Journal of Machine Learning **ISSN**: 2790-2048(e), 2790-203X(p)

Mean-Field Neural Networks-Based Algorithms for McKean-Vlasov Control Problems

Huyên Pham * [1] and Xavier Warin † [2]

[1]Laboratoire de Probabilités, Statistique et Modélisation, Université Paris Cité, Paris 75205, France
[2]Electricité de France, Research and Development, Palaiseau 91120, France

Abstract. This paper is devoted to the numerical resolution of McKean-Vlasov control problems via the class of mean-field neural networks introduced in our companion paper [Pham and Warin, Neural Netw., 168, 2023] in order to learn the solution on the Wasserstein space. We propose several algorithms either based on dynamic programming with control learning by policy or value iteration, or backward stochastic differential equation SDE from stochastic maximum principle with global or local loss functions. Extensive numerical results on different examples are presented to illustrate the accuracy of each of our eight algorithms. We discuss and compare the pros and cons of all the tested methods.

Keywords:
McKean-Vlasov control,
Mean-field neural networks,
Learning on Wasserstein space,
Dynamic programming,
Backward SDE.

Article Info.:
Volume: 3
Number: 2
Pages: 176 - 214
Date: /2024
doi.org/10.4208/jml.230106

Article History:
Received: 06/01/2023
Accepted: 17/03/2024

Communicated by:
Arnulf Jentzen

1 Introduction

This paper is concerned with the numerical resolution of McKean-Vlasov (MKV) control, also called mean-field control (MFC) problems over finite horizon. The dynamics of the controlled state process $X = (X_t)_t$ valued in $\mathbb{R}^d$ is driven by the mean-field SDE (stochastic differential equation)

$$\mathrm{d}X_t = b(X_t, \mathbb{P}_{X_t}, \alpha_t)\mathrm{d}t + \sigma(X_t, \mathbb{P}_{X_t}, \alpha_t)\mathrm{d}W_t, \quad 0 \le t \le T, \quad X_0 \sim \mu_0,$$

where W is a d-dimensional Brownian motion on a filtered probability space $(\Omega, \mathcal{F}, \mathbb{F} = (\mathcal{F}_t)_t, \mathbb{P})$, the initial distribution μ_0 of X_0 lies in $\mathcal{P}_2(\mathbb{R}^d)$, the Wasserstein space of square-integrable probability measures, $\alpha \in \mathcal{A}$ is a control process, i.e an $\mathbb{F}$-progressively measurable process valued in $A \subset \mathbb{R}^m$, and $\mathbb{P}_{X_t}$ denotes the law of X_t, valued on $\mathcal{P}_2(\mathbb{R}^d)$, under standard assumptions on the coefficients b, σ defined on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, and valued respectively in $\mathbb{R}^d$ and $\mathbb{R}^{d\times d}$. The objective is to minimize over controls $\alpha \in \mathcal{A}$, a cost functional of the form

$$J(\alpha) = \mathbb{E}\left[\int_0^T f(X_t, \mathbb{P}_{X_t}, \alpha_t)\mathrm{d}t + g(X_T, \mathbb{P}_{X_T})\right] \ \rightarrow \ v(\mu_0) = \inf_{\alpha\in\mathcal{A}} J(\alpha), \tag{1.1}$$

*Corresponding author. pham@lpsm.paris
†xavier.warin@edf.fr

where f is a running cost function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d) \times A$, and g is a terminal cost function on $\mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$.

The theory and applications of mean-field control problems that study models of large population of interacting agents controlled by a social planner, have generated a vast literature in the last decade, and we refer to the monographs [4, 6, 7] for a comprehensive treatment of this topic. As analytical solutions to MFC are rarely available, it is crucial to design efficient numerical schemes for solving such problem, and the main challenging issue is the infinite dimensional feature of MFC coming from the distribution law state variable.

Following the tremendous impact of machine learning methods for solving high-dimensional partial differential equations (PDEs) and control problems, see e.g. the survey papers [3, 16], and the link to the website `deeppde.org`, some recent works have proposed deep learning schemes for MFC, based on neural network approximations of the feedback control and/or the value function solution to the Hamilton-Jacobi-Bellman equation or backward stochastic differential equations (BSDEs). In these articles, the authors consider either approximate feedback controls by standard feedforward neural networks with input the time and the state variable X_t in $\mathbb{R}^d$ by viewing the law of X_t as a deterministic function of time (see [9, 12, 14, 24, 26, 27]), or consider a particle approximation of the MFC for reducing the problem to a finite-dimensional problem that is numerically solved by means of symmetric neural networks, see [13]. However, the outputs obtained by these deep learning schemes only provide an approximation of the solution for a given initial distribution of the state process. Hence, for another distribution μ_0 of the initial state, these algorithms have to be run again.

In this paper, we aim to compute the minimal cost function $v(\mu_0)$ for any $\mu_0 \in \mathcal{P}_2(\mathbb{R}^d)$, and to find the optimal control, which can be searched w.l.o.g. in the class of feedback controls, i.e. of the form $\alpha_t = \mathfrak{a}(t, X_t, \mathbb{P}_{X_t}), 0 \leq t \leq T$, for some measurable function $\mathfrak{a}$ on $[0, T] \times \mathbb{R}^d \times \mathcal{P}_2(\mathbb{R}^d)$. In other words, our goal is to learn the value function and the optimal feedback control on the Wasserstein space. We shall rely on a new class of neural networks, introduced in our companion paper [25], called mean-field neural networks with input a probability measure in order to approximate mappings on the Wasserstein space. We then develop several numerical schemes based either on dynamic programming (DP) or stochastic maximum principle (SMP). We first propose, in the spirit of [17, 18] a global learning of the feedback control approximated by a mean-field neural network. In the DP approach, we then propose two algorithms inspired by [20]: The first one learns the control by policy iteration while the second one learns sequentially the control and the value function by value iteration. In the SMP approach, we exploit the backward SDE characterization of the solution, and propose five different algorithms in line with recent methods developed in the context of standard BSDE (see [11, 15, 21]) that we extend to MKV BSDE with various choices of global or local loss functions to be minimized in the training of mean-field neural networks. We then provide extensive numerical experiments on three examples: a mean-field systemic risk model, a min/max linear quadratic model, and the classical mean-variance problem. We compare and discuss the advantages and drawbacks of all our algorithms.