Deep Learning for Population-Dependent Controls in Mean Field Control Problems with Common Noise

Extended Abstract

Gökçe Dayanıklı
Univ. of Illinois Urbana-Champaign
Urbana, IL, USA

Mathieu Laurière
Shanghai New York Univ.
Shanghai, PRC

Jiacheng Zhang
Univ. of California Berkeley
Berkeley, CA, USA

ABSTRACT

In this paper, we propose several approaches to learn the optimal population-dependent controls in order to solve mean field control problems (MFC). Such policies enable us to solve MFC problems with forms of common noises at a level of generality that was not covered by existing methods. We analyze rigorously the theoretical convergence of the proposed approximation algorithms. Of particular interest for its simplicity of implementation is the N-particle approximation. The effectiveness and the flexibility of our algorithms is supported by numerical experiments comparing several combinations of distribution approximation techniques and neural network architectures. We use three different benchmark problems from the literature: a systemic risk model, a price impact model, and a crowd motion model. We first show that our proposed algorithms converge to the correct solution in an explicitly solvable MFC problem. Then, we show that population-dependent controls outperform state-dependent controls. Along the way, we show that specific neural network architectures can improve the learning further.

KEYWORDS

mean field control, deep learning, stochastic optimal control

ACM Reference Format:
Gökçe Dayanıklı, Mathieu Laurière, and Jiacheng Zhang. 2024. Deep Learning for Population-Dependent Controls in Mean Field Control Problems with Common Noise: Extended Abstract. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

This article is an extended abstract of [11].

1 INTRODUCTION

Optimal control problems have found a wide range applications from engineering to finance and robotics. In most cases, the system is subject to random disturbances which means that one has to find optimal controls in a stochastic setting. In the present work, the motivation is to study very large populations of strategic identical agents who cooperate to minimize a social cost both under idiosyncratic and common noise. In order to approximate the problem, we will use mean-field approximation which consists in replacing individual interactions by the interaction of a representative agent with the distribution of the population. This leads to more tractable models and more efficient algorithms. This setting is often referred to as mean field control (MFC for short).

The numerical schemes to solve MFC problems require to solve forward backward partial differential equations (PDEs) (either with finite scheme methods [1, 2, 6] or deep learning [3, 8]) or forward backward stochastic differential equations (SDEs) (with discretization [4, 10] or deep learning [5, 9, 12]). Here, we focus on an approach that learns the optimal control without using backward PDEs or SDEs. One of the main advantages is the fact that it does not require any dynamic programming principle. This approach was used for standard stochastic optimal control problems in [13, 14] and in the mean field setup in [9, 12] with controls that were independent of the population. In this work, we extend the approach to population-aware controls, which allows us to tackle MFC with common noise. Indeed, when there is only idiosyncratic noise, randomness vanishes in the mean-field limit and it is sufficient to learn controls that are functions of only the representative agent's state and time. However, when there is common noise, the evolution of the distribution cannot be predicted with certainty and, to be optimal, the control should be a function of the distribution.

2 MFC WITH COMMON NOISE

Our aim is to minimize over closed-loop controls v the total expected cost:

$$J(v) := \mathbb{E}\left[\int_0^T f\big(t, X_t^v, A_t^v, \mu_t^v\big)\mathrm{d}t + g\big(X_T^v, \mu_T^v\big)\right],$$

where X_t^v is the state of the representative player, $\mu_t^v = \mathrm{Law}(X_t^v|\mathcal{F}_t^0)$ is the conditional state distribution of the population given a common noise W^0, $A_t^v = v(t, X_t^v, \mu_t^v)$ is the representative player's action, and for $t \geq 0$:

$$dX_t^v = b(t, X_t^v, A_t^v, \mu_t^v)dt + \sigma dW_t + \sigma_0 dW_t^0, \quad X_0^v \sim \mu_0. \quad (1)$$

On the theoretical side, we make **three main contributions**, summarized as follows. **1.** Firstly we prove a bound on the difference between the optimal value of the MFC problem in which an approximated distribution is used and the optimal value of the original MFC problem. This error bound is a function of the Wasserstein distance between the original distribution and the approximated distribution. Therefore, our theorem shows that if the approximate distribution is close enough to the original distribution, the optimal values of the two MFC problems (with the common noise) will be also close. This relies on the stability of the problem with respect to distribution approximation. **2.** As a corollary, we show that the

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand. © 2024 International Foundation for Autonomous Agents and Multiagent Systems (www.ifaamas.org).

optimal value of the MFC problem where an empirical distribution is implemented instead of the theoretical distribution will be close to the original MFC problem. **3.** Finally, since representing a distribution can be challenging, we analyze the stability of our problem with respect to a restricted admissible control set. The new restricted admissible control set allows controls that takes the embedding of the distribution as an input. Our motivation here is to represent the distribution in some natural ways, such as using empirical approximation, moment approximation, and histogram approximation. This last theoretical result shows that the optimal value of the MFC problem where the optimal control is searched in the restricted admissible set is close to the optimal value of the original MFC problem.

Next, we propose numerical methods to learn the optimal control with distribution approximation. We stress that the control's input is potentially very high dimensional because it contains an approximation of the population distribution. This motivates us to use deep learning methods.

3 NUMERICAL METHOD & EXPERIMENTS

Our numerical method uses deep learning for learning the control as a function of time, state, and distribution and Monte Carlo simulations for the time-discretized state dynamics of the particles in the population. In order to approximate the distribution, we implement a two step approach. First we summarize the distribution by using either empirical, histogram, or moments approximation. Then, we use this as an input for a parameterized distribution embedding function which is implemented as a neural network, whose output will be passed to a second neural network for the control. For the first neural network in charge of the distribution embedding, we consider **three different architectures**: feedforward (**FFNN**), convolutional (**CNN**), or symmetric (**SYM**) neural networks. After approximating the distribution with the distribution embedding step, we use it as an input for a second neural network that aims to learn the control as a function of time, state, and (approximated) distribution. In the implementation, we use Monte Carlo simulations for an interacting system of N particles based on the state dynamics in (1). We use a loss function the time-discretized and population-averaged counterpart to the total cost in (1). We illustrate our results on **three classical examples from the literature**.

In the **price impact** experiment, we extend [7] to model many traders who control the trading rates of two stocks. The representative player's state is the inventory for both stocks and they minimize their trading costs while interacting with each other through the average trading rates in the population which creates a price impact. In the **crowd motion** experiment, agents choose their velocity and their state is their position on the Euclidean axis. They interact with each other through the congestion cost, i.e., it will be more costly to be at more crowded places. Our method is used to solve both of these high dimensional problems for the distribution embedding and we showed that they outperform the case where the distribution is not used as an input to the control (*nodist*). This shows that population-dependent controls outperform the state-dependent control. In the more complex crowd motion example, we see that more sophisticated architectures such as CNN with histogram approximation and SYM with empirical approximation outperform FFNN with histogram and empirical approximations.

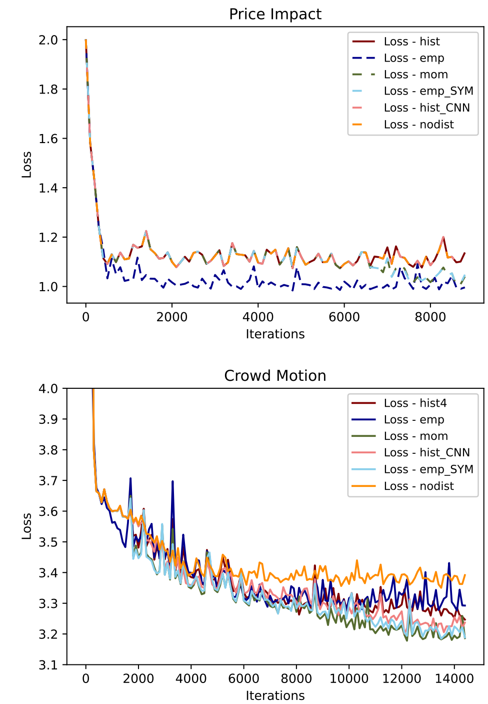

Figure 1: Loss comparison of different distribution approximations in price impact (top) and crowd motion (bottom) experiments.

4 CONTRIBUTIONS AND CONCLUSIONS

The main contribution of this article is three-fold. First, we prove theoretical approximation guarantees for the population-dependent controls in MFC with common noise. Second, we present an algorithm which trains a neural network control to minimize the social cost, and we propose several variants of distribution approximation (empirical, moments, histogram) and neural network architectures (feedforward fully connected, convolutional, symmetric). Third, we illustrate the performance of various combinations of distribution approximation and neural network architectures on three examples from the literature. We show that, in the presence of common noise, population-dependent controls outperform population-independent controls, and that the choice of approximation and architecture helps to improve the learning.

ACKNOWLEDGMENTS

Gökçe Dayanıklı is affiliated with the Department of Statistics at UIUC. Mathieu Lauriere is affiliated with the Shanghai Frontiers Science Center of Artificial Intelligence and Deep Learning and the NYU-ECNU Institute of Mathematical Sciences, NYU Shanghai, 567 West Yangsi Road, Shanghai.

REFERENCES

[1] Yves Achdou and Italo Capuzzo-Dolcetta. 2010. Mean field games: numerical methods. *SIAM J. Numer. Anal.* 48, 3 (2010), 1136–1162.

[2] Yves Achdou and Mathieu Laurière. 2015. On the system of partial differential equations arising in mean field type control. *Discrete and Continuous Dynamical Systems* 35, 9 (2015), 3879–3900.

[3] Ali Al-Aradi, Adolfo Correia, Danilo Naiff, Gabriel Jardim, and Yuri Saporito. 2018. Solving nonlinear and high-dimensional partial differential equations via deep learning. *arXiv preprint arXiv:1811.08782* (2018).

[4] Andrea Angiuli, Christy V Graves, Houzhi Li, Jean-François Chassagneux, François Delarue, and René Carmona. 2019. Cemracs 2017: numerical probabilistic approach to MFG. *ESAIM: Proceedings and Surveys* 65 (2019), 84–113.

[5] Alexander Aurell, Rene Carmona, Gökçe Dayanıklı, and Mathieu Laurière. 2022. Optimal incentives to mitigate epidemics: a Stackelberg mean field game approach. *SIAM Journal on Control and Optimization* 60, 2 (2022), S294–S322.

[6] Luis M Briceno-Arias, Dante Kalise, and Francisco J Silva. 2018. Proximal methods for stationary mean field games with local couplings. *SIAM Journal on Control and Optimization* 56, 2 (2018), 801–836.

[7] René Carmona and François Delarue. 2018. *Probabilistic Theory of Mean Field Games with Applications I: Mean Field FBSDEs, Control, and Games*. Springer International Publishing.

[8] René Carmona and Mathieu Laurière. 2021. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games I: the ergodic case. *SIAM J. Numer. Anal.* 59, 3 (2021), 1455–1485.

[9] René Carmona and Mathieu Laurière. 2022. Convergence analysis of machine learning algorithms for the numerical solution of mean field control and games: II—the finite horizon case. *The Annals of Applied Probability* 32, 6 (2022), 4065–4105.

[10] Jean-François Chassagneux, Dan Crisan, and François Delarue. 2019. Numerical method for FBSDEs of McKean–Vlasov type. *The Annals of Applied Probability* 29, 3 (2019), 1640–1684.

[11] Gökçe Dayanıklı, Mathieu Laurière, and Jiacheng Zhang. 2023. Deep Learning for Population-Dependent Controls in Mean Field Control Problems with Common Noise. arXiv:2306.04788 [math.OC]

[12] Jean-Pierre Fouque and Zhaoyu Zhang. 2020. Deep learning methods for mean field control problems with delay. *Frontiers in Applied Mathematics and Statistics* 6 (2020), 11.

[13] Emmanuel Gobet and Rémi Munos. 2005. Sensitivity Analysis Using Itô–Malliavin Calculus and Martingales, and Application to Stochastic Optimal Control. *SIAM Journal on control and optimization* 43, 5 (2005), 1676–1713.

[14] Jiequn Han and Weinan E. 2016. Deep Learning Approximation for Stochastic Control Problems. *Deep Reinforcement Learning Workshop, NIPS, arXiv preprint arXiv:1611.07422* (2016).