Master Project

Conjugate Dynamic Programming

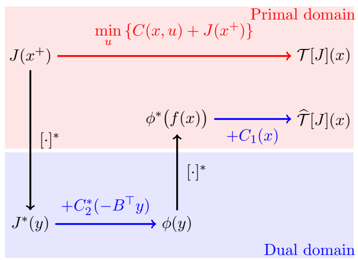
Charalampos Rodopoulos, Systems and Control MSc, TU Delft
C.Rodopoulos@student.tudelft.nl

Amin Sharifi Kolarijani, Delft Center for Systems and Control, TU Delft
M.A.SharifiKolarijani@tudelft.nl

Gyula Félix Max, Delft Center for Systems and Control, TU Delft
G.F.Max@tudelft.nl

Peyman Mohajerin Esfahani, Delft Center for Systems and Control, TU Delft
P.MohajerinEsfahani@tudelft.nl

Context



Dynamic Programming (DP) provides a powerful tool to compute optimal policies in a wide range of applications. Be it inventory management, LQ control or even stock market, the solution to the DP gives an optimal action to take. The basic idea of DP is to solve the Bellman equation

$$J_t(x_t) = \min_{u_t}\{C(x_t, u_t) + J_{t+1}(x_{t+1})\},$$

backward in time t for the costs-to-go J_t, where $C(x_t, u_t)$ is the cost of taking the control action u_t at the state x_t. Arguably, the most important drawback of DP is in its high computational cost in solving problems with a large scale *finite* state-action space. For problems with a *continuous* state space, which is commonly the case in engineering applications, solving the Bellman equation amounts to solving infinite number of optimization problems. This usually renders the exact implementation of the DP operation impossible, except for a few cases with an available closed form solution, e.g., Linear Quadratic Regulator. Therefore, efficient numerical approaches are a necessary alternative. In [1], the advantageous nature of sovling the DP operation in the Legendre domain has been shown. The proposed algorithm involves discretization of the state and input spaces and provides an algorithmic transformation of the minimization in DP operation to addition via discrete conjugation. In particular, for a specific class of problems with separable cost $C(x,u) = C_1(x) + C_2(u)$ and input-affine dynamics $x^+ = f(x) + Bu$, the typical time complexity of the DP operation is reduced from $\mathcal{O}(XU)$ to $\mathcal{O}(X + U)$, where X

and U denote the size of the discrete state and input spaces, respectively. Inspired by this, we aim to expand on that knowledge both from a theoretical and computational point.

The focus of this project revolves around two domains. First, the computational aspect of the problem should be investigated. This could involve improvements in the current implementation in order to achieve a more concrete advantage over traditional methods. Furthermore, having already shown the computational edge of working in the convex conjugate domain, we aim to expand the theoretical setting under which utilising the Legendre domain becomes advantageous. The second objective is to investigate the relation, if any, between working in the Legendre domain and similar domains from the literature in order to uncover further improvements in either the theoretical setting or the implementation of the algorithm.

Project tasks

This master thesis project is aimed at improving the setting under which the Legendre-Fenchel transform can be used to solve the DP problem. The tasks proposed in this thesis are:

1. Extract the optimal control action from within the convex conjugate domain for improved computational efficiency.

2. Extend the theoretical setting of the current algorithm in order to include the discounted infinite-horizon problems.

3. Uncover advantageous relation between the Legendre-Fenchel transform and other transform in the literature.

References

[1] M. A. S. Kolarijani and P. Mohajerin Esfahani. “Fast Approximate Dynamic Programming for Input-Affine Dynamics”, arXiv:2008.10362 (Aug. 2020).