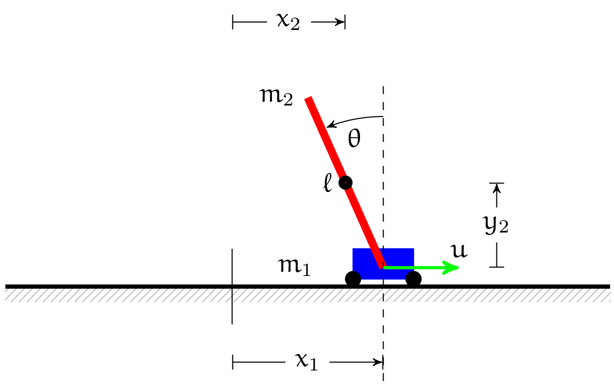
The Cart and Pole Task

Carl Edward Rasmussen

July 3rd 2014



The cart and pole dynamical system consists of a cart (mass m_1) and an attached pedulum (mass m_2, length ℓ) which swings freely in the plane. The pendulum angle, θ, is measured anti-clockwise from hanging down (diagram would be useful). The cart can move horizontally, with an applied external force u, and coefficient of friction r. Typical values are: $m_1 = 0.5$kg, $m_2 = 0.5$kg, $\ell = 0.6$m and $r = 0.1$N/m/s. The coordinates x_2 and y_2 of the midpoint of the pendulum are

$$x_2 \;=\; x_1 - \tfrac{1}{2}\ell\sin\theta, \qquad y_2 \;=\; \tfrac{1}{2}\ell\cos\theta,$$

and the squared velocity of the cart and the midpoint of the pendulum are

$$v_1^2 \;=\; \dot{x}_1^2, \qquad v_2^2 \;=\; \dot{x}_2^2 + \dot{y}_2^2 \;=\; \dot{x}_1^2 + \tfrac{1}{4}\ell^2\dot{\theta}^2 - \ell\dot{x}_1\dot{\theta}\cos\theta.$$

The system Lagrangian is

$$L \;=\; T - V \;=\; \tfrac{1}{2}m_1v_1^2 + \tfrac{1}{2}m_2v_2^2 + \tfrac{1}{2}I\dot{\theta}^2 - m_2gy_2 \;\Rightarrow$$
$$\boxed{L \;=\; \tfrac{1}{2}(m_1+m_2)\dot{x}_1^2 + \tfrac{1}{6}m_2\ell^2\dot{\theta}^2 - \tfrac{1}{2}m_2\ell(\dot{x}_1\dot{\theta}+g)\cos\theta}$$

where we have used the angular moment of inertia around the pendulum midpoint is $I = \frac{1}{12}m\ell^2$, and $g = 9.82\text{m/s}^2$ is the accelleration of gravity.
The equations of motion are

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} \;=\; Q_i,$$

where Q_i are the non-conservative forces. In our case

$$\frac{\partial L}{\partial \dot{x}_1} = (m_1+m_2)\dot{x}_1 - \tfrac{1}{2}m_2\ell\dot{\theta}\cos\theta, \qquad \frac{\partial L}{\partial x_1} = 0, \tag{1}$$

$$\frac{\partial L}{\partial \dot{\theta}} = \tfrac{1}{3}m_2\ell^2\dot{\theta} - \tfrac{1}{2}m_2\ell\dot{x}_1\cos\theta, \qquad \frac{\partial L}{\partial \theta} = \tfrac{1}{2}m_2\ell(\dot{x}_1\dot{\theta}+g)\sin\theta, \tag{2}$$

leading to the equations of motion

$$\boxed{(m_1+m_2)\ddot{x}_1 - \tfrac{1}{2}m_2\ell\ddot{\theta}\cos\theta + \tfrac{1}{2}m_2\ell\dot{\theta}^2\sin\theta = u - r\dot{x}_1, \qquad 2\ell\ddot{\theta} - 3\ddot{x}_1\cos\theta - 3g\sin\theta = 0.}$$

Collecting the four variables $z = (x_1, \theta, \dot{x}_1, \dot{\theta})$ the equations of motion can be conveniently expressed as four coupled differential equations

$$\frac{dz}{dt} = \begin{cases} z_3 \\ z_4 \\ \dfrac{-2m_2\ell z_4^2\sin z_2 + 3m_2 g\sin z_2\cos z_2 + 4u - 4rz_3}{4(m_1+m_2) - 3m_2\cos^2 z_2} \\ \dfrac{-3m_2\ell z_4^2\sin z_2\cos z_2 + 6(m_1+m_2)g\sin z_2 + 6(u - rz_3)\cos z_2}{4\ell(m_1+m_2) - 3m_2\ell\cos^2 z_2} \end{cases},$$

which can be simulated numerically.

Linearized Dynamics

Linearizing the dynamics around the goal state, we can write the following approximation

$$\frac{dz}{dt} \simeq q^{-1}(Az + Bu), \qquad A = \begin{bmatrix} 0 & 0 & q & 0 \\ 0 & 0 & 0 & q \\ 0 & 3\ell m_2 g & -4\ell r & 0 \\ 0 & 6(m_1+m_2)g & -6r & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} 0 \\ 0 \\ 4\ell \\ 6 \end{bmatrix}$$

where $q = \ell(4m_1 + m_2)$.

Loss Function

The instantaneous loss is given by

$$F = 1 - \exp(-\frac{d^2}{2a^2}),$$

were d^2 is the squared distance between the tip of the pendulum and the point at distance ℓ above the origin, and a is the width parameter. Note that the instantaneous loss does not depend on the speed variables, $\dot{x}_1$ and $\dot{\theta}$. The squared distance is

$$d^2 = x^2 + (\ell - y)^2 = (x_1 - \ell\sin\theta)^2 + (\ell - \ell\cos\theta)^2 = (\tilde{z} - \mu)^\top Q(\tilde{z} - \mu),$$

where $\tilde{z}$ is z augmented by two coordinates $\sin\theta$ and $\cos\theta$ and

$$Q = C^\top C, \qquad C = \begin{bmatrix} 1 & 0 & 0 & 0 & -\ell & 0 \\ 0 & 0 & 0 & 0 & 0 & \ell \end{bmatrix}, \qquad \mu = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}^\top.$$

The expected loss, averaging over the possibly uncertain states is therefore

$$\mathbb{E}[F(\tilde{z})] = 1 - \int F(\tilde{z}) p(\tilde{z}) d\tilde{z},$$

which can be evaluated in closed form for Gaussian $p(\tilde{z})$, see reward.pdf. Since the augmented state $\tilde{z}$ will not generally be Gaussian even if z is Gaussian, we project on to the closest Gaussian by matching first and second moments.