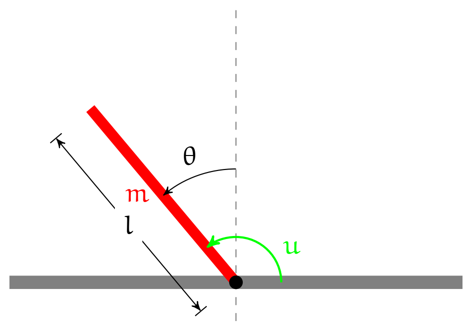
The Pendulum Task

Jonas Umlauft

July 10th 2014



General

The pendulum system consists of pendulum (uniformly distributed mass m, length l) which is attached frictionless on one end. The pendulum angle, θ, is measured anti-clockwise from upright. A external moment u acting on the pendulum at the origin can be applied. Typical values are: $m = 1$kg, $l = 1$m. Two task can be considered

- For the pendulum starting in the upright position $\theta = 0$, the task is to stabilize it in this position. This can be achieved using a linear controller or a nonlinear controller.
- For the pendulum starting in the hanging down position $\theta = \pi$, the task is to swing-up in the upright position and stabilise there. This can only be achieved using a nonlinear controller.

Dynamics

The following equilibrium of moments acting on the pendulum must be fulfilled at all times:

$$\frac{d^2\theta}{dt^2} I = \frac{1}{2} mgl \sin(\theta) + u$$

where $I = \frac{1}{3}ml^2$ is the moment of inertia.
The state vector is defined as $z = \begin{bmatrix}\dot{\theta} & \theta\end{bmatrix}^T$, thus the differential equation is given as following

$$\frac{dz}{dt} = \begin{cases} \frac{3}{ml^2}\left(\frac{mgl}{2}\sin(z_2) + u\right) \\ z_1 \end{cases},$$

Linearized Dynamics

Linearising the dynamics around the goal state $z = \begin{bmatrix} 0 & 0 \end{bmatrix}^\top$, we can write the following approximation

$$\frac{dz}{dt} \simeq (Az + Bu), \qquad A = \begin{bmatrix} 0 & \frac{3g}{2l} \\ 1 & 0 \end{bmatrix}, \qquad B = \begin{bmatrix} \frac{3}{ml^2} \\ 0 \end{bmatrix}$$

Loss Function

Currently exist two implementations for the loss functions

- The instantaneous loss is given by

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

  were a is the width parameter of the cost function and d is the Cartesian distance between the tip of the pendulum and the point at distance ℓ above the origin, The squared distance is

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

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

  $$Q = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & l^2 & 0 \\ 0 & 0 & 0 & l^2 \end{bmatrix}, \qquad \mu = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}.$$

  Note that the instantaneous loss does not depend on the speed variable $\dot{\theta}$. This is currently implemented in loss.m

- The instantaneous loss is given by

  $$F = (\cos(\theta) + 1)/2$$

  i.e. the angular distance to the upright position. This is currently implemented in loss2.m.

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.