运筹学

4. 线性规划的单纯形解法

李 力
清华大学

Email: li-li@tsinghua.edu.cn

2023.10.

复习：线性规划模型和解

标准模型顶点的数学描述

标准模型可行集 $\Omega=\left\{Y, Y\in R^n \,\middle|\, \sum_{j=1}^{n} P_j y_j=\vec{b},\ Y\geq 0\right\}$

其中 $A=\left(P_1\ P_2\cdots P_n\right),\ P_j\in R^m, \forall j=1,2,\cdots,n$

等价的多面体模型可行集

$$\Omega:\ \sum_{j=1}^{n} P_j y_j\geq \vec{b},\ -\sum_{j=1}^{n} P_j y_j\geq -\vec{b},\ Y\geq 0$$

对任意 $X\in\Omega$ 可进行如下划分

$$x_j>0,\ \ j=k(1),\cdots,k(\hat{m}),\quad x_j=0,\ \ j=k(\hat{m}+1),\cdots,k(n)$$

结论：当且仅当 $\sum_{t=1}^{\hat{m}} P_{j(t)} y_t=\vec{b}$ 的解唯一时，X 是顶点

标准模型顶点的等价数学描述（教材定理2.2.3、2.2.4）

如果把 $X \in \Omega = \left\{ Y, Y \in R^n \,\middle|\, \sum_{j=1}^{n} P_j y_j = \vec{b},\ Y \geq 0 \right\}$ 的非零分量

称为正分量，那么任何可行解是顶点的充要条件为：

其正分量对应的系数向量（P_j）线性无关

即，如果 $X \in \Omega$ 划分为

存在一些等价的表述方式，可能作为判断题出现

$$x_j > 0,\ \ j = k(1), \cdots, k(\hat{m}),\quad x_j = 0,\ \ j = k(\hat{m}+1), \cdots, k(n)$$

其为顶点的充要条件是 $P_{k(t)}, t = 1, 2, \cdots, m$ 线性无关

（行满秩）标准模型顶点的数学描述

如果 $(P_1,\cdots,P_n)$ 是行满秩矩阵，那么 $X=(x_1,\cdots,x_n)^T$ 是

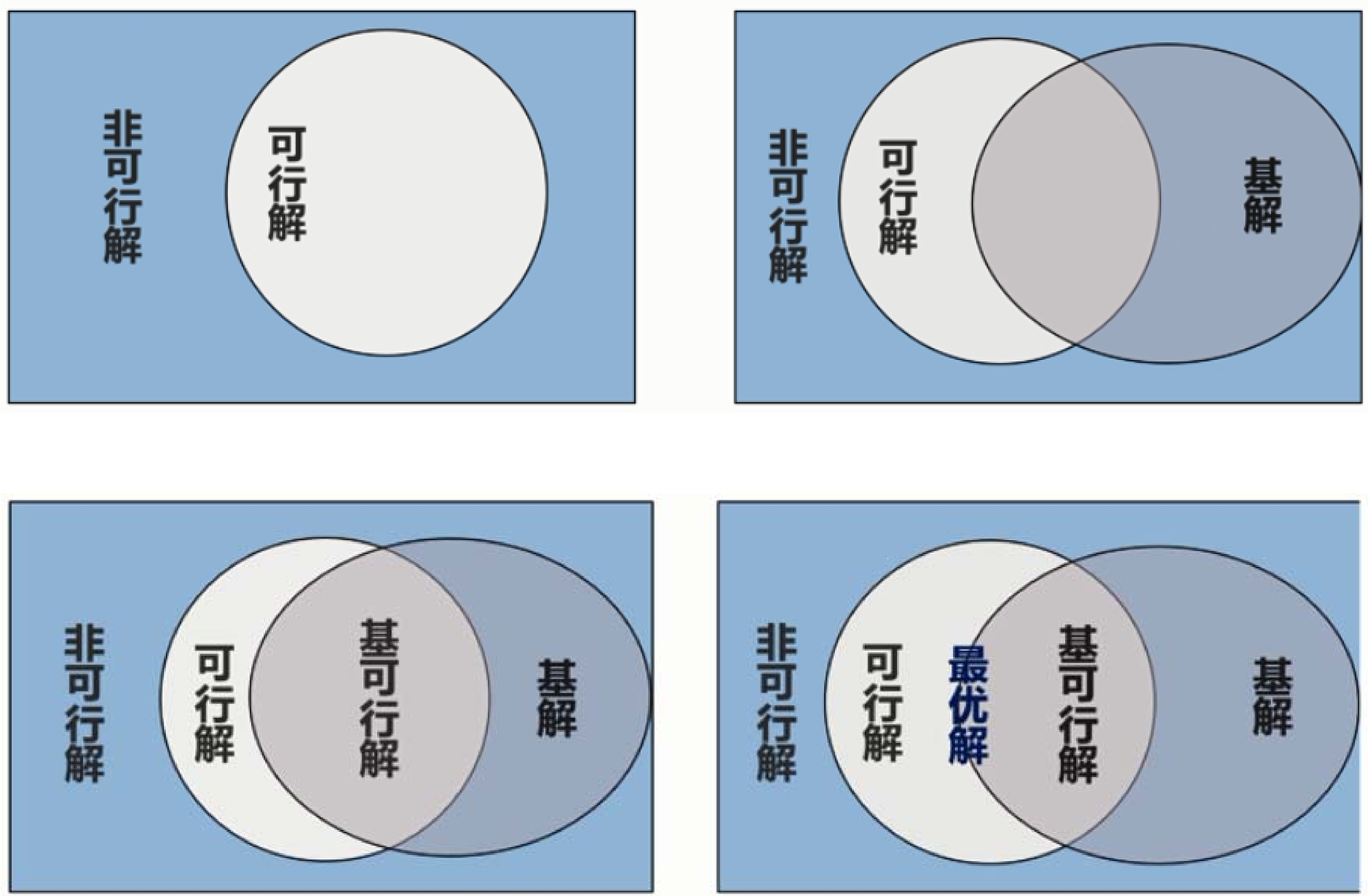
$$\Omega=\left\{Y, Y\in R^n \,\middle|\, \sum_{j=1}^{n} P_j y_j=\vec{b}, Y\geq 0\right\}$$

顶点的充要条件是：存在 $\left(k(1),\cdots,k(m)\right)\subseteq(1,\cdots,n)$ 满足

$$\begin{pmatrix} x_{k(1)} \\ \vdots \\ x_{k(m)} \end{pmatrix}=\left(P_{k(1)},\cdots,P_{k(m)}\right)^{-1}\vec{b}\geq 0,\quad x_{k(j)}=0, \forall m+1\leq j\leq n$$

（注意：前面结论中的 $\hat{m}$ 可以小于 m ）

称满足 $\left(P_{k(1)},\cdots,P_{k(m)}\right)^{-1}\vec{b}\geq 0$ 的矩阵为可行基矩阵


非可行解
可行解
非可行解
可行解
基解
非可行解
可行解
基可行解
基解
非可行解
可行解
最优解
基可行解
基解

关于标准模型顶点的两点说明

1）假定标准模型 $(P_1, \cdots, P_n)$ 是行满秩矩阵不失一般性

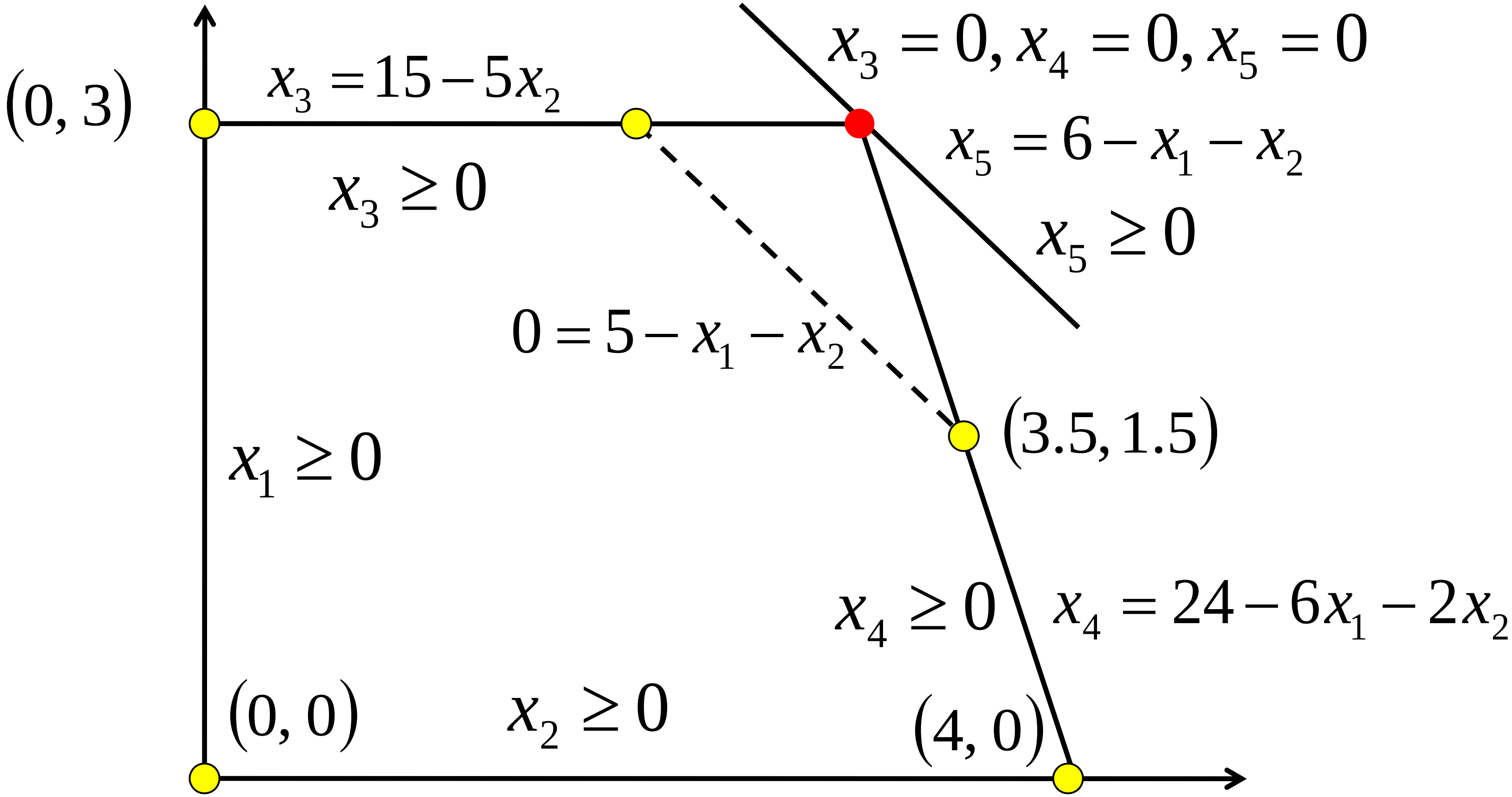
不满足该假定只有以下两种可能：

无可行解（不用考虑） 或 有多余约束（删除）

2）给定可行基矩阵可唯一确定一个顶点，反之不一定

若给定顶点有 m 个非零分量（非退化顶点），只有一个可行基矩阵可确定该顶点；否则（称为退化顶点），可能有多个可行基矩阵确定同一个顶点

退化情况


$x_3 = 0, x_4 = 0, x_5 = 0$
$(0, 3)$
$x_3 = 15 - 5x_2$
$x_5 = 6 - x_1 - x_2$
$x_3 \geq 0$
$x_5 \geq 0$
$0 = 5 - x_1 - x_2$
$(3.5, 1.5)$
$x_1 \geq 0$
$x_4 \geq 0$
$x_4 = 24 - 6x_1 - 2x_2$
$(0, 0)$
$x_2 \geq 0$
$(4, 0)$

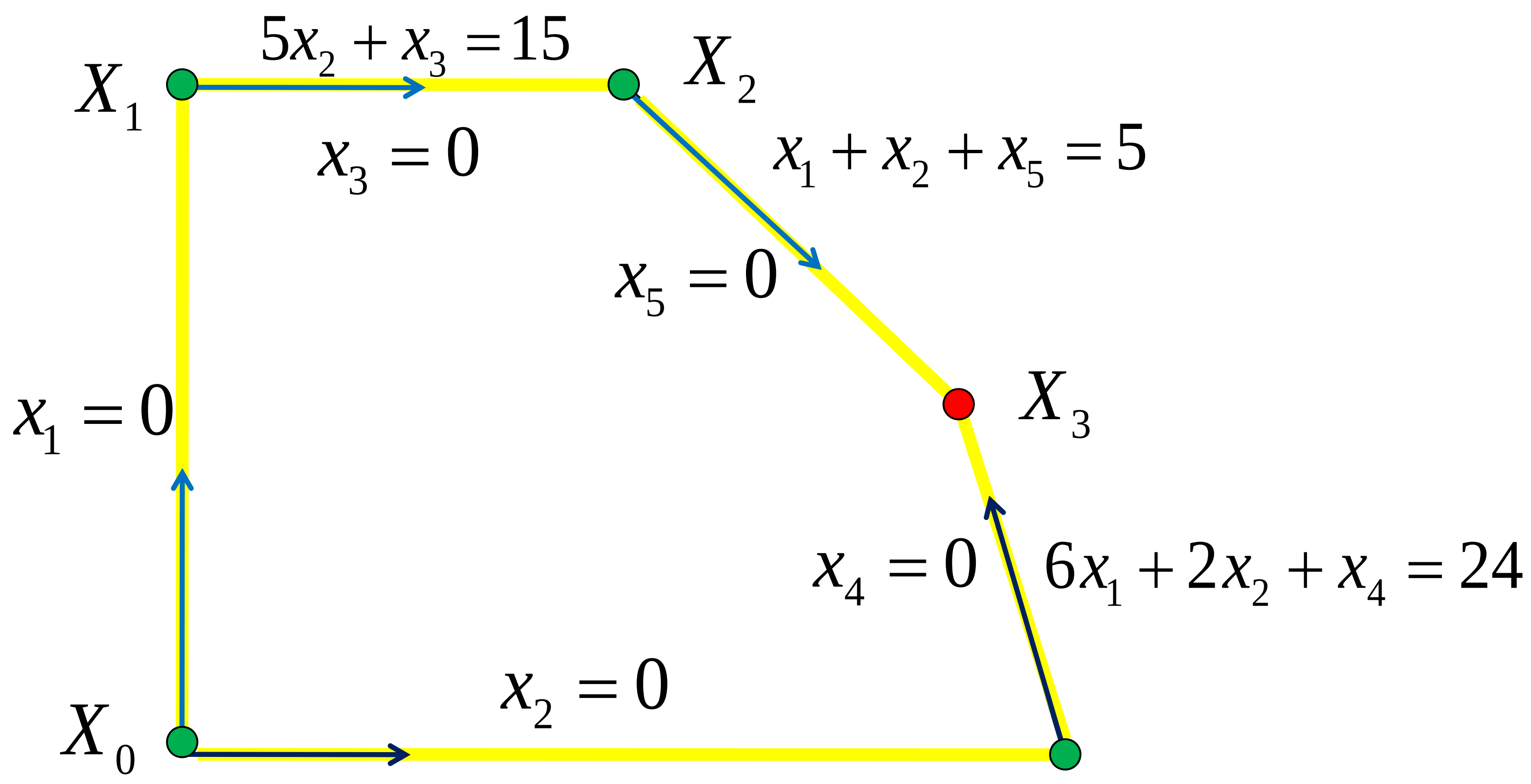
单纯型算法

（在顶点集搜索最优解的有效算法）

单纯型算法的基本想法

单纯形法最早由 **George Dantzig**于**1947**年提出，近**70**年来，虽有许多变形体已经开发，但却保持着同样的基本观念。如果线性规划问题的最优解存在，则一定可以在其可行区域的顶点中找到。基于此，单纯形法的基本思路是：先找出可行域的一个顶点，据一定规则判断其是否最优；若否，则转换到与之相邻的另一顶点，并使目标函数值更优；如此下去，直到找到某最优解为止

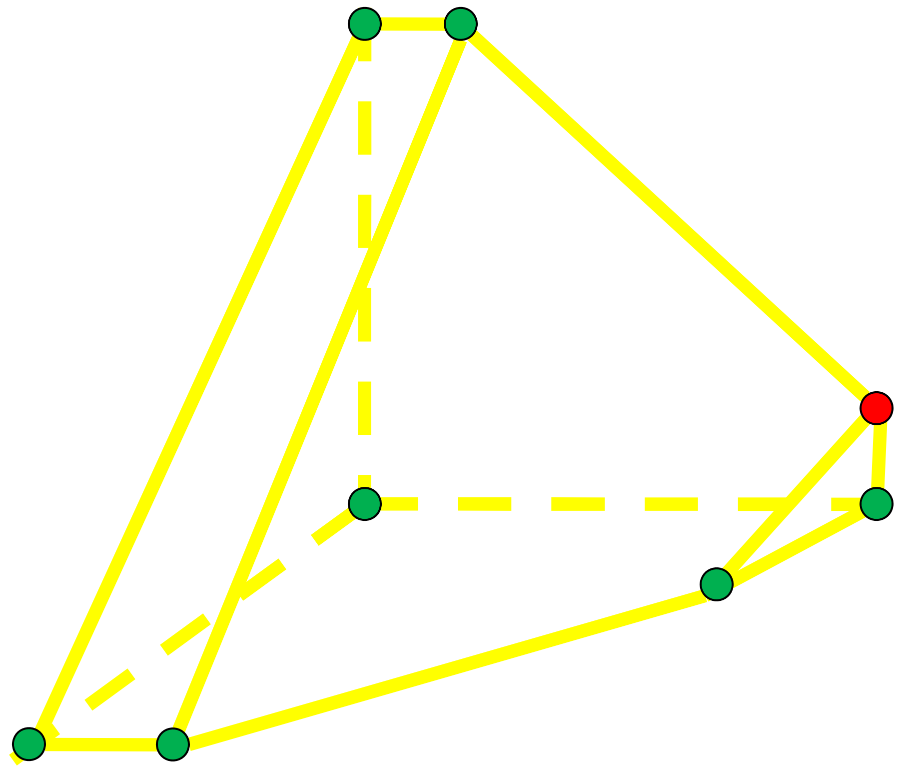
二维相邻顶点



1） 若顶点不是最优解，其相邻顶点集中有更好的顶点

2） 让一个非基变量增加到最大可行值就得到相邻顶点，该非基变量变成新顶点的基变量，称其为<u>进基变量</u>

三维相邻顶点



1）若顶点不是最优解，其相邻顶点集中有更好顶点

2）让一个非基变量变成基变量就得到一个相邻顶点

顶点集搜索方式：从任意顶点出发，让某个非基变量增加到最大值变成基变量，就可得到更好的相邻顶点

要解决的问题

如何确定一个初始顶点 4

⇓

1 如何计算选定进基变量后的顶点

⇓

2 如何选择进基变量使目标函数改进

⇓

3 如何判断已经找到最优顶点

如何计算选定进基变量后的顶点

例1 $\max \ 2x_1 + x_2 + 0x_3 + 0x_4 + 0x_5$

$$\text{s.t.} \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} x_4 + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} x_5 = \begin{pmatrix} 15 \\ 24 \\ 5 \end{pmatrix}$$

$$x_j \geq 0, \ j = 1,2,\cdots,5$$

其等式约束可写成 $$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 15 \\ 24 \\ 5 \end{pmatrix}$$

由此可看出 $\hat{X} = (0,0,15,24,5)^T$ 是一个<u>基本可行解</u>

约定b
非负

我们将<u>每个基变量只在一个等式中出现</u>，<u>且所有右边项均非负</u>的等式约束称为对应的<u>基本可行解</u>的<u>表示式</u>

假设我们想要让 x_2 变成基变量，即选择 x_2 为进基变量，根据基本可行解的表示式，必须让 x_2 只出现在一个等式约束中，可用以下消元法达到目的

在 $\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 15 \\ 24 \\ 5 \end{pmatrix}$ 的各行除以 x_2 的系数

可得

$$\begin{pmatrix} 0.2x_3 \\ 0.5x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 12 \\ 5 \end{pmatrix}$$

在 $\begin{pmatrix} 0.2x_3 \\ 0.5x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 12 \\ 5 \end{pmatrix}$ 中选定一行，用其它行减去该行，即可达到只有一行有 x_2 的目的，例如，用一、二行减去第三行可以得到

$$\begin{pmatrix} 0.2x_3 - x_5 \\ 0.5x_4 - x_5 \\ x_5 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} -2 \\ 7 \\ 5 \end{pmatrix}$$

整理后可得 $\begin{pmatrix} 0.2x_3 \\ 0.5x_4 \\ x_2 \end{pmatrix} + \begin{pmatrix} -1 \\ 2 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} x_5 = \begin{pmatrix} -2 \\ 7 \\ 5 \end{pmatrix}$

问题：第一个方程的右边出现负数！

为了避免前面的问题，在方程

$$\begin{pmatrix} 0.2x_3 \\ 0.5x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 12 \\ 5 \end{pmatrix}$$

中，只能用其它行减去右边常数最小的第一行

由此可得

$$\begin{pmatrix} 0.2x_3 \\ 0.5x_4 - 0.2x_3 \\ x_5 - 0.2x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 9 \\ 2 \end{pmatrix}$$

如此选择之后，出基变量也自动产生了

对 $\begin{pmatrix} 0.2x_3 \\ 0.5x_4 - 0.2x_3 \\ x_5 - 0.2x_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 9 \\ 2 \end{pmatrix}$ 整理后得到

$$\begin{pmatrix} x_2 \\ 0.5x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0.2 \\ -0.2 \\ -0.2 \end{pmatrix} x_3 = \begin{pmatrix} 3 \\ 9 \\ 2 \end{pmatrix}$$

再将第二行除以0.5得到新的基本可行解的表示式

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix} x_3 = \begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$$

该基本可行解是 $\overline{X} = (0,3,0,18,2)^T$，$x_3$ 变成非基变量

它是原来的基本可行解在保留 x_2 的行的基变量

由于 $$\begin{pmatrix} 0.2x_3 \\ 0.5x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 3 \\ 12 \\ 5 \end{pmatrix} = \begin{pmatrix} 15/5 \\ 24/2 \\ 5/1 \end{pmatrix}$$

来自当前的表示式 $$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 15 \\ 24 \\ 5 \end{pmatrix}$$

所以选择保留 x_2 的行的方法可以总结为选择达到

$\min\left\{ \frac{15}{5}, \frac{24}{2}, \frac{5}{1} \right\}$ 的行，一旦选定这样的行，在该行

的基变量将变成非基变量，从而确定了出基变量

现在已知基本可行解 $\overline{X}=(0,3,0,18,2)^T$ 的表示式为

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}+\begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix}x_1+\begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix}x_3=\begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$$

如果我们又想让 x_3 进基，由于

$$\min\left\{\frac{3}{0.2},\frac{18}{-0.4},\frac{2}{-0.2}\right\}=\frac{18}{-0.4}$$

采用前面总结的规则应该保留第二行的 x_3，但是若将第二行的 x_3 前的系数变成 **1**，必须在第二行除以 -0.4，此时右边系数将变成负数，所以，<u>只能选择进基变量系数非负的行</u>保留进基变量

必要条件

由于 $\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix} x_3 = \begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$，让 x_3 进基只能

用第一行的 x_3 消去其它行的 x_3，对于 x_3 的系数不是正数的行，我们需要将第一行乘以一个合适的正数加到相应的行，这种操作不会使右边项的数变成为负数，因此，在选择保留进基变量所在行的时候<u>不用考虑进基变量的系数不是正数的行</u>

总结前面的讨论，可得到下面的一般规则：

假设某基本可行解的表示式是

$$X_B + \hat{P}_{j(m+1)} x_{j(m+1)} + \cdots + \hat{P}_{j(n)} x_{j(n)} = \hat{X}_B$$

其中

$$X_B = \begin{pmatrix} x_{j(1)} \\ \vdots \\ x_{j(m)} \end{pmatrix}, \quad \hat{P}_{j(t)} = \begin{pmatrix} \hat{p}_{1\,j(t)} \\ \vdots \\ \hat{p}_{m\,j(t)} \end{pmatrix}, \forall t, \quad \hat{X}_B = \begin{pmatrix} \hat{x}_{j(1)} \\ \vdots \\ \hat{x}_{j(m)} \end{pmatrix}$$

如果要选 $x_{j(t)}$（$m+1 \le t \le n$）进基，则应该仅保留第 l 行的 $x_{j(t)}$，即 $x_{j(l)}$ 出基，其中 l 满足

$$\frac{\hat{x}_{j(l)}}{\hat{p}_{l\,j(t)}} = \min_{\hat{p}_{i\,j(t)}>0} \frac{\hat{x}_{j(i)}}{\hat{p}_{i\,j(t)}}$$

为获得 $x_{j(t)}$ 进基、$x_{j(l)}$ 出基后的基本可行解的表示式，需要对原来的表示式

$$X_B + \hat{P}_{j(m+1)} x_{j(m+1)} + \cdots + \hat{P}_{j(n)} x_{j(n)} = \hat{X}_B$$

进行行等价变换，使 $x_{j(t)}$ 前面的系数向量

$$\hat{P}_{j(t)} = \begin{pmatrix} \hat{p}_{1\,j(t)} \\ \vdots \\ \hat{p}_{l\,j(t)} \\ \vdots \\ \hat{p}_{m\,j(t)} \end{pmatrix} \quad \text{变成} \quad \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix}$$

具体做法：先在第 l 行除以 $\hat{p}_{l\,j(t)}$，再将第 l 行分别乘以 $-\hat{p}_{i\,j(t)}$ 加到第 $i\ (i = 1,2,\cdots,m, i \neq l)$ 行

可以利用数据表完成换基运算

$\hat{X} = (0,0,15,24,5)^T$ 的表示式

$$\begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 5 \\ 2 \\ 1 \end{pmatrix} x_2 = \begin{pmatrix} 15 \\ 24 \\ 5 \end{pmatrix}$$

设表时，我们的等号出现在这里

由下面的数据表完全确定

（基变量） （右边项）

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_3	0	5	1	0	0	15
x_4	6	2	0	1	0	24
x_5	1	1	0	0	1	5

让 x_2 进基是对数据表进行如下运算：

$$\min\left\{\frac{15}{5},\frac{24}{2},\frac{5}{1}\right\} \Rightarrow$$

BV	x_1	x_2	x_3	x_4	x_5	RHS	
x_3	0	5	1	0	0	15	①
x_4	6	2	0	1	0	24	②
x_5	1	1	0	0	1	5	③

		BV	x_1	x_2	x_3	x_4	x_5	RHS	
① ÷ **5**	$\Rightarrow$	x_2	0	1	0.2	0	0	3	④
④ × **(-2)** + ②	$\Rightarrow$	x_4	6	0	-0.4	1	0	18	
④ × **(-1)** + ③	$\Rightarrow$	x_5	1	0	-0.2	0	1	2	

特殊情况： $\hat{P}_{j(t)} \le 0 \quad \Rightarrow \quad$ 无法计算 $\dfrac{\hat{x}_{j(l)}}{\hat{p}_{l\,j(t)}} = \min\limits_{\hat{p}_{i\,j(t)}>0} \dfrac{\hat{x}_{j(i)}}{\hat{p}_{i\,j(t)}}$

任取正数 δ，定义 n 维向量 $X(\delta)$ 的分量如下

$$X_B(\delta) = \hat{X}_B - \hat{P}_{j(t)}\delta,$$

基变量

$$x_{j(t)}(\delta) = \delta, \quad x_{j(i)}(\delta) = 0, \quad \forall m+1 \le i \le n,\ i \ne t$$

非基变量

$$\hat{P}_{j(t)} \le 0 \quad \Rightarrow \quad X(\delta) \ge 0$$

$$X_B(\delta) + \hat{P}_{j(m+1)} x_{j(m+1)}(\delta) + \cdots + \hat{P}_{j(n)} x_{j(n)}(\delta) = \hat{X}_B - \hat{P}_{j(t)}\delta + \hat{P}_{j(t)}\delta = \hat{X}_B$$

$\Rightarrow \quad X(\delta)$ 是可行解

<u>在其它非基变量不变的情况下可令 $x_{j(t)}(\delta)$ 趋于正无穷</u>

小结

假定已知基本可行解 $\hat{X}$ 的表示式为

$$X_B + \hat{P}_{j(m+1)} x_{j(m+1)} + \cdots + \hat{P}_{j(n)} x_{j(n)} = \hat{X}_B$$

任取 $m+1 \le t \le n$，则有以下结论：

1）如果 $\hat{P}_{j(t)} \le 0$，变量 $x_{j(t)}$ 在可行集可趋于无穷大，同时保持其它非基变量不变

2）只要 $\hat{P}_{j(t)}$ 有一个分量大于 **0**，就可以通过行变换让 $x_{j(t)}$ 进基，形成一个新的基本可行解

如何选择进基变量使目标函数改进

对例1 $\max\ 2x_1 + x_2 + 0x_3 + 0x_4 + 0x_5$

$$\text{s.t.}\quad \begin{pmatrix}0\\6\\1\end{pmatrix}x_1 + \begin{pmatrix}5\\2\\1\end{pmatrix}x_2 + \begin{pmatrix}1\\0\\0\end{pmatrix}x_3 + \begin{pmatrix}0\\1\\0\end{pmatrix}x_4 + \begin{pmatrix}0\\0\\1\end{pmatrix}x_5 = \begin{pmatrix}15\\24\\5\end{pmatrix}$$

$$x_j \geq 0,\ j = 1,2,\cdots,5$$

已经得到两个基本可行解，即

$$\hat{X} = (0,0,15,24,5)^T \qquad \bar{X} = (0,3,0,18,2)^T$$

记

$$f(X) = 2x_1 + x_2 + 0x_3 + 0x_4 + 0x_5$$

则

$$f(\hat{X}) = 0,\ \ f(\bar{X}) = 3$$

如何找到其目标函数值大于 $f(\bar{X})$ 的基本可行解？

已知 $\bar{X}=(0,3,0,18,2)^T$ 的表示式为

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix} x_3 = \begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$$

将上式确定的<u>基变量对非基变量的函数关系</u>代入目标函数 $f(X)=2x_1+x_2+0x_3+0x_4+0x_5$ ，可以得到

$$f(X)=3+2x_1-0.2x_3=f(\bar{X})+2x_1-0.2x_3$$

由于每个变量都不能小于**0**，由上式可知，当且仅当 x_1 取正数（等价于让其进基）时，才能获得比 $f(\bar{X})$ 更大的目标函数值

如前所述，让 x_1 进基是对数据表进行如下运算：

BV	x_1	x_2	x_3	x_4	x_5	RHS	
x_2	0	1	0.2	0	0	3	①
x_4	6	0	-0.4	1	0	18	②
x_5	1	0	-0.2	0	1	2	③

$$\min\left\{\frac{18}{6},\frac{2}{1}\right\} \Rightarrow$$

③ × **(-6)** + ② $\Rightarrow$

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	0	0	0.8	1	-6	6
x_1	1	0	-0.2	0	1	2

根据数据表

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	0	0	-0.8	1	-6	6
x_1	1	0	-0.2	0	1	2

马上可知新的基本可行解为 $\tilde{X}=(2,3,0,6,0)^T$

将以上表格确定的基变量对非基变量的函数关系代入目标函数 $f(X)=f(\bar{X})+2x_1-0.2x_3$ 又可得新目标函数式

$$f(X)=f(\bar{X})+4+0.2x_3-2x_5=f(\tilde{X})+0.2x_3-2x_5$$

其中 $f(\tilde{X})=f(\bar{X})+4=7$ 是 $\tilde{X}$ 对应的目标函数值

用 z 表示线性规划标准型的目标函数，它和 X 之间的函数关系完全由以下线性方程组所确定

$$P_1x_1+P_2x_2+\cdots+P_nx_n=\vec{b}$$

$$c_1x_1+c_2x_2+\cdots+c_nx_n=z$$

可将其写成下面的扩充的等式约束形式

$$\begin{pmatrix}P_1\\c_1\end{pmatrix}x_1+\begin{pmatrix}P_2\\c_2\end{pmatrix}x_2+\cdots+\begin{pmatrix}P_n\\c_n\end{pmatrix}x_n=\begin{pmatrix}\vec{b}\\z\end{pmatrix}$$

例如，对于例**1**，其扩充的等式约束为

$$\begin{pmatrix}0\\6\\1\\2\end{pmatrix}x_1+\begin{pmatrix}5\\2\\1\\1\end{pmatrix}x_2+\begin{pmatrix}1\\0\\0\\0\end{pmatrix}x_3+\begin{pmatrix}0\\1\\0\\0\end{pmatrix}x_4+\begin{pmatrix}0\\0\\1\\0\end{pmatrix}x_5=\begin{pmatrix}15\\24\\5\\z\end{pmatrix}$$

对原来的等式约束进行行变换得到 $\bar{X}=(0,3,0,18,2)^T$ 的表示式

$$\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}+\begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix}x_1+\begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix}x_3=\begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$$

相当于对扩充的等式约束的前三行进行变换获得

$$\begin{pmatrix} 0 \\ 6 \\ 1 \\ 2 \end{pmatrix}x_1+\begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}x_2+\begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \\ 0 \end{pmatrix}x_3+\begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}x_4+\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}x_5=\begin{pmatrix} 3 \\ 18 \\ 2 \\ z \end{pmatrix}$$

将 $\begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix} + \begin{pmatrix} 0 \\ 6 \\ 1 \end{pmatrix} x_1 + \begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \end{pmatrix} x_3 = \begin{pmatrix} 3 \\ 18 \\ 2 \end{pmatrix}$ 所确定的基变量对非基变量的函数关系代入 $z = 2x_1 + x_2 + 0x_3 + 0x_4 + 0x_5$ 以获得仅含非基变量的 $z = 3 + 2x_1 - 0.2x_3$，相当于利用下面前三行等式<u>将第四行的基变量的系数变成**0**</u>

$$\begin{pmatrix} 0 \\ 6 \\ 1 \\ 2 \end{pmatrix} x_1 + \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} x_2 + \begin{pmatrix} 0.2 \\ -0.4 \\ -0.2 \\ 0 \end{pmatrix} x_3 + \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} x_4 + \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix} x_5 = \begin{pmatrix} 3 \\ 18 \\ 2 \\ z \end{pmatrix}$$

对于扩充约束

$$\begin{pmatrix}0\\6\\1\\2\end{pmatrix}x_1+\begin{pmatrix}1\\0\\0\\1\end{pmatrix}x_2+\begin{pmatrix}0.2\\-0.4\\-0.2\\0\end{pmatrix}x_3+\begin{pmatrix}0\\1\\0\\0\end{pmatrix}x_4+\begin{pmatrix}0\\0\\1\\0\end{pmatrix}x_5=\begin{pmatrix}3\\18\\2\\z\end{pmatrix}$$

将第一行乘以**-1**加到第四行就可以得到

$$\begin{pmatrix}0\\6\\1\\2\end{pmatrix}x_1+\begin{pmatrix}1\\0\\0\\0\end{pmatrix}x_2+\begin{pmatrix}0.2\\-0.4\\-0.2\\-0.2\end{pmatrix}x_3+\begin{pmatrix}0\\1\\0\\0\end{pmatrix}x_4+\begin{pmatrix}0\\0\\1\\0\end{pmatrix}x_5=\begin{pmatrix}3\\18\\2\\z-3\end{pmatrix}$$

我们将其称为基本可行解 $\bar{X}$ 的扩充表示式

从扩充表示式

$$\begin{pmatrix}0\\6\\1\\2\end{pmatrix}x_1+\begin{pmatrix}1\\0\\0\\0\end{pmatrix}x_2+\begin{pmatrix}0.2\\-0.4\\-0.2\\-0.2\end{pmatrix}x_3+\begin{pmatrix}0\\1\\0\\0\end{pmatrix}x_4+\begin{pmatrix}0\\0\\1\\0\end{pmatrix}x_5=\begin{pmatrix}3\\18\\2\\z-3\end{pmatrix}$$

可以获得下述信息：

1） $\overline{X}=(0,3,0,18,2)^T$ 是基本可行解

2） $\overline{X}$ 的目标函数值满足 $0=z-3$，即 $z=3$

3） 目标函数可以写成 $z=3+2x_1-0.2x_3$，因此让 x_1 进基能够增加目标函数值

将约束矩阵 A 划分为基矩阵 B 和非基矩阵 N，则目标函数和约束矩阵可以写成

$$\begin{cases} Bx_B + Nx_N = b \\ z = c_N^T x_N + c_B^T x_B \end{cases},$$

将（1）中第一个等式的两边同时乘以 B^{-1}，可以得到

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

把结果带入（1）中第二个等式，可以得到

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

定义检验数 $\sigma^T = c_N^T - c_B^T B^{-1}N$，则 $z + \sigma^T x_N = c_B^T B^{-1}b$。所以，式（1）可以写成

$$\begin{cases} x_B + \quad B^{-1}Nx_N = B^{-1}b \\ 0 + (c_N^T - c_B^T B^{-1}N)x_N = z - c_B^T B^{-1}b \end{cases}$$

前面由 $\bar{X}=(0,3,0,18,2)^T$ 的表示式获得其扩充表示式的过程可利用下面扩充的数据表（单纯型表）完成

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	6	0	-0.4	1	0	18
x_5	1	0	-0.2	0	1	2
	2	1	0	0	0	z

其中前面三行数据由 $\bar{X}$ 的表示式确定，最后一行是目标函数和变量间的（任意一种）约束式

对前面的单纯型表通过行变换将最后一行的基变量前面的系数变成**0**就得到下面的单纯型表

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	6	0	-0.4	1	0	18
x_5	1	0	-0.2	0	1	2
	2	0	-0.2	0	0	$z-3$

该表能够完全确定基本可行解 $\overline{X}$ 的扩充表示式，

我们将其称为 $\overline{X}$ 的单纯型表

利用 $\overline{X}$ 的单纯型表，很容易获得让 x_1 进基后的基本可行解的单纯型表，即先由右边项和 x_1 前面的系数的比值确定出基变量为 x_5

$$\min\left\{\frac{18}{6},\frac{2}{1}\right\} \Rightarrow$$

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	6	0	-0.4	1	0	18
x_5	1	0	-0.2	0	1	2
	2	0	-0.2	0	0	$z-3$

然后通过行变换将 x_1 所在列除了第三行以外的系数变成**0**即可得到新的基本可行解对应的单纯型表

新的基本可行解对应的单纯型表为

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	0	0	0.8	1	-6	6
x_1	1	0	-0.2	0	1	2
	0	0	0.2	0	-2	$z-7$

据此可知：

1） $\tilde{X}=(2,3,0,6,0)^T$ 是基本可行解

2） $\tilde{X}$ 的目标函数值满足 $0=z-7$，即 $z=7$

3） 让 x_3 进基能够增加目标函数值

如何判断已经找到最优顶点

由 $\tilde{X}$ 的单纯型表马上可确定出基变量为 x_4

$$\min\left\{\frac{3}{0.2},\frac{6}{0.8}\right\} \Rightarrow$$

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0.2	0	0	3
x_4	0	0	0.8	1	-6	6
x_1	1	0	-0.2	0	1	2
	0	0	0.2	0	-2	$z-7$

再通过行变换将 x_3 所在列除了第二行以外的系数变成**0**就可得到新的基本可行解对应的单纯型表

新的基本可行解对应的单纯型表为

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0	-0.25	1.5	1.5
x_3	0	0	1	1.25	-7.5	7.5
x_1	1	0	0	0.25	-0.5	3.5
	0	0	0	-0.25	-0.5	$z-8.5$

据此可知：

1） $\breve{X}=(3.5,1.5,7.5,0,0)^T$ 是基本可行解

2） $\breve{X}$ 的目标函数值满足$0=z-8.5$，即 $z=8.5$

3） 任何非基变量进基都不能增加目标函数值

已知 $\breve{X}=(3.5,1.5,7.5,0,0)^T$ 对应的单纯型表为

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0	-0.25	1.5	1.5
x_3	0	0	1	1.25	-7.5	7.5
x_1	1	0	0	0.25	-0.5	3.5
	0	0	0	-0.25	-0.5	$z-8.5$

由最后一行可知，任何可行解的目标函数都要满足 $z=8.5-0.25x_4-0.5x_5$，由于 $x_4\geq 0, x_5\geq 0$，任何可行解的目标函数值都不会大于 $\breve{X}$ 的目标函数值 **8.5**，所以 $\breve{X}$ 是该问题的最优解（<u>最优解的充分条件</u>）

上述例 **1**求解方法就是线性规划单纯型法，现在以下面的一般性线性规划标准型为对象总结其基本步骤

这里是求最大问题

$$\begin{aligned} &\max \quad z \\ &\text{s.t.} \quad P_1x_1 + P_2x_2 + \cdots + P_nx_n = \vec{b} \\ &\qquad\ c_1x_1 + c_2x_2 + \cdots + c_nx_n = z \\ &\qquad\ x_j \geq 0,\ \forall 1 \leq j \leq n \end{aligned}$$

出发点：

已知一个可逆方阵 $B = \left(P_{j(1)}, P_{j(2)}, \cdots, P_{j(m)}\right)$ 满足

$$B^{-1}\vec{b} \geq 0$$

其中 $1 \leq j(i) \leq n, \forall 1 \leq i \leq m$ ，即 B 是线性规划标准型的可行基阵

利用基阵 $B=\left(P_{j(1)},\cdots,P_{j(m)}\right)$对等式作以下等价变换

$$P_1x_1+\cdots+P_nx_n=\vec{b}$$

$$\Leftrightarrow\quad \left(P_{j(1)},\cdots,P_{j(m)}\right)\begin{pmatrix}x_{j(1)}\\ \vdots\\ x_{j(m)}\end{pmatrix}+P_{j(m+1)}x_{j(m+1)}+\cdots+P_{j(n)}x_{j(n)}=\vec{b}$$

$$\Leftrightarrow\quad \begin{pmatrix}x_{j(1)}\\ \vdots\\ x_{j(m)}\end{pmatrix}+\left(B^{-1}P_{j(m+1)}\right)x_{j(m+1)}+\cdots+\left(B^{-1}P_{j(n)}\right)x_{j(n)}=B^{-1}\vec{b}$$

由此得到一个基本可行解的表示式

记

$$X_B=\begin{pmatrix} x_{j(1)} \\ \vdots \\ x_{j(m)} \end{pmatrix},\ \hat{P}_j=B^{-1}P_j=\begin{pmatrix} \hat{p}_{1j} \\ \vdots \\ \hat{p}_{mj} \end{pmatrix},\forall 1\le j\le n+1,\quad \hat{P}_{n+1}=B^{-1}\vec{b}$$

对应表示式为 $X_B+\hat{P}_{j(m+1)}x_{j(m+1)}+\cdots+\hat{P}_{j(n)}x_{j(n)}=\hat{P}_{n+1}$

再记 $C_B=\left(c_{j(1)},\cdots,c_{j(m)}\right)^T$，由上式可得

$$C_B^T X_B+C_B^T\hat{P}_{j(m+1)}x_{j(m+1)}+\cdots+C_B^T\hat{P}_{j(n)}x_{j(n)}=C_B^T\hat{P}_{n+1}$$

和目标约束 $C_B^T X_B+c_{j(m+1)}x_{j(m+1)}+\cdots+c_{j(n)}x_{j(n)}=z$ 相减可得

$$z-C_B^T\hat{P}_{n+1}=\left(c_{j(m+1)}-C_B^T\hat{P}_{j(m+1)}\right)x_{j(m+1)}+\cdots+\left(c_{j(n)}-C_B^T\hat{P}_{j(n)}\right)x_{j(n)}$$

记为 $$z-\hat{z}=\sigma_{j(m+1)}x_{j(m+1)}+\cdots+\sigma_{j(n)}x_{j(n)}$$

将前面获得的数据填入下述表格

BV	x_1	$\cdots$	x_k	$\cdots$	x_n	RHS
$x_{j(1)}$	$\hat{p}_{11}$	$\cdots$	$\hat{p}_{1k}$	$\cdots$	$\hat{p}_{1n}$	$\hat{p}_{1n+1}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$x_{j(m)}$	$\hat{p}_{m1}$	$\cdots$	$\hat{p}_{mk}$	$\cdots$	$\hat{p}_{mn}$	$\hat{p}_{mn+1}$
	σ_1	$\cdots$	σ_k	$\cdots$	σ_n	$z-\hat{z}$

其中 $\left(\hat{P}_{j(1)},\cdots,\hat{P}_{j(m)}\right)=I_m$， $\hat{z}=C_B^T\hat{P}_{n+1}=C_B^TB^{-1}\vec{b}$

$$\sigma_j=c_j-C_B^T\hat{P}_j=c_j-C_B^TB^{-1}P_j,\quad \forall 1\le j\le n$$

称 $\sigma_1,\cdots,\sigma_n$ 为<u>检验数</u>，可看出基变量检验数等于**0**

由于目标函数可以表示为

$$z=\hat{z}+\sigma_{j(m+1)}x_{j(m+1)}+\sigma_{j(m+2)}x_{j(m+2)}+\cdots+\sigma_{j(n)}x_{j(n)}$$

令 $\sigma_{j(t)}=\max\left\{\sigma_{j(m+1)},\sigma_{j(m+2)},\cdots,\sigma_{j(n)}\right\}$

如果 $\sigma_{j(t)}\le 0$

则有 $\sigma_{j(i)}\le 0,\quad \forall m+1\le i\le n$

请注意这里取最大不是为了让最大的条件数对应的基进基，参见下面说明

则根据上式有 $z\le\hat{z}$ 可知 $\hat{z}$ 已经是最优目标函数值

如果 $\sigma_{j(t)}>0$

如果是min问题呢

由最上面表达式可知让 $x_{j(t)}$ 进基能增加目标函数值

对于单纯型表

$$\sigma_j = c_j - C_B^T \hat{P}_j$$
$$= c_j - C_B^T B^{-1} P_j$$

BV	x_1	$\cdots$	x_k	$\cdots$	x_n	RHS
$x_{j(1)}$	$\hat{p}_{11}$	$\cdots$	$\hat{p}_{1k}$	$\cdots$	$\hat{p}_{1n}$	$\hat{p}_{1n+1}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$x_{j(m)}$	$\hat{p}_{m1}$	$\cdots$	$\hat{p}_{mk}$	$\cdots$	$\hat{p}_{mn}$	$\hat{p}_{mn+1}$
	σ_1	$\cdots$	σ_k	$\cdots$	σ_n	$z-\hat{z}$

如果 $\hat{p}_{ij(t)} \le 0, \ \forall 1 \le i \le m$，是讨论进基方法时遇到的特殊情况，此时可让非基变量 $x_{j(t)}$ 单独趋于无穷，由 $z = \hat{z} + \sigma_{j(m+1)} x_{j(m+1)} + \cdots + \sigma_{j(n)} x_{j(n)} = \hat{z} + \sigma_{j(t)} x_{j(t)}$ 以及 $\sigma_{j(t)} > 0$ 可断定该问题没有有限的最优目标值

请注意，这里我们对于条件数的定义和刁在筠老师书上恰好相反，我们实际定义的是

$$z = \hat{z} + \sigma_{j(m+1)} x_{j(m+1)} + \cdots + \sigma_{j(n)} x_{j(n)}$$

刁在筠老师书上实际定义的条件数是 $-\sigma_{j(m+1)}$

$$z = \hat{z} - (-\sigma_{j(m+1)}) x_{j(m+1)} - \cdots - (-\sigma_{j(n)}) x_{j(n)}$$

或者等价的

$$z - \sigma_{j(m+1)} x_{j(m+1)} - \cdots - \sigma_{j(n)} x_{j(n)} = \hat{z}$$

因此我们这里的条件数和刁在筠老师书上的条件数判断条件反号

排除掉前面的情况，可得

$1 \le l \le m$ 满足

$$\frac{\hat{p}_{l\,n+1}}{\hat{p}_{lj(t)}} = \min_{\hat{p}_{ij(t)}>0} \frac{\hat{p}_{i\,n+1}}{\hat{p}_{ij(t)}}$$

由此可知出基变量为 $x_{j(l)}$，因此要通过行变换将单纯型表中 $x_{j(t)}$ 列的数据完全变成 $x_{j(l)}$ 列的数据，如右边所示

$$\begin{array}{ccc}
x_{j(t)} & \cdots & x_{j(l)} \\
\hat{p}_{1j(t)} & \cdots & 0 \\
\vdots & \cdots & \vdots \\
\hat{p}_{l-1\,j(t)} & \cdots & 0 \\
\boxed{\hat{p}_{lj(t)} \quad \cdots \quad 1} \\
\hat{p}_{l+1\,j(t)} & \cdots & 0 \\
\vdots & \cdots & \vdots \\
\hat{p}_{mj(t)} & \vdots & 0 \\
\sigma_{j(t)} & \cdots & 0
\end{array}$$

参照右边两列数，很容易获得计算新单纯型表的公式：

第 l 行

$$\hat{p}'_{lj} = \hat{p}_{lj} / \hat{p}_{lj(t)} \text{ ，} \forall 1 \le j \le n+1$$

第 i 行（ $1 \le i \le m, i \ne l$ ）

$$\hat{p}'_{ij} = \hat{p}_{ij} - \hat{p}'_{lj} \hat{p}_{ij(t)} \text{ ，} \forall 1 \le j \le n+1$$

第 $m+1$ 行

$$\sigma'_j = \sigma_j - \hat{p}'_{lj} \sigma_{j(t)} \text{ ，} \forall 1 \le j \le n+1$$

$$\hat{z}' = \hat{z} + \hat{p}'_{l\,n+1} \sigma_{j(t)}$$

$$\begin{array}{ccc}
x_{j(t)} & \cdots & x_{j(l)} \\
\hat{p}_{1j(t)} & \cdots & 0 \\
\vdots & \cdots & \vdots \\
\hat{p}_{l-1j(t)} & \cdots & 0 \\
\hat{p}_{lj(t)} & \cdots & 1 \\
\hat{p}_{l+1j(t)} & \cdots & 0 \\
\vdots & \cdots & \vdots \\
\hat{p}_{mj(t)} & \vdots & 0 \\
\sigma_{j(t)} & \cdots & 0
\end{array}$$

将出基变量 $x_{j(l)}$ 换成进基变量 $x_{j(t)}$，即重新令

$$j(l) = j(t)$$

可以得到下面的新数据表

BV	x_1	$\cdots$	x_k	$\cdots$	x_n	RHS
$x_{j(1)}$	$\hat{p}'_{11}$	$\cdots$	$\hat{p}'_{1k}$	$\cdots$	$\hat{p}'_{1n}$	$\hat{p}'_{1n+1}$
$\vdots$	$\vdots$	$\cdots$	$\vdots$	$\cdots$	$\vdots$	$\vdots$
$x_{j(m)}$	$\hat{p}'_{m1}$	$\cdots$	$\hat{p}'_{mk}$	$\cdots$	$\hat{p}'_{mn}$	$\hat{p}'_{mn+1}$
	σ'_1	$\cdots$	σ'_k	$\cdots$	σ'_n	$z - \hat{z}'$

对于除了$x_{j(l)}$之外的原来的基变量，即

$$x_k, \forall k \in \{\, j(1),\cdots, j(l-1), j(l+1),\cdots, j(m) \,\}$$

由于 $\hat{p}_{lk}=0$ ，有 $\hat{p}'_{lk}=\hat{p}_{lk}/\hat{p}_{lt}=0$，所以

$$\hat{p}'_{ik}=\hat{p}_{ik}-\hat{p}'_{lk}\hat{p}_{it}=\hat{p}_{ik},\ \forall 1 \le i \le m, i \ne l$$

$$\sigma'_k=\sigma_k-\hat{p}'_{lk}\sigma_t=\sigma_k$$

说明它们的数据没有变化，因此还是基变量，由此可知，新数据表就是用 $x_{j(t)}$ 替换原来的 $x_{j(l)}$ 所形成的基本可行解对应的单纯型表，因此可以以此为起点重复前面的过程

可以用单纯形表表示为:

	x_B	x_N	RHS
x_B	I	$B^{-1}N$	$B^{-1}b$
	0	σ^T	$z-c_B^TB^{-1}b$

此时，令 $x_N=0$ ，可以得到 $x_B=B^{-1}b$，即当前的基本可行解。此时目标函数的值为

$z=c_B^TB^{-1}b$ 。

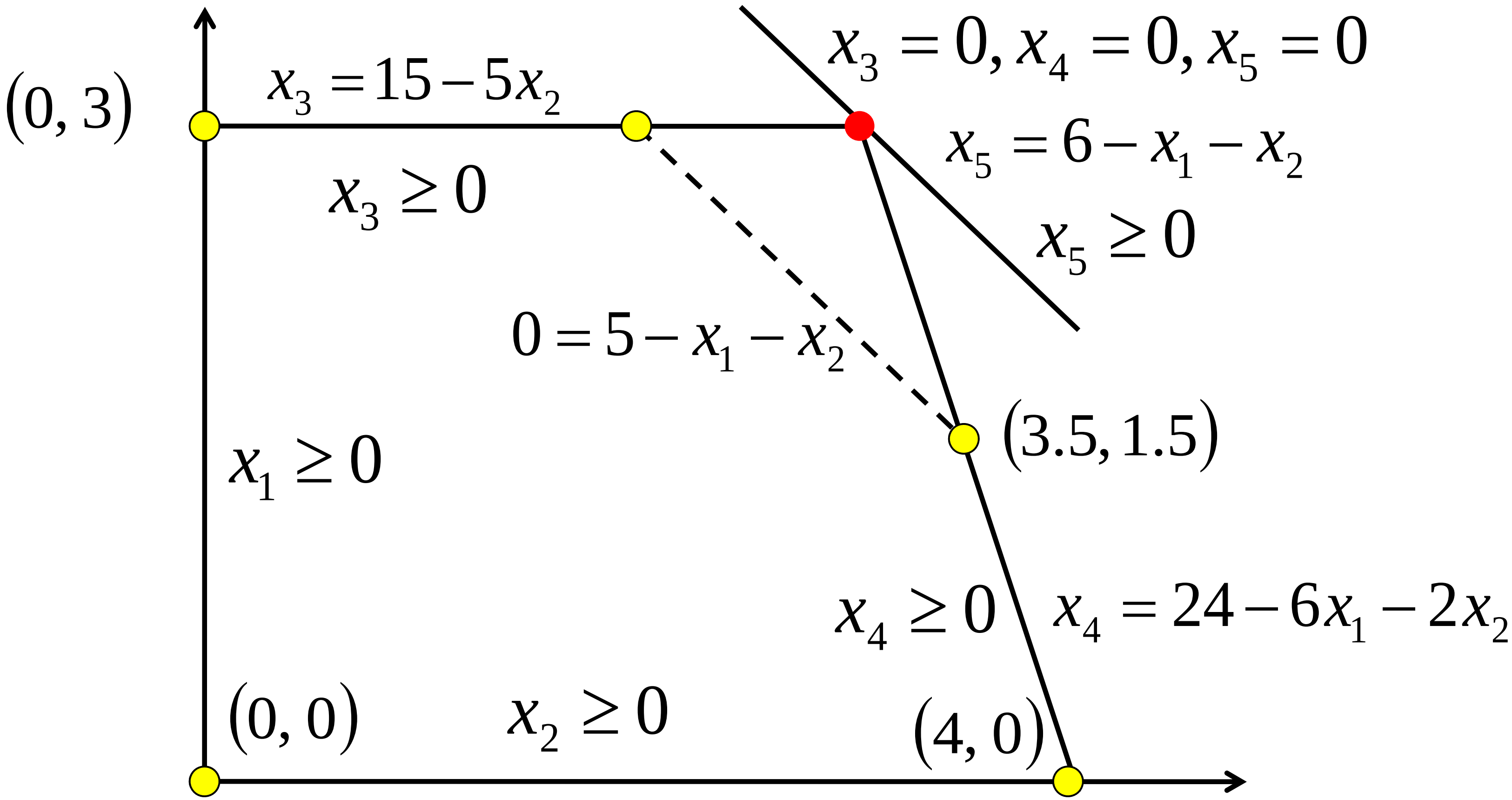
保证单纯型算法收敛的充分条件

如果在迭代过程中始终有 $\hat{P}_{n+1} = B^{-1}\vec{b} > 0$ ，即迭代过程中产生的每个基本可行解的基变量数值都严格大于**0**（称其为非退化条件），此时一定有

$$\hat{p}'_{l\,n+1} = \hat{p}_{l\,n+1} / \hat{p}_{lj(t)} > 0,\ \ \hat{z}' = \hat{z} + \hat{p}'_{l\,n+1}\sigma_{j(t)} > \hat{z}$$

即每步迭代都能保证目标函数严格增加，由于基本可行解的数目是有限的，上述过程不会无限进行，因此一定在有限次迭代后出现所有检验数都不大于 **0** 的情况，从而得到最优的基本可行解

退化情况



红点对应的基本可行解有三个基阵，可以写成

$$(P_1,P_2,P_3)^{-1}\vec{b} \quad (P_1,P_2,P_4)^{-1}\vec{b} \quad (P_1,P_2,P_5)^{-1}\vec{b}$$

退化情况产生的问题

退化情况的本质是多个可行基阵对应于一个顶点，此时经过一次进出基迭代后得到的是同一个顶点，因此有可能出现单纯型算法在一个顶点的几个基矩阵之间循环迭代的情况

退化导致（最大检验数规则）不收敛的例子（**Beale, E.M.L，1955**）

$$\max \frac{3}{4}x_4 - 20x_5 + \frac{1}{2}x_6 - 6x_7$$

$$\text{s.t. } x_1 + \frac{1}{4}x_4 - 8x_5 - x_6 + 9x_7 = 0$$

$$x_2 + \frac{1}{2}x_4 - 12x_5 - \frac{1}{2}x_6 + 3x_7 = 0$$

$$x_3 + x_6 = 1$$

$$x_i \geq 0, i = 1, 2, \cdots, 7$$

初始基变量 $\{x_1, x_2, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	0	0	$1/4$	-8	-1	9	0
x_2	0	1	0	$1/2$	-12	$-1/2$	3	0
x_3	0	0	1	0	0	1	0	1
	0	0	0	$3/4$	-20	$1/2$	-6	0

一次迭代后的基变量 $\{x_4, x_2, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_4	4	0	0	1	-32	-4	36	0
x_2	-2	1	0	0	4	$3/2$	-15	0
x_3	0	0	1	0	0	1	0	1
	-3	0	0	0	4	$7/2$	-33	0

二次迭代后的基变量 $\{x_4, x_5, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_4	-12	8	0	1	0	8	-84	0
x_5	$-1/2$	$1/4$	0	0	1	$3/8$	$-15/4$	0
x_3	0	0	1	0	0	1	0	1
	-1	-1	0	0	0	2	-18	0

三次迭代后的基变量 $\{x_6, x_5, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_6	$-2/3$	1	0	$1/8$	0	1	$-21/2$	0
x_5	$1/16$	$-1/8$	0	$-3/64$	1	0	$3/16$	0
x_3	$3/2$	-1	1	$-1/8$	0	0	$21/2$	1
	2	-3	0	$-1/4$	0	0	3	0

四次迭代后的基变量 $\{x_6, x_7, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_6	2	–6	0	–5/2	56	1	0	0
x_7	1/3	–2/3	0	–1/4	16/3	0	1	0
x_3	–2	6	1	5/2	–56	0	0	1
	1	–1	0	1/2	–16	0	0	0

五次迭代后的基变量 $\{x_1, x_7, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	–3	0	–5/4	28	1/2	0	0
x_7	0	1/3	0	1/6	–4	–1/6	1	0
x_3	0	0	1	0	0	1	0	1
	0	2	0	7/4	–44	–1/2	0	0

六次迭代后的基变量 $\{x_1, x_2, x_3\}$

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_1	1	0	0	1/4	–8	–1	9	0
x_2	0	1	0	1/2	–12	–1/2	3	0
x_3	0	0	1	0	0	1	0	1
	0	0	0	3/4	–20	1/2	–6	0

在整个迭代过程中，虽然可行基矩阵不断改变，但对应的基本可行解始终是 $X=(0,0,1,0,0,0,0)^T$，没有变化

可以保证不循环的**Bland规则**（**Robert G. Bland，1977**）

选择可进（出）基变量中下标最小的变量进（出）基

按照该规则，在前面求解**Beale**例子的第四次迭代（对应下表），应选择 x_1 进基，选择 x_5 出基，可跳出循环

BV	x_1	x_2	x_3	x_4	x_5	x_6	x_7	RHS
x_6	$-2/3$	1	0	$1/8$	0	1	$-21/2$	0
x_5	$1/16$	$-1/8$	0	$-3/64$	1	0	$3/16$	0
x_3	$3/2$	-1	1	$-1/8$	0	0	$21/2$	1
	2	-3	0	$-1/4$	0	0	3	0

结论：

从任意基本可行解出发，采用 **Bland** 规则进行单纯型迭代，必在有限次迭代后停止于以下两种情况之一：

1）得到一个最优的基本可行解；

2）确定目标函数没有有限的最优值

如何确定初始顶点

基本方法：添加人工变量，通过在迭代过程中把人工变量换出可行基获得原问题的可行基

原问题的约束

$$\sum_{j=1}^{n} a_{ij} x_j = b_i,\ \ i = 1,2,\cdots,m$$

$$x_j \geq 0,\ j = 1,2,\cdots,n$$

添加人工变量后的约束

$$\sum_{j=1}^{n} a_{ij} x_j + x_{n+i} = b_i,\ \ i = 1,2,\cdots,m$$

$$x_j \geq 0,\ j = 1,2,\cdots,n+m$$

把人工变量换出可行基的办法

1）大 *M* 法

把目标函数变为 $\max \sum_{j=1}^{n} c_j x_j + \sum_{i=1}^{m} (-M) x_{n+i}$

其中 ***M*** 是个很大的正数

2）两阶段法

先把目标函数设为 $\max \sum_{i=1}^{m} (-1) x_{n+i}$

迭代到该目标函数等于零时再用原目标函数

1）大 M 法

在一个线性规划问题的约束条件中加进人工变量后，要求人工变量对目标函数取值不受影响；为了让在后面的基变换中，能够将人工变量从基变量中替换出去，当目标函数是求最大时，需要假定人工变量在目标函数中的系数为$-M$（M为任意大的正数)这样目标函数要实现最大化时，必须把人工变量从基变量换出，否则目标函数不可能实现最大化。

因为M是正无穷大数，$-M$是负无穷大，当人工变量是基变量时，其为正，也就是说人工变量乘以（$-M$）会是负无穷大，加到目标函数中，目标函数就不可能实现最大化。另一方面，当目标函数是求最小化时，需要假定人工变量在目标函数中的系数为M。

2）两阶段法

第一阶段：不考虑原问题是否存在基可行解；给原线性规划问题加入人工变量，并构造仅含人工变量的目标函数和要求实现最小化。也就是把目标函数换成仅仅只有人工变量$x_{n+1}, x_{n+2}, \dots, x_{n+m}$并且目标函数是求最小，并且令所有其他的变量的系数都为0。约束条件不变。

目标函数 $\max \omega = -x_{n+1} - \cdots - x_{n+m} + 0x_1 + 0x_2 + \cdots + 0x_n$

$$
\text{约束条件}\begin{cases}
a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_{n+1} = b_1 \\
a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n + x_{n+2} = b_2 \\
\cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots\ \cdots \\
a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n + x_{n+m} = b_m \\
x_1, x_2, \cdots, x_n, x_{n+1}, x_{n+m} \ge 0
\end{cases}
$$

用单纯形法求解上述模型，步骤同前面介绍的一般的单纯形法相同。当获得最优解时，如果ω=0，说明此时基变量中无人工变量，这说明原问题存在基可行解，可以进行第二段计算。否则原问题无可行解，应停止计算。

两阶段法中，新得到的问题一定是有界的，因为目标函数的上确界就是0。

逆矩阵迭代单纯型法

单纯型算法每步迭代的实质工作

前提：已知可行基矩阵 $B=\left(P_{j(1)},P_{j(2)},\cdots,P_{j(m)}\right)$

1）求出非基变量检验数

$$\sigma_{j(k)}=c_{j(k)}-C_B^T B^{-1}P_{j(k)},\ \ m+1\le k\le n$$

2）确定进基变量 $\sigma_{j(t)}=\max\left\{\sigma_{j(m+1)},\sigma_{j(m+2)},\cdots,\sigma_{j(n)}\right\}$

原PPT如此

3）确定出基变量

$$\hat{P}_{j(t)}=B^{-1}P_{j(t)},\ \hat{P}_{n+1}=B^{-1}\vec{b}\ \ \Rightarrow\ \ \frac{\hat{p}_{l\,n+1}}{\hat{p}_{lj(t)}}=\min_{\hat{p}_{ij(t)}>0}\frac{\hat{p}_{i\,n+1}}{\hat{p}_{ij(t)}}$$

结果：得到新的可行基矩阵

$$\tilde{B}=\left(P_{j(1)},\cdots,P_{j(l-1)},P_{j(t)},P_{j(l+1)},\cdots,P_{j(m)}\right)$$

采用单纯型表的迭代实质

前提：已知 $\hat{P}_{n+1} = B^{-1}\vec{b}, \hat{P}_j = B^{-1}P_j, \sigma_j = c_j - C_B^T B^{-1} P_j, 1 \le j \le n$

1） 确定进基变量 $\sigma_{j(t)} = \max\left\{ \sigma_{j(m+1)}, \sigma_{j(m+2)}, \cdots, \sigma_{j(n)} \right\}$

2） 确定出基变量 $$\frac{\hat{p}_{l\,n+1}}{\hat{p}_{lj(t)}} = \min_{\hat{p}_{ij(t)}>0} \frac{\hat{p}_{i\,n+1}}{\hat{p}_{ij(t)}}$$

3） 计算 $\hat{p}'_{lj} = \hat{p}_{lj} / \hat{p}_{lj(t)}$, $\sigma'_j = \sigma_j - \hat{p}'_{lj}\sigma_{j(t)}$, $1 \le j \le n+1$

$$\hat{p}'_{ij} = \hat{p}_{ij} - \hat{p}'_{lj}\hat{p}_{ij(t)} , \ 1 \le j \le n+1, 1 \le i \le m, i \ne l$$

其中 $\hat{P}'_{n+1} = \tilde{B}^{-1}\vec{b}, \hat{P}'_j = \tilde{B}^{-1}P_j, \sigma'_j = c_j - C_{\tilde{B}}^T \tilde{B}^{-1} P_j, 1 \le j \le n$

$$\tilde{B} = \left(P_{j(1)}, \cdots, P_{j(l-1)}, P_{j(t)}, P_{j(l+1)}, \cdots, P_{j(m)} \right)$$

结果：得到下次迭代的前提条件

采用单纯型表迭代的优缺点

优点：没有显式计算 B^{-1}

缺点：每次迭代只需要

$$\sigma_{j(k)}, m+1 \le k \le n,\ \hat{P}_{j(t)} = B^{-1}P_{j(t)},\ \hat{P}_{n+1} = B^{-1}\vec{b}$$

计算所有的 $\hat{P}_j, 1 \le j \le n$ 只是为下次迭代做准备

某些数据可能始终用不上

特别不利的情况：n 远远大于 m 的情况

实现单纯型算法的另外一种迭代方法：逆矩阵迭代

前提：已知 $B^{-1}=\left(P_{j(1)},P_{j(2)},\cdots,P_{j(m)}\right)^{-1}$

1）计算（对偶变量）$\hat{Y}_B^T=C_B^T B^{-1}$

2）计算检验数 $\sigma_{j(k)}=c_{j(k)}-C_B^T B^{-1}P_{j(k)}=c_{j(k)}-\hat{Y}_B^T P_{j(k)},\ \forall k$

3）确定进基变量 $\sigma_{j(t)}=\max\left\{\sigma_{j(m+1)},\sigma_{j(m+2)},\cdots,\sigma_{j(n)}\right\}$

4）确定出基变量

$$\hat{P}_{j(t)}=B^{-1}P_{j(t)},\ \hat{P}_{n+1}=B^{-1}\vec{b}\quad\Rightarrow\quad\frac{\hat{p}_{l\,n+1}}{\hat{p}_{lj(t)}}=\min_{\hat{p}_{ij(t)}>0}\frac{\hat{p}_{i\,n+1}}{\hat{p}_{ij(t)}}$$

5）计算 $\tilde{B}^{-1}=\left(P_{j(1)},\cdots,P_{j(l-1)},P_{j(t)},P_{j(l+1)},\cdots,P_{j(m)}\right)^{-1}$

结果：得到下次迭代的前提条件

核心问题：如何利用 B^{-1} 用很少的计算量确定 $\tilde{B}^{-1}$ ？

记 $\vec{e}_i = (0,\cdots,0,1,0,\cdots,0)^T$ 其中 1 在第 i 个位置

因为 $B^{-1}\left(P_{j(1)}, P_{j(2)}, \cdots, P_{j(m)}\right) = \left(\vec{e}_1, \vec{e}_2, \cdots, \vec{e}_m\right)$

$$B^{-1}P_{j(t)} = \hat{P}_{j(t)}$$

所以 $B^{-1}\tilde{B} = B^{-1}\left(P_{j(1)}, \cdots, P_{j(l-1)}, P_{j(t)}, P_{j(l+1)}, \cdots, P_{j(m)}\right)$

$$= \left(\vec{e}_1, \cdots, \vec{e}_{l-1}, \hat{P}_{j(t)}, \vec{e}_{l+1}, \cdots, \vec{e}_m\right)$$

若能确定 $\Gamma = \left(\vec{e}_1, \cdots, \vec{e}_{l-1}, \hat{P}_{j(t)}, \vec{e}_{l+1}, \cdots, \vec{e}_m\right)^{-1}$

就能得到 $\tilde{B}^{-1} = \Gamma B^{-1}$

求逆矩阵本质都是在解线性方程组，不过招式可能有不同

因为 $\Gamma\left(\vec{e}_1,\cdots,\vec{e}_{l-1},\hat{P}_{j(t)},\vec{e}_{l+1},\cdots,\vec{e}_m\right)=\left(\vec{e}_1,\vec{e}_2,\cdots,\vec{e}_m\right)$

所以 $\Gamma\vec{e}_i=\vec{e}_i,\ \forall 1\le i\le m, i\ne l,\quad \Gamma\hat{P}_{j(t)}=\vec{e}_l$

记 $\Gamma=\left(\vec{\gamma}_1,\vec{\gamma}_2,\cdots,\vec{\gamma}_m\right)$

可得 $\vec{\gamma}_i=\vec{e}_i,\ \forall 1\le i\le m, i\ne l$

$$\sum_{\substack{i=1\\ i\ne l}}^{m}\hat{p}_{ij(t)}\vec{e}_i+\hat{p}_{lj(t)}\vec{\gamma}_l=\vec{e}_l$$

于是 $$\vec{\gamma}_l=-\frac{1}{\hat{p}_{lj(t)}}\left(\sum_{\substack{i=1\\ i\ne k}}^{m}\hat{p}_{ij(t)}\vec{e}_i-\vec{e}_l\right)=-\frac{1}{\hat{p}_{lj(t)}}\begin{pmatrix}\hat{p}_{1j(t)}\\ \vdots\\ \hat{p}_{l-1j(t)}\\ -1\\ \hat{p}_{l+1j(t)}\\ \vdots\\ \hat{p}_{mj(t)}\end{pmatrix}$$

从而 $\tilde{B}^{-1}=\left(\vec{e}_1,\cdots,\vec{e}_{l-1},\vec{\gamma}_l,\vec{e}_{l+1},\cdots,\vec{e}_m\right)B^{-1}$

小结：逆矩阵迭代单纯型法的计算过程

$$B^{-1}$$

$$\downarrow$$

$$\hat{Y}_B^T = C_B^T B^{-1},\ \sigma_k = c_k - \hat{Y}_B^T P_k,\ \forall k$$

$$\downarrow$$

$$\sigma_{j(t)} = \max\left\{\sigma_{j(m+1)}, \sigma_{j(m+2)}, \cdots, \sigma_{j(n)}\right\}$$

$$\downarrow$$

$$\hat{P}_{n+1} = B^{-1}\vec{b},\ \hat{P}_{j(t)} = B^{-1}P_{j(t)},\quad \frac{\hat{p}_{l\,n+1}}{\hat{p}_{lj(t)}} = \min_{\hat{p}_{ij(t)}>0} \frac{\hat{p}_{i\,n+1}}{\hat{p}_{ij(t)}},\ \vec{\gamma}_l = -\frac{1}{\hat{p}_{lj(t)}}\begin{pmatrix} \hat{p}_{1j(t)} \\ \vdots \\ \hat{p}_{l-1\,j(t)} \\ -1 \\ \hat{p}_{l+1\,j(t)} \\ \vdots \\ \hat{p}_{mj(t)} \end{pmatrix}$$

$$\downarrow$$

$$\widetilde{B}^{-1} = \left(\vec{e}_1, \cdots, \vec{e}_{l-1}, \vec{\gamma}_l, \vec{e}_{l+1}, \cdots, \vec{e}_m\right)B^{-1}$$

能否从最终单纯型表中找出 B^{-1}？

初始单纯型表

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_3	0	5	1	0	0	15
x_4	6	2	0	1	0	24
x_5	1	1	0	0	1	5
	2	1	0	0	0	z

$$(P_2, P_3, P_1)^{-1} \times$$

$$\Downarrow$$

最终单纯型表

BV	x_1	x_2	x_3	x_4	x_5	RHS
x_2	0	1	0	-0.25	1.5	1.5
x_3	0	0	1	1.25	-7.5	7.5
x_1	1	0	0	0.25	-0.5	3.5
	0	0	0	-0.25	-0.5	$z-8.5$

$$(P_2, P_3, P_1)^{-1} = \begin{pmatrix} 0 & -0.25 & 1.5 \\ 1 & 1.25 & -7.5 \\ 0 & 0.25 & -0.5 \end{pmatrix} \Leftarrow$$

一般情况：若在标准型初始矩阵 $A=(P_1,P_2,\cdots,P_n)$ 中存在单位阵，比如

$$(P_{\tau(1)},P_{\tau(2)},\cdots,P_{\tau(m)})=\begin{pmatrix}1 & & \\ & \ddots & \\ & & 1\end{pmatrix}$$

由于单纯型表的各列向量等于 $\hat{P}_j=B^{-1}P_j$， $\forall 1\le j\le n$

所以 $$B^{-1}=B^{-1}(P_{\tau(1)},P_{\tau(2)},\cdots,P_{\tau(m)})$$
$$=(B^{-1}P_{\tau(1)},B^{-1}P_{\tau(2)},\cdots,B^{-1}P_{\tau(m)})=(\hat{P}_{\tau(1)},\hat{P}_{\tau(2)},\cdots,\hat{P}_{\tau(m)})$$

只要将最初的单位向量所在列的最终数据按单位向量在单位矩阵中的位置排好就可以得到 B^{-1}

注意：B 中各 P_i 所在列数取决于基变量所在行数！

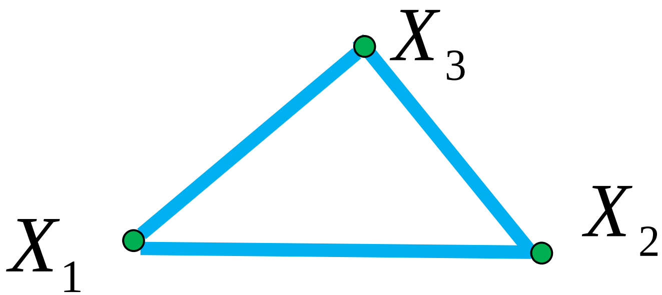
为什么叫<u>单纯型</u>算法？

单纯形定义

一维

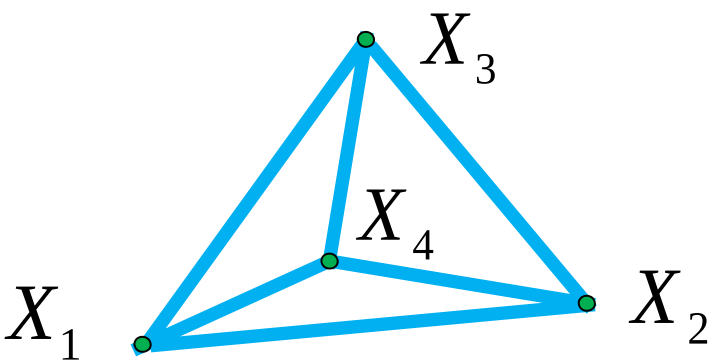
$$X=\theta_1 X_1+\theta_2 X_2$$
$$\theta_1+\theta_2=1, \theta_i\ge 0$$

二维



$$X=\theta_1 X_1+\theta_2 X_2+\theta_3 X_3$$
$$\theta_1+\theta_2+\theta_3=1, \theta_i\ge 0$$

三维



$$X=\theta_1 X_1+\theta_2 X_2+\theta_3 X_3+\theta_4 X_4$$
$$\theta_1+\theta_2+\theta_3+\theta_4=1,\ \ \theta_i\ge 0$$

n维

$$X=\sum_{i=1}^{n+1}\theta_i X_i, \sum_{i=1}^{n+1}\theta_i=1,\ \ \theta_i\ge 0, \forall$$

单纯形定义

We call the convex hull of any set of $n+1$ points in R^n which do not lie on a hyperplane a simplex.

The convex hull of a finite number of points is the set of all convex combinations of these points.

The convex hull of a set Ω is defined to be the intersection of all convex sets that contain Ω.

单纯形算法名称的由来

"*The term simplex method arose out of a discussion with T. Motzkin who felt that the approach that I was using, when viewed in the geometry of the columns, was best described as a movement from one simplex to a neighboring one.*"

- G. B. Dantzig, "Linear Programming," *Operations Research*, vol. 50, no. 1, pp. 42-47, 2002.

单纯形算法名称的由来

What exactly Motzkin had in mind is anyone's guess, but the interpretation provided by this lecture video of Prof. Craig Tovey (credit to Samarth) is noteworthy. In it, he explains that any finitely bounded problem,

$$\begin{gathered}\min c^T x \\ Ax = b, \\ 0 \leq x \leq u,\end{gathered}$$

can be scaled to $e^T u = 1$ without loss of generality. Then by rewritting all upper bound constraints to equations, $x_j + r_j = u_j$ for slack variables $r_j \geq 0$, we have that the sum of all variables (original and slack) equals $e^T u$ equals one. Hence, all finitely bounded problems can be cast to a formulation of the form

https://or.stackexchange.com/questions/7831/why-is-it-called-the-simplex-algorithm-method

单纯形算法名称的由来

can be scaled to $e^T u = 1$ without loss of generality. Then by rewritting all upper bound constraints to equations, $x_j + r_j = u_j$ for slack variables $r_j \geq 0$, we have that the sum of all variables (original and slack) equals $e^T u$ equals one. Hence, all finitely bounded problems can be cast to a formulation of the form

$$\begin{gathered} \min c^T x \\ Ax = b, \\ e^T x = 1, \\ x \geq 0, \end{gathered}$$

where the feasible set is simply described as the set of convex combinations of columns in A that equal b. If $A \in \mathbb{R}^{m \times n}$, then a feasible basic solution corresponds to finding $m+1$ columns of A such that the *simplex* generated by these $m+1$ columns contains b. In the simplex method we thus start from a simplex containing b and iterate by moving to a neighboring simplex containing b. This sounds a lot like what Motzkin was talking about.

https://or.stackexchange.com/questions/7831/why-is-it-called-the-simplex-algorithm-method