线性规划

线性规划模型

例1 生产计划问题

用 3 种原料混合配制 2 种清洁剂

	原料1	原料2	原料3	售价(万元/吨)
清洁剂 A	0.25	0.50	0.25	12
清洁剂 B	0.50	0.50		15
存量 (吨)	120	150	50	

这 2 种清洁剂应各配制多少才能使总价值最大?

例1(续)

设清洁剂 A 和 B 分别配制 x 和 y

$$\max z = 12x + 15y$$ 目标函数

$$\begin{aligned} \text{s.t.}\quad & 0.25x + 0.50y \le 120 \\ & 0.50x + 0.50y \le 150 \\ & 0.25x \le 50 \\ & x \ge 0,\ y \ge 0 \end{aligned}$$

约束条件

投资组合问题

例2 10亿元投资 5 个项目, 其中项目1和项目2是高新技术产业的企业债, 项目3和项目4是基础工业的企业债, 项目5是国债和地方政府债.

预测它们的年收益率(%)分别为 8.1, 10.5, 6.4, 7.5 和 5.0.

基于风险的考虑, 要求投资组合满足下述条件:

(1) 每个项目不超过 3 亿元.

(2) 高新技术产业的投资不超过总投资的一半, 即 5 亿元, 其中项目2又不超过高新技术产业投资的一半.

(3) 国债和地方政府债不少于基础工业项目投资的 40%.

试确定投资组合中各项目的投资额, 使年收益率最大.

例2 (续)

设项目 i 的投资额为 x_i 亿元，$i = 1, 2, 3, 4, 5.$

$$\max z = 8.1x_1 + 10.5x_2 + 6.4x_3 + 7.5x_4 + 5.0x_5$$

s.t. $x_1 \le 3,\quad x_2 \le 3$

$x_3 \le 3,\quad x_4 \le 3$

$x_5 \le 3$

每个项目不超过3亿元

$x_1 + x_2 \le 5$

高新技术产业的投资不超过总投资的一半，即5亿

$x_2 \le 0.5(x_1+x_2)$，即 $x_1 - x_2 \ge 0$

项目2又不超过高新技术产业投资的一半

$x_5 \ge 0.4(x_3+x_4)$，即 $0.4x_3 + 0.4x_4 - x_5 \le 0$

国债和地方政府不少于基础工业项目投资的 40%

$x_1 + x_2 + x_3 + x_4 + x_5 = 10$

总投资为 10 亿

$x_i \ge 0,\ i = 1, 2, 3, 4, 5$

运输问题

例3

	分销中心1	分销中心2	分销中心3	产量
工厂1	3	2	7	5000
工厂2	7	5	2	6000
需求量	6000	4000	1000	11000

产销平衡. 试制定供销方案, 使总运费最小.

设工厂 i 供应分销中心 j 的数量为 x_{ij}, $i = 1, 2$; $j = 1, 2, 3$.

$$\min\ z = 3x_{11} + 2x_{12} + 7x_{13} + 7x_{21} + 5x_{22} + 2x_{23}$$

$$\begin{aligned} \text{s.t.}\quad & x_{11} + x_{12} + x_{13} = 5000 \\ & x_{21} + x_{22} + x_{23} = 6000 \\ & x_{11} + x_{21} = 6000, \qquad x_{12} + x_{22} = 4000 \\ & x_{13} + x_{23} = 1000, \qquad x_{ij} \geq 0,\quad i = 1, 2;\ j = 1,2,3 \end{aligned}$$

饲料配方问题

例4 每头每天至少需要b_i个单位的营养素 i, $1\le i\le n$. 有n种饲料, 饲料 j 每千克含有 a_{ij}个单位的营养素 i, 售价 c_j 元, $1\le j\le m$. 要保证动物有足够营养且饲料成本最低, 应如何配方?

设每头每天的饲料中含 x_j 千克饲料 j, $1\le j\le n$.

$$\min z=\sum_{j=1}^{n} c_j x_j$$

$$\text{s.t}\quad \sum_{j=1}^{n} a_{ij}x_j \ge b_i\ ,\quad 1\le i\le m$$

$$x_j\ge 0,\quad 1\le j\le n$$

线性规划的一般形式

$$\min(\max)\ z = \sum_{j=1}^{n} c_j x_j \qquad \text{目标函数}$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j \le (=, \ge) b_i, \qquad i=1,2,\ldots,m \qquad \text{约束条件}$$

$$x_j \ge 0, \quad j \in J \subseteq \{1,2,\ldots,n\} \qquad \text{非负条件}$$

$$x_j \text{ 任意}, \ j \in \{1,2,\ldots,n\} - J \qquad \text{自由变量}$$

可行解　满足约束条件和非负条件的变量

可行域　全体可行解

最优解　目标函数值最小(最大)的可行解

最优值　最优解的目标函数值

二维线性规划的图解法

例1 (续)

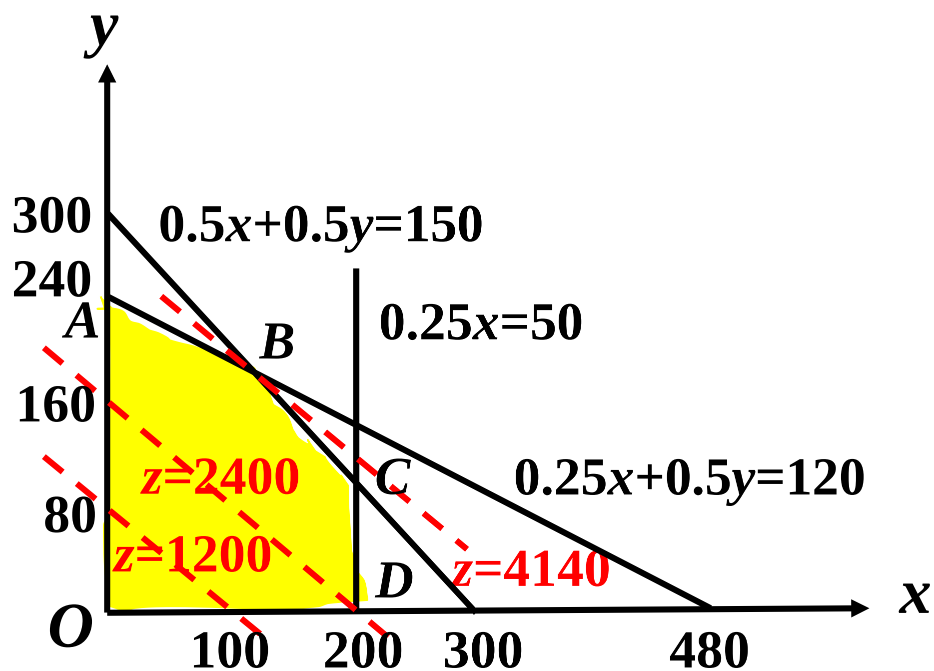
$\max z = 12x + 15y$

s.t. $0.25x + 0.50y \le 120$

$0.50x + 0.50y \le 150$

$0.25x \le 50$

$x \ge 0, y \ge 0$



$O(0,0)$, $A(0,240)$,

$B(120,180)$, $C(200,100)$,

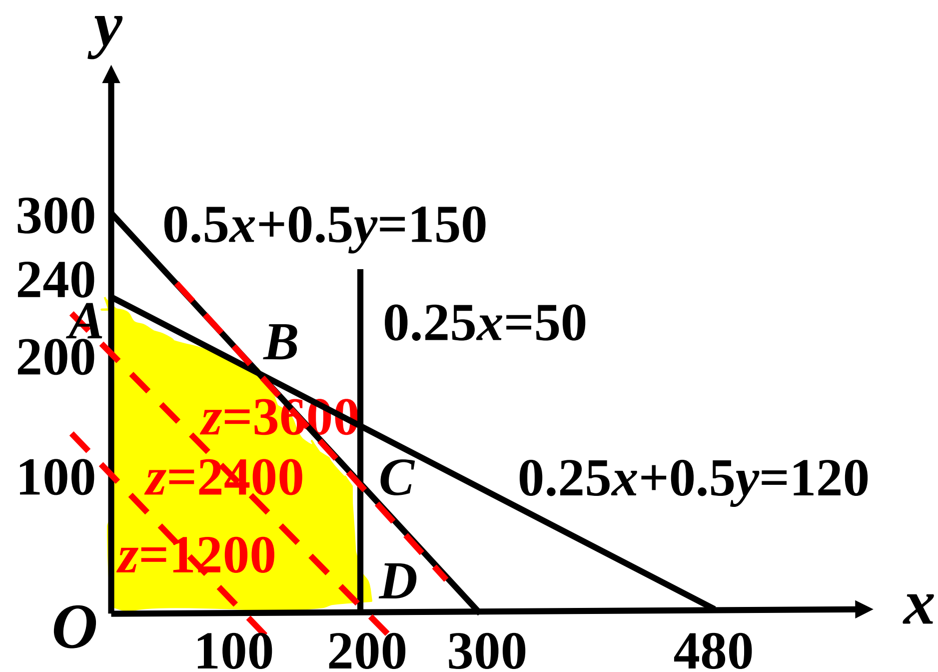
$D(200)$

最优解 $x^*=120$, $y^*=180$ (点B) 最优值 $z^*=4140$.

例5

例1中的目标函数改为

$\max z = 12x + 12y$



最优解 $x^* = 120t + 200(1-t) = 200 - 80t$

$y^* = 180t + 100(1-t) = 100 + 80t$, $0 \le t \le 1$ 线段BC

最优值 $z^* = 3600$.

例6

$\min z=x-2y$

s.t. $2x+y\geq 2$

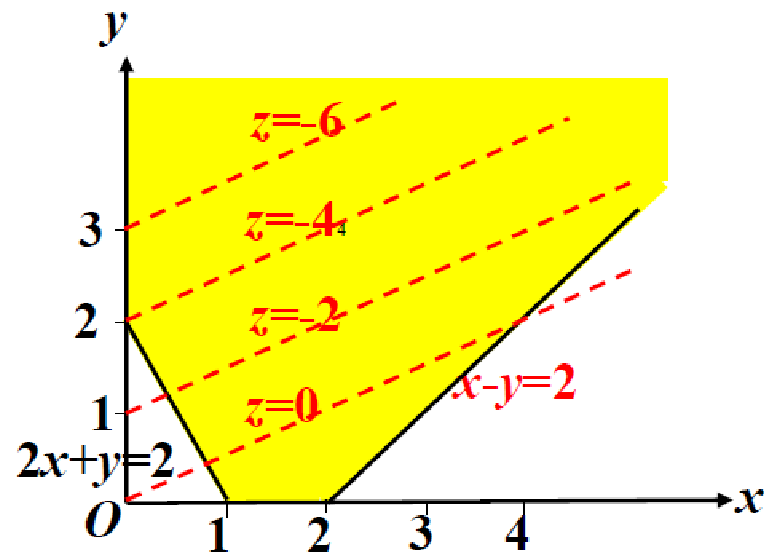
$x-y\leq 2$

$x\geq 0,\ y\geq 0$

有可行解

目标函数值可以任意小

无最优解.



$2x+y\geq 2$ 改为 $2x+y\leq 2$,

$x-y\leq 2$ 改为 $x-y\geq 2$

可行域为空集, 无可行解

性质

(1) 解有4种可能

(a) 有唯一的最优解.

(b) 有无穷多个最优解.

(c) 有可行解, 但无最优解 (目标函数值无界).

(d) 无可行解, 更无最优解.

(2) 可行域是一个凸多边形 (可能无界, 也可能是空集).

如果有最优解, 则一定可以在凸多边形的顶点取到.

一般的 n 维线性规划也是如此.

标准形

$$\min z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij} x_j = b_i \geq 0, \quad i = 1,2,\ldots,m$$

$$x_j \geq 0, \quad j = 1,2,\ldots,n$$

化成标准形

(1) 把 $\max z$ 替换成$\min z' = -z$, 即取 $c_j' = -c_j$.

(2) $b_i<0$. 两边同时变号, $\le$ 改变成 $\ge$, $\ge$ 改变成 $\le$.

(3) $\sum_{j=1}^{n} a_{ij} x_j \le b_i$. 引入松弛变量 $y_i \ge 0$, 替换成

$$\sum_{j=1}^{n} a_{ij} x_j + y_i = b_i$$

(4) $\sum_{j=1}^{n} a_{ij} x_j \ge b_i$. 引入剩余变量 $y_i \ge 0$, 替换成

$$\sum_{j=1}^{n} a_{ij} x_j - y_i = b_i$$

(5) 自由变量 x_j 替换成 $x_j' - x_j''$, $x_j' \ge 0$, $x_j'' \ge 0$.

例7

写出下述线性规划的标准形

$$\max z = 3x_1 - 2x_2 + x_3$$

$$\text{s.t.}\quad x_1 + 3x_2 - 3x_3 \le 10$$

$$4x_1 - x_2 - 5x_3 \le -30$$

$x_1 \ge 0,\ x_2 \ge 0,\ x_3$ 任意

松弛变量

解

$$\min z' = -3x_1 + 2x_2 - x_3' + x_3''$$

剩余变量

$$\text{s.t.}\quad x_1 + 3x_2 - 3x_3' + 3x_3'' + x_4 = 10$$

$$-4x_1 + x_2 + 5x_3' - 5x_3'' - x_5 = 30$$

$$x_1 \ge 0,\ x_2 \ge 0,\ x_3' \ge 0,\ x_3'' \ge 0,\ x_4 \ge 0,\ x_5 \ge 0,$$

标准形

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad c=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

矩阵形式

$$\begin{aligned} &\min\ z=c^T x \\ &\text{s.t.}\quad Ax=b \\ &\qquad\quad x\geq 0 \end{aligned}$$

向量形式

$$\begin{aligned} &\min\ z=\sum_{j=1}^{n} c_j x_j \\ &\text{s.t.}\quad \sum_{j=1}^{n} P_j x_j = b \\ &\qquad\quad x_j\geq 0,\ j=1,2,\ldots,n \end{aligned}$$

$$P_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

标准形的可行解的性质

定义 设A的秩为m, A的m个线性无关的列向量称作标准形的基. 给定基$B=(P_{i_1},P_{i_2},\cdots,P_{i_m})$, 对应基中列向量的变量 $x_{i_1},x_{i_2},\cdots,x_{i_m}$ 称作基变量, 其余的变量称作非基变量.

基变量构成的向量记作x_B, 非基变量构成的向量记作x_N. 令$x_N=0$, 等式约束变成

$$Bx_B=b$$

解得$x_B=B^{-1}b$. 这个向量x满足约束$Ax=b$且非基变量全为0, 称作关于基B的基本解. 如果x是一个基本解且$x\geq 0$, 则称x是一个基本可行解, 对应的基B为可行基.

例8

$$\min z = -12x_1 - 15x_2$$

$$\text{s.t.}\quad \begin{cases} 0.25x_1 + 0.50x_2 + x_3 = 120 \\ 0.50x_1 + 0.50x_2 + x_4 = 150 \\ 0.25x_1 + x_5 = 50 \\ x_i \geq 0, \quad i = 1,2,\ldots,5 \end{cases}$$

$$A = \begin{bmatrix} 0.25 & 0.50 & 1 & 0 & 0 \\ 0.50 & 0.50 & 0 & 1 & 0 \\ 0.25 & 0 & 0 & 0 & 1 \end{bmatrix}$$

取基$B_1=(P_1,P_2,P_3)$. 基变量x_1, x_2, x_3,非基变量x_4, x_5. 令$x_4=0$, $x_5=0$, 得

$$\begin{aligned} 0.25x_1 + 0.50x_2 + x_3 &= 120 \\ 0.50x_1 + 0.50x_2 &= 150 \\ 0.25x_1 &= 50 \end{aligned}$$

解得 $x_1=200$, $x_2=100$, $x_3=20$.

$x^{(1)}=(200,100,20,0,0)^T$是基本可行解, B_1是可行基.

例8 (续)

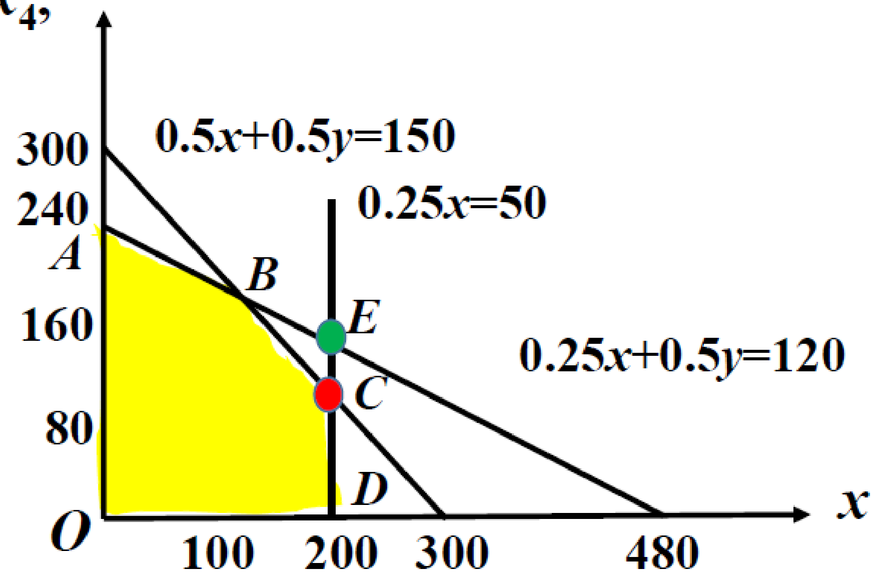
取基$B_2=(P_1,P_2,P_4)$. 基变量 x_1,x_2,x_4, 非基变量 x_3,x_5. 令$x_3=0$, $x_5=0$, 由

$$\begin{aligned} 0.25x_1+0.50x_2 \quad &=120 \\ 0.50x_1+0.50x_2+x_4 &=150 \\ 0.25x_1 \qquad\qquad &=50 \end{aligned}$$

解得$x_1=200$, $x_2=140$, $x_4=-20$.

$x^{(2)}=(200,140,0,-20,0)^T$是基本解, 但不是基本可行解.

B_2不是可行基.



这个线性规划是例1中的标准形, $x^{(1)}$是例1图中的顶点C. $x^{(2)}$是直线$0.25x+0.5y=120$与$0.25x=50$的交点, 不在可行域内.

基本可行解的性质

引理1 $Ax=b$的解α是基本解$\Leftrightarrow$ α中非零分量对应的列向量线性无关.

证 必要性 根据基本解的定义, 这是显然的.

充分性 设α的非零分量为 $\alpha_{j_1},\alpha_{j_2},\cdots,\alpha_{j_r}$, 对应的列向量

$P_{j_1},P_{j_2},\cdots,P_{j_r}$ 线性无关. A的秩为m, 必存在 $P_{j_{r+1}},\cdots,P_{j_m}$

使得 $P_{j_1},P_{j_2},\cdots,P_{j_m}$线性无关, 构成一个基, 记作$B$. α是方程$Bx_B=b$的解, 而这个方程的解是惟一的, 故α是关于B的基本解.

基本可行解的性质

定理1 如果标准形有可行解, 则必有基本可行解.

证 设 α 是一个可行解, 从α开始, 构造出一个基本可行解.

设α的非零分量为 $\alpha_1, \alpha_2, \ldots, \alpha_r$, $r \le n$. 如果对应的列向量 $P_1, P_2, \ldots, P_r$ 线性无关, 则 α 是一个基本可行解.

否则, 存在不全为 0 的 $\lambda_1, \lambda_2, \ldots, \lambda_r$ 使

$$\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_r P_r = 0$$

取 $\lambda_{r+1} = \ldots = \lambda_n = 0$, 有

$$\lambda_1 P_1 + \lambda_2 P_2 + \ldots + \lambda_r P_n = 0$$

于是, 对任意的δ有

$$\sum_{j=1}^{n} (\alpha_j + \delta\lambda_j) P_j = \sum_{j=1}^{n} \alpha_j P_j + \delta \sum_{j=1}^{n} \lambda_j P_j = b$$

定理1 (续)

记 $\lambda=(\lambda_1,\lambda_2,\ldots,\lambda_n)^T$, 为使$\alpha+\delta\lambda$ 成为一个可行解, 要求所有 $\alpha_j+\delta\lambda_j\geq 0$.

当 $\lambda_j=0$ 时, 不等式自然成立.

当 $\lambda_j>0$ 时, 要求 $\delta\geq -\alpha_j/\lambda_j$; 当 $\lambda_j<0$时, 要求 $\delta\leq -\alpha_j/\lambda_j$.

综上所述, 要求当 $\lambda_j\neq 0$ 时, $\delta\leq|\alpha_j/\lambda_j|$.

设 $\left|\alpha_{j_0}/\lambda_{j_0}\right|=\min\left\{\alpha_j/\lambda_j:\lambda_j\neq 0\right\}$, $1\leq j_0\leq r$

取 $\delta^*=-\alpha_{j_0}/\lambda_{j_0}$,令$\beta_j=\alpha_j+\delta^*\lambda_j$ ($j=1, 2, \ldots, n$), 则

$$\beta_1P_1+\ldots+\beta_nP_n=b$$

且 $\beta_j\geq 0$ ($j=1, 2, \ldots, n$), $\beta_{j_0}=0$, $\beta_{r+1}=\ldots=\beta_n=0$.

从而, $\beta=(\beta_1,\beta_2,\ldots,\beta_n)$是可行解且比$\alpha$至少少一个非零分量.

上述过程至多进行r-1次一定可以得到一个基本可行解.

基本可行解的性质

定理2 如果标准形有最优解, 则必存在一个基本可行解是最优解.

证 补充证明: 在定理1证明中, 当α是最优解时, β也是最优解. 由$\alpha_j = 0 \Rightarrow \lambda_j = 0$, 对足够小的$\delta > 0$, $\alpha + \delta\lambda$ 和 $\alpha - \delta\lambda$ 都是可行解. 从而

$$\sum_{j=1}^{n} c_j\alpha_j \le \sum_{j=1}^{n} c_j(\alpha_j + \delta\lambda_j) = \sum_{j=1}^{n} c_j\alpha_j + \delta\sum_{j=1}^{n} c_j\lambda_j$$

$$\sum_{j=1}^{n} c_j\alpha_j \le \sum_{j=1}^{n} c_j(\alpha_j - \delta\lambda_j) = \sum_{j=1}^{n} c_j\alpha_j - \delta\sum_{j=1}^{n} c_j\lambda_j$$

得 $\sum_{j=1}^{n} c_j\lambda_j = 0.$

定理2 (续)

于是

$$\sum_{j=1}^{n} c_j\beta_j \le \sum_{j=1}^{n} c_j(\alpha_j+\delta^*\lambda_j)=\sum_{j=1}^{n} c_j\alpha_j+\delta^*\sum_{j=1}^{n} c_j\lambda_j=\sum_{j=1}^{n} c_j\alpha_j$$

得证 $\beta=\alpha+\delta^*\lambda$ 也是最优解.

根据定理2, 解线性规划问题只需考虑标准形的基本可行解. A有m行n列, 至多有 C_n^m个基, 故至多有 C_n^m 个基本解. 从而, 线性规划成为一个组合优化问题.

单纯形法

基本步骤

(1) 确定初始基本可行解.

(2) 检查当前的基本可行解.

若是最优解或无最优解, 计算结束;

否则作基变换, 用一个非基变量替换一个基变量, 得到一个新的可行基和对应的基本可行解, 且使目标函数值下降(至少不升).

(3) 重复(2).

确定初始基本可行解

暂时只考虑最简单的情况, 设约束条件为

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n \le b_i \ge 0, \quad i = 1, 2, \ldots, m$$

其中 $b_i \ge 0$ $(i = 1, 2, \ldots, m)$.

引入 m 个松弛变量 $x_{n+i} \ge 0$ $(i = 1, 2, \ldots, m)$,

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{in}x_n + x_{n+i} = b_i, \quad i = 1, 2, \ldots, m$$

取 x_{n+i} $(i = 1, 2, \ldots, m)$ 作为基变量, 初始基本可行解为

$$x^{(0)} = (0, 0, \ldots, 0, b_1, b_2, \ldots, b_m)^T$$

例

例1 (续)

$$\max z = 12x+15y$$

$$\text{s.t.}\quad \begin{aligned} 0.25x + 0.50y &\le 120 \\ 0.50x + 0.50y &\le 150 \\ 0.25x \qquad\quad &\le 50 \\ x \ge 0,\ y &\ge 0 \end{aligned}$$

标准形

$$\min z' = -12x_1-15x_2$$

$$\text{s.t.}\quad \begin{aligned} 0.25x_1 + 0.50x_2 + x_3 \qquad\qquad &= 120 \\ 0.50x_1 + 0.50x_2 \qquad + x_4 \qquad &= 150 \\ 0.25x_1 \qquad\qquad\qquad\qquad + x_5 &= 50 \\ x_i \ge 0,\quad i = 1, 2, \ldots, 5 & \end{aligned}$$

取 x_3, x_4, x_5作为基变量

$$x^{(0)} = (0, 0, 120, 150, 50)^T$$

最优性检验

给定可行基 $B=(P_{\pi(1)}, P_{\pi(2)},\ldots, P_{\pi(m)})$, $Ax=b$ 两边同乘B^{-1}, 得 $B^{-1}Ax=B^{-1}b$. 记A中对应非基变量的列构成的矩阵为 N,

$$x_B + B^{-1}Nx_N = B^{-1}b$$

解得 $$x_B = B^{-1}b - B^{-1}Nx_N$$

代入目标函数 $$z = c^Tx$$

$$= c_B^T x_B + c_N^T x_N$$

$$= c_B^T (B^{-1}b - B^{-1}Nx_N) + c_N^T x_N$$

$$= c_B^T B^{-1}b + (c_N^T - c_B^T B^{-1}N)x_N$$

基本可行解 $x_B^{(0)}=B^{-1}b$, $x_N^{(0)} = 0$, 目标函数值 $z_0 = c_B^T B^{-1}b$

最优性检验

$$z = c^T x$$

$$= z_0 + (c_N^T - c_B^T B^{-1} N) x_N$$

$$= z_0 + (c_B^T - c_B^T B^{-1} B) x_B + (c_N^T - c_B^T B^{-1} N) x_N$$

$$= z_0 + (c^T - c_B^T B^{-1} A) x$$

记 $\lambda^T = c^T - c_B^T B^{-1} A$ 检验数

$$z = z_0 + \lambda^T x$$ 简化的目标函数

最优性检验

记 $B^{-1}A=(\alpha_{ij})_{m\times n}$, $P_j'=B^{-1}P_j$ $(1\le j\le n)$, $\beta=B^{-1}b$.

定理3 给定基本可行解 $x^{(0)}$, 若所有检验数大于等于0, 则 $x^{(0)}$ 是最优解. 若存在检验数 $\lambda_k<0$且所有$\alpha_{ik}\le 0$ $(1\le i\le m)$, 则无最优解.

证 如果 $\lambda\ge 0$, 则对任意可行解 $x\ge 0$, $z\ge z_0$, 故$x^{(0)}$是最优解. 如果存在检验数 $\lambda_k<0$ (λ_k必对应非基变量) 且所有$\alpha_{ik}\le 0$ $(1\le i\le m)$, 取 $x_k=M>0$, 其余非基变量 $x_j=0$, 代入

$$x_B=B^{-1}b-B^{-1}Nx_N$$

解得 $x_{\pi(i)}=\beta_i-\alpha_{ik}M\ge 0,\qquad 1\le i\le m$

这是一个可行解，其目标函数值为

$$z=z_0+\lambda_k M$$

当 $M\to+\infty$ 时，$z\to-\infty$. 得证无最优解.

基变换

给定可行基 $B=(P_{\pi(1)},P_{\pi(2)},\ldots,P_{\pi(m)})$, 设 $\lambda_k< 0$且$\alpha_{lk} > 0$, x_k必是非基变量.

基变换: 用非基变量 x_k 替换基变量 $x_{\pi(l)}$, 用 P_k 替换 B 中的 $P_{\pi(l)}$, 新的基为 $B' = (P_{\pi(1)},\ldots, P_{\pi(l-1)}, P_k, P_{\pi(l+1)},\ldots, P_{\pi(m)})$.
称 x_k为换入变量, $x_{\pi(l)}$为换出变量.

(1) 要证 B' 是一个基, 即$P_{\pi(1)},\ldots, P_{\pi(l-1)}, P_k, P_{\pi(l+1)},\ldots, P_{\pi(m)}$是线性无关的. 由于$P_{\pi(1)},P_{\pi(2)},\ldots,P_{\pi(m)}$是线性无关的, 只需证$P_{\pi(l)}$可表成 $P_{\pi(1)},\ldots,P_{\pi(l-1)}, P_k, P_{\pi(l+1)},\ldots,P_{\pi(m)}$ 的线性组合.
由于 $(P'_{\pi(1)}, P'_{\pi(2)}, \ldots , P'_{\pi(m)}) = B^{-1}B = E$,

$$P'_k = \sum_{i=1}^{m} \alpha_{ik} P'_{\pi(i)} \quad \xrightarrow{\text{两边同乘}B} \quad P_k = \sum_{i=1}^{m} \alpha_{ik} P_{\pi(i)}$$

基变换（续）

解得
$$P_{\pi(l)} = \frac{1}{\alpha_{lk}} P_k - \sum_{\substack{i=1 \\ 且 \neq l}}^{m} \frac{\alpha_{ik}}{\alpha_{lk}} P_{\pi(i)}$$

得证 B' 是一个基.

(2) 要保证 B' 是可行基.

$$Ax = b$$

$$B^{-1}Ax = B^{-1}b = \beta$$

$$x_B + B^{-1}Nx_N = \beta$$

$B^{-1}A=(P_1',P_2',\ldots,P_m')$中对应 x_B 的列 (第$\pi(1),\ldots,\pi(n)$列) 构成单位矩阵. 用 P_k 替换 $P_{\pi(l)}$得到 B', 将 x_B 中的 $x_{\pi(l)}$ 替换成 x_k , 即解出第 l 个方程中的 x_k. 这只需用 α_{lk} 除第 l 个方程, 再用第 l 个方程消去其它方程中的 x_k

基变换（续）

$$
H=\begin{bmatrix}
1 & & & \alpha_{1k} & & & \\
 & \ddots & & \vdots & & & \\
 & & 1 & \alpha_{l-1,k} & & & \\
 & & & \alpha_{lk} & & & \\
 & & & \alpha_{l+1,k} & 1 & & \\
 & & & \vdots & & \ddots & \\
 & & & \alpha_{mk} & & & 1
\end{bmatrix}
$$

第 l 列

H 是把单位矩阵的第 l 列换成 P'_k. 由上面 P_k 的表达式，有 $B'=BH$. 于是，

$$B'^{-1}Ax = B'^{-1}b$$

可写成

$$H^{-1}B^{-1}Ax = H^{-1}B^{-1}b = H^{-1}\beta$$

基变换（续）

即，把基变量 $x_{\pi(l)}$ 换成非基变量 x_k 相当于用 H^{-1} 乘式(6.5)的两边．而

$$
H^{-1}=\begin{bmatrix}
1 & & & -\alpha_{1k}/\alpha_{lk} & & & \\
 & \ddots & & \vdots & & & \\
 & & 1 & -\alpha_{l-1,k}/\alpha_{lk} & & & \\
 & & & 1/\alpha_{lk} & & & \\
 & & & -\alpha_{l+1,k}/\alpha_{lk} & 1 & & \\
 & & & \vdots & & \ddots & \\
 & & & -\alpha_{mk}/\alpha_{lk} & & & 1
\end{bmatrix}
$$

第 l 列

故用 H^{-1} 乘式(6.5)的两边就是将式(6.5)中的第 l 个方程除以 α_{lk}，第 i 个方程减第 l 个方程的 α_{ik}/α_{lk} 倍($1\leqslant i\leqslant m$ 且 $i\neq l$)．这样做的结果是把式(6.5)中第 l 个方程中 x_l 的系数变成 1，再消去其他方程中的 x_l．计算公式如下，

基变换（续）

计算公式 $\alpha_{lj}' = \alpha_{lj}/\alpha_{lk}, \quad 1 \le j \le n$

$\alpha_{ij}' = \alpha_{ij} - \alpha_{ik}\alpha_{lj}/\alpha_{lk}, \quad 1 \le i \le m$ 且 $i \ne l, 1 \le j \le n$

$\beta_l' = \beta_l/\alpha_{lk}$

$\beta_i' = \beta_i - \alpha_{ik}\beta_l/\alpha_{lk}, \quad 1 \le i \le m$ 且 $i \ne l$

要保证 B' 是可行的, 只需

$$\beta_i' = \beta_i - \alpha_{ik}\beta_l/\alpha_{lk} \ge 0, \quad 1 \le i \le m \text{ 且 } i \ne l$$

注意到 $\beta_i \ge 0$, $\beta_l \ge 0$, $\alpha_{lk} > 0$, 当 $\alpha_{ik} \le 0$ 时不等式自然成立.
于是, 只需当 $\alpha_{ik} > 0$ 时 有 $\beta_l/\alpha_{lk} \le \beta_i/\alpha_{ik}$

应取 l 使得 $\beta_l/\alpha_{lk} = \min\{\beta_i/\alpha_{ik} \mid \alpha_{ik} > 0, 1 \le i \le m\}$

用第 l 个方程消去简化的目标函数中的 x_k,

$$\lambda_j' = \lambda_j - \lambda_k\alpha_{lj}/\alpha_{lk}, \quad 1 \le j \le m$$

$$z_0' = z_0 + \lambda_k\beta_l/\alpha_{lk}$$

单纯形法

算法 单纯形法 (针对最小化)

1. 设初始可行基 $B = (P_{\pi(1)}, P_{\pi(2)}, \ldots, P_{\pi(m)})$, $\alpha = B^{-1}A$,
   $\beta = B^{-1}b$, $\lambda^T = c^T - c_B^T B^{-1}A$, $z_0 = c_B^T B^{-1}b$.
2. 若所有 $\lambda_j \geq 0$ $(1 \leq j \leq n)$, 则 $x_B = \beta$, $x_N = 0$是最优解, 计算结束
3. 取 $\lambda_k < 0$. 若所有 $\alpha_{ik} \leq 0 (1 \leq i \leq m)$, 则无最优解, 计算结束.
4. 取 l 使得

$$\beta_l / \alpha_{lk} = \min\{ \beta_i / \alpha_{ik} \mid \alpha_{ik} > 0,\ 1 \leq i \leq m \}$$

5. 以 x_k 为换入变量、$x_{\pi(l)}$为换出变量做基变换.
6. 转2.

对最大化, 2中 $\lambda_j \geq 0$ 改为 $\lambda_j \leq 0$, 3中 $\lambda_k < 0$ 改为 $\lambda_k > 0$.

标准形

$$A=\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ & \cdots\cdots & & \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \quad b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} \quad c=\begin{bmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{bmatrix} \quad x=\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

矩阵形式

$$\min z=c^T x$$
$$\text{s.t.} \quad Ax=b$$
$$x\geq 0$$

向量形式

$$\min z=\sum_{j=1}^{n} c_j x_j$$
$$\text{s.t.} \quad \sum_{j=1}^{n} P_j x_j = b$$
$$x_j\geq 0, \ j=1,2,\ldots,n$$

$$P_j=\begin{bmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{bmatrix}$$

单纯形表

			c_1	c_2	$\cdots$	c_n	
c_B	x_B	b	x_1	x_2	$\cdots$	x_n	θ
$c_{\pi(1)}$	$x_{\pi(1)}$	β_1	α_{11}	α_{12}	$\cdots$	α_{1n}	
$c_{\pi(2)}$	$x_{\pi(2)}$	β_2	α_{21}	α_{22}	$\cdots$	α_{2n}	$\theta_i = \beta_i / \alpha_{ik}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\cdots$	$\vdots$	
$c_{\pi(m)}$	$x_{\pi(m)}$	β_m	α_{m1}	α_{m2}	$\cdots$	α_{mn}	
	$-z$	$-z_0$	λ_1	λ_2	$\cdots$	λ_n	

$$-z+\lambda_1 x_1+\lambda_2 x_2+\ldots+\lambda_n x_n = -z_0$$

例1

$$\min z' = -12x_1 - 15x_2$$

$$\text{s.t.}\quad \begin{aligned} 0.25x_1+0.50x_2+x_3 &= 120 \\ 0.50x_1+0.50x_2+x_4 &= 150 \\ 0.25x_1+x_5 &= 50 \\ x_j \ge 0,\quad i=1,2,\ldots,5 \end{aligned}$$

			-12	-15	0	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	120	0.25	0.50	1	0	0	
0	x_4	150	0.50	0.50	0	1	0	
0	x_5	50	0.25	0	0	0	1	
	$-z$	0	-12	-15	0	0	0	

例1 (续)

			-12	-15	0	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	120	0.25	0.50	1	0	0	240
0	x_4	150	0.50	0.50	0	1	0	300
0	x_5	50	0.25	0	0	0	1	
	$-z$	0	-12	-15	0	0	0	

1. 若所有的检验数非负，则当前的基本可行解是最优解
2. 若存在检验数λ_k<0且所有α_{ik}≤0(1≤i≤m), 则无最优解
3. 否则，取某个λ_k<0且不是所有α_{ik}≤0
4. 对每一个$\alpha_{ik} > 0$，计算相应θ；当$\alpha_{ik} \le 0$，θ为空
5. 设θ_l最小，则在α_{lk}上加圈，取x_k为换入变量，$x_{\pi(l)}$为换出变量
6. 第ℓ行（不含$c_{\pi(l)}$和$x_{\pi(l)}$，以下相同）除以α_{lk}，第i行减第ℓ行的α_{ik}倍，把α_{ik}消成0（$1 \le i \le m$且$i \neq$l）
7. 最下面的一行减第ℓ行的λ_k倍，把λ_k消成0
8. 最后把x_B列的$x_{\pi(l)}$换成x_k,把c_B列的$c_{\pi(l)}$换成c_k

例1 (续)

			-12	-15	0	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	120	0.25	0.50	1	0	0	240
0	x_4	150	0.50	0.50	0	1	0	300
0	x_5	50	0.25	0	0	0	1	
	$-z$	0	-12	-15	0	0	0	
-15	x_2	240	0.50	1	2	0	0	480
0	x_4	30	0.25	0	-1	1	0	120
0	x_5	50	0.25	0	0	0	1	200
	$-z$	3600	-4.5	0	30	0	0	
-15	x_2	180	0	1	4	-2	0	
-12	x_1	120	1	0	-4	4	0	
0	x_5	20	0	0	1	- 1	1	
	$-z$	4140	0	0	12	18	0	

例9

用单纯形法解下述线性规划

$$
\begin{aligned}
\min z &= x_1 - 2x_2 \\
\text{s.t.}\quad & x_1 - x_2 \le 1 \\
& -2x_1 + x_2 \le 4 \\
& x_1 \ge 0,\ x_2 \ge 0
\end{aligned}
$$

解 引入2个松弛变量 x_3, x_4, 得到标准形

$$
\begin{aligned}
\min z &= x_1 - 2x_2 \\
\text{s.t.}\quad & x_1 - x_2 + x_3 \quad = 1 \\
& -2x_1 + x_2 \quad + x_4 = 4 \\
& x_j \ge 0,\ j = 1, 2, 3, 4
\end{aligned}
$$

例9 (续)

			1	-2	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	θ
0	x_3	1	1	-1	1	0	
0	x_4	4	-2	1	0	1	4
	$-z$	0	1	-2	0	0	
0	x_3	5	-1	0	1	1	
-2	x_2	4	-2	1	0	1	
	$-z$	8	-3	0	0	2	

目标函数值没有下界, 无最优解

人工变量和两阶段法

现考虑剩余的两种情况:

(1) $\sum_{j=1}^{n} a_{ij} x_j \geq b_i$

(2) $\sum_{j=1}^{n} a_{ij} x_j = b_i$

其中$b_i \geq 0$. 对于(1), 引入剩余变量转化成(2).

对(2)引入 人工变量 $y_j \geq 0$,

$$\sum_{j=1}^{n} a_{ij} x_j + y_j = b_i$$

取所有松弛变量和人工变量作为基变量, 得到初始可行基.

例10

$$\min z = -3x_1 + x_2 + x_3$$

$$\text{s.t.}\quad \begin{cases} x_1 - 2x_2 + x_3 \le 11 \\ -4x_1 + x_2 + 2x_3 \ge 3 \\ -2x_1 + x_3 = 1 \\ x_j \ge 0,\ j = 1,2,3 \end{cases}$$

引入松弛变量 x_4, 剩余变量 x_5, 标准形为

$$\min z = -3x_1 + x_2 + x_3$$

$$\text{s.t.}\quad \begin{cases} x_1 - 2x_2 + x_3 + x_4 = 11 \\ -4x_1 + x_2 + 2x_3 - x_5 = 3 \\ -2x_1 + x_3 = 1 \\ x_j \ge 0,\ 1 \le j \le 5 \end{cases}$$

例10 (续)

再引入人工变量 x_6, x_7

$$\min z = -3x_1 + x_2 + x_3$$

$$\text{s.t.}\quad \begin{aligned} x_1 - 2x_2 + x_3 + x_4 \qquad\qquad\qquad &= 11 \\ -4x_1 + x_2 + 2x_3 \qquad - x_5 + x_6 \qquad &= 3 \\ -2x_1 \qquad + x_3 \qquad\qquad\qquad + x_7 &= 1 \\ x_j \ge 0,\quad 1 \le j \le 7 \end{aligned}$$

取基变量 x_4, x_6, x_7, $x^{(0)} = (0,0,0,11,0,3,1)^T$

问题：$x^{(0)}$不对应标准形的可行解.

只有当所有人工变量等于0 时, 才能舍去人工变量得到标准形的可行解.

两阶段法

设问题

$$\min\ z = \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.}\quad \sum_{j=1}^{n} a_{ij} x_j = b_i\ , 1 \le i \le m$$

$$x_j \ge 0,\qquad 1 \le j \le n$$

其中 $b_i \ge 0,\quad 1 \le i \le m$

引入人工变量 $y_1, y_2, \dots, y_m$

辅助问题

$$\min\ w = \sum_{i=1}^{m} y_i$$

$$\text{s.t.}\quad \sum_{j=1}^{n} a_{ij} x_j + y_i = b_i\ ,\ 1 \le i \le m$$

$$x_j \ge 0,\qquad 1 \le j \le n$$

$$y_i \ge 0,\qquad 1 \le i \le m$$

两阶段法（续）

由于$w \geq 0$, 辅助问题必有最优解. 设最优值为 w^*. 有3种可能：

(1) $w^* > 0$. 原问题无可行解.

假如不然, 设 $(x_1, \ldots, x_n)^T$ 是原问题的可行解, 则 $(x_1, \ldots, x_n, 0, \ldots, 0)^T$是辅助问题的可行解, 对应的 $w = 0$. 与 $w^* > 0$ 矛盾.

(2) 在最优解中所有的人工变量都是非基变量. 此时, 人工变量都等于0, $w^* = 0$, 删去人工变量得到是原问题的基本可行解.

(3) $w^* = 0$, 但基变量中含有人工变量. 这种情况可以进一步转化成情况(2).

两阶段法（续）

此时, 所有人工变量都等于0. 设 y_k 是基变量,

由$x_B + B^{-1}Nx_N = B^{-1}b$，得

$$y_k + \sum_{j=1}^{n} \alpha_{ij} x_j + \sum_{\substack{t=1 \\ \text{且} \neq k}}^{m} \alpha'_{it} y_t = 0$$

且 y_k 不出现在其它约束等式中.

(a) 若所有$\alpha_{ij} = 0$ $(1 \le j \le n)$, 则原问题中 m 个约束等式不是线性无关的, 可以把这个等式删去，从而删去了y_k.

(b) 存在某个$\alpha_{il} \neq 0$ (可正可负). 用 x_l 作换入变量, y_k作换出变量, 做基变换. 由于$\beta_i = 0$, 经过基变换, β 的所有值均不改变, 从而新的基本解是可行解且 $w = 0$ 不变.

总之，可以使基变量中的人工变量少一个, 且保持 $w = 0$. 重复进行，最终总能变成情况(2).

两阶段法（续）

阶段一 引入人工变量, 写出辅助问题, 用单纯形法求解. 若为情况(1), 则原问题无可行解, 计算结束. 若为情况(2), 则进入阶段二.

阶段二 删去人工变量, 得到原问题的一个基本可行解. 以这个解为初始基本可行解, 用单纯形法解原问题.

例10 (续) 用两阶段法. 阶段一 辅助问题为

$$
\begin{aligned}
\min\ w &= x_6 + x_7 \\
\text{s.t.}\ \ x_1 - 2x_2 + x_3 + x_4 \qquad\qquad &= 11 \\
-4x_1 + x_2 + 2x_3 \qquad - x_5 + x_6 \qquad &= 3 \\
-2x_1 \qquad + x_3 \qquad\qquad\qquad + x_7 &= 1 \\
x_j \geq 0,\ \ j = 1, 2, \ldots, 7 &
\end{aligned}
$$

例10(续)

			0	0	0	0	0	1	1	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	θ
0	x_4	11	1	-2	1	1	0	0	0	11
1	x_6	3	-4	1	2	0	-1	1	0	1.5
1	x_7	1	-2	0	1	0	0	0	1	1
	$-w$	-4	6	-1	-3	0	1	0	0	
0	x_4	10	3	-2	0	1	0	0	-1	–
1	x_6	1	0	1	0	0	-1	1	-2	1
0	x_3	1	-2	0	1	0	0	0	1	–
	$-w$	-1	0	-1	0	0	1	0	3	
0	x_4	12	3	0	0	1	-2	2	-5	
0	x_2	1	0	1	0	0	-1	1	-2	
0	x_3	1	-2	0	1	0	0	0	1	
	$-w$	0	0	0	0	0	0	0	1	

例10(续)

阶段二

			-3	1	1	0	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_4	12	3	0	0	1	-2	4
1	x_2	1	0	1	0	0	-1	–
1	x_3	1	-2	0	1	0	0	–
	$-z$	-2	-1	0	0	0	1	
-3	x_1	4	1	0	0	1/3	-2/3	
1	x_2	1	0	1	0	0	-1	
1	x_3	9	0	0	1	2/3	-4/3	
	$-z$	2	0	0	0	1/3	1/3	

最优解 $x_1^*=4$, $x_2^*=1$, $x_3^*=9$, 最优值 $z^*= -2$

例11

$$\min z = 3x_1 - 2x_2$$
$$\text{s.t. } 2x_1 + x_2 \le 4$$
$$x_1 - x_2 \ge 3$$
$$x_1 \ge 0,\ x_2 \ge 0$$

标准形

$$\min z = 3x_1 - 2x_2$$
$$\begin{aligned} \text{s.t. } & 2x_1 + x_2 + x_3 \qquad = 4 \\ & x_1 - x_2 \qquad - x_4 = 3 \\ & x_j \ge 0, \quad 1 \le j \le 4 \end{aligned}$$

阶段一　辅助问题

$$\min w = x_5$$
$$\begin{aligned} \text{s.t. } & 2x_1 + x_2 + x_3 \qquad\qquad = 4 \\ & x_1 - x_2 \qquad - x_4 + x_5 = 3 \\ & x_j \ge 0, \quad 1 \le j \le 5 \end{aligned}$$

例11 (续)

			0	0	0	0	1	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	θ
0	x_3	4	2	1	1	0	0	2
1	x_5	3	1	-1	0	-1	1	3
	$-w$	-3	-1	1	0	1	0	
0	x_1	2	1	1/2	1/2	0	0	
1	x_5	1	0	-3/2	-1/2	-1	1	
	$-w$	-1	0	3/2	1/2	1	0	

$w^*=1 > 0$, 原问题没有可行解.

例12

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

$$\begin{aligned} \text{s.t. } & 3x_1 + 6x_2 + 2x_3 - x_4 = 12 \\ & 2x_1 \qquad\quad + x_3 \qquad\quad = 4 \\ & 3x_1 - 6x_2 + x_3 + x_4 = 0 \\ & x_j \geq 0,\ 1 \leq j \leq 4 \end{aligned}$$

阶段一　辅助问题

$$\min w = x_5 + x_6 + x_7$$

$$\begin{aligned} \text{s.t. } & 3x_1 + 6x_2 + 2x_3 - x_4 + x_5 = 12 \\ & 2x_1 \qquad\quad + x_3 \qquad\qquad + x_6 = 4 \\ & 3x_1 - 6x_2 + x_3 + x_4 \qquad\qquad + x_7 = 0 \\ & x_j \geq 0,\ 1 \leq j \leq 7 \end{aligned}$$

例12 (续)

			0	0	0	0	1	1	1	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	θ
1	x_5	12	3	6	2	-1	1	0	0	4
1	x_6	4	2	0	1	0	0	1	0	2
1	x_7	0	3	-6	1	1	0	0	1	0
	$-w$	-16	-8	0	-4	0	0	0	0	
1	x_5	12	0	12	1	-2	1	0	-1	1
1	x_6	4	0	4	1/3	-2/3	0	1	-2/3	1
0	x_1	0	1	-2	1/3	1/3	0	0	1/3	–
	$-w$	-16	0	-16	-4/3	8/3	0	0	8/3	
1	x_5	0	0	0	0	0	1	-3	1	
0	x_2	1	0	1	1/12	-1/6	0	1/4	-1/6	
0	x_1	2	1	0	1/2	0	0	1/2	0	
	$-w$	0	0	0	0	0	0	4	0	

例12 (续)

$w^*=0$, 人工变量 x_5 是基变量且$\alpha_{11}=\alpha_{12}=\alpha_{13}=\alpha_{14}=0$, β_1必为 0. 原规划中第1个约束等式是另两个的线性组合, 可以删去.

阶段二

			1	3	-2	0	
c_B	x_B	b	x_1	x_2	x_3	x_4	θ
3	x_2	1	0	1	1/12	-1/16	12
1	x_1	2	1	0	1/2	0	4
	$-z$	-5	0	0	-11/4	1/2	
3	x_2	2/3	-1/6	1	0	-1/6	
-2	x_3	4	2	0	1	0	
	$-z$	6	11/2	0	0	1/2	

最优解 $x_1^*=0$, $x_2^*=2/3$, $x_3^*=4$, $x_4^*=0$, 最优值 $z^*=-6$.

例13

求解线性规划：

$$\min z = x_1 + x_2 + x_3 - x_4$$

$$\text{s.t.} \quad \begin{aligned} 6x_1 + 3x_2 - 4x_3 + 3x_4 &= 12 \\ -x_2 \qquad\quad + 3x_4 &= 6 \\ -6x_1 \qquad + 4x_3 + 3x_4 &= 0 \\ x_j \geq 0,\ 1 \leq j \leq 4 & \end{aligned}$$

例13 (续)

阶段一　引入人工变量 x_5, x_6, x_7.

			0	0	0	0	1	1	1	
c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	θ
1	x_5	12	6	3	-4	3	1	0	0	4
1	x_6	6	0	-1	0	3	0	1	0	2
1	x_7	0	-6	0	4	3	0	0	1	0
	$-w$	-18	0	-2	0	-9	0	0	0	
1	x_5	12	12	3	-8	0	1	0	-1	1
1	x_6	6	6	-1	-4	0	0	1	-1	1
0	x_4	0	-2	0	4/3	1	0	0	1/3	–
	$-w$	-18	-18	-2	12	0	0	0	3	

例13 (续)

c_B	x_B	b	x_1	x_2	x_3	x_4	x_5	x_6	x_7	θ
0	x_1	1	1	1/4	-2/3	0	1/12	0	-1/12	
1	x_6	0	0	-5/2	0	0	-1/2	1	-1/2	
0	x_4	2	0	1/2	0	1	1/6	0	1/6	
	$-w$	0	0	5/2	0	0	3/2	0	3/2	
0	x_1	1	1	0	-2/3	0	1/30	1/10	-2/15	
0	x_2	0	0	1	0	0	1/5	-2/5	1/5	
0	x_4	2	0	0	0	1	1/15	1/5	1/15	
	$-w$	0	0	0	0	0	1	1	1	

在倒数第2个表中, $w = 0$, 但 x_6 是基变量且$\alpha_{22} \neq 0$, 取 x_6 为换出变量、 x_2为换入变量作基变量.

例13 (续)

阶段二

			1	1	1	-1	
c_B	x_B	b	x_1	x_2	x_3	x_4	θ
1	x_1	1	1	0	-2/3	0	
1	x_2	0	0	1	0	0	
-1	x_4	2	0	0	0	1	
	$-z$	1	0	0	2/3	0	

最优解 $x_1^*= 1, x_2^* = 0, x_3^* = 0, x_4^* = 2$, 最优值 $z^* = -1$.

单纯形法的有限终止

定义 如果基本可行解中基变量的值都大于0, 则称这个基本可行解是非退化的, 否则称作退化的.

如果线性规划的所有基本可行解都是非退化的, 则称这个线性规划是非退化的.

如果线性规划有可行解并且是非退化的, 则在计算的每一步

$$z_0' = z_0 + \lambda_k \beta_l / \alpha_{lk} < z_0,$$

可行基不会重复出现, 因此单纯形法在有限步内终止.

如果不是非退化的, 当$\beta_l = 0$且取 $x_{\pi(l)}$为换出变量时, 基变换不改变目标函数值. 这就可能使计算出现循环, 计算永不终止.

单纯形法出现循环的例子

一个例子

$$\min z = -0.75x_1 + 20x_2 - 0.5x_3 + 6x_4$$

$$\begin{aligned}
\text{s.t. } 0.25x_1 - 8x_2 - x_3 + 9x_4 + x_5 \quad &= 0 \\
0.5x_1 - 12x_2 - 0.5x_3 + 3x_4 \quad + x_6 \quad &= 0 \\
x_3 \quad + x_7 &= 1
\end{aligned}$$

$$x_j \geq 0, \quad 1 \leq j \leq 7$$

取 x_5, x_6, x_7 作为初始基变量, 并规定: 当有多个 $\lambda_j < 0$ 时, 设 $|\lambda_k| = \max \{ |\lambda_j| : \lambda_j<0 \}$, 取 x_k 作为换入变量; 当有多个 θ_i 同时取到最小值时, 取对应的下标最小的基变量作为换出变量. 计算经过 6 次基变换回到初始可行基, 从而计算出现循环, 永不终止.

避免循环的方法

Bland规则

规则1. 当有多个 $\lambda_j < 0$ 时, 取对应的非基变量中下标最小的作为换入变量.

规则2. 当有多个 $\theta_i = \beta_i / \alpha_{ik}\,(\alpha_{ik} > 0)$ 同时取到最小值时, 取对应的基变量中下标最小的作为换出变量.

对偶性

对偶线性规划

再看例1 公司甲用 3 种原料混合成 2 种清洁剂.
这 2 种清洁剂应各配制多少才能使总价值最大?

	原料1	原料2	原料3	售价(万元/吨)
清洁剂A	0.25	0.50	0.25	12
清洁剂B	0.50	0.50		15
存量 (吨)	120	150	50	

公司乙急需这 3 种原料, 打算向公司甲购买, 应出什么价钱?

实例

设清洁剂 A 和 B 分别配制 x 和 y

$$\max z = 12x + 15y$$

$$\begin{aligned} \text{s.t.}\quad & 0.25x + 0.50y \le 120 \\ & 0.50x + 0.50y \le 150 \\ & 0.25x \le 50 \\ & x \ge 0,\ y \ge 0 \end{aligned}$$

设公司乙出价原料每吨分别为y_1, y_2, y_3万元. 希望总价尽可能小, 但又不 低于公司甲用这些原料生产清洁剂所产生的价值

$$\min w = 120y_1 + 150y_2 + 50y_3$$

$$\begin{aligned} \text{s.t.}\quad & 0.25y_1 + 0.50y_2 + 0.25y_3 \ge 12 \\ & 0.50y_1 + 0.50y_2 \ge 15 \\ & y_1 \ge 0,\ y_2 \ge 0,\ y_3 \ge 0 \end{aligned}$$

对偶线性规划

定义 原始线性规划 (P)

$$\max\ c^T x$$
$$\text{s.t. } Ax \le b$$
$$x \ge 0$$

对偶线性规划(D)

$$\min\ b^T y$$
$$\text{s.t. } A^t y \ge c$$
$$y \ge 0$$

定理4 对偶的对偶是原始线性规划.

证 (D) 可改写成 (D′)

$$\max\ -b^T y$$
$$\text{s.t. } -A^T y \le -c$$
$$y \ge 0$$

(D′) 的对偶为

$$\min\ -c^T x$$
$$\text{s.t. } (-A^T)^T x \ge -b$$
$$x \ge 0$$

例14

写出下述线性规划的对偶

$$\max\ 2x_1 - x_2 + 3x_3$$
$$\begin{aligned} \text{s.t.}\quad & x_1 + 3x_2 - 2x_3 \le 5 \\ & -x_1 - 2x_2 + x_3 = 8 \\ & x_1 \ge 0,\ x_2 \ge 0,\ x_3\text{任意} \end{aligned}$$

对偶规划为

$$\min\ 5y_1 + 8y_2' - 8y_2''$$
$$\begin{aligned} \text{s.t.}\quad & y_1 - y_2' + y_2'' \ge 2 \\ & 3y_1 - 2y_2' + 2y_2'' \ge -1 \\ & -2y_1 + y_2' - y_2'' \ge 3 \\ & 2y_1 - y_2' + y_2'' \ge -3 \\ & y_1 \ge 0, y_2' \ge 0, y_2'' \ge 0 \end{aligned}$$

令 $x_3 = x_3' - x_3''$,

$A=B$ 等价于 $A\le B$ 和 $-A \le -B$,

$$\max\ 2x_1 - x_2 + 3x_3' - 3x_3''$$
$$\begin{aligned} \text{s.t.}\quad & x_1 + 3x_2 - 2x_3' + 2x_3'' \le 5 \\ & -x_1 - 2x_2 + x_3' - x_3'' \le 8 \\ & x_1 + 2x_2 - x_3' + x_3'' \le -8 \\ & x_1 \ge 0, x_2 \ge 0, x_3' \ge 0, x_3'' \ge 0 \end{aligned}$$

令 $y_2 = y_2' - y_2''$, 合并后2个不等式

$$\min\ 5y_1 + 8y_2$$
$$\begin{aligned} \text{s.t.}\quad & y_1 - y_2 \ge 2 \\ & 3y_1 - 2y_2 \ge -1 \\ & -2y_1 + y_2 = 3 \\ & y_1 \ge 0,\ y_2\text{任意} \end{aligned}$$

对偶规划的一般形式

原始规划	对偶规划
$\max \sum_{j=1}^{n} c_j x_j$	$\min \sum_{i=1}^{m} b_i y_i$
$\sum_{j=1}^{n} a_{ij} x_j \leqslant b_i$，$1 \leqslant i \leqslant s$	$y_i \geqslant 0$，$1 \leqslant i \leqslant s$
$\sum_{j=1}^{n} a_{ij} x_j = b_i$，$s+1 \leqslant i \leqslant m$	y_i 任意，$s+1 \leqslant i \leqslant m$
$x_j \geqslant 0$，$1 \leqslant j \leqslant t$	$\sum_{i=1}^{m} a_{ij} y_i \geqslant c_j$，$1 \leqslant j \leqslant t$
x_j 任意，$t+1 \leqslant j \leqslant n$	$\sum_{i=1}^{m} a_{ij} y_i = c_j$，$t+1 \leqslant j \leqslant n$

性质

定理5 设 x 是原始规划 (P) 的可行解, y是对偶规划 (D) 的可行解, 则恒有

$$c^T x \le b^T y$$

证

$$c^T x \le (A^T y)^T x = y^T (Ax) \le y^T b = b^T y$$

定理6 如果 x 和 y 分别是原始规划 (P) 和对偶规划 (D) 的可行解, 且$c^T x = b^T y$, 则 x 和 y 分别是它们的最优解.

定理7 如果原始规划 (P) 有最优解, 则对偶规划 (D) 也有最优解, 且它们的最优值相等. 反之亦然.

定理7证明

证 引入松弛变量 u, 将 (P) 写成

$$\max c^T x \qquad A\text{是 } m\times n \text{ 矩阵}$$

$$\text{s.t. } Ax+Eu=b \qquad E\text{是 } m\times m \text{ 单位矩阵}$$

$$x\ge 0,\ u\ge 0 \qquad u \text{ 是 } m \text{ 维向量}$$

设最优解基为B, 基变量 $x_B = B^{-1}b$, 检验数 $\lambda \le 0$ (最大化问题). λ分成两部分, 对应 x 的 λ_1和对应 u 的 λ_2. u 在目标函数中的系数都为 0, 有

$$\lambda_1^{\mathrm{T}} = c^{\mathrm{T}} - c_B^{\mathrm{T}}B^{-1}A \le 0$$

$$\lambda_2^{\mathrm{T}} = -c_B^{\mathrm{T}}B^{-1}E = -c_B^{\mathrm{T}}B^{-1} \le 0$$

令 $y^T = c_B^{\mathrm{T}}B^{-1}$, 有 $y \ge 0$, $c^{\mathrm{T}} - y^{\mathrm{T}}A \le 0$，即 $A^{\mathrm{T}}y \ge c$

从而 y 是 (D)的可行解. 又

$$w = b^T y = y^T b = c_B^{\mathrm{T}}B^{-1}b = c_B^{\mathrm{T}}x_B = z$$

得证 y 是 (D) 的最优解.

原始规划和对偶规划解的可能情况

(1) 都有最优解, 且最优值相等.

(2) 一个有可行解且目标函数值无界, 而另一个无可行解.

(3) 都没有可行解.

对偶规划 \ 原始规划	有最优解	有可行解且无界	无可行解
有最优解	(1)	×	×
有可行解且无界	×	×	(2)
无可行解	×	(2)	(3)

互补松弛性

定理8 设 x 和 y 分别是原始规划 (P) 和对偶规划 (D) 的可行解, 则 x 和 y 分别是它们的最优解当且仅当

$$(b_i - \sum_{j=1}^{n} a_{ij} x_j) y_i = 0, \qquad 1 \le i \le m \qquad (*)$$

$$x_j (\sum_{i=1}^{m} a_{ij} y_i - c_j) = 0, \qquad 1 \le j \le n \qquad (**)$$

证 $u_i = (b_i - \sum_{j=1}^{n} a_{ij} x_j) y_i \ge 0, \quad v_j = x_j (\sum_{i=1}^{m} a_{ij} y_i - c_j) \ge 0$

(*)和(**)成立 $\Leftrightarrow \sum_{i=1}^{m} u_i + \sum_{j=1}^{n} u_j = 0$

定理8证明(续)

$$\sum_{i=1}^{m} u_i + \sum_{j=1}^{n} u_j = \sum_{i=1}^{m}(b_i - \sum_{j=1}^{n} a_{ij} x_j) y_i + \sum_{j=1}^{n} x_j (\sum_{i=1}^{m} a_{ij} y_i - c_j)$$

$$= \sum_{i=1}^{m} b_i y_i - \sum_{j=1}^{n} c_j x_j$$

得证

(*)和(**)成立 $\Leftrightarrow \sum_{j=1}^{n} c_j x_j = \sum_{i=1}^{m} b_i y_i$

$\Leftrightarrow$ x是 (P) 的最优解, y是 (D) 的最优解.

对偶单纯形法

原始规划 (P)

$$\min z = c^T x$$
$$\text{s.t.}\ \ Ax = b$$
$$x \geq 0$$

对偶规划 (D)

$$\max w = b^T y$$
$$\text{s.t.}\ \ A^T y \leq c$$
y 任意

设 B是 (P) 的一个可行基, 对应的可行解 $x_B = B^{-1}b,\ x_N = 0,$
$\lambda^{\mathrm{T}}=c^{\mathrm{T}}-c_B^{\mathrm{T}}B^{-1}A,\ z_0=c_B^{\mathrm{T}}B^{-1}b$. 令 $y^{\mathrm{T}}=c_B^{\mathrm{T}}B^{-1}$, 恒有

$$w_0 = b^{\mathrm{T}}y = y^{\mathrm{T}}b = c_B^{\mathrm{T}}B^{-1}b = z_0$$

只要 y是 (D) 的可行解, 则 x 和 y 分别是 (P) 和 (D)的最优解.

由 $\lambda^{\mathrm{T}} = c^{\mathrm{T}}-c_B^{\mathrm{T}}B^{-1}A = c^{\mathrm{T}}-y^T A$, 有

$$y\text{是 (D) 的可行解} \Leftrightarrow \lambda \geq 0$$

对偶单纯形法

定义 设 B是一个基, 如果 $\lambda \geq 0$, 则称 B 是**正则的**.

如果 B 是正则的, 那么 y 是 (D)的可行解, 从而只要 x 是 (P) 的可行解, 亦即 $x_B = B^{-1}b \geq 0$, 则 x 和 y 分别是 (P) 和 (D) 的最优解.

单纯形法 保持 x 是 (P) 的可行解 (保持B是可行基), 即保持 $B^{-1}b \geq 0$, 通过基变换使 y 逐步成为 (D) 的可行解 (B变成正则基), 即逐步使 $\lambda \geq 0$.

对偶单纯形法 保持 y 是 (D) 的可行解 (保持B是正则基), 即保持 $\lambda \geq 0$, 通过基变换使 x 逐步成为 (P) 的可行解 (B 变成可行基), 即逐步使 $B^{-1}b \geq 0$.

对偶单纯形法

设$\lambda \geq 0$, $\beta_l < 0$, 若所有$\alpha_{lj} \geq 0$ $(1 \leq j \leq n)$, 则不存在 $x \geq 0$使得 $\sum_{j=1}^{n} \alpha_{lj} x_j = \beta_l$, (P)无可行解. 若存在$\alpha_{lk} < 0$, 则以 $x_{\pi(l)}$ 为换出变量、 x_k为换入变量做基变换, 必须保证

$$\lambda_j' = \lambda_j - \lambda_k \alpha_{lj} / \alpha_{lk} \geq 0, \qquad 1 \leq j \leq n$$

注意到 $\lambda_j \geq 0$, $\lambda_k \geq 0$, $\alpha_{lk} < 0$, 当 $\alpha_{lj} \geq 0$时, 不等式自然成立.
于是, 只要当 $\alpha_{lj} < 0$ 时,

$$\lambda_j / \alpha_{lj} \leq \lambda_k / \alpha_{lk}$$

故应取 k 使得

$$|\lambda_k / \alpha_{lk}| = \min\{ | \lambda_j / \alpha_{lj} | : \alpha_{lj} < 0\}$$

对偶单纯形法

算法2 对偶单纯形法

1. 取正则基 B.

2. 如果 $\beta \geq 0$, 则 x 是最优解, 计算结束.

3. 取 $\beta_l < 0$. 若所有 $\alpha_{lj} \geq 0$ $(1 \leq j \leq n)$, 则无可行解, 计算结束.

4. 取 k 使得

$$|\lambda_k / \alpha_{lk}| = \min\{\, |\lambda_j / \alpha_{lj}| : \alpha_{lj} < 0 \}$$

5. 以 $x_{\pi(l)}$ 为换出变量、 x_k 为换入变量做基变换.

6. 转2.

整数线性规划的分支限界算法

整数线性规划 在线性规划上对变量增添整数的要求

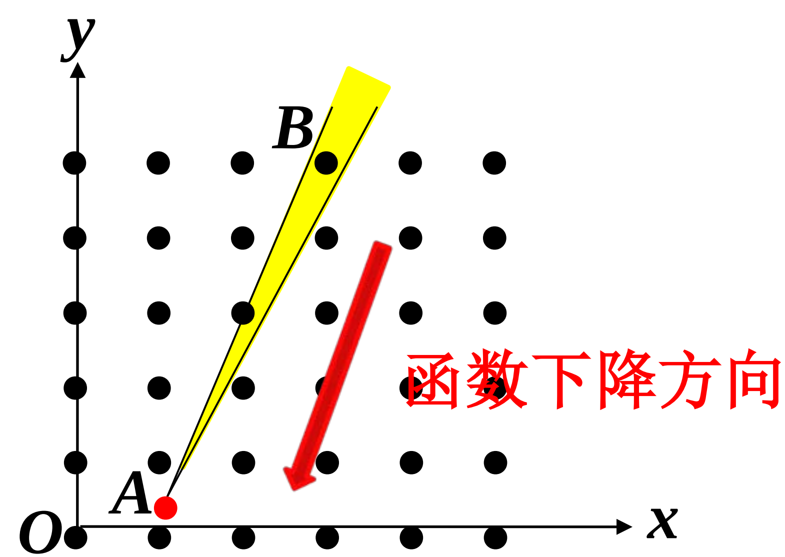
纯整数线性规划 (全整数线性规划) 要求所有变量是整数

混合整数线性规划 只要求部分变量是整数

0-1型整数线性规划 要求所有变量是 0 或 1

松弛规划 (简称松弛) 删去整数要求后得到的线性规划

松弛规划的最优值是原整数规划的最优值的界限(最小化的下界,最大化的上界), 但通常不是原整数规划的最优解



分支限界法

记整数线性规划为 **ILP**, 其松弛为 **LP**.

如果 **LP** 的最优解 α 满足整数要求, 则 α 是 **ILP**的最优解.

否则, 设α_1不满足整数要求, 在 **LP** 上分别添加 $x_1 \le \lfloor \alpha_1 \rfloor$ 和 $x_1 \ge \lfloor \alpha_1 \rfloor + 1$, 记作 $\mathbf{LP}_1$ 和 $\mathbf{LP}_2$. 如果 $\mathbf{LP}_1$ 或 $\mathbf{LP}_2$ 的最优解符合整数要求, 那么这个解也是 **ILP** 的可行解, 得到 **ILP** 的最优值的一个界限 (最小化的上界, 最大化的下界), 该子问题的计算结束.

如果子问题的最优解不满足整数要求, 则继续分支计算.

如果子问题的最优值超过界限 (最小化大于界限, 最大化小于界限), 则往下计算不可能得到 **ILP** 的最优解, 计算结束.

当没有待计算的子问题时, 所有可行解中最好的是 **ILP** 的最优解.

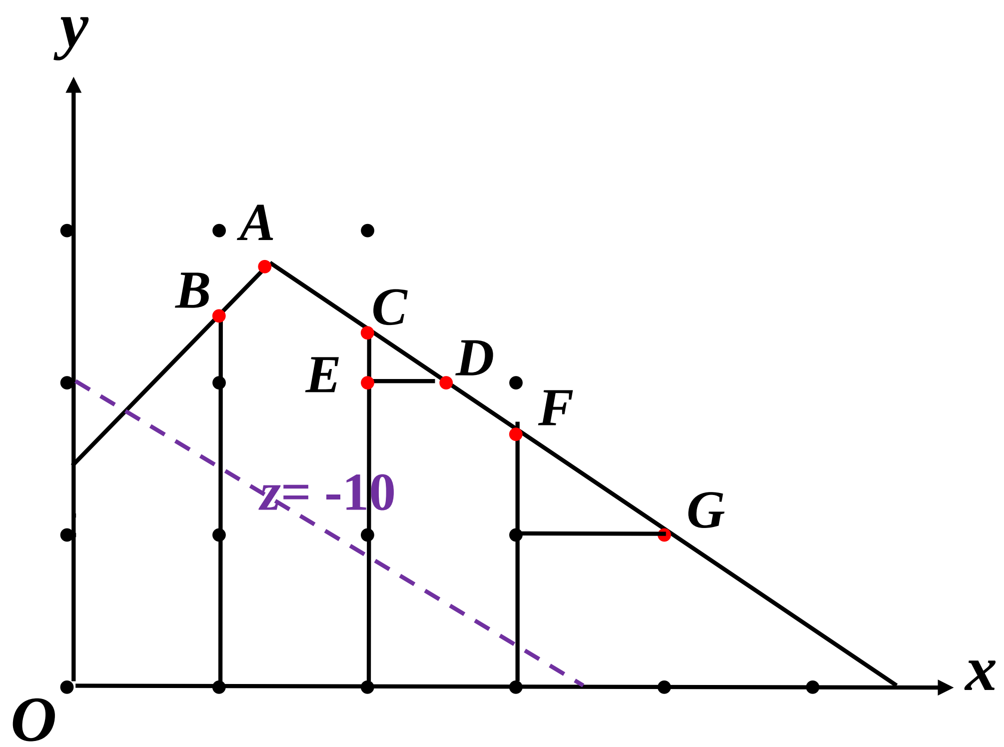
例15

$$\min z = -3x - 5y$$

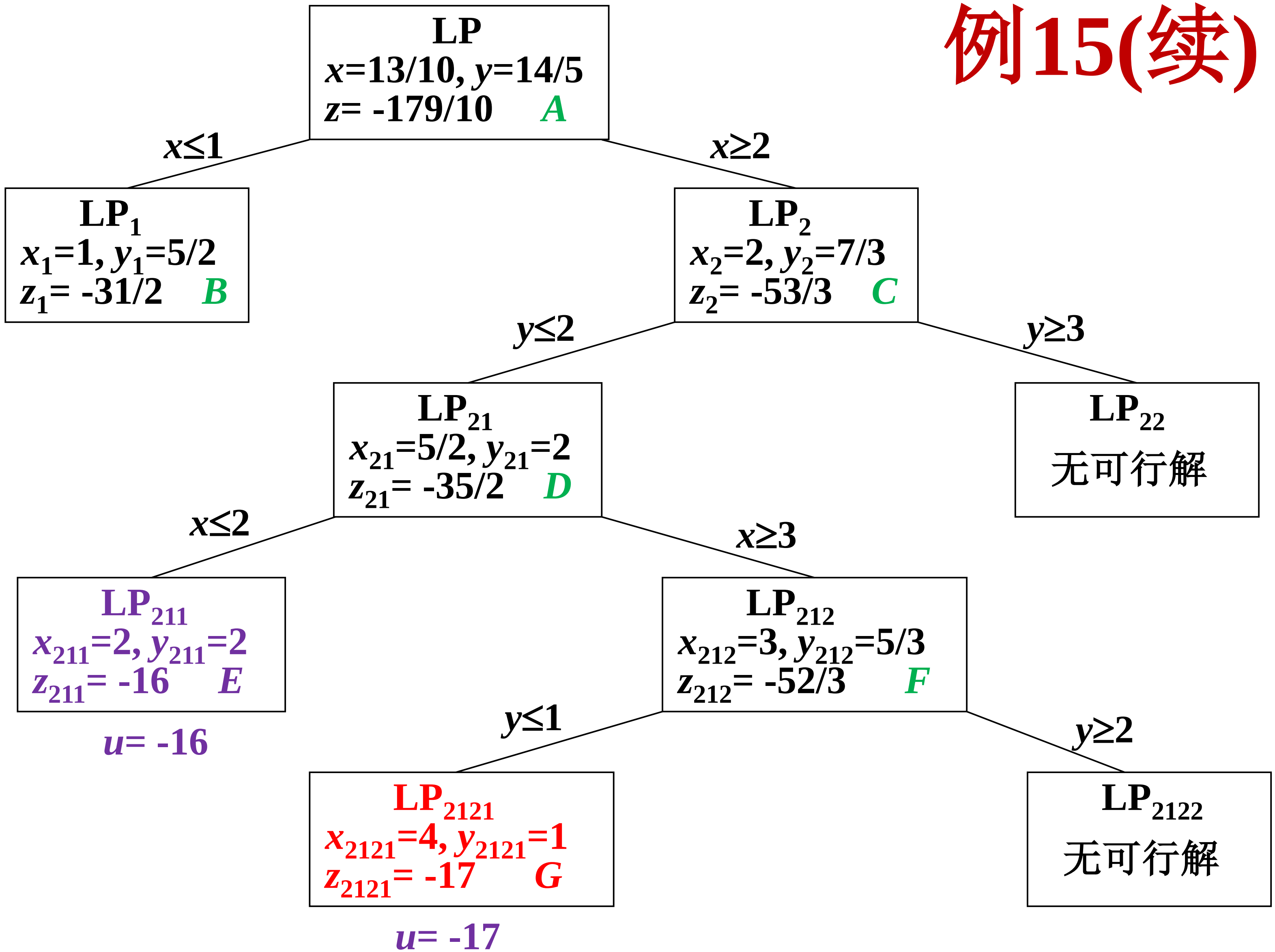
$$\text{s.t.}\quad -x + y \le 3/2$$

$$2x + 3y \le 11$$

$$x, y \ge 0, \text{整数}$$



例15(续)



LP
x=13/10, y=14/5
z= -179/10 A
x≤1
x≥2
LP1
x1=1, y1=5/2
z1= -31/2 B
LP2
x2=2, y2=7/3
z2= -53/3 C
y≤2
y≥3
LP21
x21=5/2, y21=2
z21= -35/2 D
LP22
无可行解
x≤2
x≥3
LP211
x211=2, y211=2
z211= -16 E
u= -16
LP212
x212=3, y212=5/3
z212= -52/3 F
y≤1
y≥2
LP2121
x2121=4, y2121=1
z2121= -17 G
u= -17
LP2122
无可行解

LP 的计算

			-3	-5	0	0	
c_B	x_B	b	x	y	u_1	u_2	θ
0	u_1	1.5	-1	1	1	0	1.5
0	u_2	11	2	3	0	1	11/3
	$-z$	0	-3	-5	0	0	
-5	y	1.5	-1	1	1	0	-
0	u_2	6.5	5	0	-3	1	1.3
	$-z$	7.5	-8	0	5	0	
-5	y	2.8	0	1	0.4	0.2	
-3	x	1.3	1	0	-0.6	0.2	
	$-z$	17.9	0	0	0.2	1.6	

LP_1 的计算

			-3	-5	0	0	0
c_B	x_B	b	x	y	u_1	u_2	v_1
-5	y	2.8	0	1	0.4	0.2	0
-3	x	1.3	1	0	-0.6	0.2	0
		1	1	0	0	0	1
	$-z$	17.9	0	0	0.2	1.6	0
-5	y	2.8	0	1	0.4	0.2	0
-3	x	1.3	1	0	-0.6	0.2	0
0	v_1	-0.3	0	0	0.6	-0.2	1
	$-z$	17.9	0	0	0.2	1.6	0
	θ		-	-	-	-8	-
-5	y	2.5	0	1	1	0	1
-3	x	1	1	0	0	0	1
0	u_2	1.5	0	0	-3	1	-5
	$-z$	15.5	0	0	5	0	8

对$x \leq 1$引入松弛变量 v_1,

$$x + v_1 = 1$$

第3行减第2行，消去 x (x是基变量), 取 v_1作为基变量.

用对偶单纯形法求解.

LP_2 的计算

			-3	-5	0	0	0
c_B	x_B	b	x	y	u_1	u_2	v_2
-5	y	2.8	0	1	0.4	0.2	0
-3	x	1.3	1	0	-0.6	0.2	0
		-2	-1	0	0	0	1
	$-z$	17.9	0	0	0.2	1.6	0
-5	y	2.8	0	1	0.4	0.2	0
-3	x	1.3	1	0	-0.6	0.2	0
0	v_2	-0.7	0	0	-0.6	0.2	1
	$-z$	17.9	0	0	0.2	1.6	0
	θ		-	-	-1/3	-	-
-5	y	7/3	0	1	0	1/3	2/3
-3	x	2	1	0	0	0	-1
0	u_1	7/6	0	0	1	-1/3	-1/6
	$-z$	53/3	0	0	0	5/3	1/3

对$x \geq 2$引入剩余变量v_2,

$$-x + v_2 = -2$$

第 3行加第 2行, 消去 x, 取v_2作为基变量.

用对偶单纯形法求解.

LP_{22}的计算

			-3	-5	0	0	0	0
c_B	x_B	b	x	y	u_1	u_2	v_2	v_{22}
-5	y	7/3	0	1	0	1/3	2/3	0
-3	x	2	1	0	0	0	-1	0
0	u_1	7/6	0	0	0	-1/3	-1/6	0
		-3	0	-1	0	0	0	1
	-z	53/3	0	0	0	5/3	1/3	0
-5	y	7/3	0	1	0	1/3	2/3	0
-3	x	2	1	0	0	0	-1	0
0	u_1	7/6	0	0	0	-1/3	-1/6	0
0	v_{22}	-2/3	0	0	0	1/3	2/3	1
	-z	53/3	0	0	0	5/3	1/3	0

无可行解

应用：最小顶点覆盖

- **顶点覆盖问题**

  给定图$G=(V,E)$，G的顶点覆盖是顶点子集$S\subseteq V$，使得每条边至少有一个端点属于S. 求G的最小的顶点覆盖.

- **转化为线性规划问题**

  令$V=\{1, 2, \ldots, n\}$，$\forall e\in E$，存在 $i, j\in V$，使得 $e=(i, j)$

  $\forall i\in V$，定义变量 $x_i = 0,1$，且 $x_i = 1 \Leftrightarrow i\in S$

  $\forall e = (i, j)\in E,\ x_i + x_j \geq 1$

$$\min \sum_{i\in V} x_i$$

$$\begin{aligned} \text{s.t.}\ \ & x_i + x_j \geq 1 \qquad & (i, j)\in E \\ & x_i = 0,1 \qquad & i\in V \end{aligned}$$

设计算法

- 顶点覆盖是整数规划问题，属于NP难问题.
- 近似算法的设计思想：

  1. 放松 $x_i = 0,1$ 的约束条件，令 x_i 为$[0,1]$区间的任意实数，转化为线性规划问题.

  2. 用线性规划算法找到一组 $x_i \in [0,1]$，$i=1,2,\ldots,n$，使得其和达到最小.

  3. 令$S=\{ i \mid x_i \geq 1/2\}$.

- 算法分析

  可以证明上述 S 是 G 的顶点覆盖，且 $|S| \leq 2|S^*|$，其中 S^* 为最优解.

应用：负载均衡问题

- 负载均衡问题

  给定作业集合$J=\{1, 2, \ldots, n\}$，作业 j 加工时间为 t_j，$j = 1, 2, \ldots, n$. 机器集合 $M =\{1, 2, \ldots, m\}$，对每个作业分配一台机器，作业 j 可分配的机器集合为M_j. J_i是分配到机器 i上的作业集合. 机器 i 的负载是 L_i。设分配方案的负载为L，其中

$$L = \max_{i\in\{1,2,\ldots,m\}} L_i \, , \qquad L_i = \sum_{j\in J_i} t_j$$

问题：求分配方案 使得 L达到最小。

转变为线性规划

$$\min\ L$$

$$\sum_i x_{ij} = t_j \quad \forall j \in J$$ 任务 j 在分配机器的负载之和等于加工时间

$$\sum_j x_{ij} \le L \quad \forall i \in M$$ 任何机器的负载总量不超过 L

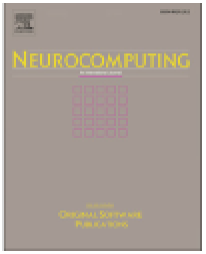
$$x_{ij} \ge 0 \qquad \forall j \in J, i \in M_j$$

- 如果上述线性规划有值不超过 L 的解，那么最优负载的值至少是 L.
- 线性规划的最优解有可能把一个作业分配到多台机器上，即负载是分数．需要调整这个解，以满足原问题的需求：每个作业只能分配到一台机器上.

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom



A deep learning approach for solving linear programming problems

Dawen Wu *, Abdel Lisser

Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systémes, 91190 Gif-sur-Yvette, France

A R T I C L E   I N F O

Article history:
Received 24 May 2022
Revised 4 October 2022
Accepted 15 November 2022
Available online 19 November 2022

Keywords:
Linear programming
ODE systems
Neural networks
Deep learning

A B S T R A C T

Finding the optimal solution to a linear programming (LP) problem is a long-standing computational problem in Operations Research. This paper proposes a deep learning approach in the form of feed-forward neural networks to solve the LP problem. The latter is first modeled by an ordinary differential equations (ODE) system, the state solution of which globally converges to the optimal solution of the LP problem. A neural network model is constructed as an approximate state solution to the ODE system, such that the neural network model contains the prediction of the LP problem. Furthermore, we extend the capability of the neural network by taking the parameter of LP problems as an input variable so that one neural network can solve multiple LP instances in a one-shot manner. Finally, we validate the proposed method through a collection of specific LP examples and show concretely how the proposed method solves the example.

$$\mathbf{y}(t) = (\mathbf{x}^T(t), \mathbf{u}^T(t))^T$$

An LP problem

$$\min_{\mathbf{x}} \mathbf{D}^T\mathbf{x}$$

s.t.

$$\mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{0}$$

The KKT conditions

$$\mathbf{D} + \mathbf{A}^T\mathbf{u} = 0$$

$$\mathbf{u}^T(\mathbf{A}\mathbf{x} - \mathbf{b}) = 0$$

$$\mathbf{A}\mathbf{x} - \mathbf{b} \leq \mathbf{0}$$

$$\mathbf{u}^T \geq \mathbf{0}$$

The ODE system

$$\frac{d\mathbf{y}}{dt} = \Phi(\mathbf{y})$$

$$\Phi(\mathbf{y}) = \begin{bmatrix} -\left(\mathbf{D} + \mathbf{A}^T(\mathbf{u} + \mathbf{A}\mathbf{x} - \mathbf{b})^+\right) \\ (\mathbf{u} + \mathbf{A}\mathbf{x} - \mathbf{b})^+ - \mathbf{u} \end{bmatrix}$$

The initial condition

$$\mathbf{y}(0) = \mathbf{0}$$

The time range

$$t \in [0, T]$$

Approximate state solution

$$\hat{\mathbf{y}}(t; \mathbf{w}) = (1 - e^{-t})\mathbf{NN}(t; \mathbf{w})$$

Approximate solution to the KKT conditions

$$\hat{\mathbf{y}}(T; \mathbf{w}) = (1 - e^{-T})\mathbf{NN}(T; \mathbf{w})$$

Linear programming → **ODE system** → **Neural network**

weights $\mathbf{w}$. The multiplier $(1 - e^{-t})$ is used to guarantee the initial condition $\hat{\mathbf{y}}(0; \mathbf{w}) = 0$ must be satisfied regardless of weights $\mathbf{w}$.

According to Theorem 1, the optimal solution to the LP problem is equivalent to the equilibrium points of the corresponding ODE system. According to Theorem 2, for any given initial point, the state solution of the ODE system will always converge to an optimal solution of the LP problem as the time t goes to infinity, i.e., $\lim_{t\to\infty}\mathbf{y}(t) = \mathbf{y}^*$. We refer the reader to [29,31,43] for the proof of the theorems and other details.

Algorithm 1: Solving an LP problem by one neural network

Require: A time range $[0, T]$
Input : An LP problem
Output : Solution to the LP problem

Function Main:
  Derive the ODE system Φ of the LP problem.
  Construct a neural network model $\hat{\mathbf{y}}(t; \mathbf{w}) = (1 - e^{-t})\mathbf{NN}(t; \mathbf{w})$.
  while True **do**
    $\mathbb{T} \sim U(0, T)$ Uniformly sample a batch of t from the interval $[0, T]$
    Forward propagation: Compute $E(\mathbf{w})$ based on $\mathbb{T}$
    Backward propagation: Train the model weight $\mathbf{w}$ by $\nabla E(\mathbf{w})$
    Stopping criteria check.
  end
end

Since the initial condition $\hat{\mathbf{y}}(0;\mathbf{w}) = 0$ is always satisfied thanks to the multiplier construction, The training objective is to make $\hat{\mathbf{y}}(t;\mathbf{w})$ satisfying the corresponding ODE system by adjusting model's weights $\mathbf{w}$. Thus, the loss function is defined as

$$E(\mathbf{w}) = \frac{1}{|\mathbb{T}|}\sum_{t_i\in\mathbb{T}}\ell\left(\frac{\partial\hat{\mathbf{y}}(t_i;\mathbf{w})}{\partial t_i}, \Phi(\hat{\mathbf{y}}(t_i;\mathbf{w}))\right), \tag{9}$$

where Φ refers to the ODE system associated with the LP problem to be solved. $\mathbb{T} \subset \mathbb{R}_+$ refers to a set of collocation times to be trained,

where each collocation time t_i is sampled by a uniform distribution over the interval $[0, T]$. $\frac{\partial\hat{\mathbf{y}}(t_i;\mathbf{w})}{\partial t_i} \in \mathbb{R}^{n+m}$ represents the partial derivative of $\hat{\mathbf{y}}$ with respect to t_i. $\Phi(\hat{\mathbf{y}}(t_j;\mathbf{w})) \in \mathbb{R}^{n+m}$ represents the partial derivative expected to be according to the ODE system. $\ell(\cdot,\cdot)$ is an error metric used to measure the difference between these two vectors.

After the training, the neural network model $\hat{\mathbf{y}}(t;\mathbf{w})$ is expected to be well approximate the true state solutiong $\mathbf{y}(t)$ on interval $[0, T]$. Furthermore, the endpoint $\hat{\mathbf{y}}(T;\mathbf{w}) = (\hat{\mathbf{x}}, \hat{\mathbf{u}})$ is expected to solve the KKT conditions, where $\hat{\mathbf{x}}$ is the final prediction to the LP problem.