Journal of Computational Mathematics
Vol.42, No.3, 2024, 638–661.

http://www.global-sci.org/jcm
doi:10.4208/jcm.2112-m2021-0084

APPROXIMATING THE STATIONARY BELLMAN EQUATION BY HIERARCHICAL TENSOR PRODUCTS*

Mathias Oster[1], Leon Sallandt and Reinhold Schneider
Technische Universität Berlin, Strasse des 17. Juni 135, 10623 Berlin, Germany
Emails: oster@math.tu-berlin.de, sallandt@math.tu-berlin.de, schneidr@math.tu-berlin.de

Abstract

We treat infinite horizon optimal control problems by solving the associated stationary Bellman equation numerically to compute the value function and an optimal feedback law. The dynamical systems under consideration are spatial discretizations of non linear parabolic partial differential equations (PDE), which means that the Bellman equation suffers from the curse of dimensionality. Its non linearity is handled by the Policy Iteration algorithm, where the problem is reduced to a sequence of linear equations, which remain the computational bottleneck due to their high dimensions. We reformulate the linearized Bellman equations via the Koopman operator into an operator equation, that is solved using a minimal residual method. Using the Koopman operator we identify a preconditioner for operator equation, which deems essential in our numerical tests. To overcome computational infeasibility we use low rank hierarchical tensor product approximation/tree-based tensor formats, in particular tensor trains (TT tensors) and multi-polynomials, together with high-dimensional quadrature, e.g. Monte-Carlo. By controlling a destabilized version of viscous Burgers and a diffusion equation with unstable reaction term numerical evidence is given.

Mathematics subject classification: 49L20, 15A69, 49M41, 49N35.
Key words: Feedback control, Dynamic programming, Hamilton-Jacobi-Bellman, Tensor product approximation, Variational Monte-Carlo.

1. Introduction

In optimal control theory finding a feedback law enables us to get a robust online control for dynamical systems. One prominent approach to find an optimal feedback law is calculating the value function, which can be done by solving either the Bellman equation or the Hamilton-Jacobi-Bellman equation (HJB). Popular numerical solutions to this problem are semi-Lagrangian methods [18, 23, 67], Domain splitting algorithms [24], variational iterative methods [36], data based methods with Neural Networks [48, 50] or Policy Iteration with Galerkin ansatz [37, 46].

The dynamical systems under consideration are spatial discretizations of non linear parabolic partial differential equations (PDE). The dimension of the HJB and the Bellman equation equals the size of the spatial discretization of the PDE, which in theory is infinite and in practice is extremely high. Two principal difficulties are the non linearity and that it may be posed in

* Received March 19, 2021 / Revised version received July 28, 2021 / Accepted December 23, 2021 / Published online April 11, 2023 /
1) Corresponding author

high spatial dimensions. One can address the non linearity by the Policy Iteration. Fixing a policy reduces the non linear Bellman equation to a linear equation. The policy is updated by an optimality condition. The solution of the linearized Bellman is the remaining numerical bottleneck and is equivalent to a linear hyperbolic PDE, a linearized HJB equation. We stress that in order to synthesize the optimal feedback control a function representation of the value function is needed. Indeed, it is not sufficient to have point-evaluations of the value function.

Most methods for the numerical solution of hyperbolic PDEs struggle with the curse of dimensionality, i.e. the exponential growth of complexity with respect to the dimension of the underlying dynamical system. To alleviate this problem different methods have been proposed, like combinations of Proper Orthogonal Decomposition (POD) and semi-Lagrangian methods [1], POD and tree structures [2], efficient polynomial Galerkin approximation and model reduction [37] or, recently, tensor based approaches [19, 34].

By using the Koopman/composition operator our approach transforms the linearized Bellman equation into an operator equation. A solution is approximated via a Least-Squares/minimal residual method on a finite dimensional ansatz space, e.g. multi dimensional polynomials. Even for ordinary differential equation (ODE) systems with considerably few variables this ansatz space becomes huge. By solving the Least-Squares problem on the non linear manifold of Tensor Trains with fixed ranks we reduce the complexity from an exponential to a polynomial dependency on the dimensions of the underlying system. Tensor trains are particular cases of hierarchical (Tucker) or tree-based tensor representations [5, 27]. In principle, the Least-Squares method requires the evaluation of high-dimensional integrals, which is practically infeasible. Therefore, the integration must be replaced by numerical quadrature. Monte-Carlo and quasi Monte-Carlo methods are a canonical choice, since they do not suffer from the curse of dimensionality. We call the resulting discrete approach variational Monte-Carlo (VMC) [21]. The term originates in quantum physics, where one minimizes the energy of a quantum system to find the ground state [14, 49]. Shortly after, [71] introduced the ideas of empirical risk minimization, which are related to the variational Monte-Carlo concept. In [21] both approaches are unified.

It turns out that the Koopman operator can be evaluated point-wise in certain quadrature points by computing trajectories of the underlying dynamical system. In contrast to the linearized HJB the linearized Bellman equation can be solved model-free, i.e. without explicit knowledge of the underlying dynamical system. Finally, we remark that the present treatment of high-dimensional operators in the tensor setting differs essentially from direct treatments as in [5, 27]. In there, the underlying partial differential operator has an explicit representation or approximation in a low rank tensor form. However, the issue of finding an explicit representation/approximation of underlying operators in low rank tensor form compromises the class of treatable problems severely. Using the model-free VMC approach the underlying operators only have to be evaluated at sample points and thus circumvent the issue of representing the operator. In particular, this means that we do not represent the Koopman operator in a low rank tensor format and it is unclear whether such a representation exists. Instead, we only evaluate its action at certain points.

Analogously to [42] one can also incorporate control constraints in terms of projection operators. The generalization of the present approach to stochastic control problems and finite horizon problems is discussed in [22, 54]. In particular, there is a stochastic counterpart of the deterministic Koopman operator, which is the semi group generated by the backward Kolmogorov operator [16, 40, 45]. We would like to mention recent groundbreaking progress in the