

数据包络分析 及其应用

吴文江 著

DEA

C²GS² C²R

34	750	655	3
35	75		
36	716	650	3
37	704	713	3
38	76	712	4
39	777	710	3
43	774	685	3
41	764	712	

中国统计出版社
China Statistics Press



- 评价同类单位间相对有效性的有力工具
- 决策分析、管理、经济等领域中一个重要的方法
- 供管理干部、工程技术人员、社会科学工作者、大专院校有关专业的师生阅读

责任编辑 / 吕 军

封面设计 / 张建民

内容简介

数据包络分析是评价同类单位间相对有效性的有力工具，还可用来研究与具有多输入、多输出的边界生产函数有关的问题，如狭义技术进步率、生产率指标、预测、规模收益分析、最小成本问题及最大收益问题等，它已成为决策分析、管理、经济等领域中一个重要的方法。

本书简明地介绍了 C^2GS^2 模型、 C^2R 模型及混合DEA模型的概念、原理及应用案例。

本书可供管理干部、工程技术人员、社会科学工作者、大专院校有关专业的师生阅读。

ISBN 7-5037-3717-4



9 787503 737176 >

ISBN 7-5037-3717-4/O·43

定 价：18.80 元

6MW

1995 Press

数据包络分析 及其应用

吴文江 著

DEA

C^2GS^2 C^2R

中国统计出版社
China Statistics Press



中国统计出版社
北京

(京)新登字 041 号

图书在版编目(CIP)数据

数据包络分析及其应用/吴文江著.

—北京:中国统计出版社,2002.11

ISBN 7-5037-3717-4

I. 数…

II. 吴…

III. 统计分析—数学模型

IV. O212

中国版本图书馆 CIP 数据核字(2002)第 003780 号

作 者/吴文江

责任编辑/刘国宁 吕 军

封面设计/张建民

出版发行/中国统计出版社

通信地址/北京市西城区月坛南街 75 号 邮政编码/100826

办公地址/北京市丰台区西三环南路甲 6 号

电 话/(010)63459084、63266600—22500(发行部)

印 刷/北京科伦克三莱印务(北京)有限公司

经 销/新华书店

开 本/850×1168mm 1/32

字 数/190 千字

印 张/7.625

印 数/1—2000 册

版 别/2002 年 11 月第 1 版

版 次/2002 年 11 月第 1 次印刷

书 号/ISBN 7-5037-3717-4/O·43

定 价/18.80 元

中国统计版图书,版权所有,侵权必究。

中国统计版图书,如有印装错误,本社发行部负责调换。

前 言

数据包络分析(Data Envelopment Analysis,简称 DEA)是评价同类部门或单位间的相对有效性的有用的决策方法,可用于各行各业。例如用它对某行业的诸上市公司进行相对有效性评价;又如某企业中以每天的输入与输出作为一个生产活动,用 DEA 对若干天的生产活动进行相对有效性评价;再如根据轴功率、排水量及航速对诸船舶进行评价。DEA 也是在经济中用来研究具有多输入、多输出的边界生产函数的有力工具,因而可用它来研究与边界生产函数有关的问题,如狭义技术进步率、生产率指标、预测、规模收益分析、资金分配问题、最小成本问题及最大收益问题等。DEA 已成为经济、管理、决策分析等领域中一个重要的方法。

DEA 中第一个模型—— C^2R 模型产生于 1978 年。在改革开放的大好形势下,我国于 1986 年就有了介绍 DEA 的文章。1988 年魏权龄教授所著的《评价相对有效性的 DEA 方法》的出版,促使 DEA 方法在国内普及、推广。在解放思想、实事求是路线指引下,国内有关 DEA 的理论研究不断发展,应用领域不断扩大,已出版有关 DEA 的专著三本,在国内外发表文章三百多篇。

为了有助于 DEA 在我国进一步普及与应用,作者总结了多年的教学实践与科研成果、学习同行学者发表的文章并引用了学生们的应用成果,在这基础上写成了本书。它有下列特点:

(1)选择常用的基本内容,以便了解线性规划的读者就能看懂(除了 § 3.5、§ 3.9、§ 5.7 外)。

(2)注意从具体到抽象、从特殊到一般的方法来编写,力求比较完整地说明问题的来源。例如,从有单输入、单输出的生产函数谈起,结合平面图形来介绍 DEA 的概况。

(3)注重理论联系实际。前面五章中,只叙述定理并介绍定理的应用。在第七章才对前面五章的定理作证明。

由于作者水平有限,错误和不妥之处在所难免,敬请读者批评指正。

作 者

2001 年 9 月

目 录

前言

第一章 绪论..... (1)

§ 1.1 基本的 DEA 模型的产生 (1)

§ 1.2 从有单输入、单输出的生产函数谈起..... (2)

第二章 弱 DEA 有效性(C^2GS^2)及其应用..... (11)

§ 2.1 基本概念及有关定理 (11)

§ 2.2 有时间参数的单输出的边界生产函数的估计 (22)

§ 2.3 有多输出的生产函数的狭义技术进步率 (24)

§ 2.4 生产率指标的计算 (30)

§ 2.5 资金分配问题 (33)

第三章 DEA 有效性(C^2GS^2)及其应用 (36)

§ 3.1 基本概念及有关定理 (36)

§ 3.2 使决策单元(DMU)变为 DEA 有效的偏差和最小法 (42)

§ 3.3 判断 DEA 有效性(C^2GS^2)的简便方法 (47)

§ 3.4 在预测方面的应用 (48)

§ 3.5 基于加性模型与灰色预测的加权预测法 (50)

§ 3.6 利润最大问题 (54)

§ 3.7 决策单元过多的处理 (55)

§ 3.8 只有输出(入)的评价问题 (57)

§ 3.9	指标个数过多的处理	(59)
第四章	弱 DEA 有效性(C^2R)及其应用	(64)
§ 4.1	基本概念及有关定理	(64)
§ 4.2	使 DMU 的最优效率评价指数取给定值	(71)
§ 4.3	公司的平均投入利润率	(71)
§ 4.4	公司的下属厂上交利润的确定	(73)
§ 4.5	规模收益分析	(74)
§ 4.6	最小成本(最大收益)问题	(77)
§ 4.7	不变输入的利用率	(81)
§ 4.8	评审项目投资估算精度	(84)
§ 4.9	用图解法判断 DMU 的弱 DEA 有效性(C^2R)	(86)
第五章	DEA 有效性(C^2R)及其应用	(92)
§ 5.1	DEA 有效性(C^2R)及有关定理	(92)
§ 5.2	具有足够小的正数 ϵ 的 C^2R 模型	(102)
§ 5.3	船舶的效率评价	(105)
§ 5.4	规模收益分析与 DMU 变为 DEA 有效(C^2R)	(106)
§ 5.5	权有附加的约束的 C^2R 模型	(109)
§ 5.6	水泥厂的生产效率评价	(114)
§ 5.7	输入或输出为模糊数的评价问题	(116)
§ 5.8	有财产保险业务的公司的效率评价	(125)
§ 5.9	有关 C^2R 模型或 C^2GS^2 模型的灵敏度分析	(128)
第六章	混合的 DEA 模型及其应用	(134)
§ 6.1	DMU ₀ 非 DEA 有效(C^2R)的含意	(134)
§ 6.2	DEA 有效(M)的概念	(135)
§ 6.3	有关 DEA 有效(M)的定理	(139)
§ 6.4	混合的 DEA 模型的最优解与最优值	(146)
§ 6.5	使 DMU 变为 DEA 有效(M)	(149)
§ 6.6	地方水泥企业的效率评价	(158)
§ 6.7	沪市股票上市公司的经营效率评价	(159)

§ 6.8	建筑卫生陶瓷企业的效率评价	(161)
§ 6.9	权有附加约束的混合 DEA 模型	(163)
§ 6.10	有关混合的 DEA 模型的灵敏度分析	(166)
第七章	有关 DEA 有效性(C^2R 或 C^2GS^2)的理论	(185)
§ 7.1	有关的线性规划对偶理论	(185)
§ 7.2	有关弱 DEA 有效(C^2R 或 C^2GS^2)的理论	(186)
§ 7.3	有关 DEA 有效(C^2R 或 C^2GS^2)的理论	(207)
参考文献	参考文献	(232)

第一章 绪 论

本章从基本的 DEA 模型的产生与从单输入、单输出的生产函数谈起来介绍 DEA 的由来与概况

§ 1.1 基本的 DEA 模型的产生

在生产活动和社会活动中经常会遇到这样的问题:在经过一段时间后需要对同类的部门或单位进行评价,其中每个部门或单位称为决策单元(Decision Making Unit, 简记为 DMU), 评价的依据,一种是它在活动中需要耗费的量,所谓输入指标,另一种是表明该活动的成效的量,所谓输出指标。例如要对若干个建筑卫生陶瓷企业的经济效益进行评价,每个企业就是一个决策单元,输入指标是:产品工厂总成本、全部职工年平均人数、全年标准煤总耗量,输出指标是:卫生陶瓷产量、墙地砖产量、釉面砖产量、陶瓷锦砖产量。根据多个输入指标和多个输出指标的数据,要评价相对于所给的诸决策单元来说某决策单元的优劣,这就是评价决策单元间的相对有效性。

1978年由著名的运筹学家 A. Charnes、W. W. Cooper 和 E. Rhodes 提出的第一个 DEA 模型被命名为 C^2R 模型,用来评价决策单元间的相对有效性。1985年由 A. Charnes、W. W. Cooper、B. Golany、L. Seiford 和 J. Stutz 给出了另一个 DEA 模型,称为 C^2GS^2 模型,用来确定对应点位于生产函数图象上的决策单元。上

述两个模型就是本书所要介绍的 DEA 模型,此外,还有新的模型在产生和发展。

本书第二、三章中将介绍有关 C^2GS^2 模型的定理及其应用,第四、五章中将介绍有关 C^2R 模型的定理及其应用,为了解决实际问题的需要第六章中将介绍混合的 DEA 模型,第七章为 DEA 的理论,将给出前面五章中的定理的证明。

DEA 要研究多输入、多输出的情况,由于认识问题总是从特殊到一般,因而先来考虑单输入、单输出的情况。

§ 1.2 从有单输入、单输出的生产函数谈起

考虑投入一种生产要素来生产一种产品的情况。我们知道,在一定技术水平下,任给生产要素的输入(投入)量 $x \geq 0$,有最大可能的输出(产出)量 Y 与它对应,这个对应关系 $Y=f(x)$ 即生产函数,也称为边界生产函数。其图象如图 1.1 中的曲线 C 所示。

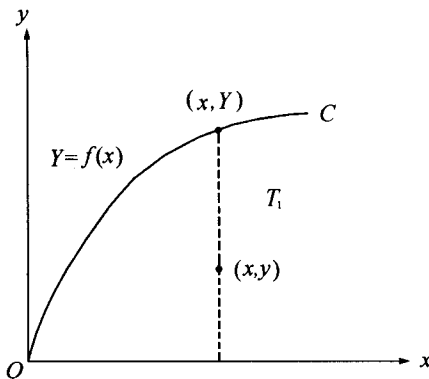


图 1.1

若 $0 \leq y \leq Y$, 则 y 为可能的输出,故 (x, y) 为可能的生产活动。可能的生产活动的全体 T_1 称为生产可能集,其图如图 1.1 中曲线 C 与 x 轴所围成的区域所示。称曲线 C 为区域 T_1 的上包络

线。

给定输入 $x \geq 0$, 对属于生产可能集 T_1 的生产活 (x, y) , 如前所述, 输出 y 有最大值 Y , 可写为

$$Y = \max y \quad (1.1)$$

$$s. t. (x, y) \in T_1$$

点 (x, Y) 在生产函数的图象、即 T_1 的上包络线 C 上。

今有 n 个同类的厂, 其中第 j 个厂有输入 x_j 、输出 y_j , 即有生产活动 $A_j(x_j, y_j)$, $j=1, \dots, n$ 。由其组成的集合

$$\hat{T} = \{(x_1, y_1), \dots, (x_n, y_n)\}$$

$n=6$ 时点 $A_j(x_j, y_j)$ ($j=1, \dots, 6$) 如图 1.2 所示。能否构造生产可能集, 并且在给定输入 x 时类似 (1.1) 式来确定生产函数的对应值?

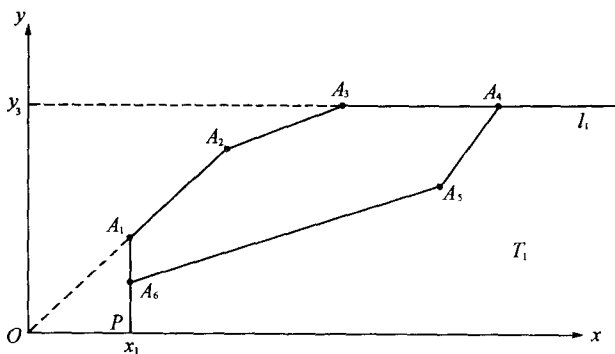


图 1.2

已知, 输入 x_1 对应有输出 y_1 、输入 x_2 对应有输出 y_2 , 因此输入 $(1/3)x_1 + (2/3)x_2$ 对应可能的输出 $(1/3)y_1 + (2/3)y_2$, 即

$$((1/3)x_1 + (2/3)x_2, (1/3)y_1 + (2/3)y_2)$$

$$= (1/3)(x_1, y_1) + (2/3)(x_2, y_2)$$

(对应点在线段 A_1A_2 上) 为可能的生产活动。

一般, 对任意两个可能的生产活动 (x, y) 与 (x', y') , 任给满足

$\lambda_1 + \lambda_2 = 1$ 的 $\lambda_1 \geq 0$ 与 $\lambda_2 \geq 0$, 输入 $\lambda_1 x + \lambda_2 x'$ 对应可能的输出 $\lambda_1 y + \lambda_2 y'$, 即

$$(\lambda_1 x + \lambda_2 x', \lambda_1 y + \lambda_2 y') = \lambda_1(x, y) + \lambda_2(x', y')$$

为可能的生产活动, 称此性质为凸性。

于是, 对 n 个生产活动 $(x_1, y_1), \dots, (x_n, y_n)$, 任给满足 $\sum_{j=1}^n \lambda_j = 1$ 的 $\lambda_j \geq 0 (j=1, \dots, n)$, 则

$$\left(\sum_{j=1}^n \lambda_j x_j, \sum_{j=1}^n \lambda_j y_j \right) = \sum_{j=1}^n \lambda_j (x_j, y_j)$$

(称为 $(x_1, y_1), \dots, (x_n, y_n)$ 的凸组合) 为可能的生产活动, 其全体记为

$$C(\hat{T}) = \left\{ \left(\sum_{j=1}^n \lambda_j x_j, \sum_{j=1}^n \lambda_j y_j \right) \mid \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n \right\}$$

$n=6$ 时其图为图 1.2 中以 $A_1, A_2, A_3, A_4, A_5, A_6$ 为顶点的凸多边形。

对可能的生产活动 (x_0, y_0) , 若 $x \geq x_0, 0 \leq y \leq y_0$, 则 (x, y) 为可能的生产活动, 即有较多的输入或有较少的输出进行的生产总是可能的, 称此性质为无效性。因此, 对可能的生产活动 $(\sum_{j=1}^n \lambda_j x_j,$

$\sum_{j=1}^n \lambda_j y_j) \in C(\hat{T})$, 若 y 非负

$$x \geq \sum_{j=1}^n \lambda_j x_j, y \leq \sum_{j=1}^n \lambda_j y_j$$

则 (x, y) 为可能的生产活动。于是

$$T_1 = \left\{ (x, y) \mid x \geq \sum_{j=1}^n \lambda_j x_j, y \leq \sum_{j=1}^n \lambda_j y_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n \right\} \quad (1.2)$$

为生产可能集。

$n=6$ 时, 其图为图 1.2 中 x 轴、线段 PA_1 与折线 l_1 (由折线 $A_1 A_2 A_3 A_4$ 及 $A_3 A_4$ 的延长线所组成) 所围成的区域。称 T_1 的上边

界线 l_1 为 T_1 的上包络线。

生产可能集 T_1 在 x 轴与 y 轴的投影分别为

$$T_{1x} = \{x | x \geq \sum_{j=1}^n x_j \lambda_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n\}$$

$$T_{1y} = \{y | y \leq \sum_{j=1}^n y_j \lambda_j, \sum_{j=1}^n \lambda_j = 1, \lambda_j \geq 0, j = 1, \dots, n\}$$

T_{1x} 为可能输入的集合, T_{1y} 为可能输出的集合。

在图 1.2 中 $T_{1x} = \{x | x \geq x_1\}$, $T_{1y} = \{y | y \leq y_3\}$ 。

$n = 6$ 时由图 1.3 可见: 设 $y_0 > 0$ 。任给 $x_0 \in T_{1x}$, 有 $\alpha > 0$ 使得 $(x_0, \alpha y_0) \in T_1$, 满足这条件的输出 αy_0 有最大值 $\alpha_0 y_0$, 即

$$\alpha_0 y_0 = \max(\alpha y_0) \quad (1.3)$$

$$s. t. (x_0, \alpha y_0) \in T_1$$

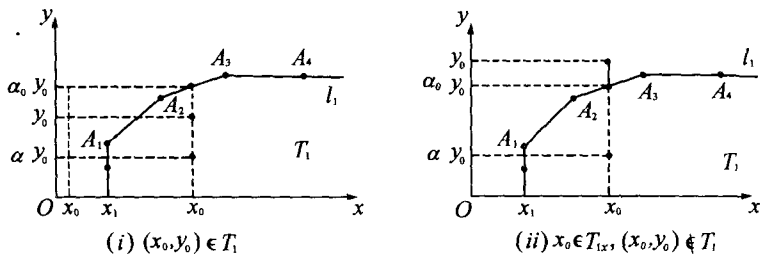


图 1.3

它与(1.1)类似。因 $y_0 > 0$, $(x_0, \alpha y_0) \in T_1$ 时 αy_0 的最大值为 $\alpha_0 y_0$ 即 α 的最大值为 α_0 , 即问题 (F_1)

$$\max \alpha \quad (1.4)$$

$$s. t. (x_0, \alpha y_0) \in T_1$$

的最大值为 α_0 。

如前所述, 并参看图 1.3, 有下列三个结论。

结论 1 设 $y_0 > 0$ 。任给输入 $x_0 \in T_{1x}$, 满足 $(x_0, \alpha y_0) \in T_1$ 的输出 αy_0 有最大值 $\alpha_0 y_0$ 与它对应, 这个对应关系为生产函数。点 $(x_0, \alpha_0 y_0)$ 在上包络线 l_1 上, 因此 l_1 为生产函数的图象。也就是, 若

α_0 为问题 (F_1) 的最大值, 则点 $(x_0, \alpha_0 y_0)$ 在生产函数的图象 l_1 上。

结论 2 由结论 1 可知, 当 $\alpha_0 = 1$, 即问题 (F_1) 的最大值为 1 时, (x_0, y_0) 在生产函数的图象 (即 T_1 的上包络线) l_1 上。

结论 3 设 $x_0 \in T_{1x}, y_0 > 0$ 。由图 1.3 的 (i) 可见, 当 $(x_0, y_0) \in T_1$ 时 $\alpha_0 y_0 \geq y_0$, 故 $\alpha_0 \geq 1$; 由图 1.3 的 (ii) 可见, 当 $(x_0, y_0) \notin T_1$ 时 $\alpha_0 y_0 < y_0$, 故 $\alpha_0 < 1$ 。因而, $x_0 \in T_{1x}$ 且 $y_0 > 0$ 时, $(x_0, y_0) \in T_1$ 的充分必要条件为问题 (F_1) 的最大值 $\alpha_0 \geq 1$ 。若 $x_0 \notin T_{1x}$, 由图 1.3 可见 $x_0 < x_1, (x_0, \alpha y_0) \notin T_1$, 故问题 (F_1) 无解。

以上考虑的问题是: 给定输入, 在生产可能集 T_1 中求输出最大的生产活动。自然产生另一个问题: 给定输出, 在生产可能集 T_1 中求输入最小的生产活动。

$n = 6$ 时由图 1.4 可见: 给 $y_0 \in T_{1y}$ 且 $x_0 > 0$, 满足 $(\theta x_0, y_0) \in T_1$ 的输入 θx_0 的最小值为 $\theta_0 x_0$, 即

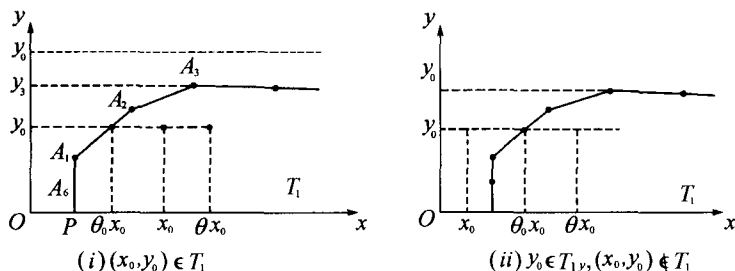


图 1.4

$$\theta_0 x_0 = \min(\theta x_0) \quad (1.5)$$

$$s. t. \quad (\theta x_0, y_0) \in T_1$$

由于 $x_0 > 0$, 即问题 (D_1)

$$\min \theta \quad (1.6)$$

$$s. t. \quad (\theta x_0, y_0) \in T_1$$

的最小值为 θ_0 。

如前所述, 并参看图 1.4, 有下列三个结论。

结论 4 设 $x_0 > 0$ 、任给 $y_0 \in T_{1y}$, $(\theta x_0, y_0) \in T_1$ 时有输入 $x = \theta x_0$ 的最小值 $\theta_0 x_0$ 与它对应, 且点 $(\theta_0 x_0, y_0)$ 在 T_1 的左边界线 l_2 (即折线 $PA_1A_2A_3$, 称它为 T_1 的左包络线) 上, 也就是, 若 θ_0 为问题 (D_1) 的最小值, 则点 $(\theta_0 x_0, y_0)$ 在 T_1 的左包络线 l_2 上。

结论 5 由结论 4 可知, 当 $\theta_0 = 1$, 即问题 (D_1) 的最小值为 1 时, 点 (x_0, y_0) 在 T_1 的左包络线 l_2 上。

结论 6 设 $y_0 \in T_{1y}$ 且 $x_0 > 0$ 。由图 1.4 的 (i) 可见, 当 $(x_0, y_0) \in T_1$ 时 $\theta_0 x_0 \leq x_0$, 故 $\theta_0 \leq 1$; 由图 1.4 的 (ii) 可见, 当 $(x_0, y_0) \notin T_1$ 时 $\theta_0 x_0 > x_0$, 故 $\theta_0 > 1$ 。因而 $y_0 \in T_{1y}$ 且 $x_0 > 0$ 时, $(x_0, y_0) \in T_1$ 的充分必要条件为 $\theta_0 \leq 1$ 。若 $y_0 \notin T_{1y}$, 由图 1.4 可见 $y_0 > y_3$, $(\theta x_0, y_0) \notin T_1$, 故问题 (D_1) 无解。

最小值与最大值统称为最优值。问题 (F_1) 与 (D_1) 都称为 C^2GS^2 模型。

结论 7 对由若干个生产活动所构成的生产可能集 T_1 及某厂对应的生产活动 $(x_0, y_0) \in T_1 (x_0 > 0, y_0 > 0)$, 若问题 (F_1) 的最优值 $\alpha_0 = 1$ 或问题 (D_1) 的最优值 $\theta_0 = 1$ (由结论 2 与结论 5 可知, 点 (x_0, y_0) 或位于 T_1 的上包络线 l_1 上、或位于 T_1 的左包络线 l_2 上, 即位于 T_1 的包络线 $l_1 \cup l_2$ 上), 称 (x_0, y_0) 对应的厂为弱 DEA 有效 (C^2GS^2)。这样就将对应点在生产可能集中的厂分为两类, 对应点在 T_1 的包络线 $l_1 \cup l_2$ 上的厂为弱 DEA 有效 (C^2GS^2) 的厂, 其余的厂为非弱 DEA 有效 (C^2GS^2)。

由图 1.2 可见, $A_j(x_j, y_j) (j = 1, \dots, 6)$ 中除了 A_5 外都在 T_1 的包络线 $l_1 \cup l_2$ 上。这包络线由线段或射线所组成, 其中每条都对应该某个斜率为非负的直线方程。以下由图 1.5 来观察在 T_1 的包络线 $l_1 \cup l_2$ 上的点 (x_0, y_0) 的特征。

(a) 设点 (x_0, y_0) 所在的直线 $L: \omega_0 x - \mu_0 y + \beta_0 = 0$, 即 $\omega_0 x_0 - \mu_0 y_0 + \beta_0 = 0$ 。

若 $y_0 > 0$ 且点 (x_0, y_0) 位于 T_1 的上包络线上, 若 $\mu_0 \neq 0$, 则直线方程可写为

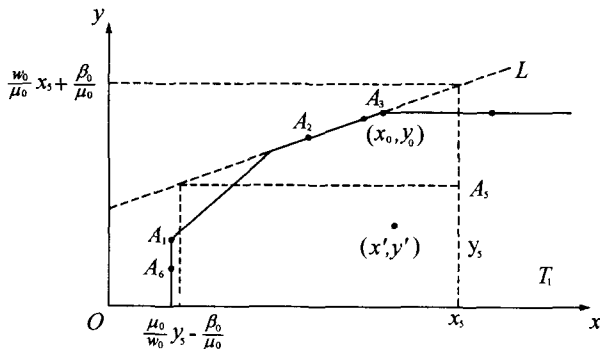


图 1.5

$$y = \left(\frac{\omega_0}{\mu_0} \right) x + \beta_0 / \mu_0, \text{ 斜率 } \omega_0 / \mu_0 \geq 0.$$

取 $\mu_0 > 0$, 则 $\omega_0 \geq 0$ 且 $\mu_0 y_0 > 0$

若 $x_0 > 0$ 且点 (x_0, y_0) 位于 T_1 的左包络线上, 若 $\omega_0 \neq 0$, 则直线方程可写为

$$x = (\mu_0 / \omega_0) y - \beta_0 / \omega_0, \quad \mu_0 / \omega_0 \geq 0$$

取 $\omega_0 > 0$, 则 $\mu_0 \geq 0$ 且 $\omega_0 x_0 > 0$.

(b) 由图 1.5 可见, 点 $(x_j, y_j) (j = 1, \dots, n = 6)$ 都不在直线 L 的上方或左方, 即

$$y_j \leq (\omega_0 / \mu_0) x_j + \beta_0 / \mu_0$$

(如 $y_5 < (\omega_0 / \mu_0) x_5 + \beta_0 / \mu_0$), $y_2 = (\omega_0 / \mu_0) x_2 + \beta_0 / \mu_0$)

或

$$x_j \geq (\mu_0 / \omega_0) y_j - \beta_0 / \omega_0$$

(如 $x_5 > (\mu_0 / \omega_0) y_5 - \beta_0 / \omega_0$)

统一写为: $\omega_0 x_j - \mu_0 y_j + \beta_j \geq 0, j = 1, \dots, n$.

此外, 由图 1.5 可见, 任意 $(x', y') \in T_1$ 都不在此直线 L 的上方或左方, 同理有

$$\omega_0 x' - \mu_0 y' + \beta_0 \geq 0$$

总之, 有下列结论。