

Solving Elliptic HJB Equations in A Value Space

Jiamin JIAN

Jungang Bu

Yiyang Che

Abstract

In this short report, we introduce the HJB equation and give a benchmark problem. To solve this problem, we will use the reinforcement approach such as value iteration and Q-learning. And we also want to apply the deep neural network to get the numerical solution of the HJB equation. What's more, we will consider what is the highest dimension we can solve.

1 Introduction

Stochastic control theory has a long history and it has been widely used in various fields. Generally we introduce the corresponding Hamilton-Jacobi-Bellman equations when we want to solve the stochastic control problem. As the regularity of the system is often not known a prior due to the control process, and since the solutions of the HJB equations are usually highly nonlinear, closed-form solution are rarely obtainable. Thus sometimes it is more convenient and efficient to consider the numerical solution of HJB equation. Numerical methods such as finite element method, finite difference method and spectral method are often used to get the numerical solution of partial differential equations.

Nowadays, with the strengthening of computer computing ability and the continuous development of machine learning and especially deep learning, we have an effective tool to solve some problems. Artificial neural networks can sometimes get unexpected great results in solving high dimensional problems. And there are increasing interests of using the reinforcement learning and deep learning technologies to solve the PDE problems. In this project we want to apply the reinforcement learning and deep learning approach based on the Markov chain approximation methods to solve a type a HJB equations. It involves two steps. Firstly, we approximate the solutions of the underlying stochastic control problems and recast them into discrete time Markov decision processes. Then, we solve the MDPs by using reinforcement learning and deep learning approaches.

2 Problem Setup

In this section we discuss the HJB equation associated to the control problem and give a benchmark problem for the demonstration of the computational method.

2.1 HJB equation

We consider an elliptic nonlinear HJB equation in the d -dimensional Euclidean space \mathbb{R}^d and \mathcal{O} is the domain,

$$F(x, Du(x), D^2u(x)) = 0, \forall x \in \mathcal{O} \quad (1)$$

with its boundary condition

$$u(x) = g(x), \forall x \in \mathcal{O}^c, \quad (2)$$

where $F : \mathbb{R}^d \times \mathbb{R}^d \times \mathbb{R}^{d \times d} \mapsto \mathbb{R}$ is a given continuous function of the type

$$F(x, p, q) = \otimes_a \{b(x, a) \cdot p + \frac{1}{2} \text{tr}(q) + l(x, a)\}. \quad (3)$$

In the above, the \otimes is an operator which satisfies the following conditions:

- $\otimes_a [c\phi_1(x, a) + \phi_2(x)] = c \otimes_a [\phi_1(x, a)] + \phi_2(x)$
- $\otimes_a^x \phi_1(x, a) \leq \otimes_a \phi_2(x, a)$, whenever $\phi_1 \leq \phi_2$
- $\otimes_a \phi_1(x, a) - \otimes_a \phi_2(x, a) \leq K \max_a |\phi_1(x, a) - \phi_2(x, a)|$,

where ϕ_1, ϕ_2 are real valued functions and c, K are constant. Many frequently used operators satisfy the aforementioned conditions. For instance, $\max_a \phi(x, a)$ and $\min_a \phi(x, a)$ are related to standard HJB equations and stochastic control problems.

2.2 Benchmark problem

To illustrate the effect of our numerical method, we consider a benchmark problem, on which we can compare the numerical solutions of differential equations with the exact ones. We consider the following d -dimensional HJB equation:

- Domain:

$$\mathcal{O} = \{x \in \mathbb{R}^d : 0 < x_i < 1, i = 1, 2, \dots, d\}.$$

- Equation on \mathcal{O} :

$$\frac{1}{2} \Delta u + \inf_{|a| \leq 3} (a \cdot \nabla u + l(x, a)) = 0,$$

where $l(x, a) = d + 2|x|^2 + \frac{1}{2}|a|^2$.

- Dirichlet data on the boundary $\partial\mathcal{O}$:

$$u(x) = -|x|^2.$$

And we can get that the exact solution of the above problem is $u(x) = -|x|^2$ with $a = 2x$, and F can be written as

$$F(x, p, q) = \inf_{|a| \leq 3} \left\{ \frac{1}{2} \text{tr}(q) + a \cdot p + l(x, a) \right\}.$$

3 Methodology and plan

We want to solve the HJB equation by reinforcement learning approach. The goals of our project are listed as follows:

- Recover the value iteration on central finite difference method.
- Do the value iteration on upwind finite difference scheme.
- Using Q-learning to solve MDPs generated by central finite difference and upwind finite difference scheme.
- Consider what is the highest dimension we can compute.