中国运筹学会
中国工业与应用数学学会
丛书

应用数学译丛
第3号

最优化方法

[日]茨木俊秀 福岛雅夫著
曾道智译

世界图书出版公司

最优化方法

〔日〕茨木俊秀　福岛雅夫　著

曾道智　译

世界图书出版公司

北京・广州・上海・西安

1997

本书既考虑到这样的新发展，又注意到通俗易懂地说明了最优化的基本知识到最新成果的内容．也就是，从单纯法开始，精选叙述了

线性规划

网络最优化

组合最优化

非线性规划

的各个领域中有代表性的方法以及最近的话题．这些话题里有上述内点法，也有网络上针对最小成本流的强多项式算法，针对旅行商问题的多面体方法，非线性规划里的逐次二次规划法及接近法等等．另外，还特地用了一章解说最近引人注目的神经网络上的最优化方法．

最后，我们不仅从数学观点上对这些最优化方法进行了理论研究，还强调了它们作为实用工具在现实中被广泛应用的事实．大多数有代表性的最优化算法已有可以方便使用的软件包．但是，有效利用这些成果是以有待解决的问题已被模型化成最优化问题的形式为前提的．这一过程未必简单．在收集客观数据的同时，要求有深刻的洞察力和综合能力．运筹学兼蓄有针对性的知识及技能，但本书完全没涉及到．仅于此指出其重要性，想等别的机会详谈．

尽管仅靠最优化的方法不能解决所有问题，但是我们认为，这方法对解决最优化问题中可以模型化的部分是必不可少的，今后也会继续起到问题解决方法的核心作用．在理解最优化方法方面本书若能有所贡献，即便一点点，我们也很高兴．

茨木俊秀　福岛雅夫

1993 年 6 月　于京都，奈良

图书在版编目(CIP)数据

最优化方法/(日)茨木俊秀，(日)福岛雅夫著；曾道智译. －北京：世界图书出版公司北京公司，1997.3
(应用数学译丛/章祥荪等主编)
ISBN 7－5062－2853－X

Ⅰ.最… Ⅱ.①茨… ②福… ③曾… Ⅲ.最佳化 Ⅳ.0224

中国版本图书馆 CIP 数据核字(97)第 05869 号

茨木俊秀　福岛雅夫
最适化の方法
共立出版株式会社，1993

最优化方法
〔日〕茨木俊秀　福岛雅夫著
曾道智　译
责任编辑　高　蓉
世界图书出版公司北京公司出版
北京朝阳门内大街 137 号
北京昌平百善印刷厂印刷
新华书店北京发行所发行　各地新华书店经售
*
1997 年 4 月第　1　版　　开本：850×1168　1/32
1997 年 4 月第 1 次印刷　　印张：8
印数：0001－3000　　字数：18 万字
ISBN：7－5062－2853－X/O·175
著作权合同登记　图字：01－96－0117 号
定价：15.00 元

应用数学译丛

中文版前言

中国和日本的的确确是近邻，飞机仅需两个多小时就可以到．我们已经有幸多次访问过中国．北京的故宫、八达岭、西安的兵马俑、曲阜的孔府等等，让我们深深感受到中国四千年历史的源远流长．我们记忆尤新并感到骄傲的是，徒步登上泰山而不是坐揽车．

然而，和这些文化遗产以及美丽大自然比起来，中国年青人奋发图强，刻苦学习的精神面貌给我们留下了更为深刻的印象．这本书能译成中文出版，那怕是能满足一点点这些年青人的求知欲，我们也会因此感到由衷地高兴．

担任翻译工作的曾道智君是位生气勃勃的后起之秀，在京都大学我们研究室里刚获得博士学位．相信他一定能够准确译出专业术语(非常遗憾的是，我们连确认的中文知识都没有)．

茨木俊秀

福岛雅夫

1996 年 4 月　于京都大学

原著前言

如何使用计算机来解决现实中各种各样问题？信息数学积累了这方面的知识并不断深入发展. 被称为**最优化** (optimization) 或**数学规划** (mathematical programming) 的领域讨论的是为找出对象问题的最优解决策略而采取的模型化及其方法. 其过程是，先把待解决问题用最优化形式描述为在给定的约束条件下找出使某个目标函数达最大 (小) 的解 ", 然后再采用数学上严密的算法来求解. 再广泛些，这一领域也被认为是运筹学 (OR) 的一部分.

所谓问题解决方法，除了运筹学和数学规划之外，人们还提出了人工智能 (AI), 专家系统，系统理论，模糊集合，神经元，遗传算法等等各种典型方法，并使它们实用化. 这些方法各有各的特征，承担起解决问题某个方面的任务，但决不否定数学规划的作用. 这是因为把要解决的问题描述成最优化问题后，没有其它方法能像数学规划那样深刻解析对象问题并给出精密解.

普遍认为，数学规划领域的诞生是在 1947 年. 当时 G.B. Dantzig 提出了线性规划的基本算法 — 单纯形算法. 至今已有近半个世纪的历史. 理论上虽已显成熟，但我们也可以说它还是个正在继续茁壮成长的领域. 在这个意义上最近成为热门话题的是始自 1984 年 N. Karmarkar 发表的内点法. 这个新方法不仅促进了算法的高速化，还因为这种想法本身获得了美国的专利，成为围绕着知识产权而展开的话题，给人们留下了很深的印象. 经内点法刺激的单纯形法也取得了进步，在实用上现在可以解有自数万到数十万个变量的线性规划问题.

第 0 章 绪 论

最优化(optimization) 一般是指在某种状况下作出最好的决策，或者是从几个候选者中选出最好的．这种问题经常用下面的数学模型描述：

> “在给定的 **约束条件**(constraint) 下，找出一个 **决策变量** (decision variable) 的值，使得被称为 **目标函数**(objective function) 的表达愿望尺度的函数达到最小或最大值．”

一般说来决策变量 (以下简称为变量) 有多个，因此用 n 维向量 $x = (x_1, \ldots, x_n)$ 来表示，把问题写成下式[1]．

$$\begin{array}{ll} \text{目标函数：} & f(x) \to \text{最小} \\ \text{约束条件：} & x \in S \end{array} \tag{1}$$

在此，目标函数 f 是定义在包含 S 的适当集合上的实值函数．进一步，S 是该问题变量 x 的可取值之集合，称为问题 (1) 的**可行域** (feasible region)．一般来说，可行域 S 用与变量 x 相关的等式及不等式表示，但也有很多情况下包含有难以用数学式子表达的组合条件．

1) 目标函数最大化等价于把它乘以 -1 后最小化，因此本书在没有特地指出的情况下仅处理最小化问题．

满足约束条件 $x \in S$ 的 x 称为问题 (1) 的 **可行解** (feasible solution), 满足

$$f(x^*) \leq f(x) \quad (x \in S) \tag{2}$$

的可行解 $x^* \in S$ 称为问题 (1) 的 **最优解** (optimal solution). 另外，在包含可行解 $x^* \in S$ 的适当邻域 $U(x^*)$ 里，当

$$f(x^*) \leq f(x) \quad (x \in S \cap U(x^*)) \tag{3}$$

成立时，称 x^* 为问题 (1) 的 **局部最优解** (local optimal solution). 不少问题的目标函数或约束条件很复杂，要找出在全可行域中使目标函数达最小的解非常困难，这时面临的目标就成为求出局部最优解. 为区别 (2) 的最优解与局部最优解，有时特地称其为 **全局最优解** (global optimal solution).

根据变量，目标函数，约束条件的类型，最优化问题 (1) 可分成几类. 首先，根据变量，大致分为变量取连续实数的 **连续最优化问题** (continuous optimization problem) 以及取整数或者类似 0, 1 的离散值的 **离散最优化问题** (discrete optimization problem). 后者因多用组合性质来表达，也称为 **组合优化问题** (combinatorial optimization problem).

连续最优化问题分为目标函数是线性，约束条件是线性方程式或 (和) 不等式组时的 **线性规划问题** (linear programming problem) 及目标函数和约束条件未必是线性的 **非线性规划问题** (nonlinear programming problem). 后者可进一步分为目标函数是二次，约束条件是线性的 **二次规划问题** (quadratic programming problem), 目标函数是凸函数，约束条件是凸集合 [1] 的 **凸规划问题** (convex programming problem) 等. 凸规划问题有局部最优

---

1) 凸集合及凸函数的定义请参看附录 A.1.

解必定是全局最优解的性质，线性规划问题是凸规划问题的特殊情形.

离散最优化问题有各式各样的问题类，把它们具体特征化的主要原因在于其约束条件的结构. 一般说来，有几个变量被限制取整数时的问题称为 **整数规划问题** (integer programming problem), 特别是，当变量的值是 0 或 1 的时候称为 0-1 **规划问题** (0-1 programming problem). 另外，约束条件由关联图或网络给出的问题也很多. 例如，处理网络的流的各种 **网络流问题** (network flow problem) 及成为困难组合最优化问题典型的 **旅行商问题** (traveling salesman problem).

上述线性规划问题，非线性规划问题，整数规划问题等统称为 **数学规划问题** (mathematical programming problem).

本书论及各种最优化问题，着重于解说它们的解法 (算法). 虽然通读本书是以大学一二年级程度的数学知识为前提，但是书末的附录综述了常用的有关线性代数，多变量函数，图论及网络的基础知识，请参考.

第 1 章 线性规划：单纯形法

本章介绍解线性规划问题的有代表性的方法 — 单纯形法之基本思路. 有好几种执行单纯形法的计算方法，这里特别介绍被称为修正单纯形法的方法. 另外，也讲解线性规划里的对偶定理及灵敏度分析.

1.1 线性规划问题

目标函数为一次函数，约束条件为一次等式或不等式所表示的问题称为**线性规划问题** (linear programming problem, LP problem). 讨论线性规划问题的时候，为方便，考虑如下称为**标准形** (standard form) 的特殊形式问题.

$$\begin{aligned}&\text{目标函数}: c^T x \longrightarrow \text{最小}\\&\text{约束条件}: Ax = b,\ x \geq 0\end{aligned} \tag{1.1}$$

这里，变量 x 为 n 维实向量，A 为 $m \times n$ 常数矩阵，b 和 c 分别为 m 维及 n 维常数向量，T 为转置记号. 以下所有的向量全为列向量，同维的两个向量 x, y 所对应的各分量当 $x_i \geq y_i$ 成立时，记为 $x \geq y$.

虽然实际应用里出现各种线性规划问题，它们都可以变换成标准形 (1.1). 比如，函数 $c^T x$ 的最大化和 $(-c)^T x$ 的最小化等

价，不等式约束条件

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

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

在引进新变量 y_i 后，分别转化为等式约束条件和变量非负条件

$$\sum_{j=1}^{n} a_{ij}x_j + y_i = b_i,\ \ y_i \geq 0 \tag{1.2}$$

$$\sum_{j=1}^{n} a_{ij}x_j - y_i = b_i,\ \ y_i \geq 0 \tag{1.3}$$

另外，不受符号约束的变量 x_j 可像 $x_j = x_j' - x_j''$ 一样表示为两个非负变量 $x_j' \geq 0, x_j'' \geq 0$ 之差，因此，通过适当组合以上步骤，可把任意问题变换为与其等价的标准形. 一般把式 (1.2) 的 y_i 称为 **松弛变量** (slack variable), 式 (1.3) 的 y_i 称为 **剩余变量** (surplus variable).

1.2 基解和最优解

为讨论方便，以下假定问题 (1.1) 的变量个数 n 比约束条件 (变量非负条件除外) 个数 m 大，约束条件的系数矩阵 A 满足 $\operatorname{rank} A = m$[1]. 前面的假定在利用松弛变量和剩余变量变换后的标准形里通常成立，后面的假定是说在约束条件里没有 (像能通过组合其它几个条件而得到的) 多余条件式子.

1) 矩阵 A 的列向量中线性无关者的最大个数称为 A 的 **秩** (rank), 记为 $\operatorname{rank} A$.

考虑问题 (1.1) 的等式约束条件

$$Ax = b \tag{1.4}$$

这里，向量 x 的 n 个分量 $x_j\,(j=1,2,...,n)$ 分成 m 个和 $n-m$ 个两组，把它们以适当顺序排列后分别记为 x_B, x_N. 进一步，对应向量 x 的分组，把 $m\times n$ 矩阵 A 也分成 $m\times m$ 矩阵和 $m\times(n-m)$ 矩阵，前面的矩阵记为 B, 后面的记为 N[1]. 以下，把这种向量和矩阵划分表示为 $x=[x_B,x_N]$, $A=[B,N]$.

对应于某个划分 $x=[x_B,x_N]$, $A=[B,N]$, 式 (1.4) 可写为

$$Bx_B + Nx_N = b, \tag{1.5}$$

因此，如果 B 为非奇异矩阵，通过令 $x_N=0$, x_B 的值可由式 (1.5) 唯一确定为 $x_B = B^{-1}b$. 这样得到的

$$x = \begin{bmatrix} x_B \\ x_N \end{bmatrix} = \begin{bmatrix} B^{-1}b \\ 0 \end{bmatrix} \tag{1.6}$$

是方程式 (1.4) 的一个特殊解，称为问题 (1.1) 的 **基解** (basic solution). 特别是，在式 (1.6) 里 $x_B = B^{-1}b \geq 0$ 成立时，该基解满足问题 (1.1) 的约束条件，因此称为 **基可行解** (basic feasible solution). 一般来说，问题 (1.1) 的可行解有无数个，而基可行解仅有有限个[2]. 对应于式 (1.6) 的基解，称 B 为 **基矩阵** (basic matrix), 向量 x_B 的各分量称为 **基变量** (basic variable), N 称为

---

1) 在 $n=5, m=3$, $x=(x_1,...,x_5), A=(a_1,...,a_5)$ 时，如果 $x_B=(x_2,x_5,x_1), x_N=(x_3,x_4)$ 则 $B=(a_2,a_5,a_1)$, $N=(a_3,a_4)$. 这里，a_i 为矩阵 A 的第 i 列向量.

2) 因为可能的划分 $x=[x_B,x_N]$ 最多有 $\begin{pmatrix} n \\ m \end{pmatrix} = \dfrac{n!}{m!(n-m)!}$ 种，基可行解的个数不可能超过这个数.

非基矩阵 (nonbasic matrix), 向量 x_N 的各分量称为 **非基变量** (nonbasic variable).

作为具体的例子，考虑如下的约束条件.

$$\begin{aligned} &2x_1 + x_2 \leq 6 \\ &x_1 + 2x_2 \leq 6 \\ &x_1 \geq 0, \;\; x_2 \geq 0 \end{aligned} \tag{1.7}$$

因它不是标准形，用前节所述的方法变换成如下的标准形.

$$\begin{aligned} &2x_1 + x_2 + x_3 = 6 \\ &x_1 + 2x_2 + x_4 = 6 \\ &x_i \geq 0 \qquad (i = 1, 2, 3, 4) \end{aligned} \tag{1.8}$$

这里， x_3, x_4 是松弛变量. 式 (1.8) 里有下面 (a)–(f) 的六个基解. 这些基解里面，除 (b) 和 (e) 以外，其余四个为基可行解.

(a) $x_B = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$

$x_N = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

(b) $x_B = \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$

$x_N = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

(c) $x_B = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$

$x_N = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$$(d)\ x_B = \begin{bmatrix} x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 2 & 0 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 3 \\ 3 \end{bmatrix}$$

$$x_N = \begin{bmatrix} x_1 \\ x_4 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(e)\ x_B = \begin{bmatrix} x_2 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ -6 \end{bmatrix}$$

$$x_N = \begin{bmatrix} x_1 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$(f)\ x_B = \begin{bmatrix} x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 6 \\ 6 \end{bmatrix} = \begin{bmatrix} 6 \\ 6 \end{bmatrix}$$

$$x_N = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

式 (1.8) 里变量的维数是 4, 无法直接用图表示，但是与式 (1.8) 等价的原不等式组 (1.7) 在 (x_1, x_2)- 平面上可表示为图 1.1. 图 1.1 中，注意包括纵轴和横轴四条直线分别表示 $x_j = 0$ $(j = 1, 2, 3, 4)$. 于是，这四条直线中每两条之交点分别对应于上述六个基解，特别是，表示可行域四边形的顶点成为基可行解. 该例中可行域为四边形，一般情况下，线性规划问题可行域为 n 维欧几里得空间 R^n 的凸多面体，基可行解对应于该凸多面体的顶点 [1].

1) 空间 R^n 里，通过适当移动 $(n-1)$ 维子空间得到的集合称为 **超平面** (hyperplane). 三维空间 R^3 内的平面以及二维空间 R^2 内的直线都是 (各自空间里的) 超平面. 超平面可用某个向量 $a \neq 0$ 及实数 b 表示成 $\{x \in R^n \mid a^T x = b\}$, 进一步，一个超平面决定两个半空间 $\{x \in R^n \mid a^T x \geq b\}$ 和 $\{x \in R^n \mid a^T x \leq b\}$. 几个半空间的共同部分称为 **凸多面体** (polyhedral convex set).

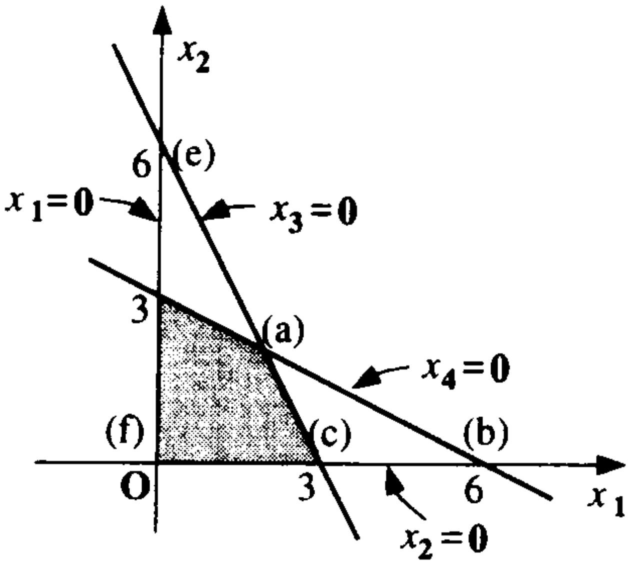
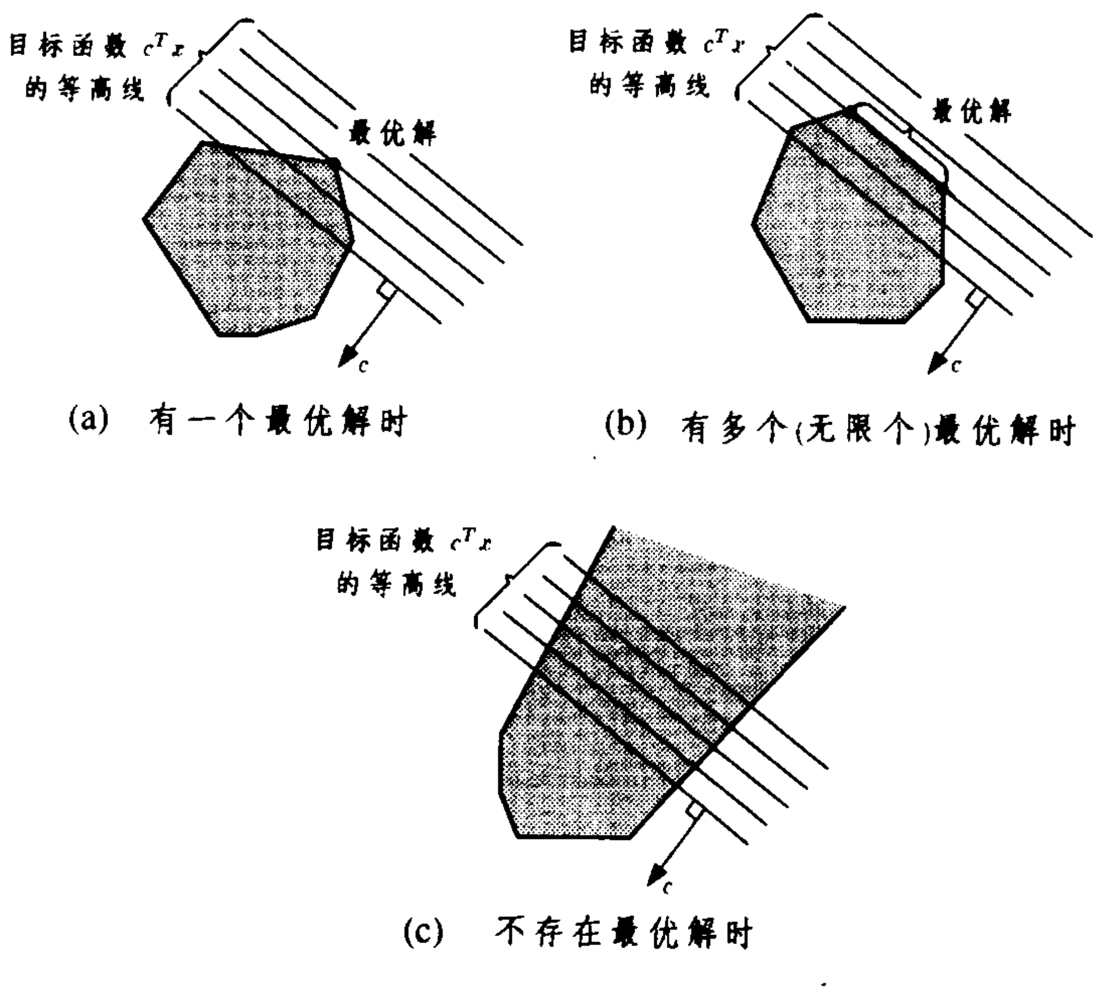


图 1.1 基 底 解



(a) 有一个最优解时

(b) 有多个(无限个)最优解时

(c) 不存在最优解时

图 1.2 线性规划问题的最优解

在空间 R^n 中作为目标函数的一次函数 c^Tx 的等高线是平

行超平面，因此线性规划问题的最优解的集合，要么像图 1.2 (a) 一样由一个顶点构成，要么像图 1.2 (b) 一样由可行域 (凸多面体) 的一个面构成，或者像图 1.2 (c) 那样可行域内目标函数值可以任意小. 特别是，当问题有最优解时，可行域的顶点之中一定存在有成为最优解者. 把以上各项综合一下就得到下面的定理.

定理 1.1 (线性规划问题的最优解的特征) 线性规划问题若有最优解，则基可行解中一定存在有最优解.

根据这个定理知道，为了找出线性规划问题的最优解没必要考虑 (无限个) 可行解的全体，可限定以具有式 (1.6) 的形式的 (有限个) 基可行解为对象进行搜索. 基于这个想法进行有效的最优基解搜索的是 1947 年 G.B. Dantzig 所设计的**单纯形法** (simplex method).

1.3 单纯形法

单纯形法是逐步生成有较小目标函数值的基可行解，最终到达最优解的迭代法. 这里，首先假定对标准形问题 (1.1) 得到了一个给出基可行解的划分 $x = [x_B, x_N]$, $A = [B, N]$ (求基可行解的一般方法在 1.6 节叙述). 此时，根据式 (1.5), 因为

$$x_B = B^{-1}b - B^{-1}Nx_N \tag{1.9}$$

目标函数的系数向量也同样划分为 $c = [c_B, c_N]$, 若用 z 表示目标函数值，则 z 可用仅含有非基变量 x_N 的形式表示如下[1].

1) 式 (1.10) 的 $j \in N$ 表达式里 N 被看成是非基变量的下标集合. 以下也同样，N 不仅表示非基矩阵，也用来表示非基变量下标集合.