组合优化与凸优化作业 1

胡冠宇

1. a) 不是凸集, 理由如下:

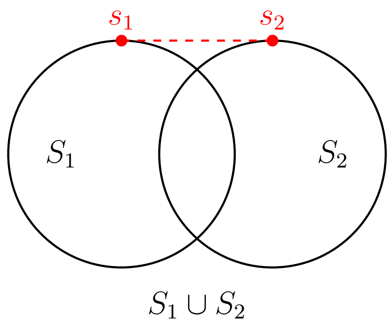


图 1

如图 1 所示, 取图中 $s_1 \in S_1$, $s_2 \in S_2$ 两点, $\lambda s_1 + (1-\lambda)s_2$ 为图中红色虚线段. 由于该线段不在 $S_1 \cup S_2$ 内, 因此 $S_1 \cup S_2$ 不是凸集.

b) 是凸集, 理由如下:

$\forall s_1, s_2 \in S_1 + S_2$, s.t.

$$s_1 = x_1 + y_1, \quad s_2 = x_2 + y_2, \quad x_1, x_2 \in S_1, \quad y_1, y_2 \in S_2.$$

$\forall \lambda \in [0,1]$, s.t.

$$\begin{aligned}\lambda s_1 + (1-\lambda)s_2 &= \lambda(x_1 + y_1) + (1-\lambda)(x_2 + y_2) \\ &= [\lambda x_1 + (1-\lambda)x_2] + [\lambda y_1 + (1-\lambda)y_2],\end{aligned}$$

由于 S_1, S_2 均为凸集, 因此 $\lambda x_1 + (1-\lambda)x_2 \in S_1$, $\lambda y_1 + (1-\lambda)y_2 \in S_2$, 则

$$\lambda s_1 + (1-\lambda)s_2 \in S_1 + S_2,$$

因此 $S_1 + S_2$ 是凸集.

c) 是凸集, 理由如下:

$\forall s_1, s_2 \in S_1 - S_2$, s.t.

$$s_1 = x_1 - y_1, \quad s_2 = x_2 - y_2, \quad x_1, x_2 \in S_1, \quad y_1, y_2 \in S_2.$$

$\forall \lambda \in [0,1]$, s.t.

$$\begin{aligned}\lambda s_1 + (1-\lambda)s_2 &= \lambda(x_1 - y_1) + (1-\lambda)(x_2 - y_2)\\ &= [\lambda x_1 + (1-\lambda)x_2] - [\lambda y_1 + (1-\lambda)y_2],\end{aligned}$$

由于 S_1, S_2 均为凸集, 因此 $\lambda x_1 + (1-\lambda)x_2 \in S_1$, $\lambda y_1 + (1-\lambda)y_2 \in S_2$, 则

$$\lambda s_1 + (1-\lambda)s_2 \in S_1 - S_2,$$

因此 $S_1 - S_2$ 是凸集.

2. a) 不是凸集, 理由如下:

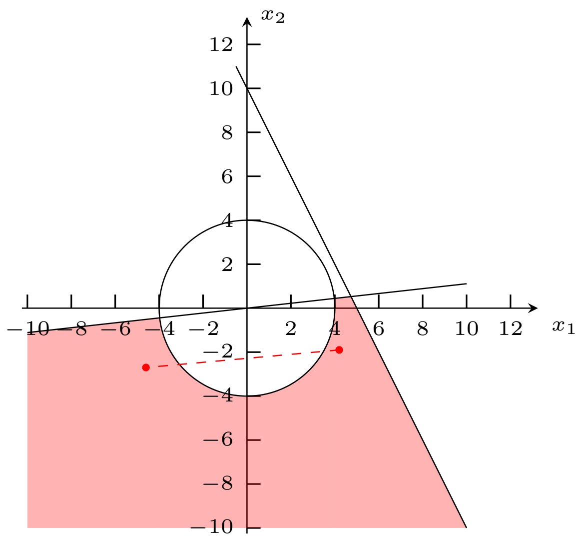


图 2

如图 2 , 取图中两点连成线段, 该线段不在 S 内, 因此 S 不是凸集.

b) 是凸集, 理由如下:

$\forall s_1 = (x_1, y_1), s_2 = (x_2, y_2) \in S$, 要证 S 是凸集, 即证 $\forall \lambda \in [0,1]$, s.t.

$$\lambda s_1 + (1-\lambda)s_2 = (x,y) = (\lambda x_1 + (1-\lambda)y_1, \lambda x_2 + (1-\lambda)y_2) \in S.$$

因为

$$x + y = \lambda(x_1 + y_1) + (1-\lambda)(x_2 + y_2) \leq 6,$$

$$-2x + 3y = \lambda(-2x_1 + 3y_1) + (1-\lambda)(-2x_2 + 3y_2) \geq 2,$$

$$4x - y = \lambda(4x_1 - y_1) + (1-\lambda)(4x_2 - y_2) \leq 12,$$

则 $(x,y) \in S$, 因此 S 是凸集.

c) 是凸集, 理由如下:

$\forall s_1 = (x_1, y_1), s_2 = (x_2, y_2) \in S$, 要证 S 是凸集, 即证 $\forall \lambda \in [0, 1]$, s.t.

$$\lambda s_1 + (1-\lambda)s_2 = (x, y) = (\lambda x_1 + (1-\lambda)y_1, \lambda x_2 + (1-\lambda)y_2) \in S.$$

易证 $x + y \geq 3, x \geq 1$, 因此只需证明 $-(x-1)^2 + y \geq 1$ 即可.
令 $f(x) = -(x-1)^2$, 则 $f''(x) = -2 \leq 0$, 因此函数 $f(x)$ 是凹函数, 即

$$-(x-1)^2 \geq -\lambda(x_1 - 1)^2 - (1-\lambda)(x_2 - 1)^2.$$

又因为 $y = \lambda x_2 + (1-\lambda)y_2$, 所以

$$-(x-1)^2 + y \geq \lambda[-(x_1-1)^2 + y_1] + (1-\lambda)[-(x_2-1)^2 + y_2] \geq 1.$$

因此 S 是凸集.

d) 不是凸集, 理由如下:

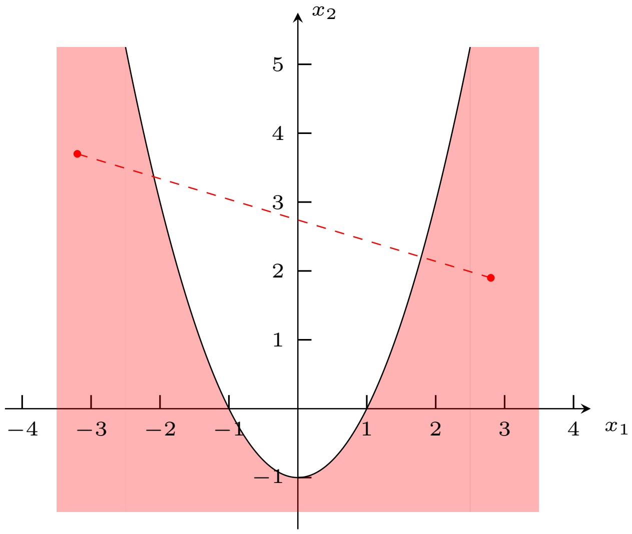


图 3

如图 3 , 取图中两点连成线段, 该线段不在 S 内, 因此 S 不是凸集.

3. a) 是凸函数, 理由如下:

$\forall x_1, x_2 \in \mathbb{R}, \lambda \in [0,1]$, 则

$$\begin{aligned} g(\lambda x_1 + (1-\lambda)x_2) &= \max_{1\le i\le n, i\in\mathbb{Z}} \{f_i(\lambda x_1 + (1-\lambda)x_2)\} \\ &\le \max_{1\le i\le n, i\in\mathbb{Z}} \{\lambda f_i(x_1) + (1-\lambda) f_i(x_2)\} \\ &\le \lambda \max_{1\le i\le n, i\in\mathbb{Z}} f_i(x_1) + (1-\lambda) \max_{1\le i\le n, i\in\mathbb{Z}} f_i(x_2) \\ &= \lambda g(x_1) + (1-\lambda) g(x_2), \end{aligned}$$

因此 $g(x)$ 是凸函数.

b) 不是凸函数, 理由如下:

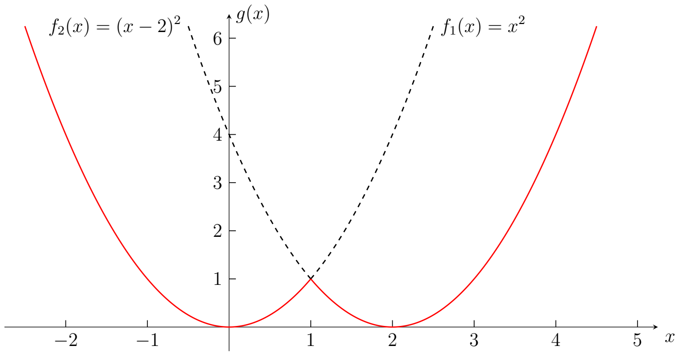


图 4

如图 4 所示, 取 $g(x) = \min\{x^2, (x-2)^2\}$, 即图中红色部分. $\forall x_1, x_2 \in [0,2]$ 时,

$$g(\lambda x_1 + (1-\lambda)x_2) \ge \lambda g(x_1) + (1-\lambda) g(x_2), \quad \lambda \in [0,1],$$

因此 $g(x)$ 不是凸函数.

c) 是凸函数, 理由如下:

$\forall \boldsymbol{x}^{(1)}, \boldsymbol{x}^{(2)} \in \mathbb{S}$, $\mathbb{S} = \left\{ [x_1, x_2, \cdots, x_n]^{\mathrm{T}} \,\middle|\, x_i > 0, \sum_{i=1}^{n} x_i = 1 \right\}$, $\lambda \in [0,1]$, 有

$$\begin{aligned} \mathrm{LHS} &= g\left(\lambda \boldsymbol{x}^{(1)} + (1-\lambda)\boldsymbol{x}^{(2)}\right) \\ &= \sum_{i=1}^{n} \left(\lambda x_i^{(1)} + (1-\lambda) x_i^{(2)}\right) \log \left(\lambda x_i^{(1)} + (1-\lambda) x_i^{(2)}\right), \end{aligned}$$

$$\begin{aligned} \mathrm{RHS} &= \lambda g\left(\boldsymbol{x}^{(1)}\right) + (1-\lambda) g\left(\boldsymbol{x}^{(2)}\right) \\ &= \lambda \sum_{i=1}^{n} x_i^{(1)} \log x_i^{(1)} + (1-\lambda) \sum_{i=1}^{n} x_i^{(2)} \log x_i^{(2)}. \end{aligned}$$

令 $h(x) = x\log x$, 易知 $h''(x) = \dfrac{1}{x} > 0$, 则 $h(x)$ 为凸函数, 因此有

$$h(\lambda x_1 + (1-\lambda)x_2) \le \lambda h(x_1) + (1-\lambda)h(x_2),$$

则

$$\begin{aligned}\text{LHS} &\le \sum_{i=1}^{n} \lambda x_i^{(1)} \log x_i^{(1)} + (1-\lambda)x_i^{(2)} \log x_i^{(2)} \\ &= \text{RHS},\end{aligned}$$

因此 $g(x)$ 是凸函数.

d) 不是凸函数, 理由如下:

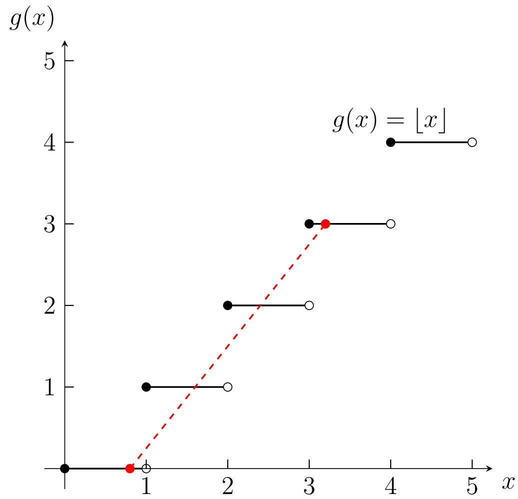


图 5

如图 5 所示, 取 $g(x)$ 上的点 $(0.8, 0)$ 和 $(3.2, 3)$ 并连成如图中所示的红色虚线段, 可观察到在 $[0.8, 3.2]$ 中有存在于线段上方的值, 因此 $g(x)$ 不是凸函数.

e) 是凸函数, 理由如下:

函数 $g(x)$ 可等价地表示为

$$g(x) = \max_{\substack{\mathcal{S}\subseteq \mathbb{Z}\cap[1,n] \\ |\mathcal{S}|=k}} \sum_{x\in\mathcal{S}} x_i.$$

又因为 $f_1(x) = \max x$ 与 $f_2(x) = x$ 均为凸函数, 则 $g(x)$ 也为凸函数.

f) 是凸函数, 理由如下:

$\forall x_1, x_2 \in \mathbb{R}, \lambda \in [0,1]$, 有

$$\begin{aligned}\text{LHS} &= g(\lambda x_1 + (1-\lambda)x_2)\\ &= (\lambda x_1 + (1-\lambda)x_2)^2\\ &= \lambda^2 x_1^2 + (1-\lambda)^2 x_2^2 + 2\lambda(1-\lambda)x_1x_2,\end{aligned}$$

$$\begin{aligned}\text{RHS} &= \lambda g(x_1) + (1-\lambda)g(x_2)\\ &= \lambda x_1^2 + (1-\lambda)x_2^2,\end{aligned}$$

$$\begin{aligned}\text{LHS} - \text{RHS} &= \left[\lambda^2 x_1^2 + (1-\lambda)^2 x_2^2 + 2\lambda(1-\lambda)x_1x_2\right] - \left[\lambda x_1^2 + (1-\lambda)x_2^2\right]\\ &= -\lambda(1-\lambda)(x_1+x_2)^2\\ &\leq 0,\end{aligned}$$

因此 $g(x)$ 是凸函数.

4. a) 将问题化为标准型:

$$\begin{cases}\max & 2x_1' + x_2 - 2x_3' + 2x_3''\\ \text{s.t.} & x_1' + x_2 + x_3' - x_3'' = 4,\\ & x_1' + x_2 - x_3' + x_3'' + x_4 = 6,\\ & x_1', x_2, x_3', x_3'', x_4 \geq 0,\end{cases}$$

则 $\boldsymbol{A} = (\boldsymbol{P}_1\ \ \boldsymbol{P}_2\ \ \boldsymbol{P}_3\ \ \boldsymbol{P}_4\ \ \boldsymbol{P}_5) = \begin{pmatrix}1 & 1 & 1 & -1 & 0\\ 1 & 1 & -1 & 1 & 1\end{pmatrix}$, $\boldsymbol{b} = (4\ \ 6)^{\mathrm{T}}$.

讨论过程如表 1 所示, 最优解为 $(5\ \ 0\ \ 0\ \ 1\ \ 0)^{\mathrm{T}}$, 最优值为 12.

b) 将问题化为标准型:

$$\begin{cases}\max & -2x_1 + x_2' - 3x_3 - x_4' + x_4''\\ \text{s.t.} & x_1 - x_2' + x_3 + x_4' - x_4'' + x_5 = 7,\\ & -2x_1 - 3x_2' - 5x_3 = 8,\\ & x_1 - 2x_3 + 2x_4' - 2x_4'' - x_6 = 1,\\ & x_1, x_2', x_3, x_4', x_4'', x_5, x_6 \geq 0,\end{cases}$$

表 1

$\boldsymbol{B}$	$\boldsymbol{B}^{-1}$	$\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}$	基本解	值
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2\end{pmatrix}, \begin{pmatrix}\boldsymbol{P}_3 & \boldsymbol{P}_4\end{pmatrix}$	-	-	-	-
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_3\end{pmatrix}, \begin{pmatrix}\boldsymbol{P}_2 & \boldsymbol{P}_3\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix}$	$\begin{pmatrix}5 & -1\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}5 & 0 & -1 & 0 & 0\end{pmatrix}^{\mathrm{T}}$ $\begin{pmatrix}0 & 5 & -1 & 0 & 0\end{pmatrix}^{\mathrm{T}}$	-
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_4\end{pmatrix}, \begin{pmatrix}\boldsymbol{P}_2 & \boldsymbol{P}_4\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}1 & 1\\ -1 & 1\end{pmatrix}$	$\begin{pmatrix}5 & 1\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}5 & 0 & 0 & 1 & 0\end{pmatrix}^{\mathrm{T}}$ $\begin{pmatrix}0 & 5 & 0 & 1 & 0\end{pmatrix}^{\mathrm{T}}$	12
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_5\end{pmatrix}, \begin{pmatrix}\boldsymbol{P}_2 & \boldsymbol{P}_5\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ -1 & 1\end{pmatrix}$	$\begin{pmatrix}4 & 2\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}4 & 0 & 0 & 0 & 2\end{pmatrix}^{\mathrm{T}}$ $\begin{pmatrix}0 & 4 & 0 & 0 & 2\end{pmatrix}^{\mathrm{T}}$	8
$\begin{pmatrix}\boldsymbol{P}_3 & \boldsymbol{P}_5\end{pmatrix}$	$\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$	$\begin{pmatrix}4 & 10\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}0 & 0 & 3 & 0 & 10\end{pmatrix}^{\mathrm{T}}$	-8
$\begin{pmatrix}\boldsymbol{P}_4 & \boldsymbol{P}_5\end{pmatrix}$	$-\begin{pmatrix}1 & 0\\ 1 & 1\end{pmatrix}$	$\begin{pmatrix}-4 & -10\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}0 & 0 & 0 & -4 & -10\end{pmatrix}^{\mathrm{T}}$	-

则 $\boldsymbol{A}=(\boldsymbol{P}_1\ \ \boldsymbol{P}_2\ \ \boldsymbol{P}_3\ \ \boldsymbol{P}_4\ \ \boldsymbol{P}_5\ \ \boldsymbol{P}_6\ \ \boldsymbol{P}_7)=\begin{pmatrix}1 & -1 & 1 & 1 & -1 & 1 & 0\\ -2 & -3 & -5 & 0 & 0 & 0 & 0\\ 1 & 0 & -2 & 2 & -2 & 0 & -1\end{pmatrix}$, $\boldsymbol{b}=(7\ \ 8\ \ 1)^{\mathrm{T}}$. 由约束可知, 当 $x_1, x_2', x_3 \geq 0$ 时, $-2x_1-3x_2'-5x_3=8$ 不成立, 因此该问题无解.

5. a) 如图 6 所示, 可行域为空, 因此无解.

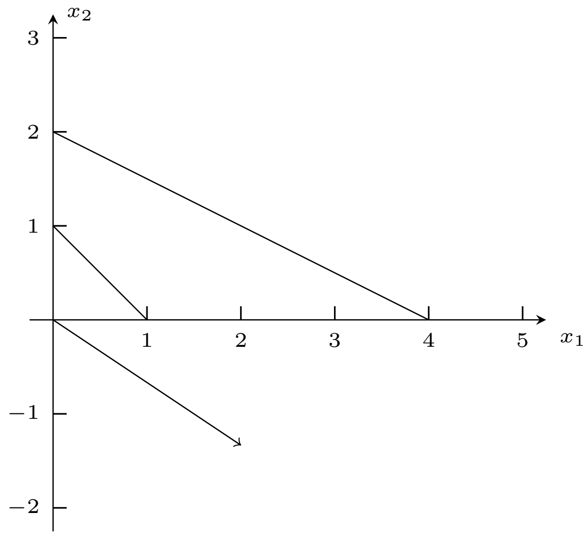


图 6

b) 如图 7 所示, 可行域为红色区域. $f=-x_1+3x_2$ 在可行域内可以无限小, 因此无解.

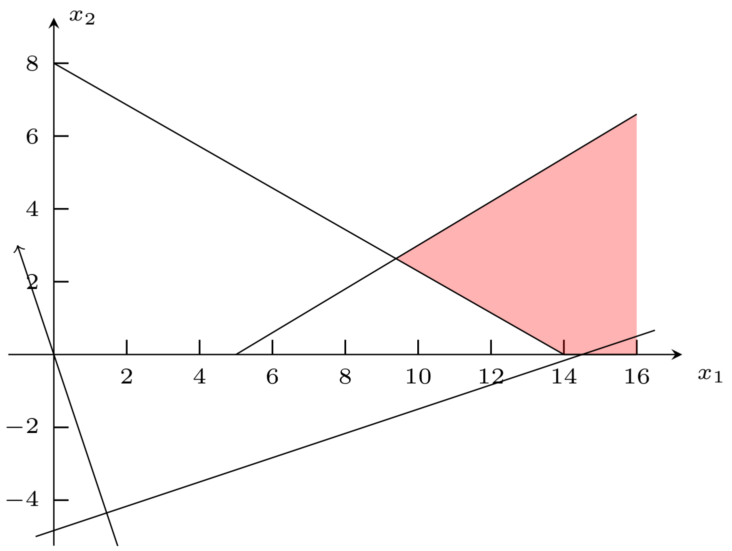


图 7

6. 问题可写成

$$\begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ -1 & 1 & -3 & 2 & -1 \end{pmatrix} \boldsymbol{x} = \begin{pmatrix} 4 \\ -5 \end{pmatrix},$$

基本解见表 2 .

表 2

$\boldsymbol{B}$	$\boldsymbol{B}^{-1}$	$\boldsymbol{x}_B$	基本解
$\begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 \end{pmatrix}$	$\frac{1}{3}\begin{pmatrix} 1 & -2 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} \frac{14}{3} & -\frac{1}{3} \end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix} \frac{14}{3} & -\frac{1}{3} & 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}$
$\begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_4 \end{pmatrix}$	$\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}$	$\begin{pmatrix} 3 & -1 \end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix} 3 & 0 & 0 & -1 & 0 \end{pmatrix}^{\mathrm{T}}$
$\begin{pmatrix} \boldsymbol{P}_2 & \boldsymbol{P}_3 \end{pmatrix}$	$\frac{1}{9}\begin{pmatrix} 3 & 3 \\ 1 & -2 \end{pmatrix}$	$\begin{pmatrix} -\frac{1}{3} & -\frac{14}{9} \end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix} 0 & -\frac{1}{3} & -\frac{14}{9} & 0 & 0 \end{pmatrix}^{\mathrm{T}}$

7. 由题可知, $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 & \boldsymbol{P}_3 & \boldsymbol{P}_4 \end{pmatrix} = \begin{pmatrix} 2 & 4 & 2 & 2 \\ 4 & 10 & 6 & 8 \end{pmatrix}$, $\boldsymbol{b} = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$. 基本解及 $\boldsymbol{B}$, $\boldsymbol{B}^{-1}$, $\boldsymbol{N}$, $\boldsymbol{x}_B$, $\boldsymbol{x}_N$ 见表 3 .

8. **必要性**: 若 $\boldsymbol{x} = \begin{pmatrix} x_1 & \cdots & x_k & 0 & \cdots 0 \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^n$ 为基本可行解, 且 $x_1, \cdots, x_k > 0$, 对应的 $\boldsymbol{B} = \begin{pmatrix} \boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_k \end{pmatrix}$. 假设 $\boldsymbol{P}_1, \boldsymbol{P}_2, \cdots, \boldsymbol{P}_k$ 线性相关, 则 $\exists \boldsymbol{d} \neq \boldsymbol{0}$, s.t. $\boldsymbol{A}\boldsymbol{d} = \boldsymbol{0}$. 将 $\boldsymbol{x}$ 沿 $\boldsymbol{d}$ 的方向移动 λ, 得到 $\boldsymbol{x} + \lambda \boldsymbol{d}$, 则

$$\boldsymbol{A}(\boldsymbol{x} + \lambda \boldsymbol{d}) = \boldsymbol{A}\boldsymbol{x} + \lambda \boldsymbol{A}\boldsymbol{d} = \boldsymbol{b},$$

表 3

$\boldsymbol{B}$	$\boldsymbol{N}$	$\boldsymbol{B}^{-1}$	$\boldsymbol{x}_B$	$\boldsymbol{x}_N$	基本解
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2\end{pmatrix}$	$\begin{pmatrix}\boldsymbol{P}_3 & \boldsymbol{P}_4\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}5 & -2\\ -2 & 1\end{pmatrix}$	$\begin{pmatrix}-2 & 2\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}0 & 0\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}-2 & 2 & 0 & 0\end{pmatrix}^{\mathrm{T}}$
$\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_3\end{pmatrix}$	$\begin{pmatrix}\boldsymbol{P}_2 & \boldsymbol{P}_4\end{pmatrix}$	$\frac{1}{2}\begin{pmatrix}3 & -1\\ -2 & 1\end{pmatrix}$	$\begin{pmatrix}0 & 2\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}0 & 0\end{pmatrix}^{\mathrm{T}}$	$\begin{pmatrix}0 & 0 & 2 & 0\end{pmatrix}^{\mathrm{T}}$

这说明 $\boldsymbol{x}+\lambda\boldsymbol{d}$ 仍然满足原约束, 只需 $\boldsymbol{x}+\lambda\boldsymbol{d}\geq\boldsymbol{0}$, 这说明 $\boldsymbol{x}$ 不唯一, 因此不是基本可行解.

充分性: 若 $\boldsymbol{x}$ 为可行解, 且 $\boldsymbol{P}_1,\boldsymbol{P}_2,\cdots,\boldsymbol{P}_k$ 线性无关, 则可令

$$\boldsymbol{B}=\begin{pmatrix}\boldsymbol{P}_1 & \boldsymbol{P}_2 & \cdots & \boldsymbol{P}_k\end{pmatrix},$$

此时 $\boldsymbol{x}$ 可分为 $\boldsymbol{x}_B$ 和 $\boldsymbol{x}_N$ 两部分, 且 $\boldsymbol{x}_B=\boldsymbol{B}^{-1}\boldsymbol{b}$. 由于 $\boldsymbol{x}_B$ 唯一确定, 则 $\boldsymbol{x}_N$ 也唯一确定且 $\boldsymbol{x}_N=\boldsymbol{0}$, 因此 $\boldsymbol{x}$ 是基本可行解.

9. a) 将问题化为标准型:

$$\begin{cases}\max & 6x_1+14x_2+13x_3\\ \text{s.t.} & x_1+4x_2+2x_3+x_4=48,\\ & x_1+2x_2+4x_3+x_5=60,\\ & x_1,x_2,x_3,x_4,x_5\geq 0.\end{cases}$$

使用单纯形表, 如表 4 所示. 因此原问题的最优解为 $\boldsymbol{x}=\begin{pmatrix}36 & 0 & 6\end{pmatrix}^{\mathrm{T}}$, 最优值为 294.

b) 将问题化为标准型:

$$\begin{cases}\max & -3x_1+2x_2+4x_3\\ \text{s.t.} & 4x_1+5x_2-2x_3+x_4=22,\\ & x_1-2x_2+x_3+x_5=30,\\ & x_1,x_2,x_3,x_4,x_5\geq 0.\end{cases}$$

使用单纯形表, 如表 5 所示. 因此原问题的最优解为 $\boldsymbol{x}=\begin{pmatrix}0 & 82 & 194\end{pmatrix}^{\mathrm{T}}$, 最优值为 940.

表 4

$\boldsymbol{c}_B$	$\boldsymbol{x}_B$	$\boldsymbol{b}$	6	14	13	0	0	$\boldsymbol{\theta}_i$
			x_1	x_2	x_3	x_4	x_5	
0	x_4	48	1	(4)	2	1	0	12
0	x_5	60	1	2	4	0	1	30
$-z$		0	6	14*	13	0	0	
14	x_2	12	$\frac{1}{4}$	1	$\frac{1}{2}$	$\frac{1}{4}$	0	24
0	x_5	36	$\frac{1}{2}$	0	(3)	$-\frac{1}{2}$	1	12
$-z$		-168	$\frac{5}{2}$	0	6*	$-\frac{7}{2}$	0	
14	x_2	6	$(\frac{1}{6})$	1	0	$\frac{1}{3}$	$-\frac{1}{6}$	36
13	x_3	12	$\frac{1}{6}$	0	1	$-\frac{1}{6}$	$\frac{1}{3}$	72
$-z$		-240	$\frac{3}{2}$*	0	0	$-\frac{5}{2}$	-2	
6	x_1	36	1	6	0	2	-1	
13	x_3	6	0	-1	1	$-\frac{1}{2}$	$\frac{1}{2}$	
$-z$		-294	0	-9	0	$-\frac{11}{2}$	$-\frac{1}{2}$	

c) 将问题化为标准型:

$$\begin{cases} \max & x_1 + x_2 + x_3 \\ \text{s.t.} & -x_1 - x_3 + x_4 = 5, \\ & 2x_1 - 3x_2 + x_3 + x_5 = 3, \\ & 2x_1 - 5x_2 + 6x_3 + x_6 = 5, \\ & x_1, x_2, x_3, x_4, x_5, x_6 \geq 0. \end{cases}$$

使用单纯形表, 如表 6 所示. 因此最优解不存在.

10. 使用大 M 法, 得到

$$\begin{cases} \max & 4x_1 + 2x_2 + 8x_3 - Mx_5 \\ \text{s.t.} & 2x_1 - x_2 + 3x_3 + x_4 = 30, \\ & x_1 + 2x_2 + 4x_3 + x_5 = 40, \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{cases}$$

使用单纯形表, 如表 7 所示.

因此原问题的最优解为 $\boldsymbol{x} = \begin{pmatrix} 20 & 10 & 0 \end{pmatrix}$, 最优值为 100.

表 5

c_B	x_B	b	-3	2	4	0	0	θ_i
			x_1	x_2	x_3	x_4	x_5	
0	x_4	22	4	5	-2	1	0	
0	x_5	30	1	-2	(1)	0	1	30
$-z$		0	-3	2	4*	0	0	
0	x_4	82	6	(1)	0	1	2	82
4	x_3	30	1	-2	1	0	1	
$-z$		-120	-7	10*	0	0	-4	
2	x_2	82	6	1	0	1	2	
4	x_3	194	13	0	1	2	5	
$-z$		-940	-67	0	0	-10	-24	

表 6

c_B	x_B	b	1	1	1	0	0	0	θ_i
			x_1	x_2	x_3	x_4	x_5	x_6	
0	x_4	5	-1	0	1	1	0	0	
0	x_5	3	(2)	-3	1	0	1	0	$\frac{3}{2}$
0	x_6	5	2	-5	6	0	0	1	$\frac{5}{2}$
$-z$		0	1*	1	1	0	0	0	
0	x_4	$\frac{13}{2}$	0	$-\frac{3}{2}$	$\frac{3}{2}$	1	$\frac{1}{2}$	0	
1	x_1	$\frac{3}{2}$	1	$-\frac{3}{2}$	$\frac{1}{2}$	0	$\frac{1}{2}$	0	
0	x_6	2	0	-2	5	0	-1	1	
$-z$		0	0	$\frac{5}{2}$*	$\frac{1}{2}$	0	$-\frac{1}{2}$	0	

使用两阶段法, 第一阶段得到

$$\begin{cases} \max & -x_4 - x_5 \\ \text{s.t.} & 2x_1 - x_2 + 3x_3 + x_4 = 30, \\ & x_1 + 2x_2 + 4x_3 + x_5 = 40, \\ & x_1, x_2, x_3, x_4, x_5 \geq 0. \end{cases}$$

如表 8 所示, 使用单纯形表得到原问题的基本可行解 $\boldsymbol{x} = \begin{pmatrix} 0 & 0 & 10 \end{pmatrix}$.

如表 9 所示, 使用单纯形表继续计算, 得到原问题的最优解 $\boldsymbol{x} = \begin{pmatrix} 20 & 10 & 0 \end{pmatrix}$, 最优

表 7

$\boldsymbol{c}_B$	$\boldsymbol{x}_B$	$\boldsymbol{b}$	4	2	8	0	$-M$	$\boldsymbol{\theta}_i$
			x_1	x_2	x_3	x_4	x_5	
0	x_4	30	2	-1	3	1	0	10
$-M$	x_5	40	1	2	(4)	0	1	10
$-z$		0	$M+4$	$2M+2$	$4M+8*$	0	0	
0	x_4	0	$(\frac{5}{4})$	$-\frac{5}{2}$	0	1	$-\frac{3}{4}$	0
8	x_3	10	$\frac{1}{4}$	$\frac{1}{2}$	1	0	$\frac{1}{4}$	40
$-z$		-80	$2*$	-2	0	0	$-M-2$	
4	x_1	0	1	-2	0	$\frac{4}{5}$	$-\frac{3}{5}$	
8	x_3	10	0	(1)	1	$-\frac{1}{5}$	$\frac{2}{5}$	10
$-z$		-80	0	$2*$	0	$-\frac{8}{5}$	$-M-\frac{4}{5}$	
4	x_1	20	1	0	2	$\frac{2}{5}$	$\frac{1}{5}$	
2	x_2	10	0	1	1	$-\frac{1}{5}$	$\frac{2}{5}$	
$-z$		-100	0	0	-2	$-\frac{6}{5}$	$-M-\frac{8}{5}$	

值为 100.

11. 对于线性规划问题

$$(LP)\quad \begin{cases} \max & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{cases}$$

其对偶问题为

$$(DP)\quad \begin{cases} \min & \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y} \geq \boldsymbol{c}, \\ & \boldsymbol{y} \geq \boldsymbol{0}, \end{cases}$$

则 DP 的对偶问题为

$$\begin{cases} \max & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \leq \boldsymbol{b}, \\ & \boldsymbol{x} \geq \boldsymbol{0}, \end{cases}$$

证毕.

12. 原问题与其对偶问题之间的核心联系体现在对偶定理.

表 8

c_B	x_B	b	0	0	0	-1	-1	θ_i
			x_1	x_2	x_3	x_4	x_5	
-1	x_4	30	2	-1	3	1	0	10
-1	x_5	40	1	2	(4)	0	1	10
$-z$		70	3	1	7*	0	0	
0	x_3	10	$\frac{2}{3}$	$-\frac{1}{3}$	1	$\frac{1}{3}$	0	
-1	x_5	0	$-\frac{5}{3}$	$\left(\frac{10}{3}\right)$	0	$-\frac{4}{3}$	1	0
$-z$		0	$-\frac{5}{3}$	$\frac{10}{3}*$	0	$-\frac{7}{3}$	0	
0	x_3	10	$\frac{1}{2}$	0	1	$\frac{1}{5}$	$\frac{1}{10}$	
0	x_2	0	$-\frac{1}{2}$	1	0	$-\frac{2}{5}$	$\frac{3}{10}$	
$-z$		0	0	0	0	-1	-1	

表 9

c_B	x_B	b	4	2	8	θ_i
			x_1	x_2	x_3	
8	x_3	10	$\left(\frac{1}{2}\right)$	0	1	20
2	x_2	0	$-\frac{1}{2}$	1	0	
$-z$		-80	1*	0	0	
4	x_1	20	1	0	2	
2	x_2	10	0	1	1	
$-z$		-100	0	0	-2	

定理 1 (弱对偶定理). *对于任何原问题的可行解 $\boldsymbol{x}$ 及其对偶问题的可行解 $\boldsymbol{y}$, 总有*

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \leq \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$$

定理 2 (强对偶定理). *若原问题及其对偶问题均有可行解且分别为 $\boldsymbol{x}^*$ 和 $\boldsymbol{y}^*$, 则*

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}^* = \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}^*$$

原问题与其对偶问题之间的区别在于:

- 目标函数的区别
  - 原问题是最大化问题 $\max \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x}$.
  - 其对偶问题是最大化问题 $\min \boldsymbol{b}^{\mathrm{T}}\boldsymbol{y}$.

- 约束条件的转换
  - 原问题的约束右端项 $\boldsymbol{b}$ 变成其对偶目标函数的系数.
  - 原问题的目标函数系数 $\boldsymbol{c}$ 变成其对偶约束的右端项.
  - 原问题的约束方向 $\leq$ 变成对偶问题的 $\geq$.
- 变量的物理意义

  在资源分配问题中, 原问题变量 $\boldsymbol{x}$ 表示实际分配方案, 而其对偶变量 $\boldsymbol{y}$ 可以解释为资源的影子价格, 表示单位资源的价值.
- 计算难度

  若原问题有 n 个变量和 m 个约束, 当 $m \ll n$ 时求解对偶问题可能更简单.

13. a) 错.

    考虑问题

$$\begin{cases} \max & x_1 + x_2 \\ \text{s.t.} & x_1 + x_2 \leq 2, \\ & x_1, x_2 \geq 0. \end{cases}$$

    易知该问题的最优解解集合为 $\{(x_1, x_2) \mid x_1 + x_2 = 2, x_1, x_2 \geq 0\}$, 其中 $(2, 0)$ 和 $(0, 2)$ 为基本可行解, 因此最优解不一定是最优基本可行解.

    b) 对.

    线性规划问题的可行域是凸集, 且目标函数为线性. 若存在多个不同的最优解 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$, 则它们的凸组合也是最优解, 因此形成无穷多个解.

    c) 错.

    考虑问题

$$\begin{cases} \max & x_1 + x_2 \\ \text{s.t.} & x_1 + x_2 \leq 2, \\ & x_1, x_2 \geq 0. \end{cases}$$

    易知该问题的最优解解集合为 $\{(x_1, x_2) \mid x_1 + x_2 = 2, x_1, x_2 \geq 0\}$. 考虑最优解 $(1, 1)$, 它有两个变量值为正, 因此不一定存在最多 m 个变量值为正.

    d) 对.

    可行域内的顶点构成基本可行解, 其最多有 m 个正分量. 非基变量共有 $n - m$ 个, 因此基变换过程中可选择的入基变量最多有 $n - m$ 个, 因此每个顶点至多有 $n - m$ 个相邻顶点.

14. 见附件.

15. 代码见附件.

设各长度的钢管数量构成 $\boldsymbol{x} \in \mathbb{R}^3$, 使得

$$\begin{pmatrix} 2.9 & 2.1 & 1.5 \end{pmatrix} \boldsymbol{x} \le 7.4, \quad x_1, x_2, x_3 \text{为正整数}.$$

遍历所有可能的情况, 得到如表 10 中所示的 21 种可能的方案, 表中的前三列构成矩阵 $\boldsymbol{A} \in \mathrm{R}^{3\times 21}$, 第五列构成向量 $\boldsymbol{c} \in \mathrm{R}^{21}$.

表 10　制作 100 套钢架问题中可能的方案

2.9m 钢管数量	2.1m 钢管数量	1.5m 钢管数量	消耗长度 (m)	剩余长度 (m)	方案执行次数
0	0	1	5.9	1.5	0
0	0	2	4.4	3.0	0
0	0	3	2.9	4.5	0
0	0	4	1.4	6.0	0
0	1	0	5.3	2.1	0
0	1	1	3.8	3.6	0
0	1	2	2.3	5.1	0
0	1	3	0.8	6.6	0
0	2	0	3.2	4.2	0
0	2	1	1.7	5.7	0
0	2	2	0.2	7.2	0
0	3	0	1.1	6.3	0
1	0	0	4.5	2.9	0
1	0	1	3.0	4.4	0
1	0	2	1.5	5.9	0
1	0	3	0.0	7.4	30
1	1	0	2.4	5.0	0
1	1	1	0.9	6.5	0
1	2	0	0.3	7.1	50
2	0	0	1.6	5.8	0
2	0	1	0.1	7.3	10

设每种方案执行次数构成 $\boldsymbol{y} \in \mathrm{R}^{21}$, 为了使材料最省, 则只需让 $\boldsymbol{c}^{\mathrm{T}}\boldsymbol{y}$ 最小, 因此得到优化问题

$$\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{y} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{y} = 100\boldsymbol{J}, \\ & y_1, y_2, \cdots, y_{21}\text{为正整数}, \end{cases} \tag{1}$$

其中 $\boldsymbol{J} \in \mathrm{R}^{21}$ 为全 1 向量.

使用 Julia 编写程序并求解公式 (1) , 得到表 10 中的第六列, 该列表示公式 (1) 的最

优解 $\boldsymbol{y}^*$.

因此, 最优值为

$$\boldsymbol{c}^{\mathrm{T}}\boldsymbol{y}^* = 16,$$

即浪费了 16m 的钢管, 共消耗了 $\boldsymbol{J}^{\mathrm{T}}\boldsymbol{y}^* = 90$ 根钢管.

```
21× 6 DataFrame
 Row │ 2.9m   2.1m   1.5m   Spend    Remain   Count
     │ Int64  Int64  Int64  Float64  Float64  Int64
─────┼──────────────────────────────────────────────
   1 │     0      0      1      5.9      1.5      0
   2 │     0      0      2      4.4      3.0      0
   3 │     0      0      3      2.9      4.5      0
   4 │     0      0      4      1.4      6.0      0
   5 │     0      1      0      5.3      2.1      0
   6 │     0      1      1      3.8      3.6      0
   7 │     0      1      2      2.3      5.1      0
   8 │     0      1      3      0.8      6.6      0
   9 │     0      2      0      3.2      4.2      0
  10 │     0      2      1      1.7      5.7      0
  11 │     0      2      2      0.2      7.2      0
  12 │     0      3      0      1.1      6.3      0
  13 │     1      0      0      4.5      2.9      0
  14 │     1      0      1      3.0      4.4      0
  15 │     1      0      2      1.5      5.9      0
  16 │     1      0      3      0.0      7.4     30
  17 │     1      1      0      2.4      5.0      0
  18 │     1      1      1      0.9      6.5      0
  19 │     1      2      0      0.3      7.1     50
  20 │     2      0      0      1.6      5.8      0
  21 │     2      0      1      0.1      7.3     10
Total: 90
Min waste:16.0
```

图 8

程序运行输出如图 8 所示.

16. (a) 线性优化与非线性优化

线性优化形如

$$\begin{cases} \max & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} \le \boldsymbol{b}, \end{cases}$$

它的目标函数和约束条件都是线性的.

非线性优化形如

$$\begin{cases} \max & f(\boldsymbol{x}) \\ \text{s.t.} & g_i(\boldsymbol{x}) \le 0, \\ & h_j(\boldsymbol{x}) = 0. \end{cases}$$

它的目标函数或约束条件中至少有一个是非线性的.

线性优化是非线性优化的一种特例，即非线性优化更一般.

(b) 凸优化与非凸优化

凸优化需要满足三点: (1) 目标函数是凸函数; (2) 约束是凸集; (3) 具有唯一全局最优解. 非凸优化的目标函数或约束可能是非凸的.

凸优化是非凸优化的一个特例, 很多非凸问题可以通过凸化转化为凸优化问题.

(c) 光滑优化与非光滑优化

光滑优化中的目标函数可微, 并且通常具有连续的一阶或高阶导数. 非光滑优化中的目标函数不可微, 如 $|x|$ 在 0 点处不可导.

许多非光滑优化可以通过光滑化处理, 使其变成光滑优化.

(d) 线性化: 通过泰勒展开, 一阶近似等方式, 把非线性问题转化为线性问题, 如

$$f(\boldsymbol{x}) \approx f(\boldsymbol{x}_0) + \nabla^{\mathrm{T}} f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)$$

可将非线性函数近似成线性函数.

凸化: 凸化是使非凸优化问题变成凸优化问题的常见方法, 如松弛, 变量变换, 凸包逼近等.

光滑化: 通过引入平滑函数逼近非光滑函数, 如 σ 平滑

$$H(x) = \begin{cases} \dfrac{1}{2}x^2, & |x| \le \sigma, \\ \sigma\left(|x| - \dfrac{1}{2}\sigma\right), & |x| > \sigma, \end{cases}$$

或 Softmax 近似等.

(e) 近年来, 随着深度学习的快速发展, 非凸优化也更加受到重视.

神经网络的损失函数通常是非凸的, 且神经网络权重空间往往复杂, 存在多个局部最优和鞍点, 在训练过程使用的梯度下降方法并不保证找到全局最优解. 但近年来的研究发现, 即使深度学习的优化问题是非凸的, 实践中仍然可以找到可接受的解决方案, 比如有研究表明, 大规模神经网络的局部最优点往往表现良好.

17. 内点法的核心思想为通过迭代更新, 从多面体的内部逐渐接近最优解, 而不是像单纯形法那样沿着可行域的边界移动.

对于标准的线性规划问题

$$\begin{cases} \min & \boldsymbol{c}^{\mathrm{T}}\boldsymbol{x} \\ \text{s.t.} & \boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}, \\ & \boldsymbol{x} \ge \boldsymbol{0}, \end{cases}$$

它的可行域为

$$\mathcal{F}=\{\boldsymbol{x}\mid \boldsymbol{A}\boldsymbol{x}=\boldsymbol{b},\boldsymbol{x}\geq\boldsymbol{0}\}.$$

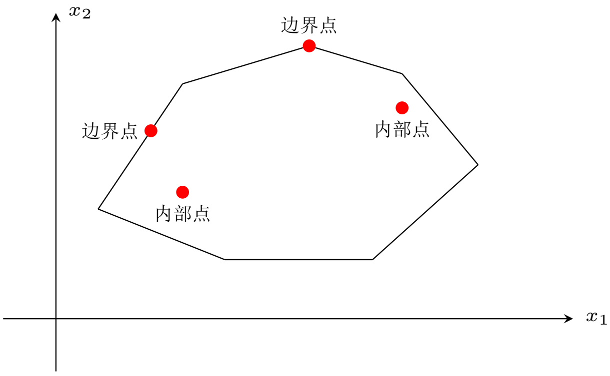
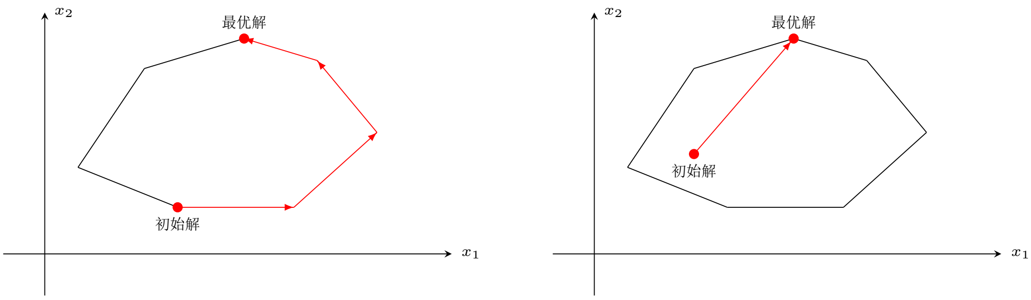


图 9

如图 9 所示, 可行域 $\mathcal{F}$ 大致分为边界和内部.



(a) 单纯形法　　(b) 内点法

图 10　不同方法搜索过程示意图

如图 10 所示, 单纯形法往往从多面体的一个顶点 (基本可行解) 出发, 沿着边界寻找最优解, 而内点法是从多面体的内部出发, 沿着搜索方向寻找最优解, 因此通常情况下, 内点法比单纯形法的效率更高.