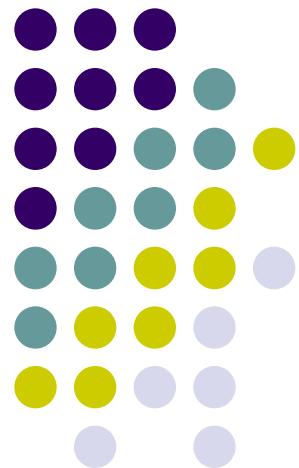
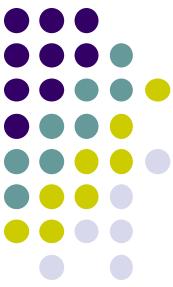


最优化方法 (II)

Hongxin Zhang

2020-04-07





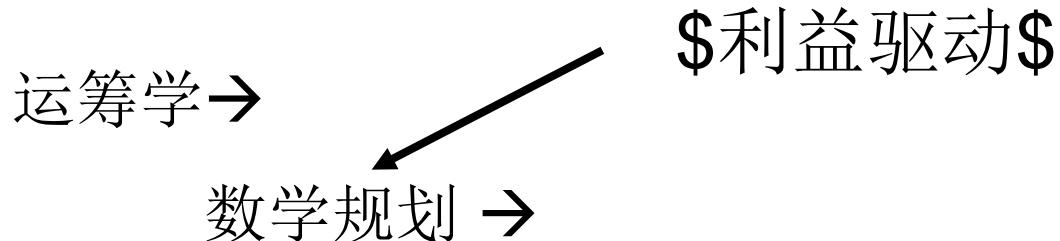
内容

- 线性规划
- 非线性优化
- 主要参考书：
 - 《线性规划》
 - 张建中, 许绍吉, 科学出版社
 - 《最优化理论与方法》
 - 袁亚湘, 孙文瑜, 科学出版社
 - 《数学规划》
 - 黄红选, 韩继业, 清华大学出版社



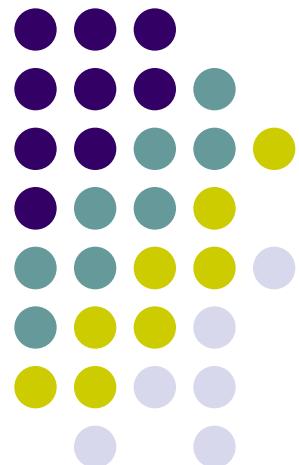
一、线性规划

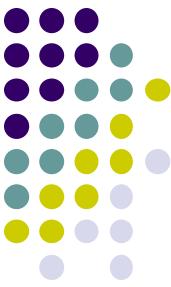
- 在人们的生产实践中，经常会遇到如何利用现有资源来安排生产，以取得最大经济效益的问题。



- 1947年G. B. Dantzig 求解线性规划的单纯形方法
- 现有的计算机能处理成千上万个约束条件和决策变量的线性规划问题之后，线性规划的适用领域更为广泛了。

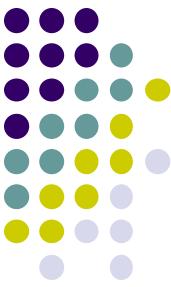
1. 基本理论





1.1 问题定义 (linear programming, LP)

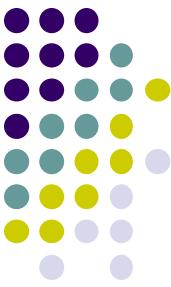
- 某计算机公司提供大型计算机系统，有A, B, C三个技术团队
 - 高端系统
 - 每套售后的利润400万元
 - 需用A、B两个技术团队协作实施
 - A团队2个月，B团队1个月；
 - 中端系统
 - 每套售后的利润300万元
 - 需用A, B, C三个技术团队协作实施
 - 每个团队各需1个月
- 2019年度，A队可工作10个月，B队8个月、C队7个月
- 问：怎样的销售策略能使总利润最大？



1.1 问题定义 (linear programming, LP)

- 设可销售 x_1 个高端系统， x_2 个中端系统
- 目标函数

$$\begin{aligned} \max \quad & z = 4x_1 + 3x_2 \\ \bullet \text{ 约束条件} \quad & \begin{cases} 2x_1 + x_2 \leq 10 \\ x_1 + x_2 \leq 8 \\ x_2 \leq 7 \\ x_1, x_2 \geq 0 \end{cases} \end{aligned}$$



1.1 问题定义 (linear programming, LP)

- 在一组线性的等式或不等式约束下，求一个线性函数的最小值或最大值。
- 形式化的定义：

$$\min c_1x_1 + \dots + c_nx_n$$

s. t.

$$a_{11}x_1 + \dots + a_{1n}x_n \leq b_1$$

...

$$a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m$$

$$x_1, \dots, x_n \geq 0$$



1.1 问题定义 (linear programming, LP)

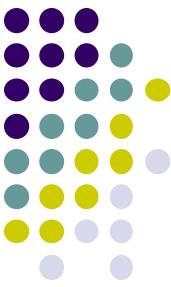
- 在一组线性的等式或不等式约束下，求一个线性函数的最小值或最大值。
- 形式化的定义：

$$\begin{aligned} & \min \mathbf{c}^T \mathbf{x} \\ & \text{s.t. } \mathbf{A}\mathbf{x} \leq \mathbf{b} \\ & \quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$



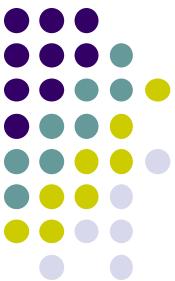
目标函数
objective function

c: 价值向量, cost vector
A: 约束矩阵, constraint matrix
b: 右端向量, right-hand-side vector



1.1 问题定义 (linear programming, LP)

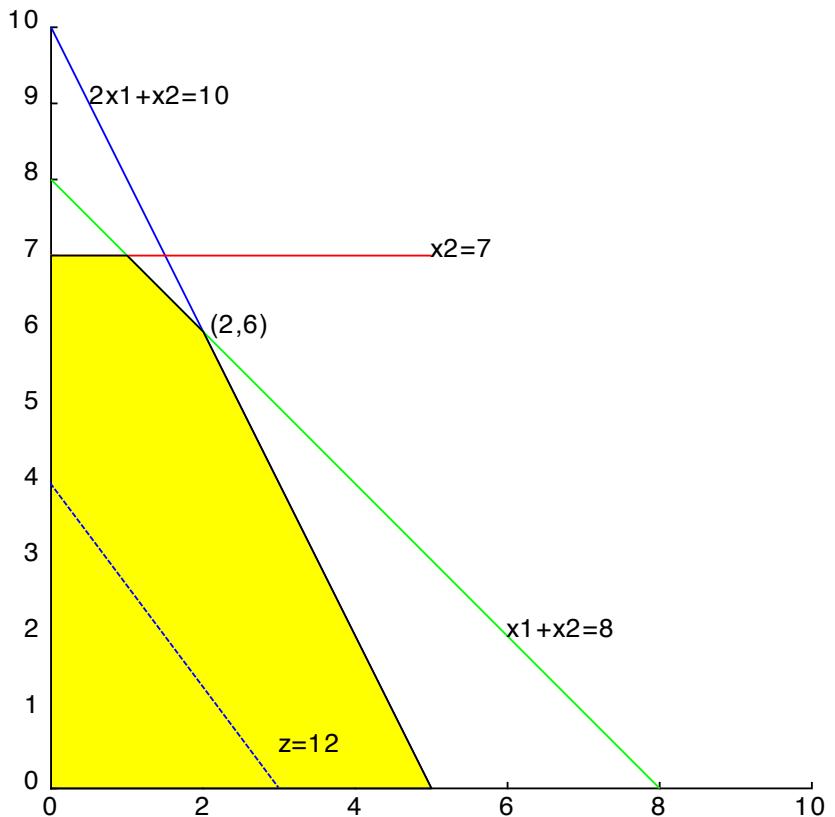
- X : 满足约束条件, 称为**可行解**或**可行点**
 - (feasible point)
- D : 所有可行点的集合, 称为**可行区域**
 - (feasible region)
- LP问题:
 - $D = \emptyset$, 无解或不可行
 - $D \neq \emptyset$, 但目标函数在 D 上无上界: 无界
 - $D \neq \emptyset$, 且目标函数有限的最优解: 有最优解

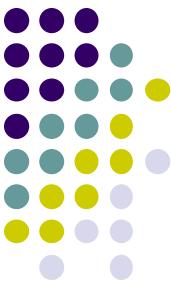


线性规划的图解法

$$\max \quad z = 4x_1 + 3x_2$$

$$\begin{cases} 2x_1 + x_2 \leq 10 \\ x_1 + x_2 \leq 8 \\ x_2 \leq 7 \\ x_1, x_2 \geq 0 \end{cases}$$





1.2 标准LP问题

- 从实际中总结出来的LP形式不完全一样：
 - 目标函数是最大值或最小值，
 - 约束条件是等式约束或不等式约束，
 - 变量有上界或下界或无界

$$\min z = \mathbf{c}^T \mathbf{x}$$

$$\text{s.t. } \mathbf{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$



标准形式的LP问题

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

- 标准化

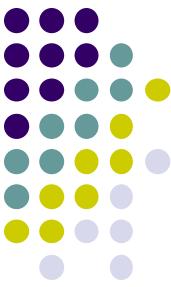
- 目标函数的转换 $\max z \rightarrow \min (-z)$
- 约束条件的转换 (引入松弛变量)

$$\sum_{j=1}^n a_{ij} x_j \leq b_i \Leftrightarrow \begin{cases} \sum_{j=1}^n a_{ij} x_j + x_{n+i} = b_i \\ x_{n+i} \geq 0 \end{cases}$$

- 变量的非负约束

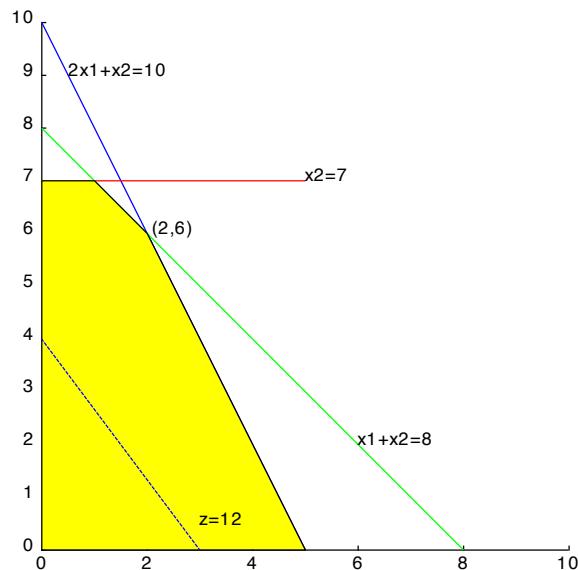
$$x_j \geq l_j \Leftrightarrow y_j \geq 0, y_j = x_j - l_j$$

$$x_j \text{自由变量} \Leftrightarrow u_j \geq 0, v_j \geq 0, x_j = u_j - v_j$$



1.3 可行区域D

- 下面讨论可行区域 $D=\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ 的结构
 - 首先讨论集合 $K=\{\mathbf{x} \mid \mathbf{Ax} = \mathbf{b}\}$





1.3 基本概念

- 仿射集:

对于集合 $S \subseteq E^n$, 如果对任意 $\mathbf{x}, \mathbf{y} \in S, \lambda \in E^1$, 都有

$$\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in S,$$

则称 S 为仿射集

- 凸集:

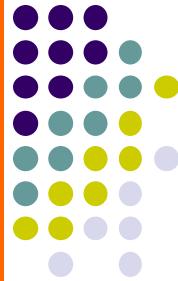
对任意的 $\mathbf{x}, \mathbf{y} \in C, \lambda \in (0,1)$, 有

$$\lambda\mathbf{x} + (1-\lambda)\mathbf{y} \in C,$$

则称 C 为凸集

=> 极点, 极方向

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



1.4 基本可行解

假定 $\text{rank}(A)=m$, 则A中必有m个线性无关的列向量——构成满秩方阵B, A中其余各列组成子阵N, 即 $A=(B, N)$ 。相应的 $\mathbf{x}=(\mathbf{x}_B, \mathbf{x}_N)$, 则 $A\mathbf{x}=\mathbf{b}$ 可改写成

$$A\mathbf{x} = B\mathbf{x}_B + N\mathbf{x}_N = \mathbf{b}.$$

因为B为满秩, 则 B^{-1} 存在, 那么

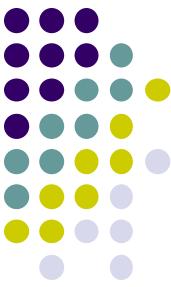
$$\mathbf{x}_B = B^{-1}\mathbf{b} - B^{-1}N\mathbf{x}_N$$

任给一组值 \mathbf{x}_N , 则可得到相应的 \mathbf{x}_B ,

$$(\mathbf{x}_B, \mathbf{x}_N)$$

为 $A\mathbf{x}=\mathbf{b}$ 的一个解。

令 $\mathbf{x}_N=0$, 则 $\mathbf{x}_B=B^{-1}\mathbf{b}$, $\mathbf{x}=(B^{-1}\mathbf{b}, 0)$ 称为约束方程组的**基本解**。



1.4 基本可行解

- 设 $\text{rank}(A)=m$
 - A 中 m 个线性无关的列向量构成满秩方阵 B
 - A 中其余各列组成子阵 N
 - $A=(B, N)$
- $x=(x_B, x_N)$
 - $Ax=b$ 可改写成 $Bx_B + Nx_N = b$

B 称为**基**， B 中的 m 个线性无关的列向量称为**基向量**， x_B 的 m 个分量称为**基变量**， 其余的变量称为**非基变量**。

基本解不一定满足非负条件，故不一定是可行解。对于非负的基本解，称为**基本可行解**，此时 B 称为**可行基**。



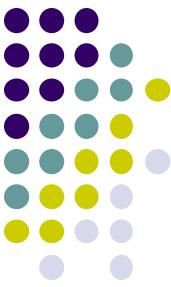
1.4 基本可行解

- 定理：
 - 可行解 \mathbf{x} 是基本可行解 $\Leftrightarrow \mathbf{x}$ 的正分量所对应的列向量线性无关
 - 可行解 \mathbf{x} 是基本可行解 $\Leftrightarrow \mathbf{x}$ 为 D 的极点



可行区域的极方向

- 极方向的代数性质： $D=\{\mathbf{x} \mid A\mathbf{x}=\mathbf{b}, \mathbf{x} \geq \mathbf{0}\}$ 的方向 \mathbf{d} 有 k 个非零分量，则 \mathbf{d} 为 D 的极方向的充要条件是 \mathbf{d} 的非零分量对应的列向量组秩为 $k-1$
- 极方向的几何性质： \mathbf{d} 为 D 的极方向的充要条件是 \mathbf{d} 为 D 的某个半直线界面方向
- 如果 D 有极方向，显然 D 为无界集；反之若 D 为无界集，则 D 有方向，且有极方向



1.5 线性规划基本定理

基本定理1（表示定理）

定理：设 $D = \{\mathbf{x} \mid A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0\}$ 的所有极点为 $\mathbf{x}^1, \dots, \mathbf{x}^k$, 所有极方向为 $\mathbf{d}^1, \dots, \mathbf{d}^l$, 则 $x \in D$ 的充要条件是存在一组 $\lambda_i (i = 1, \dots, k)$ 和 $\mu_j (j = 1, \dots, l)$ 使得

$$\mathbf{x} = \sum_{i=1}^k \lambda_i \mathbf{x}^i + \sum_{j=1}^l \mu_j \mathbf{d}^j,$$

$$\lambda_i \geq 0, i = 1, \dots, k; \mu_j \geq 0, j = 1, \dots, l$$

$$\sum_{i=1}^k \lambda_i = 1$$



基本定理2

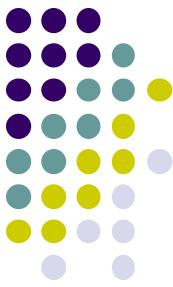
给定LP问题

$$\min z = \mathbf{c}^T \mathbf{x}$$

$$\text{s.t. } A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

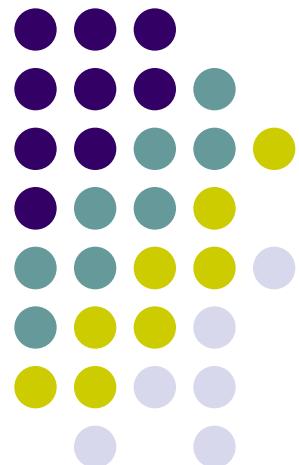
- (1) 若存在有限最优值（即有最优解），则最优值必在可行区域 D 的某个极点上达到。
- (2) 目标函数存在有限最优值的充要条件是对 D 的所有极方向 \mathbf{d}^j ，均有 $\mathbf{c}\mathbf{d}^j \geq 0$

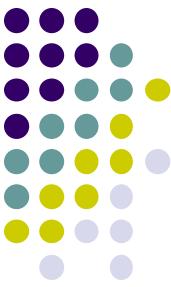


定理的注释与理解

- 定理说明：
 - 标准LP问题的目标函数的最优值必可在某个**基本可行解**处达到。
 - 求解标准LP问题，只需要在基本可行解的集合中进行搜索。
- 求出并比较所有基本可行解的方法通常不可行。且原因在于当 n 较大时，基本可行解个数将很大。
- 通常的方法根据一定的规则在基本可行解的一个子集中计算搜索，这就是单纯形法。

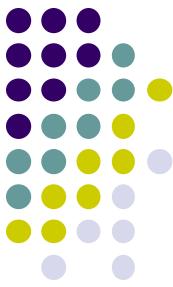
2. 单纯形方法 (Simplex Method)





2.1 基本单纯形方法

- 主要思想：先找一个基本可行解，判断它是否为最优解。若不是，就找一个更好的基本可行解，在进行检验。如此迭代，直至最后找到最优解，或判定无界。
- 两个主要问题：
 - 寻找初始解
 - 如何判别和迭代（先考虑）



确定初始的基本可行解

确定初始的基本可行解等价于确定初始的可行基，一旦初始的可行基确定了，那么对应的初始基本可行解也就唯一确定

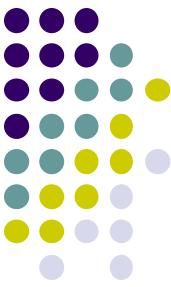
为了讨论方便，不妨假设在标准型线性规划中，系数矩阵A中前m个系数列向量恰好构成一个可行基，即

$A = (B \ N)$ ，其中

$B = (P_1, P_2, \dots, P_m)$ 为基变量 x_1, x_2, \dots, x_m 的系数列向量
构成的可行基，

$N = (P_{m+1}, P_{m+2}, \dots, P_n)$ 为非基变量 $x_{m+1}, x_{m+2}, \dots, x_n$ 的
系数列向量构成的矩阵。

$$\max Z = CX, \quad D = \{X \in R^n / AX=b, X \geq 0\}$$



$$\max Z = CX, \quad D = \{X \in R^n / AX=b, X \geq 0\}$$

所以约束方程 $AX=b$ 就可以表示为

$$AX = (BN) \begin{pmatrix} X_B \\ X_N \end{pmatrix} = BX_B + NX_N = b$$

用可行基 B 的逆阵 B^{-1} 左乘等式两端，再通过移项可推得：

$$X_B = B^{-1}b - B^{-1}NX_N$$

若令所有非基变量 $X_N = 0$ ，则基变量 $X_B = B^{-1}b$

由此可得初始的基本解 $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$

$$AX=b \rightarrow BX_B + NX_N = b \rightarrow X_B = B^{-1}b - B^{-1}NX_N \rightarrow X_N = 0, X_B = B^{-1}b$$

- 问题：

- 要判断m个系数列向量是否恰好构成一个基并不是一件容易的事。

基由系数矩阵A中m个线性无关的系数列向量构成。

但是要判断m个系数列向量是否线性无关并非易事。

- 即使系数矩阵A中找到了一个基B，也不能保证该基恰好是可行基。

因为不能保证基变量 $X_B = B^{-1}b \geq 0$ 。

- 为了求得基本可行解 $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ ，必须求基 B 的逆阵 B^{-1} 。

但是求逆阵 B^{-1} 也是一件麻烦的事。

- 结论：

- 在线性规划标准化过程中设法得到一个m阶单位矩阵I作为初始可行基 B，



为了设法得到一个m阶单位矩阵I作为初始可行基 B , 可在线性规划标准化过程中作如下处理 :

- 若在化标准形式前 , m个约束方程都是 “≤” 的形式 ,
那么在化标准形时只需在一个约束不等式左端都加上一个松弛变量 x_{n+i} ($i=1, 2, \dots, m$)。
- 若在化标准形式前 , 约束方程中有 “≥” 不等式 ,
那么在化标准形时除了在方程式左端减去剩余变量使不等式变成等式以外 , 还必须在左端再加上一个非负新变量 , 称为人工变量 .
- 若在化标准形式前 , 约束方程中有等式方程 , 那么可以直接在等式左端添加人工变量。



$$\max Z = CX, \quad D = \{X \in R^n / AX=b, X \geq 0\}$$

■判断现行的基本可行解是否最优

假如已求得一个**基本可行解**

$$X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$$

将这一基本可行解代入目标函数，可求得相应的目标函数值

$$Z = CX = (C_B \ C_N) \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix} = C_B B^{-1}b$$

其中 $C_B = (c_1, c_2, \dots, c_m)$, $C_N = (c_{m+1}, c_{m+2}, \dots, c_n)$ 分别表示基变量和非基变量所对应的价值系数子向量。



$$\max Z = CX, \quad D = \{X \in R^n / AX = b, X \geq 0\}$$

要判定 $Z = C_B B^{-1} b$ 是否已经达到最大值，只需将 $X_B = B^{-1} b - B^{-1} N X_N$ 代入目标函数，使目标函数用非基变量表示，即：

$$Z = CX = (C_B C_N) \begin{pmatrix} X_B \\ X_N \end{pmatrix}$$

$$= C_B X_B + C_N X_N = C_B (B^{-1} b - B^{-1} N X_N) + C_N X_N$$

$$= C_B B^{-1} b + (C_N - C_B B^{-1} N) X_N$$

$$\triangleq C_B B^{-1} b + \sigma_N X_N = C_B B^{-1} b + (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n) \begin{pmatrix} X_{m+1} \\ X_{m+2} \\ \vdots \\ X_n \end{pmatrix}$$

其中 $\sigma_N = C_N - C_B B^{-1} N = (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n)$ 称为非基变量 X_N 的检验向量，它的各个分量称为检验数。若 σ_N 的每一个检验数均小于等于 0，即 $\sigma_N \leq 0$ ，那么现在的基本可行解就是最优解。



定理1：最优解判别定理

对于线性规划问题

$$\max Z = CX, \quad D = \{X \in R^n / AX=b, X \geq 0\}$$

若某个基本可行解所对应的检验向量 $\sigma_N = C_N - C_B B^{-1} N \leq 0$,

则这个基本可行解就是最优解。

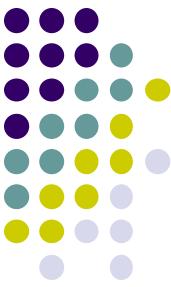
$$Z = C_B B^{-1} b + (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n) \begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$

定理2：无穷多最优解判别定理

若 $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ 是一个基本可行解，所对应的检验向量

$\sigma_N = C_N - C_B B^{-1} N \leq 0$, 其中存在一个检验数 $\sigma_{m+k} = 0$,

则线性规划问题有无穷多最优解。



■ 基本可行解的改进

如果现行的基本可行解 X 不是最优解，即在检验向量

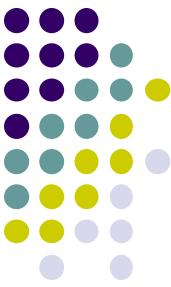
$\sigma_N = C_N - C_B B^{-1} N$ 中存在正的检验数，则需在原基本可行解 X 的基础上寻找一个新的基本可行解，并使目标函数值有所改善。具体做法是：

- 先从检验数为正的非基变量中确定一个换入变量，使它从非基变量变成基变量（将它的值从零增至正值），
- 再从原来的基变量中确定一个换出变量，使它从基变量变成非基变量（将它的值从正值减至零）。

由此可得一个新的基本可行解，由 $Z = C_B B^{-1} b + (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n)$

可知，这样的变换一定能使目标函数值有所增加。

$$\begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$



换入变量和换出变量的确定：

●换入变量的确定— **最大增加原则**

假设检验向量 $\sigma_N = C_N - C_B B^{-1} N = (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n)$

若其中有两个以上的检验数为正，那么为了使目标函数值增加得快些，通常要用“**最大增加原则**”，即选取最大正检验数所对应的非基变量为换入变量，即若

$$\max \left\{ \sigma_j / \sigma_j > 0, m+1 \leq j \leq n \right\} = \sigma_{m+k}$$

则选取对应的 x_{m+k} 为换入变量，

由于 $\sigma_{m+k} > 0$ 且为最大，

因此当 x_{m+k} 由零增至正值，

可使目标函数值 $Z = C_B B^{-1} b + (\sigma_{m+1}, \sigma_{m+2}, \dots, \sigma_n)$

最大限度的增加。

$$\begin{pmatrix} x_{m+1} \\ x_{m+2} \\ \vdots \\ x_n \end{pmatrix}$$



- 换出变量的确定— **最小比值原则**

如果确定 x_{m+k} 为换入变量，方程

$$X_B = B^{-1}b - B^{-1}N X_N \Rightarrow X_B = B^{-1}b - B^{-1}P_{m+k} x_{m+k}$$

其中 P_{m+k} 为 A 中与 x_{m+k} 对应的系数列向量。

现在需在 $X_B = (x_1, x_2, \dots, x_m)^T$ 中确定一个基变量为换出变量。

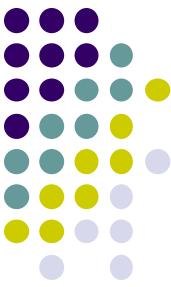
当 x_{m+k} 由零慢慢增加到某个值时 X_B 的非负性可能被打破。

为保持解的可行性，可以按最小比值原则确定换出变量：

若

$$\min \left\{ \frac{(B^{-1}b)_i}{(B^{-1}P_{m+k})_i} / (B^{-1}P_{m+k})_i > 0, 1 \leq i \leq m \right\} = \frac{(B^{-1}b)_l}{(B^{-1}P_{m+k})_l}$$

则选取对应的基变量 x_l 为换出变量。



定理3：无最优解判别定理

若 $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$ 是一个基本可行解，有一个检验数 $\sigma_{m+k} > 0$ ，

但是 $B^{-1}P_{m+k} \leq 0$ ，则该线性规划问题无最优解。

证：令 $x_{m+k} = \lambda, (\lambda > 0)$ ，则得新的可行解

将上式代入 $X_B = B^{-1}b - B^{-1}P_{m+k}x_{m+k} = B^{-1}b - B^{-1}P_{m+k}\lambda$

$$Z = C_B B^{-1}b + (\sigma_{m+1}, \dots, \sigma_{m+k}, \dots, \sigma_n) \begin{pmatrix} x_{m+1} \\ \vdots \\ \lambda \\ \vdots \\ x_n \end{pmatrix} = C_B B^{-1}b + \sigma_{m+k} \lambda$$

因为 $\sigma_{m+k} > 0$ ，故当 $\lambda \rightarrow +\infty$ 时， $Z \rightarrow +\infty$



■ 用初等变换求改进了的基本可行解

假设 B 是线性规划 $\max Z = CX, AX = b, X \geq 0$ 的可行基，则

$$AX = b \Rightarrow (BN) \begin{pmatrix} X_B \\ X_N \end{pmatrix} = b \Rightarrow (I, B^{-1}N) \begin{pmatrix} X_B \\ X_N \end{pmatrix} = B^{-1}b$$

令非基变量 $X_N = 0$ ，则基变量 $X_B = B^{-1}b$

可得基本可行解 $X = \begin{pmatrix} B^{-1}b \\ 0 \end{pmatrix}$

- 用逆阵 B^{-1} 左乘约束方程组的两端，等价于对方程组施以一系列的初等“行变换”。变换的结果是将系数矩阵 A 中的可行基 B 变换成单位矩阵 I ，把非基变量系数列向量构成的矩阵 N 变换成 $B^{-1}N$ ，把资源向量 b 变换成 $B^{-1}b$



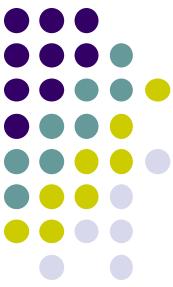
由于行初等变换后的方程组 $(I, B^{-1}N) \begin{pmatrix} X_B \\ X_N \end{pmatrix} = B^{-1}b$

与原约束方程组 $AX=b$ 或 $(B, N) \begin{pmatrix} X_B \\ X_N \end{pmatrix} = b$ 同解

且改进了的基本可行解 X' 只是在 X 的基变量的基础上用一个换入变量替代其中一个换出变量，其它的基变量仍然保持不变。这些基变量的系数列向量是单位矩阵 I 中的单位向量。为了求得改进的基本可行解 X' ，只需对增广矩阵

$$(I, B^{-1}N, B^{-1}b)$$

施行初等行变换，将换入变量的系数列向量转换成换出变量所对应的单位向量即可。



判别和迭代

- 假定 $\text{rank}(A) = m < n$, 且假定已找到一个非退化的基本可行解, 即找到一个基 B 。则 $A\mathbf{x} = \mathbf{b}$ 可写成

$$\mathbf{x}_B + B^{-1}N\mathbf{x}_N = B^{-1}\mathbf{b} \quad (1)$$

记

$$B = (a_1, \dots, a_m)$$

$$\bar{\mathbf{a}}_j = B^{-1}\mathbf{a}_j = (\bar{a}_{1j}, \dots, \bar{a}_{mj})^T, j = 1, \dots, n$$

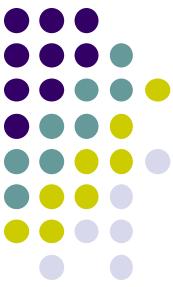
$$\bar{\mathbf{b}} = B^{-1}\mathbf{b} = (\bar{b}_1, \dots, \bar{b}_m)^T$$

$$\bar{N} = B^{-1}N$$

则上式(1)可写成 $\mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}}$ (2)

显然, 不同的基对应的方程表达式也不同。

式(2)称为基 B 的典则方程组, 简称典式。



判别和迭代

$$\mathbf{x}_B + \bar{N}\mathbf{x}_N = \bar{\mathbf{b}} \quad (2-2)$$

显然，不同的基对应的方程表达式也不同。

式(2)称为基 B 的典则方程组，简称典式。

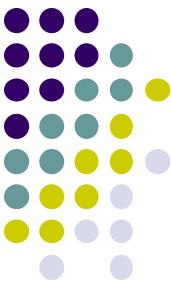
若 $\bar{\mathbf{b}} \geq 0$, 则式(2)就对应着基本可行解 $\bar{\mathbf{x}} = \begin{pmatrix} \bar{\mathbf{b}} \\ 0 \end{pmatrix}$.

将目标函数也做相应的变换，

$$z = \mathbf{c}\mathbf{x} = \mathbf{c}_B\bar{\mathbf{b}} - (\mathbf{c}_B\bar{N} - \mathbf{c}_N)\mathbf{x}_N = \mathbf{c}_B\bar{\mathbf{b}} - \sum_{j=m+1}^n (\mathbf{c}_B\bar{\mathbf{a}}_j - c_j)\mathbf{x}_j \quad (3)$$

用 z_0 表示在基本可行解 $\bar{\mathbf{x}}$ 处的目标函数值，

则有 $z_0 = \mathbf{c}\bar{\mathbf{x}} = \mathbf{c}_B\bar{\mathbf{b}}$ (即上式的常数项)



判别和迭代

因为 $\bar{\mathbf{a}}_1, \dots, \bar{\mathbf{a}}_m$ 是线性无关向量，

故当 $j = 1, \dots, m$ 时

$$\mathbf{c}_B \bar{\mathbf{a}}_j - c_j = 0.$$

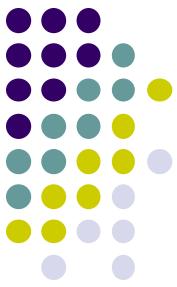
引入记号

$$\zeta_j = \mathbf{c}_B \bar{\mathbf{a}}_j - c_j, j = 1, \dots, n$$

或用向量形式

$$\zeta = \mathbf{c}_B B^{-1} A - \mathbf{c} = (\zeta_B, \zeta_N) = (\mathbf{0}, \mathbf{c}_B B^{-1} N - \mathbf{c}_N)$$

则式(3)改写成 $z = \mathbf{c}_B \bar{\mathbf{b}} - \zeta \mathbf{x}$.



判别和迭代

$$z = \mathbf{c}_B \bar{\mathbf{b}} - \boldsymbol{\zeta} \mathbf{x}, \quad \boldsymbol{\zeta} = \mathbf{c}_B B^{-1} A - \mathbf{c} = (\zeta_B, \zeta_N) = (\mathbf{0}, \mathbf{c}_B B^{-1} N - \mathbf{c}_N)$$

经过变量变换后，LP问题可叙述成

$$\min z = z_0 - \boldsymbol{\zeta} \mathbf{x} \tag{4}$$

$$s.t. \quad \mathbf{x}_B + B^{-1} N \mathbf{x}_N = \bar{\mathbf{b}}, \tag{5}$$

$$\mathbf{x} \geq \mathbf{0}.$$

定理2.1：若 (4) 中 $\boldsymbol{\zeta} \leq 0$ ，则 $\bar{\mathbf{x}}$ 为最优解。

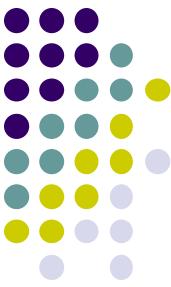
定理2.2：若 (4) 中 $\boldsymbol{\zeta}$ 存在某一分量 $\zeta_k > 0$ ，且 $\bar{\mathbf{a}}_k \leq \mathbf{0}$ ，
则原问题无解。

定理2.3：若 (4) 中有 $\zeta_k > 0$ ，且 $\bar{\mathbf{a}}_k$ 至少存在一个正分量，
则能找到基本可行解 $\hat{\mathbf{x}}$ ，使得 $\mathbf{c}\hat{\mathbf{x}} < \mathbf{c}\bar{\mathbf{x}}$ 。



单纯形法的计算步骤

1. 找初始的可行基;
2. 求出对应的典式;
3. 求 $\zeta_k = \max \{\zeta_j \mid j = 1, \dots, n\}$ 称之为检验向量。
在迭代过程中，如果 ζ 有不止一个正分量，则为了使目标函数下降得较快，一般取最大的分量 ζ_k 所对应的列向量 \mathbf{a}_k 进入基。
4. 若 $\zeta_k \leq 0$, 停止。已找到最优解 $\mathbf{x} = \begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix}$ 及最优值 $z = \mathbf{c}_B \bar{\mathbf{b}}$
5. 若 $\bar{\mathbf{a}}_k \leq \mathbf{0}$, 停止, 原问题无界。
6. 求最小比 $\frac{\bar{b}_r}{\bar{a}_{rk}} = \min \left\{ \frac{\bar{b}_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0 \right\}$
7. 以 \mathbf{a}_k 代替 \mathbf{a}_{B_r} 得到新的基, 回到2.



例1 $\max Z = 5x_1 + 2x_2 + 3x_3 - x_4 + x_5$

$$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 8 \\ 3x_1 + 4x_2 + x_3 + x_5 = 7 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases} \quad C = (5, 2, 3, -1, 1)$$

解：

$$A = \begin{pmatrix} 1 & 2 & 2 & 1 & 0 \\ 3 & 4 & 1 & 0 & 1 \end{pmatrix} \quad b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

(1) 确定初始的基本可行解

$$B = (P_4 P_5) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ 基变量 } X_4, X_5, \text{ 非基变量 } X_1, X_2, X_3$$

$$X_B = \begin{pmatrix} x_4 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, C_B = (-1, 1), C_N = (5, 2, 3), b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

$$X_N = 0 \rightarrow X_B = B^{-1}b = \begin{pmatrix} 8 \\ 7 \end{pmatrix} \Rightarrow X = (0, 0, 0, 8, 7)^T$$

$$Z = C_B B^{-1}b = (-1, 1) \begin{pmatrix} 8 \\ 7 \end{pmatrix} = -1$$

$$X_B = \begin{pmatrix} X_4 \\ X_5 \end{pmatrix}, X_N = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, C_B = (-1, 1), C_N = (5, 2, 3), b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

(2) 检验 $X=(0,0,0,8,7)^T$ 是否最优

$$\begin{aligned} \text{检验向量 } \sigma_N &= C_N - C_B B^{-1} N = (5, 2, 3) - (-1, 1) \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix} \\ &= (5, 2, 3) - (2, 2, -1) = (3, 0, 4) \\ &\quad \uparrow \quad \uparrow \quad \uparrow \\ &\quad \sigma_1 \quad \sigma_2 \quad \sigma_3 \end{aligned}$$

因为 $\sigma_1 = 3, \sigma_3 = 4$ 均大于零 ,

所以 $X=(0,0,0,8,7)^T$ 不是最优解

$$X_B = \begin{pmatrix} x_4 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, C_B = (-1, 1), C_N = (5, 2, 3), b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

$$\sigma_N = (\sigma_1, \sigma_2, \sigma_3) = (3, 0, 4)$$

(3) 基本可行解 $X=(0,0,0,8,7)^T$ 的改进

① 选取换入变量

因为 $\max\{3, 4\}=4$, 取 x_3 为换入变量。

② 选取换出变量

$$B^{-1}b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}, B^{-1}P_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \leq 0 \quad \text{且} \quad \min \left\{ \frac{8}{2}, \frac{7}{1} \right\} = \frac{8}{2}$$

选取 x_4 为换出变量。

$$\begin{pmatrix} x_4 \\ x_5 \end{pmatrix} = B^{-1}b - B^{-1}P_3x_3 = \begin{pmatrix} 8 \\ 7 \end{pmatrix} - \begin{pmatrix} 2 \\ 1 \end{pmatrix}x_3$$

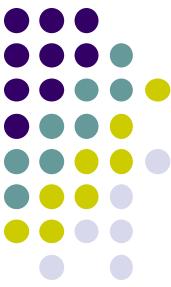
$$X_B = \begin{pmatrix} x_4 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} 1 & 2 & 2 \\ 3 & 4 & 1 \end{pmatrix}, C_B = (-1, 1), C_N = (5, 2, 3), b = \begin{pmatrix} 8 \\ 7 \end{pmatrix}$$

(4) 求改进了的基本可行解 X'

对约束方程组的增广矩阵施以初等行变换，使换入变量 x_3 所对应的系数列向量 $P_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$ 变换成换出变量 x_4 所对应的单位向量 $P_4 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ 。
注意保持基变量 x_5 的系数列向量 $P_5 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ 为单位向量不变

$$\left(\begin{array}{cc|cc|cc} 1 & 2 & 2 & 1 & 0 & 8 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{array} \right) \xrightarrow{\text{第一行除以 } 2} \left(\begin{array}{cc|cc|cc} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{array} \right)$$

$$\xrightarrow{\text{第二行减去第一行}} \left(\begin{array}{cc|cc|cc} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & -\frac{1}{2} & 1 & 3 \end{array} \right)$$



$$C = (5, 2, 3, -1, 1)$$

$$C = (5, 2, 3, -1, 1)$$

$$\left(\begin{array}{ccccc|c} 1 & 2 & 2 & 1 & 0 & 8 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{array} \right) \Rightarrow \left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & -\frac{1}{2} & 1 & 3 \end{array} \right)$$

可得改进的基本可行解。

$$B = (P_3 P_5) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ 基变量 } X_3, X_5 \quad \text{非基变量 } X_1, X_2, X_4$$

$$X_B = \begin{pmatrix} X_3 \\ X_5 \end{pmatrix}, X_N = \begin{pmatrix} X_1 \\ X_2 \\ X_4 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & -\frac{1}{2} \end{pmatrix}, C_B = (3, 1), C_N = (5, 2, -1), b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

$$X_N = 0 \rightarrow X_B = B^{-1}b = \begin{pmatrix} 4 \\ 3 \end{pmatrix} \Rightarrow \text{基本可行解}$$

$$\text{目标函数值 } Z = C_B B^{-1}b = (3, 1) \begin{pmatrix} 4 \\ 3 \end{pmatrix} = 15$$

易见目标函数值比原来的Z=-1增加了，再转向步骤(2)

$$X_B = \begin{pmatrix} X_3 \\ X_5 \end{pmatrix}, X_N = \begin{pmatrix} X_1 \\ X_2 \\ X_4 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & -\frac{1}{2} \end{pmatrix}, C_B = (3, 1), C_N = (5, 2, -1), b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

(2) 检验 $X=(0,0,4,0,3)^T$ 是否最优。

$$\text{检验向量 } \sigma_N = C_N - C_B B^{-1} N = (5, 2, -1) - (3, 1) \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & -\frac{1}{2} \end{pmatrix}$$

$$= (5, 2, -1) - (4, 6, 1) = (1, -4, -2)$$

↑ ↑ ↑

因为 $\sigma_1 = 1 > 0$, $\sigma_1 \sigma_2 \sigma_4$

所以 $X=(0,0,4,0,3)^T$ 仍不是最优解。

$$X_B = \begin{pmatrix} x_3 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \\ x_4 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & -\frac{1}{2} \end{pmatrix}, C_B = (3, 1), C_N = (5, 2, -1), b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

(3) 基本可行解 $X = (0, 0, 4, 0, 3)^T$ 的改进

① 选取换入变量

因为 $\sigma_1 = 1 > 0$, 取 x_1 为换入变量

② 选取换出变量

$$B^{-1}b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}, B^{-1}P_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{5}{2} \end{pmatrix} \leq 0 \text{ 且 } \min \left\{ \frac{4}{1/2}, \frac{3}{5/2} \right\} = \frac{3}{5/2}$$

选取 x_5 为换出变量

$$\begin{pmatrix} x_3 \\ x_5 \end{pmatrix} = B^{-1}b - B^{-1}P_1x_1 = \begin{pmatrix} 4 \\ 3 \end{pmatrix} - \begin{pmatrix} 1/2 \\ 5/2 \end{pmatrix}x_1$$

$$X_B = \begin{pmatrix} x_3 \\ x_4 \\ x_5 \end{pmatrix}, X_N = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{1}{2} & 1 & \frac{1}{2} \\ \frac{5}{2} & 3 & -\frac{1}{2} \end{pmatrix}, C_B = (3, 1), C_N = (5, 2, -1), b = \begin{pmatrix} 4 \\ 3 \end{pmatrix}$$

(4) 求改进了的基本可行解 X''

对约束方程组的增广矩阵施以初等行变换，使换入变量 x_1 所对应

的系数列向量 $P_1 = \begin{pmatrix} \frac{1}{2} \\ \frac{2}{5} \\ \frac{5}{2} \end{pmatrix}$ 变换成换出变量 x_5 所对应的单位向量 $P_5 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

$$\left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & -\frac{1}{2} & 1 & 3 \end{array} \right) \xrightarrow{\text{第二行乘以 } 2/5} \left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ 1 & \frac{6}{5} & 0 & -\frac{1}{5} & \frac{2}{5} & \frac{6}{5} \end{array} \right)$$

$$\xrightarrow{\text{第一行减以第二行的 } 1/2 \text{ 倍}} \left(\begin{array}{ccccc|c} 0 & \frac{2}{5} & 1 & \frac{3}{5} & -\frac{1}{5} & \frac{17}{5} \\ 1 & \frac{6}{5} & 0 & -\frac{1}{5} & \frac{2}{5} & \frac{6}{5} \end{array} \right)$$



$$C = (5, 2, 3, -1, 1)$$

$$C = (5, 2, 3, -1, 1)$$

$$\left(\begin{array}{ccccc|c} \frac{1}{2} & 1 & 1 & \frac{1}{2} & 0 & 4 \\ \frac{5}{2} & 3 & 0 & -\frac{1}{2} & 1 & 3 \end{array} \right) \Rightarrow \left(\begin{array}{ccccc|c} 0 & \frac{2}{5} & 1 & \frac{3}{5} & -\frac{1}{5} & \frac{17}{5} \\ 1 & \frac{6}{5} & 0 & -\frac{1}{5} & \frac{2}{5} & \frac{6}{5} \end{array} \right)$$

可得改进的基本可行解。

$$B = (P_3 P_1) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \text{ 基变量 } x_3, x_1, \text{ 非基变量 } x_2, x_4, x_5$$

$$X_B = \begin{pmatrix} x_3 \\ x_1 \end{pmatrix}, X_N = \begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{2}{5} & \frac{3}{5} & -\frac{1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \end{pmatrix}, C_B = (3, 5), C_N = (2, -1, 1), b = \begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix}$$

$$X_N = 0 \rightarrow X_B = B^{-1}b = \begin{pmatrix} \frac{17}{5} \\ \frac{5}{5} \\ \frac{6}{5} \\ \frac{5}{5} \end{pmatrix} \Rightarrow \text{基本可行解 } X = \left(\frac{6}{5}, 0, \frac{17}{5}, 0, 0 \right)^T$$

目标函数值 $Z = C_B B^{-1}b = (3, 5) \begin{pmatrix} \frac{17}{5} \\ \frac{5}{5} \\ \frac{6}{5} \\ \frac{5}{5} \end{pmatrix} = \frac{81}{5}$ **比Z=15增加了，再转向步骤(2)**

$$X_B = \begin{pmatrix} x_3 \\ x_2 \\ x_1 \end{pmatrix}, X_N = \begin{pmatrix} x_2 \\ x_4 \\ x_5 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, N = \begin{pmatrix} \frac{2}{5} & \frac{3}{5} & \frac{-1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}, C_B = (3, 5), C_N = (2, -1, 1), b = \begin{pmatrix} \frac{17}{5} \\ \frac{6}{5} \end{pmatrix}$$

(2) 检验 $X'' = (\frac{6}{5}, 0, \frac{17}{5}, 0, 0)^T$ 是否最优。

检验向量

$$\sigma_N = C_N - C_B B^{-1} N = (2, -1, 1) - (3, 5) \begin{pmatrix} \frac{2}{5} & \frac{3}{5} & \frac{-1}{5} \\ \frac{6}{5} & \frac{-1}{5} & \frac{2}{5} \\ \frac{1}{5} & \frac{1}{5} & \frac{1}{5} \end{pmatrix}$$

$$= (2, -1, 1) - (\frac{36}{5}, \frac{4}{5}, \frac{7}{5}) = (\frac{-26}{5}, \frac{-9}{5}, \frac{-2}{5})$$

↑ ↑ ↑

因为所有检验数均小于零，

$\sigma_2 \ \sigma_4 \ \sigma_5$

所以 $X^* = X'' = (\frac{6}{5}, 0, \frac{17}{5}, 0, 0)^T$ 是最优解， $Z^* = \frac{81}{5}$



2.2 单纯形表

根据 θ 和 $\hat{\mathbf{x}}$ 的构造方式可知，得到改进的基本可行解——即使原来的非基变量 x_k 变成取正值的基本变量，同时使原来的基本变量 x_r 值为零（变成非基变量）。

等价地，基 $B = [\mathbf{a}_1, \dots, \mathbf{a}_m]$ 变成另一个基 $\hat{B} = [\mathbf{a}_1, \dots, \mathbf{a}_r, \mathbf{a}_k, \mathbf{a}_{r+1}, \dots, \mathbf{a}_m]$

*换基变换

* x_k 称为进基变量， x_r 为离基变量。

在基 B 下，典式的系数增广矩阵为

$$\begin{array}{cccccccccc} x_1 & \dots & x_r & \dots & x_m & x_{m+1} & \dots & x_k & \dots & x_n & \\ 1 & \dots & 0 & \dots & 0 & \bar{a}_{1,m+1} & \dots & \bar{a}_{1,k} & \dots & \bar{a}_{1,n} & \bar{b}_1 \\ \vdots & \ddots & \vdots & & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \dots & 1 & \dots & 0 & \bar{a}_{r,m+1} & \dots & \bar{a}_{r,k} & \dots & \bar{a}_{r,n} & \bar{b}_r \\ \vdots & & \ddots & \vdots & \vdots & \vdots & & \vdots & & \vdots & \vdots \\ 0 & \dots & 0 & \dots & 1 & \bar{a}_{m,m+1} & \dots & \bar{a}_{m,k} & \dots & \bar{a}_{m,n} & \bar{b}_m \end{array}$$

x_r 与 x_k 的地位互换。

第 k 列通过初等变换变成单位向量 \mathbf{e}_r 。

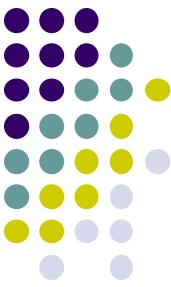
1) 第 r 行 $\bar{\mathbf{a}}_r^T$ 除以 \bar{a}_{rk} , $\hat{\mathbf{a}}_r^T = \bar{\mathbf{a}}_r^T / \bar{a}_{rk}$

2) 对第 i 行 ($i \neq r$) $\bar{\mathbf{a}}_i^T$, $\hat{\mathbf{a}}_r^T = \bar{\mathbf{a}}_i^T - \hat{\mathbf{a}}_r^T \bar{a}_{ik}$

即第 i 行减去新的第 r 行 $\hat{\mathbf{a}}_r^T$ 乘以 \bar{a}_{ik}

最后一列正是基 \hat{B} 所对应的基本可行解。

$$\hat{b}_r = \bar{b}_r / \bar{a}_{rk}, \hat{b}_i = \bar{b}_i - (\bar{b}_r / \bar{a}_{rk}) \bar{a}_{ik}, i \neq r$$



换基时，目标函数 $z = \mathbf{c}_B \bar{\mathbf{b}} - (\mathbf{c}_B \bar{N} - \mathbf{c}_N) \mathbf{x}_N$ 也要做相应的调整，即将 B 换成 \hat{B} 。
将上式的等价形式

$$z + (\mathbf{c}_B B^{-1} N - \mathbf{c}_N) \mathbf{x}_N = \mathbf{c}_B \bar{\mathbf{b}}, \text{ 或 } z + \zeta_N \mathbf{x}_N = z_0$$

看成一个方程，放在典式中一同进行初等变换。

z	x_1	\cdots	x_r	\cdots	x_m	x_{m+1}	\cdots	x_k	\cdots	x_n	RHS	
z	1	0	\cdots	0	\cdots	0	ζ_{m+1}	\cdots	ζ_k	\cdots	ζ_n	z_0
x_1	0	1	\cdots	0	\cdots	0	$\bar{a}_{1,m+1}$	\cdots	$\bar{a}_{1,k}$	\cdots	$\bar{a}_{1,n}$	\bar{b}_1
\vdots	\vdots	\vdots	\ddots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
x_r	0	0	\cdots	1	\cdots	0	$\bar{a}_{r,m+1}$	\cdots	$\bar{a}_{r,k}$	\cdots	$\bar{a}_{r,n}$	\bar{b}_r
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots	\vdots						
x_m	0	0	\cdots	0	\cdots	1	$\bar{a}_{m,m+1}$	\cdots	$\bar{a}_{m,k}$	\cdots	$\bar{a}_{m,n}$	\bar{b}_m

单纯形表

转轴列

转轴行

转轴元

单纯形表可简记为

z	\mathbf{x}_B	\mathbf{x}_N	RHS
z	1	$\mathbf{0}$	ζ_N
\mathbf{x}_B	$\mathbf{0}$	I	\bar{N}

上述变换又称之为旋转 (pivot)，类似于解线性方程组的主元消去法。



口表格单纯形法

通过例 1 我们发现，在单纯形法的求解过程中，有下列重要指标：

- 每一个基本可行解的检验向量 $\sigma_N = C_N - C_B B^{-1} N$

根据检验向量可以确定所求得的基本可行解是否为最优解。如果不是最优又可以通过检验向量确定合适的换入变量。

- 每一个基本可行解所对应的目标函数值 $Z = C_B B^{-1} b$

通过目标函数值可以观察单纯形法的每次迭代是否能使目标函数值有效地增加，直至求得最优目标函数为止。

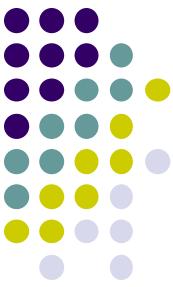
➤ 在单纯形法求解过程中，每一个基本可行解 X 都以某个经过初等行变换的约束方程组中的单位矩阵 I 为可行基。

当 $B = I$ 时， $B^{-1} = I$ ，易知： $\sigma_N = C_N - C_B N$ $Z = C_B b$

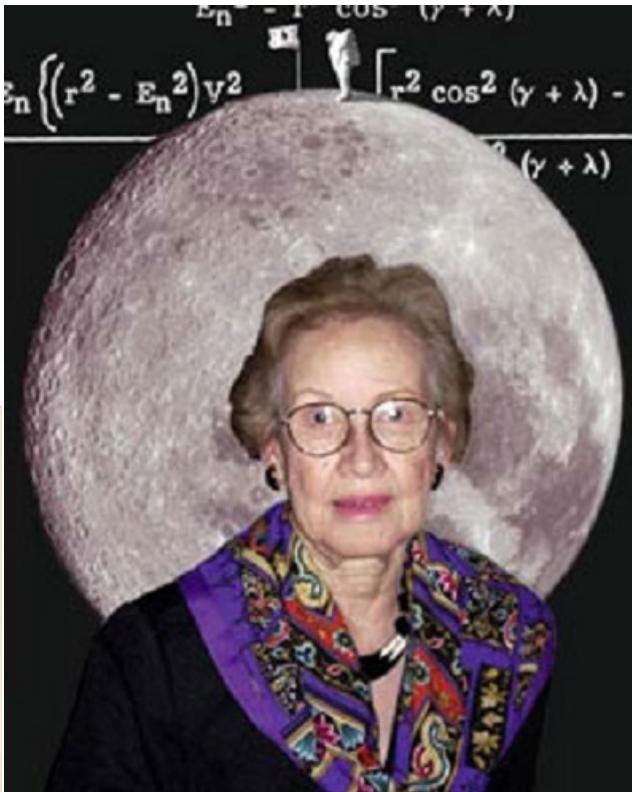


可将这些重要结论的计算设计成如下一个简单的表格，
即单纯形表来完成：

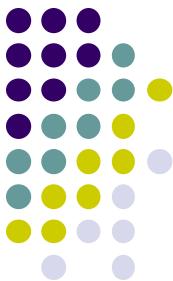
C			C _B	C _N	θ
C _B	X _B	b	X ₁ X ₂ … X _m	X _{m+1} X _{m+2} … X _n	
C ₁	X ₁	b ₁			θ ₁
C ₂	X ₂	b ₂	I	N	θ ₂
.	.	.			.
.	.	.			.
C _m	X _m	b _m			θ _m
Z		C _B b	0	C _N - C _B N	



美国影片《隐藏人物》



手算飞船轨道的非裔天才女数学家，她是**NASA**背后的隐藏英雄



美国影片《隐藏人物》



“穿裙子的计算机”们

1961年5月5日，水星计划的“自由7号”将美国第一位宇航员艾伦·谢菲德送上太空，这艘飞船的运行轨迹也正是凯瑟琳计算的



例2、试利用单纯形表求例1中的最优解解：

$$\max Z = 5x_1 + 2x_2 + 3x_3 - x_4 + x_5$$

$$C = (5, 2, 3, -1, 1)$$

$$\begin{cases} x_1 + 2x_2 + 2x_3 + x_4 = 8 \\ 3x_1 + 4x_2 + x_3 + x_5 = 7 \\ x_1, x_2, x_3, x_4, x_5 \geq 0 \end{cases}$$

$$(Ab) = \begin{pmatrix} 1 & 2 & 2 & 1 & 0 & 8 \\ 3 & 4 & 1 & 0 & 1 & 7 \end{pmatrix}$$

得初始的单纯形表：

C			5	2	3	-1	1	Θ
C_B	x_B	b	x_1	x_2	x_3	x_4	x_5	
-1	x ₄	8	1	2	2	1	0	
1	x ₅	7	3	4	1	0	1	
Z			-1	3	0	4	0	0

$$Z = C_B b$$

$$\sigma_N = C_N - C_B N$$

$$\text{初始基本可行解 } X = (0, 0, 0, 8, 7)^T, Z = -1,$$

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
-1	x_4	8	1	2	2	1	0	8/2
1	x_5	7	3	4	1	0	1	7/1
Z		-1	3	0	4	0	0	

x_3 换入变量 , x_4 换出变量 , 2为主元进行旋转变换

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	4	1/2	1	1	1/2	0	
1	x_5	3	5/2	3	0	-1/2	1	
Z		15	1	-4	0	-2	0	

基本可行解 $X=(0,0,4,0,3)^T$, $Z= 15$,

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	4	1/2	1	1	1/2	0	4/1/2
1	x_5	3	5/2	3	0	-1/2	1	3/5/2
Z			15	1	-4	0	-2	0

x_1 换入变量 , x_5 换出变量 , $5/2$ 为主元进行旋转变换

C			5	2	3	-1	1	Θ
C_B	X_B	b	x_1	x_2	x_3	x_4	x_5	
3	x_3	17/5	0	2/5	1	3/5	-1/5	
5	x_1	6/5	1	6/5	0	-1/5	2/5	
Z			81/5	0	-26/5	0	-9/5	-2/5

$$\sigma_N = C_N - C_B N \leq 0 \text{ 最优解 } X^* = \left(\frac{6}{5}, 0, \frac{17}{5}, 0, 0 \right)^T \text{ 最优值 } Z^* = \frac{81}{5}$$



2.3 初始解:两阶段法

设原问题为

$$\min \mathbf{c}\mathbf{x}$$

$$s.t. \quad A\mathbf{x} = \mathbf{b} \quad (\mathbf{b} \geq \mathbf{0})$$

$$\mathbf{x} \geq \mathbf{0}$$

引入 m 个人工变量 $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, 考虑如下辅助问题

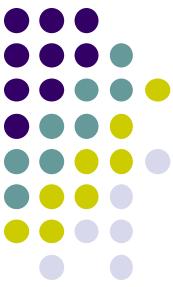
$$\min g = \mathbf{1}\mathbf{x}_a$$

$$s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$\mathbf{x}, \mathbf{x}_a \geq \mathbf{0}$$

其中 $\mathbf{1} = (1, \dots, 1)$ 。设原问题的可行域为 D , 辅助问题的可行域为 D' ,

显然 $\mathbf{x} \in D \Leftrightarrow \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \in D'$ 。而 $\begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix} \in D' \Leftrightarrow \min g = 0$.



2.3 初始解:两阶段法

m 个人工变量 $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, 考虑如下辅助问题

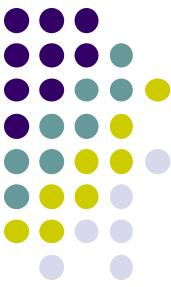
$$\begin{aligned} \min g &= \mathbf{1}\mathbf{x}_a \\ s.t. \quad A\mathbf{x} + \mathbf{x}_a &= \mathbf{b} \\ \mathbf{x}, \mathbf{x}_a &\geq 0 \end{aligned}$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

对于辅助问题来说, 一个基本可行解是 $\begin{pmatrix} \mathbf{x} \\ \mathbf{x}_a \end{pmatrix} = \begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix}$,

于是可进行单纯形迭代求解, 其结果有两种可能:

- 1) $\min g > 0$ 。说明 $D = \emptyset$.
- 2) $\min g = 0$ 。这是自然有 $\mathbf{x}_a = 0$, 把它除去后即得到原问题的一个可行解。



2.3 初始解:两阶段法

m 个人工变量 $\mathbf{x}_a = (x_{n+1}, \dots, x_{n+m})^T$, 得到辅助问题

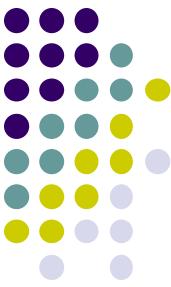
$$\begin{aligned} \min g &= \mathbf{1}\mathbf{x}_a \\ s.t. \quad A\mathbf{x} + \mathbf{x}_a &= \mathbf{b} \\ \mathbf{x}, \mathbf{x}_a &\geq 0 \end{aligned}$$

$$\begin{pmatrix} \mathbf{x} \\ 0 \end{pmatrix} \in D' \Leftrightarrow \min g = 0.$$

- * 若此时所有人工变量都是非基变量，则解为基本可行解。
- * 否则，进行旋转变换消去基变量中的人工变量。

设基变量 x_r 为人工变量，

- ▷ 取第 r 行的前 n 个元素中不为零的元素 \bar{a}_{rs} 为转轴元进行变换，
则非人工变量 x_s 进入基变量，同时从基本量消除了人工变量 x_r 。
- ▷ 若第 r 行的前 n 个元素都为零，则直接去掉该行以及对应的人工变量。



2.4 初始解:大M法

设原问题为

$$\min \mathbf{c}\mathbf{x}$$

$$s.t. \quad A\mathbf{x} = \mathbf{b} \quad (\mathbf{b} \geq \mathbf{0})$$

$$\mathbf{x} \geq \mathbf{0}$$

引入 $\mathbf{x}_a \in E^m$ 及足够大的正数 M , 得到新问题如下

$$\min \mathbf{c}\mathbf{x} + M\mathbf{1}\mathbf{x}_a$$

$$s.t. \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$\mathbf{x}, \mathbf{x}_a \geq \mathbf{0}$$

其中 $\mathbf{1} = (1, \dots, 1)$ 。只要 M 取得足够大, 那么可以说

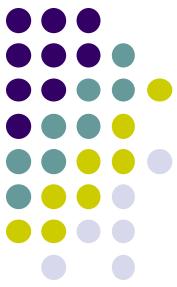
x 是原问题的最优解 $\Leftrightarrow \begin{pmatrix} \mathbf{x} \\ \mathbf{0} \end{pmatrix}$ 是新问题的最优解。

而新问题有初始解 $\begin{pmatrix} \mathbf{0} \\ \mathbf{b} \end{pmatrix}$, 因此总可以用单纯形法求解。



2.5 退化与防止循环

- 若单纯形表中右端向量存在分量为0，即出现退化解的情况下，可能出现循环。为了避免循环，需要再补充一些旋转规则。
 - 字典序方法
 - Bland规则



字典序方法

非零向量 \mathbf{x} , 并且第一个非零分量是非负的, 称为按字典序非负, 记为 $\mathbf{x} \succ = \mathbf{0}$ 。
对向量 \mathbf{x}, \mathbf{y} , 若 $\mathbf{x} - \mathbf{y} \succ = \mathbf{0}$, 则称 \mathbf{x} 按字典序大于等于 \mathbf{y} .

若一组 \mathbf{x}^i , 有 \mathbf{x}^t , 使得对所有 \mathbf{x}^i , 均有 $\mathbf{x}^i \succ = \mathbf{x}^t$, 则称 \mathbf{x}^t 为该组向量中按字典序最小的。
记为: $\mathbf{x}^t = \text{lex min } \mathbf{x}^i$.

字典序法就是在选定了进基变量 x_k 后, 令

$$\frac{\mathbf{p}_r}{\bar{a}_{rk}} = \text{lex min} \left\{ \frac{\mathbf{p}_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0, i = 1, \dots, m \right\}, \text{其中 } \mathbf{p}_i \text{ 为 } (\bar{\mathbf{b}}, B^{-1}) \text{ 的第 } i \text{ 行。}$$

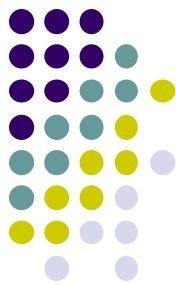


Bland规则

- 1) 选下标最小的正检验数 ζ_k 所对应的非基变量 x_k 作为进基变量。
- 2) 离基变量 x_l 的确定：如果同时有几个 $\frac{b_r}{\bar{a}_{rk}}$ 达到最小，就选其中下标最小的那个基变量作为离基变量。即

$$l = \min \left\{ r \left| \frac{b_r}{\bar{a}_{rk}} = \min \left\{ \frac{b_i}{\bar{a}_{ik}} \mid \bar{a}_{ik} > 0 \right\} \right. \right\}$$

- Bland规则的优点是简单易行，但缺点是只考虑最小下标，而不考虑目标函数值下降的快慢。故其效率比字典序或原来单纯形法低。
- 在实际中，退化是常见的，但退化不一定产生循环。事实上，产生循环是较罕见的。



2.6 修改单纯形法

- 当LP问题的规模很大时，如何减少存储量和计算时间，就是一个必须加以考虑的问题。在实际中，大多采用修改单纯形法。



逆矩阵法

一般单纯形表为

$$\left[\begin{array}{ccccc} z & \mathbf{x}_B & \mathbf{x}_N & & RHS \\ z & 1 & \mathbf{0} & \mathbf{c}_B B^{-1} N - \mathbf{c}_N & \mathbf{c}_B B^{-1} \mathbf{b} \\ \mathbf{x}_B & \mathbf{0} & I & B^{-1} N & B^{-1} \mathbf{b} \end{array} \right]$$

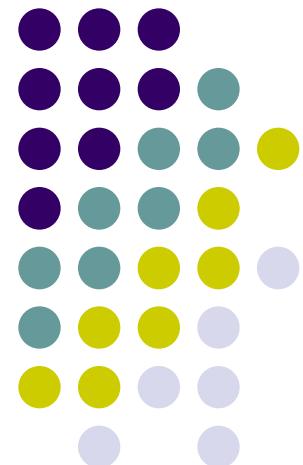
逆矩阵法对每一张单纯形表，仅存贮下列数据

$$\left[\begin{array}{cc} \mathbf{w} & z_o \\ B^{-1} & \bar{\mathbf{b}} \end{array} \right] = \left[\begin{array}{cc} \mathbf{c}_B B^{-1} & \mathbf{c}_B B^{-1} \mathbf{b} \\ B^{-1} & B^{-1} \mathbf{b} \end{array} \right], \text{ 该表称为修改单纯形表。}$$

根据这张表和原始数据进行迭代计算。由 $\zeta_k = \max\{\mathbf{w}\mathbf{a}_j - c_j \mid x_j \text{ 非基变量}\}$ 确定 x_k 为进基变量
再求 $\bar{\mathbf{a}}_k = B^{-1} \mathbf{a}_k$ ，用最小比确定离基变量 x_r ，得到新的基 \hat{B} ，最后构造对应于 \hat{B} 的修改单纯形表。
如果每一次迭代都必须从构造 B^{-1} ，则计算量显然难以令人满意。

定理：对应基 B 的修改单纯形表的右边添加一列 $\begin{pmatrix} \zeta_k \\ \bar{\mathbf{a}}_k \end{pmatrix}$ ，以 $\bar{\mathbf{a}}_{rk}$ 为转轴元旋转，就得到了对应于基 \hat{B} 的修改单纯形表。

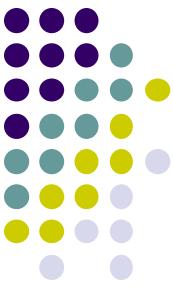
3. 最优性条件 和对偶理论





对偶性

- 对于每一个线性规划 P ，总存在着另一个线性规划 D ，两者之间存在着密切的联系，人们可以通过求解对偶规划 D 来获得原规划 P 的最优解。



3.1 Karush-Kuhn-Tucker条件

- 定理

设 A 为 $m \times n$ 阶矩阵, $\mathbf{b} \in E^m$, $\mathbf{c} \in E^n$, $\mathbf{x} \in E^n$, 则 \mathbf{x}^* 为LP问题

$$\min \mathbf{c}^T \mathbf{x}$$

$$s.t. \quad A\mathbf{x} \geq \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

的最优解的充要条件是存在 $\mathbf{w} \in E^m$, $\mathbf{v} \in E^n$, 使得下列KKT条件得到满足

$$A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0}$$

$$\mathbf{c} - \mathbf{w}^T A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0}$$

$$\mathbf{w}^T (A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^T \mathbf{x}^* = 0$$

若将 \mathbf{v} 消去 (利用 $\mathbf{c} - \mathbf{w}^T A - \mathbf{v} = \mathbf{0}$) , 可得KTT条件的另一种形式

$$A\mathbf{x}^* \geq \mathbf{b}, \quad \mathbf{x}^* \geq \mathbf{0}$$

$$\mathbf{w}A \leq \mathbf{c}, \quad \mathbf{w} \geq \mathbf{0}$$

$$\mathbf{w}(A\mathbf{x}^* - \mathbf{b}) = 0, (\mathbf{w}A - \mathbf{c})\mathbf{x}^* = 0$$

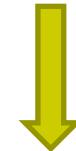


3.1 Karush-Kuhn-Tucker条件

- 证明:
- 充分性

$$Ax^* \geq \mathbf{b}, x^* \geq \mathbf{0} \longrightarrow x^* \text{ 是可行解}$$

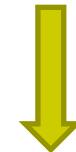
$$\begin{aligned} \mathbf{c} - \mathbf{w}^T A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0} &\longrightarrow (\mathbf{c} - \mathbf{w}^T A - \mathbf{v})^T (x - x^*) = 0 \\ \mathbf{w}^T (Ax^* - \mathbf{b}) = 0, \mathbf{v}^T x^* = 0 \end{aligned}$$



$$(\mathbf{c} - \mathbf{w}^T A - \mathbf{v})^T (x - x^*) = \mathbf{c}^T (x - x^*) - \mathbf{w}^T (Ax - \mathbf{b}) - \mathbf{v}^T x$$



$$Ax \geq \mathbf{b}, x \geq \mathbf{0} \longrightarrow \mathbf{c}^T x \geq \mathbf{c}^T x^*$$



x^* 是最优解



3.1 Karush-Kuhn-Tucker条件

- 证明:
- 必要性 \mathbf{x}^* 是最优解 $\longrightarrow \mathbf{x}^*$ 是可行解 $\longrightarrow A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0}$

$$I = \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}, I_1 \mathbf{x}^* > 0, I_2 \mathbf{x}^* = 0$$

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}, \mathbf{b} = \begin{pmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \end{pmatrix} \longrightarrow A_1 \mathbf{x}^* = \mathbf{b}_1, A_2 \mathbf{x}^* > \mathbf{b}_2$$

$$\begin{aligned} \mathbf{c} - \mathbf{w}^T A - \mathbf{v} &= \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0} \\ \mathbf{w}^T (A\mathbf{x}^* - \mathbf{b}) &= 0, \mathbf{v}^T \mathbf{x}^* = 0 \end{aligned}$$

$$B = \begin{pmatrix} A_1 \\ I_2 \end{pmatrix}, B\mathbf{d} \geq 0, \text{ 充分小的正数 } \lambda > 0 \quad A(\mathbf{x} + \lambda \mathbf{d}) \geq \mathbf{b}, \mathbf{x} + \lambda \mathbf{d} \geq 0$$

$$\mathbf{c}^T \mathbf{x}^* \leq \mathbf{c}^T (\mathbf{x}^* + \lambda \mathbf{d}) \quad \mathbf{c}^T \mathbf{d} \geq 0 \quad \text{不等式组: } \mathbf{c}^T \mathbf{d} < 0, B\mathbf{d} \geq 0, \text{ 无解!}$$

根据Farkas引理, 存在非负向量 $\mathbf{r} \geq 0$, 使得 $B^T \mathbf{r} = (A_1^T, I_1^T) \mathbf{r} = \mathbf{c}$



3.1 Karush-Kuhn-Tucker条件

- 证明:
- 必要性 \mathbf{x}^* 是最优解 $\longrightarrow \mathbf{x}^*$ 是可行解 $\longrightarrow A\mathbf{x}^* \geq \mathbf{b}, \mathbf{x}^* \geq \mathbf{0}$

根据Farkas引理，存在非负向量 $\mathbf{r} \geq 0$ ，使得 $B^T \mathbf{r} = (A_1^T, I_1^T) \mathbf{r} = \mathbf{c}$

$\mathbf{r} = (\mathbf{r}_1, \mathbf{r}_2)$, 分块方式与B对应，容易构造两个非负向量

$$\mathbf{w} = \begin{pmatrix} \mathbf{r}_1 \\ 0 \end{pmatrix}, \mathbf{v} = \begin{pmatrix} 0 \\ \mathbf{r}_2 \end{pmatrix}, \text{使得:}$$

$$A^T \mathbf{w} + \mathbf{v} = B^T \mathbf{r} = \mathbf{c}$$



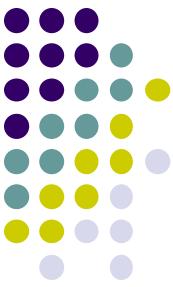
$$\mathbf{c} - \mathbf{w}^T A - \mathbf{v} = \mathbf{0}, \mathbf{w} \geq \mathbf{0}, \mathbf{v} \geq \mathbf{0}$$

容易验证 $\mathbf{v}^T \mathbf{x}^* = 0$



$$\mathbf{w}^T (A\mathbf{x}^* - \mathbf{b}) = 0, \mathbf{v}^T \mathbf{x}^* = 0$$

$$A\mathbf{x}^* - \mathbf{b} = \begin{pmatrix} 0 \\ A_2 \mathbf{x}^* - \mathbf{b}_2 \end{pmatrix}$$



3.1 Kuhn Tucker条件

对于标准形式的LP问题的K-T条件

只要将 $A\mathbf{x} = \mathbf{b}$ 改写成 $\begin{pmatrix} A \\ -A \end{pmatrix}\mathbf{x} \geq \begin{pmatrix} \mathbf{b} \\ -\mathbf{b} \end{pmatrix}$, 可推出相应的K-T条件

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

$$A^T \mathbf{w} + \mathbf{v} = \mathbf{c}, \mathbf{v} \geq \mathbf{0}$$

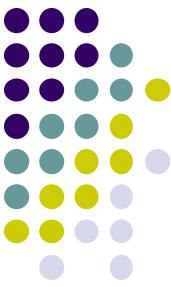
$$\mathbf{v}^T \mathbf{x} = 0$$

或等价的

$$A\mathbf{x} = \mathbf{b}, \mathbf{x} \geq \mathbf{0}$$

$$\mathbf{w}A \leq \mathbf{c}$$

$$(\mathbf{w}A - \mathbf{c})\mathbf{x} = 0.$$



3.1 Kuhn Tucker条件

一个基本可行解 $\begin{pmatrix} \mathbf{x}_B \\ \mathbf{x}_N \end{pmatrix} = \begin{pmatrix} \bar{\mathbf{b}} \\ \mathbf{0} \end{pmatrix}$ ($\bar{\mathbf{b}} > \mathbf{0}$)为最优的充要条件是检验向量 $\zeta \leq \mathbf{0}$,
和 $K-T$ 是什么关系呢?

由 $\mathbf{v}\mathbf{x} = \mathbf{v}_B\mathbf{x}_B + \mathbf{v}_N\mathbf{x}_N = 0$, 且 $\mathbf{x}_B > 0$, 可知 $\mathbf{v}_B = 0$.

条件 $\mathbf{c} - \mathbf{w}A - \mathbf{v} = \mathbf{0}$ 可改写为 $(\mathbf{c}_B, \mathbf{c}_N) - \mathbf{w}(B, N) - (\mathbf{v}_B, \mathbf{v}_N) = (\mathbf{0}, \mathbf{0})$, 于是有

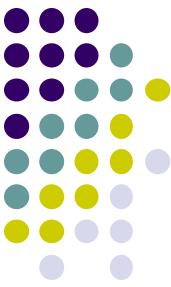
$$\begin{cases} \mathbf{c}_B - \mathbf{w}B = \mathbf{0} \\ \mathbf{c}_N - \mathbf{w}N - \mathbf{v}_N = \mathbf{0} \end{cases} \Rightarrow \begin{cases} \mathbf{w} = \mathbf{c}_B B^{-1} \\ \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1}N \end{cases}$$

这就是说, 若给一个基本可行解 x , 如果取

$$\mathbf{v}_B = \mathbf{0}, \mathbf{w} = \mathbf{c}_B B^{-1}, \mathbf{v}_N = \mathbf{c}_N - \mathbf{w}N = \mathbf{c}_N - \mathbf{c}_B B^{-1}N,$$

的 $K-T$ 条件除了 $\mathbf{v} \geq \mathbf{0}$ 外, 其余条件都已满足。

而根据 v 的取法, 易知 $v \geq \mathbf{0} \Leftrightarrow \zeta \leq \mathbf{0}$.



3.2 对偶理论

- 对偶问题可被看作是原始问题的“行列转置”

$$\min \quad \mathbf{c}^T \mathbf{x} \quad \text{s.t.} \quad \mathbf{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$$

$$\min \quad c^T x \quad \text{s.t.} \quad \begin{bmatrix} A \\ -A \end{bmatrix} x \geq \begin{bmatrix} b \\ -b \end{bmatrix}, x \geq 0$$

- 对偶问题是：

$$\max \quad [b^T \quad -b^T] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \quad \text{s.t.} \quad [A^T \quad -A^T] \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} \leq c$$

- 其中 y_1 和 y_2 分别表示对应于约束 $\mathbf{A}\mathbf{x} \geq \mathbf{b}$ 和 $-\mathbf{A}\mathbf{x} \leq \mathbf{B}$ 的对偶变量组。
令 $y=y_1-y_2$

$$\max \quad \mathbf{b}^T \mathbf{y} \quad \text{s.t.} \quad \mathbf{A}^T \mathbf{y} \leq \mathbf{c}$$



3.2 对偶理论

对偶规划生成规则：

- (1) $\min \rightarrow \max;$
- (2) $\mathbf{c} \rightarrow \mathbf{b};$ (3) $A \rightarrow A^T;$
- (4) 按以下规则添上不等号

原问题		对偶问题	
变 量	≥ 0 ≤ 0 无限制	行 约 束	\leq \geq =
行 约 束	\geq \leq =	变 量	≥ 0 ≤ 0 无限制

LP问题(P)

$$\begin{aligned} & \min \mathbf{c}\mathbf{x} \\ & s.t. \quad A\mathbf{x} \geq \mathbf{b} \\ & \quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$

对偶问题(D)为

$$\begin{aligned} & \max \mathbf{w}\mathbf{b} \\ & s.t. \quad \mathbf{w}A \leq \mathbf{c} \\ & \quad \mathbf{w} \geq \mathbf{0} \end{aligned}$$



3.2 对偶理论

对偶规划生成规则：

- (1) $\min \rightarrow \max;$
- (2) $\mathbf{c} \rightarrow \mathbf{b};$ (3) $A \rightarrow A^T;$
- (4) 按以下规则添上不等号

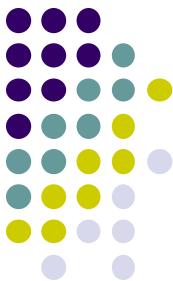
原问题		对偶问题	
变 量	≥ 0 ≤ 0 无限制	行 约 束	\leq \geq =
行 约 束	\geq \leq =	变 量	≥ 0 ≤ 0 无限制

标准LP问题

$$\begin{aligned} & \min \mathbf{c}\mathbf{x} \\ & s.t. \quad A\mathbf{x} = \mathbf{b} \\ & \quad \mathbf{x} \geq \mathbf{0} \end{aligned}$$

对偶问题为

$$\begin{aligned} & \max \mathbf{w}\mathbf{b} \\ & s.t. \quad \mathbf{w}A \leq \mathbf{c} \end{aligned}$$



3.2 对偶理论

对偶规划生成规则：

- (1) $\min \rightarrow \max;$
- (2) $\mathbf{c} \rightarrow \mathbf{b};$ (3) $A \rightarrow A^T;$
- (4) 按以下规则添上不等号

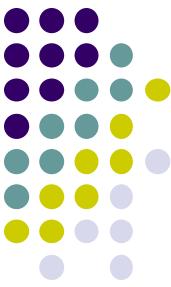
原问题		对偶问题	
变 量	≥ 0 ≤ 0 无限制	行 约 束	\leq \geq =
行 约 束	\geq \leq =	变 量	≥ 0 ≤ 0 无限制

LP问题

$$\begin{aligned}
 & \min \mathbf{c}_1 \mathbf{x}_1 + \mathbf{c}_2 \mathbf{x}_2 \\
 s.t. \quad & A_{11} \mathbf{x}_1 + A_{12} \mathbf{x}_2 \geq \mathbf{b}_1 \\
 & A_{21} \mathbf{x}_1 + A_{22} \mathbf{x}_2 = \mathbf{b}_2 \\
 & A_{31} \mathbf{x}_1 + A_{32} \mathbf{x}_2 \leq \mathbf{b}_3 \\
 & \mathbf{x}_1 \geq \mathbf{0}
 \end{aligned}$$

对偶问题为

$$\begin{aligned}
 & \max \mathbf{w}_1 \mathbf{b}_1 + \mathbf{w}_2 \mathbf{b}_2 + \mathbf{w}_3 \mathbf{b}_3 \\
 s.t. \quad & \mathbf{w}_1 A_{11} + \mathbf{w}_2 A_{21} + \mathbf{w}_3 A_{31} \leq \mathbf{c}_1 \\
 & \mathbf{w}_1 A_{12} + \mathbf{w}_2 A_{22} + \mathbf{w}_3 A_{32} = \mathbf{c}_2 \\
 & \mathbf{w}_1 \geq \mathbf{0}, \mathbf{w}_3 \leq \mathbf{0}
 \end{aligned}$$



对偶定理

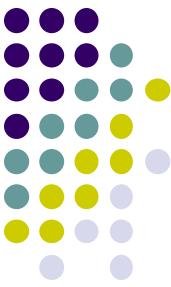
- 设 \mathbf{x} 和 \mathbf{w} 分别为(P)和(D)的可行解，则有 $\mathbf{c}\mathbf{x} \geq \mathbf{w}\mathbf{b}$.
- 设 \mathbf{x}^* 和 \mathbf{w}^* 分别为(P)和(D)的可行解，则 \mathbf{x}^* 和 \mathbf{w}^* 分别为(P)和(D)的最优解的充要条件是 $\mathbf{c}\mathbf{x}^* = \mathbf{w}^*\mathbf{b}$.
- 推论：若(P)有最优解 \mathbf{x}^* ，则(D)有最优解 \mathbf{w}^* ，且 $\mathbf{c}\mathbf{x}^* = \mathbf{w}^*\mathbf{b}$ ；若(P)无界，则(D)无解。反之亦然。
- 互补松弛性：设 \mathbf{x}^* 和 \mathbf{w}^* 分别为(P)和(D)的最优解，则有 $\mathbf{w}^*(A\mathbf{x}^* - \mathbf{b}) = \mathbf{0}$, $(\mathbf{c} - \mathbf{w}^*A)\mathbf{x}^* = \mathbf{0}$ 。

*K-T*条件：

$$A\mathbf{x}^* \geq \mathbf{b}, \quad \mathbf{x}^* \geq \mathbf{0} \quad (\text{原始可行性})$$

$$\mathbf{w}A \leq \mathbf{c}, \quad \mathbf{w} \geq \mathbf{0} \quad (\text{对偶可行性})$$

$$\mathbf{w}(A\mathbf{x}^* - \mathbf{b}) = \mathbf{0}, \quad (\mathbf{w}A - \mathbf{c})\mathbf{x}^* = \mathbf{0} \quad (\text{互补松弛性})$$



对偶定理

- 从K-T条件与单纯形方法中的最优准则的关系中可以看出：在单纯形方法中，除了K-T条件的 $v \geq 0$ 外，其余都得到满足。即单纯形法就是保持了原始可行性、互补松弛性、 $\mathbf{c} - \mathbf{w}A - \mathbf{v} = \mathbf{0}$ 得到满足的条件下逐步改善基本可行解，使得对偶可行性中的 $v = \mathbf{c} - \mathbf{w}A \geq 0$ 得到满足。
- 启发：也可以在保证对偶可行性和互补松弛性满足的条件下，改善解使得原始可行性得到满足。



对偶单纯形法

$$(P) \min \mathbf{c}\mathbf{x}$$

$$s.t. A\mathbf{x} = b$$

$$\mathbf{x} \geq 0$$

$$(D) \max \mathbf{w}\mathbf{b}$$

$$s.t. \mathbf{w}^T A \leq \mathbf{c}$$

D 的基本可行解：设 $A=(B, N)$, 其中 B 为满秩方阵, 则 $\mathbf{w}B=\mathbf{c}_B$ 的解 $\bar{\mathbf{w}}=\mathbf{c}_B B^{-1}$ 为 (D) 的基本解, 若 $\bar{\mathbf{w}}N \leq \mathbf{c}_N$, 则称 $\bar{\mathbf{w}}$ 为 (D) 的基本可行解。

P 的正则解：若原问题 (P) 的一个基本解 $\mathbf{x} = \begin{pmatrix} B^{-1}\mathbf{b} \\ 0 \end{pmatrix}$ 对应的检验向量 $\zeta = (\mathbf{0}, \mathbf{c}_B B^{-1}N - \mathbf{c}_N) \leq \mathbf{0}$, 则称 \mathbf{x} 为问题 (P) 的正则解。

此时的基 B 称为正则基。容易验证, D 的基本可行解和 P 的正则解是一一对应的。

同单纯形法一样, 求解对偶规划 (D) 从 (D) 的一个基本可行解迭代到另一个基本可行解, 使得目标函数值增加。等价地, 求解原规划 (P) 从一个正则解开始, 迭代到另一个正则解, 使目标函数 $z = \mathbf{w}\mathbf{b} = \mathbf{c}_B B^{-1}\mathbf{b} = \mathbf{c}\mathbf{x}$ 增加, 直到 $B^{-1}\mathbf{b} \geq 0$, 即正则解满足原始可行性时, 也就找到最优解。这种方法称为对偶单纯形法。



对偶单纯形法

$$(P) \min \mathbf{c}\mathbf{x}$$

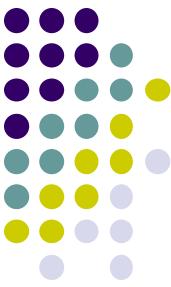
$$s.t. A\mathbf{x} = b$$

$$\mathbf{x} \geq 0$$

$$(D) \max \mathbf{w}\mathbf{b}$$

$$s.t. \mathbf{w}^T A \leq \mathbf{c}$$

- 1。找一个正则基 B , 建立单纯形表。
- 2。若 $\bar{\mathbf{b}} = B^{-1}\mathbf{b} \geq 0$, 停止, 已找到原问题最优解;
否则令 $\bar{b}_r = \min\{\bar{b}_i \mid i = 1, \dots, m\}$.
- 3。若 $\bar{\mathbf{a}}^r \geq 0$, 停止, 原问题无解;
否则令 $\frac{\zeta_k}{\bar{a}_{rk}} = \min\left\{\frac{\zeta_j}{\bar{a}_{rj}} \mid a_{rj} < 0\right\}$
- 4。以 \bar{a}_{rk} 为旋转轴元旋转, 返回2。



3.4 原始一对偶单纯形法

- 同时从对偶问题和第一阶段辅助问题求解
- 从对偶问题的任意可行解开始
 - 同时在迭代过程中保持对偶可行性和互补松弛性
 - 以及 $x \geq 0$ 的情况下通过剔除人工变量使 $Ax = b$ 成立

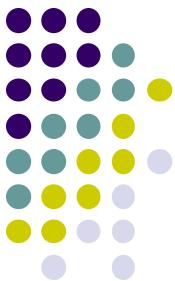
$$(P) \min \mathbf{c}\mathbf{x}$$

$$s.t. A\mathbf{x} = b$$

$$\mathbf{x} \geq 0$$

$$(D) \max \mathbf{w}\mathbf{b}$$

$$s.t. \mathbf{w}^T A \leq \mathbf{c}$$



$$(P) \min \mathbf{c}\mathbf{x}$$

$$s.t. \quad A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \geq \mathbf{0}$$

$$(D) \max \mathbf{w}\mathbf{b}$$

$$s.t. \quad \mathbf{w}A \leq \mathbf{c}$$

在(P)中引入人工变量后，得到辅助问题

$$\min g = \mathbf{1}\mathbf{x}_a$$

$$s.t \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \quad (\mathbf{b} \geq \mathbf{0})$$

$$\mathbf{x}, \mathbf{x}_a \geq \mathbf{0}$$

若已知(D)的一个可行解 $\bar{\mathbf{w}}$ ，为保持互补松弛性($\mathbf{x}(\bar{\mathbf{w}}A - \mathbf{c}) = \mathbf{0}$)，将令

$x_j = 0$ (当 $\bar{\mathbf{w}}\mathbf{a}_j \neq c_j$)，于是辅助问题为

$$(P') \min g = \mathbf{1}\mathbf{x}_a$$

$$s.t \quad A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$x_j = 0, j \notin Q$$

$$x_j \geq 0, j \in Q$$

$$\mathbf{x}_a \geq \mathbf{0}$$

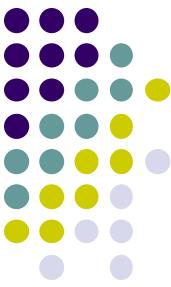
其中 $Q = \{j \mid \bar{\mathbf{w}}\mathbf{a}_j = c_j, j = 1, \dots, n\}$ 。问题(P')称为对应于 $\bar{\mathbf{w}}$ 的限定问题。

求解问题(P')，得到最优解 $\begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}_a^* \end{pmatrix}$ ，则它是(P)的辅助问题的基本可行解。

若 $\mathbf{x}_a^* = \mathbf{0}$ ，则 \mathbf{x}^* 为问题(P)的最优解

(因为 \mathbf{x}^* 和 $\bar{\mathbf{w}}$ 分别是(P)和(D)的可行解，且满足互补松弛性)。

若 $\mathbf{x}_a^* \neq \mathbf{0}$ ，找(D)的另一个基本可行解 $\hat{\mathbf{w}}$ ，使得(D)目标函数值有所增加，同时对应于 $\hat{\mathbf{w}}$ 的限定问题的最优值将较对应于 $\bar{\mathbf{w}}$ 的限定问题的最优值有所减少(等价于剔除人工变量)。



3.4 原始一对偶单纯形法

考慮(P')的对偶問題(D')

$$(D') \max \mathbf{v}\mathbf{b}$$

$$s.t. \quad \mathbf{v}\mathbf{a}_j \leq 0, j \in Q$$

$$\mathbf{v} \leq \mathbf{1}$$

令 \mathbf{v}^* 为其最优解。如对于所有的 $j = 1, \dots, n$, 都要 $\mathbf{v}^*\mathbf{a}_j \leq \mathbf{0}$, 则说明 \mathbf{v}^* 还是辅助

问题的对偶问题的最优解, 于是 $\begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}_a^* \end{pmatrix}$ 就是辅助问题的最优解。由于 $\mathbf{x}_a^* \neq \mathbf{0}$, 故

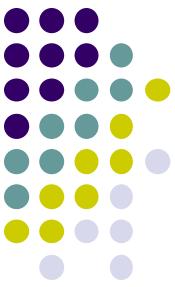
原問題(P)无解。

构造 $\hat{\mathbf{w}} = \bar{\mathbf{w}} + \theta \mathbf{v}^*$, 其中

$$\theta = \min \left\{ -\frac{\bar{\mathbf{w}}\mathbf{a}_j - c_j}{\mathbf{v}^*\mathbf{a}_j} \mid \mathbf{v}^*\mathbf{a}_j > 0, j = 1, \dots, n \right\}.$$

可以证明 $\hat{\mathbf{w}}$ 满足要求。同时, $\begin{pmatrix} \mathbf{x}^* \\ \mathbf{x}_a^* \end{pmatrix}$ 还是对应于 $\hat{\mathbf{w}}$ 的限定問題的一个基本可行解, 因此

可以使用它来作为初值来求解对应于 $\hat{\mathbf{w}}$ 的限定問題。如此循环。。。



3.4 原始一对偶单纯形法

1。变求解的规划为

$$\min \mathbf{c}\mathbf{x}$$

$$s.t. A\mathbf{x} = b$$

$$\mathbf{x} \geq 0$$

的形式。找其对偶规划的初始解 \mathbf{w} , 满足 $\mathbf{w}A \leq \mathbf{c}$ 。 (当 $\mathbf{c} \geq 0$ 时, 可取 $\mathbf{w}=0$ 作为初始解)

2。令 $Q = \{j \mid \bar{\mathbf{w}}\mathbf{a}_j = c_j, j = 1, \dots, n\}$ 。求解 \mathbf{w} 的限定问题

$$\min g = \mathbf{1}\mathbf{x}_a$$

$$s.t. A\mathbf{x} + \mathbf{x}_a = \mathbf{b}$$

$$x_j = 0, j \notin Q$$

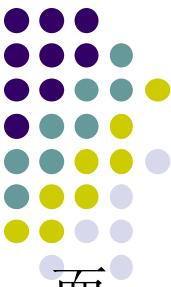
$$x_j \geq 0, j \in Q$$

$$\mathbf{x}_a \geq \mathbf{0}$$

若它的最优值 $g = 0$, 停止, 以及找到原问题的最优解; 否则, 求其对偶问题并令其最优解为 \mathbf{v} 。

3。若 $\mathbf{v}A \leq \mathbf{0}$, 则停止, 原问题无解; 否则令

$$\theta = \min \left\{ -\frac{\bar{\mathbf{w}}\mathbf{a}_j - c_j}{\mathbf{v}^*\mathbf{a}_j} \mid \mathbf{v}^*\mathbf{a}_j > 0, j = 1, \dots, n \right\}, \text{ 并且构造 } \mathbf{w} = \mathbf{w} + \theta \mathbf{v}^*, \text{ 返回2。}$$



3.5 对偶初始解

- 对偶单纯形法和原始一对偶单纯形法都需要一个对偶可行解。而且对偶单纯形法还要求是一个基本可行解。

对应任意一个LP问题，总可以使用高斯消去法找到一个基 B ，并化为典式。

此时若 $\zeta_N \leq 0$ ，则说明已经找到对偶问题的一个基本可行解 $w = c_B B^{-1}$ 。

否则，增加一个约束 $\mathbf{1}x_N \leq M$, M 是一个充分大的参数（从而这一约束不影响原问题）

增加约束后，称为扩充问题。（可证明：原问题有可行解 \Leftrightarrow 对于充分大的 M , 扩充问题有可行解）

z	\mathbf{x}_B	\mathbf{x}_N	x_{n+1}	RHS
z	1	0	ζ_N	0
\mathbf{x}_B	$\mathbf{0}$	I	\bar{N}	$\bar{\mathbf{b}}$
x_{n+1}	0	0	1	M

设 $\zeta_k = \max\{\zeta_i\}$ ，以 $\bar{a}_{m+1,k}$ 为转轴旋转得到新的单纯形表。因为 $\zeta \leq 0$ ，则新单纯形表对应一个正则解。于是可以开始用对偶单纯形法或原始一对偶单纯形法开始求解。

计算结果有两种可能：

- 扩充问题无解 \Rightarrow 原问题无解。
- 扩充问题有最优解 $\bar{\mathbf{x}} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M$ ，最优值为 $z_0 = z_1 + z_2 M$ 。这时原问题一定有可行解 $\hat{\mathbf{x}}$ ，且 $z_1 + z_2 M \leq \mathbf{c}\hat{\mathbf{x}}$ 。由于 M 充分大，故一定有 $z_2 \leq 0$ 。（否则 $\mathbf{c}\hat{\mathbf{x}}$ 会变得可以任意充分大了）

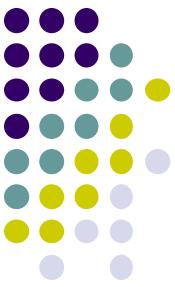
(i) $z_2 < 0$ 。当 $M \rightarrow \infty$, $z_0 \rightarrow -\infty$ 。问题无界

(ii) $z_2 = 0$ 。 $z_0 = z_1$ 为原问题最优值， $\bar{\mathbf{x}} \geq 0$ 为原问题的最优解。若 $\bar{\mathbf{x}}_2 = 0$, $\bar{\mathbf{x}}$ 还是一个基本可行解。否则 ($\bar{\mathbf{x}}_2 \neq 0$)，令 $M_o = \min\{M \mid \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M \geq 0\}$ ，此时 $\bar{\mathbf{x}}_0 = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M_0$ 也是基本可行解。而 $\{\mathbf{x} \mid \mathbf{x} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M, M \geq M_0\}$ 则表示可行区域的一个半直线界面，其上任意一点均是最优解。



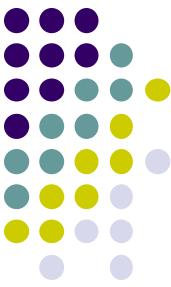
SVM与对偶方法

- http://www.cad.zju.edu.cn/home/zhx/csmath/lib/exe/fetch.php?media=svm_cjlin_dm.pdf



问题

- 单纯形法是不是一个好方法?
 - 性能
 - 对数据的适应性
 - 可实现性



算法复杂性

- **最坏性能**
 - 在最坏情况下，算法表现出来的性能。
- **平均性能**
 - 在各种可能情况，算法表现出来的期望性能。
- **问题(problem):**
 - 需要解决的数学问题，通常包含一组参数和未知的自由变量，并加以描述：
 - 对所有参数的描述
 - 对解所需要满足的特性的描述
- **实例(instance):**
 - 具体给定参数和描述的问题实例。
- **算法:**
 - 是一组含义明确的简单指令，遵循这些指令可以解决一个问题。
 - 问题的可解性，指存在一个算法，在有限时间和存储空间内，可求解该问题的任意实例。“停机问题”是算法不可解的。



算法复杂性

- 衡量标准：一个算法在运行时所需的初等运算次数。
- 输入规模：实例的编码长度
- 算法复杂性函数： $f: N \rightarrow R^+$
 - 多项式时间算法，增长速度是规模的多项式函数
 - $f(k) = O(k^3)$
 - 指数时间算法，增长速度是规模的指数函数
 - $f(k) = O(2^k)$



单纯形方法的复杂性

- 不是多项式时间的，最坏情况是指数时间的！

$$\max \quad c = \sum_{i=1}^n w_i$$

$$s.t. \quad w_i + 2 \sum_{j < i} w_j \leq \theta^{i-1}$$

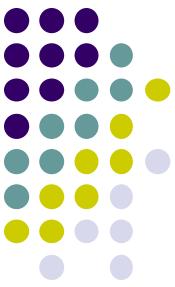
$$w_i \geq 0, i = 1, \dots, n$$

- 根据最小比值迭代需要 (MRT) 规则，需要 $2^n - 1$ 次迭代

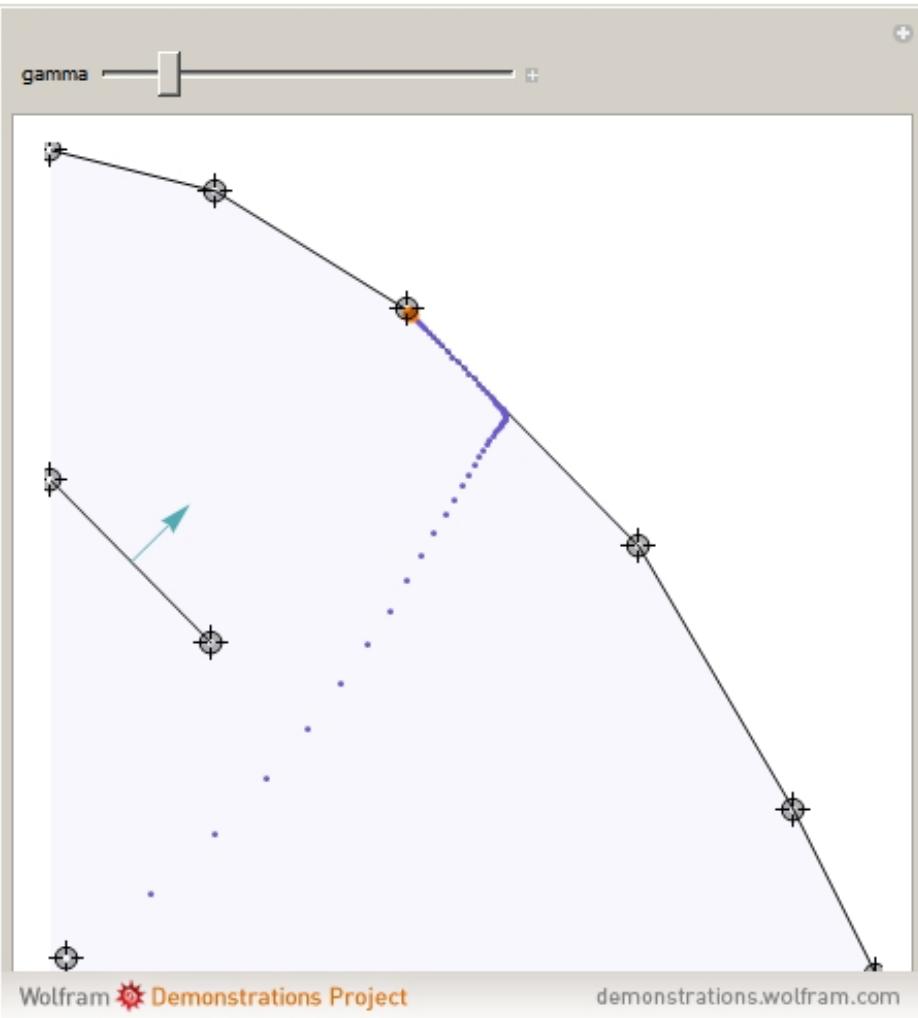
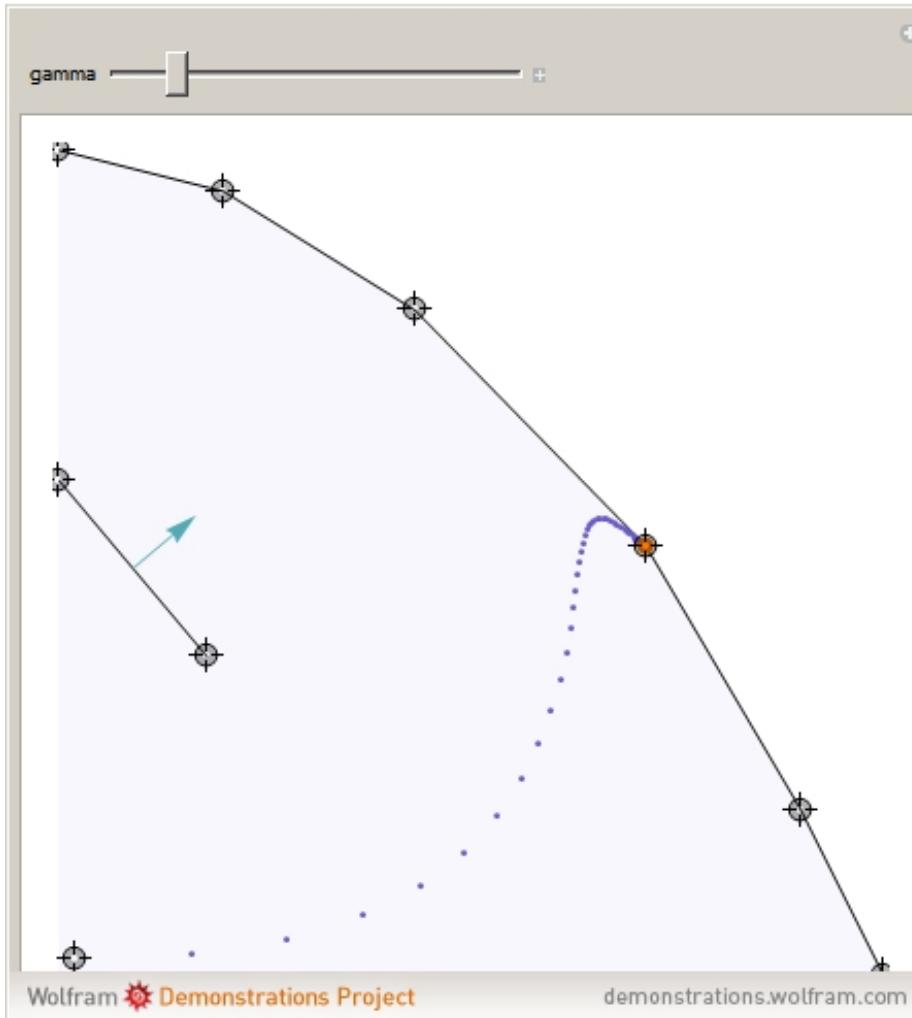


Karmarkar投影尺度算法

- 两个基本事实：
 - 如果一个内点位于多胞形的中心，沿着目标函数的最速下降方向是比较好的方向
 - 在不改变基本特性的条件下，存在一个适当的变换，可将可行域中的给定内点置于变换后的中心。
- 基本思想： 内点→投影尺度变换→中心→移动逆变换→新的内点→投影尺度变换→中心。 . .



Karmarkar投影尺度算法





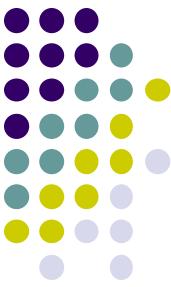
Karmarkar投影尺度算法

1. 出发点：与单纯形法不同，不沿边走，而从内部寻优。
1995年Frisch曾构造函数为：

$$\min z = \frac{1}{r_k} C^T X - \sum_{j=1}^n \log x_j$$

$$\text{s.t } AX=b$$

当 $x_j \rightarrow 0$, $z \rightarrow \infty$, 而永不靠边走。但存在问题, 收效慢
(中间点寻优方法属梯度法)



Karmarkar投影尺度算法

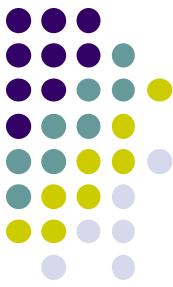
2. Karmarkar解决2个大问题 (1987 AT&T, but 1967 Dikin)

①定义目标势函数，按几何级数收敛，（属P算法）
变换原规划的最优解为0，使之第 k 次迭代值为：

$$C^T X^k \leq \left(\frac{1}{f}\right)^k C^T X^0$$

构造势函数为：

$$\begin{aligned} f(x) = f(x, c) &= n \log C^T X - \sum_{j=1}^n \log x_j \\ &= \sum_{j=1}^n \log\left(\frac{C^T X}{x_j}\right) \end{aligned}$$



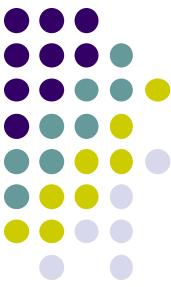
Karmarkar投影尺度算法

使变换中: $f(x^k) \leq f(x^0) - k\delta$

②从 X^k 点找下一点 X^{k+1} 点的关键是投影变换。记:

$$\Omega = \{x | Ax = b\} \quad P_+ = \{x | x \geq 0\} \quad P = \{x | \Omega \cap P_+\}$$

设 $a=(a_1, a_2, \dots, a_n)^T$ 是 P 中任一点 (Karmarka算法中是取某个可行点), 设法把 P_+ 投影到 $n+1$ 维空间的 n 维单纯形去, 且使 a 落到形心。



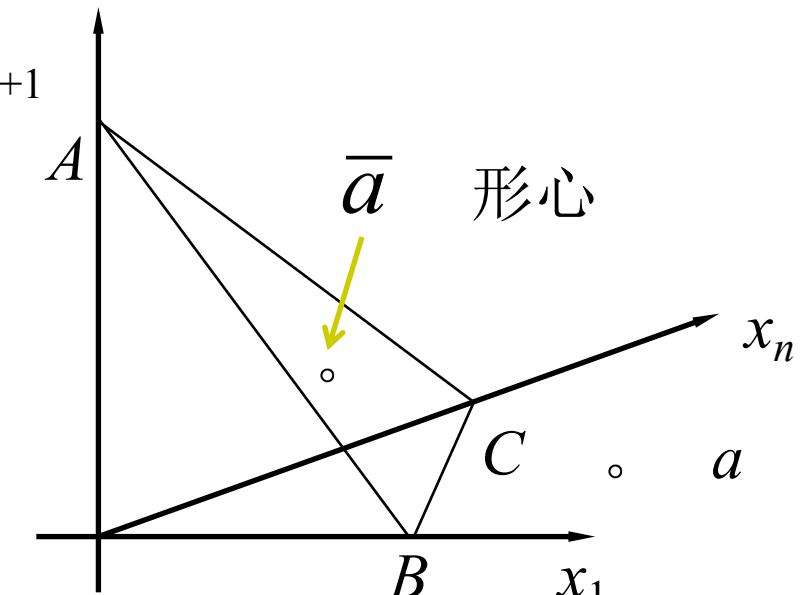
Karmarkar投影尺度算法

对于任意点 X , 投到 n 维单纯形后的坐标为:

$$\bar{x}_i = \frac{\frac{x_i}{a_i}}{\sum_{j=1}^n \left(\frac{x_j}{a_j}\right) + 1}$$

$$\bar{x}_{n+1} = \frac{1}{\sum_{j=1}^n \left(\frac{x_j}{a_j}\right) + 1}$$

($i=1, 2, \dots, n$)





Karmarkar投影尺度算法

Consider a Linear Programming problem in matrix form:

maximize $c^T x$

subject to $Ax \leq b$.

Algorithm Affine-Scaling

Input: $A, b, c, x^0, \text{stopping criterion}, \gamma$.

```
k ← 0
do while stopping criterion not satisfied
     $v^k \leftarrow b - Ax^k$ 
     $D_v \leftarrow \text{diag}(v_1^k, \dots, v_m^k)$ 
     $h_x \leftarrow (A^T D_v^{-2} A)^{-1} c$ 
     $h_v \leftarrow -A h_x$ 
    if  $h_v \geq 0$  then
        return unbounded
    end if
     $\alpha \leftarrow \gamma \cdot \min\{-v_i^k / (h_v)_i \mid (h_v)_i < 0, i = 1, \dots, m\}$ 
     $x^{k+1} \leftarrow x^k + \alpha h_x$ 
    k ← k + 1
end do
```

The End

新浪微博: @浙大张宏鑫

微信公众号:

