Lecture 34

Linear-Quadratic Optimal Control

In this lecture we introduce the optimal state-feedback and the linear-quadratic regulator (LQR) problem. Utilizing dynamic programming necessary optimality conditions are derived. It turns out that for the LQR problem the state-feedback solution can be determined by solving a differential matrix Riccati equation. The presented theory is taken from the book [12].

34.1 The problem formulation

The goal is to find a state-feedback control law of the form

$$u(t) = -Kx(t) \quad \text{for } t \in [0,T]$$

with $u : [0,T] \to \mathbb{R}^{m_u}$, $x : [0,T] \to \mathbb{R}^{m_x}$, $K \in \mathbb{R}^{m_u \times m_x}$ so that u minimizes the quadratic cost functional

$$J(x,u) = \int_0^T x(t)^\top Qx(t) + u(t)^\top Ru(t)\,\mathrm{d}t + x(T)^\top Mx(T), \tag{34.1a}$$

where the state x and the control u are related by the linear initial value problem

$$\dot{x}(t) = Ax(t) + Bu(t) \text{ for } t \in (0,T] \quad \text{and} \quad x(0) = x_0. \tag{34.1b}$$

In (34.1a) the matrices Q, $M \in \mathbb{R}^{m_x \times m_x}$ are symmetric, positive semi-definite, $R \in \mathbb{R}^{m_u \times m_u}$ is symmetric, positive definite and in (34.1b) we have $A \in \mathbb{R}^{m_x \times m_x}$, $B \in \mathbb{R}^{m_x \times m_u}$ and $x_0 \in \mathbb{R}^{m_x}$. The final time T is fixed, but the final state $x(T)$ is free. Thus, we aim to track the state to the state $\bar{x} = 0$ as good as possible. The terms $x(t)^T Qx(t)$ and $x(T)^T Mx(T)$ are measures for the control accuracy and the term $u(t)^T Ru(t)$ measures the control effort. Problem (34.1) is called the *linear-quadratic regulator problem* (*LQR problem*).

34.2 The Hamilton-Jacobi-Bellman equation

In this section we derive first-order necessary optimality conditions for the LQR problem. Since generalizing the problem to a non-linear problem does not cause more difficulties in the deviation, we consider the problem to find a state-control feedback control law

$$u(t) = \Phi(x(t),t), \quad t \in [0,T],$$

such that the cost-functional

$$J_t(x,u) = \int_t^T \ell(x(s),u(s),s)\,\mathrm{d}s + g(x(T)) \tag{34.2a}$$

is minimized subject to the non-linear system dynamics

$$\dot{x}(s) = f(x(s),u(s),s) \text{ for } s \in (0,T] \quad \text{and} \quad x(t) = x_t. \tag{34.2b}$$

We suppose that the functions $\ell : \mathbb{R}^{m_x} \times \mathbb{R}^{m_u} \times [0,T] \to [0,\infty)$ and $g : \mathbb{R}^{m_x} \to [0,\infty)$ satisfy

$$\ell(0,0,s) = 0 \text{ for } s \in [0,T] \quad \text{and} \quad g(0) = 0$$

Moreover, let $f : \mathbb{R}^{m_x} \times \mathbb{R}^{m_u} \times [0,T] \to \mathbb{R}^{m_x}$ be continuous and locally Lipschitz-continuous with respect to the variable x. Moreover, $x_t \in \mathbb{R}^{m_x}$ holds. To derive optimality conditions we use the so-called *Bellman principle* (or *dynamic programming principle*). The essential assumption is that the system can be characterized by its state $x(t)$ at the time $t \in [0,T]$ which completely summarizes the effect of all $u(s)$ for $0 \le s \le t$. The dynamic programming principle was first proposed by Bellman [5].

Theorem 34.2.1 (Bellman principle)**.** *Let $t \in [0,T]$. If $u^*(s)$ is optimal for $s \in [t,T]$ and x^* is the associated optimal state, starting at the state $x_t \in \mathbb{R}^{m_x}$, then $u^*(s)$ is also optimal over the subinterval $[t+\Delta t, T]$ for any $\Delta t \in [0, T-t]$ starting at $x_{t+\Delta t} = x^*(t+\Delta t)$.*

Proof. We show Theorem 34.2.1 by contradiction. Suppose that there exists a control u^{**} so that

$$\begin{aligned}&\int_{t+\Delta t}^{T} \ell(x^{**}(s), u^{**}(s), s)\,\mathrm{d}s + g(x^{**}(T))\\ &\qquad < \int_{t+\Delta t}^{T} \ell(x^{*}(s), u^{*}(s), s)\,\mathrm{d}s + g(x^{*}(T)),\end{aligned} \tag{34.3}$$

where

$$\dot{x}^*(s) = f(x^*(s), u^*(s), s) \quad \text{and} \quad \dot{x}^{**}(s) = f(x^{**}(s), u^{**}(s), s)$$

hold for $s \in [t+\Delta t, T]$. We define the control

$$u(s) = \begin{cases} u^*(s) & \text{if } s \in [t, t+\Delta t],\\ u^{**}(s) & \text{if } s \in (t+\Delta t, T]. \end{cases} \tag{34.4}$$

By $x(s)$ we denote the state satisfying $\dot{x}(s) = F(x(s), u(s), s)$ for $s \in [t,T]$ and $x(t) = x_t$. Then we derive from (34.3) and (34.4) that

$$\begin{aligned}&\int_t^T \ell(x(s), u(s), s)\,\mathrm{d}s + g(x(T))\\ &= \int_t^{t+\Delta t} \ell(x^*(s), u^*(s), s)\,\mathrm{d}s + \int_{t+\Delta t}^{T} \ell(x^{**}(s), u^{**}(s), s)\,\mathrm{d}s + g(x^{**}(T))\\ &< \int_t^{t+\Delta t} \ell(x^*(s), u^*(s), s)\,\mathrm{d}s + \int_{t+\Delta t}^{T} \ell(x^{*}(s), u^{*}(s), s)\,\mathrm{d}s + g(x^{*}(T))\\ &= \int_t^T \ell(x^*(s), u^*(s), s)\,\mathrm{d}s + g(x^*(T)).\end{aligned} \tag{34.5}$$

Recall that $u^*(s)$ is optimal for $s \in [t,T]$ by assumption. From (34.5) it follows that the control u given by (34.4) yields a smaller value of the cost functional. This is a contradiction. ☐

Next we derive the Hamilton-Jacobi-Bellman equation for (34.2). Let $V^* : \mathbb{R}^{m_x} \times [0,T] \to \mathbb{R}$ denote the minimal value function given by

$$\begin{aligned}&V^*(x_t, t)\\ &= \min_{u:[t,T]\to\mathbb{R}^{m_u}} \Big\{ J_t(x,u) \,\Big|\, \dot{x}(s) = f(x(s), u(s), s),\ s \in (t,T] \text{ and } x(t) = x_t \Big\}\end{aligned} \tag{34.6}$$

for $(x_t, t) \in \mathbb{R}^{m_x} \times [0,T]$, where

$$J_t(x,u) = \int_t^T \ell(x(s), u(s), s)\,\mathrm{d}s + g(x(T)).$$

From the linearity of the integral and (34.6) we conclude

$$\begin{aligned}&V^*(x_t, t)\\ &= \min_{u:[t,t+\Delta t]\to\mathbb{R}^{m_u}} \bigg\{ \int_t^{t+\Delta t} \ell(x(s), u(s), s)\,\mathrm{d}s + V^*(x(t+\Delta t), t+\Delta t) \,\Big|\\ &\qquad\qquad \dot{x}(s) = f(x(s), u(s), s),\ s \in (t, t+\Delta t] \text{ and } x(t) = x_t \bigg\}\end{aligned} \tag{34.7}$$

for $(x_t, t) \in \mathbb{R}^{m_x} \times [0, T - \Delta t]$, where we have used the Bellman principle. Thus, by using the Bellman principle the problem of finding an optimal control over the interval $[t, T]$ has been reduced to the problem of finding an optimal control over the interval $[t, t + \Delta t]$.

Now we replace the integral in (34.7) by $\ell(x(t), u(t), t)\Delta t$, perform a Taylor approximation for $V^*(x(t+\Delta t), t+\Delta t)$ about the point $(x_t, t) = (x(t), t)$ and approximate$x(t + \Delta t) - x(t)$ by $f(x(t), u(t), t)\Delta t$. Then we find

$$\begin{aligned} V^*(x_t, t) &= \min_{u_t \in \mathbb{R}^{m_u}} \Big\{ \ell(x_t, u_t, t)\Delta t + V^*(x_t, t) + \frac{\partial V^*}{\partial t}(x_t, t)\, \Delta t \\ &\qquad + \nabla V^*(x_t, t)^\top f(x_t, u_t, t)\Delta t + \mathcal{O}(\Delta t) \Big\} \\ &= V^*(x_t, t) + \frac{\partial V^*}{\partial t}(x_t, t)\Delta t \\ &\quad + \Delta t \min_{u_t \in \mathbb{R}^{m_u}} \left\{ \ell(x_t, u_t, t) + \nabla V^*(x_t, t)^\top f(x_t, u_t, t) + \frac{\mathcal{O}(\Delta t)}{\Delta t} \right\} \end{aligned}$$

for any $\Delta t > 0$. Thus,

$$-\frac{\partial V^*}{\partial t}(x_t, t) = \min_{u_t \in \mathbb{R}^{m_u}} \left\{ \ell(x_t, u_t, t) + \nabla V^*(x_t, t)^\top f(x_t, u_t, t) + \frac{\mathcal{O}(\Delta t)}{\Delta t} \right\}.$$

Taking the limit $\Delta t \to 0$ and using $V^*(x_t, T) = g(x_t)$ we obtain

$$-\frac{\partial V^*}{\partial t}\,(x_t, t) = \min_{u_t \in \mathbb{R}^{m_u}} \left\{ \ell(x_t, u_t, t) + \nabla V^*(x_t, t)^\top f(x_t, u_t, t) \right\} \tag{34.8a}$$

for all $(x_t, t) \in \mathbb{R}^{m_x} \times [0, T)$ and

$$V^*(x_t, T) = g(x_t) \tag{34.8b}$$

for all $x_t \in \mathbb{R}^{m_x}$. System (34.8) is called the *Hamilton-Jacobi-Bellman (HJB) equations.*

To solve (34.8) we proceed in two steps. First we compute a solution u_t to

$$u^*(t) = \operatorname*{argmin}_{u_t \in \mathbb{R}^{m_u}} \left\{ \ell(x_t, u_t, t) + \nabla V^*(x_t, t)^\top f(x_t, u_t, t) \right\}$$

and set

$$\Psi(\nabla V^*(x_t, t), x_t, t) = u^*(t), \tag{34.9}$$

which gives us a control law. Then we insert (34.9) into (34.8a) and solve

$$-\frac{\partial V^*}{\partial t}\,(x_t, t) = \ell(x_t, \Psi(\nabla V^*(x_t, t), x_t, t), t) + \nabla V^*(x_t, t)^\top f(x_t, \Psi(\nabla V^*(x_t, t), x_t, t), t)$$

for all $(x_t, t) \in \mathbb{R}^{m_x} \times [0, T)$. Finally, we can compute the gradient $\nabla V^*(x_t, t)$ and deduce the state-feedback law

$$u^*(t; x_t) = \Phi(x_t, t) = \Psi(\nabla V^*(x_t, t), x_t, t) \quad \text{for all } (x_t, t) \in \mathbb{R}^{m_x} \times [0, T).$$

Remark 34.2.2. 1) In general, it is not possible to solve (34.8) analytically. However, for the LQR problem we can derive an explicit solution for the state-feedback law.

2) Note that the HJB equation are only necessary optimality conditions.

◊

34.3 The state-feedback law for the linear quadratic problem

For the LQR problem we have

$$\ell(x_t, u_t, t) = x_t^\top Q x_t + u_t^\top R u_t, \quad g(x_t) = x_t^\top M x_t, \quad f(x_t, u_t, t) = A x_t + B u_t$$

for $(x_t, u, t) \in \mathbb{R}^{m_x} \times \mathbb{R}^{m_u} \times [0, T]$. For brevity, we focus on the situation, where the matrices A, B, Q, M, R are time-invariant. However, most of the presented theory also holds for the time-varying case.

First we minimize

$$x_t^\top Q x_t + u_t^\top R u_t + \nabla V^*(x_t,t)^\top \big(Ax_t + Bu_t\big)$$

with respect to u_t. First-order necessary optimality conditions are given by

$$u_t^\top R\tilde u_t + \tilde u_t^\top R u_t + \nabla V^*(x_t,t)^\top B\tilde u_t = 0 \quad \text{for all } \tilde u_t \in \mathbb{R}^{m_u} \text{ and } (x_t,t) \in \mathbb{R}^{m_x} \times [0,T).$$

By assumption, R is symmetric and positive definite. Then we find

$$\big(2Ru_t + B^\top \nabla V^*(x_t,t)\big)^\top \tilde u_t = 0 \quad \text{for all } \tilde u_t \in \mathbb{R}^{m_u} \text{ and } (x_t,t) \in \mathbb{R}^{m_x} \times [0,T)$$

and

$$\bar u_t = -\frac{1}{2} R^{-1} B^T \nabla V^*(x_t,t) \quad \text{for all } (x_t,t) \in \mathbb{R}^{m_x} \times [0,T). \tag{34.10}$$

For the minimal value function V^* we make the quadratic ansatz

$$V^*(x_t,t) = x_t^\top P(t) x_t \text{ for } (x_t,t) \in \mathbb{R}^{m_x} \times [0,T), \quad P(t) \in \mathbb{R}^{m_x \times m_x} \text{ symmetric.} \tag{34.11}$$

Then, we have $\nabla V^*(x_t,t) = 2P(t)x$ so that

$$\bar u_t = -R^{-1} B^\top P(t) x_t \quad \text{for all } (x_t,t) \in \mathbb{R}^{m_x} \times [0,T).$$

Note that for all $(x_t,t) \in \mathbb{R}^{m_x} \times [0,T)$

$$\begin{aligned}
\frac{\partial V^*}{\partial t}(x_t,t) &= x_t^\top \dot P(t) x_t,\\
\ell(x_t, -R^{-1}B^\top P(t)x_t, t) &= x_t^\top Q x_t + x_t^\top P(t) B R^{-1} B^\top P(t) x_t\\
&= x_t^\top \big(Q + P(t)BR^{-1}B^\top P(t)\big) x_t,\\
f(x_t, -R^{-1}B^\top P(t)x_t, t) &= Ax_t - BR^{-1}B^\top P(t)x_t = \big(A - BR^{-1}B^\top P(t)\big)x_t,\\
\nabla V^*(x_t,t) &= 2P(t)x_t.
\end{aligned}$$

Consequently,

$$\begin{aligned}
-\, x_t^\top \dot P(t) x_t &= -\frac{\partial V^*}{\partial t}(x_t,t)\\
&= x_t^\top \big(Q + P(t)BR^{-1}B^\top P(t)\big)x_t + \big(2P(t)x_t\big)^\top \big(A - BR^{-1}B^\top P(t)\big)x_t
\end{aligned}$$

for all $(x_t,t) \in \mathbb{R}^{m_x} \times [0,T)$, which yields

$$\begin{aligned}
&-\, x_t^\top \dot P(t) x_t\\
&\quad = x_t^\top \big(Q + P(t)BR^{-1}B^\top P(t) + 2P(t)A - 2P(t)BR^{-1}B^\top P(t)\big)x_t\\
&\quad = x_t^\top \big(2P(t)A + Q - P(t)BR^{-1}B^\top P(t)\big)x_t
\end{aligned}$$

for all $(x_t,t) \in \mathbb{R}^{m_x} \times [0,T)$. From $P(t) = P(t)^\top$ we deduce that

$$2x_t^\top P(t)Ax_t = x_t^\top P(t)Ax_t + x_t^\top A^T P(t)x_t = x_t^\top \big(A^\top P(t) + P(t)A\big)x_t.$$

Using $V^*(x_t,T) = x_t^T P(T) x_t$ and (34.8b) we get

$$-x_t^\top \dot P(t) x_t = x_t^\top \big(A^\top P(t) + P(t)A + Q - P(t)BR^{-1}B^\top P(t)\big)x_t, \quad t \in [0,T), \tag{34.12a}$$

$$x_t^\top P(T) x_t = x_t^\top M x_t. \tag{34.12b}$$

Since (34.12) holds for all $x_t \in \mathbb{R}^{m_x}$ we obtain the following *matrix Riccati equation*

$$-\dot P(t) = A^\top P(t) + P(t)A + Q - P(t)BR^{-1}B^\top P(t), \quad t \in [0,T), \tag{34.13a}$$

$$P(T) = M. \tag{34.13b}$$

Finally, the optimal state-feedback is given by

$$\bar u(t) = -K(t)x(t) \quad \text{and} \quad K(t) = R^{-1}B^\top P(t) \quad \text{for all } t \in [0,T).$$

Example 34.3.1. Let us consider the problem

$$\min \int_0^T |x(t)|^2 + |u(t)|^2 \,\mathrm{d}t \quad \text{s.t.} \quad \dot{x}(t) = u(t) \text{ for } t \in (0,T].$$

Choosing $m_x = m_u = 1$, $A = M = 0$ and $B = Q = R = 1$ the matrix Riccati equation has the form

$$-\dot{P}(t) = 1 - P(t)^2 \text{ for } t \in [0,T) \quad \text{and} \quad P(T) = 0.$$

This scalar ordinary differential equation can be solved by separation of variables. Its solution is

$$P(t) = \frac{1 - e^{-2(T-t)}}{1 + e^{-2(T-t)}} \quad \text{for } t \in [0,T)$$

with the optimal control $\bar{u}(t) = -P(t)x(t)$. ◊