经济与管理复杂自适应系统：

个体行为、系统状态及演化

郑小京　著

科学出版社

经济与管理复杂自适应系统：个体行为、系统状态及演化

郑小京　著

本书受湖北省徐绪松复杂科学管理研究基金会资助

科学出版社

北京

内 容 简 介

本书对真实的经济与管理复杂自适应系统进行抽象建模，从时间长短、空间宏微观共四个维度上将经济与管理复杂自适应系统抽象成若干在不同时间尺度中相互作用的若干个子系统——局域，通过分析确定出系统中个体行为在群体中的最优决策策略及演化规律，以及个体行为对于系统整体状态影响特征。

本书适合经济与管理复杂系统领域的科研工作者及研究生阅读。

图书在版编目(CIP)数据

经济与管理复杂自适应系统：个体行为、系统状态及演化/郑小京著．—北京：科学出版社，2014

ISBN 978-7-03-040329-2

Ⅰ．①经… Ⅱ．①郑… Ⅲ．①经济管理—自适应控制系统 Ⅳ．①F2

中国版本图书馆 CIP 数据核字(2014)第 062892 号

责任编辑：李 楠 / 责任校对：彭立军
责任印制：阎 磊 / 封面设计：无极书装

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

北京市文林印务有限公司 印刷

科学出版社发行 各地新华书店经销

*

2014 年 8 月第 一 版 开本：720×1000 B5
2014 年 8 月第一次印刷 印张：17 3/4
字数：357 000

定价：76.00 元

(如有印装质量问题，我社负责调换)

序

复杂系统越来越受到科学界的关注，不断地有一些科学家将其研究对象置于复杂系统的研究框架中，并且得到了许多非常有趣的结论。这种研究思路彻底打破了经典还原论的禁锢，以新的视角分析问题，揭示了许多从未获得的有趣现象及相对深刻的结论，这在极大程度上推动了复杂系统的研究。

由于其错综复杂的关系及系统中人的“活性”，经济与管理应该是一个系统行为能够主动适应外部环境的特征及其变化且系统行为能够改变环境属性的复杂自适应系统，由于这一点，该系统被很多经济与管理科学家所认可。然而，近年来在该领域的研究成果，尤其是在国际上有显著影响力的成果锐减，这使得大多数科学家对复杂自适应系统这一理论产生了质疑。究其原因，大致可以归纳为以下几个方面。第一，经济与管理复杂自适应系统建模的难度很大。复杂自适应系统错综复杂的关系建立在积木机制的基础上，而系统的创造性主要来自于系统中各个主体之间的非线性相互作用及与环境之间的闭环相互作用关系，因此对其建模的难度很大。第二，对经济与管理复杂自适应系统的分析手段滞后。由于其分析的复杂性，目前大部分科学家采用计算实验的方法，构造各种不同的计算机实验模型对所研究的经济与管理复杂自适应系统进行建模，同时按照计算实验的方法运行该模型，得到对应系统的某些特征。然而这一方法的局限性在于对参数的依赖性很强，以及对揭示系统表现特征背后所掩藏规律的能力偏弱。目前其实缺乏实验之后理论上升华的一些成果，也需要通过数学的方法对该领域进行更为深刻的研究，从而获得更加深刻的结论，这是复杂自适应系统研究过程中较大的瓶颈之一。

综上所述，为了在复杂自适应系统研究过程中取得更大的突破，迫切需要理论与方法体系的比较深刻的结果给予支撑。郑小京博士这部关于经济与管理复杂自适应系统中 Agent 的行为及其演化这一领域的力作是人们所期待的有深度的专著。

这部著作的作者多年前就从事复杂自适应系统的研究，作为一本方法论的书，它解决了复杂自适应系统中的一些问题，其相关结果可以直接拿出来应用。作者从宏观与微观两个不同的时间维度分析了复杂自适应系统及其主体行为的特征，并且将这两个不同粒度的时间进行了很好的接续，充分体现了宏观与微观的结合，描述了该类系统的智能性、自治性和社会性等特征，也描述了随着环境变化时整个系统及其主体行为特征及其变化规律，同时还刻画了系统内主体行为对于外部环境所造成的影响，并且，其中大部分理论可以转化成更为直接的决策结果，这给经

济与管理的决策带来了科学依据。最令人关注的是,该专著所提供的复杂自适应系统解析的方法体系,能够给各种实际的经济与管理系统的研究提供直接的研究工具,这正是这部著作价值体现最大的地方。

这部专著,不仅概述了一系列经典的、重要的结果,同时也包括了郑小京博士近年来的研究成果,有些是首次发表。该书主要有以下内容:一、对经济与管理复杂自适应系统进行了比较深刻的描述;二、将复杂自适应系统在短的 time-scale 和长的 time-scale 两个粒度上进行分析,分别对应了复杂自适应系统结构固定与结构时变两种情况,分别描述了各自对应的系统中主体及系统自身的行为及性质,以及对应的最优策略及其演化的不变分布,这一分析结果特别引人注目;三、描述了在外部环境扰动时系统的变化特征,以及系统行为对于环境的影响这一复杂自适应系统固有特色下系统的特征,并将二者进行了有机的融合,这一结果是建立在管理对象与环境的相互影响基础上的,更加符合经济与管理系统的特征,具有特别高的价值;四、确定了复杂自适应系统的无标度特性,进而分析了复杂自适应系统在遇到随机攻击时表现出的鲁棒性及在蓄意攻击时表现出的脆弱性,并确定了系统逾渗的临界状态,作者通过这一分析确定了经济与管理复杂自适应系统突变的临界特征,对于经济与管理系统的最优决策具有重要的价值。

纵观全书,作者在坚实的数学基础之上,对复杂自适应系统给出了很有特色的分析。本书写作严谨,文字流畅,结果深刻,适用广泛,有关领域的广大科研工作者能够得益于这部专著,为复杂科学管理做出自己的贡献。

牛东晓

2014 年 5 月于北京

前　言

21 世纪以来，复杂系统的研究成为科学家最主要的研究领域之一，很多科学家将不同的研究对象确定为对应的复杂系统，通过对其建模、分析来把握这些对应系统的特征与性质，并研究出大量的科学成果。

本质上讲，经济与管理复杂自适应系统其实是在自适应系统的基础上添加了一个正反馈，从而使得系统与环境之间存在一个闭环的反馈，这一特征使得整个系统的分析变得异常困难。传统的研究方法是采用实验的方法，针对某一系统的特性，构建其对应的实验模型，然后再根据系统的状态及各个变量的特征，对系统进行仿真实验，发现系统的规律，从而进行有效的管理。然而，这一类研究仅仅适用于特定的研究对象，其结论不具有普适性。

如果从空间及时间维度上以不同的粒度对系统进行分割，就可以构建出一个合适的模型，对这一模型进行求解，即可获得系统运作的规律。空间上我们可以将系统分割成若干个子系统，在本书中称之为局域；时间上我们可以将系统的发展过程按照短的时间范围及长的时间范围两种粒度分割。具体地，空间 scale 中分割及 Agent 行为的假设可以描述为：①微观上，假定同一子系统每一 Agent 知道系统中的所有信息；任何一个 Agent 不能与所有人相互作用，而只能与其博弈半径内子集内的 Agent 相互作用，且其最优策略依赖于系统/局部的结构。同一子系统内的 Agent，从更高级的层面上看，是同质 Agent，且他们之间的关系是合作随机微分博弈。不同子系统中的 Agent 属于异质 Agent，他们之间的相互作用关系是非合作随机微分博弈。②宏观上各个子系统实际上是由子系统内同质 Agent 通过相互作用涌现出来的 Super-Agent，各个 Super-Agent 属于异质 Agent，他们之间的相互作用是非合作随机微分博弈。进一步而言，时间 scale 中的长短分割及 Agent 行为的假设可以描述为：①短的 time-scale 中，Agent 之间的关系为随机微分博弈，最优解取决于系统的结构、环境特征、博弈半径内其他 Agent 的策略及自身的历史策略。②长的 time-scale 中，由于 Agent 的智能性、社会性和自治性的作用，Agent 能够自主选择与调整博弈的对象，使得系统结构不断演化，从而形成以优先连接与增长两个机制共同作用的 6 个子过程：调整 Agent 自身的行为、与同一子系统内的新 Agent 创建博弈关系、与不同子系统内的新 Agent 创建博弈关系、断开旧有的博弈关系、与新进入系统的成员创建博弈关系、退出系统。

故此，本书从时间与空间两个维度以及长（大）与短（小）两种粒度上分析复杂自适应系统中各个 Agent 及整个系统行为的属性。具体地，就是小的 spatial-scale

以及短的 times-cale 中，考虑各个 Agent 在同一局域之内的合作随机微分博弈及其最优策略的轨迹，在大的 spatial-scale 及短的 time-scale 中，考虑各个局域之间的非合作随机微分博弈，并且考虑了两种不同 spatial-scale 之间的接续情况。当 time-scale 比较短时，我们考虑了系统在大小与两种 spatial-scale 状态下 Agent 的行为状态；当 time-scale 比较长时，我们假定系统的行为是用若干个不同的短的 time-scale 顺序组织而成的。在考虑两种不同 time-scale 之间接续条件的前提下，分析了系统结构随着 Agent 行为而自适应性调整的情况时系统内 Agent 最优策略的不变分布特征及受到外部干扰时系统结构的状态与特征。通过分析，解决了时间与空间不同粒度上复杂自适应系统的状态与特征及其演化规律。具体地：①在短的 time-scale 中，对于局域内 Agent 的合作随机微分博弈，首先将整个局域看成一个整体，考虑该局域收益最优化下各个 Agent 的最优策略，然后再将所设计的调整后动态 Shapley 收益分配向量作为收益分配机制，以达到局域内 Agent 行为的同步。对于局域之间 Agent 的非合作随机微分博弈，将其抽象成局域与局域之间相互作用，将局域抽象成 Super-Agent，考虑各个 Super-Agent 的非合作随机微分博弈的反馈 Nash 均衡解。在合作/非合作随机微分博弈最优策略及其最佳收益确定的前提下，构造了一个非线性算子，使得局域内 Agent 的合作随机微分博弈的 Pareto 最优策略轨迹与局域之间 Super-Agent 之间非合作随机微分博弈的 Nash 最优策略轨迹有效地耦合在一起，从而确定系统中 Agent 的最优策略轨迹及相应的收益特征。随即讨论了这一最优策略轨迹的收敛性，确定了这一短 time-scale 的确定系统拓扑结构内最优策略轨迹的吸引子。②在长的 time-scale 中，将 Agent 相互作用的拓扑空间结构、Agent 行为、环境特征综合考虑形成一个行为-结构共演化的随机过程，针对复杂管理系统的特征，将这一随机过程分成了 Agent 行为更新、在同一局域内创建新的相互作用、与其他局域内 Agent 创建新的相互作用、删除系统已有的相互作用、与新进入系统的 Agent 建立相互作用、删除系统中的 Agent 等 6 个子过程，将这 6 个子过程通过以 Agent 收益驱动的优先连接机制与易变机制支配下形成的在长的 time-scale 中对应的总随机过程，分析这一共演化过程的分布特征，从而确定系统中最优策略轨迹吸引子的分布特征。③分析了受到外部干扰时系统个体 Agent 从系统中消失而给系统拓扑结构造成的影响，采用逾渗理论，分析随机删除一定数量的 Agent 及按照收益从大到小重排序之后蓄意删除收益最大的一定数量的 Agent 之后系统的连通性，并根据系统的连通程度确定系统逾渗的临界状态及临界概率，从而给出了系统可持续稳定发展的基本条件。

本书将会给读者展示以下结果：①在短的 time-scale 中，系统中的 Agent 在合作/非合作随机微分博弈这一过程中确定自己的最优策略轨迹。②在长的 time-scale 中，系统中的 Agent 按照自身收益的大小，在优先连接与易变机制驱动下按

照最优策略的不变分布特征确定自己的策略行为，从而使得系统发展变化。③这一系统在演化过程中，面对随机攻击，系统表现出很强的鲁棒性，面对蓄意攻击，系统表现出很强的脆弱性。不仅如此，本书从更深的层次给出以下结论：①在短的time-scale中，通过非线性耦合算子确定的合作/非合作随机微分博弈的最优策略轨迹既能够保证在同一个局域内的Agent行为达到协调，使得整个局域收益最大，也能使得系统中各个局域处于均衡状态；系统中Agent之间的合作/非合作随机微分博弈耦合下的最优策略收敛到一个确定稳定的吸引子，这一吸引子决定着在这一短的time-scale中系统Agent的最优均衡策略的变化趋势。②在较长time-scale中，复杂管理系统是Agent行为和系统拓扑结构共演化下的一个系统，这个系统表现出马尔科夫属性，其不变分布由系统中Agent的行为、系统的拓扑结构及系统中Agent的总量和噪声所决定；当系统中Agent行为的噪声趋向于无穷小时，不变分布以大偏离原则收敛到以一个特定的率函数有关的一个区域中，此时，与所有的Agent策略有关的系统状态处于一个(β,ρ)相关均衡状态；而在系统中Agent总量趋向于无穷大这一极限下，不变分布以大偏离原则收敛到以率函数$-r^{\beta}(\sigma,q)$有关的一个区域上。③当系统受到随机攻击时，系统中至少存在两个大组分将系统连通；而系统受到蓄意攻击时，其删除概率有一个临界点，在临界点之下，系统中至少有两个大组分使得系统得以连通，而在临界点之上，系统中只能有一个大组分，这一大组分内绝大多数Agent孤立，并且这一临界点与系统中Agent收益的特征有关。

在撰写本书的过程中，一直得到我的恩师——武汉大学徐绪松教授的帮助，并有幸得到了中科院数学所郭雷院士及其他多位老师的指导，他们宝贵的建议使得我的研究能够顺利完成，在这里表示衷心的感谢！在写作的后期，我有幸得到了武汉大学和哈尔滨商业大学一些老师的教诲，拓宽了本书的写作视角，在此表示深深的谢意。

由于作者能力有限，书中难免存在不足之处，敬请广大读者批评指正。

郑小京

2014年于哈尔滨

目　录

1 基础理论方法

1.1 随机微分博弈

经典的博弈论模型,大部分都是建立在离散的基础上的,并且将整个系统当成一个静态的系统看待:仅仅考虑某一时刻点系统或系统中 Agent 的收益最佳化,并没有考虑在一段时间之上的收益最大化及这种收益的稳定性,这就使得经济组织一旦采用这个结果之后,很难综合考虑短、中、长期目标发生冲突时,组织放弃短期利益这一管理学的基本"准则"。并且,作为经济组织,他们在与其他经济组织及环境进行着相互作用,在这种相互作用过程中,他们需要不断地学习,从而调整自己的行为使其更加适应于环境。从这个意义上讲,需要考虑一个动态的、具有学习性的博弈模型。

为了描述上述特征,本书将系统中每个 Agent 的目标函数和对应的约束条件写成:

$$E_{t_0}\left\{\int_{t_0}^{T} g^j[s,x_j(s),u_j(s)]\exp\left[-\int_{t_0}^{s} r(y)\mathrm{d}y\right]\mathrm{d}s + \exp\left[-\int_{t_0}^{T} r(y)\mathrm{d}y\right]q^j(x_j(T))\right\},j\in\{1,2,\cdots,n\}\equiv\mathbb{N} \tag{1-1}$$

系统状态及其变化为

$$\mathrm{d}x_i(s)=f^i[s,x_i(s),u_i(s)]\mathrm{d}s+\sigma_i[s,x_i(s)]\mathrm{d}z_i(s),x_i(t_0)=x_i^0 \tag{1-2}$$

其中 n 维向量 $\boldsymbol{x}(s)$ 表示系统状态;$\boldsymbol{\sigma}$ 为一个 $m\times\Theta$ 维的向量,表示系统随机性,且可设 $\boldsymbol{\Omega}(s,x)=\boldsymbol{\sigma}(s,x)[\sigma(s,x)]^{\mathrm{T}}$;$z$ 为一个维纳过程,描述 Agent 的行为随机性与智能性、自适应性。$\boldsymbol{u}(s)$ 为 Agent 的控制向量,也就是其策略集形成的向量①。

式(1-1)与式(1-2)表现了如下性质:

(1) 系统中任何一个 Agent 不会追求某一时刻利益的最大化,而追求在时间范围$[t_0,T]$上的利益最大化,这就使得这一 Agent 具有了"比较远大"的目标,为

① 其中,$\boldsymbol{u}(s)$应该是一个列向量,表示 Agent i 的多种不同的策略。一般来说,假定 Agent i 的策略空间维数为ϑ,则$\boldsymbol{u}(s)$是一个ϑ维数的列向量,其最优解,正如在后文中提到的针对于时间 s 和状态$x(s)$的策略$\phi_i^*(s,x(s))$,指的是针对于某一特定纯策略的最佳策略向量,$\boldsymbol{\phi}_i^*(s,x(s))=(\phi_{i,1}^*(s,x(s)),\cdots,\phi_{i,\vartheta}^*(s,x(s)))^{\mathrm{T}}$,针对于每一个$\phi_i^*(s,x(s))$,我们都可以找到对应的最佳策略及对应的收益。若令向量$\boldsymbol{\hbar}(s,P(s))=(\hbar_1(s,P(s)),\hbar_2(s,P(s)),\cdots\hbar_\vartheta(s,P(s)))$,则$\boldsymbol{\hbar}(s,P(s))\times\boldsymbol{\phi}_i^*(s,x(s))$描述了混合策略相对应的博弈。

了实现这一长期目标，有可能在某一时刻该 Agent 可以获得更大利润，但是又损伤其中长期目标，则该 Agent 将会放弃这一利润获取，这一点可以从积分中看出来。同时，在时间范围$[t_0, T]$中任何一个 Agent 的利润随着时间随机波动，该 Agent 追求这一时期的总的期望利润最大化，因此在某一时刻，允许该 Agent 利润不会达到最大化。从这一意义上讲，本书认为管理主体不会追求利润的最大化，而是追求较大利润获取这一事件发生概率的最大化。

(2) 由于资源在系统中的运作不断随着时间发生变化，同时任意 Agent 的目标实现程度也会随着时间发生变化，这种变化需要一个随着时间的贴现因子函数来描述，从而表明系统的动态性，本书称之为行为的贴现因子函数。在本书的系统中采用了目前广泛使用的贴现因子函数，即指数分布的贴现因子函数$\hbar = \int_{t_0}^{T} \exp\left[-\int_{t_0}^{s} r(y)\mathrm{d}y\right]\mathrm{d}s$，该函数表明这种行为及目标随着时间下降的平滑变化趋势，符合概率论的基本规律①，这一贴现因子函数的使用，真正描述出来行为的动态性和渐变性。

(3) 对于任意的 Agent，其有一个初始收益，这一初始状态可以用能够表示其行为特点参数的一个算子$\exp\left[-\int_{t_0}^{T} r(y)\mathrm{d}y\right] q^j(x_j(T))$，可以看出，这一初始参数是每个 Agent 收益的初始基数，而其他的总收益可以看成随着时间变化的瞬时

---

① 更加稳妥的说法是，由于系统的动态性、多样性和自适应性，Agent 在整个$[t_0, T]$中行为可能具有不同的性质，因此，我们可以找到 m 个相对独立的指数分布，然后按照加权法集结起来，即给定权值 w_1，$w_2, \cdots, w_h$，$\hbar = \int_{t_0}^{T} \sum_{h=1}^{m} w_h \exp\left[-\int_{t_0}^{s} r_h(y)\mathrm{d}y\right]\mathrm{d}s$，这有助于描述不同识别方法的差异性，更加能够提取出来系统的共性。这时的目标函数变为

$$E_{t_0}\left\{\int_{t_0}^{T} g^j[s, x_j(s), u_j(s)] \sum_{h=1}^{m}\left\{w_h \exp\left[-\int_{t_0}^{s} r_h(y)\mathrm{d}s\right]\right\}\mathrm{d}s + \sum_{h=1}^{m}\left[w_h \exp\left[-\int_{t_0}^{T} r_h(y)\mathrm{d}y\right]\right] q^j(x_j(T))\right\}$$

如果各个 Agent 之间存在行为模仿作用，这一贴现因子函数可以描述成为 Lévy 分布，其数学语言为$\int_{t_0}^{s} -\frac{p}{2\sqrt{\pi r^3}}\exp\left[-\frac{p^2}{4r}\right]\mathrm{d}p$，其对应的目标函数变成了

$$E_{t_0}\left\{\int_{t_0}^{T} g^j[s, x_j(s), u_j(s)]\left[\int_{t_0}^{s} -\frac{p}{2\sqrt{\pi r^3}}\exp\left[-\frac{p^2}{4}\mathrm{d}p\right]\mathrm{d}s\right] + \left[\int_{t_0}^{T} -\frac{p}{2\sqrt{\pi r^3}}\exp\left[-\frac{p^2}{4r}\right]\mathrm{d}p\right] q^j(x_j(T))\right\}$$

在这些分布中，最主要的为分布性质 $S(r)$，以及密度分布函数 $\mathrm{d}(t)$，下面给出其他的一些主要的分布：在$[p, q]$上均匀分布的 $\mathrm{d}(t)$ 为 $\frac{\exp(-pt)-\exp(-qt)}{(q-p)t}$；指数分布的 $S(r)$ 为 $p\exp(-pr)$，$\mathrm{d}(t)$ 为 $\frac{\exp(-pt)-\exp(-qt)}{(q-p)t}$；Γ-Erlang 分布的 $S(r)$ 为$\frac{p^q r^{q-1}}{\Gamma(q)}\exp(-pr)$，$\mathrm{d}(t)$ 为$\frac{p^q}{(p+t)^q}$；1/2 正态分布的 $S(r)$ 为 $\sqrt{\frac{2}{\pi}}\frac{1}{p}\exp\left(-\frac{r^2}{2p^2}\right)$，$\mathrm{d}(t)$ 为 $\exp\left(\frac{p^2t^2}{2}\right)\mathrm{refc}\left(\frac{pt}{\sqrt{2}}\right)$；1/2 柯西分布的 $S(r)$ 为$\frac{2p}{\pi(1+p^2r^2)}$，$\mathrm{d}(t)$ 为$\frac{2}{\pi}\cdot Ci\left(\frac{t}{b}\right)\sin\left(\frac{t}{b}\right)+\cos\left(\frac{t}{b}\right)\left[1-\frac{2}{\pi}Ci\left(\frac{t}{b}\right)\right]$。

收益在$[t_0,T]$上的累积效果或平均值。

(4) 任意 Agent 决策后得到的收益依赖于其自身的状态、与之相互作用的 Agent 的状态与策略及环境的状态。自身的状态是指$x_i(s)$的量,及自己状态的变化程度 $\mathrm{d}x_i(s)$,其他 Agent 的状态包含在状态方程 $f^i[s,x_i(s),u_i(s)]$中①,还有其自身与其他 Agent 的决策向量$\boldsymbol{u}(s)$。环境对其进行着一定的干扰,这种干扰实际上是一个白噪声与布朗运动的相互作用,或者说是一个白噪声与 Weiner 的一个综合过程,这说明了任意 Agent 行为的随机性,以及环境对其的干扰能力。

(5) 既然每个 Agent 在 t 时刻的决策依赖于系统在 t 时刻及其之前的状态及其变化程度,则可以看出:漂移项 $\sigma_i[s,x_i(s)]\mathrm{d}z_i(s)$描述了其状态变化不稳定的状况,这就使得每个 Agent 对自身在未来的状态有一个相对应的预测或预判,本书称之为"新息",正是由于这种新息的存在,每个 Agent 可以根据自身及系统的状态预测未来的情况,当然,这种预测是在 t 时刻之前自身及系统状态的变化基础(经验)之上的。这样,每个 Agent 就具有对历史经验教训的总结,以及这种基于新息的对未来的预测②,这就赋予了每个 Agent 智能性的特征。

(6) 同其收益一样,每个 Agent 都有其初始状态,这一初始状态形成了初始的博弈拓扑结构及系统中 Agent 相互作用过程中流的拓扑结构,分别表示了系统内 Agent 相互作用的结构及博弈的相对强弱结构,这对整个系统具有十分重要的作用。这一点可以从整个复杂系统所具有的蝴蝶效应这一结论中直接得到:系统对初始状态具有十分的敏感效应。

总之,可以得到以下结果:①是考虑时间范围$[t_0,T]$上的收益;②系统状态取决于多个 Agent 的状态 x、相互作用的结构 $f(x(s))$、各个 Agent 的策略 $u_i(s)$及环境的影响 σ;③Agent 行为的不确定性和自适应性通过 z 表达;④策略依赖于前一段时间收益 $v_i(\cdot)$和 $\mathscr{B}_i(\cdot)$,具有智能性。

---

① 这里给出的是一个简写,实际上,系统中各个 Agent 之间的相互作用,可以确定为 $f[s,x_i(s),x_j(s),u_i(s)],i,j\in\mathbb{N},j\neq i$ 这一算子相关的一个 PDE,例如,在后文举例说明以努力水平为核心的系统行为协同化过程中,其状态方程就描述成

$$\text{s. t. } \mathrm{d}x_i(s)=[\alpha_i[u_i(s,x_i(s))x_i(s)]^{1/2}+\sum_{j=1,j\neq i}^{k} b_j^{[j,i]}[x_j(s)x_i(s)]^{1/2}-\delta_i x_i(s)]\mathrm{d}s+\sigma_i x_i(s)\mathrm{d}z_i(s),$$

$$x_i(t_0)=x_i^0\in X,i,j\in\mathbb{N}=\{1,2,\cdots,n\}$$

其相互作用的叠加效用通过算子 $\sum_{j=1,j\neq i}^{k} b_j^{[j,i]}[x_j(s)x_i(s)]^{1/2}$ 描述,表明了这一 Agent 与系统中其他可能能够进行相互作用的 Agent i 之间相互作用的一个综合效应,这一相互作用必须与博弈的拓扑结构相匹配,这一拓扑结构在下文中通过图的拓扑结构给予描述,即 ω。从这一角度上说,该系统表现出巨大的复杂性不言而喻。

② 在实际运行的过程中,这种"新息"通过反馈、学习的形式对其行为进行修订,从而决定着 Agent 的智能性。

我们知道，每个 Agent 都有自己的目标及对应的约束条件，这样才能够在此基础上最大化自己的利益。进一步而言，当整个系统中只有一个 Agent 时，该 Agent 只需要有效、合理地使用资源，并且在环境的影响之下最大化自己的利益；然而，当系统中不止一个 Agent 时，每个 Agent 都要最大化自己的利益，但是系统中的资源有限，并且这些资源随着 Agent 之间的相互作用，彼此之间也在发生着一定的非线性作用，这就会使得每个 Agent 最后需要“妥协”才能够使得整个系统平衡。本书称前者为优化问题，而后者为博弈问题①，从这个意义上讲，博弈与优化具有等价性，这也是多局域世界图中合作-非合作随机微分博弈最优策略存在性的依据。

更进一步，从 Agent 之间相互作用关系的性质入手，按照博弈的分类，将博弈分成合作博弈与非合作博弈两种。这二者的本质区别在于说，前者主要考虑研究系统内部各个 Agent 之间的竞争性行为，即各个 Agent 相互作用达到均衡之后的 Nash 均衡解；而后者主要考虑系统内各个 Agent 为了同一个目标，达成一个协议，然后在此契约之下，行为达到协同化，从而使得整个系统利润达到最大，然后在契约的基础上进行利益的重新分配获得的 Pareto 最优均衡解。下面就对这两种不同的随机微分博弈进行分析。

1.1.1 非合作随机微分博弈

现在考虑一个 n 人的系统，每个 Agent 的行为如式(1-1)和式(1-2)所示。对于这一博弈来说，有开环 Nash 均衡、闭环 Nash 均衡、反馈的 Nash 均衡等众多均衡解，这些解来自于博弈的特征，也表明了 Agent 在相互作用过程中最优策略的确定特性。为了真正实现 Agent 的学习特征的智能性，本书选择了反馈的 Nash 均衡解。

进一步而言，非合作博弈的反馈均衡解取决于 Agent 获得的信息结构，这种信息结构的多样性使得整个 Nash 均衡表现一定的偏离特性，可以通过某种转化使得这一 Nash 解满足反馈 Nash 均衡的属性，从而进一步约束 Nash 解的准确性。特别地，博弈者的信息结构或者是一个闭环的完美状态(closed-loop perfect state, CLPS)模式，其策略为 $\eta^i(s)=\{x(s), t_0\leqslant t\leqslant s\}$②，或者一个无记忆完美状态

---

① 实际上，优化和博弈是对等的，其区别主要存在于系统抽象时粒度的确定：对于一个拥有若干个主体的系统，如果将各个主体看成研究对象，主要考虑在一定的输入下，它们之间和环境对它们的相互作用及这些相互作用对系统特征与状态的影响变化，我们就需要考虑博弈；如果仅仅将这一系统看成一个整体作为研究对象，仅仅考虑在一定的输入下，系统环境对这一研究对象的作用及这种相互作用对于系统特征与状态的影响变化，我们就将其想象成为优化问题。它们之间的对等性，在文献(陈光亚．2009．优化和均衡的等价性．系统科学与数学，29(11)：1441—1446)中被论证。

② 仅仅考虑时刻 S 的信息状态，并且这种信息状态在一段时间内是存在的，不考虑其他因素。

(memoryless perfect state,MPS)模式,其策略为 $\eta^i(s)=\{x_0,x(s)\}$①。更进一步,这一策略需要满足下列 Nash 均衡的描述。

定义 1.1 对于 n 人拥有信息 MPS 或 CLPS 的由式(1-1)和式(1-2)组成的随机微分博弈,一个 n 元组策略 $\{u_i^*(s)=\phi_i^*(s,x)\in U^i, i\in\mathbb{N}\}$ 构成了其反馈 Nash 均衡解,对于任意的 $i\in\mathbb{N}$,如果存在一个定义在区间 $[t_0,T]\times\mathbb{R}^m$ 的函数 $V^i(t,x)$ 满足下列相关条件:

$$V^i(T,x)=q^i(x)$$

$$\begin{aligned}V^i(t,x)&=\int_t^T g^i[s,x^*(s),\phi_1^*(s,\eta_s),\phi_2^*(s,\eta_s),\cdots,\phi_n^*(s,\eta_s)]\mathrm{d}s+q^i(x^*(T))\\&\geqslant\int_t^T g^i[s,x^{[i]}(s),\phi_1^*(s,\eta_s),\phi_2^*(s,\eta_s),\cdots,\phi_{i-1}^*(s,\eta_s),\\&\quad\phi_i(s,\eta_s),\phi_{i+1}^*(s,\eta_s),\cdots,\phi_n^*(s,\eta_s)]\mathrm{d}s+q^i(x^{[i]}(T))\end{aligned}$$

$$\forall\,\phi_i(.\,,.)\in\Gamma^i, x\in\mathbb{R}^n$$

其中,在区间 $[t_0,T]$ 上,有

$$\begin{aligned}\dot{x}^{[i]}(s)=f[&s,x^{[i]}(s),\phi_1^*(s,\eta_s),\phi_2^*(s,\eta_s),\cdots,\phi_{i-1}^*(s,\eta_s),\phi_i(s,\eta_s),\phi_{i+1}^*(s,\eta_s),\cdots,\\&\phi_n^*(s,\eta_s)]\end{aligned}$$

$$x^{[1]}(t)=x$$

$$\dot{x}^*(s)=f[s,x^*(s),\phi_1^*(s,\eta_s),\phi_2^*(s,\eta_s),\cdots,\phi_n^*(s,\eta_s)],x^*(t)=x$$

且 η_s 代表了数据集 $\{x(s),x_0\}$,或者 $\{x(\tau),\tau\leqslant s\}$,它依赖于博弈的信息结构是 MPS 还是 CLPS 的特征②。显然,这一模型的求解过程十分复杂,下面给出该博弈的 Nash 最优解。

为了给出其解,首先考虑下面的优化模型:

$$\max_u E_{t_0}\left\{\int_{t_0}^T g^i[s,x(s),u(s)]\mathrm{d}s+q(x(T))\right\}\tag{1-3}$$

$$\mathrm{d}x(s)=f[s,x(s),u(s)]\mathrm{d}s+\sigma[s,x(s)]\mathrm{d}z(s),x(t_0)=x_0\tag{1-4}$$

其各参数的假设同由式(1-1)和式(1-2)组成的模型。其模型的最优解可以转换为一个随机微分方程的解,这样就大大降低了求解的难度,其解如引理 1.1③ 所述。

① 不仅考虑时刻 S 的信息状态,还需要考虑初始状态的信息状态。

② 显然,反馈 Nash 最优的概念建立在一个相对的基础上,先假定任何其他的最优,然后再考虑自己的策略在其他 Agent 策略最优的前提下取得的最优。然而,这一对其他人策略最优的假定是对于系统中的任意人,只有这样,才能说明我们获得的系统反馈 Nash 最优确实是最优的。这一概念是从博弈鞍点的充要性出发即可得到。对于 $x^{[i]}(s)$,是与策略 $\phi_i(s,\eta_s)$ 对应的,显然,这不是 Agent 的最优策略轨迹对应下的状态动力学。

③ 在引理 1.1 及以后的所有非合作博弈中,我们假定 V 就是其收益函数,而对于合作博弈来说,我们假定 W 为其博弈的收益函数,以进行区别。对于一个组织来说,我们知道有其输入、输出、转换及干扰、反馈,其转换是非常关键的,那么对于转换而言非常重要,我们的研究主要针对于经济管理组织中管理策略所对应的参数,或者可以采用的策略集及其轨迹。在文中我们用"控制"这一术语来描述这一特征。

引理 1.1　控制集 $u^*(t)=\phi^*(t,x)$ 形成了由式(1-3)和式(1-4)组成的模型的一个最优解，如果存在一个连续的微分函数 $V(t,x):[t_0,T]\times\mathbb{R}^m\to\mathbb{R}$ 满足下列偏随机微分方程：

$$-V_t(t,x)-\frac{1}{2}\sum_{h,\zeta=1}^{m}\Omega^{h\zeta}(t,x)V_{x^h x^\zeta}(t,x)=\max_u\{g[t,x,u]+V_x(t,x)f[t,x,u]\},$$

$$V(T,x)=q(x)\ ①$$

证明：将最优控制 $\phi^*(t,x)$ 代入式(1-4)，获得最优状态动力学

$$dx(s)=f[s,x(s),\phi^*(s,x(s))]ds+\sigma[s,x(s)]dz(s),x(t_0)=x_0 \tag{1-5}$$

式(1-5)的解记为 $x^*(t)$，可以表达为

$$x^*(t)=x_0+\int_{t_0}^{t}f[s,x^*(s),\psi_1^{(t_0)*}(s,x^*(s)),\cdots,\psi_n^{(t_0)*}(s,x^*(s))]ds$$
$$+\int_{t_0}^{t}\sigma[s,x^*(s)]dz(s) \tag{1-6}$$

定义在时间 t 及对应现状 x_t^* 的最大收益为如下的价值函数

$$V(t,x_t^*)=\max_u E_{t_0}\left\{\int_{t_0}^{T}g[s,x(s),u(s)]ds+q(x(T))\,|\,x(t)=x_t^*\right\}$$
$$=E_{t_0}\left\{\int_{t_0}^{T}g[s,x^*(s),\phi^*(s,x^*(s))]ds+q(x^*(T))\right\}$$

满足边界条件 $V(T,x^*(T))=q(x^*(T))$。

本书可以将 $V(t,x_t^*)$ 表示为

---

① $-V_t(t,x)-\frac{1}{2}\sum_{h,\zeta=1}^{m}\Omega^{h\zeta}(t,x)V_{x^h x^\zeta}(t,x)=\max_u\{g[t,x,u]+V_x(t,x)f[t,x,u]\}$ 这一偏随机微分方程的解对应于下列模型的解：

$$\max_u E_{t_0}\left\{\int_{t_0}^{T}g^i[s,x(s),u(s)ds+q(x(T))]\right\}$$
$$dx(s)=f[s,x(s),u(s)]ds+\sigma[s,x(s)]dz(s),x(t_0)=x_0$$

在博弈模型中，我们的决策是在时间 t_0 内给出的系统在任意时间 $t\in[t_0,T]$ 的最优解。与之对应的偏随机微分方程的解中，既考虑了随机项的特征 $\sum_{h,\zeta=1}^{m}\Omega^{h\zeta}(t,x)V_{x^h x^\zeta}(t,x)$，又考虑到在任意时刻 $t\in[t_0,T]$ 系统的状态。因此我们获得的反馈 Nash 均衡解应该是一个方程集合，是一个轨迹。

由于前面提及，$\boldsymbol{\sigma}$ 为一个 $m\times\Theta$ 维的向量，表示系统随机性，且可设 $\boldsymbol{\Omega}(s,x)=\sigma(s,x)[\sigma(s,x)]^T$；$z$ 为一个维纳过程，描述 Agent 的行为随机性与智能性、自适应性，因此在定理 1.1 中提到的 $\sum_{h,\zeta=1}^{m}\Omega^{h\zeta}(t,x)V_{x^h x^\zeta}(t,x)$ 这一项中，我们需要将这 $m\times\Theta$ 维的随机向量进行考虑，这也就是从 1 到对其求和的基本原因；然而，由于考虑到这种随机性是建立在状态变化随机性的基础上的，因此，我们必须考虑系统价值的变化程度 $V_{x^h x^\zeta}(t,x)$，这实际上是沿用了泰勒展开的基本思想。

离散的博弈，包括多人博弈，不管是确定的博弈，还是随机博弈，其解实际上是随机微分博弈的特解。