

控制与决策

Control and Decision

多目标动态优化中Pareto随机合作博弈研究综述

张维海, 彭称称, 蒋秀珊

引用本文:

张维海, 彭称称, 蒋秀珊. 多目标动态优化中Pareto随机合作博弈研究综述[J]. *控制与决策*, 2023, 38(7): 1789–1801.

在线阅读 View online: <https://doi.org/10.13195/j.kzyjc.2022.2097>

您可能感兴趣的其他文章

Articles you may be interested in

考虑碳排放和综合需求响应的电-气联合运行决策博弈

Decision-making game of power-gas joint operation considering carbon emission and integrated demand response

控制与决策. 2023, 38(7): 1979–1987 <https://doi.org/10.13195/j.kzyjc.2021.1503>

带尺寸约束的二机流水车间生产运输协调博弈调度问题

Coordinated two-machine flow-shop and transportation scheduling under size constraint and game theory consideration

控制与决策. 2023, 38(7): 1942–1950 <https://doi.org/10.13195/j.kzyjc.2021.2207>

带有交货期的比例流水车间调度问题的合作博弈

Cooperative games on proportionate flow-shop scheduling problem with due-dates

控制与决策. 2022, 37(3): 712–720 <https://doi.org/10.13195/j.kzyjc.2020.1355>

网络系统的安全决策与控制: 容错博弈研究综述

Safe decision and control of network systems: A survey on fault tolerant game

控制与决策. 2022, 37(4): 769–781 <https://doi.org/10.13195/j.kzyjc.2021.1557>

一种要素双模糊的限制交流结构合作博弈方法及应用

An allocation model of limited communication structure cooperative game with dual fuzzy elements

控制与决策. 2021, 36(2): 475–482 <https://doi.org/10.13195/j.kzyjc.2019.1048>

多目标动态优化中 Pareto 随机合作博弈研究综述

张维海^{1†}, 彭称称², 蒋秀珊³

(1. 山东科技大学 电气与自动化工程学院, 山东 青岛 266590; 2. 青岛理工大学 信息与控制工程学院, 山东 青岛 266520; 3. 中国石油大学(华东) 石大山能新能源学院, 山东 青岛 266580)

摘要: 随着经济全球化的不断深入,“合作共赢”的发展战略越来越被人们接受,进而合作博弈也被合理地应用到多个领域.与静态合作博弈相比,动态博弈的约束条件为动态方程,其具有优化行为、多个玩家共同存在、决策结果的持久性以及对环境变化的鲁棒性等特点.由于动态系统总是受到某些随机波动的干扰,将这些内部随机波动和外部随机扰动考虑到系统模型中更为实际.随机动态合作博弈同时考虑策略行为、动态演化与随机因素之间的相互作用,其可能是最复杂的决策形式之一.鉴于此,对多目标动态优化中随机合作博弈的进展进行综述:首先,回顾多目标合作博弈的研究背景,给出 Pareto 最优性的定义和基本性质;其次,综述确定性的合作博弈;再次,分别论述随机合作博弈和平均场随机合作博弈;最后,提出随机合作博弈几个未来研究方向.

关键词: Pareto 最优性; 合作博弈; 随机最优控制; 线性二次控制; 平均场理论

中图分类号: TP273

文献标志码: A

DOI: 10.13195/j.kzyjc.2022.2097

引用格式: 张维海, 彭称称, 蒋秀珊. 多目标动态优化中 Pareto 随机合作博弈研究综述[J]. 控制与决策, 2023, 38(7): 1789-1801.

Pareto stochastic cooperative games in multiobjective dynamic optimization problems: A survey

ZHANG Wei-hai^{1†}, PENG Chen-chen², JIANG Xiu-shan³

(1. College of Electrical Engineering and Automation, Shandong University of Science and Technology, Qingdao 266590, China; 2. School of Information and Control Engineering, Qingdao University of Technology, Qingdao 266520, China; 3. College of New Energy, China University of Petroleum (East China), Qingdao 266580, China)

Abstract: With the deepening of economic globalization, the concept of win-win cooperation is more and more accepted by people and the cooperative game is applied appropriately to various fields. Compared with the static cooperative game, the dynamic game is subject to dynamic systems, which possess the features of optimizing actions, the co-existence of multiple players, and enduring consequences of decisions and robustness with regard to variability in the environment. Because the dynamic system is always disturbed by some random perturbations, it is natural to consider these internal random perturbations and external random disturbances in the dynamic model. The stochastic dynamic cooperative game may be one of the most complex decision-making forms, as it considers the interaction among strategic behavior, dynamic evolution and stochastic factors. Based on the above discussions, we generalize the progress of the stochastic cooperative game in the multiobjective dynamic optimization: Firstly, the research background of the cooperative game in the multiobjective optimization problem is reviewed, and the definition and fundamental properties of Pareto optimality are presented; secondly, the deterministic cooperative game is summarized; moreover, the stochastic cooperative game and the mean-field stochastic cooperative game are summarized; finally, several future research directions of the stochastic cooperative game are proposed.

Keywords: Pareto optimality; cooperative games; stochastic optimal control; linear-quadratic control; mean-field theory

0 引言

自 20 世纪 50 年代以来,由于数字计算机实用化和空间技术发展的推动,动态系统的优化理论迅速发展,形成了一个重要的学科分支——最优控制.经过

几十年的发展,动态系统优化理论不仅突破了自动控制的传统界限,而且有了许多成功的应用,它在系统工程、空间技术、人口控制、经济管理决策等众多领域应用广泛且效果显著.因此,最优控制的发展

收稿日期: 2022-12-04; 录用日期: 2023-03-15.

基金项目: 国家自然科学基金项目(61973198, 62203247); 山东省泰山学者计划项目.

[†]通讯作者. E-mail: w_hzhang@163.com.

经久不息,至今仍是一个热门研究领域. 最优控制理论是现代控制理论的重要组成部分,例如Pontryagin等^[1]提出的最大值原理、Bellman^[2-3]创立的动态规划、Kalman^[4]建立的卡尔曼滤波和线性二次最优控制理论^[5]等. 现代控制理论和控制工程吸收了现代数学和进步技术的很多重要成果,并渗透到生产、管理、规划乃至国防等领域,发挥越来越大的作用. 在此期间,最优控制也得到了很大程度的发展,例如分布参数最优控制、大系统最优控制、随机最优控制、微分对策等^[6].

1) 博弈理论.

当一个个体(玩家)在追求一个(或多个)目标时,其他个体追求的目标与该个体追求的目标可能冲突或重叠,这便产生了博弈^[7]. 博弈理论借鉴了数学、统计学、运筹学、工程学、生物学、经济学、政治学等学科的特点,研究相互作用环境中的决策行为,已成为研究实际问题的基本工具. 博弈解决的问题是确定每个个体的最优决策,研究这些决策如何相互作用使个体之间产生均衡,以及解释这些结果的性质. 从导弹控制到市场开发、自然资源开采、竞争政策、谈判技巧、宏观经济和环境规划、资本积累以及投资管理等,博弈的理论研究和实际应用正在迅速发展. 在解决上述问题的方法中,博弈论是应用数学能够提供的最丰富和最有效的方法. 从决策和政策制定者的角度看,认识和适应人类决策的相互依赖性和相互作用变得越来越重要,而博弈理论极大地增强了对决策的理解. 在社会科学中,博弈理论很好地应用到了经济和金融领域,其不仅着眼于对经济和社会决策进行更加实际和合理的分析,而且会揭示一些有趣的新问题,特别是在管理科学领域.

实际生活中的大多数决策都是基于直觉、常识、机会等. 然而,有些领域需要数学建模或者编程,如工程和经济学等,解决的问题包括设计航天器、桥梁,规划和定价生产系统或管理环境的污染问题等. 这种情况下,传统的单目标优化方法不足以解决上述问题,这就需要新的思维方式、新的概念和新的方法——多目标优化. 在多目标优化中,由于目标函数的矛盾性和不可约性,找到一个使所有目标函数同时最优的解并不简单. 此外,由多目标优化目标函数组成的向量空间不是有序的,它只是部分有序. 例如, $(3, 3)^T$ 可以说小于 $(6, 6)^T$, 但是 $(3, 6)^T$ 与 $(6, 3)^T$ 之间的大小关系却不能比较. 基于此,Edgeworth^[8]于1881年提出支配解/非支配解的概念,它被意大利经

济和社会学家Pareto^[9-10]进一步发展. 本文考虑一种多目标合作的结果,即当部分玩家的目标函数变小,必然有其他玩家的目标函数变大,这便产生了Pareto最优性^[9-10]的概念.

现代博弈论的研究始于1921年,法国数学家Borel研究了二人零和博弈问题,提出与控制律类似的博弈问题“策略”概念. 1928年, von Neumann^[11-12]建立了二人零和博弈问题的一般理论,提出最小最大定理. 1934年,德国经济学家 von Stackelberg 在经济竞争的背景下引入玩家之间顺序的区别,称先行动的为玩家为领导者,后行动的玩家为跟随者,其思想是跟随者可以观察领导者的行为并随后采取行动. 1944年, von Neumann 与 Morgenstern^[13] 合著《Theory of Games and Economic Behavior》使博弈论受到广泛关注,标志着现代博弈论的初步形成. 1950年, Nash^[14]提出纳什均衡,并讨论多人非零和博弈问题. 博弈论虽然与控制论有密切联系,但两者在很大程度上是平行发展的^[15-19].

2) Pareto 合作博弈.

随着经济全球化的不断深入,“合作共赢”的发展战略越来越被人们接受. 进而, Pareto 合作博弈的理论成果也被合理地应用到渔业管理博弈^[20]、集团公司内子公司间的广告博弈^[21]和国土安全博弈^[22]等. 当一个玩家有多个目标函数或者有多个玩家影响着动力学系统,并且玩家之间决定通过协同它们之间的策略来实现目标函数最优化时,便产生了合作博弈,此时博弈中某个玩家的目标函数不再由自己决定. 合作博弈为涉及策略行为的决策问题提供社会最优和群体最优的解决方案,其基本特性是制定玩家的最佳行为.

一般而言,根据约束条件的不同,合作博弈分为静态合作博弈^[23-26]和动态合作博弈^[20-22,27-33]. 静态合作博弈是指在某一时刻,所有玩家一次同时做出选择,然后根据所做的选择,每个玩家都会得到相应的回报. 在这种情况下,诸如决策过程中的博弈顺序、参与者在决策时可获得的信息以及博弈的演化等重要问题被抑制. 然而,在动态合作博弈中,约束条件为动态方程,其具有优化行为、多个玩家共同存在、决策结果的持久性以及对环境变化的鲁棒性等特点. 根据约束方程的不同,动态合作博弈又可以划分为合作微分博弈^[20-21,27-29](动力学方程为微分方程)和合作差分博弈^[22,30-33](动力学方程为差分方程). 与微分方程相比,差分方程具有数值解易得、很多实际问题由

差分方程描述更准确、大部分连续系统的仿真结果可通过离散化的方式获得等优点.因此,除了合作微分博弈,深入地展开动态合作差分博弈的研究也十分有必要.

3) 随机动态博弈.

由于动态系统总是受到某些随机波动的干扰,将这些内部随机波动和外部随机扰动考虑到系统模型中更为实际.近几十年来,无论在工程、经济、金融还是在生物科学领域,随机动态系统都被广泛研究.

无论个人选择以何种方式获取数据、信息或预测结果,都无法避开未来的不确定性.由于同时考虑了策略行为、动态演化与随机因素之间的相互作用,随机动态合作博弈可能是最复杂的决策形式之一.合作博弈理论已经成功地应用于运筹学、管理学、经济学、政治学等领域,将这些结果扩展到具有随机元素的动态环境中可能会更富有成效.然而,在随机动态合作博弈中,由于难以得到确定的解,借助Pareto最优性有效解决复杂随机系统的多人博弈问题仍然是一个难题.

综上所述,运用Pareto最优性研究多玩家随机合作博弈问题具有重要的理论意义和广泛的应用潜力.鉴于随机合作博弈的重要性,本文将对Pareto最优性的发展进行系统回顾,首先介绍Pareto最优性的定义和求解方法;然后阐述确定性合作博弈、随机合作博弈以及平均场随机合作博弈分别在连续时间和离散时间情形下的发展,给出随机合作差分博弈在国家间财政赤字博弈的应用;最后展望随机合作博弈理论的未来研究方向.

1 Pareto最优性的定义及求解方法

为了更好地引入和理解Pareto最优性的概念,下面以含有 N 个玩家的无限时域非线性合作差分博弈为例,给出相应的公式化描述,其动力学方程为如下非线性差分系统:

$$\begin{cases} x_{k+1} = f(k, x_k, u_{1,k}, u_{2,k}, \dots, u_{N,k}), \\ x_0 \in \mathbf{R}^n, k \in \mathcal{N} := \{0, 1, \dots\}. \end{cases} \quad (1)$$

其中: N 为正整数且 $N \geq 2$,表示博弈中玩家的个数; $u_{i,k}$ 表示玩家 i 在 k 时刻的控制策略; $x_k \in \mathbf{R}^n (k \in \mathcal{N})$ 和 $u_{i,k} \in \mathbf{R}^{m_i} (i \in \bar{N} := \{1, 2, \dots, N\})$ 分别表示状态向量和控制策略.假设每个玩家 i 都期望使如下目标函数最小:

$$\begin{aligned} & \mathcal{J}_i(x_0; u_{1,k}, u_{2,k}, \dots, u_{N,k}) := \\ & \sum_{k=0}^{\infty} g_i(k, x_k, u_{1,k}, u_{2,k}, \dots, u_{N,k}), k \in \mathcal{N}. \end{aligned} \quad (2)$$

此外,玩家 i 的控制策略 $u_{i,k}$ 只能在其容许控制空间 \mathcal{U}_{ad}^i 中选取.由所有玩家的控制策略组成的联合控制策略集为

$$u := (u_{1,k}, u_{2,k}, \dots, u_{N,k}) \in \mathcal{U}_{ad}^1 \times \mathcal{U}_{ad}^2 \times \dots \times \mathcal{U}_{ad}^N = \mathcal{U}_{ad} \in \mathbf{R}^m, \quad (3)$$

其中 $m = \sum_{i=1}^N m_i$.

定义1^[8] 称 $u^* = (u_{1,k}^*, u_{2,k}^*, \dots, u_{N,k}^*)$ 生成的由目标函数组成的向量

$$\begin{aligned} & (\mathcal{J}_1(x_0; u_{1,k}^*, u_{2,k}^*, \dots, u_{N,k}^*), \\ & \mathcal{J}_2(x_0; u_{1,k}^*, u_{2,k}^*, \dots, u_{N,k}^*), \dots, \\ & \mathcal{J}_N(x_0; u_{1,k}^*, u_{2,k}^*, \dots, u_{N,k}^*)), \end{aligned} \quad (4)$$

或简记为 $(\mathcal{J}_1(x_0; u^*), \mathcal{J}_2(x_0; u^*), \dots, \mathcal{J}_N(x_0; u^*))$,支配另一个决策向量 $\tilde{u} = (\tilde{u}_{1,k}, \tilde{u}_{2,k}, \dots, \tilde{u}_{N,k})$ 生成的由目标函数组成的向量 $(\mathcal{J}_1(x_0; \tilde{u}), \mathcal{J}_2(x_0; \tilde{u}), \dots, \mathcal{J}_N(x_0; \tilde{u}))$,如果如下不等式:

$$\begin{cases} \mathcal{J}_1(x_0; u^*) \leq \mathcal{J}_1(x_0; \tilde{u}), \\ \mathcal{J}_2(x_0; u^*) \leq \mathcal{J}_2(x_0; \tilde{u}), \\ \vdots \\ \mathcal{J}_N(x_0; u^*) \leq \mathcal{J}_N(x_0; \tilde{u}) \end{cases}$$

至少有一个严格成立.

由定义1可得,对于两个决策向量 \tilde{u} 和 u^* ,如果决策向量 u^* 生成的每个目标函数都比由决策向量 \tilde{u} 生成的目标函数小,则称 u^* 生成的解支配 \tilde{u} 生成的解.因此,支配性意味着所有玩家的目标函数可以同时变大或者变小.

定义2^[9-10] 称 u^* 为Pareto最优决策向量,如果不存在另一个决策向量 \tilde{u} 使得如下不等式:

$$\begin{cases} \mathcal{J}_i(x_0; \tilde{u}) \leq \mathcal{J}_i(x_0; u^*), \\ \mathcal{J}_\ell(x_0; \tilde{u}) \leq \mathcal{J}_\ell(x_0; u^*) \end{cases} \quad (5)$$

对至少一个 ℓ 严格成立.向量 $(\mathcal{J}_1(x_0; u^*), \mathcal{J}_2(x_0; u^*), \dots, \mathcal{J}_N(x_0; u^*))^T \in \mathbf{R}^N$ 组成的点称为Pareto最优解,Pareto最优解组成的集合称为Pareto边界.

在多目标合作博弈中,一个求解Pareto最优决策向量的有效方法是由Gass等^[34]于1955年以及Zadeh^[35]于1963年提出的加权方法,其核心思想是根据每个玩家在博弈中的相对重要性,给其目标函数一个权重,然后最小化权重和目标函数,每个能使权重和目标函数最小的决策向量为Pareto最优决策向量.

引理1^[34-35] 令 $\gamma \in \Gamma = \{\gamma = (\gamma_1, \gamma_2, \dots, \gamma_N) |$

$$\left. \begin{aligned} \gamma_i \geq 0, \sum_{i=1}^N \gamma_i = 1 \end{aligned} \right\}, \text{如果 } u^* \in \mathcal{U}_{ad} \text{ 满足}$$

$$u^* \in \arg \min_{u \in \mathcal{U}_{ad}} \left\{ \sum_{i=1}^N \gamma_i \mathcal{J}_i(x_0; u) \right\}, \quad (6)$$

则 u^* 是 Pareto 最优决策向量.

如果合作博弈中的容许控制空间 \mathcal{U}_{ad} 和目标函数 $\mathcal{J}_i(x_0; u) (i \in \bar{N})$ 分别关于联合控制策略 u 为凸集和凸函数, 则所有的 Pareto 最优决策向量可以通过加权方法获得^[36-37].

引理 2^[36-37] 假设容许控制空间 \mathcal{U}_{ad} 和目标函数 $\mathcal{J}_i(x_0; u) (i \in \bar{N})$ 关于控制策略 u 分别为凸集和凸函数, 如果 u^* 是一个 Pareto 最优决策向量, 则存在 $\gamma \in \Gamma$ 使式(6)成立.

由引理 1 可得, 加权方法只是求解 Pareto 最优决策向量的充分条件. 在求解 Pareto 最优解时, 加权方法既不是充分条件也不是必要条件. 显然, 若存在另一个决策向量 \hat{u} 能使每一个玩家的目标函数最小, 即

$$\hat{u} \in \arg \min_{u \in \mathcal{U}_{ad}} \{ \mathcal{J}_i(x_0; u), \forall i \in \bar{N} \}, \quad (7)$$

则根据定义 2, \hat{u} 为 Pareto 最优决策向量. 该性质由 Balandin 等^[38-40] 提出, 并将此结论应用到鲁棒控制理论. 对比式(6)和(7)可得, 引理 1 中的加权方法在求解 Pareto 最优决策向量时降低了保守性, 即对于一个固定的 $\gamma \in \Gamma$, 能使每个玩家目标函数最小的决策向量也可以使权和目标函数最小, 反之不成立. 针对博弈中每个玩家的目标函数与权和目标函数凸性之间的关系, 给出如下结论.

引理 3 若目标函数 $\mathcal{J}_i(x_0; u) (\forall i \in \bar{N})$ 为关于决策向量 u 的凸函数, 则对于任意的 $\gamma \in \Gamma$, 权和目标函数 $\mathcal{J}_\gamma(x_0; u) = \sum_{i=1}^N \gamma_i \mathcal{J}_i(x_0; u)$ 也为凸函数, 反之不成立.

在由 N 个玩家组成的合作博弈中, 通过求解 N 个优化问题可以确定所有的 Pareto 最优解, 其中求解玩家 i 的 Pareto 最优解的充要条件是玩家 i 的容许控制策略 \mathcal{U}_{ad}^i 依赖除玩家 i 外其他玩家 Pareto 最优解的损失.

引理 4^[41-42] $u^* \in \mathcal{U}_{ad}$ 是 Pareto 最优决策向量, 当且仅当定义在如下集合的 u^* 使得目标函数 $\mathcal{J}_i(x_0; u) (\forall i \in \bar{N})$ 最小化:

$$\mathcal{U}_{ad}^i := \{ u \in \mathcal{U}_{ad} | \mathcal{J}_\ell(x_0; u) \leq \mathcal{J}_\ell(x_0; u^*), \forall \ell \in \bar{N} \setminus i \}, \quad (8)$$

其中 $\bar{N} \setminus i$ 表示由所有玩家组成的集合 \bar{N} 但不包括玩家 i 的集合.

2 确定性动态合作博弈

2.1 合作微分博弈

线性二次最优控制由 Kalman^[5] 创立, 是现代控制理论最重要的组成部分之一. 在线性二次最优控制问题中, 一般假设状态向量的权重矩阵半正定, 控制向量的权重矩阵正定, 这种情况称为正则线性二次优化问题. 对于正则线性二次合作微分博弈, 目标函数关于联合控制策略 u 是凸函数, 并且容许控制集 \mathcal{U}_{ad} 在绝对可积的条件下也是凸集. 此时, 加权方法是求解 Pareto 最优决策向量的充要条件^[42]. 然而, 如果合作线性二次微分博弈的状态权重矩阵不定号(可正, 可负, 也可为零), 则即使控制向量的权重矩阵正定, 目标函数的凸性也不能保证, 那么便得不到所有的 Pareto 最优策略. 针对上述问题, Engwerda^[20] 首先推导了目标函数凸性的充要条件, 使加权方法和 Pareto 最优决策向量等价, 然后得到有限时域和无限时域的所有 Pareto 最优解. 所得理论结果被应用到渔业管理博弈, 即工厂对环境保护的投资和回报问题.

非线性多目标合作博弈的 Pareto 最优解一般通过参数化方法获得, 但是此方法的缺点是不能求出所有的 Pareto 最优解. 基于此, 通过引入适当的辅助函数, 借助带有终端约束的最大值原理, 得到了有限时域合作微分博弈的 Pareto 最优解的必要条件^[21]. 在哈密顿函数的最小值关于联合控制策略 u 为凸函数的假设下, 得到了 Pareto 最优解的充分条件, 并将得到的非线性 Pareto 最优性的充要条件应用到集团公司内子公司间的广告投入与收益问题. 此外, 上述推导的非线性结果被用来研究具有固定初值和任意初值的合作线性二次微分博弈, 其解分别通过求解两点边值问题和代数 Riccati 方程的解获得. 加权方法、包络法和射线法被提出用以研究具有固定结构和可调参数的线性微分系统的 Pareto 最优决策向量^[37].

当决策时域没有边界时便产生了无限时域合作博弈. 在一些多目标优化问题中, 博弈结束的时间要么非常遥远, 要么玩家不知道. 例如, 一家上市公司的价值是其贴现未来收益的现值, 但是没人知道这家公司什么时候倒闭. 在这种情况下, 设置终端时间 T 满足 $T = \infty$, 即为无限时域, 更能反映真实博弈时间的最佳近似值. 在实际生产生活中, 这类问题主要包括可再生资源开采、环境管理和公司股权定价等. 不同于有限时域合作博弈, 无限时域博弈需要保证动力学方程的稳定性, 因此需要一些特殊假设. 当线性系统能控和非线性系统关于容许控制偏导数的函数矩

阵满秩时,分别得到了无限时域合作线性二次微分博弈^[43]和非线性合作微分博弈^[44]的Pareto最优解的必要条件和充分条件.

对于多目标干扰衰减问题,其Pareto最优决策向量可以通过无限维参数 Q 化方法获得,通过设计观测器可以给出多目标优化线性二次调节问题的Pareto最优解^[45]. Khargonkar等^[46]基于Youla参数化方法研究了多目标 H_2 问题. 尽管以 H_2 和 H_∞ 作为目标函数的多目标最优控制问题的研究结果很多,但是目前没有学者能求出精确的Pareto最优解或者给出Pareto最优解和Pareto次优解的近似程度. 基于此,在得不到Pareto最优解的情况下, Balandin等^[38]研究了以 γ_0 (初值条件的不确定性)范数和 H_2 范数作为目标函数的Pareto次优问题. 并且, Pareto次优解与Pareto最优解之间的误差不超过 $1 - \frac{\sqrt{N}}{N}$ (N 表示目标函数的个数). 此外,以 γ_0 范数和 H_2 范数作为目标函数, Balandin等分别推导了无限时域^[39]和有限时域^[40]的Pareto最优解.

2.2 合作差分博弈

需要指出,上面得到的都是关于多目标合作微分博弈的结果,而关于合作差分博弈的结果较少,文献^[22]仅将Engwerda^[21]微分博弈的结果推广到了有限时域合作差分博弈. 对于有限时域优化问题,横截条件是最大值原理获得必要条件的一部分. 然而,由于与约束相关的协态变量和拉格朗日乘子在无限时间表现出不可控的渐近行为,无限时域优化问题的横截条件并不完备. 更准确地说,无限时域最优控制的横截条件应该是有限时域横截条件的推广,即随着时间趋于无穷,无限时域优化问题的协态变量趋于零,与约束相关的拉格朗日乘子非负,且与目标函数相关的乘子为正数. 但是,上述结论在无限时域优化问题中均不成立. 由于离散时间无限时域最优控制的可用结果较少,例如带等式和不等式约束的最大值原理,无限时域合作差分博弈的研究相对较难.

鉴于此,对一类带有指数折扣目标函数非线性自治系统的多目标优化问题^[47],分别推导无限时域合作差分博弈Pareto最优解的充分条件和必要条件. 首先,通过引入适当的辅助状态, N 个受约束的最优控制问题被转化为一个具有受限混合终端的不受约束优化问题;其次,通过定义两个实值函数,将无限时域合作差分博弈等价地转化为增广和截断的有限时域优化问题,根据离散最大值原理推导Pareto最优性的必要条件;再次,给出一般非自治系统Pareto最优解

的充分条件,得到的非线性结果用来分析线性二次情形下的适定性和可达性问题;最后,通过离散化的广告博弈和政府债务博弈验证所得理论的正确性.

3 随机合作博弈

3.1 不定号理论

随机线性二次最优控制最早由Wonham^[48]提出,并被国内外很多学者广泛研究^[49-53]. Riccati方程理论是求解线性二次最优控制一种重要且有效的方法. 与确定的线性二次最优控制问题不同,当随机线性微分系统的扩散项受控制策略影响时,即使目标函数中的状态向量和控制向量的权重矩阵均负定,该随机线性二次优化问题也可能适定^[54]. 这种与扩散项包含的深层次的不确定性有关的现象最早由Peng^[55]发现,并定义其为奇异最优控制.

引入新的微分Riccati方程,并证明该方程的可解性是不定号随机线性二次最优控制问题适定的充分条件^[54]. 随后,借助带有矩阵伪逆和代数等式/不等式约束的广义微分Riccati方程,确定不定号随机线性二次最优控制问题适定和最优策略存在的充要条件^[56]. 对于无限时域不定号随机线性二次最优控制,借助引入的Riccati方程讨论了可解性、适定性和可达性等问题^[57-59]. 对于离散时间不定号随机线性二次优化问题,基于引入的新的广义差分Riccati方程的解,得到带有两个自由度的一簇最优控制. 此外,建立了不定号随机线性二次最优控制问题的适定性、可达性,广义倒向差分Riccati方程的可解性以及引入的线性矩阵不等式可解性之间的等价关系^[60]. 通过半正定规划获得的广义代数Riccati方程的解研究无限时域不定号随机线性二次最优控制问题,并证明无限时域线性二次最优控制的适定性等价于线性矩阵不等式的可解性,广义代数Riccati方程镇定解的存在性等价于线性二次最优控制的可达性^[61-62].

3.2 随机合作微分博弈

连续时间随机最优控制发展较早,且已取得一些重要和有意义的成果,例如,1990年文献^[63]发表的关于随机非线性最大值原理在非凸条件下的一般理论、文献^[54]从1998年开始对连续时间不定号随机线性二次优化给出了一般结果等,使得随机合作微分博弈在近十余年取得了一系列成果.

对具有状态依赖噪声的一类Itô随机弱耦合大规模系统,推导了在无限时域的Pareto最优静态输出反馈控制器^[64]. 对于广义随机合作线性二次微分博弈,分别研究了有限时域^[65]和无限时域^[66]的Pareto最优

性. 再者, 随机非线性合作微分博弈 Pareto 最优解的必要条件和充分条件也分别被给出^[67]. 此外, Pareto 最优性也被用来研究鲁棒随机控制问题, 例如, 随机非线性模糊系统的最优滤波设计^[68], 随机非线性金融系统的多目标投资策略选取^[69], 具有状态、控制、干扰依赖乘性噪声的有限时域随机合作线性二次微分博弈^[70], 有限时域^[71]和无限时域^[72]受限 H_∞ 平均场随机线性二次 Pareto 最优策略等.

3.3 随机合作差分博弈

与连续时间相比, 离散时间随机最优控制的结果较少. 不带约束的有限时域单目标优化分别得到了离散随机非线性最大值原理^[73]、带有多重时滞的离散随机线性最大值原理^[74]和离散随机 LaSalle 不变原理^[75]. 然而, 要获得多目标随机合作差分博弈的 Pareto 最优性, 需要带有混合终端约束(如终端等式或者不等式约束)的最大值原理. 因此, 上述结论^[73-75]不能直接应用到随机合作差分博弈. 基于此, 加权方法被用以研究有限时域正则随机合作线性二次差分博弈^[76]. 然而, 关于不定号随机合作线性二次差分博弈, 是否可以通过 Riccati 方程的解得到 Pareto 最优性? 加权的 Riccati 方程的解能不能保证无限时域系统的稳定性? 目标函数的适定性与凸性之间是否存在联系? 需要围绕有限时域不定号随机合作线性二次差分博弈和情况更复杂的无限时域随机合作线性二次差分博弈展开研究.

鉴于此, 研究了有限时域正则和不定号随机合作线性二次差分博弈的 Pareto 最优性^[77]. 首先, 通过定义有界线性算子序列, 得到随机线性差分系统的通解及其线性性质; 其次, 在正则合作博弈权重矩阵的约束下, 证明目标函数为凸函数, 此时通过加权方法可以很好地刻画 Pareto 最优性, 对不定号合作博弈, 推导目标函数的凸性判据, 该条件下目标函数的权和最小化等价于 Pareto 最优决策向量; 再次, 提出一个计算算法, 通过求解加权的倒向差分 Riccati 方程和加权的倒向差分 Lyapunov 方程, 得到有限时域的 Pareto 最优决策向量和 Pareto 最优解; 最后, 给出两个国家财政政策相互影响的实际例子, 仿真结果表明了所提出算法的可解性.

与有限时域随机合作线性二次差分博弈不同, 无限时域合作博弈 Pareto 最优性的求解更加复杂. 一方面, Riccati 方程理论虽然对求解有限时域和无限时域 Pareto 最优决策向量均有效, 但是区别于有限时域倒向差分 Riccati 方程的解, 无限时域非线性代数 Riccati

方程解的形式更加复杂, 无限时域对应的 Riccati 方程不仅要保证解的存在性, 而且要保证动力学方程的稳定性, 为了保证系统的稳定性以及非线性代数 Riccati 方程的可解性, 需要引入均方稳定和精确能观^[78-79]的概念; 另一方面, 相对于正则合作博弈对控制策略权重矩阵正定的要求, 不定号优化问题降低了控制策略权重矩阵的保守性, 对于连续时间随机优化问题, 当状态方程的扩散项受控制策略影响时, 不定号线性二次最优控制才具有适定性. 然而, 当离散时间线性二次最优控制的目标函数控制策略的权重矩阵不定号(特殊地, 负定), 如果动态系统受到控制策略的影响, 则在确定和随机两种情况下, 不定号优化问题都可能适定.

对于随机合作线性二次动态差分博弈的无限时域多目标最优控制^[80], 首先, 根据权重矩阵的约束和动态系统的线性性质, 推导正则合作博弈目标函数的凸性, 得到目标函数权和最小化与 Pareto 最优决策向量之间的等价性; 其次, 在精确能观的条件下, 基于加权的代数 Riccati 方程的解, 得到正则合作博弈的 Pareto 最优决策向量, 通过最优反馈增益矩阵和加权的代数 Lyapunov 方程的解确定 Pareto 最优解; 再次, 提出不定号合作博弈目标函数凸性的充要条件, 根据加权的广义代数 Riccati 方程和加权的广义代数 Lyapunov 方程的解, 结合半正定规划研究带有指数折扣不定号目标函数的 Pareto 最优性; 最后, 以带有随机因素影响的渔业管理博弈为例验证了所得结果的正确性.

4 平均场随机合作博弈

平均场类型的随机微分方程由 Kac^[81]于 1956 年和 McKean 于 1966 年^[82]分别提出. 区别于经典的随机线性二次最优控制, 平均场随机线性二次优化的状态方程和目标函数均具有状态 x_k 和控制策略 u_k 的期望值 $\mathcal{E}x_k$ 和 $\mathcal{E}u_k$, 此类优化问题是平均场理论和随机线性二次理论的结合. 在各种社会学和物理动力学模型中, 建立了平均场理论来研究个体相互作用引起的集体行为^[83-84]. 当个体的数量趋于无穷大时, 平均项接近期望值. 为了更好地说明这一性质, 假设第 i ($i \in N_L := \{1, 2, \dots, L\}$) 个个体的动力学方程为

$$x_{k+1}^{i,L} = \left(A_k x_k^{i,L} + \bar{A}_k \frac{1}{L} \sum_{j=1}^L x_k^{j,L} + B_k u_k \right) + \left(C_k x_k^{i,L} + \bar{C}_k \frac{1}{L} \sum_{j=1}^L x_k^{j,L} + D_k u_k \right) w_k^i. \quad (9)$$

在适当的假设条件下,令 $L \rightarrow \infty$,由大数定律可得

$$\frac{1}{L} \sum_{j=1}^L x_k^{j,L} \text{ 近似 } \mathcal{E}x_k, \text{ 式(9)等价于}$$

$$x_{k+1} = (Ax_k + \bar{A}\mathcal{E}x_k + Bu_k) + (Cx_k + \bar{C}\mathcal{E}x_k + Du_k)w_k. \quad (10)$$

式(10)的具体推导可以由经典的Mckean-Vlasov公式^[82]获得.

4.1 平均场随机合作微分博弈

随着平均场理论不断发展,平均场随机线性二次最优控制取得了一系列成果^[85-87].由于平均场随机线性二次最优控制目标函数中状态过程和控制过程均值 $\mathcal{E}x_k$ 和 $\mathcal{E}u_k$ 的存在,可以通过对目标函数引入方差 $\text{var}(x_k)$ 和 $\text{var}(u_k)$ 降低状态过程和控制过程对随机事件的敏感性^[85],其中,典型应用是金融投资领域的Markowitz均值—方差组合选取,此问题可通过方差量化的最小化风险解决.另外,在变分法的基础上推导了一个最优控制系统,其由一个线性的平均场正倒向随机微分方程表示.利用解耦方法得到了两个耦合的Riccati微分方程组,在一定条件下该方程组具有唯一解.当状态方程和观测方程的漂移项关于状态及其均值都是线性方程时,研究了具有可观测噪声的平均场正倒向随机微分方程的优化问题^[86].对目标函数中半正定的状态向量权重矩阵和控制向量权重矩阵,提出了无限时域平均场随机微分方程最优镇定的充要条件^[87].在系统精确能观(精确能检)的假设下,得到交叉耦合的代数Riccati方程具有唯一的正定解(半正定解).

近年来,连续时间平均场随机最优控制取得了一些重要成果,例如动态规划^[88]和最大值原理^[89-92]等,这些结果的出现加快了平均场随机多目标优化的研究进程.一个与平均场多目标优化有关的实际问题是种群随机动态博弈,它们通过相互作用对个体在集体的行为进行建模.其植根于多智能体系统,旨在根据每个智能体的信息模式和目标函数设计分布式的控制策略.随着种群数量的增加,计算的复杂度也会变得越来越高.为了克服这一缺点,提出状态聚合方法用以构造分散的 ϵ -Nash 均衡,通过状态聚合(在式(9)中当 $L \rightarrow \infty$ 时,产生状态平均值的限制)形成了分散式的控制策略^[93].对于种群随机动态博弈种群数量为无限的情形,由于直接从平均场理论的角度研究优化问题,此类平均场问题称为平均场博弈^[94].通过借助正倒向常微分方程的解,刻画了一类线性二次平均场博弈解的存在唯一性,构造分散式的

控制策略得到了 ϵ -Nash 均衡^[95].此类问题也被称为平均场博弈,其中一个关键点是平均值或总体状态平均值在目标函数中呈现非线性.

另一方面,对于有限时域平均场随机合作微分博弈,给出了Pareto最优解的必要条件和充分条件,讨论了正则线性二次情形下的Pareto最优解^[96].值得注意的是,正则合作博弈的Pareto最优决策向量与加权方法等价.然而,由于不定号合作博弈的目标函数为非凸函数,借助加权方法不足以求解Pareto最优决策向量.此外,在交叉耦合的代数Riccati方程有解的假设下,讨论了无限时域正则随机合作线性二次微分博弈的Pareto最优性^[97-98],但是并未给出交叉耦合的代数Riccati方程解的存在条件.

4.2 平均场随机合作差分博弈

对于离散时间情形,Hilbert空间中的二次最优控制方法、矩阵最小值原理方法、算子线性二次方法和配方法被提出,用以研究有限时域正则平均场随机线性二次优化^[83].对于有限时域不定号平均场随机线性二次优化问题,提出了最优控制和最优目标函数,并将其应用到多期均值方差投资组合选择问题^[84].然而,不定号多目标合作博弈的适定性与可达性、广义倒向差分Riccati方程的解是否可以确定Pareto最优决策向量等这些问题并未讨论.接下来,无限时域正则^[99]、不定号^[100]和非正则^[101]平均场随机线性二次优化问题被分别研究.需要指出,系统在精确能观和精确能检条件下的平均场随机线性二次优化问题被深入研究^[87,101],但并没有给出系统精确能观、精确能检的一般性判据,也并未讨论交叉耦合的代数Lyapunov方程解的情况.

基于上述分析,对不定号平均场随机合作线性二次动态差分博弈^[102],首先,推导引入的 N 个交叉耦合的广义倒向差分Riccati方程的可解性与多目标优化问题的可解性之间的等价性;其次,因为不同玩家目标函数的权重矩阵不同,博弈中依赖于权重矩阵的所有玩家最小化某个玩家和最小化其他玩家的最优联合策略不同,因此,很难基于 N 个交叉耦合的广义倒向差分Riccati方程的解确定Pareto最优决策向量;再次,推导目标函数凸性的充要条件,使得加权方法是推导平均场Pareto最优决策向量的充要条件,基于加权的交叉耦合的广义倒向差分Riccati方程的解和加权的交叉耦合的广义倒向差分Lyapunov方程的解,给出平均场Pareto最优性算法,其在理论上可以得到所有的Pareto最优决策向量和Pareto最优解;最后,提

出一个合作的网络安全博弈以验证所得理论结果,数值仿真结果验证了算法的可解性、正确性和有效性.

在无限时域随机线性二次优化问题中,稳定性^[103-104]、精确能观、精确能检^[62,78-79]是基本假设且发挥着重要作用,例如,平均场非正则随机线性二次最优控制^[87,101]、平均场正则随机线性二次最优控制^[99]、随机线性二次 Stackelberg 差分博弈^[105]等. 对于无限时域平均场随机线性二次优化问题^[87,99,101],在平均场随机线性系统精确能观、精确能检的假设下,可以得到交叉耦合的代数 Riccati 方程的解,从而确定最优目标函数和最优控制策略. 然而,平均场随机线性系统精确能观、精确能检的 PBH 特征向量判据,尤其是该判据的充要条件还未给出.

鉴于此,对于无限时域平均场随机合作线性二次差分博弈的 Pareto 最优性^[106],首先,基于 H -表示方法^[107],得到平均场随机线性差分系统均方稳定和精确能观、精确能检 Popov-Belevith-Hautus (PBH) 特征向量判据的充要条件;其次,在系统均方稳定和精确能观、精确能检的条件下,讨论交叉耦合的广义 Lyapunov 方程的解,并在平均场随机线性系统精确能观(精确能检)的条件下,得到无限时域交叉耦合的代数 Riccati 方程有唯一的正定解(半正定解),其也是反馈镇定解;再次,给出每个玩家目标函数对应的系统和权和目标函数对应的系统关于精确能观性的关系,所有 Pareto 最优决策向量和 Pareto 最优解分别通过求解加权交叉耦合代数 Riccati 方程和加权交叉耦合代数 Lyapunov 方程的解获得;最后,通过多接入边缘计算网络中的迁移计算验证推导结果的正确性.

5 仿真算例

本节考虑将两个国家的有限时域财政政策博弈^[47]推广到无限时域,两个国家财政当局所受的约束方程为

$$x_{k+1} = a_1 x_k + b_{1,1} u_{1,k} + b_{2,1} u_{2,k} + a_2 x_k \omega_k. \quad (11)$$

其中: x_k 为两个国家的财富差, $u_{1,k}$ 和 $u_{2,k}$ 分别为由国家 1 和国家 2 的财政当局支配的财政赤字, $a_2 x_k \omega_k$ 为随机事件(如各部分政策执行能力或决策者的自身素质等)造成的持续内在影响. 两个国家的财政当局均希望使如下跨时期目标函数最小:

$$\mathcal{J}_i(x_0; u_{1,k}, u_{2,k}) = \mathcal{E} \sum_{k=0}^{\infty} (q_i x_k^2 + r_{i1} u_{1,k}^2 + r_{i2} u_{2,k}^2). \quad (12)$$

假设两个国家决定通过合作以消除因初始财富不同导致的两个国家的财富差. 选取参数 $a_1 = 0.6, b_{1,1} = 1, b_{2,1} = -1, a_2 = 0.3, q_1 = 1, q_2 = 0.5, r_{11} = 1, r_{12} = 0.1, r_{21} = 0.2, r_{22} = 2$, 初值 x_0 是期望为 1、方差为 5 的均匀分布. 基于第 1 节的加权方法和随机线性二次最优控制理论,得到仿真结果如图 1~图 3 所示.

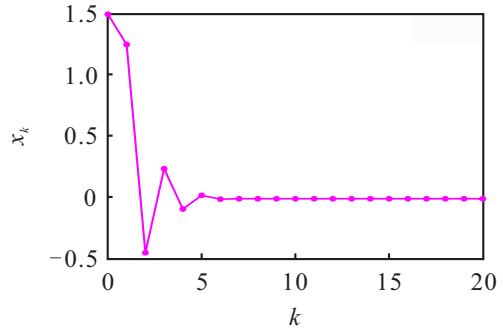


图 1 两个国家间的最优财富差

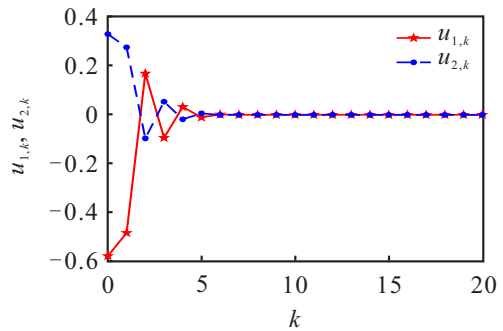


图 2 两个国家财政当局的最优控制策略

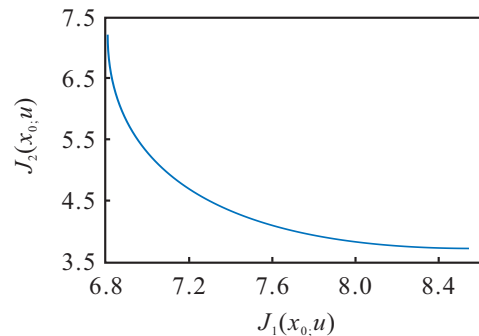


图 3 两个国家财政当局跨时期目标函数的 Pareto 边界

由仿真结果可以看出,两个国家的财政当局通过合作(如图 3 所示,两个国家财政当局跨时期目标函数的仿真结果满足 Pareto 最优性的定义),采取相应的控制策略(如图 2 所示),可以有效降低两个国家间的财富差(如图 1 所示).

6 结语

本文总结了多目标动态优化中 Pareto 合作博弈取得的研究结果,全面介绍了确定性合作博弈、随机合作博弈以及平均场随机合作博弈分别在连续时间和离散时间情形下的发展,概括为表 1.

表1 合作博弈研究文献分类

| 博弈类型 | 博弈研究内容 |
|-----------|---|
| 确定性合作博弈 | 合作微分博弈 ^[20-21,37,43-44] Pareto最优性的多目标干扰衰减 ^[38-40,45-46] 合作差分博弈 ^[22,47] |
| 随机合作博弈 | 随机合作微分博弈 ^[64-67] Pareto最优性的随机鲁棒控制 ^[68-72] 随机合作差分博弈 ^[77,80] |
| 平均场随机合作博弈 | 平均场随机合作微分博弈 ^[96] 平均场随机合作差分博弈 ^[102,106] |

由表1可见,Pareto合作博弈的研究成果较为丰富,但是关于平均场随机合作博弈以及随机合作差分博弈的研究仍十分有限,许多问题值得继续研究:

1) 随机非线性合作差分博弈. 文献[77, 80]分别研究了有限时域和无限时域在正则和不定号情形下的随机合作线性二次差分博弈. 如何推导带有混合等式、不等式终端约束的离散随机最大值原理,是研究随机非线性合作差分博弈Pareto最优性必要条件面临的一个大的挑战.

2) 无限时域不定号平均场随机合作线性二次差分博弈. 文献[106]研究了无限时域正则平均场随机合作线性二次差分博弈. 如何通过分析系统的稳定性、确定交叉耦合的广义代数Riccati方程的可解性,并求解不定号权重矩阵的适定性边界等问题,是获得无限时域不定号平均场随机合作线性二次差分博弈的所有Pareto最优决策向量和Pareto最优解需要解决的难题.

3) 放宽Riccati方程的求解条件. 无限时域正则随机合作线性二次微分和差分博弈的Pareto最优性均是基于代数Riccati方程在精确能观或精确能检条件下确定的解,借助适当放宽条件获得的Riccati方程的其他解确定Pareto最优性可以较大程度地提高随机合作博弈的适用性.

4) 平均场随机非线性多目标优化的Pareto最优性. 平均场随机合作线性二次微分博弈和非线性合作微分博弈,以及平均场随机合作线性二次差分博弈的结果均已给出,如何推导离散时间平均场随机优化理论,如最大值原理和动态规划等,是研究平均场随机非线性合作差分博弈的关键.

参考文献(References)

[1] Boltyanski V G, Gamkrelidze R V, Pontryagin L S. On the theory of optimal processes[J]. Doklady Akademii NaukaSSSR, 1956, 10: 7-10.
 [2] Bellman R. On the theory of dynamic programming[J].

Proceedings of the National Academy of Sciences of the United States of America, 1952, 38(8): 716-719.
 [3] Bellman R. Dynamic programming[M]. Princeton: Princeton University Press, 1957.
 [4] Kalman R E. A new approach to linear filtering and prediction problems[J]. Journal of Basic Engineering, 1960, 82(1): 35-45.
 [5] Kalman R E. Contributions to the theory of optimal control[J]. Boletindela Sociedad Matematica Mexicana, 1960, 5: 102-119.
 [6] 解学书. 最优控制——理论与应用[M]. 北京: 清华大学出版社, 1986.
 (Xie X S. Optimal control theory and application[M]. Beijing: Tsinghua University Press, 1986.)
 [7] 张杰, 王飞跃. 最优控制: 数学理论与智能方法[M]. 北京: 清华大学出版社, 2017.
 (Zhang J, Wang F Y. Optimal control: Mathematical theory and intelligent method[M]. Beijing: Tsinghua University Press, 2017.)
 [8] Edgeworth F Y. Mathematical psychics: An essay on the application of mathematics to the moral sciences[M]. London: C. K. Paul, 1881.
 [9] Pareto V. Cours d'economic politique[M]. Lausanne: Duncker & Humblot, 1896.
 [10] Pareto V. Manual of political economy[M]. Oxford: Oxford University Press, 1927.
 [11] von Neumann J. Zur theorie der gesellschaftsspiele[J]. Mathematische Annalen, 1928, 100(1): 295-320.
 [12] von Neumann J. On the theory of games of strategy[J]. Contributions to the Theory of Games, 1959, 4: 13-42.
 [13] von Neumann J, Morgenstern O. Theory of games and economic behavior[M]. Princeton: Princeton University Press, 1944.
 [14] Nash J F. Equilibrium points in *N*-person games[J]. Proceedings of the National Academy of Sciences of the United States of America, 1950, 36(1): 48-49.
 [15] Basar T, Olsder G J. Dynamic noncooperative game theory[M]. The 2nd edition. Philadelphia: SIAM, 1999.
 [16] Yeung D W K, Petrosyan L A. Cooperative stochastic differential games[M]. New York: Springer, 2006.
 [17] Filar J A, Gaitsgory V, Mizukami K. Advances in dynamic games and applications[M]. Boston: Birkhäuser, 2000.
 [18] 郭雷, 程代展, 冯德兴. 控制理论导论: 从基本概念到研究前沿[M]. 北京: 科学出版社, 2005.
 (Guo L, Cheng D Z, Feng D X. Introduction to control theory: From basic concepts to research frontiers[M]. Beijing: Science Press, 2005.)
 [19] 袁硕, 郭雷. 随机自适应动态博弈[J]. 中国科学: 数学, 2016, 46(10): 1367-1382.
 (Yuan S, Guo L. Stochastic adaptive dynamical

- games[J]. *Scientia Sinica: Mathematica*, 2016, 46(10): 1367-1382.)
- [20] Engwerda J C. The regular convex cooperative linear quadratic control problem[J]. *Automatica*, 2008, 44(9): 2453-2457.
- [21] Engwerda J C. Necessary and sufficient conditions for Pareto optimal solutions of cooperative differential games[J]. *SIAM Journal on Control and Optimization*, 2010, 48(6): 3859-3881.
- [22] Lin Y N, Zhang W H. Necessary/sufficient conditions for Pareto optimum in cooperative difference game[J]. *Optimal Control Applications and Methods*, 2018, 39(2): 1043-1060.
- [23] 孙浩, 杨景明, 刘醒, 等. 基于环境 Pareto 支配选择策略的有约束多目标差分进化算法[J]. *控制与决策*, 2016, 31(1): 45-51.
(Sun H, Yang J M, Liu X, et al. Differential evolution algorithm based on environment Pareto dominated selection strategy in constrained multi-objective optimization problem[J]. *Control and Decision*, 2016, 31(1): 45-51.)
- [24] Zhou X, Wang H M, Peng W, et al. Solving multi-scenario cardinality constrained optimization problems via multi-objective evolutionary algorithms[J]. *Science China Information Sciences*, 2019, 62(9): 192104.
- [25] 封文清, 巩敦卫. 基于在线感知 Pareto 前沿划分目标空间的多目标进化优化[J]. *自动化学报*, 2020, 46(8): 1628-1643.
(Feng W Q, Gong D W. Multi-objective evolutionary optimization with objective space partition based on online perception of Pareto front[J]. *Acta Automatica Sinica*, 2020, 46(8): 1628-1643.)
- [26] Hua Y C, Liu Q Q, Hao K R, et al. A survey of evolutionary algorithms for multi-objective optimization problems with irregular Pareto fronts[J]. *IEEE/CAA Journal of Automatica Sinica*, 2021, 8(2): 303-318.
- [27] Huang Y B, Zhao J. Pareto efficiency of finite horizon switched linear quadratic differential games[J]. *Journal of Systems Science and Complexity*, 2018, 31(1): 173-187.
- [28] 张晗, 杨继斌, 张继业, 等. 燃料电池有轨电车能量管理 Pareto 多目标优化[J]. *自动化学报*, 2019, 45(12): 2378-2392.
(Zhang H, Yang J B, Zhang J Y, et al. Pareto-based multi-objective optimization of energy management for fuel cell tramway[J]. *Acta Automatica Sinica*, 2019, 45(12): 2378-2392.)
- [29] Mahmoodabadi M J, Mostaghim S A. Stability of nonlinear systems using optimal fuzzy controllers and its simulation by Java programming[J]. *IEEE/CAA Journal of Automatica Sinica*, 2017, 6(6): 1519-1527.
- [30] Na X X, Cole D J. Modelling of a human driver's interaction with vehicle automated steering using cooperative game theory[J]. *IEEE/CAA Journal of Automatica Sinica*, 2019, 6(5): 1095-1107.
- [31] Lv Z M, Wang L Q, Han Z Y, et al. Surrogate-assisted particle swarm optimization algorithm with Pareto active learning for expensive multi-objective optimization[J]. *IEEE/CAA Journal of Automatica Sinica*, 2019, 6(3): 838-849.
- [32] Fu X Y, Chan F T S, Niu B, et al. A multi-objective pigeon inspired optimization algorithm for fuzzy production scheduling problem considering mould maintenance[J]. *Science China Information Sciences*, 2019, 62(7): 70202.
- [33] 丁进良, 杨翠娥, 陈立鹏, 等. 基于参考点预测的动态多目标优化算法[J]. *自动化学报*, 2017, 43(2): 313-320.
(Ding J L, Yang C E, Chen L P, et al. Dynamic multi-objective optimization algorithm based on reference point prediction[J]. *Acta Automatica Sinica*, 2017, 43(2): 313-320.)
- [34] Gass S, Saaty T. The computational algorithm for the parametric objective function[J]. *Naval Research Logistics Quarterly*, 1955, 2(1/2): 39-45.
- [35] Zadeh L A. Optimality and non-scalar-valued performance criteria[J]. *IEEE Transactions on Automatic Control*, 1963, 8(1): 59-60.
- [36] Collette Y, Siarry P. *Principles and case studies*[M]. Berlin, Heidelberg: Springer, 2004.
- [37] Menini L, Possieri C, Tornambè A. Algebraic methods for multiobjective optimal design of control feedbacks for linear systems[J]. *IEEE Transactions on Automatic Control*, 2018, 63(12): 4188-4203.
- [38] Balandin D V, Kogan M M. Pareto suboptimal controllers in multi-objective disturbance attenuation problems[J]. *Automatica*, 2017, 84: 56-61.
- [39] Balandin D V, Kogan M M. Multi-objective generalized control[J]. *Automatica*, 2019, 99: 317-322.
- [40] Balandin D V, Biryukov R S, Kogan M M. Finite-horizon multi-objective generalized H_2 control with transients[J]. *Automatica*, 2019, 106: 27-34.
- [41] Blume L, Simon C P. *Mathematics for economists*[M]. New York: Norton, 1994.
- [42] Engwerda J C. *LQ dynamic optimization and differential games*[M]. Chichester: Wiley & Sons, 2005.
- [43] Reddy P V, Engwerda J C. Pareto optimality in infinite horizon linear quadratic differential games[J]. *Automatica*, 2013, 49(6): 1705-1714.
- [44] Reddy P V, Engwerda J C. Necessary and sufficient conditions for Pareto optimality in infinite horizon

- cooperative differential games[J]. IEEE Transactions on Automatic Control, 2014, 59(9): 2536-2542.
- [45] Makila P M. On multiple criteria stationary linear quadratic control[J]. IEEE Transactions on Automatic Control, 1989, 34(12): 1311-1313.
- [46] Khargonekar P P, Rotea M A. Multiple objective optimal control of linear systems: The quadratic norm case[J]. IEEE Transactions on Automatic Control, 1991, 36(1): 14-24.
- [47] Peng C C, Zhang W H. Multiobjective dynamic optimization of cooperative difference games in infinite horizon[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(11): 6669-6680.
- [48] Wonham W M. On a matrix Riccati equation of stochastic control[J]. SIAM Journal on Control and Optimization, 1968, 6(4): 681-697.
- [49] Bismut J M. Linear quadratic optimal stochastic control with random coefficients[J]. SIAM Journal on Control and Optimization, 1976, 14(3): 419-444.
- [50] Yong J M, Zhou X Y. Stochastic controls: Hamiltonian systems and HJB equations[M]. New York: Springer, 1999.
- [51] Kohlmann M, Tang S J. Minimization of risk and linear quadratic optimal control theory[J]. SIAM Journal on Control and Optimization, 2003, 42(3): 1118-1142.
- [52] Costa O L V, de Paulo W L. Indefinite quadratic with linear costs optimal control of Markov jump with multiplicative noise systems[J]. Automatica, 2007, 43(4): 587-597.
- [53] Duncan T E. Linear-exponential-quadratic Gaussian control[J]. IEEE Transactions on Automatic Control, 2013, 58(11): 2910-2911.
- [54] Chen S P, Li X J, Zhou X Y. Stochastic linear quadratic regulators with indefinite control weight costs[J]. SIAM Journal on Control and Optimization, 1998, 36(5): 1685-1702.
- [55] Peng S G. New developments in stochastic maximum principle and related backward stochastic differential equations[C]. Proceedings of the 31st IEEE Conference on Decision and Control. Tucson, 2002: 2043-2047.
- [56] Ait Rami M, Moore J, Zhou X Y. Indefinite stochastic linear quadratic control and generalized differential Riccati equation[J]. SIAM Journal on Control and Optimization, 2001, 40(4): 1296-1311.
- [57] Ait Rami M, Zhou X Y. Linear matrix inequalities, Riccati equations, and indefinite stochastic linear quadratic controls[J]. IEEE Transactions on Automatic Control, 2000, 45(6): 1131-1143.
- [58] Ait Rami M, Chen X, Moore J B, et al. Solvability and asymptotic behavior of generalized Riccati equations arising in indefinite stochastic LQ controls[J]. IEEE Transactions on Automatic Control, 2001, 46(3): 428-440.
- [59] Ait Rami M, Zhou X Y, Moore J B. Well-posedness and attainability of indefinite stochastic linear quadratic control in infinite time horizon[J]. Systems & Control Letters, 2000, 41: 123-133.
- [60] Ait Rami M, Chen X, Zhou X Y. Discrete-time indefinite LQ control with state and control dependent noises[J]. Journal of Global Optimization, 2002, 23(3): 245-265.
- [61] Zhang W H, Li Y, Liu X K. Infinite horizon indefinite stochastic linear quadratic control for discrete-time systems[J]. Control Theory and Technology, 2015, 13(3): 230-237.
- [62] Zhang W H, Xie L H, Chen B S. Stochastic H_2/H_∞ control: A nash game approach[M]. Boca Raton: CRC Press, 2017.
- [63] Peng S G. A general stochastic maximum principle for optimal control problems[J]. SIAM Journal on Control and Optimization, 1990, 28(4): 966-979.
- [64] Mukaidani H, Xu H. Pareto optimal strategy for stochastic weakly coupled large scale systems with state dependent system noise[J]. IEEE Transactions on Automatic Control, 2009, 54(9): 2244-2250.
- [65] Zhang W H, Lin Y N, Xue L R. Linear quadratic Pareto optimal control problem of stochastic singular systems[J]. Journal of the Franklin Institute, 2017, 354(2): 1220-1238.
- [66] Lin Y N, Zhang T L, Zhang W H. Infinite horizon linear quadratic Pareto game of the stochastic singular systems[J]. Journal of the Franklin Institute, 2018, 355(10): 4436-4452.
- [67] Lin Y N, Jiang X S, Zhang W H. Necessary and sufficient conditions for Pareto optimality of the stochastic systems in infinite horizon[J]. Automatica, 2018, 94: 341-348.
- [68] Chen B S, Lee H C, Wu C F. Pareto optimal filter design for nonlinear stochastic fuzzy systems via multiobjective optimization[J]. IEEE Transactions on Fuzzy Systems, 2015, 23(2): 387-399.
- [69] Wu C F, Chen B S, Zhang W H. Multiobjective investment policy for a nonlinear stochastic financial system: A fuzzy approach[J]. IEEE Transactions on Fuzzy Systems, 2017, 25(2): 460-474.
- [70] Jiang X S, Tian S P, Zhang T L, et al. Pareto optimal strategy for linear stochastic systems with H_∞ constraint in finite horizon[J]. Information Sciences, 2020, 512: 1103-1117.
- [71] Jiang X S, Tian S P, Zhang W H, et al. Pareto-optimal strategy for linear mean-field stochastic systems with H_∞ constraint[J]. IEEE Transactions on Cybernetics, 2022,

- 52(5): 2846-2859.
- [72] Jiang X S, Su S F, Zhao D Y. Pareto optimal strategy under H_∞ constraint for the mean-field stochastic systems in infinite horizon[J]. IEEE Transactions on Cybernetics, DOI: 10.1109/TCYB.2022.3179605.
- [73] Lin X Y, Zhang W H. A maximum principle for optimal control of discrete-time stochastic systems with multiplicative noise[J]. IEEE Transactions on Automatic Control, 2015, 60(4): 1121-1126.
- [74] Zhang H S, Wang H X, Li L. Adapted and casual maximum principle and analytical solution to optimal control for stochastic multiplicative-noise systems with multiple input-delays[C]. IEEE 51st IEEE Conference on Decision and Control. Maui, 2013: 2122-2127.
- [75] Zhang W H, Lin X Y, Chen B S. LaSalle-type theorem and its applications to infinite horizon optimal control of discrete-time nonlinear stochastic systems[J]. IEEE Transactions on Automatic Control, 2017, 62(1): 250-261.
- [76] Zhu H N, Zhang C K. Finite horizon linear quadratic dynamic games for discrete-time stochastic systems with N-players[J]. Operations Research Letters, 2016, 44(3): 307-312.
- [77] Peng C C, Zhang W H. Multicriteria optimization problems of finite horizon stochastic cooperative linear-quadratic difference games[J]. Science China Information Sciences, 2022, 65(7): 172203.
- [78] Zhang W H, Chen B S. On stabilizability and exact observability of stochastic systems with their applications[J]. Automatica, 2004, 40(1): 87-94.
- [79] Zhang W H, Zheng W X, Chen B S. Detectability, observability and Lyapunov-type theorems of linear discrete time-varying stochastic systems with multiplicative noise[J]. International Journal of Control, 2017, 90(11): 2490-2507.
- [80] Peng C C, Zhang W H, Ma L M. Infinite horizon multiobjective optimal control of stochastic cooperative linear-quadratic dynamic difference games[J]. Journal of the Franklin Institute, 2021, 358(16): 8288-8307.
- [81] Kac K. Foundations of kinetic theory[C]. Proceeding of the 3rd Berkeley Symposium on Mathematical Statistics and Probability. Berkeley, 1956, 3: 171-197.
- [82] McKean H P. A class of Markov processes associated with nonlinear parabolic equations[J]. Proceedings of the National Academy of Sciences of the United States of America, 1966, 56(6): 1907-1911.
- [83] Elliott R, Li X, Ni Y H. Discrete time mean-field stochastic linear-quadratic optimal control problems[J]. Automatica, 2013, 49(11): 3222-3233.
- [84] Ni Y H, Zhang J F, Li X. Indenite mean-field stochastic linear quadratic optimal control[J]. IEEE Transactions on Automatic Control, 2015, 60(7): 1786-1800.
- [85] Yong J M. Linear-quadratic optimal control problems for mean-field stochastic differential equations[J]. SIAM Journal on Control and Optimization, 2013, 51(4): 2809-2838.
- [86] Wang G C, Xiao H, Xing G J. An optimal control problem for mean-field forward-backward stochastic differential equation with noisy observation[J]. Automatica, 2017, 86: 104-109.
- [87] Qi Q Y, Zhang H S, Wu Z. Stabilization control for linear continuous-time mean-field systems[J]. IEEE Transactions on Automatic Control, 2019, 64(8): 3461-3468.
- [88] Ahmed N U, Ding X. Controlled McKean-Vlasov equations[J]. Communications on Applied Analysis, 2001, 5: 183-206.
- [89] Buckdahn R, Djehiche B, Li J, et al. Mean-field backward stochastic differential equations: A limit approach[J]. Annals of Probability, 2009, 37(4): 1524-1565.
- [90] Andersson D, Djehiche B. A maximum principle for SDEs of mean-field type[J]. Applied Mathematics and Optimization, 2011, 63(3): 341-356.
- [91] Buckdahn R, Djehiche B, Li J. A general maximum principle for SDEs of mean-field type[J]. Applied Mathematics and Optimization, 2011, 64: 197-216.
- [92] Li J. Stochastic maximum principle in the mean-field controls[J]. Automatica, 2012, 48(2): 366-373.
- [93] Huang M Y, Caines P E, Malhame R P. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ε -Nash equilibria[J]. IEEE Transactions on Automatic Control, 2007, 52(9): 1560-1571.
- [94] Lasry J M, Lions P L. Mean field games[J]. Japanese Journal of Mathematics, 2007, 2(1): 229-260.
- [95] Bensoussan A, Sung K C J, Yam S C P, et al. Linear-quadratic mean field games[J]. Journal of Optimization Theory and Applications, 2016, 169(2): 496-529.
- [96] Lin Y N. Necessary/sufficient conditions for Pareto optimality in finite horizon mean-field type stochastic differential game[J]. Automatica, 2020, 119: 108951.
- [97] Lin Y N, Zhang T L, Zhang W H. Pareto-based guaranteed cost control of the uncertain mean-field stochastic systems in infinite horizon[J]. Automatica, 2018, 92: 197-209.
- [98] Lin Y N, Zhang W H. Essays on Pareto optimality in cooperative games[M]. Singapore: Springer Nature Singapore, 2022.
- [99] Ni Y H, Elliott R, Li X. Discrete-time mean-field stochastic linear-quadratic optimal control problems[J].

- Automatica, 2015, 57: 65-77.
- [100] Ni Y H, Li X, Zhang J F. Indefinite mean-field stochastic linear-quadratic optimal control: From finite horizon to infinite horizon[J]. IEEE Transactions on Automatic Control, 2016, 61(11): 3269-3284.
- [101] Zhang H S, Qi Q Y, Fu M Y. Optimal stabilization control for discrete-time mean-field stochastic systems[J]. IEEE Transactions on Automatic Control, 2019, 64(3): 1125-1136.
- [102] Zhang W H, Peng C C. Indefinite mean-field stochastic cooperative linear-quadratic dynamic difference game with its application to the network security model[J]. IEEE Transactions on Cybernetics, 2022, 52(11): 11805-11818.
- [103] Zhang T L, Deng F Q, Sun Y, et al. Fault estimation and fault-tolerant control for linear discrete time-varying stochastic systems[J]. Science China Information Sciences, 2021, 64(10): 200201.
- [104] Jiang X S, Tian S P, Zhang W H. p th moment exponential stability of general nonlinear discrete-time stochastic systems[J]. Science China Information Sciences, 2021, 64(10): 209204.
- [105] Mukaidani H, Xu H. Infinite horizon linear-quadratic Stackelberg games for discrete-time stochastic systems[J]. Automatica, 2017, 76: 301-308.
- [106] Peng C C, Zhang W H. Pareto optimality in infinite horizon mean-field stochastic cooperative linear-quadratic difference games[J]. IEEE Transactions on Automatic Control, DOI: 10.1109/TAC.2022.3202824.
- [107] Zhang W H, Chen B S. H -representation and applications to generalized Lyapunov equations and linear stochastic systems[J]. IEEE Transactions on Automatic Control, 2012, 57(12): 3009-3022.

作者简介

张维海(1965—),男,教授,博士生导师,从事随机系统的鲁棒控制、随机系统的算子谱分析、离散随机最大值原理和LaSalle不变原理等研究, E-mail: w_hzhang@163.com;
彭称称(1990—),男,副教授,博士,从事多目标动态优

化、博弈理论、随机最优控制等研究, E-mail: pengchenchen1029@163.com;

蒋秀珊(1992—),女,副教授,博士,从事随机稳定性分析、随机鲁棒控制、多目标动态优化等研究, E-mail: jiangxsjy@163.com.

科研团队简介

张维海教授科研团队长期专注于随机控制理论及其应用以及多目标优化等方面的研究,获得了一系列原创性和系统性的研究成果,受到了国际学术界的高度评价。在国际上率先建立了一套随机时不变系统稳定性的算子谱分析理论,从而可以更加深刻地刻画随机稳定性;率先从物理的观点精确定义了随机系统的能观性和能检测性等概念,并给出了易于验证的PBH判据;建立了离散随机乘性噪声系统的最大值原理、LaSalle不变原理,在非线性和离散随机系统的最优控制方面做出了重要贡献。张维海教授团队在国际上较早开展随机系统的合作博弈Pareto策略研究,获得了高水平的研究成果。

团队带头人张维海教授主持和承担国家及省部级项目10多项,发表SCI期刊论文200余篇,在CRC和Spring出版社出版英文专著2部。连续2年入选全球前2项尖科学家“终身科学影响力排行榜”榜单(2021年和2022年)。以首位完成人获教育部和山东省自然科学二等奖各2项。作为指导教师获山东省优秀博士学位论文6篇。张维海教授是两个聘期的山东省“泰山学者”特聘教授,山东省有突出贡献的中青年专家。现任中国自动化学会控制理论专业委员会委员、信息物理系统专业委员会委员、中国工业与应用数学学会系统控制专业委员会副主任委员、山东省自动化学会常务理事、IEEE高级会员。担任国际SCI期刊“Asian Journal of Control”“Journal of The Franklin Institute”“International Journal of Fuzzy Systems”编委,以及《控制与决策》《系统科学与数学》期刊编委。