Inverted Pendulum Optimal Control

Kaela Nelson*, Sydney Thompson*, Casey Wahl*, Joseph Ward*

December 15, 2018

Abstract

Our project aims to solve the inverted pendulum problem with an optimal solution using control theory learned in class. We derived the physics of the inverted pendulum system and then used it to create our state equations for this system. We then adjusted parameters to optimize our solution further. Our solution is robust to over 15,000 tested random state initializations. On average, our control solves the inverse pendulum problem in 89 time steps.

1 Introduction

Given a pendulum with with mass m and length l on top of a cart with mass M, we minimized the time steps required to bring the cart to a stop and balance the pendulum in an unstable vertical equilibrium. This system solves the inverse pendulum problem given an arbitrary mass and length of the pendulum, an initial position, angle, approximate change in position, and approximate change in angle. The given constraints require that the control variable $|u| < 20$, the mass of the pendulum is 0.1 kilograms, and the mass of the cart is 1 kilogram. We wrote code to interact with the provided virtual inverted pendulum through the OpenAI Gym interface.

2 Algorithm

2.1 Physics and State Equations

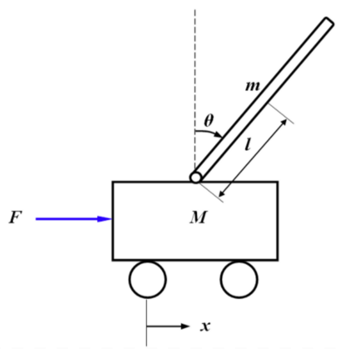

Figure 1: Inverted pendulum [1]

The kinetic energy of the system is given by

$$T = \frac{1}{2}(M+m)\dot{x}^2 + ml\cos(\theta)\dot{\theta}\dot{x} + \frac{1}{2}ml^2\dot{\theta}^2,$$

and the potential energy of the system is given by

$$U = mgl\cos\theta - Fx.$$

From these equations describing the energy system, we can reach the following state equations:

$$\ddot{\theta} = \frac{g\sin(\theta) + \cos(\theta)(\frac{-F - ml\dot{\theta}^2 \sin\theta}{M+m})}{l(\frac{4}{3} - \frac{m\cos^2(\theta)}{M+m})} \tag{1}$$

$$\ddot{x} = \frac{F + ml(\dot{\theta}^2 \sin(\theta) - \ddot{\theta}\cos(\theta))}{m+M}. \tag{2}$$

These formulas were derived from the equations given in Florian [2], which are:

$$\hat{\ddot{x}} = \frac{F + ml(\dot{\theta}\sin(\theta) - \ddot{\theta}\cos(\theta)) - F_f}{M+m}$$

$$\hat{\ddot{\theta}} = \frac{g\sin(\theta) + \cos(\theta)[\frac{-F - ml\dot{\theta}^2\sin(\theta) + F_f}{M+m}] - \frac{\mu_p \dot{\theta}}{ml}}{l[\frac{4}{3} - \frac{m\cos^2(\theta)}{M+m}]}$$

In the above equations, we notice that there are two more constants than in our final equations, F_f and μ_p. These represent the force of friction and the coefficient of friction in the articulation connecting the rod to the cart, respectively. Because we are assuming that the system is frictionless, these constants are both 0. By plugging in 0 for F_f and μ_p, we get the following equations

$$\ddot{\theta} = \frac{g\sin(\theta) + \cos(\theta)(\frac{-F - ml\dot{\theta}^2 \sin\theta}{M+m})}{l(\frac{4}{3} - \frac{m\cos^2(\theta)}{M+m})}$$

$$\ddot{x} = \frac{F + ml(\dot{\theta}^2 \sin(\theta) - \ddot{\theta}\cos(\theta))}{m+M}$$

which match the derived equations (1) and (2).

2.2 Linearization

We made the simplifying assumption that both θ and $\dot{\theta}$ are close to 0. For the optimal solution, this assumption will be true, since the unstable steady state occurs when $\theta = 0$ and $\dot{\theta} = 0$, representing the state where the pole is vertical and not rotating. Thus, we can linearize about 0 for θ and $\dot{\theta}$. Thus, $\theta \to \theta$, $\dot{\theta}^2 \to 0$, $\sin(\theta) \to \theta$, and $\cos(\theta) \to 1$. This gives the equations:

$$\ddot{\theta} = \frac{g\theta + \frac{-F}{M+m}}{l(\frac{4}{3} - \frac{m}{M+m})} = \frac{g}{l(\frac{4}{3} - \frac{m}{M+m})}\theta + \frac{\frac{-1}{M+m}}{l(\frac{4}{3} - \frac{m}{M+m})}F$$

and

$$\ddot{x} = \frac{F + ml(-\ddot{\theta})}{m+M} = \frac{1}{m+M}F + \frac{ml}{m+M}(\frac{-g}{l(\frac{4}{3} - \frac{m}{M+m})}\theta + \frac{\frac{1}{M+m}}{l(\frac{4}{3} - \frac{m}{M+m})})F$$

$$\ddot{x} = \frac{-mlg}{l(\frac{4}{3}(m+M) - m)}\theta + (\frac{1}{m+M} + \frac{m}{\frac{4}{3} - \frac{m}{m+M}})F.$$

We can write this in matrix form as $\dot{z} = Az + Bu$, where

$$z = \begin{bmatrix} x \\ \dot{x} \\ \theta \\ \dot{\theta} \end{bmatrix}, A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & \frac{-mlg}{l(\frac{4}{3}(m+M)-m)} & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{-g}{l(\frac{4}{3} - \frac{m}{M+m})} & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ \frac{1}{m+M} + \frac{m}{\frac{4}{3} - \frac{m}{m+M}} \\ 0 \\ \frac{-1}{l(\frac{4}{3}(m+M)-m)} \end{bmatrix}, u = F.$$

2.3 The LQR Method

Since we have our problem formatted in such a way that we can express it as

$$\dot{z} = Az + Bu$$

with

$$J[z] = \int_0^\infty z^T Qz + u^T Rudt,$$

by Theorem 22.3.1 in Whitehead [3], we can use the LQR method to optimally solve for our control, u.

The LQR method minimizes the following integral

$$J[z] = \int_0^\infty (q_1 x^2 + q_2 \dot{x}^2 + q_3 \theta^2 + q_4 \dot{\theta}^2 + ru^2)dt.$$

Notice that u and $\dot{z}$ can be approximated with

$$\tilde{u} = -R^{-1}B^T P\tilde{z},$$

$$\dot{\tilde{z}} = (A - BR^{-1}B^T P)\tilde{z}$$

where P is a matrix satisfying the Ricatti equation

$$P\dot{(}t) = PA + A^T P + Q - PBR^{-1}B^T P.$$

.

Note that P must satisfy the Ricatti equation because of the Hamiltonian of $J[u]$.

As described in "The Inverted Pendulum" [4], since we are considering an infinite time horizon problem, we have $\dot{P} = 0$. Thus, P is a constant matrix, and we do not have to include a time dependency. Hence, we need only satisfy the equation

$$0 = PA + A^T P + Q - PBR^{-1}B^T P.$$

Then, we can estimate $\dot{z}$ with $\dot{\tilde{z}}$ and u with $\tilde{u}$. Note that these are estimations because we use approximate solvers for P.

2.4 Optimizing the Algorithm

Since the setup of the LQR problem does not include the option to optimize over the final time, we found that we could have the most impact on our solution by changing the values of q_2, q_3, q_4 and r. This will alter the cost of various states. Note that q_1 is relative cost of not being centered in the screen, q_2 is the penalty for the cart moving horizontally, q_3 is the penalty for having a non-vertical rod, and q_4 is the penalty for the rod being in motion.

Since we are not concerned with the position of the cart, we set $q_1 = 0$. We also simultaneously want $|u| < 20$ and the cart to move quickly so that we can minimize the time to solve the problem. Thus, we set $r = 0.008$; we experimentally verified that it maintains $|u| < 20$, but does not penalize u enough that we do not apply sufficient force to quickly solve the problem. We also experimentally verified that setting $q_2 = 20, q_3 = 1$, and $q_4 = 1$ produced a very fast-converging control. We believe this is because it heavily penalizes the cart having non-zero velocity while also penalizing the rod for not being non-upright and still moving. We believe that by heavily penalizing the cart moving, we encourage the algorithm to choose a control that will effectively get the cart into position and then stop moving.

This hypothesis was supported when we reduced our average number of time steps to conversion from approximately 130 to approximately 88. This essentially transitioned from heavily penalizing non-zero θ and $\dot{\theta}$ to heavily penalizing $\dot{x}$.

It is also worthwhile to note here that we re-run the algorithm to compute the optimal control for each time step. This is because our linearizing assumption of θ and $\dot{\theta}$ being close to 0 are not very accurate for most initial conditions produced by the environment. However, our linearized solution would still be relatively close to the non-linearized solution for a small change in time.

Thus, we decided to use the linearized solution and reiterate with our new state at every time step. In this way, we were able to optimally move at each time step while using a relatively simple model. We chose to re-evaluate our control at each time step more as a matter of convenience. We chose to update after each time step because the linearized solution diverged from the non-linearized solution relatively quickly, so to reduce time to reach the unstable steady state, we chose to update our control after each time step.

3 Results

We ran our solution on 120,000 randomly generated initial states, and successfully balanced the pole each time. Below, we can see a histogram of the average number of steps to convergence for each 200 iteration block.

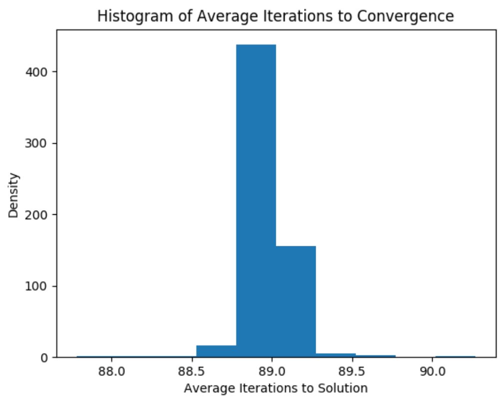

Figure 2: Histogram of average number of iterations to solve the cart pole problem over 200 random initializations.

The histogram in Fig. 2 shows the average number of iterations that our solution converged after running the code on all four of our computers. As we can see, the distribution of the average number of time steps required to solve the problem is highly concentrated between 88.5 and 89 time steps, with an average of 88.966 iterations to convergence. However, we can see from the Fig. 3 that there is a high degree of skewness in the number of time steps required to solve the inverse pendulum problem.

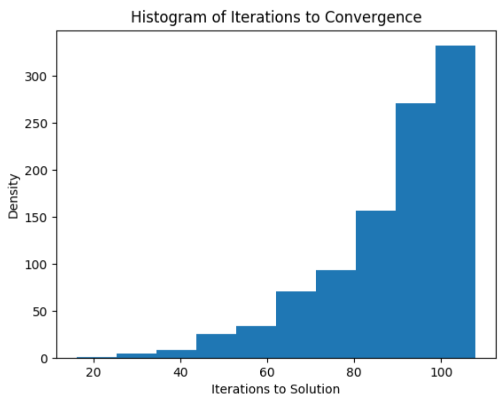

Figure 3: Histogram of time steps required to solve inverse pendulum problem with 1000 random initializations.

This indicates that there are a small number of starting states that are very close to the unsteady state, but most states require a longer time to orient the rod. We also note that for all 15,000 tested randomly initialized starting states, our control properly oriented the pendulum on the cart, and we saw no failures.

4 Conclusion

In conclusion, we used the physics of the inverted pendulum problem to construct state equations for the motion of the cart and pendulum. We linearized our state equations to simplify computation and apply the LQR method to construct an optimal control. The results show that our solution converged within an average of 88.966 iterations.

5 References

[1] https://www.intechopen.com/books/parallel_manipulators_towards_new_applications/mobility_of_spatial_parallel_manipulators
[2] Razvan V. Florian, "Correct equations for the dynamics of the cartpole system," Center for Cognitive and Neural Studies (Coneural) Str. Saturn 24, 400504 Cluj-Napoca, Romania. July 11, 2005; updated February 10, 2007 http://ctms.engin.umich.edu/CTMS/index.php?example=InvertedPendulum&section=SystemModeling
[3] Whitehead, Jared. "Chapter 22. Linear Quadratic Regulator (the 'right' way to optimize)"
[4] "The Inverted Pendulum" http://www.acme.byu.edu/wp-content/uploads/2018/03/21-Inverted-Pendulum.pdf