ECE-GY 9243 / ME-GY 7973
Optimal and learning control for robotics

Tarun Thathvik

Exercise Series 1
Finite Horizon Optimal Control

Exercise 1

Part a)

Given,

$$x_{n+1} = \begin{cases} -x_n + 1 + u_n & \text{if } -2 \geq -x_n + 1 + u_n \leq 0 \\ 2 & \text{if } -x_n + 1 + u_n > 0 \\ -2 & \text{else} \end{cases} \tag{1}$$

The possible states are $\{-2, -1, 0, 1, 2\}$ and possible control actions are $\{-1, 0, 1\}$

$$J = \left(\sum_{k=0}^{2} 2|x_k| + |u_k| \right) + x_3^2 \tag{2}$$

From this, we infer that:

$$g_i(x_i) = 2|x_i| + |u_i| \qquad \text{The stage cost} \tag{3}$$
$$g_N(x_N) = x_N^2 \qquad \text{Terminal cost} \tag{4}$$

Solving using Dynamic Programming

It is a backward recursion process, so the cost at stage $N = 3$ is calculated first:

$$g_N(x_N) = x_N^2$$

and then the cost at stage $N - 1 = 2$, and so on, using the following recursion formula:

$$J_k(x_k) = g_k(x_k, u_k) + J_{k+1}(x_{k+1}) \tag{5}$$

Substituting the equation eq.(3) in eq.(5) gives:

$$J_k(x_k) = 2|x_k| + |u_k| + J_{k+1}(x_{k+1}) \tag{6}$$

The value of x_{k+1} is found using eq.(1), and the corresponding J_{k+1} is taken from the table as it fills up.

States x	Stage 0		Stage 1		Stage 2		Stage 3
	J_0	u_0	J_1	u_1	J_2	u_2	J_3
-2	10	0	9	0	8	0	4
-1	6	-1	5	-1	4	-1	1
0	3	0, -1	2	0, -1	1	0, -1	0
1	4	0	3	0	2	0	1
2	7	1	6	1	5	0, 1	4

Table 1: Solution for the Dynamic Program

As you can see in the Table(1), in some of the stages, for the state, there are multiple optimal control values, with the same cost-to-go. This opens the possibility for various optimal control paths.

Part b)

For $x_0 = 0$, the optimal control cost is 3 with five different Optimal control solutions:

- $u^* = \{0, 0, 0\}$ and $x^* = \{0, 1, 0, 1\}$ with the final state $x_3 = 1$
- $u^* = \{0, 0, -1\}$ and $x^* = \{0, 1, 0, 0\}$ with the final state $x_3 = 0$
- $u^* = \{-1, 0, 0\}$ and $x^* = \{0, 0, 1, 0\}$ with the final state $x_3 = 0$
- $u^* = \{-1, -1, 0\}$ and $x^* = \{0, 0, 0, 1\}$ with the final state $x_3 = 1$
- $u^* = \{-1, -1, -1\}$ and $x^* = \{0, 0, 0, 0\}$ with the final state $x_3 = 0$

For $x_0 = -2$ the optimal control cost is 10 with two different Optimal control solutions:

- $u^* = \{0, 1, 0\}$ and $x^* = \{-2, 2, 0, 1\}$ with the final state $x_3 = 1$
- $u^* = \{0, 1, -1\}$ and $x^* = \{-2, 2, 0, 0\}$ with the final state $x_3 = 0$

For $x_0 = 2$ the optimal control cost is 7 with three different Optimal control solutions:

- $u^* = \{1, 0, 0\}$ and $x^* = \{2, 0, 1, 0\}$ with the final state $x_3 = 0$
- $u^* = \{1, -1, 0\}$ and $x^* = \{2, 0, 0, 1\}$ with the final state $x_3 = 1$
- $u^* = \{1, -1, -1\}$ and $x^* = \{-2, 2, 0, 0\}$ with the final state $x_3 = 0$

Part c)

$$x_{n+1} = \begin{cases} -x_n + w_n + u_n & \text{if } -2 \geq -x_n + w_n + u_n \leq 0 \\ 2 & \text{if } -x_n + w_n + u_n > 0 \\ -2 & \text{else} \end{cases} \tag{7}$$

With the same possible states and control actions as Part a, i.e. $x_n \in \{-2, -1, 0, 1, 2\}$ and $u_n \in \{-1, 0, 1\}$. Unlike eq(1), from Part a, eq(7) has a random variable w_n. Where $w_n \in \{0, 1\}$ and has a probability distribution $p(w_n = 0) = 0.3$, $p(w_n = 1) = 0.7$. And the cost function is:

$$J = E\left[\left(\sum_{k=0}^{2} 2|x_k| + |u_k|\right) + x_3^2\right] \tag{8}$$

From this, we infer that we get the same costs:

$$\begin{aligned} g_i(x_i) &= 2|x_i| + |u_i| && \text{The stage cost} && (3) \\ g_N(x_N) &= x_N^2 && \text{Terminal cost} && (4) \end{aligned}$$

Solving using Dynamic Programming

It is again a backward recursion process, so the terminal cost at stage $N = 3$ for every x_N is calculated first, using:

$$g_N(x_N) = x_N^2$$

and then the cost at stage $N - 1 = 2$, and so on, using the following probabilistic recursion formula:

$$J_k(x_k) = E\left[g_k(x_k, u_k) + J_{k+1}(x_{k+1})\right] \tag{9}$$

where only the term x_{k+1} is dependant on the random variable; and substituting the equation eq.(3) in eq.(9) gives:

$$J_k(x_k) = 2|x_k| + |u_k| + E\left[J_{k+1}(x_{k+1})\right] \tag{10}$$

The value of x_{k+1} is found using eq.(7), and the corresponding J_{k+1} is taken from the table as it fills up.

States x	Stage 0		Stage 1		Stage 2		Stage 3
	J_0	u_0	J_1	u_1	J_2	u_2	J_3
-2	10.6	0	9.3	0	8	0	4
-1	6.066	-1	4.82	-1	3.7	-1	1
0	3.066	0	1.82	0	0.7	0	0
1	4.72	0	3.6	0	2.3	0	1
2	7.72	1	6.6	1	5.3	1	4

Table 2: Solution for the Probabilistic Dynamic Program

Part d)

Although the cost function of both parts is almost the same, due to uncertainty in part 2 the costs in tab(2) seem to be a little different from the costs in tab(1).

Unlike part a, the Optimal Control of part c cannot be predetermined. The tab(2) only gives the best (optimal) control decision (action) that can be taken when a particular state is achieved at a stage. This results in the best possible outcome for the next stage, but the outcome cannot be determined due to the uncertainty, and therefore we have to wait until the next state is known to make the next control decision.