# TSRT08: Optimal Control
# Solutions

20260114

1. (a) The Hamiltonian is given by

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

Pointwise minimization yields

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

The adjoint equation is given by

$$\dot{\lambda}(t) = -\lambda(t) - 1, \quad \lambda(T) = 0$$

which is a first order linear ODE with the solution

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

and the optimal control is

$$u^*(t) = -\frac{1}{2}\lambda(t) = \frac{1 - e^{T-t}}{2}.$$

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

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

The following equations must hold

$$\dot{\lambda} = -\frac{\partial H}{\partial x}(t, x, u, \lambda) = -2x,$$
$$0 = \frac{\partial H}{\partial u}(t, x, u, \lambda) = 2u + \lambda.$$

The latter yields $\dot{\lambda} = -2\dot{u} = -2\ddot{y}$, which plugged into the first equation gives

$$-2\ddot{y} = -2x = -2y \quad \Leftrightarrow \quad \ddot{y} - y = 0,$$

which has the solution

$$y(t) = c_1 e^t + c_2 e^{-t},$$

for some constants $c_1$ and $c_2$. The boundary constraints $y(0) = 0$ and $y(1) = 1$ yields

$$\begin{pmatrix} 1 & 1 \\ e^1 & e^{-1} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

which has the solution

$$\begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \frac{1}{e^{-1} - e^1} \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$$

Thus,

$$y(t) = \frac{1}{e^{-1} - e^1}(e^{-t} - e^t),$$

is the sought after extremal.

2. The constraints $U(k, x_k)$ on the input signal may be expressed as

$$U(k, x_k) = \{u_k : ax_k^\alpha - u_k \geq 0\}.$$

The dynamic programming algorithm yields:

- $k = 2$: Then it holds that
$$V_2(x_2) = 0.$$
- $k = 1$: Now, the above yields
$$V_1(x_1) = \min_{u_1 \in U(1,x_1)} \{-\beta \log(u_1)\},$$
which takes its minimum at $u_1^* = ax_1^\alpha$. Thus,
$$V^*(x_1) = -\beta \log(ax_1^\alpha).$$
- $k = 0$: Finally, it follows that
$$\begin{aligned} V_0(x_0) &= \min_{u_0 \in U(0,x_0)} \{-\log(u_0) - \beta \log\left(a(ax_0^\alpha - u_0)^\alpha\right)\} \\ &= \min_{u_0 \in U(0,x_0)} \{\underbrace{-\log(u_0) - \beta \log(a) - \alpha\beta \log(ax_0^\alpha - u_0)}_{\triangleq h(u_0)}\}. \end{aligned}$$
The extrema are found via differentiation of $h$, i.e.,
$$0 = h'(u_0) = -\frac{1}{u_0} + \alpha\beta \frac{1}{ax_0^\alpha - u_0} = \frac{\alpha\beta u_0 - (ax_0^\alpha - u_0)}{u_0(ax_0^\alpha - u_0)} = \frac{(1+\alpha\beta)u_0 - ax_0^\alpha}{u_0(ax_0^\alpha - u_0)}.$$
Since $1 + \alpha\beta > 1$, the extremum is attained by
$$u_0^* = \frac{1}{1+\alpha\beta} ax_0^\alpha \in U(0, x_0),$$
with the corresponding optimal cost-to-go function easily calculated from the above. That this constitutes a minimum is verified by the fact that
$$h''(u_0) = \frac{1}{u_0^2} + \alpha\beta \frac{1}{(ax_0^\alpha - u_0)^2} \geq 0.$$

3. (a) In order to formulate the maximum likelihood estimation problem as an discrete-time optimal control problem, we choose to maximize $\log p(\bar{x})$ instead of maximizing $p(\bar{x})$, which we can do since $\log(\cdot)$ is a strictly increasing function. From the question description we have
$$\begin{aligned} \log p(\bar{x}) = \log P\{y(0) = y_0 | x(0) = x_0\} + \log P\{x(0) = x_0\} \\ + \sum_{k=1}^{N-1} \log P\{y(k) = y_k | x(k) = x_k\} + \log P\{x(k) = x_k | x(k-1) = x_{k-1}\} \end{aligned}$$
By introducing
$$\begin{aligned} V(\bar{x}) &= \log p(\bar{x}) \\ u_k &= x_{k-1} \\ \bar{u} &= (u_1, \dots u_N) = \bar{x} \\ \phi(x) &= \log P\{y(0) = y_0 | x(0) = x\} + \log P\{x(0) = x\} \\ f(k, x, u) &= \log P\{y(k) = y_k | x(k) = x\} + \log P\{x(k) = x | x(k-1) = u\} \\ F(k, x, u) &= u \end{aligned}$$

we can formulate the maximization problem as the following optimal control problem

$$\begin{aligned} \text{maximize}_{\bar{u}} \quad & J(\bar{u}) = \phi(x_0) + \sum_{k=1}^{N-1} f(k, x_k, u_k) \\ \text{subject to} \quad & x_{k-1} = F(k, x_k, u_k) \end{aligned}$$

The solution to the problem can be found with dynamic programming. Since the dynamics is running backwards in time, the dynamics programming recursions will run in reverse order.

$$\begin{aligned} J_0(x) &= \phi(x) \\ J_k(x) &= \max_{u \in X} \{f(k, x, u) + J_{k-1}(F(k, x, u))\}, \quad k = 1, \ldots, N-1 \end{aligned}$$

At stage $N-1$ we have optimized for all $u_1, \ldots, u_{N-1}$ except $u_N$. We then have

$$J_{N-1}(x_{N-1}) = \max_{(u_1, \ldots, u_{N-1}) \in X^{N-1}} J(\bar{u})$$

where $x_{N-1} = u_N$. Hence

$$\max_{\bar{x} \in X^N} J(\bar{x}) = \max_{\bar{u} \in X^N} J(\bar{u}) = \max_x J_{N-1}(x)$$

By identifying

$$\log V_k(x) = J_k(x) \quad \text{and} \quad \log V(\bar{x}) = J(\bar{x})$$

we get the following recursions

$$\begin{aligned} V_0(x) &= P\{y(0) = y_0 | x(0) = x\} P\{x(0) = x\} \\ V_k(x) &= P\{y(k) = y_k | x(k) = x\} \max_{u \in X} \big[P\{x(k) = x | x(k-1) = u\} V_{k-1}(u)\big], \quad k = 1, \ldots, N-1 \end{aligned}$$

$$\max_{\bar{x} \in X^N} V(\bar{x}) = \max_{x \in X} V_{N-1}(x)$$

where the optimal solution is given by

$$\begin{aligned} x_{N-1}^* &= \arg\max_{u \in X} V_{N-1}(x) \\ x_{k-1}^* &= u_k^* = \mu(k, x_k^*), \quad \text{if} \quad k < N \end{aligned}$$

with

$$\mu(k, x) = \arg\max_{u \in X} \big[P\{x(k) = x | x(k-1) = u\} V_{k-1}(u)\big]$$

(b) We have the optimization problem on the form of the previous question with $N = 3$.

**Stage** $k = 0$: The first day, the doctor is making the observation $y_0 = \beta_1$, which gives

$$\begin{aligned} V_0(\alpha_i) &= P\{y(0) = \beta_1 | x(0) = \alpha_i\} P\{x(0) = \alpha_i\} \\ &= b_{i1} P\{x(0) = \alpha_i\} \end{aligned}$$

This gives

$$\begin{aligned} V_0(\alpha_1) &= b_{11} * 0.6 = 0.5 * 0.6 = 0.3 \\ V_0(\alpha_2) &= b_{21} * 0.4 = 0.1 * 0.4 = 0.04 \end{aligned}$$

| $x$ | $V_0(x)$ | $u^*$ |
|---|---|---|
| $\alpha_1$ | 0.3 | - |
| $\alpha_2$ | 0.04 | - |

**Stage** $k = 1$: The second day, the doctor is making the observation $y_1 = \beta_2$, which gives

$$\begin{aligned} V_1(\alpha_i) &= P\{y(1) = \beta_2 | x(1) = \alpha_i\} \max_{u \in \{\alpha_i, \alpha_2\}} \left[ P\{x(1) = \alpha_i | x(0) = u\} V_0(u) \right] \\ &= b_{i2} \max_{i \in \{1,2\}} \{a_{1i} V_0(\alpha_1), a_{2i} V_0(\alpha_2)\} \end{aligned}$$

This gives

$$V_1(\alpha_1) = b_{12} * \max\{a_{11} V_0(\alpha_1), a_{21} V_0(\alpha_2)\} = 0.4 * \max\{0.7 * 0.3, 0.4 * 0.04\} = 0.4 * 0.7 * 0.3 = 0.084$$
$$V_1(\alpha_2) = b_{22} * \max\{a_{12} V_0(\alpha_1), a_{22} V_0(\alpha_2)\} = 0.3 * \max\{0.3 * 0.3, 0.6 * 0.04\} = 0.3 * 0.3 * 0.3 = 0.027$$

| $x$ | $V_1(x)$ | $u^*$ |
|---|---|---|
| $\alpha_1$ | 0.084 | $\alpha_1$ |
| $\alpha_2$ | 0.027 | $\alpha_1$ |

**Stage** $k = 2$: The third day, the doctor is making the observation $y_2 = \beta_3$. In a similar fashion as in the previous step we get

$$V_2(\alpha_i) = b_{i3} \max\{a_{1i} V_1(\alpha_1), a_{2i} V_1(\alpha_2)\}$$

This gives

$$V_2(\alpha_1) = b_{13} * \max\{a_{11} V_1(\alpha_1), a_{21} V_1(\alpha_2)\} = 0.1 * \max\{0.7 * 0.084, 0.4 * 0.027\} = 0.1 * 0.7 * 0.084$$
$$V_2(\alpha_2) = b_{23} * \max\{a_{12} V_1(\alpha_1), a_{22} V_1(\alpha_2)\} = 0.6 * \max\{0.3 * 0.084, 0.6 * 0.027\} = 0.6 * 0.3 * 0.084$$

| $x$ | $V_2(x)$ | $u^*$ |
|---|---|---|
| $\alpha_1$ | 0.07*0.084 | $\alpha_1$ |
| $\alpha_2$ | 0.18*0.084 | $\alpha_1$ |

Since $V_2(\alpha_2) > V_2(\alpha_1)$ we have that

$$x_2^* = \arg\max_{x \in \{\alpha_1, \alpha_2\}} V_2(x) = \alpha_2$$

To summarize, the system evolves as follows:

| $k$ | $x_k^*$ | $u_k^*$ | $V_k(x_k^*)$ |
|---|---|---|---|
| 2 | $\alpha_2$ | $\alpha_1$ | 0.18*0.084 |
| 1 | $\alpha_1$ | $\alpha_1$ | 0.084 |
| 0 | $\alpha_1$ | - | 0.3 |

Thus, the most likely condition for the patient is that he/she was healthy the first and second day, $x_0^* = x_1^* = \alpha_1$, and had fever the third day, $x_2^* = \alpha_2$.

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

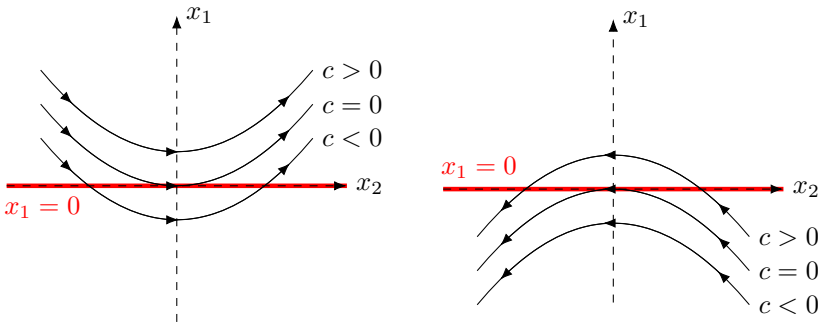
$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u + \begin{bmatrix} 0 \\ k \end{bmatrix} \tag{1}$$

The Hamiltonian of the problem is then given by

$$H(x, u, \lambda) = |u| + \lambda^T (Ax + Bu + K) \tag{2}$$

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

$$u^*(t, \lambda) = \operatorname*{argmin}_{u \in [-1,1]} H(x, u, \lambda) = \operatorname*{argmin}_{u \in [-1,1]} |u| + \lambda_2 u = \begin{cases} 0, & \text{if } \lambda_2 \in [-1, 1] \\ -1, & \text{if } \lambda_2 > 1 \\ 1, & \text{if } \lambda_2 < 1 \end{cases} \tag{3}$$




(a) $u + k > 0$ (b) $u + k < 0$

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

and the adjoint equation gives

$$\dot{\lambda} = -H_x = -A^T\lambda = \begin{bmatrix} 0 \\ -\lambda_1 \end{bmatrix} \implies \begin{cases} \lambda_1 &= c_1 \\ \lambda_2 &= -c_1 t + c_2 \end{cases} \tag{4}$$

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 $\lambda_2$ is affine in $t$, the optimal control switches at most one time, resulting in either

$$u^*(t) = \begin{cases} -1, & 0 \le t \le \tilde{t} \\ 0, & \tilde{t} \le t \le T \end{cases} \tag{5}$$

or

$$u^*(t) = \begin{cases} 1, & 0 \le t \le \tilde{t} \\ 0, & \tilde{t} \le t \le T \end{cases} \tag{6}$$

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

$$\int_0^T |u^*(t)|dt = \int_0^{\tilde{t}} 1dt = \tilde{t} \tag{7}$$

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,

$$\frac{\dot{x}_1}{\dot{x}_2} = \frac{dx_1}{dx_2} = \frac{x_2}{u + k} \implies x_1 = \frac{x_2^2}{2(u + k)} + c \tag{8}$$

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 (6) 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 (6).
A policy on the form (5) 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 $x_1 \leq 0$ will follow a trajectory which crosses the $x_2$-axis. Also $c \leq 0$ and $x_2 \leq 0$ lead to a trajectory which crosses the $x_2$-axis being followed, i.e., $u^* = 0$ is also optimal for all $x$ satisfying $x_2 \leq 0$ and $x_1 \leq \frac{x_2^2}{2k}$.
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 $x_2 = \dot{y}$),

$$u(y, \dot{y}) = \begin{cases} 0, & y \leq 0 \text{ or } y \leq \frac{\dot{y}^2}{2k} \text{ and } \dot{y} \leq 0 \\ -1, & \text{otherwise} \end{cases} \tag{9}$$

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

$$u^* = \begin{cases} -1, & 0 \leq t \leq t_1 \\ 0, & t_1 \leq t \leq t_2 \\ 1, & t_2 \leq t \end{cases} \tag{10}$$

$$u^* = \begin{cases} 1, & 0 \leq t \leq t_1 \\ 0, & t_1 \leq t \leq t_2 \\ -1, & t_2 \leq t \end{cases} \tag{11}$$

depending on the slope $c_1$ of $\lambda_2$.