

优 化 与 决 策

纳什均衡

——
动态博弈的初步讨论

王则柯

葛菲编

上海科学技术出版社

优 化 与 决 策

纳什均衡

——动态博弈的初步讨论

王则柯 葛 菲 编

上海科学技术出版社

图书在版编目 (CIP) 数据

纳什均衡：动态博弈的初步讨论/王则柯，葛菲编.
上海：上海科学技术出版社，2009.6
(优化与决策)
ISBN 978-7-5323-9755-6

I .纳... II.①王...②葛... III.对策论-通俗读物
IV.0225-49

中国版本图书馆CIP数据核字 (2009) 第017806号

上海世纪出版股份有限公司 出版、发行
上海科学技术出版社
(上海钦州南路71号 邮政编码 200235)
新华书店上海发行所经销
苏州望电印刷有限公司印刷
开本 889×1194 1/32 印张 4
字数：85千字
2009年6月第1版 2009年6月第1次印刷
印数：1-2 250
ISBN 978-7-5323-9755-6/N·280
定价：15.00元

本书如有缺页、错装或坏损等严重质量问题，
请向工厂联系调换

《优化与决策》丛书编委会

名誉主编 吴文俊 谷超豪

主 编 胡毓达

编 委 (以姓氏笔画为序)

王长钰 王兴华 王则柯 方伟武

石钟慈 史树中 刘源张 李 端

汪寿阳 张连生 陈光亚 茆诗松

姚恩瑜 袁亚湘 顾基发 徐利治

唐国春 章祥荪 越民义 韩继业

管梅谷 魏权龄

序

“人尽其能，物尽其用”，是人类进步的重要标志和社会发展的根本动力。

在现代社会中，小至个人事务的处理，大到国家政策的制定，无不需人们进行关于“人”和“物”的选优抉择，以求取好的结果。在科技日新月异和经济快速发展的 21 世纪，人们要日益面临各种愈来愈复杂的决策问题，因此，现代优化思想和科学决策知识，已是当今人们普遍需要具备的基本素养。

现代教育提倡对学生进行创新精神和综合能力的素质培养。在我国大中学教育中，让学生们了解某些现代优化方法和进行决策能力培养，也正是素质教育的重要内容。

为了向广大读者普及最优化和科学决策的思想和方法，在中国运筹学会及其决策科学分会、数学规划分会和排序分会，中国系统工程学会，中国数学会计算数学分会以及上海运筹学会的倡议和支持下，我们邀请了在相应领域卓有成就的有关专家，撰写了这套《优化与决策》系列丛书。

这套丛书具有以下特点：

选题实用求新 本丛书的重要特色是内容的实用性。各选题在扩大知识的同时，均注重联系实际结合应用展开讨论。不论是定量或定性的决策问题，进行选优建模和效益分析一般要归为用数量刻画和作数值计算，因此，数学是这套丛书各选题的基本工具。但是，与以往作为中学数学教科书内容补充的科普性数学读

物或抽象的数学专著不同,本丛书强调综合运用数学和有关知识去解决现实中的应用问题。另外,丛书的选题既考虑其内容是具代表性的,同时也注意对新领域和某些发展中问题的介绍。

表述浅出深入 本丛书着力于用通俗易懂的方式引导读者掌握现代优化和决策知识。书中特配置了形象的图画以帮助加深读者对内容的理解。我们计划,具有高中数学基础的读者,即可读懂其中的基本内容。但是,为了每一选题的系统性和完整性,也不放弃对一些最基本和著名理论结果的介绍。因此,主要想了解思想方法和借鉴应用手段的读者,阅读时可以略去其中某些理论和证明部分,而不会影响对主要内容的理解。然而,对于具相当数学素养并对理论结果同样有兴趣者,这些较深入的内容对他们是有价值的,其中有些结果即使对于同行学者也将具有重要的参考意义。因此,不同层次的读者,阅读本丛书后均会有所得益。

这套丛书的读者面是多层次和极广泛的,它既适用于各行业管理者,各级行政公务员,广大科技工作者,以及各专业大学生、研究生和教师们阅读,同时也可作为大专和高中学生的选读材料或课外读物。

写作这套既具科普性又基于一定理论分析的现代应用丛书,对于丛书作者是一种新的尝试。本丛书从酝酿组织、确定选题,直至现在与读者见面,曾经历了较长的时间。许国志院士和俞文鱿

教授生前曾积极参与出版本丛书的策划,并热情承担了写作任务,可惜未及如愿。值得一提的是,各位作者对分担撰写选题的内容都进行了精心选择和安排。特别是,许多作者专业造诣精深,但写作科普著作则是第一次,因此在可读性方面曾倾注了许多心血。对于他们的这种认真和奉献精神,谨致以衷心感谢和崇高的敬意!鉴于著述这套丛书对多数作者是一件新的工作,其中难免或有不足之处,期待读者们不吝指正。

最后,我们对吴文俊院士和谷超豪院士关心和支持本丛书的出版,并乐于担任名誉主编致以诚挚的谢忱!同时,感谢上海科学技术出版社对出版这套丛书所作的一系列努力。

胡毓达

2006年9月19日

前 言

在中文文献中,英文术语 game theory 既翻译作对策论,又翻译作博弈论。经济学界说博弈论比较多,但是也有其他一些领域,特别是这个丛书的许多作者和读者所在的领域,一些人习惯对策论的说法。至于大众媒体,自然说博弈论比较多,毕竟在以经济建设为中心的时代,经济学的语言和理念对于民众的影响更大。本书则按照主编的建议,对于作为一门学问的 game theory,采用对策论的说法,但是对于具体的 game,笔者觉得还是把它叫做博弈比较好。

对策论,主要研究决策主体的行为发生相互作用的时候主体的决策以及这些决策之间的均衡问题。也就是说,研究当一个主体(例如一个人、一个企业、一支军队或一个国家)的选择受到其他主体的选择的影响,而且反过来又影响到其他主体的选择时的决策问题以及均衡问题。

描述一个博弈,至少需要三个最基本的要素:一是博弈的参与人,也可以叫做局中人,二是这些参与人各自可以采取的策略,三是这些参与人在他们的各种策略选择对局之下之博弈所得,叫做他们参与博弈的支付。进一步,参与人同时做出策略选择的博弈,叫做静态博弈;参与人先后做出策略选择的博弈,叫做动态博弈。

纳什均衡是对策论的核心概念,刻画互动背景下的稳定对局。这是一种全新的“最优”理念,是一种依赖的“最优”理念。半个世

纪以来,纳什均衡的概念在许多学科都产生很大的影响,甚至是革命性的影响。本书主要论述动态博弈的纳什均衡。

为了轮廓式地向读者展现对策论的整个理论体系,开篇还是花了两节的笔墨,粗线条地介绍主要的静态博弈模型及其纳什均衡,作为全书的热身,让读者首先在静态框架内熟悉对策论的基础概念与方法,然后在后面再引入更深一层的动态博弈纳什均衡的讨论。

关于动态博弈,首先介绍动态博弈的博弈树表示方法,一般地讨论动态博弈的纳什均衡。然后,将进一步讨论这些纳什均衡的稳定性,并介绍子博弈精炼纳什均衡,以及寻求纳什均衡的倒推法。紧接着,还会讨论对策意味最浓的重复博弈与所谓策略性行为,深入到竞争机制这一层次来考察动态博弈。最后,本书还将介绍日常生活中应用非常普遍的霍特林模型及其两个系列的纳什均衡解。

本书写作的特点,是从故事开始,每一个内容都由一个博弈小故事引入,让读者从趣味盎然的小故事中领会抽象的对策论概念与模型,以便读者更好地掌握对策论的规律与方法。

本书的主要素材,取自笔者在中信出版社出版的《人人博弈论》,但是按照本丛书的要求做了改写。葛菲同学参与了改写的讨论和文字整理。

王则柯 识于戊子年秋

目 录

前言…… I

1. 静态博弈及其纳什均衡…… 1

2. 混合策略与反应函数交叉法…… 23

3. 动态博弈与博弈树…… 33

4. 子博弈精炼纳什均衡…… 51

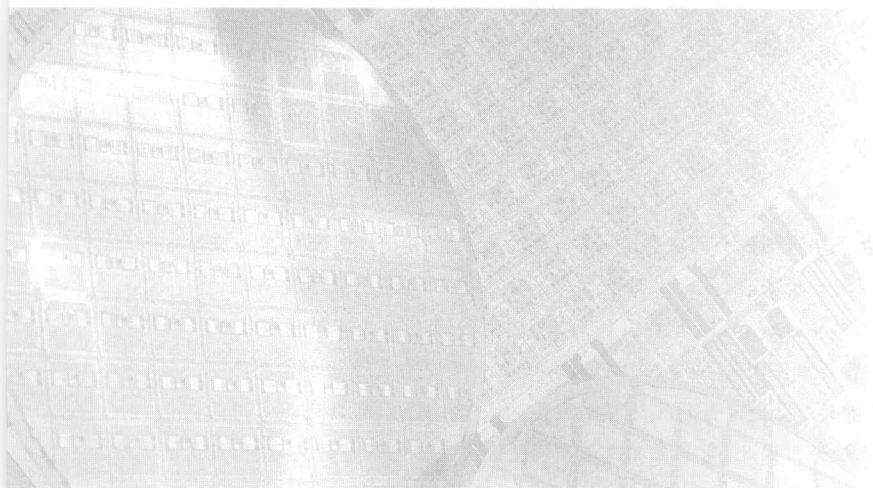
5. 重复博弈…… 65

6. 策略性行动…… 79

7. 霍特林模型…… 95

参考文献…… 114

静态博弈 及其纳什均衡





在争夺市场份额的激烈竞争中,提高商品价格一方面使得单个商品的利润会提高,但另一方面却意味着会丢掉不能忍受价格提高而选择其他商品的顾客,两者共同作用下,总体利润有可能不升反降。然而,如果那些替代商品也一同提价,那么原来经常光顾的老顾客也许不会“移情别恋”,会考虑到既然其他同类商品也在提价,那么还是常用的牌子比较好。因此,商家在考虑是否调整价格并如何调整的问题时,就不得不把竞争对手可能的决策也纳入自己的考虑范围内。

这种把对方有可能采取的策略也作为考虑因素进而作出决策的情况,在商场、战场乃至日常生活中都广泛存在着,而正文提到的“价格大战”则是静态博弈最典型的例子之一。

对策论是研究博弈参与人如何通过策略互动的考量作出各自的决策的学问,讨论的前提是理性人假设。理性人假设说的是当事人都有明确的目标,其行为模式是在自身能力和环境因素的约束之下,努力奔着自己的目标而去。理性人假设的模式本来非常宽泛,与备受误解的“自私自利”有很大距离。例如“毫不利己专门利人”的人,每次利人,就得到享受,增加满意程度;例如父母忍饥挨饿看着孩子享受美食,仍然非常受用,比自己和孩子平等消费要更加满意。

谈到理性人努力“奔着自己的目标而去”,可以概念地想像他是沿着一条虚拟的道路奔向他的目标,往前走得越远,就越满意。这样,按照沿着这条想像的道路往前的“里程数目”,理性人这种享受或者受用的程度,就能够用“里程”数值的大小表达出来,越大表示越满意。至此,按照在这条虚拟的道路上往前走得尽可能远,参与人的行为模式,就可以概括为追求自身目标利益的最大化,即在自身能力和环境因素的约束之下,追求上述“满意程度”数值的最大化。

对策论的核心概念是博弈的纳什均衡。本节将从囚徒困境这个经典的对策论模型说起,介绍静态博弈的三大类型及其纳什均

衡的分析方法,加深读者对博弈三要素和纳什均衡等概念的理解,作为进入动态博弈纳什均衡的讨论之前的热身。

囚徒困境与劣势策略消去法

先从对策论最深刻的模型囚徒困境(prisoner's dilemma)说起。

一次严重的纵火案发生后,警察在现场抓到甲乙两个犯罪嫌疑人。事实上正是他们为了报复而一起放火烧了这个仓库,但是警方没有掌握足够的证据。于是,警方把他们隔离囚禁起来,要求坦白交代。如果他们都承认纵火,每人将入狱三年;如果他们都不坦白,由于证据不充分,他们每人将只入狱一年;如果一个抵赖而另一个坦白并且愿意作证,那么抵赖者将入狱五年,而坦白者将得到宽大释放,免于刑事处罚。有理由以入狱时间长短的相反数刻画两个犯罪嫌疑人的满意程度,这样,他们面临的对策格局就如图1-1所示,每个格子中左下角的数字是甲的博弈所得,右上角的数字是乙的博弈所得。注意,现在这些数字都不是正数。

		乙	
		坦白	抵赖
甲	坦白	-3, -5	-3, 0
	抵赖	0, -1	-5, -1

图 1-1 囚徒困境

表述一个博弈(game)的基本要素有三个:

- (1) 参与人或者局中人(players);
- (2) 他们可选择的行动(actions)或策略(strategies);

(3) 所有可能的对局的博弈结果,用参与人在相应对局下的博弈所得来表示,这个博弈所得,叫做支付(payoffs)。这里注意,“支付”要理解为因为他们参与博弈所得到的支付,而不是他们付出的支付。

这里关键的一个概念是“对局”或者“策略对局”:每个参与人都选定了自己的一个策略,就确定博弈的一个对局,或者说策略对局。

在囚徒困境博弈中,博弈的两个参与人是犯罪嫌疑人甲和犯罪嫌疑人乙;他们可以选择的策略都是同样的两个,即坦白和抵赖,所以这个博弈的策略对局一共有4个,分别是(坦白,坦白)对局,(坦白,抵赖)对局,(抵赖,坦白)对局和(抵赖,抵赖)对局,括号中头一个策略是参与者甲的选择,第二个策略是参与者乙的选择;甲在各种对局下之博弈所得,是相应格子里面左下角的数字,乙在各种对局下之博弈所得,是相应格子里面右上角的数字。

这种用矩阵形式的表格表示两个参与人的博弈所得的做法,来自对策理论的一位先驱学者谢林(T. C. Schelling),按照矩阵表格行和列的不同位置关系,其中犯罪嫌疑人甲又叫做“行参与者”(row player),犯罪嫌疑人乙则叫做“列参与者”(column player)。从代数中早已熟悉的矩阵,每个位置一个数,一般用两条弧线或者方括号那样的两条竖线括住。现在这种博弈矩阵或者对策矩阵,每个格子位置有行参与人的支付和列参与人的支付这样两个数字,这样写成的表格,叫做双矩阵(bi-matrix)。

在囚徒困境中,如果两个嫌疑犯都是只为自己目标利益打算的所谓理性主体人(rational agent),两位犯罪嫌疑人博弈可能的结果会怎样呢?要是乙抵赖,那么如果甲坦白甲就可以得到宽大释放;要是乙坦白,那么如果甲也坦白的话甲要坐三年牢,但是如果

甲抵赖的话甲可要坐五年牢。可见对于甲来说,不管乙采取什么策略,他坦白自己总是比较有利的。所以两相比较,坦白是他的全面的严格的优势策略。

全面,指的是不论对方采取哪个策略,当事人的这个策略比他别的策略总是显示优势:对方坦白,我坦白比抵赖好;对方抵赖,我也是坦白比抵赖好。严格,指的是这个优势策略给当事人带来的支付确实要好一些:对方坦白,我坦白得-3确实比抵赖得-5好;对方抵赖,我坦白得0也确实比抵赖得-1好。这里,严格是说:-3不仅仅是不差于-5,而且是严格好于-5;0不仅仅是不差于-1,而且是严格好于-1。“全面的严格的优势策略”说起来拗口,可以就简称为严格优势策略(strictly dominant strategy)。优势劣势是比较而言的。在这个博弈中,既然坦白是严格优势策略,那么严格劣于它的抵赖策略就是相应的严格劣势策略(strictly dominated strategy)。

同样道理,坦白也是犯罪嫌疑人乙的全面的严格的优势策略,抵赖是相应的严格劣势策略。

理性的主体人,不会采用对自己显然不利的严格劣势策略。所以在分析博弈可能的结局的时候,如果发现一个参与人可以选择的某个策略是他的严格劣势策略的时候,应该把参与人的严格劣势策略删去。下面图中一横一竖的两条粗实线,就代表把两个参与人各自的严格劣势策略删去。

这个博弈非常简单,每个参与人只有两个可选择的策略。当博弈双方的严格劣势策略都删去以后,就只剩下左上角一个策略对局了。于是得到“坦白,坦白”这个严格优势策略均衡,在这个策略对局之下,两个参与人之所得为(-3, -3)。注意,在(-3, -3)这样的写法中,第一个数字是表格左方博弈参与人之所得,第

二个数字是表格上方博弈参与人之所得(图 1-2)。

		乙	
		坦白	抵赖
甲	坦白	-3	-5
	抵赖	-5	-1

图 1-2 劣势策略消去法

经济学习惯把市场力量对峙的稳定结局,叫做市场均衡(equilibrium)。比方说苹果的市场,供不应求将驱使价格上升,供大于求将迫使价格下降,供求力量对峙的结果,会在某个价格水平达到市场供求的均衡,即需求量和供给量相等。但是像上面这样用删去劣势策略的方法得到的由双方的严格优势策略组成的对局,作为这个博弈的均衡,叫做严格优势策略均衡(equilibrium of strictly dominant strategies)。这种通过把严格劣势策略删去的方法寻求对局结果的方法,叫做严格劣势策略消去法。如果甲乙都有三四个甚至更多的策略选择,通常需要一次又一次地把严格劣势策略删去,才能最后得到一个严格优势策略均衡。一次又一次把严格劣势策略删去以寻求对局结果的方法,叫做严格劣势策略逐次消去法(iterated elimination of strictly dominated strategies)。

为什么两个嫌疑人都会选择坦白呢?那是因为每个嫌疑人都只关心自己的目标利益。不管对方的决策怎样,自己总是采取坦白策略会更好。这就促使双方都采取坦白策略。但是,如果双方勾结起来,都抵赖,那么双方都可以获得较轻的惩罚。对于博弈的双方来说,这样一种大家都采取抵赖策略的对局形势,叫做双赢