Nonlinear Feedback Stabilization of Bilinear Systems by Model Predictive Control and Q-Learning

M. Q. Phan[1] **and S. M. B. Azad**[2]

Abstract: Model Predictive Control (MPC) emerged as one of the most powerful modern control methods for dynamical systems. MPC minimizes a finite-time receding cost function to solve an infinite time optimal control problem. As the prediction horizon increases, the MPC solution approaches the infinite time optimal solution, thus providing a balance between optimal performance and computational complexity. MPC can incorporate system identification, leading to data-based control methods that design feedback controllers for stabilization, tracking, and/or disturbance rejection directly from input-output measurements.

In parallel with the development of modern control theory, Reinforcement Learning (RL) offers a seemingly different framework to solve problems that are often not found in standard optimal control theory texbooks such as playing games or solving puzzles. However, RL can also be applied to solving standard optimal control problems. Instead of minimizing cost as virtually all of optimal control theory aim to achieve, RL maximizes reward. Although both minimizing cost and maximizing reward are ultimately based on the same principle of optimality, this simple divergence in thinking creates a conceptual gap between the two approaches that is becoming more and more difficult to bridge over time. There is a distinct need to understand reinforcement learning at a very detailed level using the language of modern control theory.

In previous work we completed the effort to relate a standard reinforcement learning method called Q-Learning to LQR (Linear Quadratic Regulator) which is a standard modern control method for linear systems. MPC is where the connection between reinforcement learning and modern control theory can be made. We derived expressions for the Q-functions employed in Q-Learning using the framework of modern control theory. We showed when and how all these methods produce identical feedback gains. Moreover, we combined model predictive control with traditional Q-Learning to produce a new algorithm called Model Predictive Q-Learning (MPQ-L). We discussed both the similarities and the differences between the model-based methods of LQR and MPC and the model-free methods of Q-Learning and MPQ-L. In this work, we present a similar set of results for bilinear systems. By Carleman's linearization, a large class of nonlinear systems can be approximated as bilinear systems. Thus this work also has direct applications to the control of nonlinear systems in general.

[1,2] Thayer School of Engineering
Dartmouth College
14 Engineering Drive, Hanover, NH 03755, USA
{*mqphan, seyed.mahdi.basiri.azad*}*@dartmouth.edu*