出发**，并保持在**可行域内部**进行搜索。

适用于下列**只有不等式约束的问题**：

$$\min\ f(x)$$

$$\text{s.t.}\ \ g_i(x) \geq 0\ \ i=1,2,\cdots,m$$

$f(x)$, $g_i(x)$是E^n上的连续函数

将可行域记作 $S=\{x|\ g_i(x) \geq 0, i=1,2,\cdots,m\}$

■ 内点法

➢障碍函数

保持迭代点含于可行域内部的方法是定义障碍函数

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

其中B(x)是连续函数，当点x趋向可行域边界时， $B(x) \to +\infty$

两种最重要的形式

$$B(x) = \sum_{i=1}^{m} \frac{1}{g_i(x)}$$ 倒数形式

$$B(x) = -\sum_{i=1}^{m} \log g_i(x)$$ 对数形式（常用）

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

■ 内点法

➢障碍函数

r是很小的正数，当x趋向边界时，函数$G(x,r)\to+\infty$；否则，由于r很小，则函数$G(x,r)$的取值近似f(x)。

求解下列问题得到原优化问题的近似解：

$$\min \quad G(x,r)$$
$$\text{s. t.} \quad x\in \text{int}S$$

由于B(x)的存在，在可行域边界形成“围墙”，因此解$\overline{x_r}$必含于可行域的内部。

由于B(x)的阻挡作用是自动实现的，因此从计算的观点看，可当作无约束问题来处理

■ 内点法

➢障碍函数

$$B(\mathbf{x}) = \sum_{i=1}^{m} \frac{1}{g_i(\mathbf{x})}$$ **倒数形式**

$$B(\mathbf{x}) = -\sum_{i=1}^{m} \log g_i(\mathbf{x})$$ **对数形式**

用内点法求解规划问题

$$\min \quad \frac{1}{12}(x_1+1)^3 + x_2$$

$$\text{s. t} \quad x_1 - 1 \geq 0$$

$$x_2 \geq 0$$

解： 罚函数：

$$G(x, r_k) = \frac{1}{12}(x_1+1)^3 + x_2 + r_k\left(\frac{1}{x_1-1} + \frac{1}{x_2}\right)$$

■ 外点法

➢障碍函数

用解析法求解

$$\min\ G(x, r_k)$$
$$\text{s. t.}\quad x \in S$$

$$\frac{dG}{dx_1}=\frac{1}{4}(x_1+1)^2-\frac{r_k}{(x_1-1)^2}=0$$

$$\frac{dG}{dx_2}=1-\frac{r_k}{x_2^2}=0$$

$$\bar{x}_{r_k}=(x_1, x_2)=\left(\sqrt{1+2\sqrt{r_k}}, \sqrt{r_k}\right)$$

$r_k \to 0$时， $\bar{x}_{r_k} \to \bar{x}=(1, 0)$，$\bar{x}$是问题的最优解

■ 作业

413：

1(4)、3(2)

目录

- **Lagrange方法**
- **起作用集方法**

目录

- Lagrange方法
- 起作用集方法

■ Lagrange方法

- 面向什么规划问题
- Lagrange函数的构造
- Lagrange方法的最优解

■ Lagrange方法

➢面向二次规划

考虑二次规划问题

$$\min \frac{1}{2}x^T Hx + c^T x$$

$$\text{s.t. } Ax = b \qquad (14.1.1)$$

其中**H**是**n**阶对称矩阵，$\boldsymbol{A}$是**m**×**n**矩阵，$\boldsymbol{x} \in \mathbf{R}^{\mathbf{n}}$，$\boldsymbol{b}$是**m**维列向量

■ Lagrange方法

➢Lagrange函数

下面推导用Lagrange乘子法求解二次规划问题的公式

定义Lagrange函数

$$L(x,\lambda) = \frac{1}{2}x^T Hx + c^T x - \lambda^T (Ax - b)$$

令 $\nabla_x L(x,\lambda)=0,\ \nabla_\lambda L(x,\lambda)=0$

得到方程组 $$Hx + c - A^T\lambda = 0$$

$$-Ax + b = 0$$

将此方程组写成

$$\begin{bmatrix} H & -A^T \\ -A & 0 \end{bmatrix}\begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ -b \end{bmatrix}$$

Lagrange矩阵

■ Lagrange方法

➢Lagrange函数

设上述Lagrange矩阵可逆，则可以表示为

$$\begin{bmatrix} H & -A^T \\ -A & 0 \end{bmatrix}^{-1} = \begin{bmatrix} Q & -R^T \\ -R & S \end{bmatrix}$$

由式

$$\begin{bmatrix} H & -A^T \\ -A & 0 \end{bmatrix}\begin{bmatrix} Q & -R^T \\ -R & S \end{bmatrix} = I_{m+n}$$

推得

$$\begin{cases} HQ + A^TR = I_n \\ -HR^T - A^TS = 0_{m+n} \\ -AQ = 0_{m+n} \\ AR^T = I_m \end{cases}$$

■ Lagrange方法

➢Lagrange方法的最优解

$$\begin{cases} HQ + A^TR = I_n \\ -HR^T - A^TS = 0_{m+n} \\ -AQ = 0_{m+n} \\ AR^T = I_m \end{cases}$$

假设逆矩阵H^{-1}存在，得到矩阵Q,R,S的表达式

$$\begin{cases} \mathbf{Q} = H^{-1} - H^{-1}A^T\left(AH^{-1}A^T\right)^{-1}AH^{-1} & (14.1.4) \\ \mathbf{R} = (AH^{-1}A^T)^{-1}AH^{-1} & (14.1.5) \\ \boldsymbol{S} = -\left(AH^{-1}A^T\right)^{-1} & (14.1.6) \end{cases}$$

$\boldsymbol{Q}, \boldsymbol{s}$对称矩阵

$$\begin{bmatrix} H & -A^T \\ -A & 0 \end{bmatrix} \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} -c \\ -b \end{bmatrix}$$

将上式等号两端乘以Lagrange矩阵的逆，则得到问题的解

$$\begin{cases} \bar{x} = -Qc + R^Tb & (14.1.7) \\ \bar{\lambda} = Rc - Sb & (14.1.8) \end{cases}$$

■ Lagrange方法

➢Lagrange方法的最优解

$$\begin{cases}\bar{\boldsymbol{x}} = -\boldsymbol{Qc} + \mathbf{R}^{\mathrm{T}}\boldsymbol{b} & (14.1.7)\\ \bar{\boldsymbol{\lambda}} = \boldsymbol{Rc} - \boldsymbol{Sb} & (14.1.8)\end{cases}$$

$\bar{\boldsymbol{x}}$和$\bar{\boldsymbol{\lambda}}$的另一种表达式：

设$\mathbf{x}^{(\mathbf{k})}$是任一可行解，即$\mathbf{x}^{(\mathbf{k})}$满足$\mathbf{Ax}^{(\mathbf{k})} = \boldsymbol{b}$. 在此点目标函数的梯度

$$\mathbf{g_k} = \nabla\mathbf{f}\left(\mathbf{x}^{(\mathbf{k})}\right) = \mathbf{Hx}^{(\mathbf{k})} + \mathbf{c}$$

利用$\mathbf{x}^{(\mathbf{k})}$和$\mathbf{g_k}$，可将（14.1.7）和（14.1.8）改写成

$$\begin{cases}\bar{\boldsymbol{x}} = \mathbf{x}^{(\mathbf{k})} - \boldsymbol{Q}\mathbf{g_k} & (14.1.9)\\ \bar{\boldsymbol{\lambda}} = \boldsymbol{R}\mathbf{g_k} & (14.1.10)\end{cases}$$

■ Lagrange方法

$$\min \frac{1}{2}x^T Hx + c^T x$$
$$\text{s.t. } Ax = b$$

➢例子

$$\min \quad x_1^2 + 2x_2^2 + x_3^2 - 2x_1x_2 + x_3$$
$$\text{s.t.} \quad x_1 + x_2 + x_3 = 4$$
$$2x_1 - x_2 + x_3 = 2$$

解：易知

$$H=\begin{bmatrix} 2 & -2 & 0 \\ -2 & 4 & 0 \\ 0 & 0 & 2 \end{bmatrix}, c=\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & 1 \end{bmatrix}, b = \begin{bmatrix} 4 \\ 2 \end{bmatrix},$$

H的逆矩阵H^{-1}=
$$\begin{bmatrix} 1 & \frac{1}{2} & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$$

■ Lagrange方法

➢例子

$$\begin{cases} \bar{x} = -Qc + R^{T} b \\ \bar{\lambda} = Rc - Sb \end{cases}$$

由（14.1.4）至（14.1.6）得

$$Q=\frac{4}{11}\begin{bmatrix} \frac{1}{2} & \frac{1}{4} & -\frac{3}{4} \\ \frac{1}{4} & \frac{1}{8} & -\frac{3}{8} \\ -\frac{3}{4} & -\frac{3}{8} & \frac{9}{8} \end{bmatrix} R=\frac{4}{11}\begin{bmatrix} \frac{3}{4} & \frac{7}{4} & \frac{1}{4} \\ \frac{3}{4} & -1 & \frac{1}{4} \end{bmatrix} S=-\frac{4}{11}\begin{bmatrix} 3 & -\frac{5}{2} \\ -\frac{5}{2} & 3 \end{bmatrix}$$

把Q，R代入（14.1.7），得到问题的最优解

$$\bar{x}=\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}=\begin{bmatrix} \frac{21}{11} \\ \frac{43}{22} \\ \frac{3}{22} \end{bmatrix}$$

■ Lagrange方法

➢例子

$$\bar{x}=\begin{bmatrix}x_1\\x_2\\x_3\end{bmatrix}=\begin{bmatrix}\frac{21}{11}\\\frac{43}{22}\\\frac{3}{22}\end{bmatrix}$$

$$Q=\frac{4}{11}\begin{bmatrix}\frac{1}{2} & \frac{1}{4} & -\frac{3}{4}\\ \frac{1}{4} & \frac{1}{8} & -\frac{3}{8}\\ -\frac{3}{4} & -\frac{3}{8} & \frac{9}{8}\end{bmatrix},\quad R=\frac{4}{11}\begin{bmatrix}\frac{3}{4} & \frac{7}{4} & \frac{1}{4}\\ \frac{3}{4} & -1 & \frac{1}{4}\end{bmatrix}$$

$$H=\begin{bmatrix}2 & -2 & 0\\ -2 & 4 & 0\\ 0 & 0 & 2\end{bmatrix},c=\begin{bmatrix}0\\0\\1\end{bmatrix}$$

$$\begin{cases}\bar{x}=\mathbf{x}^{(k)}-Q\mathbf{g}_k\\ \bar{\lambda}=R\mathbf{g}_k\\ \mathbf{g}_k=\nabla f\left(\mathbf{x}^{(k)}\right)=\mathbf{H}\mathbf{x}^{(k)}+\mathbf{c}\end{cases}$$

目录

- Lagrange方法
- 起作用集方法

■ 起作用集方法

- 面向什么**规划问题**
- 起作用集方法的**基本思想**
- 起作用集方法的**计算步骤**

■ 起作用集方法

➢面向的规划问题

考虑具有不等式约束二次规划问题

$$\min \quad f(\mathbf{x}) \overset{\text{def}}{=} \frac{1}{2}x^T H x + c^T x$$

$$\text{s. t} \quad Ax \geq b \qquad (14.2.1)$$

其中H是n阶对称矩阵，c是n维列向量，A是m×n矩阵，A的秩为m, b是m维列向量, $x \in R^n$.

■ 起作用集方法

➢基本思想

由于不等式约束的出现，不能直接使用消元法和Lagrange方法求解，解决这个问题的策略之一，是用起作用集方法将它转化为求解等式约束问题。

基本思路：

在每次迭代中，以已知的可行点为起点，把在该点起作用约束作为等式约束，在此约束下极小化目标函数f(x),而其余的约束暂且不管，求得新的比较好的可行点后，再重复以上做法。

■ 起作用集方法

➢搜索方向和步长

设在第k次迭代中，已知可行点$x^{(k)}$，在该点起作用约束指标集用$I^{(k)}$表示，这时需要求解等式约束问题

$$
\begin{aligned}
\min \quad & f(x) \\
s.t \quad & a^{(i)}x=b_i,\ i\in I^{(k)}
\end{aligned}
\qquad (14.2.2)
$$

其中$a^{(i)}$是矩阵A的第i行。

$$\min \quad f(x)$$
$$\text{s. t} \quad a^{(i)}x=b_i\ ,\ i\in I^{(k)}$$

■ 起作用集方法

➢搜索方向和步长

为方便起见，现将坐标原点移至$x^{(k)}$，令

$$\delta = x - x^{(k)},$$

则$f(x)=\frac{1}{2}(\delta+x^{(k)})^T H(\delta+x^{(k)})+c^T(\delta+x^{(k)})$

$$=\frac{1}{2}\delta^T H\delta+(Hx^{(k)})^T\delta+\frac{1}{2}x^{(k)T}Hx^{(k)}+c^T\delta+c^T x^{(k)}$$

$$=\frac{1}{2}\delta^T H\delta+\nabla f(x^{(k)})^T\delta+f(x^{(k)}) \qquad (14.2.3)$$

于是问题（14.2.2）转化成求校正量$\delta^{(k)}$的问题

$$\min \quad \frac{1}{2}\delta^T H\delta+\nabla f(x^{(k)})^T\delta$$
$$\text{s. t} \quad a^{(i)}\delta=0\ ,\ i\in I^{(k)} \qquad (14.2.4)$$

求出最优解$\delta^{(k)}$

■ 起作用集方法

➢搜索方向和步长

如果$x^{(k)}+\delta^{(k)}$是可行点，且$\delta^{(k)}\neq 0$，则在第k+1次迭代中，已知点取作

$$x^{(k+1)}=x^{(k)}+\delta^{(k)}$$

如果$x^{(k)}+\delta^{(k)}$不是可行点，则令方向$d^{(k)}=\delta^{(k)}$,并沿$d^{(k)}$搜索，令$x=x^{(k)}+\alpha d^{(k)}$

确定沿$d^{(k)}$方向的步长α_k，根据保持可行性的要求，α_k的取值应使得对于每个$i\notin I^{(k)}$成立

$$a^{(i)}(x^{(k)}+\alpha d^{(k)})\geq b_i \qquad (14.2.5)$$

■ 起作用集方法

$$a^{(i)}(x^{(k)}+\alpha \mathrm{d}^{(k)}) \geq b_i$$

➢搜索方向和步长

由于$x^{(k)}$是可行点，$a^i x^{(k)} \geq \mathrm{b_i}$

当$a^i d^{(k)} \geq 0$时，对任意的非负数α_k，（14.2.5）总成立；

当$a^i d^{(k)} < 0$时，只要取正数

$$\alpha_k \leq min\left\{\frac{\mathrm{b_i}-a^i x^{(k)}}{a^i d^{(k)}}\middle|\mathrm{i} \notin \mathrm{I}^{(k)}, a^i d^{(k)} < 0\right\},$$

对于每个$\mathrm{i} \notin \mathrm{I}^{(k)}$，（14.2.5）成立。

记 $$\hat{\alpha}_k = min\left\{\frac{\mathrm{b_i}-a^i x^{(k)}}{a^i d^{(k)}}\middle|\mathrm{i} \notin \mathrm{I}^{(k)}, a^i d^{(k)} < 0\right\}$$

■ 起作用集方法

$$a^{(i)}(x^{(k)}+\alpha \mathbf{d}^{(k)}) \geq b_i$$

➢搜索方向和步长

综上，为在第k次迭代中得到较好可行点，应取

$$\alpha_k = \min\{1,\ \hat{\alpha}_k\} \qquad (14.2.6)$$

于是 $x^{(k+1)} = x^{(k)} + \alpha_k \mathbf{d}^{(k)}$

如果 $\alpha_k = \dfrac{b_p - a^p x^{(k)}}{a^p d^{(k)}} < 1 \qquad (14.2.7)$

则在点$x^{(k+1)}$，有$a^p x^{(k+1)} = a^p\left(x^{(k)} + \alpha_k \mathbf{d}^{(k)}\right) = b_p$

因此，在$x^{(k+1)}$处，$a^p \mathbf{x} \geq \mathrm{b}_p$为**起作用约束**。这时，把指标p加入$\mathrm{I}^{(k)}$，得到在$x^{(k+1)}$处的起作用**约束指标集**$\mathrm{I}^{(k+1)}$

■ 起作用集方法

➢迭代终止条件

$$\min \ f(x)$$
$$\text{s.t} \quad a^{(i)}x=b_i, \ i\in I^{(k)}$$

如果$\delta^{(k)}=0$，则$x^{(k)}$是问题14.2.2的最优解。这时应判断$x^{(k)}$是否为原约束问题的最优解：

为此，需计算起作用约束的乘子$\lambda_i^{(k)}(i\in I^{(k)})$。若$\lambda_i^{(k)}\geq 0$，则点$x^{(k)}$是问题的K-T点，也是最优解。

如果存在$q\in I^{(k)}$,使得$\lambda_q^{(k)}<0$，则$x^{(k)}$不可能是最优解。

当$\lambda_q^{(k)}<0$时，把下标q从$I^{(k)}$中删除。如果有几个乘子同时为负数，令$\lambda_q^{(k)}=\min\left\{\lambda_i^{(k)} \middle| i\in I^{(k)}\right\}$

将对应$\lambda_q^{(k)}$的约束从起作用约束集中去掉，再解转化问题

■ 起作用集方法

$$\min \quad \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$\text{s.t} \quad a^{(i)}\delta = 0, \quad i\in I^{(k)}$$

➢计算步骤

(1).给定初始可行解$x^{(1)}$，相应的起作用约束指标集为$I^{(k)}$，置k=1。

(2).求解问题（14.2.4），设其最优解为$\delta^{(k)}$，若$\delta^{(k)}=0$，则进行步骤(5)；否则，进行步骤(3)。

(3).令$d^{(k)}=\delta^{(k)}$，由（14.2.6）确定α_k，令

$$x^{(k+1)}= x^{(k)} + \alpha_k d^{(k)}$$

计算$\nabla f(x^{(k+1)})$

(4)若$\alpha_k<1$,则置$I^{(k+1)} = I^{(k)}\cup\{p\}$，k:=k+1，返回步骤（2）；若$\alpha_k=1$，记点$x^{(k+1)}$处起作用约束指标集为$I^{(k+1)}$，置k:=k+1，进行步骤（5）。

(5)用（14.1.10）计算对应起作用约束的Lagrange乘子$\lambda^{(k)}$，设 $\lambda_q^{(k)}=\min\left\{\lambda_i^{(k)}\middle| i\in I^{(k)}\right\}$，若$\lambda_q^{(k)}\ge 0$，则停止计算，得到最优解$x^{(k)}$；否则，从$I^{(k)}$中删除q，返回步骤(2)

■ 起作用集方法

$$\min \quad \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$s.t \quad a^{(i)}\delta = 0, \quad i \in I^{(k)}$$

➢例子

用起作用集方法求解下列二次规划问题：

$$\min \quad f(x) \overset{\text{def}}{=} x_1^2 - x_1x_2 + 2x_2^2 - x_1 - 10x_2$$
$$s.t \quad -3x_1 - 2x_2 \geq -6$$
$$x_1, \quad x_2 \geq 0$$

解：目标函数f(x)写成

$$f(x) = \frac{1}{2}(x_1, \quad x_2)\begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + (-1, \quad -10)\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

由此可知

$$H = \begin{bmatrix} 2 & -1 \\ -1 & 4 \end{bmatrix}, \quad c = \begin{bmatrix} -1 \\ -10 \end{bmatrix}$$

■ 起作用集方法

$$\min \ \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$

$$\text{s. t} \quad a^{(i)}\delta = 0\ ,\ \ i \in I^{(k)}$$

➢例子

取初始可行点$x^{(1)} = (0,0)^T$，在点$x^{(1)}$，起约束作用指标集$I^{(1)}$ ={2，3}，求解相应的转化问题

$$\min \ f(x) \overset{\text{def}}{=} \delta_1^2 - \delta_1\delta_2 + 2\delta_2^2 - \delta_1 - 10\delta_2$$

$$\text{s. t} \ \ \delta_1 = 0$$

$$\delta_2 = 0$$

得解$\delta^{(1)} = (0,0)^T$

因此$x^{(1)}$是相应转化问题的最优解，计算Lagrange乘子，由$I^{(1)}$ ={2，3}知

$$g_k = \nabla f\left(x^{(k)}\right) = Hx^{(k)} + c$$

$$R = (AH^{-1}A^T)^{-1}AH^{-1}$$

$$\bar{\lambda} = Rg_k$$

A= $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$，b= $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$

算得乘子$\lambda_2^{(1)} = -1$，$\lambda_3^{(1)} = -10$，$x^{(1)}$不是例题的最优解。

■ 起作用集方法

$$\min \ \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$\text{s.t} \quad a^{(i)}\delta = 0, \ i \in I^{(k)}$$

➢例子

将$\lambda_3^{(1)}$对应的约束，即原来问题的第3个约束，从起作用约束集中去掉，置$I^{(1)}=\{2\}$，再解相应的转化问题

$$\min \ f(x) \overset{\text{def}}{=} \delta_1^2 - \delta_1\delta_2 + 2\delta_2^2 - \delta_1 - 10\delta_2$$
$$\text{s. t.} \ \delta_1 = 0$$

得解$\delta^{(1)} = (0, \frac{5}{2})^T$。

由于$\delta^{(1)} \neq 0$，需要计算α_1。由于 $\alpha_1 = \min\{1, \frac{6}{5}\} = 1$

令$x^{(2)} = x^{(1)} + \alpha_1\delta^{(1)} = \begin{bmatrix}0\\0\end{bmatrix} + 1\cdot\begin{bmatrix}0\\\frac{5}{2}\end{bmatrix} = \begin{bmatrix}0\\\frac{5}{2}\end{bmatrix}$,算出$\nabla f(x^{(2)}) = (-\frac{7}{2}, 0)^T$

■ 起作用集方法

$$\min \quad \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$\text{s.t} \quad a^{(i)}\delta = 0, \quad i \in I^{(k)}$$

➢例子

由于$\alpha_1=1$，置$I^{(2)}=\{2\}$，在点$x^{(2)}$处，计算相应的Lagrange乘子。

此时$A = (1,0), b = 0$

算得$\lambda_2^{(2)} = -\frac{7}{2}$，因此$x^{(2)}$不是问题的最优解。

将指标2从$I^{(2)}$中剔除，于是$I^{(2)}=\phi$，再解相应的转化问题

$$\min \quad \delta_1^2 - \delta_1\delta_2 + 2\delta_2^2 - \frac{7}{2}\delta_1$$

得解$\delta^{(2)} = (2, \frac{1}{2})^T$。

由于$\delta^{(2)} \neq 0$，需要计算α_2。由（14.2.6），有

$$\alpha_2 = \min\{1, \frac{1}{7}\} = \frac{1}{7}$$

■ 起作用集方法

$$\min \quad \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$s.t \quad a^{(i)}\delta = 0\ ,\ i \in I^{(k)}$$

➢例子

令 $x^{(3)} = x^{(2)} + \alpha_2\delta^{(2)} = \begin{bmatrix} 0 \\ \frac{5}{2} \end{bmatrix} + \frac{1}{7}\cdot\begin{bmatrix} 2 \\ \frac{1}{2} \end{bmatrix} = \begin{bmatrix} \frac{2}{7} \\ \frac{18}{7} \end{bmatrix}$

算出$\nabla f(x^{(3)}) = (-3, 0)^T$

在点$x^{(3)}$，第1个约束起作用约束，这时$I^{(3)}=\{1\}$，解相应的转化问题

$$\min \quad \delta_1^2 - \delta_1\delta_2 + 2\delta_2^2 - \frac{7}{2}\delta_1$$
$$s.\ t \quad -3\delta_1 - 2\delta_2 = 0$$

得解$\delta^{(3)} = (\frac{3}{14}, -\frac{9}{28})^T$。计算$\alpha_3$:

$$\alpha_3 = \min\{1,\ 8\} = 1$$

■ 起作用集方法

$$\min \quad \frac{1}{2}\delta^T H\delta + \nabla f(x^{(k)})^T\delta$$
$$\text{s.t} \quad a^{(i)}\delta = 0, \quad i\in I^{(k)}$$

➢例子

令$x^{(4)} = x^{(3)} + \alpha_3\delta^{(3)} = \begin{bmatrix} \frac{2}{7} \\ \frac{18}{7} \end{bmatrix} + 1\cdot\begin{bmatrix} \frac{3}{14} \\ -\frac{9}{28} \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \\ \frac{9}{4} \end{bmatrix}$

算出$\nabla f(x^{(4)}) = (-\frac{9}{4}, -\frac{3}{2})^T$

在点$x^{(4)}$， $I^{(4)}=\{1\}$，计算相应的Lagrange乘子，得到$\lambda_1^{(4)} = \frac{3}{4}$， 点$x^{(4)} = (\frac{1}{2}, \frac{9}{4})^T$是所求的最优解。