Actor-Critic learning for mean-field control in continuous time

Huyen Pham

We develop policy gradient and actor-critic algorithms for solving mean-field control problems within a continuous time model-free a.k.a. reinforcement learning setting. Our approach leverages a gradient-based representation of the value function, employing parametrized randomized policies. The learning for both the actor (policy) and critic (value function) is facilitated by a class of moment neural network functions on the Wasserstein space of probability measures, and the key feature is to sample directly trajectories of distributions. A central challenge addressed in this study pertains to the computational treatment of an operator specific to the mean-field framework. To illustrate the effectiveness of our methods, we provide a comprehensive set of numerical results. These encompass diverse examples, including multi-dimensional settings and nonlinear quadratic mean-field control problems with controlled volatility.