TSRT08: Optimal Control
Solutions

20240110

1. (a) Define $x = y, u = \dot{y}$ and $f(t, x, u) = x^2 + u^2 - 2x \sin t$, then the Hamiltonian is given by

$$H(t, x, u, \lambda) = x^2 + u^2 - 2x \sin t + \lambda u$$

The following equations must hold

$$y(0) = 0 \tag{1.1a}$$
$$0 = \frac{\partial H}{\partial u}(t, x, u, \lambda) = 2u + \lambda \tag{1.1b}$$
$$\dot{\lambda} = -\frac{\partial H}{\partial x}(t, x, u, \lambda) = -2x + 2 \sin t \tag{1.1c}$$
$$\lambda(1) = 0 \tag{1.1d}$$

Equation (1.1b) gives $\dot{\lambda} = -2\dot{u} = -2\ddot{y}$ and with (1.1c) we have

$$\ddot{y} - y = -\sin t$$

with the solution

$$y(t) = c_1 e^t + c_2 e^{-t} + \frac{1}{2} \sin t, \quad c_1, c_2 \in \mathbb{R}$$

where (1.1a) gives the relationship of c_1 and c_2 as

$$c_2 = -c_1$$

also from (1.1b) and (1.1d), $u(1) = 0$, and since $\dot{y}(t) = u(t)$, therefore $\dot{y}(1) = 0$ which gives that

$$\dot{y}(1) = c_1 e - c_2 e^{-1} + \frac{1}{2} \cos 1 \Rightarrow c_2 = c_1 e^2 + \frac{e}{2} \cos 1$$

Consequently, we have

$$y(t) = -\frac{e \cos 1}{2\left(e^2 + 1\right)} \left(e^t - e^{-t}\right) + \frac{1}{2} \sin t$$

(b) Define $x = y, u = \dot{y}$ and $f(t, x, u) = u^2/t^3$, then the Hamiltonian is given by

$$H(t, x, u, \lambda) = u^2 t^{-3} + \lambda u$$

The following equations must hold

$$y(0) = 0 \tag{1.2a}$$
$$0 = \frac{\partial H}{\partial u}(t, x, u, \lambda) = 2ut^{-3} + \lambda \tag{1.2b}$$
$$\dot{\lambda} = -\frac{\partial H}{\partial x}(t, x, u, \lambda) = 0 \tag{1.2c}$$
$$\lambda(1) = 0 \tag{1.2d}$$

Equation (1.2c) gives $\lambda = c_1, c_1 \in \mathbb{R}$, and (1.2d) that $c_1 = 0$. Then (1.2b) gives $u(t) = 0$ which implies $y(t) = c_2, c_2 \in \mathbb{R}$. Finally, with (1.2a) this gives

$$y(t) = 0$$

(c) Analogous to (a). The system equation is

$$\ddot{y} - y = e^t$$

with the solution

$$y(t) = c_1 e^t + c_2 e^{-t} + \frac{1}{2} t e^t$$

where

$$c_2 = -c_1$$

and where

$$c_2 = e^2 \left(c_1 + 1\right)$$

which gives

$$y(t) = \frac{-e^2}{e^2 + 1} \left(e^t - e^{-t}\right) + \frac{1}{2} t e^t$$

2. (a) The discrete optimization problem is given by

$$\begin{array}{ll} \underset{u_0,u_1}{\text{minimize}} & r\left(x_2-T\right)^2+\sum_{k=0}^{1} u_k^2 \\ \text{subject to} & x_{k+1}=(1-a)x_k+au_k, k=0,1, \\ & x_0 \text{ given}, \\ & u_k \in \mathbb{R}, k=0,1. \end{array}$$

(b) With $a = 1/2, T = 0$ and $r = 1$ this can be cast on standard form with $N = 2, \phi(x) = x^2, f(k,x,u) = u^2$, and $F(k,x,u) = \frac{1}{2}x + \frac{1}{2}u$. The DP algorithm gives us:

Stage $k = N = 2$

$$V_2(x) = x^2$$

Stage $k = 1$:

$$\begin{aligned} V_1(x) &= \min_u \left\{u^2 + V_2(F(1,x,u))\right\} \\ &= \min_u \left\{u^2 + \left(\frac{1}{2}x + \frac{1}{2}u\right)^2\right\} \end{aligned} \tag{2.1}$$

The minimization is done by setting the derivative, w.r.t. u, to zero, since the function is strictly convex in u. Thus, we have

$$2u + \left(\frac{1}{2}x + \frac{1}{2}u\right) = 0$$

which gives the control function

$$u_1^* = \mu_1(x) = -\frac{1}{5}x$$

Note that we now have computed the optimal control for each possible state x. By substituting the optimal u_1^* into (2.1) we obtain

$$V_1(x) = \frac{1}{5}x^2 \tag{2.2}$$

Stage $k = 0$:

$$\begin{aligned} V_0(x) &= \min_u \left\{u^2 + V_1(F(0,x,u)\right\} \\ &= \min_u \left\{u^2 + V_1\left(\frac{1}{2}x + \frac{1}{2}u\right)\right\} \end{aligned}$$

Substituting $V_1(x)$ by (2.2) and minimizing by setting the derivative, w.r.t. u, to zero gives (after some calculations) the optimal control

$$u_0^* = \mu_0(x) = -\frac{1}{21}x \tag{2.3}$$

and the optimal cost

$$V_0(x) = \frac{1}{21}x^2 \tag{2.4}$$

(c) This can be cast on standard form with $N = 2, \phi(x) = r(x-T)^2$, $f(k,x,u) = u^2$, and $F(k,x,u) = (1-a)x + au$. The DP algorithm gives us:

Stage $k = N = 2$:

$$V_2(x) = r(x-T)^2$$

Stage $k = 1$:

$$\begin{aligned} V_1(x) &= \min_u \left\{u^2 + V_2(F(1,x,u))\right\} \\ &= \min_u \left\{u^2 + r((1-a)x + au - T)^2\right\} \end{aligned} \tag{2.5}$$

The minimization is done by setting the derivative, w.r.t. u, to zero, since the function is strictly convex in u. Thus, we have

$$2u + 2ra((1-a)x + au - T) = 0$$

which gives the control function

$$u_1^* = \mu_1(x) = \frac{ra(T-(1-a)x)}{1+ra^2}$$

Note that we now have computed the optimal control for each possible state x. By substituting the optimal u_1^* into (2.5) we obtain (after some work)

$$V_1(x) = \frac{r((1-a)x-T)^2}{1+ra^2} \tag{2.6}$$

Stage $k = 0$:

$$\begin{aligned} V_0(x) &= \min_u \left\{u^2 + V_1(F(0,x,u)\right\} \\ &= \min_u \left\{u^2 + V_1((1-a)x + au)\right\} \end{aligned}$$

Substituting $V_1(x)$ by (2.6) and minimizing by setting the derivative, w.r.t. u, to zero gives (after some calculations) the optimal control

$$u_0^* = \mu_0(x) = \frac{r(1-a)a\left(T-(1-a)^2x\right)}{1+ra^2\left(1+(1-a)^2\right)} \tag{2.7}$$

and the optimal cost

$$V_0(x) = \frac{r\left((1-a)^2x-T\right)^2}{1+ra^2\left(1+(1-a)^2\right)} \tag{2.8}$$

3. (a) The Q-function makes it possible to solve optimal control problems without knowing an explicit model of the system to control. This is possible since the Q-function depends on both the state and the control signal.

   (b) Instead of using reinforcement learning one can first perform system identification using experiments to obtain a model of the system to control, and then traditional optimal control can be used based on the model obtained. Also adaptive control can be used.

   (c) The key assumption is that a repeated task should be carried out, typically following the same trajectory over and over again.

4. The problem can be rewritten to standard form as

$$\begin{array}{ll} \underset{u(\cdot)}{\text{minimize}} & -x\,(T) - \int_0^T (1-u(t))x(t)dt \\ \text{subject to} & \dot{x}(t) = \alpha u(t)x(t), 0 < \alpha < 1 \\ & x(0) = x_0 > 0, \\ & 0 \le u(t) \le 1, \forall t \in [0,T]\,. \end{array} \tag{4.1}$$

The Hamiltonian is given by

$$H(t,x,u,\lambda) \triangleq f(t,x,u) + \lambda^T F(t,x,u) = -(1-u)x + \lambda\alpha ux$$

Pointwise optimization of the Hamiltonian yields

$$\begin{aligned} \tilde{\mu}(t,x,\lambda) &\triangleq \underset{u\in[0,1]}{\arg\min} H(t,x,u,\lambda) \\ &= \underset{u\in[0,1]}{\arg\min}\{-(1-u)x + \lambda\alpha ux\} \\ &= \underset{u\in[0,1]}{\arg\min}\{(1+\lambda\alpha)ux\} \\ &= \begin{cases} 1, & (1+\lambda\alpha)x < 0 \\ 0, & (1+\lambda\alpha)x > 0 \\ \tilde{u}, & (1+\lambda\alpha)x = 0 \end{cases}, \end{aligned}$$

where $\tilde{u}$ is an arbitrary value in $[0,1]$. To be able to find an analytical solution, it is important to remove the variable x from the pointwise optimal solution above. Otherwise we are going to end up with a PDE, which is difficult to solve. In this particular case it is simple. Since $x_0 > 0, \alpha > 0$ and $u > 0$ in (4.1), it follows that $x(t) > 0$ for all $t \in [0,T]$. Hence, the optimal control is given by

$$u^*(t) \triangleq \tilde{\mu}(t,x(t),\lambda(t)) = \begin{cases} 1, & (1+\lambda\alpha) < 0 \\ 0, & (1+\lambda\alpha) > 0 \\ \tilde{u}, & (1+\lambda\alpha) = 0 \end{cases}$$

and we define the switching function as

$$\sigma(t) \triangleq 1 + \lambda(t)\alpha$$

The adjoint equations are now given by

$$\begin{aligned} \dot{\lambda}(t) &\triangleq -\frac{\partial H}{\partial x}(t,x(t),u^*(t),\lambda(t)) \\ &= (1-u^*(t)) - \alpha\lambda(t)u^*(t) \\ &= \begin{cases} -\alpha\lambda(t), & (1+\lambda\alpha) < 0, (u^*(t) = 1) \\ 1, & (1+\lambda\alpha) > 0, (u^*(t) = 0) \\ 1, & (1+\lambda\alpha) = 0 \end{cases} \end{aligned} \tag{4.2}$$

The boundary constraints are

$$\lambda(T) - \frac{\partial \phi}{\partial x}(T, x(T)) \perp S_T(T) \quad \Longleftrightarrow \quad \lambda(T) + 1 = 0$$

which implies that $\lambda(T) = -1$. Thus

$$\sigma(T) = 1 + \lambda(T)\alpha = 1 - \alpha > 0$$

so that $u^*(T) = 0$. What remains to be determined is how many switches occurs. A hint of the number of switches can often be found by considering the value of $\dot{\sigma}(t)|_{\sigma(t)=0}$. From (4.2), it follows that

$$\dot{\sigma}(t)\big|_{\sigma(t)=0} = \dot{\lambda}(t)\alpha\big|_{1+\lambda(t)\alpha=0} = \alpha > 0$$

Hence, there can only be at most one switch, since we can pass $\sigma(t) = 0$ only once. Since $u^*(T) = 0$, it is not possible that $u^*(t) = 1$ for all $t \in [0, T]$. Thus,

$$u^*(t) = \begin{cases} 1, & 0 \le t \le t' \\ 0, & t' \le t \le T \end{cases}$$

for some unknown switching time $t' \in [0, T]$ The switching occurs when

$$0 = \sigma(t') = 1 + \lambda(t')\alpha$$

and to find the value of t' we need to determine the value of $\lambda(t')$. From (4.2), it holds that during the period $t' \le t \le T$ where $u^*(t) = 0$ we have that

$$\dot{\lambda}(t) = 1, \lambda(T) = -1$$

which has the solution $\lambda(t) = t - T - 1$. Since $\sigma(t') = 0$ is equivalent to $\lambda(t') = -1/\alpha$, we get

$$t' = T + 1 - \frac{1}{\alpha}$$

and the optimal control is thus given by

$$u^*(t) = \begin{cases} 1, & 0 \le t \le T + 1 - \frac{1}{\alpha} \\ 0, & T + 1 - \frac{1}{\alpha} \le t \le T \end{cases}$$

It is worth noting that when α is small, t' will become negative and the optimal control law is $u^*(t) = 0$ for all $t \in [0, T]$.