Online Lecture Notes

Prof. Boris Houska

June 7, 2022

1 Linear Quadratic Regulator

During the last two lectures we have learned how to solve linear-quadratic optimal control problems of the form

$$\begin{aligned} J(x_0) \quad &\overset{\text{def}}{=} \quad \min_{x,u} \qquad \int_0^T x(t)^\intercal Q x(t) + u(t)^\intercal R u(t) \, \mathrm{d}x + x(T)^\intercal P_T x(T) \\ &\qquad \text{s.t.} \qquad \left\{ \begin{array}{l} \forall t \in [0,T], \\ \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0 \ . \end{array} \right. \end{aligned} \tag{1}$$

Here, $x_0 \in \mathbb{R}^{n_x}$ denotes the intial state, $A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times n_u}$ are the system matrices, $Q \in \mathbb{R}^{n_x \times n_x}$ is a symmetric positive semi-definite weighting matrix penalizing the state deviation and $R \in \mathbb{R}^{n_u \times n_u}$ is a symmetric positive definite control penalty matrix. The matrix $P_T \in \mathbb{R}^{n_x \times n_x}$ denotes a positive semi-definite terminal weight. As usual $x : \mathbb{R} \to \mathbb{R}^{n_x}$ denotes our state trajectory and $u : \mathbb{R} \to \mathbb{R}^{n_u}$ denotes our control input, which both depend on the free variable t (usually time). The above optimal control problem is called a *Linear Quadratic Regulation* (LQR) problem.

1.1 Solution to the LQR Problem

We have learned that the solution to the LQR problem can be found by dynamic programming. It turns our that the objective value function

$$J(x_0) = x_0^\intercal P(0) x_0$$

is a quadratic form in x_0, which can be found by solving the backward Riccati differential equation

$$\begin{aligned} \forall t \in [0,T], \qquad -\dot{P}(t) \quad &= \quad P(t)A + A^\intercal P(t) + Q - P(t)BR^{-1}B^\intercal P(t) \\ P(T) \quad &= \quad P_T \ . \end{aligned}$$

We have also learned that the corresponding optimal feedback control is linear and given in the form

$$u(t) = K(t)x(t) \qquad \text{with} \qquad K(t) = -R^{-1}B^\intercal P(t) \ .$$

This means that we can compute the optimal feedback gain by simulating the Riccati differential equation backward in time and then substituting $P(t)$ to find the above time varying control gain $K(t) \in \mathbb{R}^{n_u \times n_x}$. The optimal input and state trajectory can then be found by a forward simulation

$$\forall t \in [0,T], \qquad \dot{x}(t) = [A + BK(t)]\, x(t) \quad \text{with} \quad x(t) = x_0 \ .$$

1.2 Example: Scalar Control System

Let us start by analyzing a "simple" scalar control system, for instance,

$$\dot{x}(t) = u(t) \qquad \text{with} \qquad x(0) = x_0 \ .$$

Let us also introduce the objective function

$$\int_0^T \left[q \cdot x(t)^2 + r \cdot u(t)^2 \right] \, \mathrm{d}t$$

with $q, r > 0$ being positive (scalar) weights. For simplicity we assume here that our terminal weight is zero. Remember that we have to go through three steps in order to work out what the optimal x and u are:

1. Solve the Riccati differential equation backward in time,
2. compute the optimal feedback gain $K(t)$,
3. simulate the closed-loop system forward in time.

Let us do all three steps explicitly for this particular example. We start by solving the Riccati differential equation

$$\begin{aligned} -\dot{p}(t) &= q - \frac{1}{r} p(t)^2 \\ p(T) &= 0 \ . \end{aligned} \tag{2}$$

If we want to solve this differential equation explicitly, we need to use the concept of separation of variables, which yields:

$$\int_t^T \frac{\dot{p}(\tau)}{\frac{1}{r}p(\tau)^2 - q} \, \mathrm{d}\tau \ = \ \int_t^T 1 \, \mathrm{d}\tau \ = \ T - t \tag{3}$$

Next, we need to work out the integral

$$\begin{aligned} \int_t^T \frac{\dot{p}(\tau)}{\frac{1}{r}p(\tau)^2 - q} \, \mathrm{d}\tau &= r \int_t^T \frac{\dot{p}(\tau)}{p(\tau)^2 - qr} \, \mathrm{d}\tau \\ &= r \int_t^T \frac{\dot{p}(\tau)}{(p(\tau) - \sqrt{qr})(p(\tau) + \sqrt{qr})} \, \mathrm{d}\tau \end{aligned}$$

In order to simplify this expression further. let us introduce the shorthand

$$\gamma = \sqrt{qr} > 0 \ .$$

Next, we can use that

$$\frac{1}{(p(\tau)-\gamma)(p(\tau)+\gamma)} \quad = \quad \frac{\frac{1}{2\gamma}}{p(\tau)-\gamma} + \frac{-\frac{1}{2\gamma}}{p(\tau)+\gamma}$$

Thus, we have

$$\begin{aligned}
r\int_t^T \frac{\dot{p}(\tau)}{(p(\tau)-\gamma)(p(\tau)+\gamma)}\,\mathrm{d}\tau \quad &= \quad \frac{r}{2\gamma}\int_t^T \left[\frac{\dot{p}(\tau)}{p(\tau)-\gamma} + \frac{-\dot{p}(\tau)}{p(\tau)+\gamma}\right] \\
&= \quad \frac{r}{2\gamma}\left|\log(|p(\tau)-\gamma|) - \log(|p(\tau)+\gamma|)\right|_t^T \\
&= \quad \frac{r}{2\gamma}\left|\log\left(\frac{\gamma-p(\tau)}{\gamma+p(\tau)}\right)\right|_t^T \\
&= \quad \frac{-r}{2\gamma}\log\left(\frac{\gamma-p(t)}{\gamma+p(t)}\right) \qquad (4)
\end{aligned}$$

Next, we can find an explicit expression for $p(t)$ by solving the equation

$$\frac{-r}{2\gamma}\log\left(\frac{\gamma-p(t)}{\gamma+p(t)}\right) = T-t$$

This yields

$$\gamma - p(t) = \exp\left(\frac{2\gamma}{r}(t-T)\right)[\gamma+p(t)]$$

and then

$$\left[1+\exp\left(\frac{2\gamma}{r}(t-T)\right)\right]p(t) = \gamma\left[1-\exp\left(\frac{2\gamma}{r}(t-T)\right)\right]$$

This means that we find the explicit expression

$$p(t) = \gamma\frac{1-\exp\left(\frac{2\gamma}{r}(t-T)\right)}{1+\exp\left(\frac{2\gamma}{r}(t-T)\right)}$$

This means that we find the optimal control gain

$$k(t) = -\frac{\gamma}{r}\frac{1-\exp\left(\frac{2\gamma}{r}(t-T)\right)}{1+\exp\left(\frac{2\gamma}{r}(t-T)\right)}$$

In order to understand the properties of this expression for $k(t)$, we analyze the case that T is very large and $t \ll T$. In this case we have

$$\text{if } t \ll T, \quad k(t) \approx -\frac{\gamma}{r} = -\sqrt{\frac{q}{r}}$$

This means that if we choose $q \gg r$ then our control gain will be very large, since we have almost no control cost, but a very large penalty on the state deviation. This leads to very extreme control reactions! The other way around, if we choose $r \ll q$, then the control gain will be very small (and $\dot{x}(t) \approx 0$), this means that we have almost no control reaction and the state decays only very

very slowly to zero. In practice, we would probably choose $q \approx r$ such that we get a reasonable trade-off. Notice that the limit case $T = \infty$ correspond the so-called infinite horizon case. In this case, the control gain becomes a constant; that is we have

$$k_\infty(t) = -\sqrt{\frac{q}{r}} \ .$$

1.3 Infinite Horizon LQR Control

In practice, we are often interested in solving the infinite horizon optimal control problem

$$\begin{aligned} J_\infty(x_0) \quad &\overset{\text{def}}{=} \quad \min_{x,u} \quad \int_0^\infty x(t)^\intercal Q x(t) + u(t)^\intercal R u(t) \, \mathrm{d}x \\ &\qquad\quad \text{s.t.} \quad \begin{cases} \forall t \in [0,T], \\ \dot{x}(t) = Ax(t) + Bu(t) \\ x(0) = x_0 \ . \end{cases} \end{aligned} \tag{5}$$

This problem can be solved by taking the limit for $T \to \infty$. This means that we need to solve the backward Riccati equation

$$-\dot{P}(t) = P(t)A + A^\intercal P(t) + Q - P(t)B^\intercal R^{-1} B P(t)$$

on an infinite time horizon. This is only possible, if P converges to a steady-state P_∞, which satisfies the steady-state condition

$$0 = P_\infty A + A^\intercal P_\infty + Q - P_\infty B^\intercal R^{-1} B P_\infty \ .$$

This equation is called the *Algebraic Riccati Equation* (ARE). One theorem that we won't have time to prove in this lecture is the following:

> The above algebraic Riccati equation has a unique positive definite solution P_∞ if and only if there exists a control that asymptotically stabilizes the system (A, B).

Notice that in the infinite horizon case, the optimal feedback gain is time-invariant,

$$K_\infty = -R^{-1} B^\intercal P_\infty \ .$$

Thus, there are also three steps for solving the infinite horizon LQR problem, which are

1. Compute the positive definite solution P_∞ of the algebraic Riccati equation,
2. compute the optimal infinite horizon feedback law K_∞,
3. solve the associated linear closed-loop system explcitly by using matrix exponentials

1.4 Example

Let us consider the optimal control problem

$$\min_{x,u} \int_0^\infty x(t)^\intercal x(t) + u(t)^\intercal u(t) \, \mathrm{d}t \quad \text{s.t.} \quad \left\{ \begin{array}{l} \forall t \in [0,\infty), \\ \dot{x}(t) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} u(t) \\ x(0) = x_0 \end{array} \right.$$

This corresponds a controlled double-integrator system. Let us work out the algebraic Riccati equation by working out the terms

$$A^\intercal P = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ P_{11} & P_{12} \end{pmatrix}$$

This implies

$$PA + A^\intercal P = \begin{pmatrix} 0 & P_{11} \\ P_{11} & 2P_{12} \end{pmatrix}$$

Moreover, we have

$$\begin{aligned} PB^\intercal R^{-1} BP &= \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} \\ &= \begin{pmatrix} P_{11} & P_{12} \\ P_{12} & P_{22} \end{pmatrix} \begin{pmatrix} 0 & 0 \\ P_{12} & P_{22} \end{pmatrix} \\ &= \begin{pmatrix} P_{12}^2 & P_{12}P_{22} \\ P_{12}P_{22} & P_{22}^2 \end{pmatrix} . \end{aligned} \tag{6}$$

This means that we find that the ARE has the form

$$\begin{aligned} 0 &= PA + PA^\intercal + Q - PBR^{-1}BP \\ &= \begin{pmatrix} 1 - P_{12}^2 & P_{11} - P_{12}P_{22} \\ P_{11} - P_{12}P_{22} & 1 + 2P_{12} - P_{22}^2 \end{pmatrix} \end{aligned} \tag{7}$$

We can solve the upper left equation with respect to P_{12}, which yields

$$P_{12} = \pm 1 \; .$$

Let us substitute this into the lower right equation, which yields

$$P_{22}^2 = 1 \pm 2$$

This equation only has a real solution if we set $P_{12} = +1$. This means that

$$P_{22} = \sqrt{3} \; .$$

Next, we find $P_{11} = P_{12}P_{22} = \sqrt{3}$. This yields

$$P = \begin{pmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix} \quad \text{and} \quad K = -(0,1) \begin{pmatrix} \sqrt{3} & 1 \\ 1 & \sqrt{3} \end{pmatrix} = \begin{pmatrix} -1, & -\sqrt{3} \end{pmatrix}$$

This is an explicit expression for the optimal infinite horizon control gain. Exercise: work out the corresponding closed-loop system gain and work out the optimal solution trajectory by using matrix exponentials.