East Asian Journal on Applied Mathematics
doi: 10.4208/eajam.190420.200420

Vol. **10**, No. 4, pp. 635-658
November 2020

A Stochastic Gradient Descent Approach for Stochastic Optimal Control

Richard Archibald[1], Feng Bao[2,*] and Jiongmin Yong[3]

[1]*Computational Science and Mathematics Division, Oak Ridge National Laboratory, Oak Ridge, Tennessee, USA.*
[2]*Department of Mathematics, Florida State University, Tallahassee, Florida, USA.*
[3]*Department of Mathematics, University of Central Florida, Orlando, Florida, USA.*

Received 19 April 2020; Accepted (in revised version) 20 April 2020.

Abstract. In this work, we introduce a stochastic gradient descent approach to solve the stochastic optimal control problem through stochastic maximum principle. The motivation that drives our method is the gradient of the cost functional in the stochastic optimal control problem is under expectation, and numerical calculation of such an expectation requires fully computation of a system of forward backward stochastic differential equations, which is computationally expensive. By evaluating the expectation with single-sample representation as suggested by the stochastic gradient descent type optimisation, we could save computational efforts in solving FBSDEs and only focus on the optimisation task which aims to determine the optimal control process.

AMS subject classifications: 65K10, 49M37, 49M25

Key words: Stochastic optimal control, stochastic gradient descent, maximum principle, forward backward stochastic differential equations.

1. Introduction

Stochastic optimal control is an important and active research topic in applied mathematics, and it has extensive applications in numerous areas, including engineering, finance and economics, biology, public health, communication networks, to mention a few [17,36]. In the past half century, fundamental results of stochastic optimal control theory have been established: Pontryagin type maximum principle (MP, for short) [12, 29, 30], Bellman dynamic programming principle (DPP, for short) [8, 9] and Hamilton-Jacobi-Bellman (HJB, for short) equation theory [13], and linear-quadratic (LQ, for short) optimal control and Riccati equation theory [23, 34]. These are three well-recognised mile stones of stochastic optimal control theory.

*Corresponding author. *Email address:* `bao@math.fsu.edu` (F. Bao)

It is known that except for some limited special cases, such as LQ problems, one-dimensional linear state equation with convex/concave performance index (such as Merton type problem in mathematical finance), most stochastic optimal control problems are not explicitly solvable and therefore numerical algorithms to generate approximate solutions are needed. One of the most widely used approaches to solve the stochastic optimal control problem is the above-mentioned DPP, mainly due to Bellman [8, 9]. The main idea of the DPP approach is to consider a family of optimal control problems with different initial states and times, and establish relationships among these problems through the HJB equation, which is a fully nonlinear PDE. Taking the advantage of well-established numerical schemes for solving PDEs, many computational methods for stochastic optimal control are developed under the DPP approach [15, 16, 22, 33]. Although all of these methods solve the control problem successfully, due to the complexity of numerical approximations for solutions of PDEs and the nonlinearity of the HJB equation, methods that follow the DPP approach are computationally expensive, and even infeasible when the dimension exceeds 3, although, in recent years, some efforts have been made to pursue the relaxation of the dimensional restriction [7]. Another disadvantage of DPP is that the optimal control problem considered could not have any state constraints. The presence of state constraints will lead to the discontinuity of the value function, for which the suitable HJB equation theory is not available as of today.

Another important approach to solve the stochastic optimal control problem is the stochastic maximum principle. The classic deterministic maximum principle was first introduced by Pontryagin and his students [12,30]. Stochastic version was developed by several researchers since 1960s [10, 11, 21, 24, 29]. The central idea of maximum principle is that any optimal control problem must satisfy an optimisation condition of a function called the Hamiltonian, and it is much easier to optimise a Hamiltonian than solving the original optimal control problem, which is infinite-dimensional. The MP approach for stochastic optimal control problems has three major advantages over the DPP approach: First, there is no dimension restriction; Second, it allows to have some state constraints, especially some finite dimensional terminal state constraints; Third, it allows to have random coefficients in the state equation and/or in the performance functional (to be optimised).

The goal of this paper is to introduce a stochastic gradient descent approach to solve stochastic optimal control problems under the stochastic maximum principle framework, which leads to a stochastic Hamiltonian system that consists of two forward backward stochastic differential equations (FBSDEs) [26]. In this way, solving stochastic optimal control problems through stochastic maximum principle involves solving FBSDEs that meet certain optimisation condition, which is typically achieved by gradient descent based approaches. It can be shown (under appropriate assumptions) that the gradient process of the optimisation condition can be expressed by a FBSDE system, and one may also convert the stochastic optimal control into second order FBSDEs [40]. Therefore, numerical implementation of stochastic maximum principle requires solving FBSDEs repeatedly to reach the optimisation condition. However, since computational methods for solving FBSDEs are not as well developed as those for solving PDEs, numerical studies for stochastic optimal control through stochastic maximum principle are only beginning.