Optimal Control for CartPole Probelm

Jia-Wei Liao

August 2025

1 Problem Formulation

Objective. We aim to control a cart by applying a horizontal force $u(t)$ such that an inverted pendulum mounted on top of the cart remains stabilized near the *upright* equilibrium position. The system state is defined as

$$\mathbf{x} := \begin{bmatrix} x & \dot{x} & \theta & \dot{\theta} \end{bmatrix}^\mathsf{T},$$

where x denotes the cart position, $\theta = 0$ corresponds to the upright vertical position of the pole, and $\dot{\square}$ denotes differentiation with respect to time. The control input u is the horizontal force applied to the cart (positive to the right).

Parameters. Let m_c be the mass of the cart, m_p the mass of the pole, $M = m_c + m_p$ the total mass, and g the gravitational acceleration. The pole has total length 2ℓ, with its center of mass located a distance ℓ above the pivot. Friction and damping are neglected.

2 Nonlinear Dynamics

We derive the nonlinear equations of motion using a *free-body diagram* combined with translational and rotational dynamics, without resorting to the Euler–Lagrange formulation.

2.1 Geometry and Center-of-Mass Acceleration

The pivot point is located at $(x, 0)$, and the coordinates of the pole's center of mass are

$$\begin{aligned} x_c &= x + \ell \sin\theta, \\ y_c &= \ell \cos\theta. \end{aligned}$$

The velocities are

$$\begin{aligned} \dot{x}_c &= \dot{x} + \ell \cos\theta \dot{\theta}, \\ \dot{y}_c &= -\ell \sin\theta \dot{\theta}. \end{aligned}$$

The accelerations are

$$\begin{aligned} \ddot{x}_c &= \ddot{x} - \ell \sin\theta \dot{\theta}^2 + \ell \cos\theta \ddot{\theta}, \\ \ddot{y}_c &= -\ell \cos\theta \dot{\theta}^2 - \ell \sin\theta \ddot{\theta}. \end{aligned}$$

2.2 Translational and Rotational Dynamics of the Pole

Let (H, V) denote the horizontal and vertical reaction forces at the pivot (positive to the right and upward, respectively). The gravitational force acts at the pole's center of mass as $(0, -m_p g)$. Newton's second law for the pole's center of mass yields

$$m_p \ddot{x}_c = H, \qquad m_p \ddot{y}_c = V - m_p g. \tag{1}$$

The rotational dynamics about the center of mass are given by

$$I_{\text{com}} \ddot{\theta} = \ell \left(H \cos\theta - V \sin\theta \right), \tag{2}$$

where the moment of inertia for a uniform thin rod about its center is $I_{\text{com}} = \frac{1}{12} m_p (2\ell)^2 = \frac{1}{3} m_p \ell^2$.

2.3 Translational Dynamics of the Cart

The cart is subject to the control input u and the horizontal reaction force from the pole, $-H$:

$$m_c \ddot{x} = u - H. \tag{3}$$

2.4 Elimination to Obtain the Main Equations

From (1), we have

$$\begin{aligned} H &= m_p \left(\ddot{x} - \ell \sin\theta \dot{\theta}^2 + \ell \cos\theta \ddot{\theta} \right), \\ V &= m_p \left(-\ell \cos\theta \dot{\theta}^2 - \ell \sin\theta \ddot{\theta} + g \right). \end{aligned}$$

Substituting into (2) and simplifying gives the coupled equation

$$\frac{4}{3} m_p \ell^2 \ddot{\theta} + m_p \ell \cos\theta \ddot{x} + m_p g \ell \sin\theta = 0. \tag{4}$$

Substituting H into (3) yields

$$M\ddot{x} + m_p \ell \cos\theta \ddot{\theta} - m_p \ell \sin\theta \dot{\theta}^2 = u. \tag{5}$$

Solving (4)–(5) leads to the standard nonlinear closed-form:

$$\begin{aligned} \ddot{\theta} &= \frac{g \sin\theta - \cos\theta \frac{u + m_p \ell \dot{\theta}^2 \sin\theta}{M}}{\ell \left(\frac{4}{3} - \frac{m_p}{M} \cos^2\theta \right)}, \\ \ddot{x} &= \frac{u + m_p \ell \dot{\theta}^2 \sin\theta}{M} - \frac{m_p \ell \cos\theta}{M} \ddot{\theta}. \end{aligned}$$

3 Linearization around the Upright Equilibrium

Linearizing about $\theta \approx 0$ and $\dot{\theta} \approx 0$ using $\sin\theta \approx \theta$, $\cos\theta \approx 1$, and neglecting higher-order terms, we define

$$D := \ell \left(\frac{4}{3} - \frac{m_p}{M} \right), \qquad \alpha := \frac{\frac{m_p}{M}}{\frac{4}{3} - \frac{m_p}{M}} = \frac{3m_p}{4M - 3m_p}.$$

The approximated dynamics become

$$\begin{aligned} \ddot{\theta} &\approx \frac{g}{D}\theta - \frac{1}{DM}u, \\ \ddot{x} &\approx -\alpha g \theta + \frac{1+\alpha}{M}u. \end{aligned}$$

Let $\mathbf{x} = [x \quad \dot{x} \quad \theta \quad \dot{\theta}]^\mathsf{T}$. This yields the linear state-space system $\dot{\mathbf{x}} = A\mathbf{x} + Bu$ with

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & -\alpha g & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & \frac{g}{D} & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ \frac{1+\alpha}{M} \\ 0 \\ -\frac{1}{DM} \end{bmatrix}.$$

4 Linear–Quadratic Regulator (LQR)

4.1 Continuous-Time LQR

The quadratic cost is

$$J_c = \int_0^\infty \left(\mathbf{x}(t)^\mathsf{T} Q \mathbf{x}(t) + u(t)^\mathsf{T} R u(t)\right) \mathrm{d}t, \quad Q \succeq 0, R \succ 0.$$

The optimal control law is $u(t) = -K\mathbf{x}(t)$, where

$$\text{CARE:} \quad A^\mathsf{T} P + PA - PBR^{-1}B^\mathsf{T} P + Q = 0, \quad K = R^{-1}B^\mathsf{T} P.$$

4.2 Discrete-Time LQR

Zero-Order Hold. For a simulation or implementation with fixed timestep Δt, the discrete-time matrices (A_d, B_d) are obtained via zero-order hold (ZOH):

$$\begin{bmatrix} A_d & B_d \\ 0 & I \end{bmatrix} = \exp\left(\begin{bmatrix} A & B \\ 0 & 0 \end{bmatrix} \Delta t\right),$$

resulting in the discrete-time model $\mathbf{x}_{k+1} = A_d \mathbf{x}_k + B_d u_k$.

The quadratic cost is

$$J_d = \sum_{k=0}^\infty \left(\mathbf{x}_k^\mathsf{T} Q \mathbf{x}_k + u_k^\mathsf{T} R u_k\right),$$

with optimal law $u_k = -K\mathbf{x}_k$, where

$$\text{DARE:} \quad P = A_d^\mathsf{T} P A_d - A_d^\mathsf{T} P B_d (R + B_d^\mathsf{T} P B_d)^{-1} B_d^\mathsf{T} P A_d + Q,$$
$$K = (R + B_d^\mathsf{T} P B_d)^{-1} B_d^\mathsf{T} P A_d.$$

4.3 Bang–Bang Control

In OpenAI Gym `CartPole-v1`, the action space is binary (left/right with fixed magnitude). Given the continuous LQR output $u^* = -K\mathbf{x}$, a bang–bang mapping is applied:

$$\texttt{action} = \begin{cases} 1, & \text{if } u^* > 0, \\ 0, & \text{otherwise.} \end{cases}$$

5 Summary

In this note, we have derived the nonlinear dynamics of the cart–pole system using Newtonian mechanics, linearized them near the upright equilibrium, discretized the model via zero-order hold, and designed an LQR controller via CARE/DARE to yield the optimal linear state-feedback law $u = -K\mathbf{x}$. This process provides a complete bridge from *physical modeling* to *linear optimal control design*, suitable for both implementation and educational purposes.