

Optimal Control WS20/21: Homework 2

Daniel Bergmann

January 19, 2021

Problem 1

a) Formulating the problem as an discrete-time, infinite-horizon o. c. problem yields

$$\begin{aligned} \min_{u_0, u_1, \dots} \quad & \sum_{k=0}^{\infty} \alpha^k f_0(x_k, u_k) = \sum_{k=0}^{\infty} 0.9^k f_0(x_k, u_k) \\ \text{subject to} \quad & x_{k+1} = f(x_k, u_k), \\ & x_k \in \mathcal{X} = \{\xi_1, \dots, \xi_8\}, \\ & u_k \in \mathcal{U} = \{0, 1, 2\}, \\ & x_0 = \xi_1, \end{aligned} \tag{1}$$

where the dynamics $f : \mathcal{X} \times \mathcal{U} \rightarrow \mathcal{X}$ are defined by the arrows in the graph.

The Value function iteration is defined by the Bellman operator

$$V_{k+1}(x) = TV_k(x) = \min_{u \in \mathcal{U}} \{f_0(x, u) + \alpha V_k(x)\} \quad \text{with} \quad \alpha = 0.9. \tag{2}$$

Evaluated for this particular problem, this yields

$$\begin{aligned} V_{k+1}(\xi_1) &= \min_{u \in \mathcal{U}} \{1 + 0.9V_k(\xi_2), 1 + 0.9V_k(\xi_2)\}, \\ V_{k+1}(\xi_2) &= \min_{u \in \mathcal{U}} \{3 + 0.9V_k(\xi_7), 6 + 0.9V_k(\xi_5), 3 + 0.9V_k(\xi_4)\}, \\ V_{k+1}(\xi_3) &= \min_{u \in \mathcal{U}} \{1 + 0.9V_k(\xi_4), 2 + 0.9V_k(\xi_6), 3 + 0.9V_k(\xi_5)\}, \\ V_{k+1}(\xi_4) &= \min_{u \in \mathcal{U}} \{2 + 0.9V_k(\xi_7), 6 + 0.9V_k(\xi_8), 3 + 0.9V_k(\xi_6)\}, \\ V_{k+1}(\xi_5) &= \min_{u \in \mathcal{U}} \{0 + 0.9V_k(\xi_4), 0 + 0.9V_k(\xi_4), 1 + 0.9V_k(\xi_6)\}, \\ V_{k+1}(\xi_6) &= \min_{u \in \mathcal{U}} \{5 + 0.9V_k(\xi_1), 1 + 0.9V_k(\xi_7), 1 + 0.9V_k(\xi_8)\}, \\ V_{k+1}(\xi_7) &= \min_{u \in \mathcal{U}} \{2 + 0.9V_k(\xi_8), 2 + 0.9V_k(\xi_8), 2 + 0.9V_k(\xi_8)\}, \\ V_{k+1}(\xi_8) &= \min_{u \in \mathcal{U}} \{0 + 0.9V_k(\xi_8), 0 + 0.9V_k(\xi_8), 0 + 0.9V_k(\xi_8)\}, \end{aligned} \tag{3}$$

starting with an arbitrary initial value function $V^0 : \mathcal{X} \rightarrow \mathbb{R}$.

b) Value function $V : \mathbb{R}^8 \rightarrow \mathbb{R}$, obtained after 1000 value function iterations:

$$\begin{aligned} V(\xi_1) &= 3.61 \\ V(\xi_2) &= 4.80 \\ V(\xi_3) &= 2.90 \\ V(\xi_4) &= 3.80 \\ V(\xi_5) &= 1.90 \\ V(\xi_6) &= 1.00 \\ V(\xi_7) &= 2.00 \\ V(\xi_8) &= 0.00. \end{aligned} \tag{4}$$

The values in (4) induce the following state feedback:

$$\begin{aligned}
k(\xi_1) &= u_2 \\
k(\xi_2) &= u_1 \\
k(\xi_3) &= u_1 \\
k(\xi_4) &= u_1 \\
k(\xi_5) &= u_2 \\
k(\xi_6) &= u_2 \\
k(\xi_7) &= u_0 \\
k(\xi_8) &= u_0.
\end{aligned} \tag{5}$$

For the initial state $x_0 = \xi_1$, the state-feedback given in (5) yields the optimal input sequence

$$u^* = (2 \quad 1 \quad 2),$$

that steers the state to the terminal state ξ_8 via the route $\xi_1, \xi_3, \xi_6, \xi_8$.

c) We suppose

$$V^0(\xi) \leq TV^0(\xi) \quad \forall \xi \in \mathcal{X}, \tag{6}$$

for an initial function V^0 for a value function iteration with the Bellman Operator, defined in (2). From the lectures, we have the following properties.

$$\|TV^1 - TV^2\|_\infty \leq \alpha \|V^1 - V^2\|_\infty \tag{Contraction} \tag{7}$$

$$V^1(\xi) \leq V^2(\xi) \quad \forall \xi \in \mathcal{X} \implies TV^1(\xi) \leq TV^2(\xi) \quad \forall \xi \in \mathcal{X} \tag{Monotonicity} \tag{8}$$

V^* solves the Bellman Equation V^* is the Value function, since we have discounted costs. (9)

First we show, that the iteration converges to the value function of the problem. Since (7) holds and $\alpha \in (0, 1)$, we deduce the convergence of $\{V_k\}_{k \in \mathbb{N}}$ by the Banach-Fixed-Point-Theorem. Since we have such a limit $\lim_{k \rightarrow \infty} V_k \longrightarrow V^*$, it has to satisfy the fixed point property $TV^* = V^*$. Hence it satisfies the Bellman equation and (9) verifies V^* as the value function of our problem.

Second, it remains to show that $V^0(\xi) \leq V^*(\xi) \forall \xi \in \mathcal{X}$ holds.

By assumption (6) and the monotonicity property (8), we infer inductively, that $V_k(\xi) \leq V_{k+1}(\xi)$ for all $\xi \in \mathcal{X}$. Thus, the sequence $\{V_k(\xi)\}_{k \in \mathbb{N}_0}$ is non-decreasing for any $\xi \in \mathcal{X}$. This proves the claim $V^0(\xi) \leq V^*(\xi) \forall \xi \in \mathcal{X}$.

d) Maximization Problem:

$$\begin{aligned}
&\max_{V(\xi_1), \dots, V(\xi_n)} \sum_{k=0}^n V(\xi_k) \\
&\text{subject to} \quad V(\xi_i) \leq TV(\xi_i) \quad \forall i=1, \dots, n.
\end{aligned} \tag{10}$$

To show: V^* is a solution to the maximization problem given above. Proof by contradiction. Suppose, there exists an admissible Function $\tilde{V} : \mathbb{R}^n \longrightarrow \mathbb{R}$ with

$$\tilde{V}(\xi_i) > V^*(\xi_i) \tag{11}$$

for at least one $i \in \{1, \dots, n\}$. Since \tilde{V} is admissible for the given max. problem, we know that $\tilde{V}(\xi_i) \leq T\tilde{V}(\xi_i) \quad \forall i \in \{1, \dots, n\}$. From exercise c) we know, that hence $\tilde{V}(\xi_i) \leq T^k \tilde{V}(\xi_i) \leq V^*(\xi_i)$ holds for all $i \in \{1, \dots, n\}$ and for all $k \in \mathbb{N}$. This clearly contradicts (11). Thus, V^* maximizes the objective of the given problem.

e) Transform (10) into

$$\begin{aligned}
V^* = \arg \max_{V(\xi_1), \dots, V(\xi_n)} & \sum_{k=0}^n V(\xi_k) \\
\text{subject to} & V(\xi_i) \leq f_0(\xi_i, u) + \alpha V(f(\xi, u)) \\
& \forall u \in \mathcal{U}(\xi_i) \forall i=1, \dots, n.
\end{aligned} \tag{12}$$

As shown in d), V^* maximizes the objective of problem (10). It remains to show, that the constraints of problems (10) and (14) are equivalent. The operator T is defined by the Bellman Equation, as written in (2). We have

$$\begin{aligned}
V(\xi_i) \leq TV(\xi_i) &= \min_{u \in \mathcal{U}} \{f_0(\xi_i, u) + \alpha V_k(\xi_i)\} \\
&\leq f_0(\xi_i, \tilde{u}) + \alpha V_k(\xi_i)
\end{aligned} \tag{13}$$

for any input signal $\tilde{u} \in \mathcal{U}$. This arises directly from the definition of the minimum. Thus, (10) and (14) denote the same problem.

It remains to formulate this as a linear program

$$\begin{aligned}
V^* = \arg \max_V & c^T V \\
\text{subject to} & AV \leq b.
\end{aligned} \tag{14}$$

With $V = (V(\xi_1), \dots, V(\xi_n))^T$, the objective $\max. \sum_{k=0}^n V(\xi_k)$ is the same as $\min. c^T V$ with

$$c = -(1, \dots, 1)^T \in \mathbb{R}^n.$$

Now we rewrite the constraints as a linear inequality $AV = b$.

For a state $\xi_i \in \mathcal{X}$, the inequality

$$\begin{aligned}
V(\xi_i) &\leq f_0(\xi_i, u) + \alpha V(f(\xi_i, u)) \\
\iff V(\xi_i) - \alpha V(f(\xi_i, u)) &\leq f_0(\xi_i, u)
\end{aligned} \tag{15}$$

has to hold for each possible input $u \in \mathcal{U}$, which leads to three scalar inequalities per state. Such an inequality can be represented with a line in the matrix A , with an 1 in the i -th column and $-\alpha$ in the $f(\xi_i, u)$ -th column. The corresponding entry of the vector b is $f_0(\xi_i, u)$. The resulting matrices A, b for the particular problem can be found in the solutions-sheet or can be generated by the matlab script.

One can determine the optimal input u^* in any state ξ_i by using the input that makes inequality (13) an equality. This minimizes the V -value of the next state that will be reached.

Problem 2

a)

b)