Game Model Learning for Mean Field Games

Extended Abstract

Yongzhao Wang
University of Michigan
Ann Arbor, USA
wangyzh@umich.edu

Michael P. Wellman
University of Michigan
Ann Arbor, USA
wellman@umich.edu

ABSTRACT

We present an approach to learning models for mean field games from simulation data with a *coarse coding* scheme that abstracts away the time-dependent complexity and dramatically simplifies the input representation.

KEYWORDS

Mean Field Games; Game Model Learning

ACM Reference Format:
Yongzhao Wang and Michael P. Wellman. 2023. Game Model Learning for Mean Field Games: Extended Abstract. In *Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), London, United Kingdom, May 29 – June 2, 2023*, IFAAMAS, 3 pages.

1 INTRODUCTION

Mean field games (MFGs) [2, 3] describe systems with a conceptually infinite number of interacting strategic players. Players within a population are treated atomistically and identically, and the state of the system can be captured by a probability distribution over player states (the "mean field"). The strategic interaction in MFGs can be characterized by the the optimal behavior of a single representative player against the full population, as represented by the mean field. Under certain general assumptions, it can be proved that a class of MFGs is the limit of N-player games as N approaches infinity [1] and the solution of MFGs approximates the solution of the corresponding finite game. Therefore, an MFG modeling enables game-theoretic analysis for games with a large but finite number of players that would be intractable with a standard modeling.

In this work, we propose a game model learning approach for MFGs, which is essentially a form of regression that learns a utility function over a restricted set of strategies and distributions derived by these strategies. We study a general setup of MFGs where strategies and distributions are both time-dependent (i.e., non-stationary), and hence encoding them explicitly as inputs to a learner entails impractically high dimensionality. To handle the time-dependency, we propose a *coarse coding* scheme and learn a game model (i.e., the utility function) that takes as inputs sufficiently-statistical representations of strategies and distributions, and outputs a utility value.

To obtain samples in high-dimensional spaces of strategies and mean fields, we propose a combination of two sampling schemes: grid sampling and sampling from the Dirichlet distribution with varying concentration parameters. By combining coarse coding with our data sampling methods, we demonstrate that our approach can successfully achieve effective generalization and accurate predictions on utilities. We also show that the learned game model can support game-theoretic analysis, that is, both fictitious play (FP) [4] and replicator dynamics (RD) [5] empirically converge to NE with the model.

Proc. of the 22nd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2023), A. Ricci, W. Yeoh, N. Agmon, B. An (eds.), May 29 – June 2, 2023, London, United Kingdom.

2 METHODS

2.1 Coarse Coding

We handle the time-dependency by learning a black-box version of the true utility function $\tilde{u}$. Mathematically, consider a restricted strategy set $\Lambda \subseteq S$. Let $I : \Lambda \to \mathbb{Z}_{+}$ be a function that index each strategy $s \in \Lambda$ with a positive integer. Let σ be the mixed strategy that induces the distributions μ^{σ}. Since the distribution induction function Φ is deterministic, it is sufficient for σ to determine μ^{σ} given a fixed initial distribution $\mu_0 \in \Delta(X)$. Instead of learning $\tilde{u}(s, \mu^{\sigma})$ with time-dependent inputs, we learn a black-box utility function $\hat{u} : I(\Lambda) \times \Delta(\Lambda) \to \mathbb{R}$ as a game model using sufficient representations $I(s)$ and σ of s and μ^{σ}. We refer to this representation as *coarse coding*.

Our object is to predict the true utility $\tilde{u}(s, \mu^{\sigma})$ by $\hat{u}(I(s), \sigma)$ and thus minimizing the mean square loss $E[(\tilde{u}(s, \mu^{\sigma}) - \hat{u}(I(s), \sigma))^2]$. Our regression is based on neural networks.

2.2 Data Sampling

For regression in our case, a data point constitutes an index of a pure strategy $I(s)$, a mixed strategy σ, and a true utility $\tilde{u}(s, \mu^{\sigma})$. To collect these data points, the basic requirement is that the sampled mixed strategies σ's should uniformly distribute in the restricted strategy space so as to endow the learner with the ability of generalization across the space of induced distributions. For a large MFG, a game model typically contains dozens of strategies, which makes the sample space high-dimensional. To handle this issue, we combine grid sampling and sampling from symmetric Dirichlet distributions.

3 EXPERIMENTAL RESULTS

3.1 Approximating NE with a Game Model

In Figure 1 and Figure 2, we plot the regret curves of FP and RD with the true utility function and the game model respectively in three MFGs. In all cases of FP, we observe that the regret curve generated with our game model can quickly coincide with the one using the true utility function, and both successfully converge to 0 (i.e., reaching a NE). For RD, we apply re-sampling near the equilibrium point (indicated by the red vertical line) and observe the convergence to NE as well.

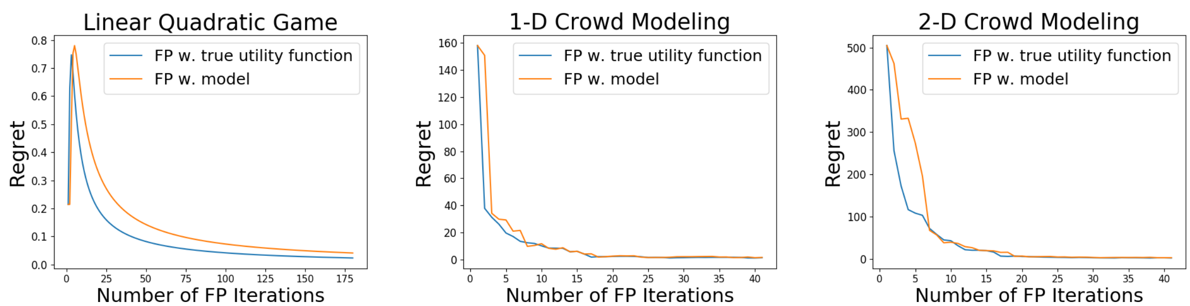

Figure 1 Regret curves with FP.

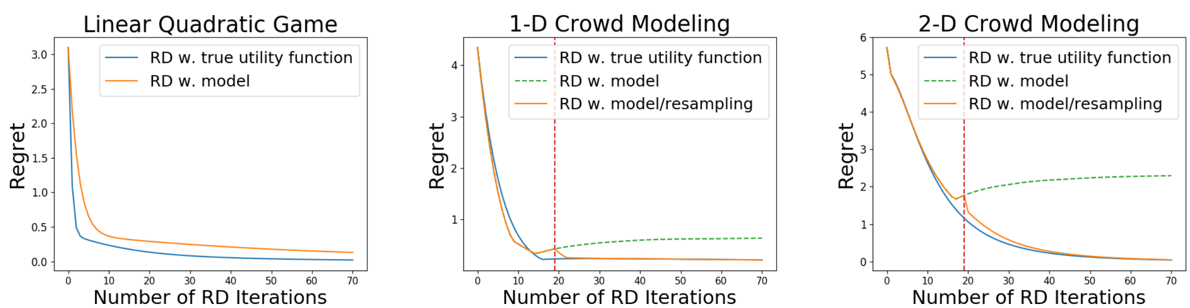

Figure 2 Regret curves with RD.

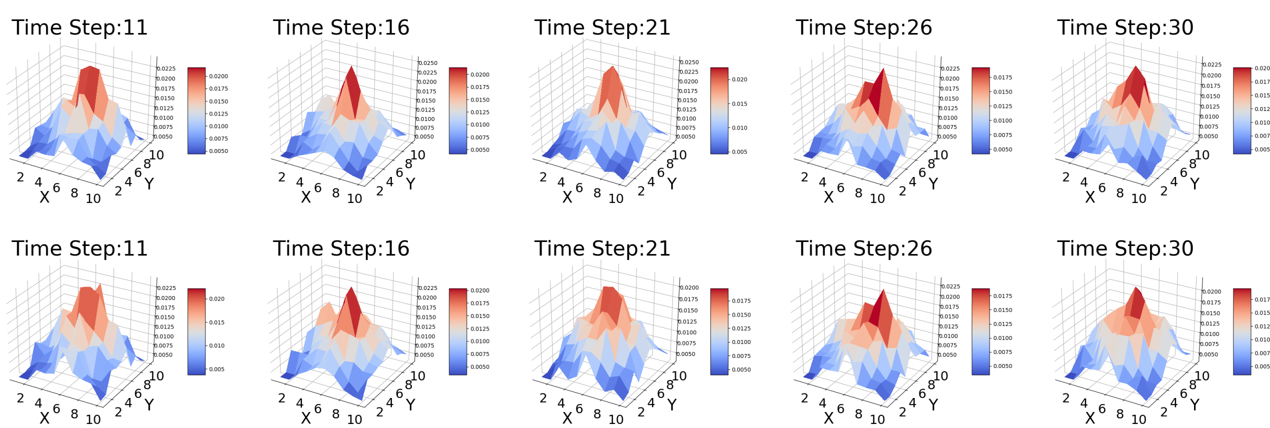

Figure 3 Distribution estimation in 2-D crowd modeling: (top) true utility function; (bottom) game model.

3.2 Mean Field Estimation

To verify that the learned game model can estimate the mean field (i.e., distributions) accurately, we plot the time-dependent distributions induced by the equilibrium strategies computed with the true utility function and the game model in Figure 3. By comparing plots on the top and at the bottom in Figure 3 respectively, we observe that the distributions generated with the game model are almost indistinguishable by inspection from those generated with the true utility function. This accuracy can be quantified by Wasserstein distance, which as we report in Table 1 are all quite tiny (< 0.0005) though with a tendency to increase over time.

MFG	$t = 11$	$t = 16$	$t = 21$	$t = 26$	$t = 30$
2-D Crowd	3.9	3.9	4.9	4.0	4.6

Table 1 Wasserstein distances ($\times 10^{-4}$) in the 1-D and 2-D crowd modeling games.

REFERENCES

[1] Alain Bensoussan, Jens Frehse, Phillip Yam, et al. 2013. *Mean field games and mean field type control theory*. Vol. 101. Springer.

[2] Minyi Huang, Roland P Malhamé, Peter E Caines, et al. 2006. Large population stochastic dynamic games: Closed-loop McKean-Vlasov systems and the Nash certainty equivalence principle. *Communications in Information & Systems* 6, 3 (2006), 221–252.

[3] Jean-Michel Lasry and Pierre-Louis Lions. 2007. Mean field games. *Japanese Journal of Mathematics* 2, 1 (2007), 229–260.

[4] Sarah Perrin, Julien Pérolat, Mathieu Laurière, Matthieu Geist, Romuald Elie, and Olivier Pietquin. 2020. Fictitious play for mean field games: Continuous time analysis and applications. *34st Annual Conference on Neural Information Processing Systems* (2020).

[5] Yongzhao Wang and Michael P Wellman. 2021. Empirical Game-Theoretic Analysis in Mean Field Games. *arXiv preprint arXiv:2112.00900* (2021).