Class Tutorial 4

1. The linear quadratic regulator (LQR)

Consider the following deterministic discrete-time linear system:

$$\mathbf{x}_{t+1} = A\mathbf{x}_t + B\mathbf{u}_t, \qquad (1)$$

Where $\boldsymbol{x}_t \in \mathbb{R}^n, \boldsymbol{u}_t \in \mathbb{R}^m, A \in \mathbb{R}^{n\times n}, B \in \mathbb{R}^{n\times m}$ and x_0 is known in advance.

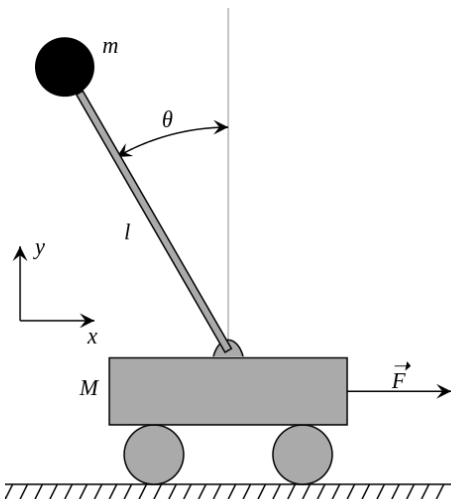
The goal is to choose controls $U = \{\boldsymbol{u}_0, .., \boldsymbol{u}_{T-1}\}$ that minimize the following criterion

$$\boldsymbol{J}(U) = \sum_{t=0}^{T-1} (\boldsymbol{x}_t^T Q \boldsymbol{x}_t + \boldsymbol{u}_t^T Q \boldsymbol{u}_t) + \boldsymbol{x}_T^T Q \boldsymbol{x}_T$$

Where $Q = Q^\top \geq 0, R = R^\top \geq 0, Q_f = Q_f{}^\top \geq 0$ are given state-cost, control-cost, and final-cost matrices.

Example – inverted pendulum on a cart (from Wikipedia)

Consider the following system:



The equations of motion for this system are

$$(M+m)\ddot{x} - ml\ddot{\theta}\cos\theta + ml\dot{\theta}^2\sin\theta = F$$
$$l\ddot{\theta} - g\sin\theta = \ddot{x}\cos\theta$$

The goal is to balance the pendulum in its upright position.

a. Linearize and discretize the system, and write its state evolution in the form (1).

b. Choose Q, Q_f, R that correspond to the desired control goal.

Dynamic programming solution:

Define the value function $V_t(z)$ as

$$V_\tau(z) = \min_{u_\tau,\ldots,u_{T-1}} \left\{ \sum_{t=\tau}^{T-1} \left(\mathbf{x}_t^\top Q \mathbf{x}_t + \mathbf{u}_t^\top R \mathbf{u}_t \right) + \mathbf{x}_T^\top Q_f \mathbf{x}_T \right\} \tag{2}$$

Where $x_\tau = z$ in (2) and the states evolve according to (1).

c. Write down a Bellman equation for $V_\tau(z)$. Could standard dynamic programming be applied to solve the equation? What would be the complexity of a discretization approach?

d. Assume that $V_{\tau+1}(z)$ is of the following form: $V_{\tau+1}(z) = z^\top P_{\tau+1} z$, where $P_{\tau+1} = P_{\tau+1}^\top \geq 0$. Solve the minimization in the Bellman equation explicitly, and show that $V_\tau(z)$ can be written as $V_\tau(z) = z^\top P_\tau z$.

e. Justify the assumption in (d).

f. Write down an algorithm for computing the optimal LQR controller. What is the complexity of the algorithm?

g. Usually, for $\tau \ll T$, P_τ converges rapidly. In practice, often a steady-state P is used. Write an equation for P_{ss} - the steady-state value of P_τ. This equation is known as the (discrete time) algebraic Riccati equation (ARE). What is the steady-state controller?

Solution:

a. We linearize around the upright position $\theta = 0$, and set $\sin\theta \approx \theta, \cos\theta \approx 1, \dot{\theta}^2 \approx 0$

$$\begin{aligned} &(M+m)\ddot{x} - ml\ddot{\theta} = F \\ &l\ddot{\theta} - g\theta = \ddot{x} \end{aligned}$$

We define the control $u = F$ and the state-space vector $\mathbf{x}$ as

$$\mathbf{x} = \begin{pmatrix} x \\ \dot{x} \\ \theta \\ \dot{\theta} \end{pmatrix},$$

And we have the differential equation

$$\dot{\mathbf{x}} = \bar{A}\mathbf{x} + \bar{B}u$$

Where

$$\bar{A}=\begin{pmatrix}0 & 1 & 0 & 0\\ 0 & 0 & \frac{m}{M}g & 0\\ 0 & 0 & 0 & 1\\ 0 & 0 & \frac{g}{l}\left(1+\frac{m}{M}\right) & 0\end{pmatrix},\bar{B}=\begin{pmatrix}0\\ \frac{1}{M}\\ 0\\ \frac{1}{Ml}\end{pmatrix}.$$

Using a first order approximation of the differentiation we have $\mathbf{x}_{t+dt}=\left(I+\bar{A}dt\right)\mathbf{x}_t+\bar{B}dt{\cdot}u_t$, therefore we set $A=I+\bar{A}dt,\quad B=\bar{B}dt$ in (1).

b. For example, let $Q_f=Q=\begin{pmatrix}w_1 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 0 & w_2 & 0\\ 0 & 0 & 0 & 0\end{pmatrix}, R=w_3$. Changing the weights w_1,w_2,w_3 balances between the control-cost and state-cost.

c. The Bellman equation is

$$V_\tau\left(z\right)=z^\top Qz+\min_u\left\{u^\top Ru+V_{\tau+1}\left(Az+Bu\right)\right\}$$

Standard DP doesn't apply since the state and action spaces are continuous. A discretization approach would scale exponentially in n - this is the well-known curse of dimensionality.

d. We have

$$\begin{aligned}V_\tau\left(z\right)&=z^\top Qz+\min_u\left\{u^\top Ru+\left(Az+Bu\right)^\top P_{\tau+1}\left(Az+Bu\right)\right\}\\ &=z^\top Qz+z^\top A^\top P_{\tau+1}Az+\min_u\left\{u^\top\left(B^\top P_{\tau+1}B+R\right)u+2u^\top B^\top P_{\tau+1}Az\right\}\end{aligned}$$

And solving the minimization (by setting the gradient to zero) gives

$$u_{min}=-\left(B^\top P_{\tau+1}B+R\right)^{-1}B^\top P_{\tau+1}Az\,.$$

Substituting $u_{\min}$ and simplifying gives

$$V_\tau\left(z\right)=z^\top\left(Q+A^\top P_{\tau+1}A-A^\top P_{\tau+1}B\left(R+B^\top P_{\tau+1}B\right)^{-1}B^\top P_{\tau+1}A\right)z\doteq z^\top P_\tau z\,.$$

e. For $\tau=T$ we have $V_T\left(z\right)=z^\top Q_f z$. It may be seen that P_τ calculated in (d) satisfies $P_\tau=P_\tau^\top\geq 0$. By induction the result holds for all τ .

f. The LQR algorithm:

- Set $P_T = Q_f$
- Recursively calculate
  $P_\tau = Q + A^\top P_{\tau+1} A - A^\top P_{\tau+1} B\left(R + B^\top P_{\tau+1} B\right)^{-1} B^\top P_{\tau+1} A$
- Define the controller gain $K_t = -\left(B^\top P_{\tau+1} B + R\right)^{-1} B^\top P_{\tau+1} A$
- The optimal LQR controller is the linear feedback controller $u_t^* = K_t x_t$

The complexity is $\mathcal{O}\left(Tn^3\right)$.

g. We have

$$P_{ss} = Q + A^\top P_{ss} A - A^\top P_{ss} B\left(R + B^\top P_{ss} B\right)^{-1} B^\top P_{ss} A,$$

And the controller is a constant linear feedback controller of the form

$$u_t = K_{ss} x_t, \qquad K_{ss} = -\left(B^\top P_{ss} B + R\right)^{-1} B^\top P_{ss} A.$$