最优化方法 (II)

张宏鑫

2010-06-17

浙江大学计算机学院

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

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。

$利益驱动$

运筹学→

数学规划 →

线性规划(Linear Programming，LP)

- 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。

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
  - 2009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？

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

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

- 约束条件

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

附加条件：x_1和x_2是整数

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：

$$
\begin{aligned}
&\min\ c_1x_1 + \ldots + c_nx_n \\
&\text{s. t.} \\
&\qquad a_{11}x_1 + \ldots + a_{1n}x_n \le b_1 \\
&\qquad \ldots \\
&\qquad a_{m1}x_1 + \ldots + a_{mn}x_n \le b_m \\
&\qquad x_1, \ldots, x_n \ge 0
\end{aligned}
$$

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：

目标函数
objective function

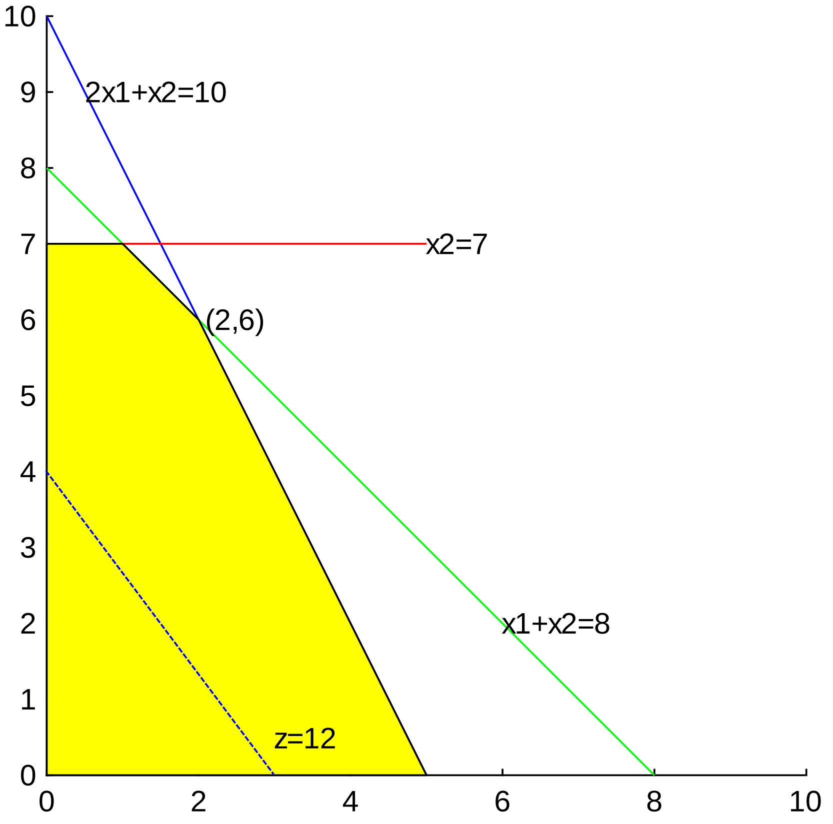
$$\min \mathbf{c}^T\mathbf{x}$$
$$\text{s.t. } A\mathbf{x} \le \mathbf{b}$$
$$\mathbf{x} \ge \mathbf{0}$$

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

线性规划的图解法

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

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



线性规划在Matlab中的实现方法

- Matlab中规定线性规划的标准形式为:

$$\min_{\mathbf{x}} \mathbf{c}^T\mathbf{x} \quad \text{such that} \quad \mathbf{A}\mathbf{x} \le \mathbf{b}$$

- 其中:
  - $\mathbf{c}$和$\mathbf{x}$为n维列向量,
  - $\mathbf{b}$为m维列向量,
  - $\mathbf{A}$为$m \times n$矩阵

线性规划在Matlab中的实现方法

- 基本函数形式为

  linprog(c, A, b)

  - 返回值是向量**x**的值
- 调用举例：

  [x,fval]=linprog(c, A, b, Aeq, beq, LB, UB, X0, OPTIONS)

  fval返回目标函数的值，

  Aeq、beq对应等式约束，Aeq X = beq

  LB、UB分别是变量X的下界和上界， X0是X的初始值

  OPTIONS是控制参数。

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

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

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$\mathbf{x}^1,\ldots,\mathbf{x}^k$,所有极方向为$\mathbf{d}^1,\ldots,\mathbf{d}^l$,则$x\in D$的充要条件是存在一组$\lambda_i(i=1,\ldots,k)$和$\mu_j(j=1,\ldots,l)$使得

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

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

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

线性规划基本定理

基本定理2（存在定理）

若LP问题

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

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

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

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. 并且，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$\mathbf{d}^i$, 均有$\mathbf{c}\mathbf{d}^i \geq 0$.

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）

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 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

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

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$K-T$条件得到满足

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

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

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

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

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

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

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

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$\zeta\le 0$，

和$K-T$是什么关系呢？

由$\mathbf{vx}=\mathbf{v}_B\mathbf{x}_B+\mathbf{v}_N\mathbf{x}_N=0$，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

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

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

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

3.2 对偶理论

- 对偶问题可被看作是原始问题的“行列转置”

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

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

- 对偶问题是：

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

- 其中y_1和y_2分别表示对应于约束Ax>=b和-Ax<=B的对偶变量组。令$y=y_1-y_2$

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

3.2 对偶理论

对偶规划生成规则：

(1) min $\to$ max；

(2) $\mathbf{c} \to \mathbf{b}$；　(3) $A \to A^T$；

(4) 按以下规则添上不等号

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

LP问题(P)

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

对偶问题(D)为

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

3.2 对偶理论

对偶规划生成规则：

(1) min $\to$ max；

（2）$\mathbf{c} \to \mathbf{b}$；　（3）$A \to A^T$；

（4）按以下规则添上不等号

原问题		对偶问题	
变量	≥ 0	行约束	$\leq$
	≤ 0		$\geq$
	无限制		$=$
行约束	$\geq$	变量	≥ 0
	$\leq$		≤ 0
	$=$		无限制

标准LP问题

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

对偶问题为

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

3.2 对偶理论

对偶规划生成规则：

(1) min $\to$ max；

（2）$\mathbf{c} \to \mathbf{b}$；　（3）$A \to A^T$；

（4）按以下规则添上不等号

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

*LP*问题

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

对偶问题为

$$
\begin{aligned}
\max\ & \mathbf{w}_1\mathbf{b}_1 + \mathbf{w}_2\mathbf{b}_2 + \mathbf{w}_3\mathbf{b}_3 \\
s.t.\ & \mathbf{w}_1A_{11} + \mathbf{w}_2A_{21} + \mathbf{w}_3A_{31} \leq \mathbf{c}_1 \\
& \mathbf{w}_1A_{12} + \mathbf{w}_2A_{22} + \mathbf{w}_3A_{32} = \mathbf{c}_2 \\
& \mathbf{w}_1 \geq \mathbf{0}, \mathbf{w}_3 \leq \mathbf{0}
\end{aligned}
$$

对偶定理

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

$K-T$条件：

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

$$\mathbf{w}A \le \mathbf{c}, \mathbf{w} \ge \mathbf{0} \qquad \text{（对偶可行性）}$$

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

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$\mathbf{v} \geq \mathbf{0}$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$\mathbf{v}=\mathbf{c}-\mathbf{w}A \geq \mathbf{0}$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。

3.3 对偶单纯形法

$$(P)\ \min \mathbf{cx} \qquad\qquad (D)\ \max \mathbf{wb}$$
$$s.t. A\mathbf{x} = b \qquad\qquad s.t.\ \ \mathbf{w}A \le \mathbf{c}$$
$$\mathbf{x} \ge 0$$

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$\overline{\mathbf{w}} = \mathbf{c}_B B^{-1}$为$(D)$的基本解，若$\overline{\mathbf{w}}N \le \mathbf{c}_N$，则称$\overline{\mathbf{w}}$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$\zeta = (\mathbf{0}, \mathbf{c}_B B^{-1} N - \mathbf{c}_N) \le \mathbf{0}$,

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$z = \mathbf{wb} = \mathbf{c}_B B^{-1}\mathbf{b} = \mathbf{cx}$增加，直到$B^{-1}\mathbf{b} \ge 0$，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性法。

3.3 对偶单纯形法

$$(P)\ \min \mathbf{cx} \qquad\qquad (D)\ \max \mathbf{wb}$$

$$s.t. A\mathbf{x}=b \qquad\qquad s.t.\ \ \mathbf{w}A\le \mathbf{c}$$

$$\mathbf{x}\ge 0$$

1。找一个正则基B，建立单纯形表。

2。若$\bar{\mathbf{b}}=B^{-1}\mathbf{b}\ge 0$，停止，已找到原问题最优解；

否则令$\bar{b}_r=\min\{\bar{b}_i\,|\,i=1,\dots,m\}$.

3。若$\bar{\mathbf{a}}^r\ge 0$，停止，原问题无解；

否则令$\dfrac{\zeta_k}{\bar{a}_{rk}}=\min\{\dfrac{\zeta_j}{\bar{a}_{rj}}\,|\,a_{rj}<0\}$

4。以$\bar{a}_{rk}$为转转轴元旋转，返回2。

3.4 原始—对偶单纯形法

- 同时对对偶问题和第一阶段地辅助问题求解。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。

$$(P)\ \min \mathbf{cx} \qquad\qquad (D)\ \max \mathbf{wb}$$
$$s.t.\ \ A\mathbf{x}=b \qquad\qquad s.t.\ \ \mathbf{w}A\le \mathbf{c}$$
$$\mathbf{x}\ge 0$$

在(P)中引入人工变量后，得到辅助问题

$$\min g=\mathbf{1x}_a$$
$$s.t\ \ A\mathbf{x}+\mathbf{x}_a=\mathbf{b} \quad (\mathbf{b}\ge 0)$$
$$\mathbf{x},\mathbf{x}_a\ge \mathbf{0}$$

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

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

$$(P')\ \min g=\mathbf{1x}_a$$
$$s.t\ \ A\mathbf{x}+\mathbf{x}_a=\mathbf{b}$$
$$x_j=0,\ j\notin Q$$
$$x_j\ge 0,\ j\in Q$$
$$\mathbf{x}_a\ge \mathbf{0}$$

其中$Q=\{j\mid \overline{\mathbf{w}}\mathbf{a}_j=c_j,\ j=1,\dots,n\}$。问题$(P')$称为对应于$\overline{\mathbf{w}}$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$\overline{\mathbf{w}}$分别是$(P)$和$(D)$的可行解，且满足互补松弛性）。

若$\mathbf{x}_a^*\ne\mathbf{0}$，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$\overline{\mathbf{w}}$的限定问题的最优值有所减少（等价于剔除人工变量）。

考虑(P')的对偶问题(D')

$$(D')\ \max\ \mathbf{vb}$$
$$s.t.\ \ \mathbf{va}_j \le 0,\ j \in Q$$
$$\mathbf{v} \le \mathbf{1}$$

令$\mathbf{v}^*$为其最优解。如对于所有的$j=1,\dots,n$，都要$\mathbf{v}^*\mathbf{a}_j \le \mathbf{0}$，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$\mathbf{x}_a^* \ne 0$，故原问题$(P)$无解。

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

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

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$\hat{\mathbf{w}}$的限定问题的一个基本可行解，因此可以使用它来作为初值来求解对应于$\hat{\mathbf{w}}$的限定问题。如此循环。。。

1。变求解的规划为

$$\begin{aligned} &\min \mathbf{cx} \\ &s.t.\ \ A\mathbf{x} = b \\ &\qquad \mathbf{x} \ge 0 \end{aligned}$$

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

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

$$\begin{aligned} &\min g = \mathbf{1x}_a \\ &s.t\ \ A\mathbf{x} + \mathbf{x}_a = \mathbf{b} \\ &\qquad x_j = 0, j \notin Q \\ &\qquad x_j \ge 0, j \in Q \\ &\qquad \mathbf{x}_a \ge \mathbf{0} \end{aligned}$$

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$\mathbf{v}A \le \mathbf{0}$，则停止，原问题无解；否则令

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

3.5 对偶初始解

- 对偶单纯形法和原始－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$\boldsymbol{\zeta}_N \le \mathbf{0}$,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$\mathbf{1x}_N \le M$, 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	$\boldsymbol{\zeta}_N$	0	z_0
$\mathbf{x}_B$	$\mathbf{0}$	I	$\bar{N}$	$\mathbf{0}$	$\bar{\mathbf{b}}$
x_{n+1}	0	$\mathbf{0}$	$\mathbf{1}$	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$\boldsymbol{\zeta} \le 0$，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－对偶单纯形法开始求解。

计算结果有两种可能：

1。扩充问题无解 $\Rightarrow$ 原问题无解。

2。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$z_1 + z_2 M \le \mathbf{c}\hat{\mathbf{x}}$。由于$M$充分大，故一定有$z_2 \le 0$。（否则$\mathbf{c}\hat{\mathbf{x}}$会变的可以任意充分大了）

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

(ii) z_2=0。$z_0 = z_1$为原问题最优值，$\bar{\mathbf{x}} \ge 0$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（$\bar{\mathbf{x}}_2 \ne 0$），令$M_o = \min\{M \mid \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M \ge 0\}$,此时$\bar{\mathbf{x}}_0 = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M_0$也是基本可行解。

而$\{\mathbf{x} \mid \mathbf{x} = \bar{\mathbf{x}}_1 + \bar{\mathbf{x}}_2 M, M \ge M_0\}$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。

二、二次规化

非线性最优化

- 最优化的问题的一般形式为

$$\text{Min } f(\mathbf{x}) \quad \text{s.t. } \mathbf{x} \in X$$

$f(\mathbf{x})$为目标函数，$X \subset E^n$为可行域。
如$X = E^n$，则以上最优化问题为无约束最优化问题。

约束最优化问题通常写为

$$\begin{aligned} &\text{Min } f(\mathbf{x}) \\ &\text{s.t. } c_i(\mathbf{x})=0,\ i \in E, \\ &\qquad c_i(\mathbf{x}) \geq 0,\ i \in I, \end{aligned}$$

其中E, I分别为等式约束的指标集和不等式约束的指标集，$c_i(\mathbf{x})$是约束函数。

非线性优化中的概念

- 可行点，可行域
- 极小，全局极小（总体极小点），全局严格极小，局部极小，局部严格极小
- 积极与非积极，积极约束，非积极约束
- 可行方向集，线性可行方向集，序列可行方向集

- Farkas引理与K-T条件

- 以上参见《最优化理论与方法》第八章

无约束二次最优化

$$\min f(\mathbf{x}) = \tfrac{1}{2}\,\mathbf{x}^{\mathrm{T}} H\mathbf{x}+\mathbf{c}^{\mathrm{T}}\mathbf{x},\ \mathbf{x}\in R^{n}$$

H是对称阵

基本解法：求导然后找局部极值。

二次规划的一般形式

$$\min\ f(\mathbf{x}) = \tfrac{1}{2}\,\mathbf{x}^{\mathrm{T}}H\mathbf{x}+\mathbf{c}^{\mathrm{T}}\mathbf{x},\ \mathbf{x}\in R^{n}$$

$$\text{s.t. } A\mathbf{x}\leq\mathbf{b} \qquad (1)$$

- 当H为对称矩阵时，被称为**二次规划**(Quadratic Programming，记作QP)。

- 特别，当H**正定**时，目标函数为凸函数，线性约束下可行域又是凸集。上式被称为**凸二次规划**。

二次规划的性质

- QP是一种最简单的非线性规划。QP有如下良好的性质，当H是半正定时：
  - K-T条件不仅是最优解的必要条件，而且是充分条件；
  - 局部最优解就是全局最优解。

等式约束下的二次规划

$$\min\ f(\mathbf{x}) = ½\, \mathbf{x}^{\mathrm{T}} H\mathbf{x} + \mathbf{c}^{\mathrm{T}}\mathbf{x},\ \mathbf{x} \in R^n$$

$$\text{s.t. } A\mathbf{x} = \mathbf{b} \qquad （2）$$

求解方法：Lagrange乘子法，求解以下无约束二次最优化问题。

$$L(\mathbf{x}, \boldsymbol{\lambda}) = \tfrac{1}{2}\mathbf{x}^{\mathrm{T}} H\mathbf{x} + \mathbf{c}^{\mathrm{T}}\mathbf{x} + \boldsymbol{\lambda}^{\mathrm{T}}(A\mathbf{x} - \mathbf{b})$$

令$L(\mathbf{x}, \boldsymbol{\lambda})$对$\mathbf{x}$和$\boldsymbol{\lambda}$的导数为零，得线性方程组

$$H\mathbf{x} + \mathbf{c}^{\mathrm{T}} + A^{\mathrm{T}}\boldsymbol{\lambda} = 0$$

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

可解得$\mathbf{x}$，即为上式的解。

二次规划的有效集方法

- 直观解释：将不起作用约束去掉，将起作用约束作为等式约束，通过解一系列等式约束的二次规划来实现不等式约束的优化。

- 基本原理：若$\mathbf{x}$为问题（1）的最优解，则它也是问题

$$\begin{aligned} &\min \frac{1}{2}\mathbf{x}^{\mathrm{T}}H\mathbf{x}+\mathbf{c}^{\mathrm{T}}\mathbf{x} \\ &\text{s.t. } \mathbf{a}_i^{\mathrm{T}}\mathbf{x}=b_i, i\in I \end{aligned} \tag{3}$$

- 的最优解，其中$\mathbf{a}_i$是A的第i行，I为起作用约束指标集（有效集）。
- 反之，若$\mathbf{x}$为（1）的可行解，又是（3）的K-T点，且相应的乘子$\lambda_i \geq 0$，则$\mathbf{x}$为（1）的最优解。

二次规划的有效集方法

- 算法步骤（迭代法）：
  - 设当前迭代点为$\mathbf{x}_k$， 它也是（1）的可行解。该点的有效集记作 I_k，为寻求$\mathbf{x}_k$点的迭代方向$\mathbf{d}$，用乘子法求解

$$\min \frac{1}{2}(\mathbf{x}_k+\mathbf{d})^{\mathrm{T}} H(\mathbf{x}_k+\mathbf{d})+\mathbf{c}^{\mathrm{T}}(\mathbf{x}_k+\mathbf{d})$$

$$\text{s.t. } \mathbf{a}_i^{\mathrm{T}}\mathbf{d}=0, i\in I_k$$

  - 若所得最优值$\mathbf{d}_k$=0，则$\mathbf{x}_k$是（3）的最优解。
    - 为判断它是否（1）的最优解，考察对应于有效约束的乘子 $\lambda_i \geq 0$是否成立。若成立，则$\mathbf{x}_k$是K-T点，由二次规划性质$\mathbf{x}_k$是（1）的最优解。

二次规划的有效集方法

- 算法步骤（迭代法）：

$$\min \frac{1}{2}(\mathbf{x}_k+\mathbf{d})^{\mathrm{T}} H(\mathbf{x}_k+\mathbf{d})+\mathbf{c}^{\mathrm{T}}(\mathbf{x}_k+\mathbf{d})$$

$$\text{s.t. } \mathbf{a}_i^{\mathrm{T}}\mathbf{d}=0, i\in I_k$$

  - 若最优值$\mathbf{d}_k\neq 0$，则取$\mathbf{x}_{k+1}=\mathbf{x}_k+\alpha\mathbf{d}_k$，在$\mathbf{x}_{k+1}$为可行点的条件下确定$\mathbf{d}_k$方向的步长$\alpha_k$
    - 如果存在p不在I_k中，使得$\mathbf{a}_p\mathbf{x}_{k+1}=\mathrm{b}_p$，则将$p$加入有效集

  - 如果存在I_k中的指标q，使得$\lambda_i<0$, 则$\mathbf{x}^{(k)}$不是最优解，从有效集中去掉q