Online Lecture Notes

Prof. Boris Houska

June 2, 2022

1 Dynamic Programming

Recall from last lecture that we discussed how to discretize the continuous-time linear-quadratic optimal control problem

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

We found that this optimal control problem can be approximated by its discrete version

$$\begin{aligned} \min_{y,v} \quad & \sum_{i=0}^{N-1} \{ y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i \} + y_N^\intercal P y_N \\ \text{s.t.} \quad & \left\{ \begin{array}{l} \forall i \in \{0,1,\ldots,N-1\}, \\ y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_0 = x_0 \; , \end{array} \right. \end{aligned} \tag{2}$$

where we have introduced the shorthands

$$\mathcal{Q} = hQ, \quad \mathcal{R} = hR, \quad \mathcal{A} = I + hA, \quad \text{and} \quad \mathcal{B} = hB \; .$$

We also recall that this approximation is arbitrarily accurate (up to terms of order $O(h)$; even for all integrable (not only continuous) control input functions), which means that we can later take the limit for $h \to 0$ if we want to go back to the continuous-time case. Also notice that if Q is positive semi-definite, then $\mathcal{Q}$ is positive semi-definite (since $h > 0$) and if R is positive definite, then $\mathcal{R}$ is positive definite. This means that our assumptios on Q and R carry over to $\mathcal{Q}$ and $\mathcal{R}$.

1.1 Bellman's Principle of Optimality

Our next goal is to solve the discrete-time optimal control problem (11) by using a so-called dynamic programming strategy, which is based on the principle of optimality. Here, the main is to introduce the so-called cost-to-go function

$$J_i(z_i) \;=\; \min_{y,v} \quad \sum_{k=i}^{N-1} \{y_k^\intercal \mathcal{Q} y_k + v_k^\intercal \mathcal{R} v_k\} + y_N^\intercal P y_N$$

$$\text{s.t.} \quad \begin{cases} \forall k \in \{i, i+1, \ldots, N-1\}, \\ y_{k+1} = \mathcal{A} y_k + \mathcal{B} v_k \ , \\ y_i = z_i \end{cases} \tag{3}$$

The key observation that we can break the horizon into N pieces. If the discrete-time y_i at time i is a minimizer of the full-horizon discrete-time optimal control problem, then all the remaining states from $k = i$ to $k = N$ can be found by solving (3). Notice that the functions J_i satisfy a recursion of the form

$$\begin{aligned}
J_i(z_i) \;&=\; \min_{y,v} \sum_{k=i}^{N-1} \{y_k^\intercal \mathcal{Q} y_k + v_k^\intercal \mathcal{R} v_k\} + y_N^\intercal P y_N \\
&\quad\ \text{s.t.} \begin{cases} \forall k \in \{i, i+1, \ldots, N-1\}, \\ y_{k+1} = \mathcal{A} y_k + \mathcal{B} v_k \ , \\ y_i = z_i \end{cases} \\
&=\; \min_{y,v} \ y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i + \sum_{k=i+1}^{N-1} \{y_k^\intercal \mathcal{Q} y_k + v_k^\intercal \mathcal{R} v_k\} + y_N^\intercal P y_N \\
&\quad\ \text{s.t.} \begin{cases} y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_i = z_i \\ \forall k \in \{i+1, i+2, \ldots, N-1\}, \\ y_{k+1} = \mathcal{A} y_k + \mathcal{B} v_k \ , \end{cases} \\
&=\; \min_{y_i,v_i} \ y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i + J_{i+1}(y_{i+1}) \\
&\quad\ \text{s.t.} \begin{cases} y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_i = z_i \ . \end{cases}
\end{aligned}$$

Notice that the latter equation is a recursion for the function sequence $J_0, J_1, \ldots, J_N$, which has the form

$$J_i(z_i) \;=\; \min_{y_i,v_i} \ y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i + J_{i+1}(y_{i+1})$$

$$\text{s.t.} \quad \begin{cases} y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_i = z_i \ . \end{cases} \tag{4}$$

for all $i \in \{0, 1, \ldots, N-1\}$. Moreover, we have

$$J_N(z_N) = z_N^\intercal P z_N \ . \tag{5}$$

The recursion (4) together with the boundary condition (5) is called a *dynamic programming recursion*. The main idea of dynamic programming that we can solve the discrete-time optimal control problem by starting with the given terminal cost function $J_N(z_N)$ and compute J_{N-1}, J_{N-2}, ..., J_0 by a backward recursion. In our special case that we have a linear system with a quadratic objective, it turns our that we can solve this recursion explicitly! Here, the main observation is that the functions J_i are all quadratic forms. This means that we can find matrices P_N, P_{N-1}, P_{N-2}, ..., P_0 such that

$$J_i(z_i) = z_i^\intercal P_i z_i \; . \tag{6}$$

In order to prove that this is so and in order to derive a recursion of the matrices P_i, we proceed by performing a backward induction over the index i. This mean that our induction start is given by

$$J_N(z_N) = z_i^\intercal P_N z_i = z_N P z_N$$

where we set $P_N = P$. Next, our induction assumption is that J_{i+1} has the form (6), which is the same as saying that we already found P_{i+1} such that $J_{i+1}(z_{i+1}) = z_{i+1}^\intercal P_{i+1} z_{i+1}$. Thus, our induction step takes the form

$$\begin{aligned} J_i(z_i) &= \min_{y_i, v_i} \; y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i + y_{i+1}^\intercal P_{i+1} y_{i+1} \\ &\qquad \text{s.t.} \quad \left\{ \begin{array}{l} y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_i = z_i \; . \end{array} \right. \\ &= \min_{v_i} \; z_i^\intercal \mathcal{Q} z_i + v_i^\intercal \mathcal{R} v_i + (\mathcal{A} z_i + \mathcal{B} v_i)^\intercal P_{i+1} (\mathcal{A} z_i + \mathcal{B} v_i) \\ &= \min_{v_i} \; z_i^\intercal \mathcal{Q} z_i + v_i^\intercal \mathcal{R} v_i + z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{A} z_i + 2 z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{B} v_i + v_i^\intercal \mathcal{B} P_{i+1} \mathcal{B} v_i \\ &= z_i^\intercal \left[\mathcal{Q} + \mathcal{A}^\intercal P_{i+1} \mathcal{A} \right] z_i + \min_{v_i} \left\{ v_i^\intercal \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right] v_i + 2 z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{B} v_i \right\} \end{aligned}$$

In order to solve the latter minimization problem explicitly, we can write out the stationarity condition

$$\begin{aligned} 0 &= \nabla_{v_i} \left\{ v_i^\intercal \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right] v_i + 2 z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{B} v_i \right\} \\ &= 2 \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right] v_i + 2 \mathcal{B}^\intercal P_{i+1} \mathcal{A} z_i \; . \end{aligned} \tag{7}$$

Recall that $\mathcal{R}$ is assumed to positive definite. This means that if P_{i+1} is positive semi-definite, we have

$$v_i^\star = - \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right]^{-1} \mathcal{B}^\intercal P_{i+1} \mathcal{A} z_i \; . \tag{8}$$

Before we interpret this equation as a feedback law, we continue our induction step by substituting the expression for the minimizer v_i into the above equation for J_i, which yields:

$$\begin{aligned} J_i(z_i) &= z_i^\intercal \left[\mathcal{Q} + \mathcal{A}^\intercal P_{i+1} \mathcal{A} \right] z_i + (v_i^\star)^\intercal \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right] v_i^\star + 2 z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{B} v_i^\star \\ &\overset{(8)}{=} z_i^\intercal \left[\mathcal{Q} + \mathcal{A}^\intercal P_{i+1} \mathcal{A} \right] z_i - z_i^\intercal \mathcal{A}^\intercal P_{i+1} \mathcal{B} \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right]^{-1} \mathcal{B}^\intercal P_{i+1} \mathcal{A} z_i \\ &= z_i^\intercal \left[\mathcal{Q} + \mathcal{A}^\intercal P_{i+1} \mathcal{A} - \mathcal{A}^\intercal P_{i+1} \mathcal{B} \left[\mathcal{R} + \mathcal{B} P_{i+1} \mathcal{B} \right]^{-1} \mathcal{B}^\intercal P_{i+1} \mathcal{A} \right] z_i \\ &= z_i^\intercal P_i z_i \end{aligned} \tag{9}$$

with

$$P_i \quad = \quad \mathcal{Q} + \mathcal{A}^\intercal P_{i+1}\mathcal{A} - \mathcal{A}^\intercal P_{i+1}\mathcal{B}\left[\mathcal{R} + \mathcal{B}P_{i+1}\mathcal{B}\right]^{-1}\mathcal{B}^\intercal P_{i+1}\mathcal{A} \tag{10}$$

This completes our induction proof!

1.2 Summary: Discrete-Time LQR Control

In summary, the discrete-time linear quadratic optimal control problem

$$\begin{aligned}
&\min_{y,v} \quad && \sum_{i=0}^{N-1} \{y_i^\intercal \mathcal{Q} y_i + v_i^\intercal \mathcal{R} v_i\} + y_N^\intercal P y_N \\
&\text{s.t.} && \begin{cases} \forall i \in \{0, 1, \ldots, N-1\}, \\ y_{i+1} = \mathcal{A} y_i + \mathcal{B} v_i \\ y_0 = x_0 \ , \end{cases}
\end{aligned} \tag{11}$$

can be solved by a so-called *dynamic programming recursion*. This means that we implement the so-called backward *Riccati recursion*

$$P_i \quad = \quad \mathcal{Q} + \mathcal{A}^\intercal P_{i+1}\mathcal{A} - \mathcal{A}^\intercal P_{i+1}\mathcal{B}\left[\mathcal{R} + \mathcal{B}P_{i+1}\mathcal{B}\right]^{-1}\mathcal{B}^\intercal P_{i+1}\mathcal{A} \tag{12}$$

for $i = N, N-1, \ldots, 0$, started at $P_N = P$. In order to find the matrices $P_0, P_1, \ldots, P_N$, which are the weights of our cost to go functions $J_i(z_i) = z_i^\intercal P_i z_i$. Moreover, the optimal discrete-time control inputs satisfy an equation of the form

$$v_i^\star = \mathcal{K}_i y_i \qquad \text{with} \qquad \mathcal{K}_i \ = \ -\left[\mathcal{R} + \mathcal{B}P_{i+1}\mathcal{B}\right]^{-1}\mathcal{B}^\intercal P_{i+1}\mathcal{A} \ .$$

In this form, this equation for $v_i^\star$ can be interpreted as a linear feedback control law, since $v_i^\star$—our current control input—is now a function of the discrete time state $y_i = z_i$. The corresponding optimal discrete-time state trajectory can be found be a forward simulation of the closed loop system

$$\forall i \in \{0, 1, \ldots, N-1\}, \qquad y_{i+1} \ = \ \mathcal{A} y_i + \mathcal{B}\mathcal{K}_i y_i \ = \ [\mathcal{A} + \mathcal{B}\mathcal{K}]\, y_i \tag{13}$$

$$y_0 \ = \ x_0 \ . \tag{14}$$

In summary, this means that we can solve the discrete-time optimal control problem by a backward recursion for the matrices P_i and a forward recursion for the state trajectory.

1.3 Continuous-Time Optimal Control

In order to also finally solve the continuous-time optimal control problem

$$\begin{aligned}
&\min_{x,u} \quad && \int_0^T x(t)^\intercal Q x(t) + u(t)^\intercal R u(t)\, \mathrm{d}t + x(T)^\intercal P x(T) \\
&\text{s.t.} && \begin{cases} \forall t \in [0, T], \\ \dot{x}(t) = A x(t) + B u(t) \\ x(0) = x_0 \ . \end{cases}
\end{aligned} \tag{15}$$

we still have to compute the limit for $h \to 0$. In order to compute this limit, we recall that

$$\mathcal{Q} = hQ, \quad \mathcal{R} = hR, \quad \mathcal{A} = I + hA, \quad \text{and} \quad \mathcal{B} = hB \; .$$

Thus, if we want to take the limit for $h \to 0$ of our Riccati backward recursion, we need to write this recursion in the form

$$\begin{aligned} P_i &= \mathcal{Q} + \mathcal{A}^\intercal P_{i+1}\mathcal{A} - \mathcal{A}^\intercal P_{i+1}\mathcal{B}\left[\mathcal{R} + \mathcal{B}P_{i+1}\mathcal{B}\right]^{-1}\mathcal{B}^\intercal P_{i+1}\mathcal{A} \\ &= hQ + (I + hA)^\intercal P_{i+1}(I + hA) \\ &\quad -(I + hA)^\intercal P_{i+1}h^2B\left[hR + h^2BP_{i+1}B\right]^{-1}B^\intercal P_{i+1}(I + hA) \\ &= hQ + P_{i+1} + hA^\intercal P_{i+1} + hP_{i+1}A - hP_{i+1}BR^{-1}B^\intercal P_{i+1} + O(h^2) \; . \end{aligned}$$

We can rewrite this equation in the form

$$\frac{P_i - P_{i+1}}{h} = Q + A^\intercal P_{i+1} + P_{i+1}A - P_{i+1}BR^{-1}B^\intercal P_{i+1} + O(h)$$

Next, we subsitute $P_i = \mathcal{P}(t) + O(h)$ as well as $\mathcal{P}_{i+1} = \mathcal{P}(t + h) + O(h)$ and take the limit for $h \to 0$. This yields

$$-\dot{\mathcal{P}}(t) = Q + A^\intercal\mathcal{P}(t) + \mathcal{P}(t)A - \mathcal{P}(t)BR^{-1}B^\intercal\mathcal{P}(t) \qquad (16)$$

$$\mathcal{P}(T) = P \; . \qquad (17)$$

(sorry for the inconsisten notation—on the slides P and $\mathcal{P}$ are interchanged). Similarly, we can work out the corresponding optimal continuous-time feedback gain $K(t) = \mathcal{K}_i + O(h)$ finding that

$$\begin{aligned} K(t) &= \lim_{h\to 0}\left\{-h\left[hR + h^2BP_{i+1}B\right]^{-1}B^\intercal\mathcal{P}(t)(I + hA)\right\} \\ &= -R^{-1}B^\intercal\mathcal{P}(t) \; . \qquad (18) \end{aligned}$$

This is called the continuous-time optimal LQR feedback gain. In summary, we can solve the continuous-time linear-quadratic optimal control problem by simulating the differential Riccati equation for $\mathcal{P}(t)$ backwards in time, computing the optimal control gain $K(t)$ by the above equation and simulating the closed-loop system

$$\begin{aligned} \forall t \in [0, T], \qquad \dot{x}(t) &= [A + BK(t)]x(t) \\ x(0) &= x_0 \qquad (19) \end{aligned}$$

forward in time in order to find the optimal state trajectory.