

关于具有不可分离哈密顿量的平均场控制系统的策略迭代算法的研究

唐毅雄

中国地质大学(武汉)数学与物理学院, 湖北 武汉

收稿日期: 2023年4月28日; 录用日期: 2023年5月21日; 发布日期: 2023年5月29日

摘要

平均场控制研究的是在大量智能体参与的系统中, 由一个中心控制者计算系统的帕累托最优, 且所有个体先验地使用完全相同的反馈控制。本文研究的对象是一类来自于平均场控制问题的倒向 - 正向偏微分方程组有限差分数值计算方法, 使用的算法基于最优控制问题中的策略迭代方法。本文讨论了较短时间区域内算法的收敛性和收敛速度。本文通过一些例子验证了算法的有效性。本文通过数值试验讨论了平均场博弈纳什均衡和平均场控制帕累托最优之间的总体效益比较。

关键词

平均场控制, 微分博弈, 最优控制, 数值计算方法

A Study on the Policy Iteration Method for Time-Dependent Mean Field Controls Systems with Non-Separable Hamiltonians

Yixiong Tang

School of Mathematics and Physics, China University of Geosciences, Wuhan Hubei

Received: Apr. 28th, 2023; accepted: May 21st, 2023; published: May 29th, 2023

Abstract

Mean field type control studies the Pareto optimum computed by a central planner in a system

文章引用: 唐毅雄. 关于具有不可分离哈密顿量的平均场控制系统的策略迭代算法的研究[J]. 应用数学进展, 2023, 12(5): 2364-2375. DOI: 10.12677/aam.2023.125239

with large number of agents, with each agent using a priori the same feedback control. This paper studies the finite difference numerical approximation of a type of Mean field type control problem, using algorithms based on the policy iteration method from optimal control problems. Under the assumption of a short time horizon, convergence of the algorithm and the rate of convergence are discussed. Some numerical examples are supplemented to show the efficacy of the algorithms. The comparisons between Nash equilibrium and Pareto optimum, in terms of aggregate utility, are illustrated by numerical examples.

Keywords

Mean Field Type Control, Differential Games, Optimal Control, Numerical Methods

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近年来，随着人工智能技术的发展，面对日益复杂的系统控制和决策问题，人们研究大规模多智能体系统的平均行为和博弈策略，提出平均场博弈模型(Mean Field Game，简称 MFG)和平均场控制系统(Mean Field Control，简称 MFC)来指导个体的决策和行动，从而实现系统整体的优化和协同。为了比较效率，定义“Price of Anarchy”(简称 POA)为纳什均衡下全局效用与最优全局效用之间的比率。

自 2006 年 Lasry 和 Lions [1]、Huang 等[2]首次提出平均场博弈的数学 PDE 模型以来，MFG 与 MFC 广泛应用于交通拥堵、能源市场、金融市场等领域[1]。Achdou 等[3]从 PDE 的角度对平均场控制系统进行理论研究，给出了一些例子的 PDE 的存在性和唯一性结果，并通过简单的数值模拟比较了 MFG 和 MFC。Carmona 等[4]从概率的角度研究了 MFC 的解的性质，并证明了对应 PDE 的解是 MFG 的近似纳什均衡。Bensoussan 等[5]全面介绍并比较了 MFG 与 MFC 问题。文献[6] [7]介绍了平均场博弈系统的有限差分等数值解法，提出了前向 - 后向系统的数值格式，并对其数值效果进行了验证。Cacace 等[8]对有可分离哈密顿量的 MFG 引入了的策略迭代算法。Camilli 等[9]提到了分析了 MFG 中策略迭代算法的收敛条件和收敛速度的估计。Cardaliaguet 等[10]介绍了 MFG 系统中 POA 的意义，并深入探讨了 POA 的计算和优化问题。Carmona 等[11]分析了对于社会最优成本而言 MFC 与 MFG 的区别，并讨论了影响 POA 的参数。

数值解对理论研究和实际问题很重要。但是目前的很多数值解法只适用在可分离哈密顿量的条件下，而对 MFC 的理论研究大部分针对的是非局部依赖的条件。策略迭代算法是最优控制问题中常见的求解的方法，已经成功应用于 MFG 问题中。但是将策略迭代算法扩展到不可分离哈密顿量的平均场控制问题目前尚缺乏研究。本论文的主要目标是提出一种策略迭代算法来解决非可分离哈密顿量的 MFC 问题。在一定的假设下，我们证明算法收敛于最优策略。我们使用该算法计算 POA，讨论了 MFC 问题中几类参数对 POA 的影响。

2. 主要内容

2.1. MFG 与 MFC

平均场博弈问题通常由以下方程组描述：

$$\begin{cases} -\partial_t u - v \Delta u + H(m, Du) = f(x, m) \\ \partial_t m - v \Delta m - \operatorname{div}(m H_p(m, Du)) = 0 \\ m(x, 0) = m_0(x), u(x, T) = u_T(x) \end{cases} \quad (1)$$

平均场型控制问题与 MFG 的不同之处在于，虽然没有智能体能够单独偏离其策略以最大化其个人利益，但有一个监督者要让群体合作使得共同利益最大化。所以，我们不能假设我们知道平均状态的密度分布，然后得到与该平均值对应的最优控制，因为这里平均值取决于监督者设置的控制。我们参考 Graber [12] 文章中第 5 节，用最优控制问题的帕累托最优来刻画 MFC 系统的 PDE。

平均场类型的控制问题(MFC)通常由以下方程组描述：

$$\begin{cases} -\partial_t u - v \Delta u + Du \cdot q - L(m, q) - mL_m(m, q) = 0 \\ \partial_t m - v \Delta m - \operatorname{div}(m H_p(m, Du)) = 0 \\ m(x, 0) = m_0(x), u(x, T) = u_T(x) \end{cases} \quad (2)$$

虽然 MFC 与 MFG 的可以用类似的 PDE 刻画最优解，但是 MFC 中 u 不代表个体的价值函数，它的方程只是有类似于 HJB 方程的结构，所以可以利用 HJB 相关的数学方法来解决。

2.2. 算法

首先我们对后文出现的哈密顿量 $H(m, p)$ 做出假设：

(H1) H 关于 p, m 是连续的。 $H, H_{p_i}, H_{p_i p_j}, H_{m p_i}$ 是关于 $(p, m) \in \mathbb{R}^d \times \mathbb{R}^+$ 的局部 Lipschitz 连续函数。 $H(m, p) = \frac{|p|^\gamma}{(c+m)^\beta} + f(m)$

(H2) H 关于变量 p 是严格凸的。

比如哈密顿量 $H(m, p)$ 的形式如下：

$$H(m, p) = \frac{|p|^\gamma}{(c+m)^\beta} + f(m)$$

注意到当 $q^* = H_p(m, Du)$ 时，

$$H_m(m, Du) = -L_m(m, q^*) \quad (3)$$

因为从 Cirant 等[13]的工作中已知解在 $W_r^{2,1}(Q)$ 空间中的存在唯一性，故存在某一个足够大的 R 使得 $q = H_p(m, p)$ 成立。基于对之前的文献的研究总结，我们针对 MFC 系统提出了以下两种策略迭代算法：策略迭代算法包括迭代更新密度分布函数、价值函数和控制函数。在算法中首先引入控制函数的一个上界 $R > 0$ ， $Q := \mathbb{T}^d \times [0, T]$ ， \mathbb{T}^d 代表环 $\mathbb{R}^d / \mathbb{Z}^d$ 。假设有界可测的向量场 $q^{(0)} : \mathbb{T}^d \times [0, T] \rightarrow \mathbb{R}^d$ 满足 $|q^{(0)}| \leq R$ ， $\|\operatorname{div} q^{(0)}\|_{L^r(Q)} \leq R$ 。

Algorithm 1: 具有不可分离哈密顿量的平均场控制系统的策略迭代算法(PII)

Data: 由初始的 $q^{(n)}$ ，求出密度分布 $m^{(n)}$ 和 $u^{(n)}$ 后，更新策略 $q^{(n+1)}$ ，再迭代。

[1] **for** $n = 0, 1, 2, \dots, N$ **do**

[2] 计算 $m^{(n)}$ ；

$$\begin{cases} \partial_t m^{(n)} - v \Delta m^{(n)} - \operatorname{div}(m^{(n)} q^{(n)}) = 0, & \text{in } Q \\ m^{(n)}(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases} \quad (4)$$

[3]计算 $u^{(n)}$;

$$\begin{cases} -\partial_t u^{(n)} - \nu \Delta u^{(n)} - \mathcal{L}(m^{(n)}, Du^{(n-1)}, q^{(n)}) \\ + q^{(n)} Du^{(n)} - m \mathcal{L}_m(m^{(n)}, Du^{(n-1)}, q^{(n)}) = 0 & \text{in } Q \\ u^{(n)}(x, T) = u_T(x) & \text{in } \mathbb{T}^d \end{cases} \quad (5)$$

$$\mathcal{L}(m, p, q) = p \cdot q - H(m, p) \quad (6)$$

[4]更新策略 $q^{(n+1)}$

$$q^{(n+1)}(x, t) = \arg \max_{|q| \leq R} \left\{ q \cdot Du^{(n)}(x, t) - L(m^{(n)}, q) \right\} \quad \text{in } Q$$

判断是否停止迭代；若否，则返回策略评估。

[5] end

我们注意到不可分离哈密顿量带来的一个核心变化是， $q = H_p(m, Du)$ 显性地依赖于 m 。对于可分离的形式 $q = H_p(Du)$ ，控制对 m 的依赖是非显性的，通过价值函数 u 对 m 体现。我们可以设计一个新的算法，在概率密度和价值函数的每次更新之间更新控制 q 。

Algorithm 2: 具有不可分离哈密顿量的平均场控制系统的策略迭代算法(PI2)

Data:由初始的 $q^{(n)}$ ，求出密度分布 $m^{(n)}$ 和 $u^{(n)}$ 后，更新策略 $q^{(n+1)}$ ，再迭代。

[1] for $n = 0, 1, 2, \dots, N$

[2]计算 $m^{(n)}$;

$$\begin{cases} \partial_t m^{(n)} - \nu \Delta m^{(n)} - \operatorname{div}(m^{(n)} q^{(n)}) = 0, & \text{in } Q \\ m^{(n)}(x, 0) = m_0(x) & \text{in } \mathbb{T}^d \end{cases} \quad (7)$$

[3]更新策略 $\tilde{q}^{(n)}$;

$$\tilde{q}^{(n)}(x, t) = \arg \max_{|q| \leq R} \left\{ \tilde{q} \cdot D\tilde{u}^{(n-1)}(x, t) - L(m^{(n)}, \tilde{q}) \right\} \quad \text{in } Q \quad (8)$$

[4]计算 $u^{(n)}$;

$$\begin{cases} -\partial_t \tilde{u}^{(n)} - \nu \Delta \tilde{u}^{(n)} + \tilde{q}^{(n)} D\tilde{u}^{(n)} - L(m^{(n)}, \tilde{q}^{(n)}) \\ - m \mathcal{L}_m(m^{(n)}, \tilde{q}^{(n)}) = 0 & \text{in } Q \\ \tilde{u}^{(n)}(x, T) = u_T(x) & \text{in } \mathbb{T}^d \end{cases} \quad (9)$$

[5]再次更新策略 $q^{(n+1)}$;

$$q^{(n+1)}(x, t) = \arg \max_{|q| \leq R} \left\{ q \cdot D\tilde{u}^{(n)}(x, t) - L(m^{(n)}, q) \right\} \quad \text{in } Q \quad (10)$$

判断是否停止迭代；若否，则返回策略评估。

[6] end

显然，当 H 是可分形式的哈密顿量时，则算法(PI1)与(PI2)是等价的。虽然出现在(9)和(10)的 L 是拉格朗日量，但定义并出现在(6)和(5)的项 \mathcal{L} 只是近似的拉格朗日量。它不是由勒让德变换得到。但是如果算法收敛，即 $Du^{(n-1)} \rightarrow Du^*$, $q^{(n)} \rightarrow q^*$, $m^{(n)} \rightarrow m^*$ 且 $q^* = H_p(m, Du)$ 则 $\mathcal{L}(m^{(n)}, Du^{(n-1)}, q^{(n)}) \rightarrow L(m^*, q^*)$ 。在(PI1)代码中，实际上不会用到 $Du^{(n-1)}$ ， \mathcal{L} 的意义只是为了在理论上，完全从 H 的性质出发。

2.3. 收敛性分析

除了上一节对 H 做出的假设(H1), (H2)外, 我们增加以下假设。

(I1) 当有 $r > d + 2$, $m_0 \in W_r^2(\mathbb{T}^d)$, $m_0 \geq \underline{m} > 0$ 和 $\int_{\mathbb{T}^d} m_0(x) dx = 1$ 时 $u_T \in W_\infty^2(\mathbb{T}^d)$

我们定义以下空间。

$$X_M^T = \left\{ (u, m) : u \in C^{1,0}(\mathcal{Q}) \cap W_r^{2,1}(\mathcal{Q}), m \in C^{1,0}(\mathcal{Q}) \mid \|u\|_{W_r^{2,1}(\mathcal{Q})} + |u|_{\mathcal{Q}}^{(1)} + |m|_{\mathcal{Q}}^{(1)} \leq M \right\} + \left\{ u \in C^{1,0}(\mathcal{Q}) \mid |u|_{\mathcal{Q}}^{(1)} \leq M \right\} \quad (11)$$

用 $\mathcal{T}(u, m) = (\hat{u}, \hat{m})$ 定义 X_M^T 上的算子 \mathcal{T} , 满足:

$$\begin{cases} \partial_t \hat{m} - \epsilon \Delta \hat{m} - H_p(m, Du) D \hat{m} - H_{pm}(m, Du)(Dm) \hat{m} - \sum_{i,j} H_{p_j p_i}(m, Du) (\partial_{x_j x_i}^2 u) \hat{m} = 0 \\ -\partial_t \hat{u} - \epsilon \Delta \hat{u} + H_p(m, Du) D \hat{u} - \mathcal{L}(\hat{m}, Du, H_p(m, Du)) - \hat{m} \mathcal{L}_m(\hat{m}, Du, H_p(m, Du)) = 0 \\ \hat{u}(x, T) = \hat{u}_T(x), \quad \hat{m}(x, 0) = \hat{m}_0(x) \end{cases} \quad (12)$$

定理 2.1. 假设之前的(H1), (H2)和(I1)仍成立, 取 K 满足

$$K \geq \max \left\{ \frac{2}{\underline{m}}, 2|m_0|_{\mathbb{T}^d}^{(1)}, 2|u(\cdot, T)|_{\mathbb{T}^d}^{(1)} \right\} \quad (13)$$

设

$$M_1 = 2(|m_0|_{\mathbb{T}^d}^{(1)} + |u_T|_{\mathbb{T}^d}^{(1)}) \quad (14)$$

则存在足够小的 \bar{T} 满足对任意的 $T \in (0, \bar{T}]$, 映射 \mathcal{T} 是空间

$$X_{M_1}^T \cap \left\{ (u, m) : |u|_{\mathcal{Q}}^{(1)} \leq K, 1/K \leq m \leq K \right\}$$

中的压缩映射算子。

定理 2.2. 假设之前的(H1), (H2)和(I1)仍成立, 则存在足够小的 \bar{T} 满足对任意的 $T \in (0, \bar{T}]$, MFC 系统(1.2)有唯一的解 $(u^*, m^*) \in W_r^{2,1}(\mathcal{Q}) \times W_r^{2,1}(\mathcal{Q})$ 。

定理 2.3. 假设之前的(H1), (H2)和(I1)仍成立, 根据定理 1.2.1 取 \bar{T} 。则存在一个 $\hat{T}_1 \leq \bar{T}$ 满足对任意的 $T \in (0, \hat{T}_1]$ 和 R 足够大时, 由策略迭代算法(PI1)生成的解序列 $(u^{(n)}, m^{(n)})$, 收敛到(1.2)的解 $(u^*, m^*) \in W_r^{2,1}(\mathcal{Q}) \times W_r^{2,1}(\mathcal{Q})$ 。

定理 2.1, 定理 2.2, 定理 2.3 分别对应 Lauriere, Song 与 Tang [13] 中的定理 3.1, 3.2 和 3.3, 我们在此省略证明, 本质区别在于关于 \hat{u} 的方程中所出现的

$$-\hat{m} \mathcal{L}_m(\hat{m}, Du, H_p(m, Du))$$

基于假设(H1), (H2)和(I1)不难得到与这一项相关的估计。由于篇幅限制, 我们在此省略算法 (PI2) 收敛性分析与收敛速度的证明细节。

2.4. 收敛速度分析

本节中的定理和推论完全类似于 Lauriere, Song 与 Tang [13] 中第 4 节关于平均场博弈系统的结果。我们给出相应的结论, 具体的证明细节我们会在后续的工作中给出。特别地, 我们用 (u^*, m^*) 和 q^* 分别表示方程组(1.1)的解和对应问题的最优控制。我们证明在 T 较小的假设前提下, 误差 $\|u^{(n+1)} - u^*\|_{W_r^{2,1}(\mathcal{Q})}$ 和 $\|m^{(n+1)} - m^*\|_{W_r^{2,1}(\mathcal{Q})}$ 具有线性收敛速度。

定理 2.4. 假设(H1), (H2)和(I1)有效, 取和定理 2.3 一样的 \hat{T}_1 和 R 。则存在一个对于所有 $T \in [0, \hat{T}_1]$ 有界的常数 C , 它仅取决于问题的数据。此时, 由策略迭代方法(PI1)生成的序列 $(u^{(n)}, m^{(n)})$ 满足下列不等式:

$$\|m^{(n+1)} - m^*\|_{W_r^{2,1}(\mathcal{Q})} \leq C \left(\|q^{(n+1)} - q^*\|_{L^r(\mathcal{Q})} + \|\operatorname{div}(q^{(n+1)} - q^*)\|_{L^r(\mathcal{Q})} \right) \quad (15)$$

和

$$\|u^{(n+1)} - u^*\|_{W_r^{2,1}(\mathcal{Q})} \leq CT^{\frac{1-d}{2-2r}} \left(\|u^{(n)} - u^*\|_{W_r^{2,1}(\mathcal{Q})} + \|m^{(n+1)} - m^*\|_{W_r^{2,1}(\mathcal{Q})} \right) \quad (16)$$

推论 2.5. 设 $M^{(n)} = m^{(n)} - m^*$, $U^{(n)} = u^{(n)} - u^*$ 对所有 $n \geq 0$ 成立。与定理 2.4 相同的假设下, 下列估计对 $n \geq 1$ 都成立:

$$\|U^{(n+1)}\|_{W_r^{2,1}(\mathcal{Q})} + \|M^{(n+1)}\|_{W_r^{2,1}(\mathcal{Q})} \leq CT^{\frac{1-d}{2-2r}} \left(\|U^{(n)}\|_{W_r^{2,1}(\mathcal{Q})} + \|M^{(n)}\|_{W_r^{2,1}(\mathcal{Q})} + \|U^{(n-1)}\|_{W_r^{2,1}(\mathcal{Q})} \right) \quad (17)$$

进而存在 $\ell > 1$, 足够大的 n_0 和足够小的 $\check{T} \in (0, \hat{T}_1]$ 满足对任意 $T \in (0, \check{T}]$, 我们都有线性收敛速度, 即,

$$\|U^{(n)}\|_{W_r^{2,1}(\mathcal{Q})} + \|M^{(n)}\|_{W_r^{2,1}(\mathcal{Q})} \leq \left(\frac{1}{\ell} \right)^{n-n_0} \left(\|U^{(n_0)}\|_{W_r^{2,1}(\mathcal{Q})} + \|M^{(n_0)}\|_{W_r^{2,1}(\mathcal{Q})} \right)$$

3. 实验

我们在 Lauriere 等人的[14]工作基础上对基于 MFG 的策略迭代算法做出了新的改进, 对 MFC 系统引入了新的策略迭代方法(PI1)和(PI2)。并对 MFC 系统使用有限差分方法, 给出具体的数值实验例子和分析。对于数值实验的图像与说明, 本文比较同一最优控制问题的 MFC 和 MFG 的策略迭代方法计算的数值结果差异。类似于 Carmona 等[11], 关注计算 MFG 与 MFC 中的 POA。据我们所知, 目前还没有人用策略迭代算法计算在 MFC 中, 具有局部依赖条件下的 POA。

3.1. 算法实现

我们先给出该平均场控制系统中对应的函数参数形式, 然后在具体数值实验中再赋予相关参数值。拉格朗日量是:

$$L(m, q) = \frac{1}{2} r(1+m)^\beta |q|^2 + f(m)$$

其中 $f(m) = \zeta m + \gamma \left(1 - e^{-10(x-0.25)^2} \right)$, 则:

$$L_m(m, q) = \frac{1}{2} r \beta (1+m)^{\beta-1} |q|^2 + \zeta$$

哈密顿量是:

$$\begin{aligned} H(m, Du) &= \sup_{q \in \mathbb{R}^d} \left\{ q \cdot Du - \frac{1}{2} r(1+m)^\beta |q|^2 - f(m) \right\} \\ &= \frac{1}{2r(1+m)^\beta} |Du|^2 - f(m) \end{aligned}$$

其中 β , ζ 为正常数。最后的常数 ζ 是人群厌恶成本, 表示智能体不愿意处于非常拥挤的区域。

在计算时 MFG 求解纳什均衡, MFC 求解帕累托最优问题, 分别用 J_{mfc} 和 J_{mfg} 表示两类不同的社会总成本:

$$\begin{cases} J_{mfg} = J(q; m) = \mathbb{E} \left[\int_0^T L(m, q) ds + g(X_T, m(T)) \right] \\ J_{mfc} = \int_0^T \int_{\mathbb{T}^d} L(m, q) m ds dx + \int_{\mathbb{T}^d} g(x, m(T)) m^q dx \end{cases} \quad (18)$$

但是在实际计算过程中, 因为我们是对 m 的局部依赖, 所以用 $g(X_T)$ 代替 $g(X_T, m(T))$ 计算。接着我们给出对于大规模多智能体系统控制问题的 POA 的数学计算表达式:

$$\text{POA} = \frac{J_{mfc}}{J_{mfg}}$$

理论上 POA 应该大于 1, 我们通过多组同参数下的数值实验证此结论。为了研究比较 MFG 与 MFC 系统中运行成本、群体密度分布和终止条件对计算 POA 的影响, 找到优化大规模多智能体系统的控制问题中的社会总成本的方法, 我们把对应函数都参数化。

我们在 \mathbb{T}^d 上固定一个网格 \mathcal{G} 。然后, 我们用 U 、 M 和 Q 表示 \mathcal{G} 上的矢量, 分别逼近 MFC 的解和控制策略。这里我们使用均匀网格和中心二阶有限差分来计算离散的拉普拉斯算子。而对于时间的离散化, 我们采用隐式欧拉方法来处理时间前向的 KFP 方程和时间后向的 HJB 方程。在具体实施过程中我们在空间上使用了 $I = 200$ 的节点数, 在时间上使用了 $N = 200$ 的节点数。把初始控制策略设定为: 在 \mathcal{G} 上对所有 $Q_n^{(0)} = (0, 0)$

我们给出下面的一维实例。让所有智能体在规定时间内向某个目标移动, 并根据在时间结束时智能体与目标位置的距离来惩罚智能体。智能体的状态范围是一维环形 $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ 。假设移动的目标位于 0.8 处。我们取初始的群体分布 m_0 为正态分布, 相应的 MFC 方程就是:

$$\begin{cases} -\partial_t u - \nu \Delta u + \frac{1}{2r(1+m)^\beta} |Du|^2 - \frac{m\beta}{2(1+m)^{\beta+1}} |Du|^2 = 2\zeta m + \gamma \left(1 - e^{-10(x-0.25)^2}\right) \\ \partial_t m - \nu \Delta m - \operatorname{div} \left(\frac{m Du}{(1+m)^\beta} \right) = 0 \\ u_T(x) = \left(-e^{-\omega^*(x-0.8)^2} + 1\right)\theta \\ m_0(x) = -e^{-10^*(x-0.2)^2} \end{cases} \quad (19)$$

本例中我们对下面参数赋值 $\beta = 0.5$, $r = 1$, $\nu = 0.05$, $\zeta = 0.1$, $\gamma = 1$ 和 $T = 1$, 由此给出对 MFC 的策略迭代算法(PI1)的结果。

在图 1 中, 通过绘制几个固定的 n , 我们分别展示了 MFC 数值模拟和 MFG 数值模拟的解密度分布随时间的演变。我们可以看到, 解的密度分布都向同一个目标移动, 而 MFC 方法更集中于目标附近。然而, 由于拥堵成本和人群厌恶成本, 两种方法都不能完全集中在目标上。在图 2 中, 我们展示了与迭代次数为 30 的收敛结果: 离散 MFC 和 MFG 系统的残差。这里我们使用了(PI1)算法获得的基准解。在这里我们观察到, 通过我们的两种策略迭代方法得到的数值解是一致的, 也侧面验证了之前的理论: 除了前几次迭代和由于离散系统的近似精度限制而达到下限后以外, 该收敛速度是线性的。

按照(18)计算本例中最优控制问题的社会总成本, 即: J_{mfc} 和 J_{mfg} 。然后不管我们如何改变参数我们得出 $\text{POA} = \frac{J_{mfc}}{J_{mfg}}$ 的值总是大于 1。这也从数值解的角度验证了两类不同的社会总成本的优劣。

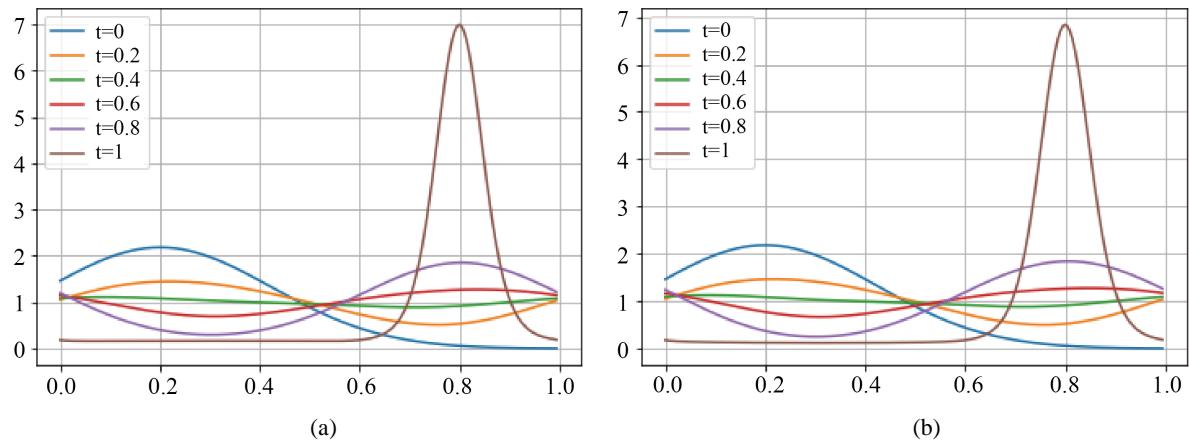


Figure 1. Numerical solutions of Algorithm (PI1) of (19) at different times. (a) MFC; (b) MFG
图 1. 算法(PI1)的(19)在不同时间下的数值解。 (a) MFC; (b) MFG

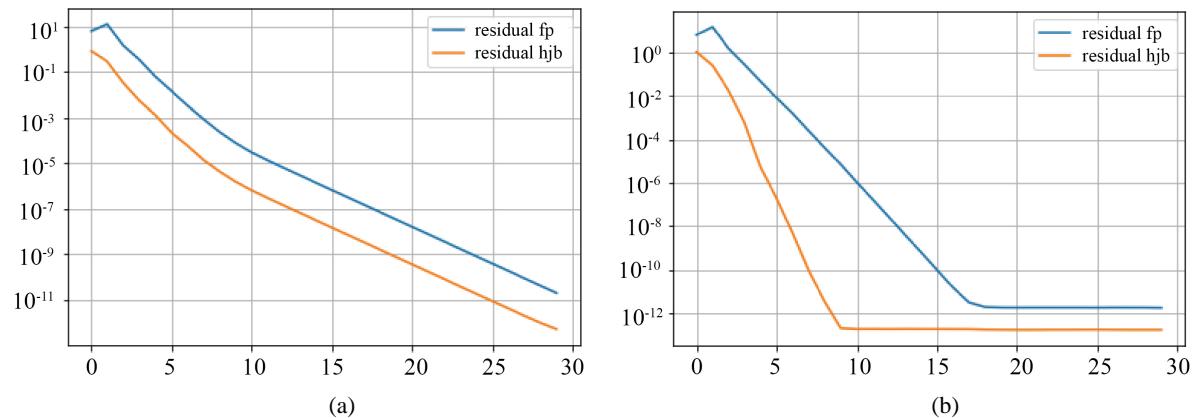


Figure 2. Convergence of Algorithm (PI1) of (19). (a) Residual of MFC; (b) Residual of MFG
图 2. 算法(PI1)的(19)的收敛性。 (a) MFC 的残差; (b) MFG 的残差

3.2. 计算 POA

Carmona 等[11]在对群体密度分布非局部依赖条件下, 分析参数变化对 POA 的影响。而我们在对群体密度分布局部依赖条件下, 使在时间 T 足够小时, 我们通过有限差分法求出对应算法的数值解就可以计算相对准确的 POA。在图 3~9 中, 我们在不同参数设定下计算出不同的 J_{mfc} 和 J_{mfg} 。然后我们可以计算 J_{mfg} 与 J_{mfc} 的比率, 它总是大于 1。为了研究的严谨性, 下面所有的例子都是基于一维实例(19), 除非另有说明, 我们让参数保持在下面的默认值, 对每类参数的影响都画出 2 组不同的 POA 图像。

$$\begin{aligned} \beta &= 0.5, \quad \zeta = 0.1, \quad \nu = 0.08, \quad r = 1, \\ \gamma &= 1, \quad T = 1, \quad \omega = 10, \quad \theta = 1 \end{aligned}$$

基于类似的不同参数情况下的对照实验, 在图 3 中当 β 变大时, POA 变小。在图 4 中当 ζ 变大时, POA 变大。我们发现在图 5 中当 r 变大时, POA 变小。

在图 6 中当 γ 变大时, POA 变小。在图 7 中当 T 变大时, POA 变大。在图 8 中当 θ 变大时, POA 变小。在图 9 中当 ν 变大时, POA 变小。我们重复不同数值下验证结果, 得到类似的现象。这与 Carmona 等[11]对 POA 影响的结果可以相互验证, 我们在 T 足够小时, 对局部依赖的条件, 也分析了两类不同参数对 POA 的影响都是单调的但是方向不同。

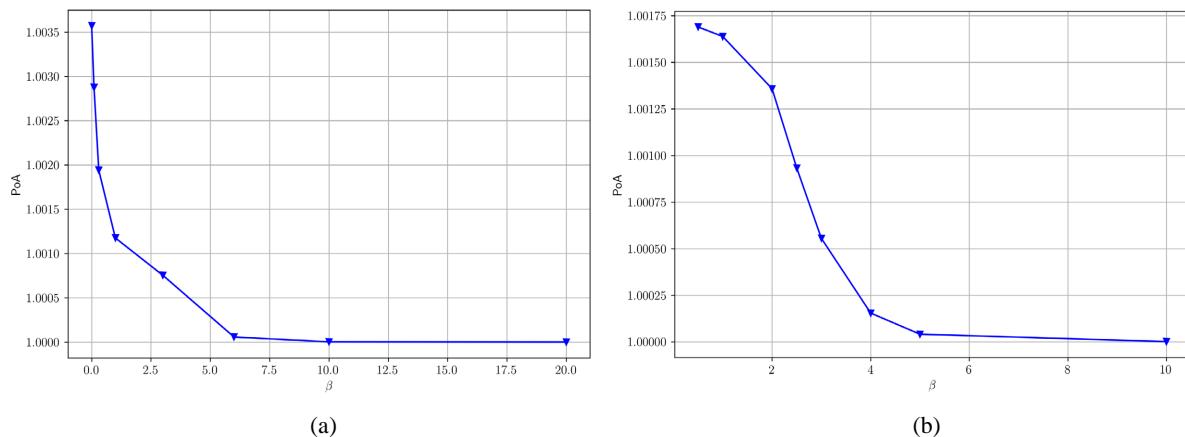


Figure 3. The change of POA with β after changing some parameters. (a) $\gamma = 2.2$; $\theta = 0.01$; (b) $\zeta = 0.2$

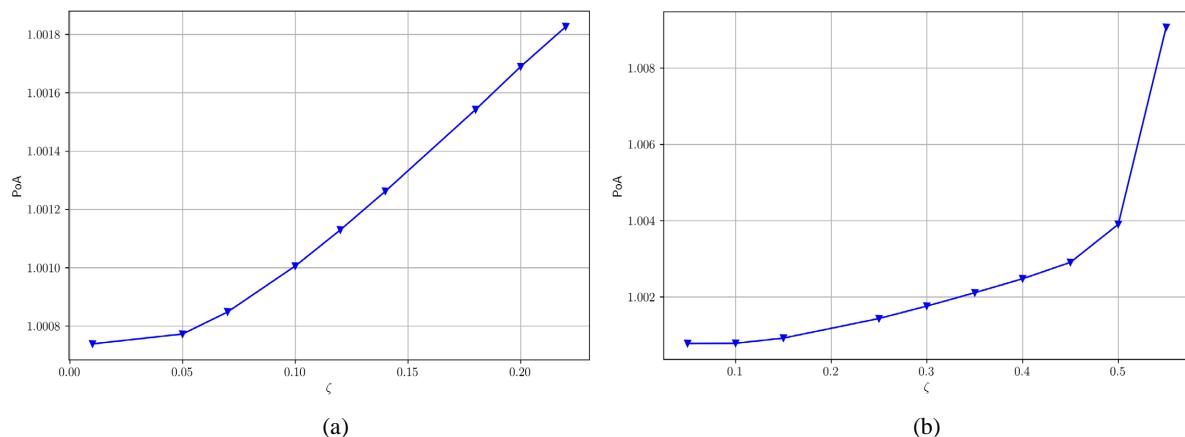


Figure 4. The change of POA with ζ after changing some parameters. (a) Default Parameters; (b) $v = 0.07$, $T = 0.3$

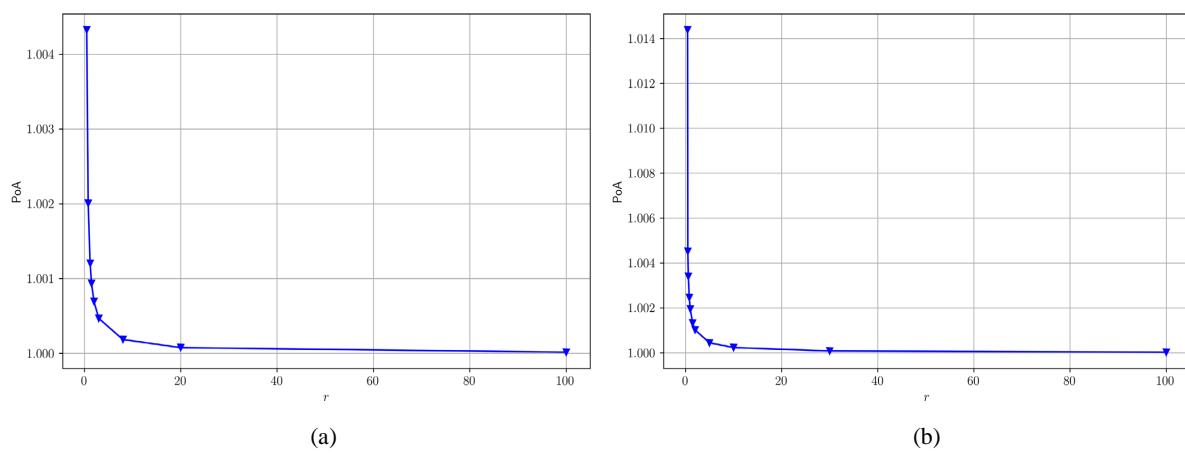


Figure 5. The change of POA with r after changing some parameters. (a) $\zeta = 0.5, \beta = 2, \gamma = 1.8, T = 0.5, \theta = 0.01$; (b) $\zeta = 0.6, \beta = 1.8, \gamma = 1.8, T = 0.4, \theta = 0.01$

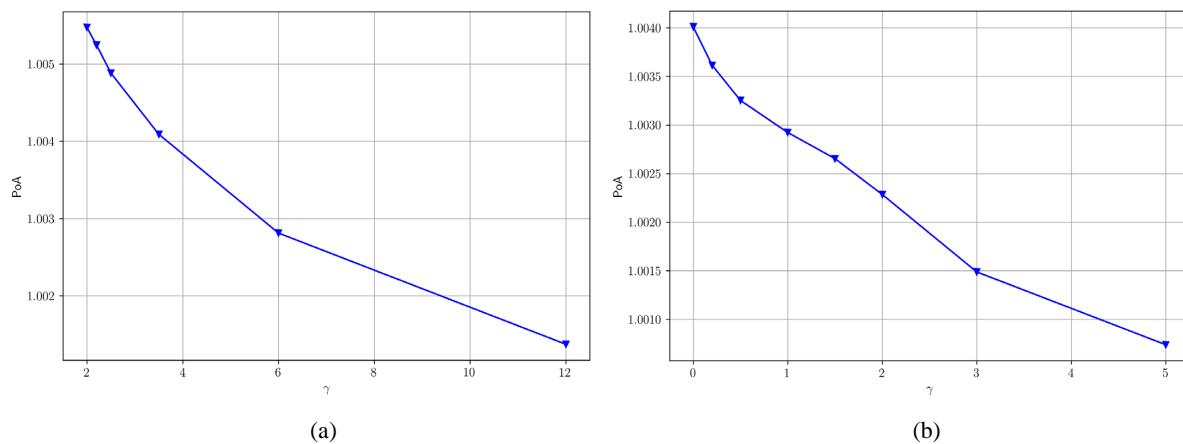


Figure 6. The change of POA with γ after changing some parameters. (a) $\zeta = 0.2, \beta = 0.8, v = 0.06, T = 0.8, \theta = 0.01$; (b) $v = 0.07, \omega = 100$

图6. 改变部分参数后, POA 随 γ 的变化。(a) $\zeta = 0.2$, $\beta = 0.8$, $v = 0.06$, $T = 0.8$, $\theta = 0.01$; (b) $v = 0.07$, $\omega = 100$

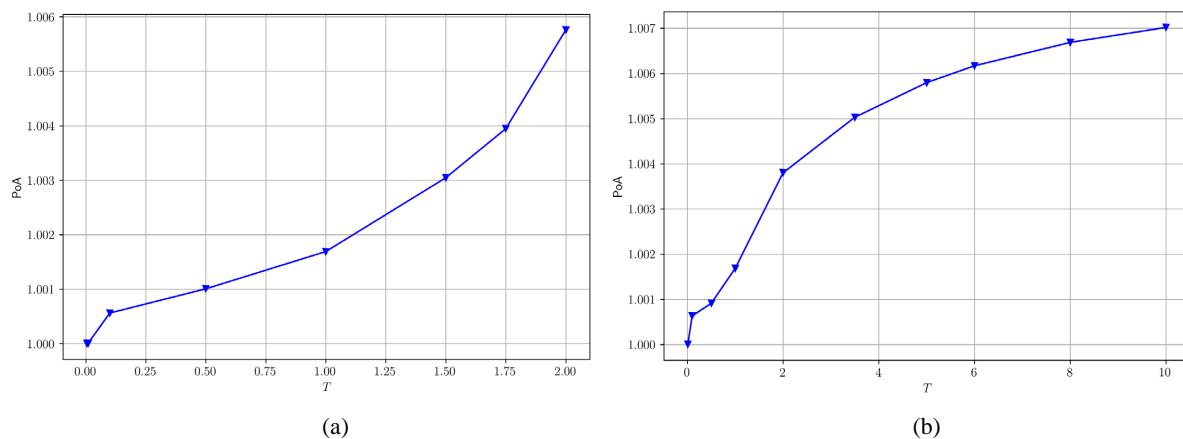


Figure 7. The change of POA with T after changing some parameters. (a) $\zeta = 0.2$; (b) $v = 0.06$

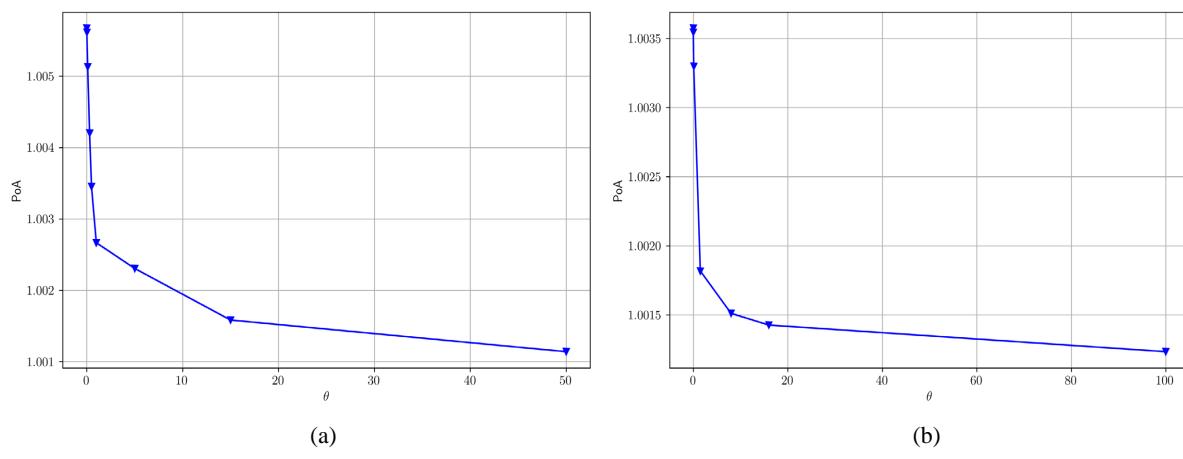


Figure 8. The change of POA with θ after changing some parameters. (a) $\beta = 0.05$, $v = 0.06$, $\gamma = 1.8$; (b) $\beta = 0.1$, $v = 0.07$, $\gamma = 2.2$

图 8. 改变部分参数后, POA 随 θ 的变化。(a) $\beta = 0.05$, $v = 0.06$, $\gamma = 1.8$; (b) $\beta = 0.1$, $v = 0.07$, $\gamma = 2.2$

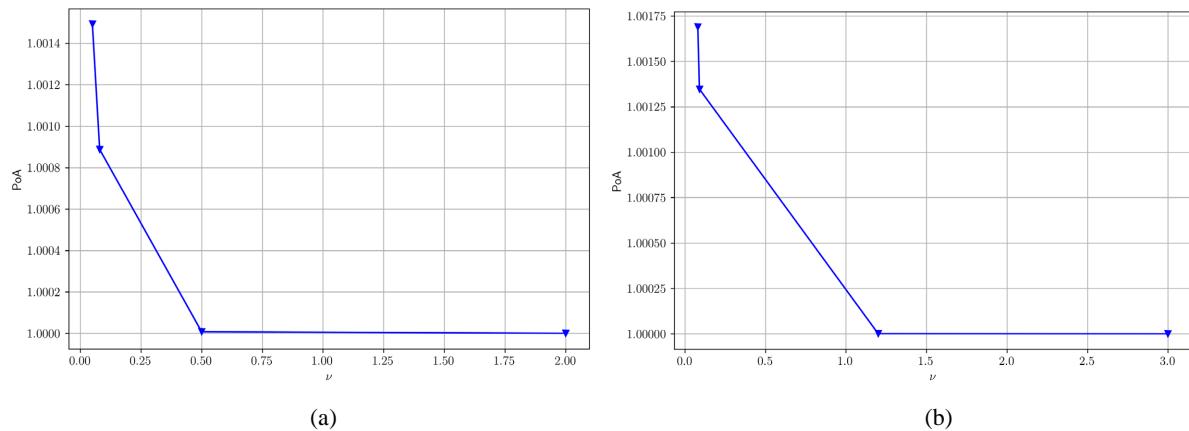


Figure 9. The change of POA with ν after changing some parameters. (a) $\beta = 1$, $T = 0.5$, $\gamma = 5$; (b) $\zeta = 0.2$

图9. 改变部分参数后, POA 随 ν 的变化。 (a) $\beta = 1$, $T = 0.5$, $\gamma = 5$; (b) $\zeta = 0.2$

4. 总结

本文研究的内容是一类来自于平均场控制问题的倒向 - 正向偏微分方程组有限差分数值计算方法，使用的算法基于最优控制问题中的策略迭代方法。其中，平均场偶合项是局部的，因此给解的正则性讨论和算法收敛性研究带来了新的困难。本文通过一些例子验证了算法的有效性。特别地，本文通过数值试验讨论了平均场博弈纳什均衡和平均场控制帕累托最优之间的总体效益比较。时间区域 T 较小的假设和相关的证明技巧导致本文提出的算法难以被从理论和实践上被用于更一般的情形，即使方程组可通过其他方法(Schauder 不动点原理等)被证明解的存在性。同时，这种方法难以适用于方程组具有非唯一解的情况。这些问题有待进一步研究。

参考文献

- [1] Lasry, J.-M. and Lions, P.-L. (2007) Mean Field Games. *Japanese Journal of Mathematics*, **2**, 229-260. <https://doi.org/10.1007/s11537-007-0657-8>
 - [2] Huang, M., Malhamé, R.P. and Caines, P.E. (2006) Large Population Stochastic Dynamic Games: Closed-Loop McKean-Vlasov Systems and the Nash Certainty Equivalence Principle. *Communications in Information and Systems*, **6**, 221-252. <https://doi.org/10.4310/CIS.2006.v6.n3.a5>
 - [3] Achdou, Y. and Lauriere, M. (2015) On the System of Partial Differential Equations Arising in Mean Field Type Control. arXiv:1503.05044. <https://doi.org/10.3934/dcds.2015.35.3879>
 - [4] Carmona, R. and Delarue, F. (2013) Probabilistic Analysis of Mean-Field Games. *SIAM Journal on Control and Optimization*, **51**, 2705-2734. <https://doi.org/10.1137/120883499>
 - [5] Bensoussan, A., Frehse, J., Yam, P., et al. (2013) Mean Field Games and Mean Field Type Control Theory. Vol. 101, Springer. <https://doi.org/10.1007/978-1-4614-8508-7>
 - [6] Lauriere, M. (2021) Numerical Methods for Mean Field Games and Mean Field Type Control. *Mean Field Games*, **78**, 221. <https://doi.org/10.1090/psapm/078/06>
 - [7] Achdou, Y., Cardaliaguet, P., Delarue, F., et al. (2019) Mean Field Games and Applications: Numerical Aspects. 249-307. <https://doi.org/10.1007/978-3-030-59837-2>
 - [8] Cacace, S., Camilli, F. and Goffi, A. (2021) A Policy Iteration Method for Mean Field Games. *ESAIM: Control, Optimisation and Calculus of Variations*, **27**, 85. <https://doi.org/10.1051/cocv/2021081>
 - [9] Camilli, F. and Tang, Q. (2022) Rates of Convergence for the Policy Iteration Method for Mean Field Games Systems. *Journal of Mathematical Analysis and Applications*, **512**, 126138. <https://doi.org/10.1016/j.jmaa.2022.126138>
 - [10] Cardaliaguet, P. and Rainer, C. (2019) On the (in)Efficiency of MFG Equilibria. *SIAM Journal on Control and Optimization*, **57**, 2292-2314. <https://doi.org/10.1137/18M1172363>
 - [11] Carmona, R., Graves, C.V. and Tan, Z. (2019) Price of Anarchy for Mean Field Games. *ESAIM: Proceedings and Surveys*,

- 65**, 349-383. <https://doi.org/10.1051/proc/201965349>
- [12] Graber, P.J. (2015) Weak Solutions for Mean Field Games with Congestion. *Analysis of PDEs*.
- [13] Cirant, M., Gianni, R. and Mannucci, P. (2020) Short-Time Existence for a General Backward-Forward Parabolic System Arising from Mean-Field Games. *Dynamic Games and Applications*, **10**, 100-119.
<https://doi.org/10.1007/s13235-019-00311-5>
- [14] Laurière, M., Song, J. and Tang, Q. (2023) Policy Iteration Method for Time-Dependent Mean Field Games Systems with Non-Separable Hamiltonians. *Applied Mathematics & Optimization*, **87**, 1-34.
<https://doi.org/10.1007/s00245-022-09925-5>