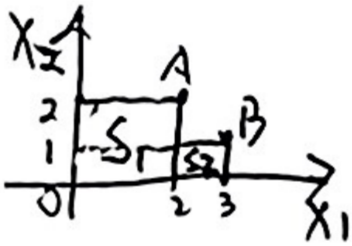
1.

(a) $S_1 \cup S_2$ 不一定为凸集
反例:
$S_1 = \{x \in R^2 | 0 \le x_1 \le 2, 0 \le x_2 \le 2\}$
$S_2 = \{x \in R^2 | 2 \le x_1 \le 3, 0 \le x_2 \le 1\}$



其中A与B的连线 不在 $S_1 \cup S_2$ 中固不为凸集

(b) $S_1 + S_2 = \{x + y | x \in S_1, y \in S_2\}$ 是凸集

证明:

对 $S_1 + S_2$ 中任意两个元素 z_1, z_2

$z_1 = x_1 + y_1, x_1 \in S_1, y_1 \in S_2 \Rightarrow z_1 \in S_1 + S_2$

$z_2 = x_2 + y_2, x_2 \in S_1, y_2 \in S_2 \Rightarrow z_2 \in S_1 + S_2$

又对于任意 $\theta \in [0, 1]$

$\theta z_1 + (1 - \theta) z_2$

$= \theta(x_1 + y_1) + (1 - \theta)(x_2 + y_2)$

$= [\theta x_1 + (1 - \theta) x_2] + [\theta y_1 + (1 - \theta) y_2]$

$\because x_1, x_2 \in S_1, y_1, y_2 \in S_2$

$\theta x_1 + (1 - \theta) x_2 \in S_1, \theta y_1 + (1 - \theta) y_2 \in S_2$

$\therefore \theta z_1 + (1 - \theta) z_2 \in S_1 + S_2$

因为 $z_1, z_2 \in S_1 + S_2$, 所以 $S_1 + S_2$ 是凸集。

(c) $S_1 - S_2 = |x - y| \quad x \in S_1, y \in S_2$ 是凸集.

与(b)相似

取 $S_1 - S_2$ 中任意两点 z_1, z_2

$z_1 = x_1 - y_1, \quad x_1 \in S_1, y_1 \in S_2$

$z_2 = x_2 - y_2, \quad x_2 \in S_1, y_2 \in S_2.$

对任点 $\theta \in [0, 1]$

$\theta z_1 + (1 - \theta) z_2 = \theta(x_1 - y_1) + (1 - \theta)(x_2 - y_2)$

$= [\theta x_1 + (1 - \theta) x_2] - [\theta y_1 + (1 - \theta) y_2]$

$\because S_1$ 和 S_2 是凸集.

$\therefore \theta x_1+(1-\theta)x_2 \in S_1$

$\theta y_1+(1-\theta)y_2 \in S_2.$

$\because [\theta x_1+(1-\theta)x_2]-[\theta y_1+(1-\theta)y_2] \in S_1-S_2$

即 $\theta z_1+(1-\theta)z_2 \in S_1-S_2$

$\because z_1, z_2 \in S_1-S_2 \quad \therefore S_1-S_2$ 为凸集.

2.

(a) $S=\{(x_1,x_2)|x_1^2+x_2^2 \geq 4, 2x_1+x_2 \leq 10, -x_1-x_2 \leq -10x_2\}$ 不为凸集
证明:
取两点 $\left(\frac{18}{\sqrt{82}}, \frac{2}{\sqrt{82}}\right)$ 和 $(2,0)$
易知两点均属于 S
取 $\theta=\frac{1}{2}$, 记构造的新点为 z, 坐标计算如下:
$\frac{1}{2} \times (\frac{18}{\sqrt{82}}, \frac{2}{\sqrt{82}})+\frac{1}{2}(2,0)$
$=(\frac{9}{\sqrt{82}}+1, \frac{1}{\sqrt{82}})$
易知 $(\frac{9}{\sqrt{82}}+1)^2+(\frac{1}{\sqrt{82}})^2=2+\frac{18}{\sqrt{82}}<2+2=4$
因可知点 z 不属于 S, 因 S 不为凸集

(b) 是凸集

取两点 (x_1,x_2) 和 (y_1,y_2) 属于 S, 且 $\theta \in [0,1]$

$\theta(x_1,x_2)+(1-\theta)(y_1,y_2)=(\theta x_1+(1-\theta)y_1, \theta x_1+(1-\theta)y_2)$

分别判断条件:

①$\theta x_1+\theta x_2+(1-\theta)y_1+(1-\theta)y_2 \leq \theta(x_1+x_2)+(1-\theta)(y_1+y_2)<6$

② $-2[\theta x_1+(1-\theta)y_1]+3[\theta x_2+(1-\theta)y_2]=\theta(-2x_1+3x_2)+(1-\theta)(-2y_1+3y_2) \geq 2$

③ $4[\theta x_1+(1-\theta)y_1]-[\theta x_2+(1-\theta)y_2]=\theta(4x_1-x_2)+(1-\theta)(4y_1-y_2) \leq 12$

综上可知构造的新点满足 S条件，根据定义可知 S 为凸集

(c) 是凸集.

证明:
取两点 (x_1,y_1) 和 (x_2,y_2) 属于 S, $\theta \in [0,1]$

构造新点 $z=(z_1,z_2)=(\theta x_1+(1-\theta)y_1, \theta x_2+(1-\theta)y_2)$

下证新点 $z \in S$

① $-(x_1-1)^2+x_2 \geq 1 \Rightarrow x_2 \geq 1+(x_1-1)^2$

由琴生不等式

$\theta[1+(x_1-1)^2]+(1-\theta)[1+(y_1-1)^2] \geq 1+[\theta x_1+(1-\theta)y_1-1]^2$

又 $x_2 \geq 1+(x_1-1)^2$

$\therefore \theta x_2+(1-\theta)y_2 \geq \theta[1+(x_1-1)^2]+(1-\theta)[1+(y_1-1)^2] \geq 1+[\theta x_1+(1-\theta)y_1-1]^2$

即 $z_2 \geq 1 + [z_1 - 1]^2 \therefore -(z_1 - 1)^2 + z_2 \geq 1$

② $z_1 + z_2 = \theta(x_1 + x_2) + (1-\theta)(y_1 + y_2) \geq 3$

③ $\theta x_1 + (1-\theta) y_1 \geq \theta + 1 - \theta = 1$

综上所述,$(z_1, z_2) \in S$

所以由定义可知 S 为凸集

(d) 不是凸集

理由:
取 $x_1 = -1, x_2 = 1, x_1, x_2 \in S$
取 $\theta = \frac{1}{2}$ 可得 $z = \theta x_1 + (1-\theta)x_2 = 0$
$z^2 = 0 < 1 \therefore z \notin S$
$\therefore S$ 不为凸集

3.

(a) 是凸函数.
取 $x_1, x_2, \lambda \in [0,1), x_3 = \lambda x_1 + (1-\lambda)x_2$
由 $f_1(x)$ 为凸函数, 有 $f_1(x_3) \leq \lambda f_1(x_1) + (1-\lambda) f_1(x_2)$
$f_2, \cdots f_m$ 与 f_1 同理
$g(x) = \max\{f_1(x), f_2(x), \cdots f_m(x)\}$ 不妨设 $g(x_3) = f_1(x_3)$
则 $g(x_3) = f_1(x_3) \leq \lambda f_1(x_1) + (1-\lambda) f_1(x_2) \leq \lambda g(x_1) + (1-\lambda) g(x_2)$
故 $g(x)$ 是凸函数

(b) 不是凸函数.

反例:
$g(x) = \min\{f_1(x) = x^2, f_2(x) = (x-1)^2\}$

显然 $g(x)$ 此时不为凸函数.

$\because$ 对 $x_1 = 0, x_2 = 1, \lambda = \frac{1}{2}$

$\lambda g(x_1) + (1-\lambda) g(x_2) = 0$

$g(\lambda x_1 + (1-\lambda)x_2) = g(\frac{1}{2}) = \frac{1}{4}$

$g(\lambda x_1 + (1-\lambda)x_2) > \lambda g(x_1) + (1-\lambda) g(x_2)$

与定义不符.

(c) 是凸函数

取 $x=(x_1,x_2,\ldots,x_n), y=(y_1,y_2,\ldots,y_n), \lambda\in[0,1)$

得 $\lambda x+(1-\lambda)y=(\lambda x_1+(1-\lambda)y_1, \lambda x_2+(1-\lambda)y_2,\ldots,\lambda x_n+(1-\lambda)y_n)$

记作 $z=(z_1,z_2,\ldots,z_n)$

易知 $z\in R^n, z_i\geq 0, \sum_{i=1}^n z_i=1$

$g(z)=\sum_{i=1}^n z_i log(z_i)$

$\lambda g(x)+(1-\lambda)g(y)=\sum_{i=1}^n \lambda x_i log(x_i)+(1-\lambda)y_i log(y_i)$

即证 $z_i log(z_i)\leq \lambda x_i log(x_i)+(1-\lambda)y_i log(y_i)$

即证 $xlog(x)$函数为凸函数

由二阶条件易知其为凸函数，固原函数为凸函数

(d)不为凸函数

取 $x_1=-\frac{1}{2}, x_2=\frac{3}{2}, \lambda=\frac{3}{4}\in(0,1)$

$g(\lambda x_1+(1-\lambda)x_2)=g(0)=0$

$\lambda g(x_1)+(1-\lambda)g(x_2)=\frac{3}{4}\times(-1)+\frac{1}{4}\times 1=-\frac{1}{2}$

$g(\lambda x_1+(1-\lambda)x_2)>\lambda g(x_1)+(1-\lambda)g(x_2)$

由定义可知 $g(x)$ 不为凸函数

(e) 为凸函数

取 x 和 $y\in R^n$，$x=(x_1,x_2,\ldots,x_k), y=(y_1,y_2,\ldots,y_k)$

记前 k 个最大$x(i)$为 $a_1,a_2,\ldots,a_k$，$1\leq k\leq n$

前 k 个最大$y(i)$为 $b_1,b_2,\ldots,b_k$，$1\leq k\leq n$

记 $\lambda x+(1-\lambda)y$ 为 z，其中 $\lambda\in(0,1)$

z的前 k 个最大$z(i)$为 $c_1,c_2,\ldots,c_k$

$g(x)=a_1+a_2+\ldots+a_k$

$g(y)=b_1+b_2+\ldots+b_k$

$g(z)=c_1+c_2+\ldots+c_k$

而 $\lambda g(x)+(1-\lambda)g(y)=\lambda(a_1+\ldots+a_k)+(1-\lambda)(b_1+\ldots+b_k)$

因为 $g(x)$ 为 x 中前 k 个最大的，$g(y)$ 为 y 中前 k 个最大的，

而 $g(z)$ 为 $(\lambda x_1+(1-\lambda)y_1,\ldots,\lambda x_n+(1-\lambda)y_n)$ 中前 k 个最大的

$\Rightarrow \lambda g(x)+(1-\lambda)g(y)\geq g(z)$（最大值相加的组合一定大于等于组合后对应最大值的相加）

固 $g(x)$ 为凸函数

(f) 是凸函数。以下是证明过程：
对于任意$x,y\in\mathbb{R}$ 和 $\lambda\in[0,1]$，需证明：

$$g(\lambda x+(1-\lambda)y)\le \lambda g(x)+(1-\lambda)g(y).$$

展开左边：

$$[\lambda x+(1-\lambda)y]^2=\lambda^2x^2+2\lambda(1-\lambda)xy+(1-\lambda)^2y^2.$$

右边为：

$$\lambda x^2+(1-\lambda)y^2.$$

计算右边减去左边：

$$\lambda x^2+(1-\lambda)y^2-[\lambda^2x^2+2\lambda(1-\lambda)xy+(1-\lambda)^2y^2]=\lambda(1-\lambda)(x-y)^2\ge 0.$$

由于 $\lambda(1-\lambda)\ge 0$ 且 $(x-y)^2\ge 0$，因此右边恒大于等于左边，即：

$$g(\lambda x+(1-\lambda)y)\le \lambda g(x)+(1-\lambda)g(y).$$

故 $g(x)=x^2$ 满足凸函数的定义。

4.

(a).

标准型:

$$\max z'=2x_1'+x_2-2(x_3'-x_4)+0x_5$$
$$s.t.\begin{cases}x_1'+x_2+x_3'-x_4=4\\ x_1'+x_2-(x_3'-x_4)+x_5=6\\ x_1',x_2,x_3',x_4,x_5\ge 0.\end{cases}$$

系数矩阵为

$$\begin{pmatrix}1&1&1&-1&0\\1&1&-1&1&1\end{pmatrix}$$

选择基

$$\begin{aligned}
&B=(P_1,P_2)\text{ 不可逆}\\
&B=(P_1,P_3)\quad X_0=(5,0,-1,0,0)\text{ 不可行}\\
&B=(P_1,P_4)\quad X_1=(5,0,0,1,0)\quad z'=12.\\
&B=(P_1,P_5)\quad X_2=(4,0,0,0,2)\quad z'=8\\
&B=(P_2,P_3)\quad X_3=(0,5,-1,0,0)\text{ 不可行}\\
&B=(P_2,P_4)\quad X_4=(0,5,0,1,0)\quad z'=7\\
&B=(P_2,P_5)\quad X_5=(0,4,0,0,2)\quad z'=4\\
&\cdots
\end{aligned}$$

列举可知 $X_1=(5,0,0,1,0)$ 为最优解，对应 $z'=12.$

(b).

标准型:

$$\max z=-2x_1+x_2'-3x_3-(x_4'-x_5)+0x_6+0x_7$$
$$s.t.\begin{cases}x_1-x_2'+x_3+(x_4'-x_5)+x_6=7\\ 2x_1+3x_2'+5x_3=-8\\ x_1-2x_3+2(x_4'-x_5)-x_7=1\\ x_1,x_2,x_3,x_4,x_5,x_6,x_7\ge 0.\end{cases}$$

系数矩阵为

$$\begin{pmatrix}1 & -1 & 1 & 1 & -1 & 1 & 0\\2 & 3 & 5 & 0 & 0 & 0 & 0\\1 & 0 & -2 & 2 & -2 & 0 & -1\end{pmatrix}$$

$B=(P_1,P_2,P_3)$不可行
$B=(P_1,P_2,P_4)$不可行
$B=(P_1,P_2,P_6)$不可行

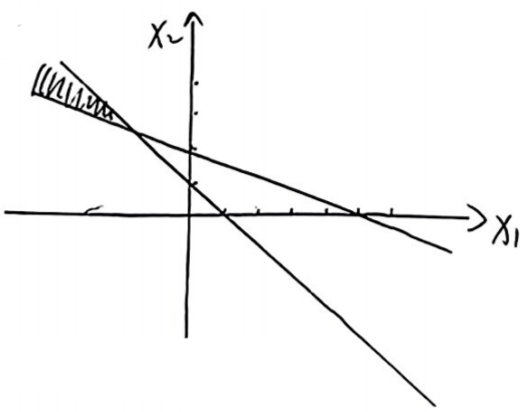
故综上可知
原问题无最优解
显然

$$\begin{cases}2x_1+3x_2'+5x_3=-8\\x_1,x_2',x_3\geq 0\end{cases}$$

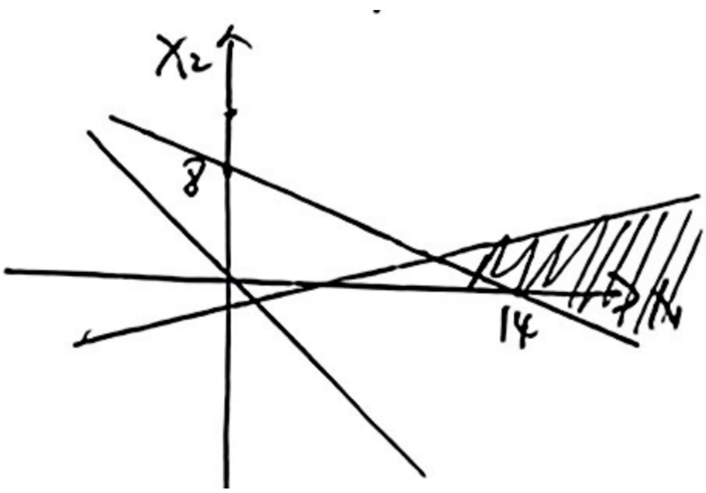
冲突

5.

(a)



由图可知无可行解
(b)



有可行解，由于阴影可行域无界，无最优解。

6.

原式可写为 $\begin{pmatrix} 1 & 2 & 3 & -1 & 1 \\ -1 & 1 & -3 & 2 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 4 \\ -5 \end{pmatrix}$

基本解有:
① 基变量为 x_1, x_2则基本解为 $(\frac{14}{3}, -\frac{1}{3}, 0, 0, 0)$
② 基变量为 x_1, x_4 则基本解为 $(3, 0, 0, -1, 0)$
③ 基变量为 x_2, x_3 则其解为 $(0, -\frac{1}{3}, \frac{14}{9}, 0, 0)$

7.

$A = \begin{pmatrix} 2 & 4 & 2 & 2 \\ 4 & 10 & 6 & 8 \end{pmatrix} \quad b = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$
基本解1:
$B = \begin{pmatrix} 2 & 4 \\ 4 & 10 \end{pmatrix} \quad B^{-1} = \begin{pmatrix} \frac{5}{2} & -1 \\ -1 & \frac{1}{2} \end{pmatrix}$
$X_B = (-2, 2) \quad X_N = (0, 0)$
基础解为 $(-2, 2, 0, 0)$
基本解2:
$B = \begin{pmatrix} 2 & 2 \\ 4 & 6 \end{pmatrix} \quad B^{-1} = \begin{pmatrix} \frac{3}{2} & -\frac{1}{2} \\ -1 & \frac{1}{2} \end{pmatrix}$
$X_B = (0, 2) \quad X_N = (0, 0)$
基础解为 $(0, 0, 2, 0)$

8.证明:

设 $X = (x_1, x_2 \cdots x_n)^T$ 是LP的一个可行解，且满足 $AX = b$
$A = (P_1, P_2 \cdots P_n)$

充分性
设 $X = (x_1, x_2 \cdots x_k, 0 \cdots 0)^T$ 为一个可行解
$x_1, x_2 \cdots x_k$ 为正分量，对应系数列向量线性无关
$R(A) = m \;\; k \leq m$

(1) 若 $k = m$，$x_i \geq 0 (i = 1, 2 \cdots n)$ 由定义知 x 为基本可行解。

(2) 若 $k < m$，$R(P_1, P_2 \cdots P_n) = m$ 则在 $P_{k+1} \cdots P_n$ 中必然存在 $m - k$ 个分量与 $P_1 \cdots P_k$ 线性无关，记作$(P_{k+1}, P_{k+2} \cdots P_m)$
则有基$B = (P_1 \cdots P_m)$ $x_i > 0 (i = 1, 2 \cdots k)$
有 $X_B = (x_1, x_2 \cdots x_k, 0 \cdots 0)^T$ 故 X 为基本可行解
必要性:
若 X 为基本可行解，设 $X_B = (x_1, x_2 \cdots x_m)^T$
$B = (P_1 \cdots P_m)$ $X_B = (x_{m+1} \cdots x_n) = 0$
设 X_B 中有 k 个正分量，$k \leq m$，对应矩阵的系数列向量为 $(q_1 \cdots q_k)$，$q_1, q_2 \cdots q_k$ 均为 B 的列向量
因由 B 可逆，故 $q_1, q_2 \cdots q_k$ 线性无关
由以上可知，原命题得证。

9.

(a)

标准化为$max 6x_1+14x_2+13x_3+0x_4+0x_5$

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

		c_j	6	14	13	0	0	
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_4	48	1	4	2	1	0	12
0	x_5	60	1	2	4	0	1	30
		σ_j	6	14	13	0	0	θ
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
		σ_j	$\frac{5}{2}$	0	6	$-\frac{7}{2}$	0	θ
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
		σ_j	$\frac{3}{2}$	0	0	$-\frac{5}{2}$	-2	θ
6	x_1	36	1	6	0	2	-1	
13	x_3	6	0	-1	1	$-\frac{1}{2}$	$\frac{1}{2}$	
			0	-9	0	$-\frac{11}{2}$	$-\frac{1}{2}$	

综上有唯一最优解。为（36，0，6，0，0）

$maxz=6*36+6*13=294$

(b)标准化为

$max-3x_1+2x_2+4x_3$

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

C_j		-3	2	4	0	0	
C_B X_B	b	x_1	x_2	x_3	x_4	x_5	θ
0 x_4	22	4	5	-2	1	0	
0 x_5	70	1	-2	1	0	1	30
σ_j		-3	2	4	0	0	θ
0 x_4	82	6	1	0	1	2	82
4 x_3	30	1	-2	1	0	1	
σ_j		-7	10	0	0	-4	
2 x_2	82	6	1	0	1	2	
4 x_3	194	13	0	1	2	5	
		-67	0	0	-10	-16	

故有唯一最优解

$x=(0,82,194,0,0)\ \max z=940$

(c)

标准型：$\max z=x_1+x_2+x_3+0x_4+0x_5+0x_6$

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

C_j		1	1	1	0	0	0	
C_B X_B	b	x_1	x_2	x_3	x_4	x_5	x_6	θ
0 x_4	5	-1	0	-2	1	0	0	-5
0 x_5	3	2	-3	1	0	1	0	$\frac{3}{2}$
0 x_6	5	2	-5	6	0	0	1	$\frac{5}{2}$
σ_j		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	
σ_j		0	$\frac{5}{2}$	$\frac{1}{2}$	0	$-\frac{1}{2}$	0	

因为检验数最大的正数对应系数列向量所有分量均小于0，该LP问题无最优解。

10.

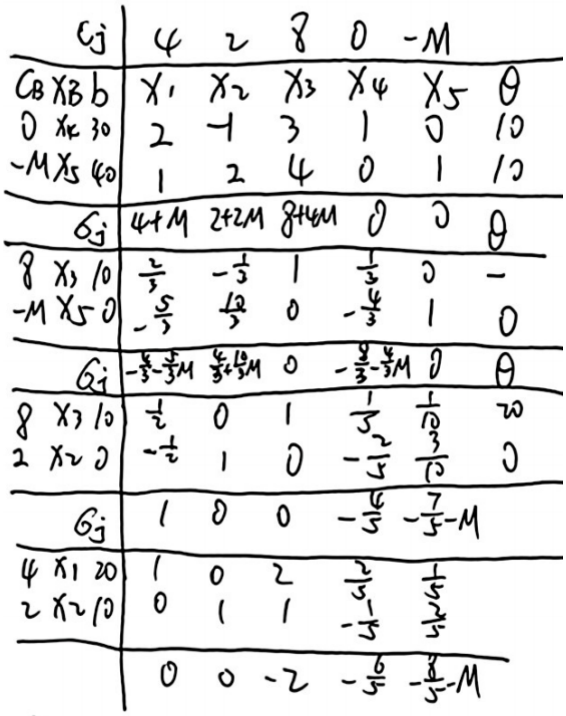
大M法:

标准化为

max $4x_1+2x_2+8x_3+0x_4$

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

引入人工变量 得 max $4x_1+2x_2+8x_3+0x_4-Mx_5$



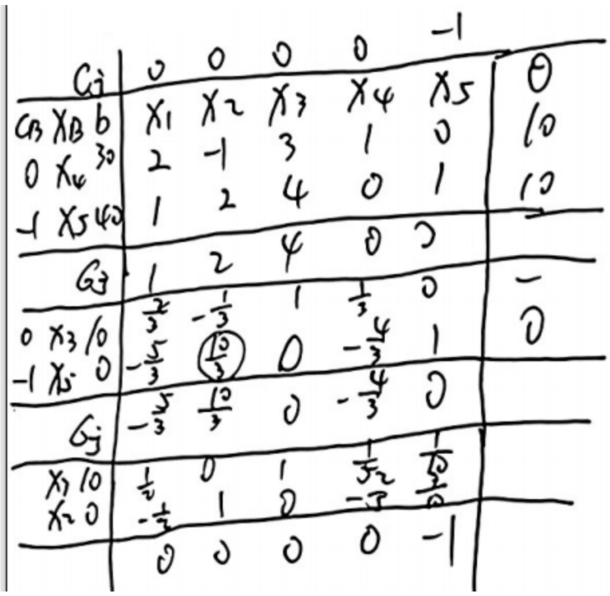
有唯一最优解 $x=(20,10,0,0,0)$

max $z=4\times 20+2\times 10=100$

两阶段法:

第一阶段 $maxz'=-x_5$

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



基本可行解为 $x=(0,0,10,0,0)$

第二阶段

		C_j	4	2	8	0	
C_B	X_B	b	x_1	x_2	x_3	x_4	θ
8	x_3	10	$\frac{1}{2}$	0	1	$\frac{1}{5}$	20
2	x_2	0	$-\frac{1}{2}$	1	0	$-\frac{2}{5}$	—
		σ_j	1	0	0	$-\frac{4}{5}$	
4	x_1	20	1	0	2	$\frac{2}{5}$	
2	x_2	10	0	1	1	$-\frac{1}{5}$	
			0	0	-2	$-\frac{6}{5}$	

唯一最优解$x=(20,10,0,0,0)$

max $z=4\times 20+2\times 10=100$

11.

证明:

设原LP问题为 $\max z=C^TX$ s.t. $\begin{cases}AX\le b\\ x_i\ge 0\end{cases}$

则对偶问题为 $\min z'=b^TY$ s.t. $\begin{cases}A^TY\ge C\\ y_i\ge 0\end{cases}$

求对偶问题的对偶

$\max z''=C^TW$ s.t. $\begin{cases}AW\le b\\ w_i\ge 0\end{cases}$

由上述可知 命题得证.

12.

联系:

①原问题的对偶问题的对偶问题是原问题

②若x,y分别是LP,DP问题的可行解,则$C^Tx\le b'y$

③若原规划LP问题有最优解,则对偶规划问题DP也存最优解,反之亦然,且两者目标函数值相等.

区别:

①原问题最大化目标函数,对偶问题最小化目标函数

②变量个数与约束个数不同

13.

(1) 错误，此时最优解一定可在顶点取得，而X不一定为基本可行解，可能它在两顶点的线段上
(2) 正确
(3) 错误
最优的基本可行解是最多有m个变量值为正
(4) 正确

14.

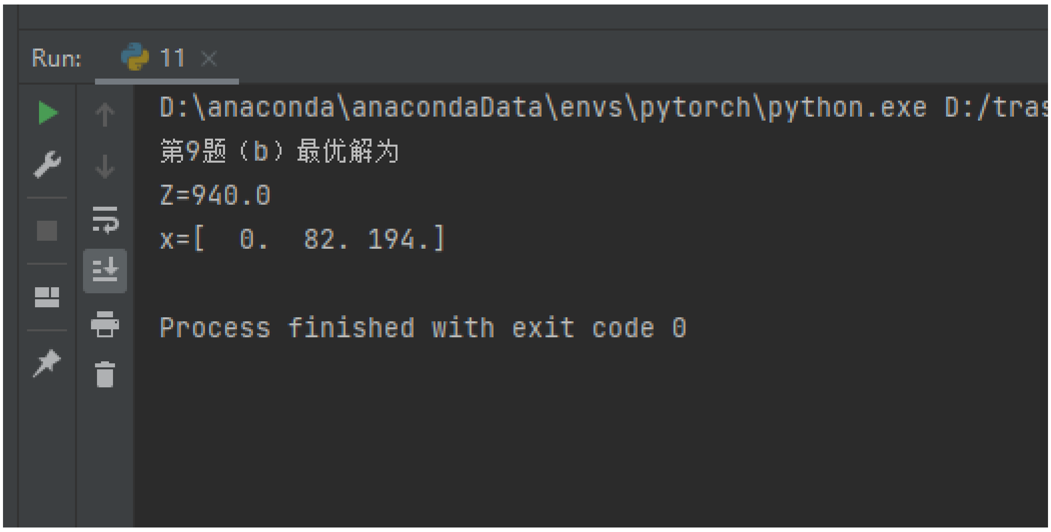
（1）单纯性方法的基本流程：

第一步：构造一个初始基本可行解：对已经标准化的模型，设法从约束矩阵中构造一个m阶的单位阵

第二步判断当前基本可行解是否为最优解：求出用非基变量x_N来表示基变量x_B和目标函数z，这个称为LP问题的典式(canonical form规范式)，将目标函数的典式中非基变量前的系数称为检验数，最大问题所有检验数为非正数，则当前解即为最优解

第三步：若当前解不是最优解，则进行基变换，迭代到下一个基本可行解：进基变量从目标函数典式中选取正的最大检验数对应的非基变量；再从当前基变量中选取一个离基变量，选取准则为除进基变量外，令其余非基变量为0，再按最小比值准则确立离基变量，然后回到第二步循环至找到最优解或者判断无最优解结束。

（2）用第九题（b）检验，结果与计算结果一致，代码在附录1中给出，以下是运行结果截图：



15.

程序运行结果如下所示：

```
Reduced MIP has 0 binaries, 5 generals, 0 SOSs, and 0 indicators.
Presolve time = 0.00 sec. (0.00 ticks)
MIP emphasis: balance optimality and feasibility.
MIP search method: dynamic search.
Parallel mode: deterministic, using up to 12 threads.
Root relaxation solution time = 0.00 sec. (0.01 ticks)

        Nodes                                         Cuts/
   Node  Left     Objective  IInf  Best Integer    Best Bound    ItCnt     Gap

*     0+    0                         16.0000        1.0000            93.75%
      0     0        cutoff           16.0000                     3    0.00%

Root node processing (before b&c):
  Real time             =    0.01 sec. (0.03 ticks)
Parallel b&c, 12 threads:
  Real time             =    0.00 sec. (0.00 ticks)
  Sync time (average)   =    0.00 sec.
  Wait time (average)   =    0.00 sec.
                          ------------
Total (root+branch&cut) =    0.01 sec. (0.03 ticks)
使用五种方法各自的钢管原材料个数:
方法 1: 30.0
方法 2: 10.0
方法 3: 0.0
方法 4: 50.0
方法 5: 0.0
```

即：分配给方法一 30根原材料；方法二 10根原材料；方法四 50 根原材料

具体代码查看后续代码附录2

16.

线性优化与非线性优化

1. **区别**:
   - **线性优化**（如线性规划，LP）：目标函数和约束均为线性函数，可行域为凸多面体。
   - **非线性优化**（如二次规划，QP）：目标函数或约束中至少有一个是非线性的，可能涉及凸或非凸问题。
2. **联系**:
   - 线性优化是凸优化的特例（线性函数是凸函数）。
   - 非线性优化包含线性优化，但求解难度更高（需处理非线性性）。

凸优化与非凸优化

1. **区别**:
   - **凸优化**：目标函数为凸函数，可行域为凸集。其核心性质是局部最优即全局最优。
   - **非凸优化**：目标函数或可行域非凸，可能存在多个局部最优解，求解困难（如NP-hard问题）。
2. **联系**:
   - 凸优化是非凸优化的特例，但非凸优化无法保证全局最优性。
   - 许多非凸问题可通过凸松弛（如半定规划）或启发式算法（如梯度下降）近似求解。

光滑优化与非光滑优化

1. **区别**:
    - **光滑优化**：目标函数和约束连续可导（如二次函数），可用梯度、牛顿法等。
    - **非光滑优化**（如L1正则化）：目标函数不可导（如绝对值、ReLU），需次梯度、近端梯度法等。
2. **联系**:
    - 非光滑问题可通过光滑化技术（如Moreau-Yosida正则化）转化为光滑问题。

转化方法

1. **线性化**:
    - 用一阶泰勒展开近似非线性函数，例如：
      $f(x) \approx f(x_0) + \nabla f(x_0)^\top (x - x_0)$.
    - 适用场景：局部近似或混合整数规划中的分段线性化。
2. **凸化**:
    - 引入辅助变量或约束，例如将非凸约束 $(xy \geq 1)$转化为$(x \geq a, y \geq b, a + b \geq 2)$。
    - 使用凸松弛（如半定松弛）放宽非凸约束。
3. **光滑化**:
    - 用可导函数逼近非光滑函数，例如用Huber损失替代绝对值损失：
      $$h_\delta(x) = \begin{cases} \frac{x^2}{2\delta} & |x| \leq \delta, \\ |x| - \frac{\delta}{2} & \text{其他}. \end{cases}$$

为什么凸优化与非凸优化是主流?

1. **凸优化的优势**:
    - 理论成熟（如对偶理论、KKT条件），可高效求得全局最优解。
    - 为实际问题提供可靠基准（如支持向量机、压缩感知）。
2. **非凸优化的现实需求**:
    - 实际问题多是非凸的（如神经网络训练、组合优化）。
    - 尽管非凸，但梯度下降等算法在实践中表现优异（如深度学习）。
3. **研究趋势**:
    - 凸优化提供理论工具（如鲁棒性分析），非凸优化推动算法创新（如随机优化、非凸正则化）。
    - 两者结合，例如用凸方法初始化非凸问题，或分析非凸问题的局部解性质。

结论

凸优化与非凸优化共同构成现代优化理论的核心：前者提供理论保障，后者应对复杂现实问题。其分析成为主流，既因实际需求驱动，也因算法与理论的持续突破。

线性规划内点法的核心思想是通过在可行域内部迭代逼近最优解，避免触碰边界，从而高效求解线性规划问题。其核心要点如下：

---

1. 基本思路

- **内部路径追踪**：不同于单纯形法的顶点遍历，内点法通过构造一条从初始内点指向最优解的**中心路径**（Central Path），并沿此路径逐步逼近最优解。
- **障碍函数引入**：通过添加**对数障碍函数**（Logarithmic Barrier）将原始线性规划问题转化为一系列无约束优化问题，强制解点远离可行域边界。

---

2. 关键步骤

1. **问题转化**：
   对于标准线性规划问题：
   $\min\ c^\top x \quad \text{s.t. } Ax = b,\ x \geq 0,$
   引入对数障碍函数处理非负约束：
   $\min\ c^\top x - \mu \sum_{i=1}^n \ln x_i \quad \text{s.t. } Ax = b,$
   其中 $(\mu > 0)$ 为障碍参数，控制解点与边界的距离。
2. **牛顿法迭代**：
   对转化后的无约束问题应用牛顿法：
   - 计算目标函数的梯度与海森矩阵。
   - 通过牛顿方向更新解点：
     $x^{k+1} = x^k + \alpha d^k,$
     其中(d^k)为牛顿方向，(α)为步长（需保证 $(x^{k+1} > 0)$）。
3. **参数更新**：
   逐步减小障碍参数$(\mu \to 0)$，使障碍项的影响减弱，解点趋近原始问题的最优解。

---

3. 核心优势

- **多项式时间复杂度**：内点法的最坏时间复杂度为$(O(n^{3.5}L))$（(L) 为输入规模），优于单纯形法的指数时间风险。
- **大规模问题高效性**：尤其适合高维稀疏问题，计算稳定性强。
- **全局收敛性**：在合理参数调整下，可保证收敛到最优解。

---

4. 主要类型

1. **路径跟踪法**（Path-Following）：
   显式跟踪中心路径，通过调整 μ 控制迭代方向。
2. **势垒函数法**（Barrier Method）：
   隐式依赖障碍函数，通过参数衰减驱动收敛。
3. **原对偶内点法**（Primal-Dual）：
   同时更新原始变量和对偶变量，利用互补松弛条件加速收敛。

---

5. 实际应用

- 内点法广泛应用于大规模线性规划（如供应链优化、金融组合管理）、二次规划及半定规划问题。
- 在机器学习中，支持向量机（SVM）的求解常采用内点法。

总结

内点法的核心在于通过障碍函数将约束问题转化为无约束优化，利用牛顿法在可行域内部高效迭代，并沿中心路径逼近最优解。其理论完备性与计算效率使其成为现代优化算法的重要支柱。

答案：
内点法通过障碍函数和牛顿迭代在可行域内部逼近最优解，具有多项式时间复杂度和全局收敛性

代码附录 1

```
import numpy as np

M = 999999


def simplex(A, b, c, base, n):
    S = np.concatenate([A, b], 1)
    Z = np.concatenate([c[:, 0], [0]], 0)
    while 1:
        # 选择入基出基变量、计算θ
        max_c = max(Z[:-1])
        if max_c <= 0:
            break
        in_base = np.argmax(Z[:-1])  # 列
        theta = [row[-1] / row[in_base] if row[in_base] >= 0 else M for row in S]
        # theta = [i if i >= 0 else M for i in theta]
        if min(theta) == M:
            print("最优解为无穷大")
            return None
        out_base = np.argmin(theta)  # 行

        # 进行基的替换
        base[out_base] = in_base
        scale = 1 / S[out_base][in_base]
        S[out_base] *= scale
        for i in range(S.shape[0]):
            if i == out_base:
                continue
            scale = S[i][in_base] * -1
            S[i] += S[out_base] * scale

        # 更新-Z
        scale = Z[in_base] * -1
        Z += S[out_base] * scale
```

```
        # 输出这一步的结果
        # print(S)
        # print(Z)
    x = np.zeros(n)
    for i in range(n):
        if i in base:
            x[i] = S[np.where(base == i)][0][-1]
    return -Z[-1], x



n2 = 3
A2 = np.array([[4, 5, -2, 1, 0],
              [1, -2, 1, 0, 1]], dtype='float64')
b2 = np.array([[22], [30]], dtype='float64')
c2 = np.array([[-3], [2], [4], [0], [0]], dtype='float64')
base2 = np.array([3, 4])
Z2, x2 = simplex(A2, b2, c2, base2, n2)
print("第9题（b）最优解为\nZ={}\nx={}".format(Z2, x2))
```

代码附录 2

```
using JuMP
using CPLEX

model = Model(CPLEX.Optimizer)

# 声明变量
@variable(model, x[1:5] >= 0, Int)

# 设置约束条件
@constraint(model, 1x[1] + 2x[2] + 1x[4] == 100)
@constraint(model, 2x[3] + 2x[4] + 1x[5] == 100)
@constraint(model, 3x[1] + 1x[2] + 2x[3] + 3x[5] == 100)

# 设置优化目标
@objective(model, Min, 0.1x[2] + 0.2x[3] + 0.3x[4] + 0.8x[5])

# 求解优化问题
optimize!(model)


println("使用五种方法各自的钢管原材料个数:")
for i in 1:5
    println("方法 ", i, ": ", value(x[i]))
end
```