From Zero to Cartpole

Stephen Tu
tu.stephenl@gmail.com

December 22, 2016

1 Problem Statement

The purpose of this note is to go through the entire process of deriving a stabilizing controller about the unstable equilibrium of the following system shown in Figure 1.

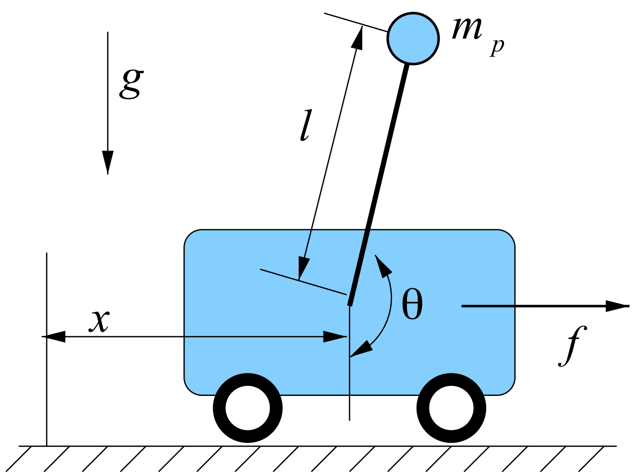

Figure 1: Cartpole free body diagram, taken from Russ Tedrake's 6.832 course notes on underactuated robotics. The mass of the cart m_c is unlabelled in the diagram. For simplicity, the surface is friction-less.

I will focus mainly on the mechanics of the derivation, taking many results for granted. For those interested in the details of *why*, note that everything I present here is standard, with lots of great references available online.

2 Equations of Motion

We start by deriving the governing ordinary differential equation for the system in Figure 1. To do this, we will use the theory of Lagrangian mechanics.

We treat both the mass of the cart and the mass of the pole as two point masses located at $q_1 \in \mathbb{R}^2$ and $q_2 \in \mathbb{R}^2$, respectively. A little bit of trigonometry yields that $q_1 = (x, 0)^\mathsf{T}$ and $q_2 = (x + l\sin(\theta), -l\cos(\theta))^\mathsf{T}$. We next define the Lagrangian $L = T - U$, where T denotes the kinetic energy of the system and U denotes the potential energy. Recalling some basic physics,

$$L = T - U = \frac{1}{2}m_c\|\dot{q}_1\|_2^2 + \frac{1}{2}m_p\|\dot{q}_2\|_2^2 - m_p g q_2(2) \;, \tag{1}$$

where for a vector $q \in \mathbb{R}^n$ we let $\|q\|_2$ denote its ℓ_2-norm and $q(i)$ denote its i-th component. Recalling that x, θ are functions of t, by the chain rule we have

$$\dot{q}_1 = \begin{bmatrix} \dot{x} \\ 0 \end{bmatrix} \;,\;\; \dot{q}_2 = \begin{bmatrix} \dot{x} + l\cos(\theta)\dot{\theta} \\ l\sin(\theta)\dot{\theta} \end{bmatrix} \;.$$

Plugging our expressions for $\dot{q}_1$ and $\dot{q}_2$ into (1), after a bit of algebra we get that

$$L = \frac{1}{2}m_c\dot{x}^2 + \frac{1}{2}m_p\left(2l\cos(\theta)(g + \dot{\theta}\dot{x}) + l^2\dot{\theta}^2 + \dot{x}^2\right) \;.$$

The Euler-Lagrange equations state that

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{x}} = \frac{\partial L}{\partial x} + f \;, \tag{2a}$$

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{\theta}} = \frac{\partial L}{\partial \theta} \;. \tag{2b}$$

From (2a) and (2b),

$$\begin{aligned}(m_c + m_p)\,\ddot{x} - f + l m_p\ddot{\theta}\cos(\theta) - l m_p\dot{\theta}^2\sin(\theta) = 0 \;, \\ g l m_p\sin(\theta) + l^2 m_p\ddot{\theta} + l m_p\cos(\theta)\ddot{x} = 0 \;.\end{aligned}$$

Solving this simultaneous set of equations for $\ddot{x}$ and $\ddot{\theta}$,

$$\ddot{x} = \frac{f + m_p\sin(\theta)(g\cos(\theta) + l\dot{\theta}^2)}{m_c + m_p\sin^2(\theta)} \;, \tag{3a}$$

$$\ddot{\theta} = -\frac{f\cos(\theta) + (m_p + m_c)g\sin(\theta) + l m_p\dot{\theta}^2\sin(\theta)\cos(\theta)}{l(m_c + m_p\sin^2(\theta))} \;. \tag{3b}$$

This is an *ordinary differential equation* (ODE). Letting $q \in \mathbb{R}^4$ be defined as

$$q = \begin{bmatrix} x \\ \theta \\ \dot{x} \\ \dot{\theta} \end{bmatrix} \;,$$

we have from (3a) and (3b) that $\dot{q} = F(q, f)$ for some $F : \mathbb{R}^4 \times \mathbb{R} \longrightarrow \mathbb{R}^4$. Furthermore, it is easy to check that $q_{\text{eq}} = (0, \pi, 0, 0)^\mathsf{T}$ is an (unstable) equilibrium point, since $F(q_{\text{eq}}, 0) = 0$. This equilibrium point corresponds to the cartpole being positioned vertically upright with no velocity and no input applied.

For the remainder of this document, we will focus on stabilizing the system about q_{eq}, under the assumption that the system starts in a state q_0 such that $\|q_0 - q_{\text{eq}}\|_2$ is sufficiently small.

3 Linearizing about the Unstable Equilibrium Point

Our next step will be to linearize the equations of motion (3a) and (3b) about q_{eq}. We do this for purely computational reasons. As we will see shortly, computing an optimal controller for linear dynamical systems is quite straightforward.

To compute the linearization, we define $e_q = q - q_{\text{eq}}$. Taking a first order Taylor expansion of F about $(q_{\text{eq}}, 0)$,

$$\begin{aligned}
\dot{e_q} = F(q,u) &= F(e_q + q_{\text{eq}}, u) \\
&= F(q_{\text{eq}}, 0) + D_q F(q_{\text{eq}}, 0) e_q + D_f F(q_{\text{eq}}, 0) u + o(\|q - q_{\text{eq}}\|_2) + o(|u|) \\
&= D_q F(q_{\text{eq}}, 0) e_q + D_f F(q_{\text{eq}}, 0) u + o(\|q - q_{\text{eq}}\|_2) + o(|u|) \,.
\end{aligned}$$

Above, $D_q F(q_{\text{eq}}, 0) \in \mathbb{R}^{4\times 4}$ is the derivative of $F(\cdot, 0)$ evaluated at $q = q_{\text{eq}}$, and $D_f F(q_{\text{eq}}, 0) \in \mathbb{R}^4$ is the derivative of $F(q_{\text{eq}}, \cdot)$ evaluated at $f = 0$. Ignoring the higher order terms, this writes $\dot{e_q}$ as the following linear dynamical system

$$\dot{e_q} = A_{\text{eq}} e_q + B_{\text{eq}} u \,, \;\; A_{\text{eq}} = D_q F(q_{\text{eq}}, 0) \,, B_{\text{eq}} = D_f F(q_{\text{eq}}, 0) \,.$$

Now we work on computing $D_q F$ and $D_f F$. Write $F(q,f) = (F_1(q,f), F_2(q,f), F_3(q,f), F_4(q,f))$, with each $F_i(q,f)$ a real-valued function. We have $F_1(q,f) = e_3^\mathsf{T} q$ and $F_2(q,f) = e_4^\mathsf{T} q$, so $D_q F_1(q,f) = e_3^\mathsf{T}$ and $D_q F_2(q,f) = e_4^\mathsf{T}$, where e_i is the i-th standard basis vector.

Next, the expressions for $D_q F_3$ and $D_q F_4$ are a bit tedious, but when we evaluate them at the equilibrium point q_{eq}, many terms simplify. Skipping a lot of tedious calculations,

$$\begin{aligned}
D_q F_3(q_{\text{eq}}, 0) &= (0, g\frac{m_p}{m_c}, 0, 0) \,, \\
D_q F_4(q_{\text{eq}}, 0) &= (0, \frac{g(m_c + m_p)}{l m_c}, 0, 0) \,.
\end{aligned}$$

We can perform the calculations for $D_f F_i(q_{\text{eq}}, 0)$ in a similar manner. Putting it all together,

$$A_{\text{eq}} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & g\frac{m_p}{m_c} & 0 & 0 \\ 0 & \frac{g(m_c+m_p)}{l m_c} & 0 & 0 \end{bmatrix} \,, \;\; B_{\text{eq}} = \begin{bmatrix} 0 \\ 0 \\ \frac{1}{m_c} \\ \frac{1}{l m_c} \end{bmatrix} \,. \tag{4}$$

Let us study a few properties of this linearized system. First, we have that the eigenvalues of A_{eq} are given by

$$\lambda(A_{\text{eq}}) = \left\{ 0, 0, \pm\sqrt{\frac{g}{l}\left(1 + \frac{m_p}{m_c}\right)} \right\} \,.$$

Hence, unsurprisingly, the system is not stable. It is however, controllable, since

$$\operatorname{rank}\left(\begin{bmatrix} B_{\text{eq}} & A_{\text{eq}} B_{\text{eq}} & A_{\text{eq}}^2 B_{\text{eq}} & A_{\text{eq}}^3 B_{\text{eq}} \end{bmatrix}\right) = \operatorname{rank}\left(\begin{bmatrix} 0 & \frac{1}{m_c} & 0 & \frac{g m_p}{l m_c^2} \\ 0 & \frac{1}{l m_c} & 0 & \frac{g(m_c+m_p)}{l^2 m_c^2} \\ \frac{1}{m_c} & 0 & \frac{g m_p}{l m_c^2} & 0 \\ \frac{1}{l m_c} & 0 & \frac{g(m_c+m_p)}{l^2 m_c^2} & 0 \end{bmatrix}\right) = 4 \,.$$

Therefore, from linear systems theory, we know we are able to design a feedback controller $K_{\text{eq}} \in \mathbb{R}^{1\times 4}$ such that the closed loop system $A_{\text{eq}} - B_{\text{eq}}K_{\text{eq}}$ is stable, and furthermore we can place the eigenvalues of the closed loop matrix $A_{\text{eq}} - B_{\text{eq}}K_{\text{eq}}$ anywhere we want in the open left half plane. We, however, do something different. Instead of pole placement techniques, we turn to optimal control techniques, as they are slightly more intuitive.

4 Optimal Control via the Linear Quadratic Regulator (LQR)

Consider the following optimization problem. Let (A, B) be controllable (weaker conditions are possible), Q positive semi-definite, and R positive definite. Consider minimizing, over all continuous functions $u \in C(\mathbb{R}_+)$, the following objective

$$\min_{u} \; J(u) = \int_0^\infty x^\mathsf{T} Q x + u^\mathsf{T} R u \; dt \; , \; \dot{x} = Ax + Bu \; , x(0) = x_0 \; . \tag{5}$$

(We know that there exists u such that $J(u) < \infty$ since (A, B) is controllable). It is a well known result that the optimal u for (5) is given by $u = -Kx$ where $K = R^{-1}B^\mathsf{T}P$ and P is a solution to the *continuous algebraic Ricatti equation* (CARE),

$$A^\mathsf{T}P + PA - PBR^{-1}B^\mathsf{T}P + Q = 0 \; . \tag{6}$$

Furthermore, there is a unique P solving (6) such that the closed loop system $A - BK$ is stable.

5 Putting the Pieces Together

The LQR theory gives us an easy way to construct a lot of stabilizing controllers by varying the matrices Q and R. The simplest choice is to set $Q = I$ and $R = I$, which corresponds to the cost function $J(u) = \int_0^\infty \|x\|_2^2 + \|u\|_2^2 \; dt$. Another common choice is to set Q and R to be diagonal matrices with positive entries. This allows one to assign relative importance to certain states.

As an example, in the cartpole case, suppose we really want to make sure that x and θ remain close to their equilibrium, but we are willing to tolerate some movement in order to achieve this. Then we might set $Q = \text{diag}(\lambda_1, \lambda_1, \lambda_2, \lambda_2)$ with $\lambda_1 > \lambda_2$.

Let us now check how the LQR controller on $(A_{\text{eq}}, B_{\text{eq}})$ behaves on the actual dynamics (3a) and (3b) near q_{eq}. Figure 2 shows the closed loop trajectories of $x(t)$ and $\theta(t)$ for four initial starting points generated by small perturbations to q_{eq}, and Figure 3 shows the closed loop trajectories of $\dot{x}(t)$ and $\dot{\theta}(t)$ for these same initial configurations. The LQR problem (5) was solved on $(A_{\text{eq}}, B_{\text{eq}})$ from (4) with $Q = I$ and $R = 1$.

From Figure 2 and Figure 3, we are able to see that despite solving the LQR controller on a linearized model of (3a) and (3b), the controller is still able to bring trajectories initiating at points close to q_{eq} towards q_{eq}.

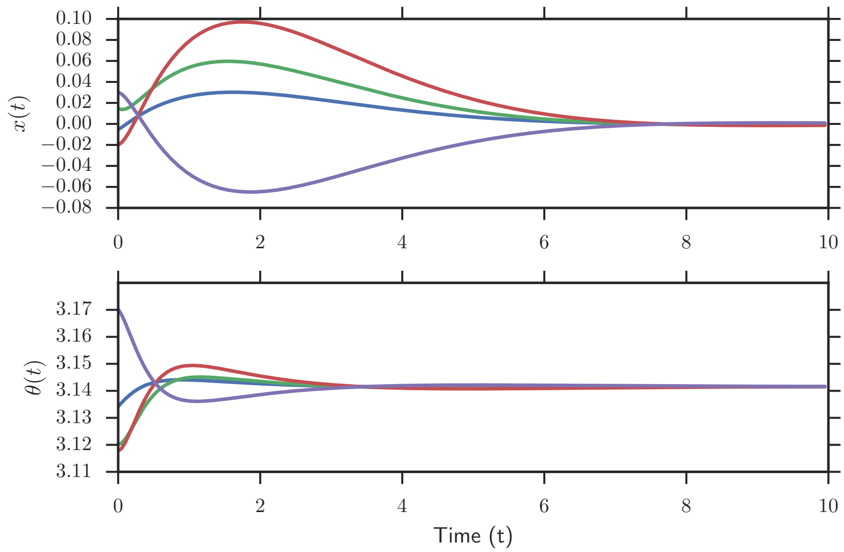

Figure 2: The top figure contains the closed loop trajectories of $x(t)$ on the actual cartpole dynamics (3a) and (3b) using the LQR controller (5) with the linearization (4) for four different starting states which are generated as random perturbations to q_{eq}. Each color corresponds to a particular starting point. The bottom figure shows the trajectory of $\theta(t)$ under the same conditions.

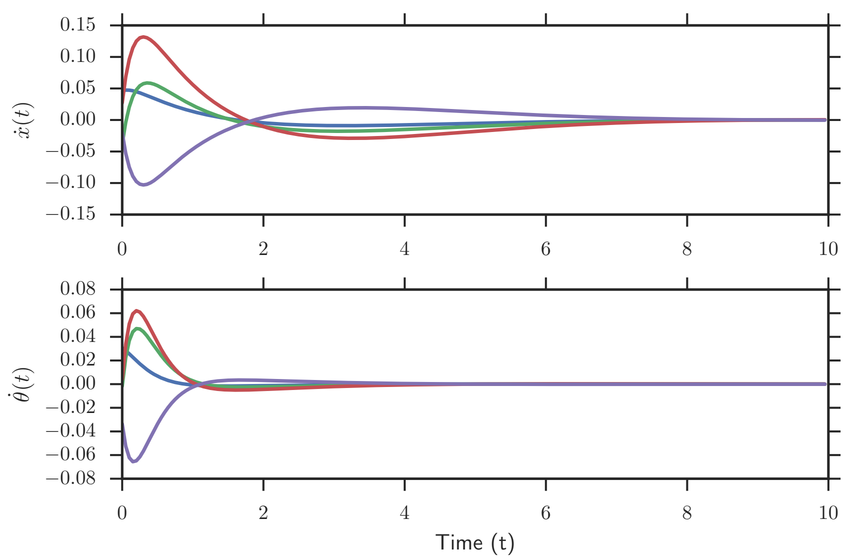

Figure 3: A similar plot to Figure 2 under the same starting conditions, except this time the trajectories of $\dot{x}(t)$ and $\dot{\theta}(t)$ are plotted on the top and bottom, respectively.