运筹与管理科学丛书 7

动态合作博弈

高红伟 〔俄〕彼得罗相 著

科学出版社
www.sciencep.com

运筹与管理科学丛书 7

动态合作博弈

高红伟 〔俄〕彼得罗相 著

国家自然科学基金(70571040, 70711120204, 70871064)
山东省研究生教育创新计划项目(SDYC08045) 资助
青岛大学学术专著出版基金(第七批)

科学出版社
北京

内 容 简 介

本书系统地介绍了零和博弈、非零和博弈、多阶段博弈以及微分博弈的基础理论，针对经典的特征函数型合作博弈、多阶段合作博弈以及合作微分博弈展开深入系统的研究，建立了比较完整的动态合作博弈的理论框架. 通过运用分配补偿程序的概念给出了各类动态合作博弈的建立具有动态稳定性(时间一致性)以及强动态稳定性解的方法. 书中介绍了大量的博弈模型，其中多数是首次出现在中文版博弈论专业书籍中.

本书的结构安排适应多层次的读者群体. 前三章可以作为基础数学、应用数学以及经济管理等专业本科生对策理论的基础教材，也可以作为博弈论初学者的入门教材；第4~6章可以作为研究生层次动态合作博弈的教材，经济、管理等领域的研究人员也能从中获得解决实际应用问题的灵感. 此外，如果把本书各章中针对追逃对策所进行的研究独立剥离出来，则在军事科技领域也具有一定的参考价值.

图书在版编目(CIP)数据

动态合作博弈/高红伟，〔俄〕彼得罗相著. —北京：科学出版社，2009

(运筹与管理科学丛书；7)

ISBN 978-7-03-023873-3

I. 动… II. ①高… ②彼… III. 合作对策 IV. O225

中国版本图书馆CIP数据核字(2009) 第001449号

责任编辑：赵彦超／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

*

2009年3月第 一 版 开本：B5(720×1000)

2009年3月第一次印刷 印张：26 1/4

印数：1—2 500 字数：510 000

定价：69.00元

(如有印装质量问题，我社负责调换〈路通〉)

《运筹与管理科学丛书》序

运筹学是运用数学方法来刻画、分析以及求解决策问题的科学. 运筹学的例子在我国古已有之, 春秋战国时期著名军事家孙膑为田忌赛马所设计的排序就是一个很好的代表. 运筹学的重要性同样在很早就被人们所认识, 汉高祖刘邦在称赞张良时就说道: "运筹帷幄之中, 决胜千里之外."

运筹学作为一门学科兴起于第二次世界大战期间, 源于对军事行动的研究. 运筹学的英文名字 Operational Research 诞生于 1937 年. 运筹学发展迅速, 目前已有众多的分支, 如线性规划、非线性规划、整数规划、网络规划、图论、组合优化、非光滑优化、锥优化、多目标规划、动态规划、随机规划、决策分析、排队论、对策论、物流、风险管理等.

我国的运筹学研究始于 20 世纪 50 年代, 经过半个世纪的发展, 运筹学队伍已具相当大的规模. 运筹学的理论和方法在国防、经济、金融、工程、管理等许多重要领域有着广泛应用, 运筹学成果的应用也常常能带来巨大的经济和社会效益. 由于在我国经济快速增长的过程中涌现出了大量迫切需要解决的运筹学问题, 因而进一步提高我国运筹学的研究水平、促进运筹学成果的应用和转化、加快运筹学领域优秀青年人才的培养是我们当今面临的十分重要、光荣、同时也是十分艰巨的任务. 我相信,《运筹与管理科学丛书》能在这些方面有所作为.

《运筹与管理科学丛书》可作为运筹学、管理科学、应用数学、系统科学、计算机科学等有关专业的高校师生、科研人员、工程技术人员的参考书, 同时也可作为相关专业的高年级本科生和研究生的教材或教学参考书. 希望该丛书能越办越好, 为我国运筹学和管理科学的发展做出贡献.

袁亚湘

2007 年 9 月

前　言

1944 年, 著名数学家 John von Neumann 和经济学家 Oskar Morgenstern 出版的《博弈论与经济行为》标志着对策论作为一门独立学科的诞生. 为了规范国内学术术语的使用, 钱学森先生曾经建议将该学科称为 "对策论", 不过在经济和管理科学等领域, 许多学者更喜欢称之为 "博弈论".

本书以动态合作博弈理论为主线, 但作者并没有对经典合作理论介绍到细致入微的程度, 而是希望引入并介绍尽可能多的对策类型, 尤其是现有中文版对策论文献中不多见的对策模型, 使读者有办法处理所遇到的各类问题. 为此, 本书删减了经典合作对策理论中的许多基础内容, 如谈判集、核 (kernel)、核仁 (nucleolus)、均衡对策、NTU 对策等. 删减的原则是为了适应本书推出的动态合作博弈理论体系, 做到够用即可, 同时希望尽量使读者开卷之后保持读下去的兴趣. 本书基础理论部分的内容主要引自 Vorobjev 和彼得罗相 (Leon A Petrosyan) 教授的著作.

许多读者是从经济博弈论开始接触非合作均衡理论体系的, 本书第 4 章采用严谨的数学语言介绍动态多阶段对策, 这对于具有一定的数学功底并且熟悉均衡理论的读者会起到开拓视野的作用. 第 5 章是在此基础上关于离散动态合作对策的研究工作, 主要的特色是以统一的观点来研究动态对策过程, 即不仅包含最大化自己支付的个体理性行为和完全合作行为, 还要考虑到各种程度的合作的可能, 以及这种合作程度根据对策进程所产生的变化. 第 6 章的模型也许比合作方式本身更具有亲和力.

本书关于时间离散或连续类型的动态对策的介绍还有另外一条主线, 在第 7, 8 章中体现得较为明显, 那就是几乎所有的理论都会围绕着追逃对策展开, 读懂它们并不困难. 因此, 本书在军事科技领域应该具有一定的参考价值.

关于时间一致性的研究是一个过去甚至现在都颇具争议的研究领域. 2002 年 8 月, 作者在一次与 Selten 教授共进晚餐的过程中, 他曾当面委婉地表露对这个方向的异议. 2004 年诺贝尔经济学奖授予挪威经济学家 Finn E Kydland 和美国经济学家 Edward C Prescott, 以表彰他们在动态宏观经济学方面做出的杰出贡献, 使得这方面的争论多少有所缓解. 事实上, 时间一致性解的观点是由彼得罗相和 Kydland 及 Prescott 几乎同时 (1977 年) 但是分别独立提出并深入研究的. 彼得罗相教授领导的研究团队沿着这个方向进行了长达 30 年不懈的研究, 在这个领域达到巅峰

的标志是 2006 年与杨荣基教授合著并在 Springer-Verlag 出版的专著 *Cooperative Stochastic Differential Games*.

此外, 2007 年初, 由中国市场出版社出版了一部由彼得罗相教授、杨荣基教授以及李颂志博士合著同样用中文撰写的《动态合作 —— 尖端博弈论》, 我和我的研究生为此书进行了校稿. 这部著作的许多读者可能对书中复杂庞大的算式望而却步, 而事实上, 在合作随机微分对策背景之下建立并运用分配补偿程序的概念是理论本身最大的难点所在. 但是, 结合本书在离散的多阶段对策以及一般形式的微分对策 (主要是追逃微分对策) 的框架内针对分配补偿程序所进行的研究, 实际上使得上述难点演变成自然的结果, 而动态稳定解 (时间一致性) 的理论体系得以完整的建立.

作为中国运筹学会对策论专业委员会的主任, 作者多次在青岛大学主办博弈论及其应用方面的学术会议, 其中包括学会的学术年会以及 2002 年国际数学家大会对策论及其应用卫星会议等. 由于举办 2002 年国际数学家大会的缘故, 这一年有很多大师级的科学家来到中国, 如物理学家史蒂芬 • 威廉 • 霍金 (Stephen William Hawking)(被认为是继爱因斯坦之后最杰出的理论物理学家), 邀请霍金来华的是数学界泰斗丘成桐教授. 2002 年 8 月, 诺贝尔经济学奖 (1994 年) 得主约翰 • 纳什教授夫妇接受作者的邀请首次来到中国. 此外, Aumann (2005 年诺贝尔经济学奖得主)、Selten (1994 年诺贝尔经济学奖得主) 以及 Shapley 教授同时出席会议, 并被授予青岛大学名誉教授.

感谢国家自然科学基金、国家留学基金、山东省研究生教育创新计划项目以及青岛大学学术专著出版基金的资助. 截至目前, 我已经先后四次主持国家自然科学基金项目, 先后两次获得国家留学基金的资助. 1991 年起师从彼得罗相教授在俄罗斯 St.Petersburg State University 公派留学并攻读博士学位, 在 2006 年底作为进修访问学者与他进行为期半年的合作研究. 本书的撰写工作正是我进修访问期间在彼得罗相教授的直接建议下开始的, 他对本书的核心内容、理论框架、主要特色提出了重要的指导性意见.

作者在此也想感谢 Vorobjev 教授对于对策论在中国的传播所做出的卓越贡献 (Vorobjev 教授是原苏联博弈论的奠基人, 20 世纪 50 年代应中国科学院的邀请来华讲授博弈论, 受到周恩来总理的接见, 他帮助中国培养了第一代博弈论领域的研究生), 当然不否认这里还有其他方面的原因, 事实上, Vorobjev 教授是彼得罗相教授的导师.

感谢我的父母、家人以及朋友, 他们的期望多年以来一直是我努力前行的动力. 还要感谢我的研究生王茜、张乐平、秦冬冬、王桂熙、杨慧敬、于琨、吕婷婷、宋

琳、李文文等, 他们在文稿输入、格式修正以及核对公式等许多方面付出了巨大的努力.

鉴于作者才疏学浅, 书中出现纰漏错误在所难免, 恳请读者批评指正.

高红伟

2008 年 12 月 18 日

目　录

第1章　矩阵对策

1.1　规范型二人零和对策的定义

定义 1.1.1　系统

$$\Gamma = (X, Y, K) \tag{1.1.1}$$

称为规范型二人零和对策, 其中 X 和 Y 是非空集合, $K : X \times Y \to R^1$ 是实值函数.

在对策 Γ 中, 元素 $x \in X$ 和 $y \in Y$ 分别称为局中人 1 和 2 的策略, 笛卡儿乘积 $X \times Y$ 的元素 (也就是策略对 (x, y), 其中 $x \in X$ 和 $y \in Y$) 称为局势, 函数 K 是局中人 1 的支付函数. 在局势 (x, y) 下局中人 2 的支付为 $(-K(x, y))$, 所以函数 K 称为对策 Γ 的支付函数, 而对策 Γ 称为零和对策. 运用对策术语, 为了给定对策 Γ, 必须定义局中人 1 和 2 的策略集 X 和 Y, 同时确定全体局势的集合 $X \times Y$ 上的支付函数 K.

对策 Γ 可以解释如下: 局中人同时并且独立地选择策略 $x \in X$ 和 $y \in Y$, 之后局中人 1 获得的支付为 $K(x, y)$, 而局中人 2 得到的支付为 $(-K(x, y))$.

定义 1.1.2　对策 $\Gamma' = (X', Y', K')$ 称为对策 $\Gamma = (X, Y, K)$ 的子对策, 如果 $X' \subset X$ 和 $Y' \subset Y$, 而函数 $K' : X' \times Y' \to R^1$ 是函数 K 在 $X' \times Y'$ 上的限制.

本章主要考察局中人的策略集合为有限集的二人零和对策.

定义 1.1.3　两个局中人均具有有限策略集合的二人零和对策称为矩阵对策.

假设对策 (1.1.1) 中局中人 1 共有 m 个策略. 对第一个局中人的策略集合 X 进行排序, 也就是在集合 $M = \{1, 2, \cdots, m\}$ 和 X 之间建立一个一一映射. 类似地, 如果局中人 2 共有 n 个策略, 可以在集合 $N = \{1, 2, \cdots, n\}$ 和 Y 之间建立一一映射. 那么, 对策 Γ 可以完全由矩阵 $A = \{a_{ij}\}$ 确定, 其中 $a_{ij} = K(x_i, y_j)$, $(x_i, y_j) \in X \times Y$, $i \in M$, $j \in N$(因此对策称为矩阵对策). 对策按如下方式进行, 局中人 1 选择行 $i \in M$, 局中人 2 选择列 $j \in N$, 此时局中人 1 和 2 是同时并且独立地进行选择. 因此局中人 1 获得支付 a_{ij}, 局中人 2 获得支付 $(-a_{ij})$. 如果支付是一个负数, 那么可以认为是局中人的实际损失.

记具有支付矩阵 A 的对策 Γ 为 Γ_A, 根据矩阵的维数, 称之为 $(m \times n)$ 对策. 如果根据上下文可以理解所考察的对策与哪一个矩阵对应, 那么可以去掉下标 A.

在矩阵对策中可以用不同的方式排列策略, 因此严格地说, 每一个顺序关系对应一个矩阵, 这样一个有限二人零和对策可以用行和列顺序不同的矩阵来描述.

例 1.1.1(城防)[1]　在文献中这个例子被称为 "布鲁多上校对策". 布鲁多上校有 m 个军团, 他的敌人有 n 个军团, 敌人需要保卫两个要地. 如果在一个要地上布鲁多上校实施进攻的军团在数量上占有优势, 他就能占领这个要地. 敌对双方需要在两个要地分布自己的军团.

定义布鲁多上校 (局中人 1) 在每个要地的支付. 在某个要地, 如果布鲁多上校比敌人 (局中人 2) 有更多的军团, 那么他的支付等于敌人的军团数加上 1(要地的占领等价于获得一个军团). 如果局中人 2 在某个要地比局中人 1 有更多的军团, 则局中人 1 失去他所有的军团并且再减去 1(因为失去这个要地); 如果在一个要地双方有相同的军团数, 那么双方打平并且各自的支付为零, 每个局中人的支付等于在两个要地支付的总和.

这个对策是零和的, 下面描述局中人的策略. 为确定起见, 假设 $m > n$, 局中人 1 有如下的策略: $x_0 = (m, 0)$, 即放置所有的军团在第一个要地; $x_1 = (m-1, 1)$, 即放置 $(m-1)$ 个军团在第一个要地, 放置 1 个军团在第二个要地; $x_2 = (m-2, 2)$, $\cdots$, $x_{m-1} = (1, m-1)$, $x_m = (0, m)$. 敌人 (局中人 2) 有如下的策略: $y_0 = (n, 0)$, $y_1 = (n-1, 1)$, $\cdots$, $y_n = (0, n)$.

假设局中人 1 选择策略 x_0, 局中人 2 选择策略 y_0, 计算局中人 1 在这种局势下的支付 a_{00}. 由于 $m > n$, 在第一个要地局中人 1 取胜, 他的支付为 $n+1$(1 表示拥有这个要地), 在第二个要地上双方打平, 因此 $a_{00} = n+1$. 再计算 a_{01}, 由于 $m > n-1$, 那么在第一个要地局中人 1 的支付为 $n-1+1=n$, 在第二个要地局中人 2 取胜, 因此局中人 1 在这个要地的损失为 1, 这样 $a_{01} = n-1$. 类似地可以得到 $a_{0j} = n-j+1-1 = n-j$, $1 \leqslant j \leqslant n$. 如果 $m-1 > n$, 则 $a_{10} = n+1+1 = n+2$, $a_{11} = n-1+1 = n$, $a_{1j} = n-j+1-1-1 = n-j-1$, $2 \leqslant j \leqslant n$. 一般情况下, 对任意的 m 和 n, 元素 a_{ij}, $i = \overline{0,m}$, $j = \overline{0,n}$ 以及支付矩阵可以如下计算:

$$a_{ij} = K(x_i, y_j) = \begin{cases} n+2, & m-i > n-j, i > j \\ n-j+1, & m-i > n-j, i = j \\ n-j-i, & m-i > n-j, i < j \\ -m+i+j, & m-i < n-j, i > j \\ j+1, & m-i = n-j, i > j \\ -m-2, & m-i < n-j, i < j \\ -i-1, & m-i = n-j, i < j \\ -m+i-1, & m-i < n-j, i = j \\ 0, & m-i = n-j, i = j \end{cases}$$

当 $m=4$ 和 $n=3$ 时, 考察所有的可能局势, 得到对策的支付矩阵 A 为

$$A=\begin{array}{c} \\ x_0 \\ x_1 \\ x_2 \\ x_3 \\ x_4 \end{array}\begin{array}{c} \begin{array}{cccc} y_0 & y_1 & y_2 & y_3 \end{array} \\ \begin{bmatrix} 4 & 2 & 1 & 0 \\ 1 & 3 & 0 & -1 \\ -2 & 2 & 2 & -2 \\ -1 & 0 & 3 & 1 \\ 0 & 1 & 2 & 4 \end{bmatrix} \end{array}$$

例 1.1.2(偏差对策)[2] 局中人 1 和 2 从集合 $\{1,\cdots,n\}$ 中选择整数 i 和 j, 此时局中人 1 赢得 $|i-j|$, 对策是零和的. 支付矩阵是 $(n\times n)$ 方阵, 其中 $a_{ij}=|i-j|$. 当 $n=4$ 时, 支付矩阵有如下形式:

$$A=\begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \end{array}\begin{array}{c} \begin{array}{cccc} 1 & 2 & 3 & 4 \end{array} \\ \begin{bmatrix} 0 & 1 & 2 & 3 \\ 1 & 0 & 1 & 2 \\ 2 & 1 & 0 & 1 \\ 3 & 2 & 1 & 0 \end{bmatrix} \end{array}$$

例 1.1.3(离散型决斗对策)[2] 局中人迎面走 n 步互相靠近, 局中人在每走一步后可能射击也可能不射击, 但是在对策进程中每个局中人仅仅可以射击一次. 设局中人在走了 k 步之后射击并击中对方的概率为 $k/n\,(k\leqslant n)$.

局中人 1 或 2 的策略是在第 $i(j)$ 步作出射击的决定. 假设 $i<j$, 并且局中人 1 在第 i 步决定进行射击, 局中人 2 在第 j 步决定进行射击. 局中人 1 的支付 a_{ij} 由下式确定:

$$a_{ij}=\frac{i}{n}-\left(1-\frac{i}{n}\right)\frac{j}{n}=\frac{n\,(i-j)+ij}{n^2}$$

这样, a_{ij} 是击中对方和被对方击中的概率之差. 在 $i>j$ 的情况下, 局中人 2 首先射击并且 $a_{ij}=-a_{ji}$. 如果 $i=j$, 设 $a_{ij}=0$. 当 $n=5$ 并且用 25 乘以这个对策的支付矩阵, 得到

$$A=\begin{bmatrix} 0 & -3 & -7 & -11 & -15 \\ 3 & 0 & 1 & -2 & -5 \\ 7 & -1 & 0 & 7 & 5 \\ 11 & 2 & -7 & 0 & 15 \\ 15 & 5 & -5 & -15 & 0 \end{bmatrix}$$

例 1.1.4(攻击防卫对策) 设局中人 1 考虑攻击分别具有正的价值 $\tau_1>0,\cdots,\tau_n>0$ 的目标 $c_1,\cdots,c_n$ 之一, 局中人 2 保护这些目标之一. 我们假设, 如果被攻击

的目标 c_i 没有防卫, 那么它一定被毁坏 (局中人 1 赢得 τ_i), 有防卫的目标以概率 $1>\beta_i>0$ 被毁坏 (目标 c_i 以概率 $1-\beta_i>0$ 避免被攻击毁坏), 即局中人 1 赢得 $\beta_i\tau_i$, $i=1,2,\cdots,n$(平均).

选择进行攻击的目标 (局中人 1) 和防卫目标 (局中人 2) 的问题可以转化为矩阵对策, 其支付矩阵为

$$A=\begin{bmatrix} \beta_1\tau_1 & \tau_1 & \cdots & \tau_1 \\ \tau_2 & \beta_2\tau_2 & \cdots & \tau_2 \\ \vdots & \vdots & & \vdots \\ \tau_n & \tau_n & \cdots & \beta_n\tau_n \end{bmatrix}$$

例 1.1.5(离散型搜索对策)　有 n 个坑, 局中人 2 在 n 个坑中之一隐藏物体, 局中人 1 希望找到它. 在寻找第 i 个坑时局中人 1 付出的努力为 $\tau_i>0$, 在第 i 个坑中找到目标 (如果它被隐藏在那里) 的概率为 $0<\beta_i\leqslant 1$, $i=1,2,\cdots,n$. 如果找到目标, 局中人 1 获得收益为 α. 在其中隐藏和搜索物体的坑的编号是局中人的策略, 局中人 1 的支付等于期望收益与寻找目标时所付出的努力之差. 隐藏和搜索目标的问题可以转化为矩阵对策, 其支付矩阵为

$$A=\begin{bmatrix} \alpha\beta_1-\tau_1 & -\tau_1 & -\tau_1 & \cdots & -\tau_1 \\ -\tau_2 & \alpha\beta_2-\tau_2 & -\tau_2 & \cdots & -\tau_2 \\ \vdots & \vdots & \vdots & & \vdots \\ -\tau_n & -\tau_n & -\tau_n & \cdots & \alpha\beta_n-\tau_n \end{bmatrix}$$

例 1.1.6(搜索 "有噪声" 的目标)　假设局中人 1 对移动目标 (局中人 2) 进行搜索, 目的是为了发现它. 局中人 2 的目的则相反, 他总是尽量避免被找到. 局中人 1 可以分别以 $\alpha_1=1$, $\alpha_2=2$, $\alpha_3=3$, 而局中人 2 相应地可以 $\beta_1=1$, $\beta_2=2$, $\beta_3=3$ 的速度移动. 局中人 1 使用的侦查工具的作用距离依赖于对策参与者移动的速度, 并在下面的矩阵中给出

$$D=\begin{array}{c} \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{array}\begin{array}{c} \begin{array}{ccc} \beta_1 & \beta_2 & \beta_3 \end{array} \\ \begin{bmatrix} 4 & 5 & 6 \\ 3 & 4 & 5 \\ 1 & 2 & 3 \end{bmatrix} \end{array}$$

移动速度是局中人的策略, 我们采用搜索效率 $a_{ij}=\alpha_i\delta_{ij}$, $i=\overline{1,3}$, $j=\overline{1,3}$ 作为局中人 1 在局势 (α_i,β_j) 下的支付, 其中 δ_{ij} 是矩阵 D 中的元素. 在搜索-逃离时局中人速度选择的问题可以表达为矩阵对策, 具有如下的矩阵

$$A = \begin{array}{c} \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{array} \begin{array}{c} \begin{array}{ccc} \beta_1 & \beta_2 & \beta_3 \end{array} \\ \begin{bmatrix} 4 & 5 & 6 \\ 6 & 8 & 10 \\ 3 & 6 & 9 \end{bmatrix} \end{array}$$

1.2 最大最小策略和最小最大策略

考察一个二人零和对策 $\Gamma = (X, Y, K)$, 其中每一个局中人通过选择策略致力于最大化自己的支付. 局中人 1 的支付由函数 $K(x, y)$ 确定, 局中人 2 的支付为 $[-K(x, y)]$, 也就是说, 局中人的目的是截然相反的. 注意到, 局中人 1(2) 的支付由对策进程中形成的局势 $(x, y) \in X \times Y$ 所确定. 在每一局势下局中人的支付不仅仅依赖于他自己的选择, 同时也依赖于他的对手选择什么样的策略, 对手的目的和他是直接相对的. 因此在致力于获得可能的最大支付时, 每一个局中人必须考虑到对方的行为.

我们利用城防对策来解释. 如果局中人 1 希望获得最大的支付, 那么他应当采用策略 x_0(或 x_4). 在这种情形下, 如果局中人 2 使用策略 $y_0(y_3)$, 那么前者得到 4 个单位的支付. 但是, 如果局中人 2 使用策略 y_3(相应地 y_0), 则局中人 1 得到的支付为 0, 也就是说, 他失去 4 个单位. 类似的讨论同样可以针对局中人 2 进行.

在对策论中, 假设局中人的行为都是理性的, 也就是说, 在考虑到对方采用的是对他自己最有利行为的情况下局中人致力于获得最大支付. 局中人 1 能给自己确保什么呢? 假设局中人 1 选择策略 x, 那么在最坏的情形下他将获得 $\min\limits_{y} K(x, y)$, 因此, 局中人 1 总是可以给自己确保支付为 $\max\limits_{x}\min\limits_{y} K(x, y)$. 如果放弃达到极值的假设, 那么局中人 1 总是可以确保自己获得的支付无限接近于数值

$$\underline{v} = \sup_{x \in X} \inf_{y \in Y} K(x, y) \tag{1.2.1}$$

称之为对策的下值. 如果式 (1.2.1) 中的外侧极值达到, 那么数 $\underline{v}$ 同样被为最大最小值, 在最大化最小支付的基础上构建策略 x 的原则称为最大最小准则, 而根据这个准则所选择的策略 x 称为局中人 1 的最大最小策略.

针对局中人 2 可以进行类似的讨论. 假设局中人 2 选择策略 y, 那么在最坏的情形下他将失去 $\max\limits_{x} K(x, y)$, 因此局中人 2 总可以确保自己失去的不超过 $\min\limits_{y}\max\limits_{x} K(x, y)$,

$$\bar{v} = \inf_{y \in Y} \sup_{x \in X} K(x, y) \tag{1.2.2}$$

称为对策 Γ 的上值, 当式 (1.2.2) 中的外侧极值达到时, 同样称 $\bar{v}$ 为最小最大值. 在最小化最大损失的基础上构建策略 y 的原则称为最小最大准则, 而根据这个准则

选择的策略 y 称为局中人 2 的最小最大策略. 应该强调, 最小最大 (最大最小) 策略的存在性由式 (1.2.2)(或 (1.2.1)) 中外侧极值的可达到性决定.

假设给定 $(m\times n)$ 矩阵对策 Γ_A. 那么在式 (1.2.1) 和 (1.2.2) 中的极值是可达到的, 而对策的下值和上值分别为

$$\underline{v}=\max_{1\leqslant i\leqslant m}\min_{1\leqslant j\leqslant n}a_{ij} \tag{1.2.3}$$

$$\bar{v}=\min_{1\leqslant j\leqslant n}\max_{1\leqslant i\leqslant m}a_{ij} \tag{1.2.4}$$

对策 Γ_A 的最小最大值和最大最小值可以通过下面的图解得到

$$\left.\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}\begin{matrix} \min\limits_j a_{1j} \\ \min\limits_j a_{2j} \\ \vdots \\ \min\limits_j a_{mj} \end{matrix}\right\}\max_i\min_j a_{ij}=\underline{v}$$

$$\underbrace{\max_i a_{i1}\quad \max_i a_{i2}\quad \cdots\quad \max_i a_{in}}_{\min\limits_j\max\limits_i a_{ij}=\bar{v}}$$

因此在对策 Γ_A 中, 当支付矩阵为

$$A=\begin{pmatrix} 1 & 0 & 4 \\ 5 & 3 & 8 \\ 6 & 0 & 1 \end{pmatrix}$$

时, 下值 (最大最小值) $\underline{v}$ 和第一个局中人的最大最小策略 i_0 分别为 $\underline{v}=3$, $i_0=2$, 而上值 (最小最大值)$\bar{v}$ 和第一个局中人的最小最大策略 j_0 分别为 $\bar{v}=3$, $j_0=2$.

对于任意的二人零和对策 $\Gamma=(X,Y,K)$, 有下面的结论.

引理 1.2.1　在二人零和对策 Γ 中,

$$\underline{v}\leqslant\bar{v} \tag{1.2.5}$$

$$\sup_{x\in X}\inf_{y\in Y}K(x,y)\leqslant\inf_{y\in Y}\sup_{x\in X}K(x,y) \tag{1.2.6}$$

证明　假设 $x\in X$ 是局中人 1 的任意策略, 那么有

$$K(x,y)\leqslant\sup_{x\in X}K(x,y)$$

因此得到

$$\inf_{y\in Y}K(x,y)\leqslant\inf_{y\in Y}\sup_{x\in X}K(x,y)$$