对策论导论

●谢政　编著

科学出版社
www.sciencep.com

对策论导论

谢 政 编著

科学出版社
北京

内 容 简 介

本书重视基础性，强调完备性，兼顾前瞻性，力求用浅显的数学理论和方法来揭示对策论的深刻内涵，通俗易懂，便于自学. 它是对策论的入门教材，所涉及的都是对策论中最基本、最重要的理论和方法.

全书共九章，包括预备知识、对策、二人零和有限对策、二人零和无限对策、决策分析、非合作 n 人对策、合作 n 人对策、对策的应用以及微分对策.

可作为运筹学专业的研究生教材，也可供应用数学、系统科学、管理科学、经济学和军事运筹学等有关专业的教师、研究生和大学高年级学生参考.

图书在版编目(CIP)数据

对策论导论/谢政编著. —北京：科学出版社，2010

ISBN 978-7-03-026015-4

I. 对… II. 谢… III. 对策论-教材 IV. O225

中国版本图书馆 CIP 数据核字(2009) 第 206663 号

责任编辑：陈玉琢／责任校对：陈玉凤

责任印制：钱玉芬／封面设计：王　浩

科学出版社 出版

北京东黄城根北街 16 号

邮政编码：100717

http://www.sciencep.com

铭浩彩色印装有限公司 印刷

科学出版社发行　各地新华书店经销

*

2010 年 1 月第　一　版　开本：B5(720×1000)

2010 年 1 月第一次印刷　印张：20 3/4

印数：1—2 500　字数：407 000

定价：58.00 元

(如有印装质量问题，我社负责调换〈环伟〉)

前　　言

朴素的对策思想在中国古代源远流长, 田忌赛马的故事就是一个重要的例证. 而现代对策论则起源于 20 世纪初, 以 Zemelo、Borel 和 Von Neumann 等人的工作为代表. 第二次世界大战为对策论的应用提供了广泛的背景, 从而加快了对策论体系的形成. Von Neumann 和 Morgenstern 在 1944 合著的 "*Theory of Game Economic Behavior*" 是一本里程碑式的重要著作, 它完善了对策论的数学理论, 使之系统化和公理化. 为对策论的发展作出最重大贡献的除了 Von Neumann, 还有他的学生 Nash, Nash 解决了非合作对策的混合平衡局势的存在性问题, 奠定了非合作对策研究的基石, 他还提出了二人合作对策的谈判解, 这些都与经济学密切相关, 因而吸引了许多经济学家和统计学家投身于对策论的研究. 由于对策论的研究对象与社会、政治、军事、经济、科学、技术各领域都密切相关, 而且处理问题的方式又具有鲜明的特色, 因此引起了人们的广泛注意, 使之成为运筹学近年来发展较快的一个重要分支.

本书是对策论的入门教材, 所涉及的都是对策论中最基本、最重要的理论和方法. 第零章为预备知识, 简单介绍了凸集、凸函数、凸集分离定理、凸集表示定理、线性规划的单纯形法和对偶理论、Stieltjes 积分、不动点理论以及可测函数等内容. 编写这一章是为了便于读者查阅, 希望不至于使读者对对策论望而生畏. 第一章至第六章系统地介绍了静态对策的基本内容, 包括静态对策的概念、二人零和有限对策、二人零和无限对策、非合作 n 人对策和合作 n 人对策, 以及与对策密切相关的决策分析. 其中 1.3 节利用图论术语来描述对策树, 显得准确、简单、自然, 并且与第四章的决策树相呼应; 3.5 节给出了求解特殊的凸连续对策中局中人乙的最优纯策略、以及特殊的凹连续对策中局中人甲的最优纯策略的一般方法; 第四章用对策论的观点和方法来讨论决策分析的内容是我们的首次尝试; 5.3 节给出了非合作双矩阵对策的一般求解方法; 6.10 节合作双矩阵对策的谈判解是对 Nash 谈判解的简化, 回避了复杂的 Nash 公理体系. 第七章介绍了静态对策在经济和军事上的应用以及元对策. 第八章初步介绍了微分对策的一些基本概念, 着重突出离散序列法的思想, 而忽视微分对策求解方法的探讨. 每章都有一定的例题和习题, 其中很多都是有着应用背景的.

重视基础性, 强调完备性, 兼顾前瞻性是我们编写的原则, 用浅显的数学理论

和方法来揭示对策论的深刻内涵是我们追求的目标.

本书的撰写和出版得到了国防科技大学研究生院和理学院数学与系统科学系的大力支持, 得到了科学出版社陈玉琢编辑的帮助, 在此表示诚挚的谢意.

谢　政

2009 年 6 月

目　录

第零章　预 备 知 识

为了读者阅读的方便, 我们把对策论将要涉及的预备知识集中于一章.

本章简明扼要地介绍凸集、凸函数、线性规则、Stieltjes 积分、不动点定理及可测函数与弱收敛等概念和性质.

0.1　凸　　性

本节讨论凸集和凸函数的概念及性质.

除特别声明外, 本书中所出现的向量均为行向量.

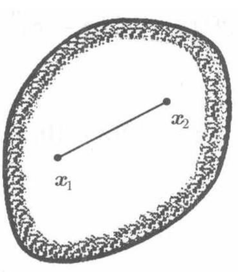
定义 0.1.1　设 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m \in \mathbb{R}^n$, 常数 $\lambda_i \geqslant 0 (i = 1, 2, \cdots, n)$, 且 $\sum\limits_{i=1}^{m} \lambda_i = 1$, 则称 $\sum\limits_{i=1}^{m} \lambda_i \boldsymbol{x}_i$ 为 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_m$ 的一个凸组合 (convex combination). 若 $\lambda_i > 0 (i = 1, 2, \cdots, m)$, 且 $\sum\limits_{i=1}^{m} \lambda_i = 1$, 则称该凸组合为严格凸组合 (strictly convex combination).

定义 0.1.2　设 $S \subseteq \mathbb{R}^n$, 且 $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, 有

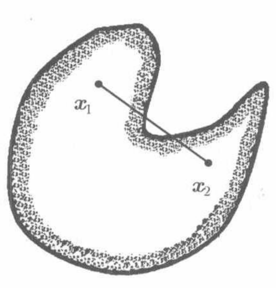
$$\lambda \boldsymbol{x}_1 + (1 - \lambda) \boldsymbol{x}_2 \in S, \quad \forall \lambda \in [0, 1],$$

则称 S 为 $\mathbb{R}^n$ 中的凸集 (convex set).

图 0.1.1 给出了凸集的正例和反例.



(a) 凸集



(b) 非凸集

图 0.1.1　凸集和非凸集

凸集的例子还很多, 如三角形、四面体、球形等. 一般来说, 凸集的并不一定是凸集, 但凸集的交仍是凸集.

定义 0.1.3　设 $S \subseteq \mathbb{R}^n$ 是凸集, $\boldsymbol{x}_0 \in S$, 若 $\boldsymbol{x}_0$ 不能表示成 S 中两个不同点的严格凸组合, 则称 $\boldsymbol{x}_0$ 是 S 的极点 (extreme point) .

例如, 在平面上, 闭三角形的三个顶点都是极点; 闭圆盘的圆周上每一点都是极点; 开圆盘没有极点; 原点是每个象限的惟一极点.

定义 0.1.4　设 $S \subseteq \mathbb{R}^n$, S 中任意两个点的所有凸组合构成的集合称为 S 的凸包 (convex hull), 记为 $H(S)$, 即

$$H(S) = \left\{ \sum_{i=1}^{m} \lambda_i \boldsymbol{x}_i | \boldsymbol{x}_i \in S, \lambda_i \geqslant 0, i = 1, 2, \cdots, m, \sum_{i=1}^{m} \lambda_i = 1, m \in \mathbb{N}_+ \right\}.$$

特别地, 若 $S = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k\}$, 则称 $H(S)$ 是由 $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots, \boldsymbol{x}_k$ 生成的凸包.

凸集有许多重要性质, 例如凸集分离定理和凸集表示定理.

定理 0.1.1(凸集分离定理)　设 $S \subseteq \mathbb{R}^n$ 为非空闭凸集, $\boldsymbol{y} \in \mathbb{R}^n \backslash S$, 则存在 $\boldsymbol{p} \in \mathbb{R}^n, \boldsymbol{p} \neq \boldsymbol{0}$ 以及 $\alpha \in \mathbb{R}$, 使得

$$\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}} \geqslant \alpha > \boldsymbol{p}\boldsymbol{y}^{\mathrm{T}}, \quad \forall \boldsymbol{x} \in S.$$

证明　因为 $S \subseteq \mathbb{R}^n$ 为非空闭凸集, $\boldsymbol{y} \notin S$, 而且范数 $\|\boldsymbol{x} - \boldsymbol{y}\|$ 关于 $\boldsymbol{x}$ 是 S 上的连续函数, 所以存在 $\bar{\boldsymbol{x}} \in S$, 使得

$$\|\bar{\boldsymbol{x}} - \boldsymbol{y}\| = \min\{\|\boldsymbol{x} - \boldsymbol{y}\| | \boldsymbol{x} \in S\} > 0.$$

由 S 为凸集可知

$$\lambda \boldsymbol{x} + (1 - \lambda)\bar{\boldsymbol{x}} \in S, \quad \forall \boldsymbol{x} \in S, \forall \lambda \in [0, 1].$$

于是

$$\begin{aligned}\|\bar{\boldsymbol{x}} - \boldsymbol{y}\|^2 &\leqslant \|\lambda \boldsymbol{x} + (1 - \lambda)\bar{\boldsymbol{x}} - \boldsymbol{y}\|^2 \\ &= \|\bar{\boldsymbol{x}} - \boldsymbol{y}\|^2 + \lambda^2 \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 + 2\lambda(\bar{\boldsymbol{x}} - \boldsymbol{y})(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}},\end{aligned}$$

从而

$$\lambda \|\boldsymbol{x} - \bar{\boldsymbol{x}}\|^2 + 2(\bar{\boldsymbol{x}} - \boldsymbol{y})(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \geqslant 0, \quad \forall \boldsymbol{x} \in S, \forall \lambda \in (0, 1).$$

令 $\lambda \to 0^+$, 得

$$(\bar{\boldsymbol{x}} - \boldsymbol{y})(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \geqslant 0, \quad \forall \boldsymbol{x} \in S.$$

记 $\boldsymbol{p} = \bar{\boldsymbol{x}} - \boldsymbol{y}$, 则 $\boldsymbol{p} \neq \boldsymbol{0}$, 且 $\boldsymbol{p}(\boldsymbol{x} - \bar{\boldsymbol{x}})^{\mathrm{T}} \geqslant 0 (\forall \boldsymbol{x} \in S)$. 又记 $\alpha = \boldsymbol{p}\bar{\boldsymbol{x}}^{\mathrm{T}}$, 则

$$\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}} \geqslant \alpha, \quad \forall \boldsymbol{x} \in S.$$

另一方面, 由于

$$\begin{aligned}\alpha-\boldsymbol{p}\boldsymbol{y}^{\mathrm{T}}&=\boldsymbol{p}\bar{\boldsymbol{x}}^{\mathrm{T}}-\boldsymbol{p}\boldsymbol{y}^{\mathrm{T}}=\boldsymbol{p}(\bar{\boldsymbol{x}}-\boldsymbol{y})^{\mathrm{T}}\\&=\|\boldsymbol{y}-\bar{\boldsymbol{x}}\|^2>0,\end{aligned}$$

$$\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}}\geqslant\alpha>\boldsymbol{p}\boldsymbol{y}^{\mathrm{T}},\quad\forall\boldsymbol{x}\in S.$$ □

凸集分离定理的几何意义可以用图 0.1.2 表示, 它说明在定理的条件下, 一定能找到一个超平面 $H=\{\boldsymbol{x}\in\mathbb{R}^n|\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}}=\alpha\}$ 严格分离 y 和 S.

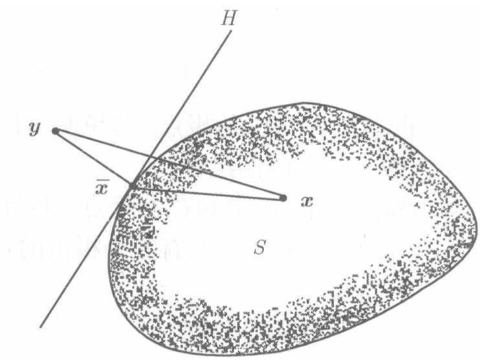


图 0.1.2 凸集分离定理的示意图

定理 0.1.2(凸集表示定理) 设 $S\subseteq\mathbb{R}^n$ 是非空有界闭凸集, K 是 S 的所有极点构成的集合, 则 $S=H(K)$.

证明 因 $K\subseteq S$, 故 $H(K)\subseteq S$. 下面对 n 用归纳法证明 $S\subseteq H(K)$, 即证 S 中每个元素都可以表示成 K 中有限个元素的凸组合.

当 n=1 时, S 就是 $\mathbb{R}$ 中有界闭区间 $[a, b]$, S 只有极点 a 和 b, 此时结论成立.

假设对于 $\mathbb{R}^{n-1}$ 任何非空有界闭凸集, 结论成立. 设 S 是 $\mathbb{R}^n$ 中非空有界闭凸集, $\boldsymbol{x}\in S$. 任取 $\boldsymbol{y}\in K$, 因 S 是有界闭凸集, 故过 $\boldsymbol{x}$ 和 $\boldsymbol{y}$ 的直线必与 S 的边界有一个交点 $\bar{\boldsymbol{y}},\bar{\boldsymbol{y}}\neq\boldsymbol{y}$. 显然 $\boldsymbol{x}$ 是 $\boldsymbol{y}$ 与 $\bar{\boldsymbol{y}}$ 的凸组合. 因此只需说明 $\bar{\boldsymbol{y}}$ 是 K 中有限个元素的凸组合即可.

因为 $\bar{\boldsymbol{y}}$ 在非空闭凸集 S 的边界上, 所以根据定理 0.1.1, 存在 $\boldsymbol{p}\in\mathbb{R}^n$, $\boldsymbol{p}\neq\mathbf{0}$ 及 $\alpha\in\mathbb{R}$, 使得

$$\boldsymbol{p}\bar{\boldsymbol{y}}^{\mathrm{T}}=\alpha,$$

$$\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}}\geqslant\alpha,\ \forall\boldsymbol{x}\in S,$$

记超平面 $H=\{\boldsymbol{x}\in\mathbb{R}^n|\boldsymbol{p}\boldsymbol{x}^{\mathrm{T}}=\alpha\}$. 易知 $\bar{\boldsymbol{y}}\in S\bigcap H$, 且 $S\bigcap H$ 是非空有界闭凸集. 由于 $S\bigcap H$ 在超平面 H 内, 因此它是 $n-1$ 维凸集. 从而由归纳法假设, $\bar{\boldsymbol{y}}$ 可以表示成 $S\bigcap H$ 中有限个极点的凸组合.

设 $\tilde{\boldsymbol{x}} \in S \bigcap H$, 即 $\boldsymbol{p}\tilde{\boldsymbol{x}}^{\mathrm{T}} = \alpha$. 若 $\tilde{\boldsymbol{x}}$ 不是 S 的极点, 则存在 $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in S$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2$, 及 $\lambda \in (0,1)$, 使得

$$\tilde{\boldsymbol{x}} = \lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2.$$

由于

$$\boldsymbol{p}\boldsymbol{x}_i^{\mathrm{T}} \geqslant \alpha, i = 1, 2,$$

因此

$$\alpha = \boldsymbol{p}\tilde{\boldsymbol{x}}^{\mathrm{T}} = \lambda \boldsymbol{p}\boldsymbol{x}_1^{\mathrm{T}} + (1-\lambda)\boldsymbol{p}\boldsymbol{x}_2^{\mathrm{T}} \geqslant \alpha,$$

从而

$$\boldsymbol{p}\boldsymbol{x}_i^{\mathrm{T}} = \alpha,\ i = 1, 2,$$

即知 $\boldsymbol{x}_1$, $\boldsymbol{x}_2 \in S \bigcap H$, 故 $\tilde{\boldsymbol{x}}$ 也不是 $S \bigcap H$ 的极点. 这说明 $S \bigcap H$ 的极点必为 S 的极点. 于是 $\bar{\boldsymbol{y}}$ 可表示成 S 有限个极点的凸组合. □

这个定理表明, 非空有界闭凸集 S 中必存在极点, 并且 S 中每个元素均可用 S 中有限个极点的凸组合来表示, S 中任意有限个极点的每个凸组合都是 S 的元素.

下面再介绍与凸集有关的两类特殊的函数.

定义 0.1.5　设 $S \subseteq \mathbb{R}^n$ 是非空凸集, $f: S \to \mathbb{R}$, 如果 $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\forall \lambda \in [0,1]$, 都有

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) \leqslant \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2),$$

则称 f 为 S 上的凸函数 (convex function). 如果 $\forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in S$, $\boldsymbol{x}_1 \neq \boldsymbol{x}_2, \forall \lambda \in (0,1)$, 都有

$$f(\lambda \boldsymbol{x}_1 + (1-\lambda)\boldsymbol{x}_2) < \lambda f(\boldsymbol{x}_1) + (1-\lambda) f(\boldsymbol{x}_2),$$

则称 f 为 S 上的严格凸函数 (strictly convex function).

如果 $-f$ 为 S 上的凸函数, 则称 f 为 S 上的凹函数 (concave function). 如果 $-f$ 为 S 上的严格凸函数, 则称 f 为 S 上的严格凹函数 (strictly concave function).

图 0.1.3 给出了凸函数、凹函数和非凸非凹函数的图形.

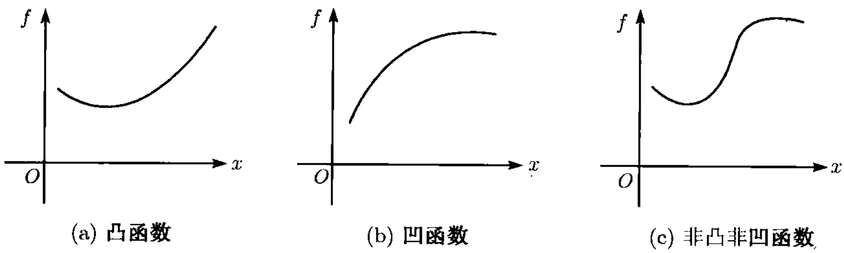


图 0.1.3　凸函数、凹函数和非凸非凹函数的图形

由凸函数和凹函数的定义易知, 线性函数

$$f(\boldsymbol{x})=\boldsymbol{a}\boldsymbol{x}^{\mathrm{T}}+b,\quad \boldsymbol{a},\boldsymbol{x}\in\mathbb{R}^n,\boldsymbol{a}\neq\boldsymbol{0},b\in\mathbb{R}$$

既是凸函数又是凹函数.

设 $S\subseteq\mathbb{R}^n$ 是非空开凸集, $f: s\to\mathbb{R}$. 如果 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处对于自变量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)^{\mathrm{T}}$ 的各分量的偏导数

$$\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_i}\quad(i=1,2,\cdots,n)$$

都存在向量, 则称

$$\nabla f(\bar{\boldsymbol{x}})=\left(\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_1},\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_2},\cdots,\frac{\partial f(\bar{\boldsymbol{x}})}{\partial x_n}\right)^{\mathrm{T}}$$

为 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的一阶导数或梯度. 如果 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处对于自变量的各分量的二阶偏导数

$$\frac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_i\partial x_j}\quad(i,j=1,2,\cdots,n)$$

都存在, 则称矩阵

$$\nabla^2 f(\bar{\boldsymbol{x}})=\begin{pmatrix}\dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1^2} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1\partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_1\partial x_n}\\ \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2\partial x_1} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2^2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_2\partial x_n}\\ \vdots & \vdots & & \vdots\\ \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n\partial x_1} & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n\partial x_2} & \cdots & \dfrac{\partial^2 f(\bar{\boldsymbol{x}})}{\partial x_n^2}\end{pmatrix}$$

为 $f(\boldsymbol{x})$ 在点 $\bar{\boldsymbol{x}}$ 处的二阶导数或 Hesse 矩阵.

下面不加证明地给出凸函数地两个充要条件.

定理 0.1.3 设 $S\subseteq\mathbb{R}^n$ 是非空开凸集, $f:S\to\mathbb{R}$ 在 S 上可微, 则 f 为 S 上的凸函数的充要条件是

$$f(\boldsymbol{x}_2)\geqslant f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),\ \forall\boldsymbol{x}_1,\boldsymbol{x}_2\in S;$$

f 为 S 上的严格凸函数的充要条件是

$$f(\boldsymbol{x}_2)>f(\boldsymbol{x}_1)+\nabla f(\boldsymbol{x}_1)^{\mathrm{T}}(\boldsymbol{x}_2-\boldsymbol{x}_1),\ \forall\boldsymbol{x}_1,\boldsymbol{x}_2\in S,\boldsymbol{x}_1\neq\boldsymbol{x}_2.$$ □

定理 0.1.4 设 $S\subset\mathbb{R}^n$ 是非空开凸集, $f:S\to\mathbb{R}$, 在 S 上具有二阶连续偏导数, 则 f 是 S 上的凸函数的充要条件是对于一切 $\boldsymbol{x}\in S$, f 在 $\boldsymbol{x}$ 处的 Hesse 矩阵 $\nabla^2 f(\boldsymbol{x})$ 是半正定矩阵.

□

0.2 线性规划

本节我们并不展开讨论线性规划, 而仅仅涉及两个问题: 一是给出单纯形法求解线性规划的算法步骤; 二是给出对偶理论中一个十分漂亮的结果: 原问题与对偶问题都有最优解的充要条件是它们都有可行解.

我们知道, 任何一个线性规划都可以化为如下的标准形式:

$$\begin{cases} \min \boldsymbol{c}\boldsymbol{x}^{\mathrm{T}} \\ \text{ s.t. } \boldsymbol{A}\boldsymbol{x}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}} \\ \quad\ \ \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases} \tag{LP}$$

其中 $\boldsymbol{c},\boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{A} \in \mathbb{R}^{m\times n}$. 矩阵 $\boldsymbol{A}$ 称为约束矩阵 (constraint matrix), 方程组 $\boldsymbol{A}\boldsymbol{x}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}$ 称为约束方程组 (system of constraint equations), $\boldsymbol{x} \geqslant \boldsymbol{0}$ 称为非负约束 (nonnegative constraint). 并且我们总是假设 $n \geqslant m \geqslant 1$, 且 $\operatorname{rank}\boldsymbol{A} = m$, 即 $\boldsymbol{A}$ 为行满秩矩阵.

首先给出线性规划解的一系列概念.

定义 0.2.1　称满足约束方程组及非负约束的向量 $\boldsymbol{x}$ 为可行解 (feasible solution), 可行解的全体称为可行域 (feasible region), 记作 K, 即

$$K = \{\boldsymbol{x} | \boldsymbol{A}\boldsymbol{x}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}, \boldsymbol{x} \geqslant \boldsymbol{0}\};$$

称使得目标函数 $\boldsymbol{c}\boldsymbol{x}^{\mathrm{T}}$ 在 K 上取得最小值的可行解为最优解 (optimal solution), 并称最优解对应的目标函数值为最优值 (optimal value).

定义 0.2.2　在问题 (LP) 中, $\boldsymbol{A}$ 的任一个 m 阶满秩子方程 $\boldsymbol{B}$ 称为基 (basis), $\boldsymbol{B}$ 中 m 个线性无关的列向量称为基向量 (basic vectors); $\boldsymbol{x}$ 中与 $\boldsymbol{B}$ 对应的分量称为关于 $\boldsymbol{B}$ 的基变量, 基变量构成的向量记作 $\boldsymbol{x_B}$; 其余的变量称为关于 $\boldsymbol{B}$ 的非基变量, 非基变量构成的向量记为 $\boldsymbol{x_N}$; 不失一般性, 假设 $\boldsymbol{x}$ 的前 m 个分量为基变量, 后 $n-m$ 个分量为非基变量, 如果让

$$\boldsymbol{x_N} = \boldsymbol{0}, \quad \boldsymbol{x_B} = (\boldsymbol{B}^{-1}\boldsymbol{b}^{\mathrm{T}})^{\mathrm{T}},$$

那么 $\bar{\boldsymbol{x}} = (\boldsymbol{x_B}, \boldsymbol{x_N})$ 是 $\boldsymbol{A}\boldsymbol{x}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}}$ 的一个解, 称为基本解 (basic solution); 若 $\boldsymbol{B}^{-1}\boldsymbol{b}^{\mathrm{T}} \geqslant \boldsymbol{0}$, 则称相应的基本解 $\bar{\boldsymbol{x}}$ 为问题 (LP) 的关于基 $\boldsymbol{B}$ 的一个基本可行解 (basic feasible solution), 相应的基 $\boldsymbol{B}$ 称为问题 (LP) 的可行基 (feasible basis).

下面给出单纯形法的算法步骤:

Step 1　找一个可行基 $\boldsymbol{B} = (\boldsymbol{p}_{j_1}^{\mathrm{T}}, \boldsymbol{p}_{j_2}^{\mathrm{T}}, \cdots, \boldsymbol{p}_{j_m}^{\mathrm{T}})$, 求出关于基 $\boldsymbol{B}$ 的单纯形表 $T(\boldsymbol{B})$, 如表 0.2.1 所示, 其中 $\boldsymbol{c_B}$ 为目标函数中基变量的系数构成的 m 维向量, $(\pi_1, \pi_2, \cdots, \pi_n) = \boldsymbol{c_B}\boldsymbol{B}^{-1}\boldsymbol{A} - \boldsymbol{c}$.

表 0.2.1 关于基 $\boldsymbol{B}$ 的单纯形表 $T(\boldsymbol{B})$

	x_1	$\cdots$	x_r	$\cdots$	x_n	
x_{j_1}	b_{11}	$\cdots$	b_{1r}	$\cdots$	b_{1n}	b_{10}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{j_s}	b_{s1}	$\cdots$	b_{sr}	$\cdots$	b_{sn}	b_{s0}
$\vdots$	$\vdots$		$\vdots$		$\vdots$	$\vdots$
x_{j_m}	b_{m1}	$\cdots$	b_{mr}	$\cdots$	b_{mn}	b_{m0}
f	π_1	$\cdots$	π_r	$\cdots$	π_n	$\bar{f}$

我们再给出 $T(\boldsymbol{B})$ 的另一种形式, 如表 0.2.2 所示, 它告诉我们 $T(\boldsymbol{B})$ 是如何求出的.

表 0.2.2 矩阵形式的单纯形表 $T(\boldsymbol{B})$

	$\boldsymbol{x}$	
$\boldsymbol{x}_{\boldsymbol{B}}^{\mathrm{T}}$	$\boldsymbol{B}^{-1}\boldsymbol{A}$	$\boldsymbol{B}^{-1}\boldsymbol{b}^{\mathrm{T}}$
f	$\boldsymbol{c}_{\boldsymbol{B}}\boldsymbol{B}^{-1}\boldsymbol{A}-\boldsymbol{c}$	$\boldsymbol{c}_{\boldsymbol{B}}\boldsymbol{B}^{-1}\boldsymbol{b}^{\mathrm{T}}$

Step 2 若 $\pi_j \leqslant 0(j=1,2,\cdots,n)$, 则关于 $\boldsymbol{B}$ 的基本可行解 $\bar{\boldsymbol{x}}$ 就是问题 (LP) 的最优解, $\bar{f}$ 为最优解, 算法结束; 否则转 Step 3.

Step 3 若有 $\pi_r > 0$, 使 $T(\boldsymbol{B})$ 中 x_r 所对应的列非正, 即

$$(b_{1r}, b_{2r}, \cdots, b_{mr})^{\mathrm{T}} \leqslant \mathbf{0},$$

则问题 (LP) 无最优解, 算法结束; 否则转 Step 4.

Step 4 令

$$r=\min\{l|\pi_l=\max_{\pi_j>0}\pi_j\},$$

将 x_r 变为基变量, 称 x_r 为进基变量; 再令

$$j_s=\min\left\{j_k\left|\frac{b_{k0}}{b_{kr}}=\min_{b_{ir}>0}\frac{b_{i0}}{b_{ir}}\right.\right\},$$

将 x_{j_s} 变为非基变量, 称 x_{j_s} 为离基变量.

Step 5 进行下列 $\{s,r\}$ 旋转变换:

(1) 将 $T(\boldsymbol{B})$ 中第 s 行同除以 b_{sr} 作为新的第 s 行, 即

$$(s\text{ 行}) := \frac{1}{b_{sr}}\cdot(s\text{ 行});$$

(2) 将表中新的第 s 行乘以 $(-b_{ir})$ 加到第 i 行 $(i\neq s)$, 得到新的第 i 行, 即

$$(i\text{ 行}) := (i\text{ 行}) - b_{ir}\cdot(s\text{ 行}) \quad (i\in\{1,2,\cdots,m\}\backslash\{s\});$$

(3) 将表中新的第 s 行乘以 $(-\pi_r)$ 加到第 $m+1$ 行, 得到新的第 $m+1$ 行, 即

$$(m+1\text{ 行}) := (m+1\text{ 行}) - \pi_r \cdot (s\text{ 行}).$$

变换后得到新单纯形表 $T(\tilde{\boldsymbol{B}})$, 用 $\tilde{\boldsymbol{B}}$ 代替 $\boldsymbol{B}$, $T(\tilde{\boldsymbol{B}})$ 代替 $T(\boldsymbol{B})$, 即 $\boldsymbol{B} := \tilde{B}$, $T(\boldsymbol{B}) := T(\tilde{\boldsymbol{B}})$, 再转 Step 2.

下面来介绍对偶理论.

定义 0.2.3　对线性规划问题

$$\begin{cases} \min \boldsymbol{c}\boldsymbol{x}^{\mathrm{T}} \\ \text{ s.t. } \boldsymbol{A}\boldsymbol{x}^{\mathrm{T}} = \boldsymbol{b}^{\mathrm{T}} \\ \qquad \boldsymbol{x} \geqslant \boldsymbol{0} \end{cases}$$

其中 $\boldsymbol{c}, \boldsymbol{x} \in \mathbb{R}^n, \boldsymbol{b} \in \mathbb{R}^m, \boldsymbol{A} \in \mathbb{R}^{m\times n}$. 构造另一个线性规划问题

$$\begin{cases} \max \boldsymbol{b}\boldsymbol{y}^{\mathrm{T}} \\ \text{ s.t. } \boldsymbol{A}^{\mathrm{T}}\boldsymbol{y}^{\mathrm{T}} \leqslant \boldsymbol{c}^{\mathrm{T}} \\ \qquad \boldsymbol{y} \geqslant \boldsymbol{0} \end{cases}$$

其中 $\boldsymbol{y} \in \mathbb{R}^m$. 我们称后者为前者的对偶问题 (dual problem), 称前者为原问题 (prime problem), 也称这两个问题互为对偶问题.

对一个一般的线性规划问题

$$\begin{cases} \min \boldsymbol{c}_1\boldsymbol{x}_1^{\mathrm{T}} + \boldsymbol{c}_2\boldsymbol{x}_2^{\mathrm{T}} \\ \text{ s.t. } \boldsymbol{A}_{11}\boldsymbol{x}_1^{\mathrm{T}} + \boldsymbol{A}_{12}\boldsymbol{x}_2^{\mathrm{T}} \geqslant \boldsymbol{b}_1^{\mathrm{T}} \\ \qquad \boldsymbol{A}_{21}\boldsymbol{x}_1^{\mathrm{T}} + \boldsymbol{A}_{22}\boldsymbol{x}_2^{\mathrm{T}} = \boldsymbol{b}_2^{\mathrm{T}} \\ \qquad \boldsymbol{x}_1 \geqslant \boldsymbol{0} \end{cases} \tag{P}$$

其中 $\boldsymbol{A}_{ij}$ 是 $m_i \times n_j$ 矩阵, $\boldsymbol{b}_j$ 为 m_i 维向量, $\boldsymbol{c}_j$ 为 n_j 维向量, $\boldsymbol{x}_j$ 也是 n_j 维向量, $i = 1, 2; j = 1, 2$. 我们不难求得其对偶问题为

$$\begin{cases} \max \boldsymbol{b}_1\boldsymbol{y}_1^{\mathrm{T}} + \boldsymbol{b}_2\boldsymbol{y}_2^{\mathrm{T}} \\ \text{ s.t. } \boldsymbol{A}_{11}^{\mathrm{T}}\boldsymbol{y}_1^{\mathrm{T}} + \boldsymbol{A}_{21}^{\mathrm{T}}\boldsymbol{y}_2^{\mathrm{T}} \leqslant \boldsymbol{c}_1^{\mathrm{T}} \\ \qquad \boldsymbol{A}_{12}^{\mathrm{T}}\boldsymbol{y}_1^{\mathrm{T}} + \boldsymbol{A}_{22}^{\mathrm{T}}\boldsymbol{y}_2^{\mathrm{T}} = \boldsymbol{c}_2^{\mathrm{T}} \\ \qquad \boldsymbol{y}_1 \geqslant \boldsymbol{0} \end{cases} \tag{D}$$

关于问题 (P) 和问题 (D) 有下面的结论.

定理 0.2.1　问题 (P) 和问题 (D) 都有最优解当且仅当它们都有可行解.

证明　只需证明充分性.