TSRT08: Optimal Control
Solutions

20230315

1. (a) The Hamiltonian is given by

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

Since the Hamiltonian is strictly convex in u pointwise minimization w.r.t u yields

$$0 = \frac{\partial H}{\partial u}(t,x,u,\lambda) = \frac{2u}{x} - \lambda \quad \Rightarrow \quad u^* = \frac{1}{2}\lambda x.$$

The adjoint equation is given by

$$\dot{\lambda}(t) = -\frac{\partial H}{\partial x}(t,x,u,\lambda) = \frac{u^2}{x^2} = \frac{1}{4}\lambda^2,$$

which is a separable ODE in λ. We have

$$\frac{d\lambda}{dt} = \frac{1}{4}\lambda^2 \quad \Leftrightarrow \quad \frac{1}{\lambda^2}d\lambda = \frac{1}{4}dt \quad \Leftrightarrow \quad -\frac{1}{\lambda} = \frac{1}{4}t + c \quad \Leftrightarrow \quad \lambda = -\frac{4}{t+4c}$$

for some constant c. Since $\lambda(T) = \frac{\partial \phi}{\partial x} = 1$, we have $c = -1 - T/4$ and thus $\lambda = -4/(t-T-4)$, which yields the optimal control signal

$$u^* = \frac{1}{2}\lambda x = -\frac{2}{t-T-4}x$$

(b) Introducing $x = y$ and $u = \dot{y}$, it holds that $\dot{x} = u$ and the Hamiltonian is given by

$$H(t,x,u,\lambda) = e^t x + u^2 t^{-1} + \lambda u.$$

The following equations must hold

$$\dot{\lambda} = -\frac{\partial H}{\partial x}(t,x,u,\lambda) = -e^t, \tag{1a}$$

$$0 = \frac{\partial H}{\partial u}(t,x,u,\lambda) = 2ut^{-1} + \lambda. \tag{1b}$$

Now, (1a) yields $\lambda = -e^t + c_1$ and this inserted into (1b) yields

$$0 = 2ut^{-1} + \lambda = 2\dot{y}t^{-1} + c_1 - e^t \quad \Leftrightarrow \quad y = \frac{1}{2}\int t(e^t - c_1)\,dt = \frac{1}{2}(t-1)e^t - \frac{c_1}{4}t^2 + c_2.$$

Finally, $y(0) = 1$, $y(1) = 0$ gives the constants $c_1 = 6$ and $c_2 = 3/2$. Hence

$$y = \frac{1}{2}\big((t-1)e^t - 3t^2 + 3\big),$$

is the sought after extrema.

2. (a) The problem gives the following Hamilton-Jacobi-equation

$$0 = \min_u (V_x u + u^2 + x^{2m}). \tag{2}$$

Minimizing with respect to u gives $u = -V_x/2$ and if you plug that into (2) one gets

$$0 = -\frac{V_x^2}{4} + x^{2m} \Rightarrow V_x = \pm 2|x|^m.$$

Since V shall be a value function we require $V(x) > 0$ for $x \neq 0$ and $V(0) = 0$. This means that the sign of V_x shall be

$$V_x = \begin{cases} 2|x|^m & x > 0, \\ -2|x|^m & x < 0, \end{cases}$$

which gives $V(x) = 2|x|^{m+1}/(m+1)$. The optimal feedback is then

$$u = \begin{cases} -|x|^m & x > 0, \\ |x|^m & x < 0. \end{cases}$$

(b) The dynamic programming recursion is given by

$$\begin{aligned} J(N, x) &= \phi(x) \\ J(n, x) &= \min_u \{f_0(n, x, u) + J(n+1, f(n, x, u))\}, \end{aligned}$$

which is our case will be

$$\begin{aligned} J(N, x) &= x^{2m} \\ J(n, x) &= \min_u \{x^{2m} + u^2 + J(n+1, x+u)\}, \end{aligned}$$

For the case $m = 1$ this is a LQ problem (with finite time horizon). This can be solved with a cost-to-go function which is quadratic in x, this means with the ansatz $J(n, x) = \alpha(n)x^2$. However, with $m > 1$, the polynomial order of $J(n, x)$ will inevitably increase as we proceed backward in time which makes it difficult to find an analytical solution.

3. (a) Neglecting the constraints on the control signal yields the Bellman equation

$$J(x) = \min_u \{f_0(x, u) + J(f(x, u))\},$$

which in our case with $J(x) = px^2$ can be written as

$$px^2 = \min_u \{\underbrace{x^2 + u^2 + p(x+u)^2}_{\triangleq h(x,u)}\}.$$

Differentiating h with respect to u and setting the derivative to zero yields

$$0 = \frac{\partial h}{\partial u}(x, u) = 2u + 2p(x+u) \quad \Longrightarrow \quad u^* = -\frac{p}{1+p}x,$$

which is a minimum since

$$\frac{\partial^2 h}{\partial u^2}(x, u^*) = 2 + 2p > 0.$$

Now, the HJBE reads

$$px^2 = x^2 + \left(-\frac{p}{1+p}x\right)^2 + p\left(x - \frac{p}{1+p}x\right)^2 = \frac{1+2p}{1+p}x^2,$$

which implies that $p^2 - p - 1 = 0$ and therefore

$$p = \frac{1+\sqrt{5}}{2}.$$

Finally, the optimal controller is given by

$$u^* = -\frac{3-\sqrt{5}}{2}x$$

(b) If $u \in (-1, 1)$, then the solution is the same as in a). Thus,

$$u = -\frac{3-\sqrt{5}}{2}x,$$

as long as x satisfies $-1 \leq (3-\sqrt{5})x/2 \leq 1$. Otherwise, the solution lies on the boundary. Hence,

$$u = \begin{cases} 1, & x \leq -\frac{2}{3-\sqrt{5}} \\ -\frac{3-\sqrt{5}}{2}x, & -\frac{2}{3-\sqrt{5}} \leq x \leq \frac{2}{3-\sqrt{5}} \\ -1, & x \geq \frac{2}{3-\sqrt{5}} \end{cases}.$$

(c) This problem may be divided into three subproblems:

(i) If $x_k \leq -2/(3-\sqrt{5})$, then $u_k = 1$ and

$$x_{k+1} = x_k + u_k = x_k + 1 > x_k,$$

which shows that the state sequence is strictly increasing. Since

$$\text{width}\left[-\frac{2}{3-\sqrt{5}}, \frac{2}{3-\sqrt{5}}\right] > 1,$$

we know that $-2/(3-\sqrt{5}) \leq x_{k+n} \leq 2/(3-\sqrt{5})$ after a finite number of steps n.

(ii) In a similar fashion, if $x_k \geq 2/(3-\sqrt{5})$, then $u_k = -1$ and the state sequence is strictly decreasing and will reach the region of no saturation in a finite number of steps.

(iii) What remains to be shown is that the nominal controller

$$u = -\frac{3-\sqrt{5}}{2}x,$$

is stabilizing. But this is a direct consequence of Theorem 2 in Chapter 2 of the course compendium.

4. (a) By introducing the states $x_1 = y$ and $x_2 = \dot{y}$, the dynamics can be expressed as

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ k \end{bmatrix} \tag{3}$$

The Hamiltonian of the problem is then given by

$$H(x, u, \lambda) = |u| + \lambda^T(Ax + Bu + K) \tag{4}$$

Pointwise minimization of the Hamiltonian w.r.t. u gives

$$u^*(t, \lambda) = \underset{u\in[-1,1]}{\text{argmin}}\ H(x, u, \lambda) = \underset{u\in[-1,1]}{\text{argmin}}\ |u| + \lambda_2 u = \begin{cases} 0, & \text{if } \lambda_2 \in [-1, 1] \\ -1, & \text{if } \lambda_2 > 1 \\ 1, & \text{if } \lambda_2 < 1 \end{cases} \tag{5}$$

and the adjoint equation gives

$$\dot{\lambda} = -H_x = -A^T\lambda = \begin{bmatrix} 0 \\ -\lambda_1 \end{bmatrix} \implies \begin{cases} \lambda_1 & = c_1 \\ \lambda_2 & = c_1 t + c_2 \end{cases} \tag{6}$$

for some constants $c_1, c_2 \in \mathbb{R}$.
Furthermore, since $x_2(T)$ is free we have that $\lambda_2(T) = 0$. Since λ_2 is affine in t, the optimal control switches at most one time, resulting in either

$$u^*(t) = \begin{cases} -1, & 0 \leq t \leq \tilde{t} \\ 0, & \tilde{t} \leq t \leq T \end{cases} \tag{7}$$

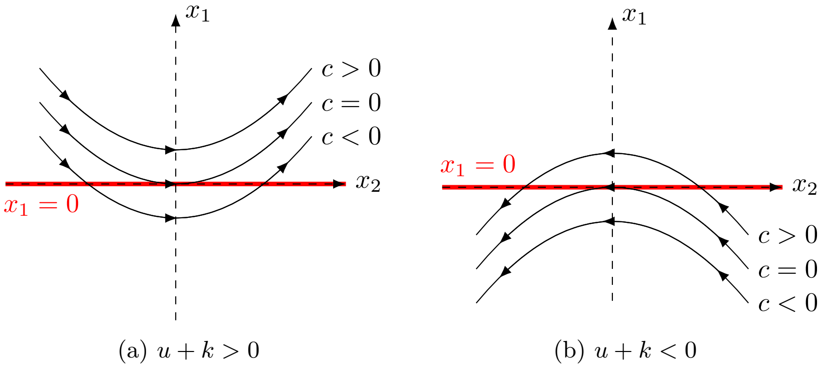

Figure 1: Phase plane for the system when u is constant.

or

$$u^*(t) = \begin{cases} 1, & 0 \le t \le \tilde{t} \\ 0, & \tilde{t} \le t \le T \end{cases} \tag{8}$$

depending on if the slope c_1 is positive or negative, respectively. We note that in both cases the optimal cost will be

$$\int_0^T |u^*(t)|dt = \int_0^{\tilde{t}} 1dt = \tilde{t} \tag{9}$$

Hence, we want to have a non-zero input for the shortest time possible while still satisfying $y(T) = x_1(T) = 0$.

Using the state dynamics, the following relationship can be established,

$$\frac{\dot{x}_1}{\dot{x}_2} = \frac{dx_1}{dx_2} = \frac{x_2}{u+k} \implies x_1 = \frac{x_2^2}{2(u+k)} + c \tag{10}$$

for some constant $c \in \mathbb{R}$. Depending on the sign of $u+k$ we get different charactersitcs in the phase plane, shown in Figure 1.

Since $0 < k < 1$ we have that $u+k > 0$, when $u \ge 0$, so the candidate policy given by (8) will only follow the characteristics of (a) in Figure 1. However, this leads to some states in, e.g., the first quadrant not following a trajectory which crosses the x_2-axis. I.e., for some initial conditions $x = x_0$, $x_1(T) = 0$ can not be satisfied for any $T > 0$ using a control on the form (8).

A policy on the form (7) makes it possible for a trajectory to combine both characteristics in Figure 1, since $u+k$ can become both negative and then positive with such a policy, and hence $x_1(T) = 0$ can always be satisfied for some $T > 0$.

Also, note that for some initial values of x already starts on a trajectory that will cross the x_2-axis and we will get $\tilde{t} = 0$ which gives $u^*(t) = 0$.

(b) As was mentioned before the optimal cost will be $\tilde{t}$ which is the switching time. Hence, we want to have a non-zero u for a time period as short as possible, idealy $\tilde{t} = 0$, while still satisfying $x_1(T) = 0$ for some $T > 0$. When $u = 0$, $u + k = k > 0$, resulting in the characteristics given by Figure 1 (a). In Figure 1 (a) it can be seen that all $x_1 \le 0$ will follow a trajectory which crosses the x_2-axis. Also $c \le 0$ and $x_2 \le 0$ lead to a trajectory which crosses the x_2-axis being followed, i.e., $u^* = 0$ is also optimal for all x satisfying $x_2 \le 0$ and $x_1 \le \frac{x_2^2}{2k}$.

All other states results in $u = -1$ for some time to force the state into the above mentioned region. In summary we get the feedback policy, (recall that $x_1 = y$ and $x_2 = \dot{y}$),

$$u(y, \dot{y}) = \begin{cases} 0, & y \le 0 \text{ or } y \le \frac{\dot{y}^2}{2k} \text{ and } \dot{y} \le 0 \\ -1, & \text{otherwise} \end{cases} \tag{11}$$

(c) For this case the constraint on $\lambda_2(T) = 0$ disappears and we get the cases

$$u^* = \begin{cases} -1, & 0 \leq t \leq t_1 \\ 0, & t_1 \leq t \leq t_2 \\ 1, & t_2 \leq t \end{cases} \tag{12}$$

$$u^* = \begin{cases} 1, & 0 \leq t \leq t_1 \\ 0, & t_1 \leq t \leq t_2 \\ -1, & t_2 \leq t \end{cases} \tag{13}$$

depending on the slope c_1 of λ_2.