优化理论与最优控制

Prof. Dr . -ing L. Wang

线性规划问题 及单纯型法

线性规划问题的标准型

$$\max \quad s = C^{T} x$$

$$\text{s.t.} \quad Ax = b$$

$$x \geq 0$$

$C = (c_1, c_2 \ldots c_n)^{T}$ 价值系数

$b_j > 0$

$b = (b_1, b_2 \ldots b_m)^{T}$ 要求向量

自由度

m 线性规划的阶数

n 线性规划的维数

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

约束方程的系数矩阵

❖ m < n

线性规划问题的标准型

$$\begin{aligned} \max \quad & s = C^{\mathrm{T}} x \\ \text{s.t.} \quad & Ax = b \\ & x \geq 0 \end{aligned}$$

特点：

(1) 目标函数求最大值

(2) 约束条件都为等式方程，且右端常数项b_i都大于或等于零

(3) 决策变量x_j为非负。

非标准型转化成标准型

若目标函数要求极小　　min　$s = C^T x$ ➡ max　$s' = -C^T x$

小于等于形式的约束　　s.t. $Ax \le b$ ➡ s.t.　$Ax +$ ж $= b$

非负的松弛变量
(slack variable)

目标函数中的
价值系数为0

大于等于形式的约束　　s.t. $Ax \ge b$ ➡ s.t.　$Ax -$ ж $= b$

可正可负的变量取值　　$x_k = x_k' - x_k''$　　$x_k' \ge 0$，　$x_k'' \ge 0$

非标准型转化成标准型

例：将下列线性规划问题化为标准形式

$$\min Z = -2x_1 + x_2 + 3x_3$$

$$\begin{cases} 5x_1 + x_2 + x_3 \le 7 \\ x_1 - x_2 - 4x_3 \ge 2 \\ -3x_1 + x_2 + 2x_3 = -5 \\ x_1, x_2 \ge 0, x_3 \text{free} \end{cases}$$

线性规划问题解的概念

s.t.
$$
\begin{aligned}
x_1 + x_2 + 2x_3 + 4x_4 + x_5 \quad &= 4 \\
x_1 + 2x_2 + 2.5x_3 + x_5 + x_6 &= 5 \\
\boxed{x_j \geq 0 \quad j = (1, 2 \ldots 6)}&
\end{aligned}
$$

n=6 维数

m=2 阶数

可行域

↑

可行解

$x^1 = (3,\ 1,\ 0,\ 0,\ 0,\ 0)$ 基本解 非退化的基本可行解 基变量 x_1, x_2 基向量 $p^1 = (1,1)^T$ $p^2 = (1,2)^T$

$x^2 = (1,\ 1,\ 0,\ 0.25,\ 1,\ 1)$

$x^3 = (0,\ 0,\ 2,\ 0,\ 0,\ 0)$

$x^4 = (5,\ 0,\ 0,\ -0.25,\ 0,\ 0)$ 基本解 退化的基本可行解 基变量 x_1, x_4

线性规划问题解的一般结论

- ✓ 所有可行解构成的集合一般是凸集，也可能是无界域
- ✓ 基本可行解与凸集的顶点相对应，最多C_n^m
- ✓ 最优解一定可以在某顶点处得到

$$\max \quad s = -x_1 + 2x_2 + 3x_3$$

$$\text{s.t.} \quad x_1 + x_2 + x_3 = 2$$

$$x_1 + 2x_2 + 4x_3 = 6$$

$$x_1 \geq 0, \quad x_2 \geq 0, \quad x_3 \geq 0$$

3 组基

$p^1 = (1,1)^T$

$p^2 = (1,2)^T$

$p^3 = (1,4)^T$

$x^{12} = (-2,\ 4,\ 0)$

$x^{23} = (0,\ 1,\ 1)$

$x^{13} = (2/3,\ 0,\ 4/3)$

$x^{23} = (0,\ 1,\ 1)$ $\quad s^{23} = 5$ *

$x^{13} = (2/3,\ 0,\ 4/3)$ $\quad s^{13} = 10/3$

单纯形法（simplex method)

基本思想

从一个已知的初始基本可行解（或者是可行域的某一顶点）出发，转换到另一个基本可行解，使目标函数逐步达到最大值。

3 个问题

初始基本可行解的产生

怎样由一个基本可行解迭代出另一个基本可行解

怎样确定可以使目标函数有较大上升的基本可行解

引例

$$\max\ f(x_1, x_2) = 40x_1 + 50x_2$$

$$
\begin{aligned}
s.\ t.\quad x_1 + 2x_2 &\le 30 \\
3x_1 + 2x_2 &\le 60 \\
2x_2 &\le 24 \\
x_1,\ x_2 &\ge 0
\end{aligned}
$$

引例

$$\max\ f(x_1, x_2, x_3, x_4, x_5) = 40x_1 + 50x_2$$

$$\begin{aligned} s.\ t.\quad & x_1 + 2x_2 + x_3 = 30 \\ & 3x_1 + 2x_2 + x_4 = 60 \\ & 2x_2 + x_5 = 24 \\ & x_1 \ldots x_5 \geq 0 \end{aligned}$$

解：(1) 确定初始可行解

n = 5 ; m=3;　　基变量3个　　　$B^{(1)} = (P_3 \ \ P_4 \ \ P_5) = I$

$$f = 0 + 40x_1 + 50x_2$$
$$x_3 = 30 - (x_1 + 2x_2)$$
$$x_4 = 60 - (3x_1 + 2x_2)$$
$$x_5 = 24 - 2x_2$$

令非基变量　$x_1 = x_2 = 0$

$$\vec{x}^{(1)} = (0, 0, 30, 60, 24)^T$$

$$f^{(1)} = 0$$

解：(2) 判定解是否最优

$$f = 0 + 40x_1 + 50x_2$$

当 x_1从0↗ 或 x_2从0↗

f从0↗

∴ $x^{(1)}$不是最优解

解：(3) 由一个基本可行解→另一个基本可行解

∵ 系数 50 ﹥ 40　选 x_2 从 0 ↗，$x_1=0$

$$x_3 = 30 - 2x_2 \geq 0 \qquad x_2 \leq 15$$

$$x_4 = 60 - 2x_2 \geq 0 \qquad x_2 \leq 30$$

$$x_5 = 24 - 2x_2 \geq 0 \qquad x_2 \leq 12$$

$$x_2 = \min(15, 30, 12) = 12$$

x_2进基变量，x_5出基变量

解：　$B^{(2)} = (P_3 \ \ P_4 \ \ P_2)$

$$f = 0 + 40x_1 + 50x_2 \quad ④$$
$$x_3 + 2x_2 = 30 - x_1 \quad ①$$
$$x_4 + 2x_2 = 60 - 3x_1 \quad ②$$
$$2x_2 = 24 - x_5 \quad ③$$

③×0.5，③代入④式，　①－③，②－③

$$f = 600 + 40x_1 - 25x_5$$
$$x_3 = 6 - x_1 + x_5$$
$$x_4 = 36 - 3x_1 + x_5$$
$$x_2 = 12 - 0.5x_5$$

令非基变量 $x_1 = x_5 = 0$　　$\vec{x}^{(2)} = (0, 12, 6, 36, 0)^T$　　$f^{(2)} = 600$

解：(2') 继续判定解是否最优

∵ 系数 $40>0$ ∴ $\vec{x}^{(2)}$不是

(3') 由一个基本可行解→另一个基本可行解

选 x_1 从0↗， $x_5=0$

$$x_3 = 6 - x_1 \geq 0$$

$$x_4 = 36 - 3x_1 \geq 0$$

$$x_2 = 12 \quad \geq 0$$

$$x_1 = \min(6, 12) = 6$$

x_1进基，x_3出基。

解： $B^{(3)} = (P_1 \ \ P_4 \ \ P_2)$

$$f = 840 - 40x_3 + 15x_5$$
$$x_1 = 6 - x_3 + x_5$$
$$x_4 = 18 + 3x_3 - 2x_5$$
$$x_2 = 12 - 0.5x_5$$

令 $x_3 = x_5 = 0$; $\vec{x}^{(3)} = (6, 12, 0, 18, 0)^T$ $f^{(3)} = 840$

(2")继续判定解是否最优

∵ 系数 15 > 0 ∴ $\vec{x}^{(3)}$ 不是

(3")由一个基本可行解→另一个基本可行解

解：　选 x_5 从0↗，　$x_3 = 0$

$$x_1 = 6 + x_5 \geq 0$$
$$x_4 = 18 - 2x_5 \geq 0$$
$$x_2 = 12 - 0.5x_5 \geq 0$$

$x_5 = \min(9, 24) = 9$

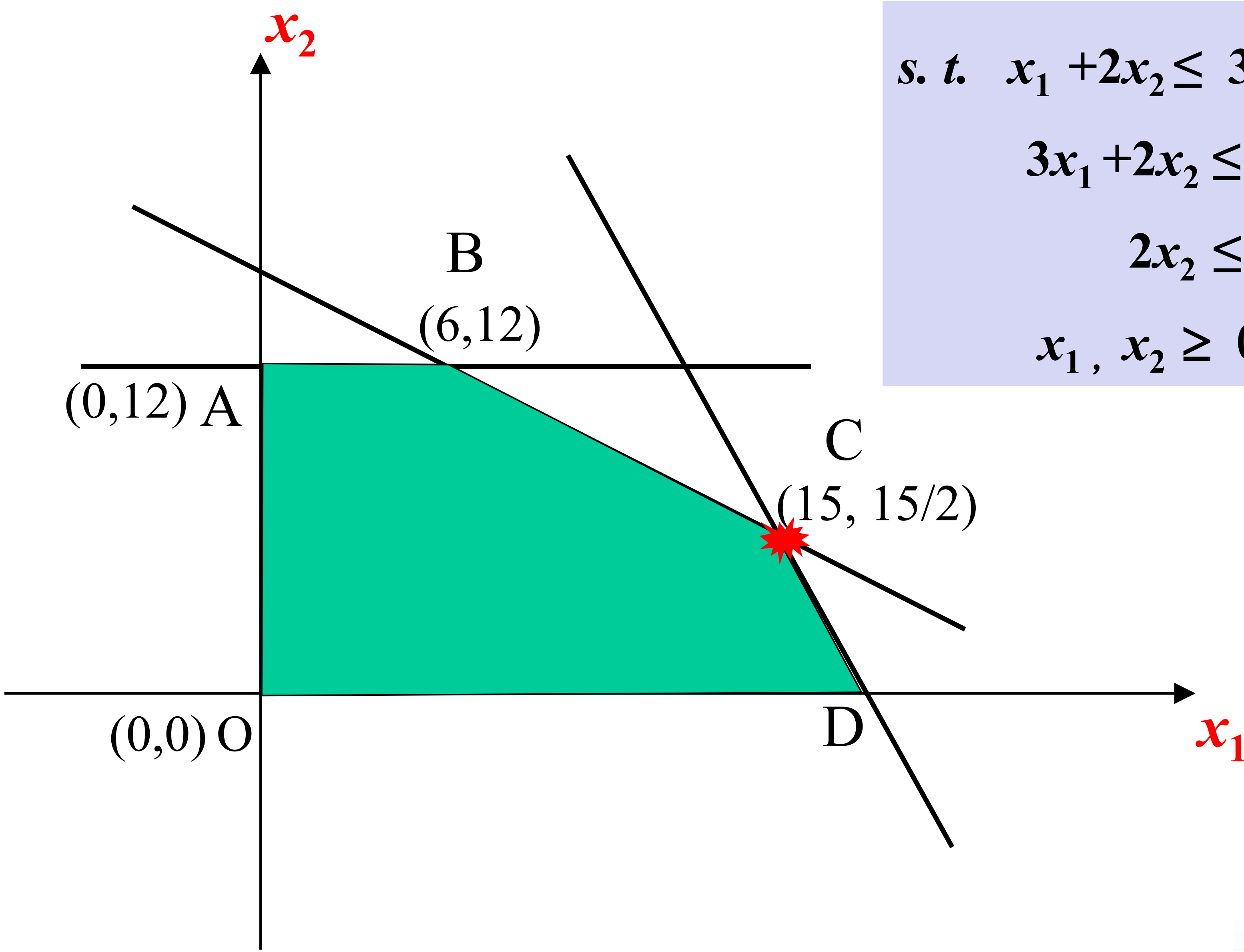
x_5进基，x_4出基。

OK!!!

$B^{(4)} = (P_1\ \ P_5\ \ P_2)$　　$f = 975 - 35/2\, x_3 - 15/2\, x_4$

$$x_1 = 15 + 1/2\, x_3 - 1/2\, x_4$$
$$x_5 = 9 + 3/2\, x_3 - 1/2\, x_4$$
$$x_2 = 15/2 - 3/4\, x_3 + 1/4\, x_4$$

令 $x_3 = x_4 = 0$　$\vec{x}^{(4)} = (15, 15/2, 0, 0, 9)^T$　　$f^{(4)} = 975$



$$
\begin{aligned}
s.\ t.\quad x_1 + 2x_2 &\le 30 \\
3x_1 + 2x_2 &\le 60 \\
2x_2 &\le 24 \\
x_1,\ x_2 &\ge 0
\end{aligned}
$$

单纯形表(只有位置，没有过程)

			40	50	0	0	0	
C_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	30	1	2	1	0	0	0
0	x_4	60	3	2	0	1	0	0
0	x_5	24	0	2	0	0	1	0
	$-s$		0	0	0	0	0	

$$\max\ f(x_1, x_2, x_3, x_4, x_5) = s = 40x_1 + 50x_2$$

$$\begin{aligned} s.\ t.\quad & 1x_1 + 2x_2 + 1x_3 = 30 \\ & 3x_1 + 2x_2 \qquad + 1x_4 = 60 \\ & \qquad 2x_2 \qquad\qquad + 1x_5 = 24 \\ & x_1 \ldots x_5 \geq 0 \end{aligned}$$

单纯形表

			40	50	0	0	0	
C_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	30	1	2	1	0	0	15
0	x_4	60	3	2	0	1	0	30
0	x_5	24	0	(2)	0	0	1	12
	$-s$	0	40	50	0	0	0	

解：(1)确定初始可行解 $\vec{x}^{(1)} = (0, 0, 30, 60, 24)^T \quad f^{(1)} = 0$

(2)判定解是否最优 $f = 0 + 40x_1 + 50x_2$

(3)由一个基本可行解→另一个基本可行解

$$
\begin{aligned}
x_3 &= 30 - 2x_2 \geq 0 & x_2 &\leq 15 \\
x_4 &= 60 - 2x_2 \geq 0 & x_2 &\leq 30 \\
x_5 &= 24 - 2x_2 \geq 0 & x_2 &\leq 12
\end{aligned}
$$

(x_2进，x_5出基变量)

单纯形表

			40	50	0	0	0	
C_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	6	(1)	0	1	0	-1	6
0	x_4	36	3	0	0	1	-1	12
50	x_2	12	0	1	0	0	1/2	
	$-s$	600	40	0	0	0	-25	

解：(1')确定初始可行解 $\vec{x}^{(2)} = (0, 12, 6, 36, 0)^T$ $f^{(2)} = 600$

(2')判定解是否最优 $f = 600 + 40x_1 - 25x_5$

(3')由一个基本可行解→另一个基本可行解

$x_3 = 6 - x_1 \geq 0$ $\quad x_1 \leq 6$

$x_4 = 36 - 3x_1 \geq 0$ $\quad x_1 \leq 12$

$x_2 = 12 \geq 0$

(x_1进，x_3出基变量)

单纯形表

			40	50	0	0	0	
C_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
40	x_1	15	1	0	-1/2	1/2	0	
0	x_5	9	0	0	-3/2	1/2	1	
50	x_2	15/2	0	1	3/4	-1/4	0	
	$-s$	975	0	0	-35/2	-15/2	0	

最优解 $x = (15, 15/2, 0, 0, 9)^T$ $f = 975$

有约束多变量问题—非线性规划问题

有约束多变量问题的最优化方法

线性等式或线性不等式 → 线性规划

非线性等式约束或非线性不等式约束 → 非线性规划

随机约束 → 随机优化

非线性有约束问题的数学模型

$$\min \quad f(x)$$

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

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

Lagrange乘子法

两类求解方法

有约束问题 转化为一系列无约束问题

有约束非线性规划问题转化为线性规划问题

可行方向法

Lagrange乘子法

$$\min \quad f(x) = -x_1 x_2$$

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

构造lagrange函数

$$L(x_1, x_2, \lambda) = f(x_1, x_2) - \lambda h(x_1, x_2,)$$

取 $L(x_1, x_2, \lambda) = 0$ 即

$$\frac{\partial L}{\partial x_1} = 0 \qquad \frac{\partial L}{\partial x_2} = 0$$

$$\frac{\partial L}{\partial \lambda} = 0$$

联立求驻点 $\quad x_1^* = x_2^*, \quad \lambda^* =$

Lagrange乘子法

严格局部极小点

验证是否为极小点

对于任意非零向量 $y \in R^n$ 且 $y^T \nabla h_j(x^*) = 0 \quad j = (1, 2 \ldots l)$

$$y^T \nabla_x^2 L(x^*, \lambda^*) y > 0$$

$$\nabla h(x_1^*, x_2^*) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$H_x(x^k, \lambda^*) = \begin{bmatrix} \dfrac{\partial^2 L(x^k)}{\partial x_1^2} & \dfrac{\partial^2 L(x^k)}{\partial x_1 \partial x_2} \\ \dfrac{\partial^2 L(x^k)}{\partial x_2 \partial x_1} & \dfrac{\partial^2 L(x^k)}{\partial x_2^2} \end{bmatrix}$$

Lagrange乘子法

困难

Lagrange乘子法根据最优性条件将等式约束问题最后转化为$n+l$个未知数，$n+l$个未知方程所构成的方程组求解。

迭代法

$$S=(\nabla_x L)^2+(\nabla_\lambda L)^2$$
$$=\sum_{i=1}^{n}\left(\frac{\partial L}{\partial x_i}\right)^2+\sum_{j=1}^{l}(h_j(x))^2$$

Marquardt法

Lagrange乘子法

处理不等式约束的方法

$$g_i(x) \geq 0 \quad i = (1, 2 \ldots m) \Rightarrow g_i(x) - (x_{n+i})^2 = 0 \quad i = (1, 2 \ldots m)$$

example

$$\min \quad f(x) = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 6x_1$$

$$\text{s.t.} \quad g_1(x) = -3x_1 - 4x_2 + 6 \geq 0$$

$$g_2(x) = x_1 - 4x_2 + 2 \geq 0$$

有约束问题的最优解

约束问题最优解的充分必要条件 --- 基本概念

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & g_i(x) \ge 0 \quad i=(1,2\ \dots\ m) \\ & h_j(x) = 0 \quad j=(1,2\ \dots\ l) \end{aligned}$$

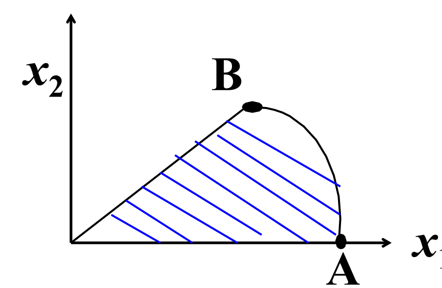
可行域 D:

$$D = \{\, x \mid g_i(x) \ge 0 \quad h_j(x) = 0 \quad i=(1,2\ \dots\ m) \quad j=(1,2\ \dots\ l) \,\}$$

$$\begin{aligned} \min \quad & f(x) \\ \text{s.t.} \quad & x \in D \end{aligned}$$

可行点

起作用约束: $I = \{\, i \mid g_i(x) = 0 \quad i=(1,2\ \dots\ m) \,\}$



任一等式约束都是关于任何一个可行点的起作用约束

约束问题最优解的充分必要条件 --- 基本概念

$$\min \quad f(x)$$
$$\text{s.t.} \quad g_i(x) \ge 0 \quad i = (1, 2 \ldots m)$$
$$h_j(x) = 0 \quad j = (1, 2 \ldots l)$$

可行域 D:

$$D = \{ x \mid g_i(x) \ge 0 \quad h_j(x) = 0 \quad i = (1, 2 \ldots m) \quad j = (1, 2 \ldots l) \}$$

$$\min \quad f(x)$$
$$\text{s.t.} \quad x \in D$$

可行点

可行方向：

设 $x^k \in D$，对于非零向量 $P^k \in R^n$ 存在 $\delta > 0$，

对于所有的 $\lambda \in [0, \delta]$ 必有 $(x^k + \lambda P^k) \in D$

则称 P^k 为区域D中从点 x^k 出发的一个可行方向

约束问题最优解的一阶必要条件 -- Kuhn-Tucker 条件

极值点 ----- 必须 $\boxed{\text{满足Kuhn-Tucker 条件}}$ 不一定是极值点

必要条件

- x^* 是一局部最优解
- $f(x)$，$g_i(x)$，$h_j(x)$ 在 x^* 点可微
- 对于所有的 $i \in \mathrm{I}$ 的 $\nabla g_i(x^*)$

  与 $\nabla h_1(x^*)$，…，$\nabla h_l(x^*)$ 线性无关

$m+l$个向量之间

约束问题最优解的一阶必要条件 -- Kuhn-Tucker 条件

那么，存在不全为零的一组数μ_1， μ_2， … μ_m， λ_1， λ_2，… λ_l，使得：

$$\nabla f(x^*) - \sum_{i=1}^{m} \mu_i \nabla g_i(x^*) - \sum_{j=1}^{l} \lambda_j \nabla h_j(x^*) = 0$$

$$\mu_i g_i(x^*) = 0 \quad (i = 1, 2, \square\ m)$$

$$\mu_i \geq 0 \quad (i = 1, 2, \square\ m)$$

Kuhn-Tucker 点不一定是最优点，然而，如果是最优点，则必须满足Kuhn-Tucker 条件

$$\min \quad f(x) = x_1$$

$$\text{s.t.} \quad g_1(x) = 16 - (x_1-4)^2 - x_2^2 \geq 0$$

$$h_1(x) = (x_1-3)^2 + (x_2-2)^2 - 13 = 0$$

$f(x)$ 在 A点（0，0）和 B点（6.4，3.2）是局部极值点

Kuhn-Tucker 条件是否成立？

a. 目标函数和约束函数的梯度表达式以及在A和B处的值？

b. 起作用约束梯度是否线性独立？

c. 不全为零的一组数μ_1， λ_1？目标函数的梯度可以表达为起作用约束的线性组合

惩罚函数法

基本策略

利用问题的目标函数和约束函数构造新的目标函数

——罚函数(penalty function)

约束最优化问题 → 无约束最优化问题

和 Lagrange乘子法 思路相同

惩罚函数法 –外惩罚函数法

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

可行域 D:

$$
\begin{cases}
\min \quad f(\mathrm{x}) \\
\text{s.t.} \quad g_i(\mathrm{x}) \geq 0, i=1,2,\mathrm{L}\ \ ,m \\
\qquad\quad h_j(\mathrm{x})=0, j=1,2,\mathrm{L}\ \ ,l
\end{cases}
\Rightarrow
F(\mathbf{x})=\mathring{F}(f(\mathbf{x}),g_i(\mathbf{x}),h_j(\mathbf{x}))
$$

具有“惩罚 性质”的辅助函数$F(\mathrm{x})$

要求$x \in D$, 当且仅当 $F(x)=f(x)$, 而 $\mathrm{x} \notin D$

时， $F(x)>f(x)$, 并且$F(x)-f(x)$随着x到D

的距离的增大而增大。

惩罚函数法 -外惩罚函数法

对于等式约束问题

$$\begin{cases} \min & f(\mathbf{x}) \\ \text{s.t.} & h_j(\mathbf{x}) = 0,\ j = 1, 2, \cdots, l \end{cases}$$

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

$$\min_{\mathbf{x} \in \mathrm{R}^n} F_1(\mathbf{x}, \sigma)$$

最优解必使所有 $h_j(\mathbf{x})(j = 1, 2, \cdots, l)$都接近0。否则 $F_1(\mathbf{x}, \sigma)$罚函数的第二项是很大的正数，与最优解取到极小值矛盾。

惩罚函数法 -外惩罚函数法

对于不等式约束问题

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

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

$$\min_{\mathbf{x} \in R^n} F_2(\mathbf{x}, \sigma)$$

最优解必使所有 $g_i(\mathbf{x})(i = 1, 2, \cdots, m)$ 都接近0或小于0。

否则，罚函数的第二项是很大的正数，与最优解取到极小值矛盾

惩罚函数法 -外惩罚函数法

对于一般约束问题

$$\begin{cases} \min & f(\mathbf{x}) \\ \text{s.t.} & h_j(\mathbf{x}) = 0, j = 1, 2, \cdots, l \\ & g_i(\mathbf{x}) \ge 0, i = 1, 2, \cdots, m \end{cases} \Rightarrow \begin{cases} F(\mathbf{x}, \mu) = f(\mathbf{x}) + \mu P(\mathbf{x}) \\ \min F(\mathbf{x}, \mu) \end{cases}$$

惩罚 因子μ很大的正数

$$P(\mathbf{x}) = \sum_{i=1}^{m} [g_i(\mathbf{x})]^2 \phi[g_i(\mathbf{x})] + \sum_{j=1}^{l} [h_j(\mathbf{x})]^2$$

$$\phi[g_i(x)] = \begin{cases} 0 & g_i(x) \ge 0 \\ 1 & g_i(x) < 0 \end{cases}$$

惩罚函数法 -外惩罚函数法

$$F(\mathbf{x},\mu)=f(\mathbf{x})+\mu P(\mathbf{x})$$

$$P(\mathbf{x})=\sum_{i=1}^{m}[g_i(\mathbf{x})]^2\,\phi[g_i(\mathbf{x})]$$

$$\phi[g_i(x)]=\begin{cases}0 & g_i(x)\ge 0\\ 1 & g_i(x)<0\end{cases}$$

例子

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

① $F(\mathbf{x},\mu)=?$

② 在***x-F***图上，令***μ*=1/4, 1/2, 1,** **画出** $F(\mathbf{x},\mu)$

③ $dF/dx=?$

④ $\mu\to\infty$ 时，***x****可得

惩罚函数法

-外惩罚函数法

$$F(\mathbf{x},\mu)=f(\mathbf{x})+\mu P(\mathbf{x})$$

$$P(\mathbf{x})=\sum_{i=1}^{m}[g_i(\mathbf{x})]^2\,\phi[g_i(\mathbf{x})]$$

$$\phi[g_i(x)]=\begin{cases}0 & g_i(x)\ge 0\\ 1 & g_i(x)<0\end{cases}$$

例子

$$\begin{cases}\min & f(\mathbf{x})=\dfrac{1}{3}(x_1+1)^3+x_2\\ \text{s.t.} & x_1-1\ge 0\\ & x_2\ge 0\end{cases}$$

惩罚函数法 -内惩罚函数法

基本策略

在迭代中总是从可行点出发，并保持在可行域内部进行搜索。因此，这种方法适用于只有不等式约束的最优化问题

惩罚函数法 -内惩罚函数法

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

$$G(\mathbf{x}, r) = f(\mathbf{x}) + rB(\mathbf{x})$$

$$\min G(\mathbf{x}, r)$$

对于不等式约束问题，其可行域 D 的内部 $\operatorname{int} D = \{\mathbf{x} \in \mathrm{R}^n \mid g_i(\mathbf{x}) > 0, i = 1, 2, \cdots, m\}$ **。**

为了保持迭代点始终含于 int D，r是很小的正数，B(x)是 int D 上的非负实值连续函数，当点 X 趋向可行域 D 的边界时，$B(\mathbf{x}) \to \infty$ **。**

显然，G(x,r) 罚函数的作用对企图脱离可行域的点给予惩罚，相当于在可行域的边界设置了障碍，不让迭代点穿越到可行域之外，因此也称为障碍函数（barrier function）。

惩罚函数法 -内惩罚函数法

常用的形式

$$G(\mathbf{x},r)=f(\mathbf{x})-r\sum_{i=1}^{m}\ln g_i(\mathbf{x})$$

$$G(\mathbf{x},r)=f(\mathbf{x})+r\sum_{i=1}^{m}\frac{1}{(g_i(\mathbf{x}))^2} \qquad r \longrightarrow 0$$

$$G(\mathbf{x},r)=f(\mathbf{x})+r\sum_{i=1}^{m}\frac{1}{g_i(\mathbf{x})}$$

$$\begin{cases}\min & G(\mathbf{x},r)=f(\mathbf{x})+rB(\mathbf{x})\\ \text{s.t.} & \mathbf{x}\in \operatorname{int} D\end{cases} \longrightarrow \begin{cases}\min & f(\mathbf{x})\\ \text{s.t.} & g_i(\mathbf{x})\geq 0, i=1,2,\cdots,m\end{cases}$$

惩罚函数法 –内惩罚函数法

例子

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

$$G(\mathbf{x}, r) = f(\mathbf{x}) + rB(\mathbf{x})$$

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

① $G(\mathbf{x}, r) = ?$

② $dG / dx = ?$

③ $r \to 0$ 时，x^*可得

可行方向法

- 可行方向法的基本思想是从可行点出发，沿可行下降方向进行搜索，求出使目标函数值下降的新的可行点。

- 算法包括选择搜索方向和确定搜索步长两个主要方面.
搜索方向的选择方式不同就形成不同的可行方向法。

- 在方向搜索过程中，通常需要利用一系列线性规划的方法进行求解。

可行方向法

- **Zoutendijk可行方向法**
- **梯度投影法(Gradient Projection Method)**
- **既约梯度法 (Reduced Gradient Method)**

可行方向法

1. Zoutendijk可行方向法是Zoutendijk于1960年提出的.

2. Zoutendijk可行方向法中选择搜索方向：

   起作用约束构造可行方向

3. Zoutendijk可行方向法可以求解线性约束优化问题和非线性约束优化问题.

可行方向法

Zoutendijk可行方向法

线性约束情形

例子

$$
\begin{cases}
\min & f(\mathbf{x}) = x_1^2 + 4x_2^2 \\
\text{s.t.} & x_1 + x_2 \geq 1 \\
 & 15x_1 + 10x_2 \geq 12 \\
 & x_1 \geq 0 \quad x_2 \geq 0
\end{cases}
$$

可行方向法

$$\begin{cases} \min & f(\mathbf{x}) = x_1^2 + 4x_2^2 \\ \text{s.t.} & x_1 + x_2 \geq 1 \\ & 15x_1 + 10x_2 \geq 12 \\ & x_1 \geq 0 \quad x_2 \geq 0 \end{cases}$$

(1) 利用起作用约束构造可行下降方向

$$\nabla f(\mathbf{x}) = (2x_1, 8x_2)$$

约束方程系数矩阵

$$\mathbf{A} = \begin{bmatrix} 1 & 1 \\ 15 & 10 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\mathbf{x}^0 = (0, 2)$$

对$\mathbf{x}^0$而言：起作用约束 $\quad \mathbf{A}_1^0 = (1, 0) \qquad \mathbf{b}_1^0 = 0$

不起作用约束

$$\mathbf{A}_2^0 = \begin{bmatrix} 1 & 1 \\ 15 & 10 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{b}_2^0 = \begin{bmatrix} 1 \\ 12 \\ 0 \end{bmatrix}$$

可行方向法

$$\begin{cases} \min & f(\mathbf{x}) = x_1^2 + 4x_2^2 \\ \text{s.t.} & x_1 + x_2 \ge 1 \\ & 15x_1 + 10x_2 \ge 12 \\ & x_1 \ge 0 \quad x_2 \ge 0 \end{cases}$$

(2) 确定可行下降方向　　$\mathbf{x}^0 = (0, 2)$

$$\nabla f(\mathbf{x}^0) = (0, 16)$$

下降方向$\mathbf{P}^0 = (p_1, p_2)^T$的线性规划

$$\mathbf{A}_1^0 = (1, 0)$$

$$\begin{cases} \min & \nabla f(\mathrm{x}^0)(\mathbf{P}^0) = 16p_2 \\ \text{s.t.} & \mathrm{A}_1^0 \mathbf{P}^0 = p_1 \ge 0 \\ & -1 \le p_1 \le 1 \\ & -1 \le p_2 \le 1 \end{cases}$$

最优解$\mathbf{P}^0 = (0, -1)^T$

$\nabla f(\mathbf{x}^0)\mathbf{P}^0 = -16 \ne 0$　不满足收敛精度，继续迭代

$$|\nabla f(\mathbf{x}^0)\mathbf{P}^0| > \varepsilon$$

可行方向法

$$\mathbf{A}_2^0=\begin{bmatrix}1&1\\15&10\\0&1\end{bmatrix}\quad \mathbf{b}_2^0=\begin{bmatrix}1\\12\\0\end{bmatrix}$$

$$\mathbf{x}^0=(0,2)\qquad \mathbf{P}^0=(0,-1)^T$$

$$\begin{cases}\min & f(\mathbf{x})=x_1^2+4x_2^2\\ \text{s.t.} & x_1+x_2\ge 1\\ & 15x_1+10x_2\ge 12\\ & x_1\ge 0\quad x_2\ge 0\end{cases}$$

(3) 计算步长

一维搜索
$$\begin{cases}\min & f(\lambda_0)=f(\mathbf{x}^0+\lambda_0\mathbf{P}^0)=4(2-\lambda_0)^2\\ \text{s.t.} & 0\le\lambda_0\le \boxed{?}\end{cases}$$

$$\boxed{?}=\lambda_0^U=\min_i\left(-\frac{\mu_i^0}{v_i^0}\right)\quad v_i^0<0$$

$$\boldsymbol{\mu}^0=\mathbf{A}_2^0\mathbf{x}^0-\mathbf{b}_2^0=(1,8,2)$$

$$\mathbf{v}^0=\mathbf{A}_2^0\mathbf{P}^0=(-1,-10,-1)$$

$$\boxed{?}=\lambda_0^U=4/5$$

最优解 $\lambda_0=4/5$ $\quad \mathbf{x}^1=\mathbf{x}^0+\lambda_0\mathbf{P}^0=(0,6/5)$

可行方向法

(4) 继续迭代 $\mathbf{x}^1 = (0, 6/5)$

对$\mathbf{x}^1$而言：起作用约束 $\mathbf{A}_1^1 = \begin{pmatrix} 15 & 10 \\ 1 & 0 \end{pmatrix}$ $\mathbf{b}_1^1 = \begin{pmatrix} 12 \\ 0 \end{pmatrix}$

不起作用约束 $\mathbf{A}_2^1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ $\mathbf{b}_2^1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

可行方向法

- **Zoutendijk可行方向法的改进(线性约束)**

  - 考虑起作用约束和不起作用约束在确定搜索方向中都起作用.

- 全作用约束方向法是Topkis和Veinott (1967) 提出 (非线性不等式约束)

可行方向法

◆ 非线性不等式：TV方法

$$\begin{cases} \min & f(\mathbf{x}) = (x_1 - 2)^2 + (x_2 - 0.5)^2 \\ \text{s.t.} & 1.5 - x_1^2 - x_2 \geq 0 \\ & -x_1 + 2x_2 \geq 0 \\ & x_1 \geq 0 \end{cases}$$

$$\varepsilon = 0.001$$

$$\begin{cases} \min & f(\mathbf{x}) = (x_1-2)^2 + (x_2-0.5)^2 \\ \text{s.t.} & 1.5 - x_1^2 - x_2 \geq 0 \\ & -x_1 + 2x_2 \geq 0 \\ & x_1 \geq 0 \end{cases}$$

可行方向法

(1)　目标函数和约束函数的梯度表达式

$$\nabla g_1(\mathbf{x}) = [-2x_1, -1]^{\mathrm{T}}$$

$$\nabla f(\mathbf{x}) = [2(x_1-2), 2(x_2-0.5)]^{\mathrm{T}} \qquad \nabla g_2(\mathbf{x}) = [-1, 2]^{\mathrm{T}}$$

$$\nabla g_3(\mathbf{x}) = [1, 0]^{\mathrm{T}}$$

$\mathbf{x}^0 = (0,1)$ **对$\mathbf{x}^0$而言：**

$$g_1(\mathbf{x}^0) = 0.5, \quad g_2(\mathbf{x}^0) = 2, \quad g_3(\mathbf{x}^0) = 0$$

$$\nabla f(\mathbf{x}^0) = [-4, 1]^{\mathrm{T}} \quad \nabla g_1(\mathbf{x}^0) = [0, -1]^{\mathrm{T}}$$

$$\nabla g_2(\mathbf{x}^0) = [-1, 2]^{\mathrm{T}} \quad \nabla g_3(\mathbf{x}^0) = [1, 0]^{\mathrm{T}}$$

$$\begin{cases} \min & f(\mathbf{x})=(x_1-2)^2+(x_2-0.5)^2 \\ \text{s.t.} & 1.5-x_1^2-x_2\geq 0 \\ & -x_1+2x_2\geq 0 \\ & x_1\geq 0 \end{cases}$$

可行方向法

(2) 确定可行下降方向 $\mathbf{x}^0=(0,1)$

$\nabla f(\mathbf{x}^0)=[-4,1]^{\mathrm{T}}$ $\nabla g_1(\mathbf{x}^0)=[0,-1]^{\mathrm{T}}$

$\nabla g_2(\mathbf{x}^0)=[-1,2]^{\mathrm{T}}$ $\nabla g_3(\mathbf{x}^0)=[1,0]^{\mathrm{T}}$

$g_1(\mathbf{x}^0)=0.5,\quad g_2(\mathbf{x}^0)=2,\quad g_3(\mathbf{x}^0)=0$

下降方向$\mathbf{P}^0=(p_1, p_2)^{\mathrm{T}}$的线性规划

$$\begin{cases} \min & f(\mathbf{P}^0,y^0)=y_0 \\ \textbf{s.t.} & \nabla f(\mathbf{x}^0)^{\mathrm{T}}\mathbf{P}^0-y_0\leq 0 \\ & -\nabla g_i(\mathbf{x}^0)^{\mathrm{T}}\mathbf{P}^0-y_0\leq g_i(\mathbf{x}^0) \\ & -1\leq p_1\leq 1 \\ & -1\leq p_2\leq 1 \end{cases}$$

最优解$\mathbf{P}^0$＝(0.75, -0.25)$^{\mathrm{T}}$

$y^0=-0.75\rightarrow\left|y^0\right|>\varepsilon$ 不满足收敛精度，继续迭代

可行方向法

$$\begin{cases} \min & f(\mathbf{x})=(x_1-2)^2+(x_2-0.5)^2 \\ \text{s.t.} & 1.5-x_1^2-x_2\ge 0 \\ & -x_1+2x_2\ge 0 \\ & x_1\ge 0 \end{cases}$$

(3) 计算步长

$$\begin{cases} \min & f(\lambda_0)=f(\mathbf{x}^0+\lambda_0\mathbf{P}^0)=(0.75\lambda_0-2)^2+(0.5-0.25\lambda_0)^2 \\ \text{s.t.} & 0\le\lambda_0\le \boxed{?} \end{cases}$$

一维搜索

$$\boxed{?}=\lambda_0^U=\max_i\{\lambda_k \mid g_i(\mathrm{x}^{(k)}+\lambda_k \mathrm{P}^{(k)})\ge 0\}$$

$$\boxed{?}=\lambda_0^U=4+4\sqrt{19}/18=1.19$$

最优解 $\lambda_0=1.19$ $\mathbf{x}^1=\mathbf{x}^0+\lambda_0\mathbf{P}^0=(0.89,0.7)$

可行方向法

(4) 继续迭代　　$\mathbf{x}^1 = (0.89, 0.7)$

$\nabla f(\mathbf{x}^1) = [-2.2, 0.4]^{\mathrm{T}}$　　$\nabla g_1(\mathbf{x}^1) = [-1.8, -1]^{\mathrm{T}}$

$\nabla g_2(\mathbf{x}^1) = [-1, 2]^{\mathrm{T}}$　　$\nabla g_3(\mathbf{x}^1) = [1, 0]^{\mathrm{T}}$

$g_1(\mathbf{x}^1) = 0,\quad g_2(\mathbf{x}^1) = 0.51,\quad g_3(\mathbf{x}^1) = 0.89$

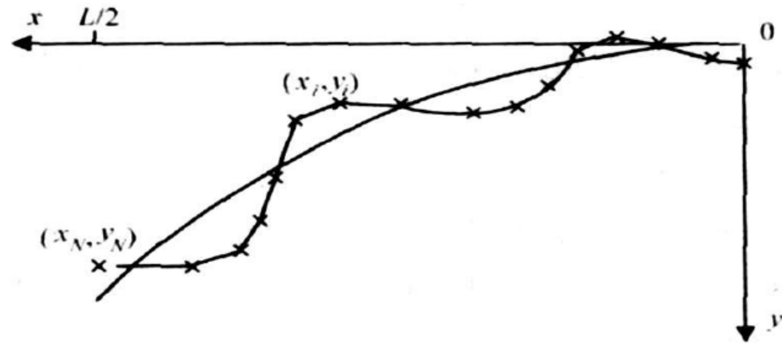
下降方向$\mathbf{P}^0 = (p_1, p_2)^{\mathrm{T}}$的线性规划

$$\begin{cases} \min & f(\mathbf{P}^1, y^1) = y_1 \\ \text{s.t.} & \nabla f(\mathbf{x}^1)^{\mathrm{T}} \mathbf{P}^1 - y_1 \le 0 \\ & -\nabla g_i(\mathbf{x}^1)^{\mathrm{T}} \mathbf{P}^1 - y_1 \le g_i(\mathbf{x}^1) \\ & -1 \le p_1 \le 1 \\ & -1 \le p_2 \le 1 \end{cases}$$

最优解$\mathbf{P}^0 = (0.026, -0.177)^{\mathrm{T}}$

可行方向法

应用实例



如图 1 所示坐标系，其中虚线为拱轴压力线，压力线上的 $N+1$ 个点为：(x_0, y_0)，(x_1, y_1)，…，(x_N, y_N)，实线为相应的拟合拱轴线，拱顶在坐标系原点。用 4 次样条函数拟合的拱轴线可表示为

$$F(x) = a_1 x^2 + a_2 x^3 + \sum_{j=0}^{N-1} a_{j+3}(x - x_j)_+^4,$$

$$x \in [0, x_N], \qquad (1)$$

式中，$(x - x_j)_+ = \begin{cases} x - x_j & x \geq x_j \\ 0 & x < x_j \end{cases}$；$a_1$，$a_2$，…，$a_{N+2}$ 为

待定参数

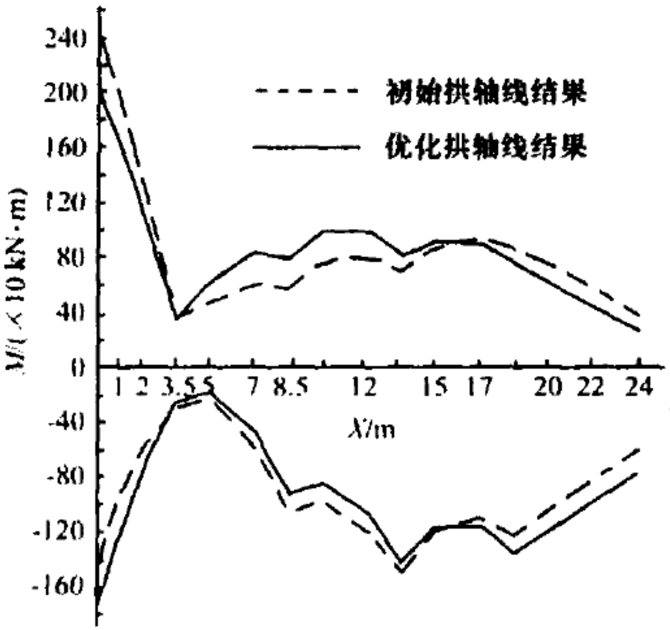
可行方向法

应用实例

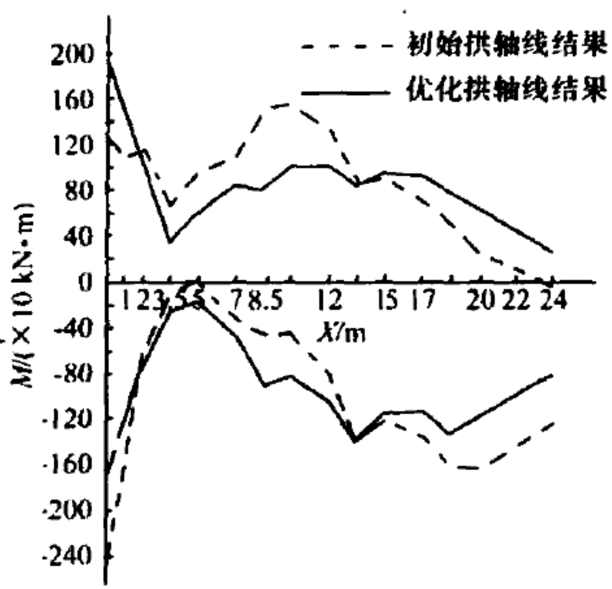
$$
\begin{cases}
\min \quad \lambda, \\
\text{s. t.} \quad a_1 + 3x_i + 6\sum_{j=0}^{i-1}(x_i - x_j)^2 a_{j+3} \geqslant 0, \\
\quad 8\left(a_1 + 6\sum_{j=0}^{i-1} x_j a_{j+3}\right)\sum_{j=0}^{i-1} a_{j+3} - 3\left(a_2 - \right. \\
\quad \left. 4\sum_{j=0}^{i-1} x_j a_{j+3}\right)^2 \geqslant 0, \\
\quad x_i^2 a_1 + x_i^3 a_2 + \sum_{j=0}^{i-1}(x_i - x_j)^4 a_{j+3} + \lambda \geqslant y_i, \\
\quad - x_i^2 a_1 - x_i^3 a_2 - \sum_{j=0}^{i-1}(x_i - x_j)^4 a_{j+3} + \\
\quad \lambda \geqslant - y_i, \\
\quad 1 \leqslant i \leqslant N,
\end{cases}
$$

可行方向法

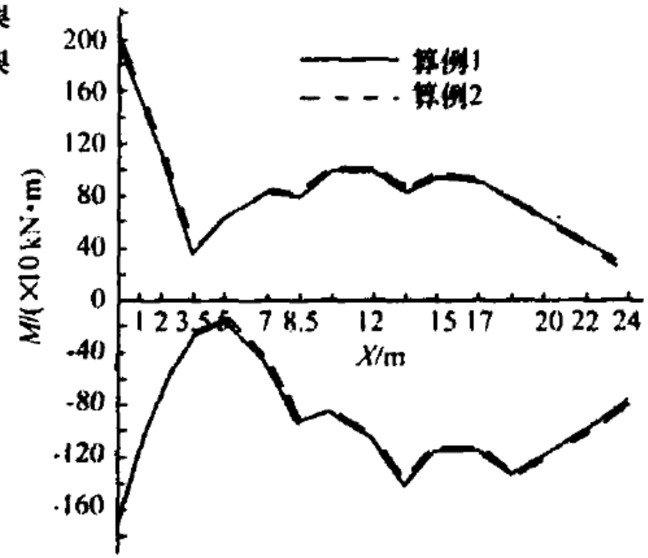
应用实例



初始拱轴线结果
优化拱轴线结果
M/(×10 kN·m)
X/m
240
200
160
120
80
40
0
-40
-80
-120
-160
1 2 3.5 5 7 8.5 12 15 17 20 22 24



初始拱轴线结果
优化拱轴线结果
M/(×10 kN·m)
X/m
200
160
120
80
40
0
-40
-80
-120
-160
-200
-240
1 2 3.5 5 7 8.5 12 15 17 20 22 24



算例1
算例2
M/(×10 kN·m)
X/m
200
160
120
80
40
0
-40
-80
-120
-160
1 2 3.5 5 7 8.5 12 15 17 20 22 24