Physics of the CartPole environment

Tim Bakker

The CartPole system is an interesting exercise in classical mechanics. Here we provide a derivation of the system's equations of motion using Lagrangian mechanics.

System definition

The CartPole system consists of a frictionlessly moving cart, connected to an inverted physical pendulum. Standing fully upright, the pendulum is in an unstable state, meaning that any perturbation will cause it to start falling.

By initialising the system already in motion, with the pendulum on some small offset from the vertical, we can train an agent to balance the pole. The agent is allowed to move left or right with some (usually predefined) force F, and its goal is to prevent the pole from falling for a given number of timesteps in the simulation.

In order to derive the equations of motion, we first define the variables in our system:

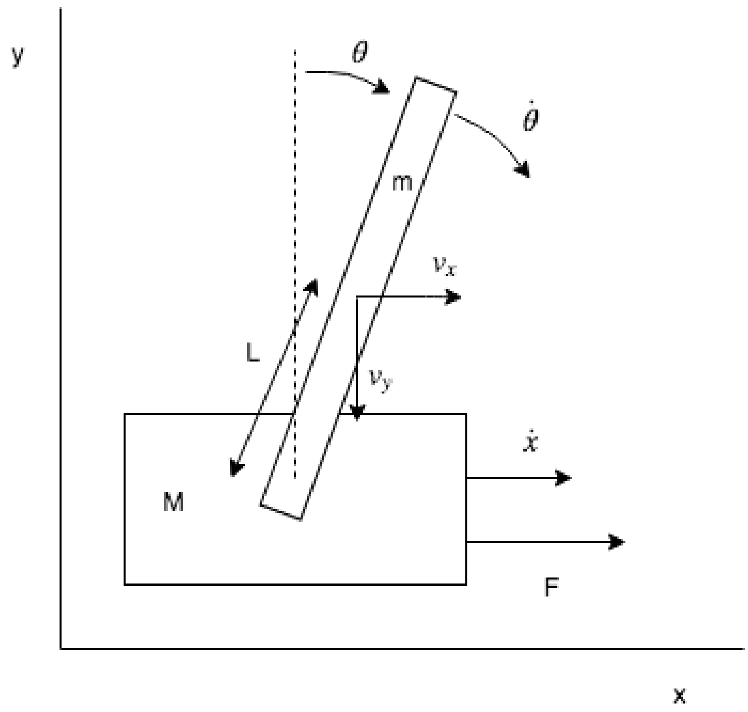

Figure 1: Diagram of CartPole system.

Where the variables indicate:

- x: the x position of the pole's center of mass.

- y: the y position of the pole's center of mass.
- M: the mass of the cart.
- m: the mass of the pole.
- L: half the length of the pole.
- F: the force on the cart (this is the force the learning agent will exert on the cart in a time step).
- θ: the angle between the y-axis and the pole.
- $\dot{\theta}$: the angular velocity of the pole.
- $\dot{x}$: the velocity of the CartPole system in the x direction.
- v_x: the velocity of the pole in the x direction as a result of rotation and the movement of the cart.
- v_y: the velocity of the pole in the y direction as a result of rotation.

Our basic reinforcement learning agent will only have access to x, θ, $\dot{x}$ and $\dot{\theta}$ as the state variable. It will apply a force F either to the left or the right of the cart each time step. This will be all that is required for the agent to balance the pole.

Setting up the Langrangian

Given the Lagrangian of this system, we can find the equations of motion by solving the Euler-Lagrange equations. The Lagrangian ($\mathcal{L}$) is given by the sum of the kinetic (K) and potential energy (V) of the system.

$$\mathcal{L} = K - V$$

The potential energy is simply the gravitational potential energy of the falling pole. For any angle θ the height h of the center of mass of the pole is given by $L\cos\theta$. The gravitational potential energy $V = mgh = mgL\cos\theta$, where g is the gravitational acceleration at the Earth's surface (approximately $9.81m/s^2$).

The kinetic energy of this system consists of two parts:

1. the kinetic energy of the cart, due to its velocity $\dot{x}$.
2. the kinetic energy of the pole, due to both the cart moving, and the rotation caused by falling.

A body of mass m moving with velocity v has translational kinetic energy T equal to $\frac{1}{2}mv^2$. In the case of our cart, this translates to a kinetic energy of $K_{cart} = T_{cart} = \frac{1}{2}M\dot{x}^2$.

Our pole has two kinetic energy components: translational (T) and rotational (R). The translational kinetic energy is given by the same formula as before. However, v now has two components: we introduce the following diagram for clarity:

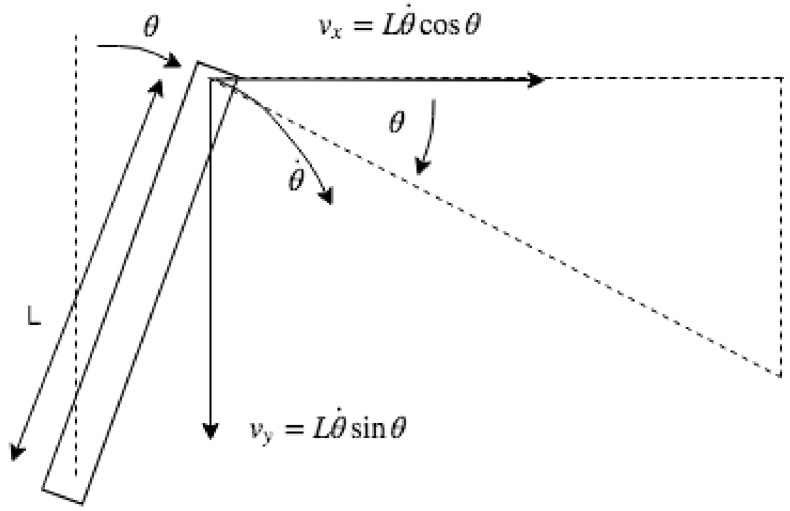

Figure 2: Rotational velocities of the pole (half of the pole pictured).

Note that the diagram doesn't yet include the translational velocity $\dot{x}$. The total translational velocity v_{pole} is given by the Pythagorean theorem as $v_{pole}^2 = v_{pole_x}^2 + v_{pole_y}^2$.

From the diagram we read that $v_{pole_y} = v_y = L\dot{\theta}\sin\theta$, and $v_{pole_x} = \dot{x} + v_x = \dot{x} + L\dot{\theta}\cos\theta$. Thus we find that:

$$v_{pole}^2 = L^2\dot{\theta}^2 + \dot{x}^2 + 2L\cos\theta\,\dot{x}\dot{\theta}$$

Such that the pole has translational kinetic energy $T_{pole} = \frac{1}{2}m(L^2\dot{\theta}^2 + \dot{x}^2 + 2L\cos\theta\,\dot{x}\dot{\theta})$.

The pole's rotational kinetic energy is given by $R_{pole} = \frac{1}{2}I\dot{\theta}^2$, where I is the moment of inertia around the axis of rotation. This is calculated as $I = \int r^2 dm$, where dm represents the mass of an infinitesimal slice of the rod, and r the distance of that slice to the axis of rotation. By noting that we can define a density $\lambda = \frac{m}{2L}$ and that $dm = \lambda dr$, with dr the width of an infinitesimal slice of the rod, we can evaluate this integral as:

$$I = \int_0^{2L} r^2 \frac{m}{2L} dr = \frac{m}{2L}\frac{1}{3}\left[r^3\right]\Big|_0^{2L} = \frac{4}{3}mL^2$$

Thus, $R_{pole} = \frac{2}{3}mL^2\dot{\theta}^2$, and we have acquired all ingredients for setting up the Lagrangian $\mathcal{L}$.

$$\mathcal{L} = \frac{1}{2}(m+M)\dot{x}^2 + mL\left(\frac{7}{6}L\dot{\theta}^2 + \dot{x}\dot{\theta}\cos\theta - g\cos\theta\right)$$

Solving the Euler-Lagrange equations

Hamilton's principle of stationary action states that the equations of motion of a physical system are given as solutions to a variational problem on the system's action functional. This action is fully described by the Lagrangian of the system, and as such the variational condition reduces to the Euler-Lagrange equations. Thus we can find the dynamics of the CartPole system using these equations:

$$\begin{cases} \dfrac{d}{dt}\dfrac{\partial \mathcal{L}}{\partial \dot{x}} - \dfrac{\partial \mathcal{L}}{\partial x} = F \\ \dfrac{d}{dt}\dfrac{\partial \mathcal{L}}{\partial \dot{\theta}} - \dfrac{\partial \mathcal{L}}{\partial \theta} = 0 \end{cases}$$

Where F is the force acting on the cart (in our case the force is applied by the agent every time step).

Plugging in the Lagrangian, we find:

$$\begin{cases} (M+m)\,\ddot{x} + mL\left(\ddot{\theta}\cos\theta - \dot{\theta}\sin\theta\right) = F \\ \dfrac{7}{6}mL^2\ddot{\theta} = mL\,(g\sin\theta - \ddot{x}\cos\theta) \end{cases}$$

Extracting

$$\ddot{x}$$

in the first equation and plugging it into the second equation, we find after some algebra:

$$\begin{cases} \ddot{x} = \dfrac{-\ddot{\theta}\cos\theta}{M+m} + C_F \\ \ddot{\theta} = \dfrac{g\sin\theta - C_F\cos\theta}{\frac{7}{6}L - \frac{m}{M+m}\cos\theta^2} \end{cases}$$

Where we have defined $C_F = \dfrac{F + mL\dot{\theta}^2\sin\theta}{M+m}$ for ease of notation.

We are now all set for simulating CartPole. We initialise the system with some state $\{x, \dot{x}, \theta, \dot{\theta}\}$. Then, every time step we calculate the accelerations $\{\ddot{x}, \ddot{\theta}\}$ based on the previous state, update the velocities $\{\dot{x}, \dot{\theta}\}$ using the accelerations, and finally find new positions $\{x, \theta\}$ using these velocities. The new position and velocities define the state for the next time step. Note that we must first calculate the rotational acceleration $\ddot{\theta}$, because the translational acceleration $\ddot{x}$ depends on it.