Optimal Control For An Inverted Pendulum

Sari Tarabay
University of Toronto
sari.tarabay@mail.utoronto.ca
1009685057

Adam Wei
University of Toronto
adam.wei@mail.utoronto.ca
1006068283

Abstract— **We explored 3 optimal control strategies, LQR, MPC, and adaptive MPC, to control an inverted pendulum on a cart. Our experiments showed that LQR could stabilize the system; however, input constraints could render it ineffective. MPC is a more powerful control strategy that can address this issue by explicitly considering input and state constraints. Our MPC controller could achieve stability, even when the maximum input magnitude was clamped to $50N$. Unfortunately, MPC's performance suffered when the it was provided an incorrect estimate of the cart's damping parameter. To remedy this, we designed an adaptive MPC controller that used gradient descent to estimate the damping online. This enabled the controller to be more robust to model error. Even when the initial estimate of the damping was incorrect, adaptive MPC could still stabilize the system by improving the accuracy of its model estimate. Overall, adaptive MPC was the most effective at satisfying our design specifications.**

I. Introduction

Optimal control is a branch of mathematical optimization that aims to find optimal control inputs for a dynamic system to satisfy a set of design specifications [1]. In this paper, we survey three different optimal controllers for an inverted pendulum on a cart. Namely, we experimented with the Linear Quadratic Regulator (LQR) [2], Model Predictive Control (MPC) [3], and adaptive MPC [4]. Each controller's performance was evaluated based on their ability to satisfy the following design specifications:

1) Stabilize the pendulum in at a displacement of 10m;
2) Satisfy input and state constraints;
3) Robustness to model error and unmodelled disturbances.

This problem is difficult since the upright equillirbrium for the inverted pendulum is unstable [5]. Therefore, precise control is required to stabilize the system and reject disturbances. Amongst the 3 designs we explored, Adaptive MPC was the most effective at satisfying the design specifications.

II. Dynamics For An Inverted Pendulum On A Cart

The system we aim to control is the inverted pendulum on a cart in Figure 1. This system has 4 states: the cart displacement x, cart velocity $\dot{x}$, pendulum angle θ, and angular velocity $\dot{\theta}$. Its equations of motion are shown in Equation 1 and the system parameters, m, M, L, g, and d, are defined in Figure 1.

$$\ddot{x} = \frac{mg\cos(\theta)\sin(\theta) - mL\dot{\theta}^2\sin(\theta) - d\dot{x} + u}{M + m\sin^2(\theta)}$$
$$\ddot{\theta} = \frac{(M+m)g\sin(\theta) - mL\dot{\theta}^2\sin(\theta)\cos(\theta) - d\cos(\theta)\dot{x} + \cos(\theta)u}{L(M + m\sin^2(\theta))} \tag{1}$$

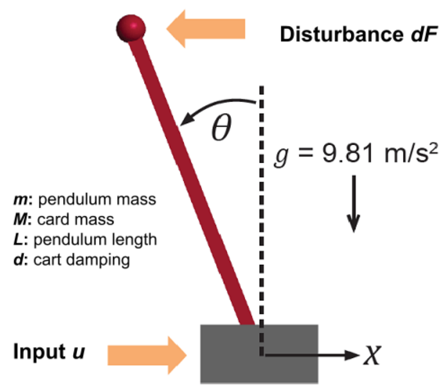

Fig. 1. A diagram of the inverted pendulum on a cart with labels for the relavent system parameters.

Let $\mathbf{x} = [x\ \theta\ \dot{x}\ \dot{\theta}]^T$ be the state vector. Then the system can be linearized about its upright equillibrium such that $\dot{\mathbf{x}} \approx A\mathbf{x} + Bu$. A and B are defined symbolically in Equation 2. This linearized system can also be discretized with period T. In the discretized system, $\mathbf{x}[k+1] = A_d\mathbf{x}[k] + B_d u[k]$, where $A_d = e^{AT}$ and $B_d = \int_0^T e^{At}dtB$ [6].

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & -\frac{d}{M} & \frac{mg}{M} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & -\frac{d}{ML} & \frac{(m+M)g}{ML} & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ \frac{1}{M} \\ 0 \\ \frac{1}{ML} \end{bmatrix} \tag{2}$$

III. Control Methodologies

A. Linear Quadratic Regulator (LQR)

Given a system with dynamics $\dot{\mathbf{x}} = A\mathbf{x} + Bu$, an LQR controller minimizes the quadratic cost function shown below [2]. In Equation 3, n and m are the number of states and inputs respectively, and $Q \in S_+^n$ and $R \in S_{++}^m$ are gain matrices. The optimal control law that minimizes $J(u(t))$ is the state feedback controller $u(t) = -K\mathbf{x}(t)$, where K is a gain matrix that can be computed by solving the Algebraic Ricatti Equation [2]. We implemented an LQR controller for the inverted pendulum system using the linearized dynamics in Equation 2.

$$J(u(t)) = \int_0^\infty \mathbf{x}(t)^T Q\mathbf{x}(t) + \mathbf{u}(t)^T R\mathbf{u}(t)dt \tag{3}$$

While LQR is sufficient for many applications, it has some limitations. Since LQR solves an unconstrained optimization problem, it cannot explicitly consider input or state constraints [7]. In the remainder of this section, we explore more powerful optimal control techniques that resolve this issue.

B. Model Predictive Control (MPC)

Model Predictive Control is a powerful control strategy that can explicitly consider a system's input and state constraints [3]. Given a system with discretized linear dynamics, MPC solves a quadratic program (QP), like the one presented in Equation 4, to determine the optimal control inputs for all N timesteps in the *prediction horizon*; however, only inputs in the *control horizon* are applied to the system [3]. Thus, the QP must be solved once per control horizon to continually supply inputs to the system. In practise, the control horizon is significantly shorter than the prediction horizon, and the QP must be solved many times per second. By continually re-planning its control strategy, the controller can more effectively react to inaccurate predictions or changes within its environment [8].

$$\begin{aligned} \min_{\mathbf{x}_k, u_{k-1}, k\in[N]} & \sum_{k=1}^{N} (\mathbf{x}_d - \mathbf{x}_k)^T Q (\mathbf{x}_d - \mathbf{x}_k) + u_k^T R u_k \\ \text{s.t. } & \mathbf{x}_{k+1} = A_d \mathbf{x}_k + B_d u_k \ \forall \ k \in [0, N-1] \\ & l \leq u_k \leq u \end{aligned} \tag{4}$$

In the QP above, Q and R are positive semidefinite gain matrices, Intuitively, MPC aims to minimize the state error while keeping its inputs small and satisfying the constraints.

C. Adaptive MPC

In adaptive MPC, the control inputs are also computed by solving the QP in Equation 4; however, at every time step, we can update the controller's system model. This is important because accurate system models can be difficult to obtain [9], while others systems are subject to online changes [4].

There are many different ways of updating the controllers model [4][10][11]. We chose to use a gradient based method to estimate the damping coefficient of the system. We chose to estimate the damping coefficient since frictional parameters tend to be challenging to identify a priori [12].

Given the state, input, and an estimate of the damping parameter, $\hat{d}$, we can compute the *predicted* state of the system at the next time step using Euler's method: $\hat{\mathbf{x}}[k+1] = \mathbf{x}[k] + T\dot{\mathbf{x}}[k]$. T is the period of the discretized system and $\dot{\mathbf{x}}$ can be computed using Equation 1. The damping estimate $\hat{d}$ is updated every iteration using gradient descent with momentum to minimize the L2 prediction error. The system's linearization is also updated every iteration to reflect the changes in $\hat{d}$.

$$\begin{aligned} v[k] &= \frac{\partial}{\partial d} \|\mathbf{x}[k] - \hat{\mathbf{x}}[k]\|_2^2 + \eta v[k-1], \ v[0] = 0 \\ \hat{d}[k+1] &= \hat{d}[k] - \alpha v[k], \ \hat{d}[0] = d_0 \end{aligned} \tag{5}$$

IV. Results

All 3 controllers were implemented and simulated in MATLAB. This section presents and discusses the results.

A. LQR Control

Figure 2 shows the response of the system using an LQR controller. Note that the pendulum is able to track the reference trajectory with reasonable transients.

An issue with LQR controllers is that they cannot consider input and state constraints; thus, these controllers can request large control inputs that may not always be feasible. For instance, at $t \approx 0$, the magnitude of the input signal is greater than $100N$. If we clamp the magnitude of the input signal

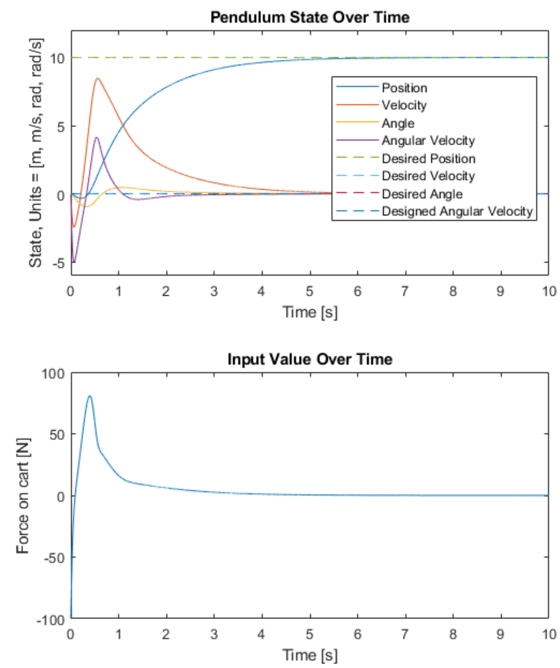

Fig. 2. System response and control input for the LQR controller with no input constraints.

to $50N$, then our LQR controller (which is unaware of the input clamp) is unable to stabilize the system. The system's response using LQR control and input clamps is shown in Figure 3. Note that the design objectives are no longer met.

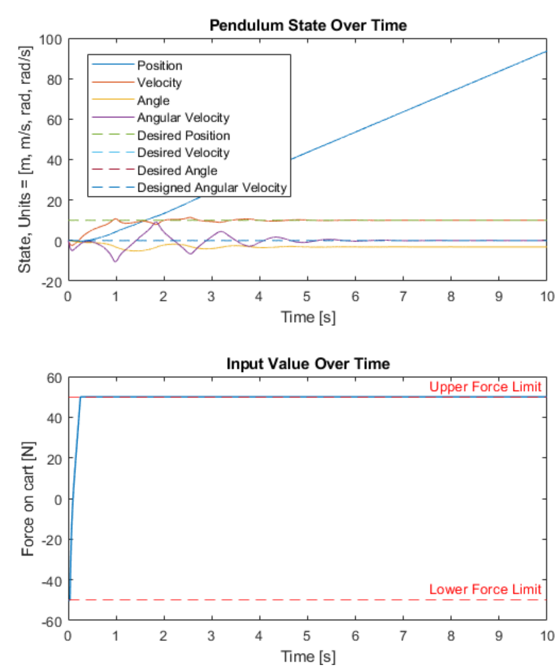

Fig. 3. System response for the LQR controller with input constraints.

B. MPC

Figure 4 shows the system response using an MPC controller. Unlike LQR, the MPC controller successfully stabilizes the system at its desired state in roughly 15s. This is because the MPC controller is aware of the input constraints and computes its control strategy accordingly.

A limitation of MPC is that it requires an accurate system model. Figure 5 shows the system response when the controller uses a damping parameter of $3\frac{Ns}{m}$ instead of the correct value of $5\frac{Ns}{m}$. The model error causes the MPC controller to drive the system out of the region where the linearized dynamics are accurate; afterwards, it is impossible to recover. This illustrates that vanilla MPC lacks robustness to model error.

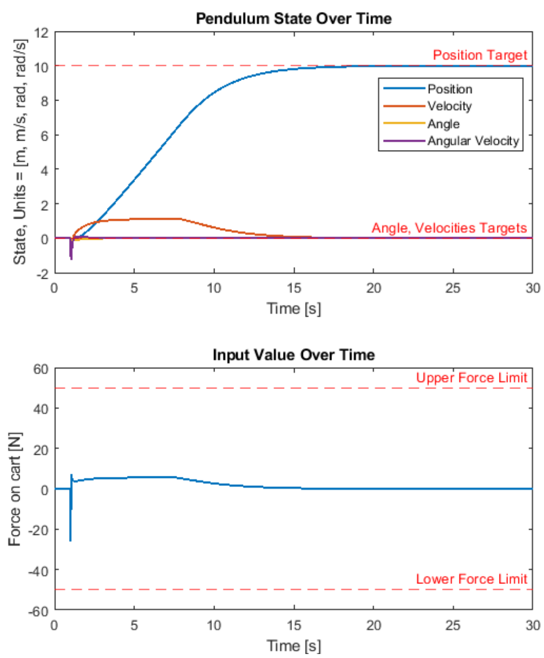

Fig. 4. System response input for the MPC controller with no model error.

C. Adaptive MPC

Figure 6 shows the response of the system of the adaptive MPC controller using the same incorrect estimate of the damping parameter as the experiment in Figure 5. The response exhibits a few interesting characteristics.

Firstly, the controller's damping estimate improves over time, eventually settling at $4.85\frac{Ns}{m}$. The final estimate is not exactly at $5\frac{Ns}{m}$ due to a subtlety in the gradient computation. Since the $\hat{\mathbf{x}}$ term in Equation 5 is computed with Euler's method, the gradient's accuracy degrades when the transients are fast. This occurs at $t \approx 12.5s$ when the system overshoots its target position and stabilizes shortly afterwards. The stablized system is unable to correct the estimation error since the gradient with respect to damping vanishes when the velocity is 0.

Finally, the transient response is slower than that of the error-free MPC controller in Figure 4. This is because the system begins with an underestimate of the damping, causing the initial control inputs to be unnecessarily conservative.

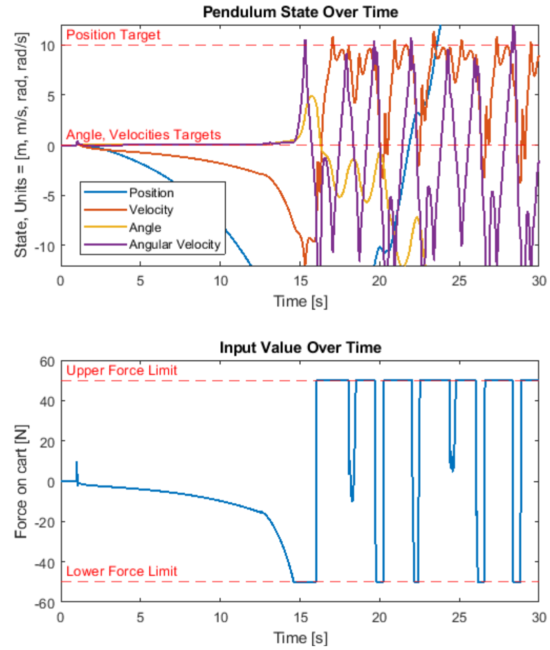

Fig. 5. System response for the MPC controller with an incorrect estimate of the damping parameter.

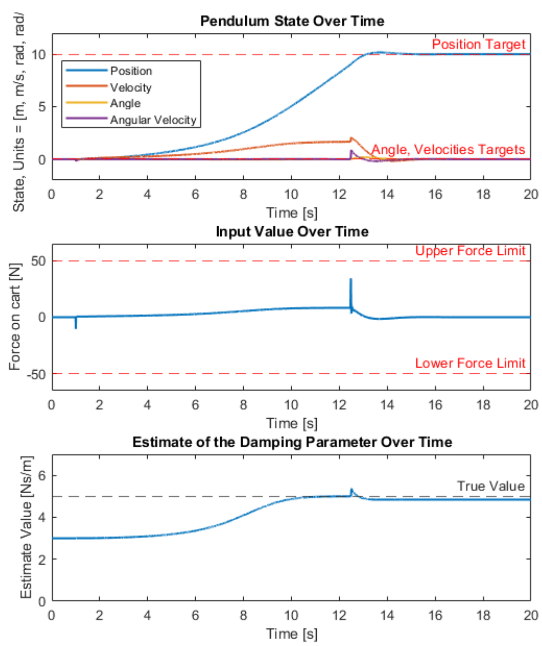

Fig. 6. System response for the Adaptive MPC controller.

V. CONCLUSIONS

We explored 3 optimal control strategies, LQR, MPC, and adaptive MPC, for the inverted pendulum on a cart. For each design, we presented simulation results using MATLAB and discussed their advantages and limitations. Out of the 3 control strategies, adaptive MPC satisfied the design specifications from Section I most effectively. In future work, we hope to evaluate our designs on more complex systems and to explore other model estimation techniques for adaptive MPC.

References

[1] I. Ross, *A Primer on Pontryagin's Principle in Optimal Control: Second Edition*. Collegiate Publishers, 2015.

[2] R. Tedrake, "Chapter 8: Linear quadratic regulators," 2022.

[3] M. Schwenzer, M. Ay, T. Bergs, and D. Abel, "Review on model predictive control: an engineering perspective," *The International Journal of Advanced Manufacturing Technology*, 2021.

[4] K. Pereida and A. P. Schoellig, "Adaptive model predictive control for high-accuracy trajectory tracking in changing conditions," pp. 7831–7837, 2018.

[5] P. Morasso, T. Nomura, Y. Suzuki, and J. Zenzeri, "Stabilization of a cart inverted pendulum: Improving the intermittent feedback strategy to match the limits of human performance," *Frontiers in Computational Neuroscience*, 4 2019.

[6] N. Karampetakis and A. Gregoriadou, "On a first order hold discretization for singular systems," *2011 International Conference on Communications, Computing and Control Applications, CCCA 2011*, 03 2011.

[7] P. Scokaert and J. Rawlings, "Constrained linear quadratic regulation," *IEEE Transactions on Automatic Control*, vol. 43, no. 8, pp. 1163–1169, 1998.

[8] B. Kouvaritakis and M. Cannon, "Model predictive control," *Switzerland: Springer International Publishing*, vol. 38, 2016.

[9] A. Nagabandi, K. Konoglie, S. Levine, and V. Kumar, "Deep dynamics models for learning dexterous manipulation," 2019.

[10] M. Tanaskovic, L. Fagiano, and V. Gligorovski, "Adaptive model predictive control for linear time varying mimo systems," *Automatica*, vol. 105, pp. 237–245, 2019.

[11] L. Wu, "An interpretative and adaptive mpc for nonlinear systems," 2022.

[12] M. Afrough and A. A. Hanieh, "Identification of dynamic parameters and friction coefficients," *Journal of Intelligent & Robotic Systems*, vol. 94, no. 1, pp. 3–13, Apr 2019.