Exercise 7.1 (student presents)

A man is fishing on his boat on a sea shore. We model his movement in discrete time. The location of the fisherman on period k is denoted by the state variable x_k. It defines how many sea miles to the right (positive value) or to the left (negative value), from the dock the fisherman is. On period $k = 0$, the fisherman starts from the dock. The fisherman knows that a school of trouts appears two sea miles to the right from the dock on period $k = 1$ and one sea mile to the right from the dock on period $k = 2$.

There is a heavy storm on the sea, which is affecting the fisherman's movement. The direction of the waves can change from period to period. Depending on the direction of the waves, they bring him either one sea mile to the right or left, with equal probability, on each period. This is modeled by a random variable w_k, which attains the values -1 and 1 with equal probabilities.

The fisherman's boat is hard to maneuver, thus on each period he can choose to stay on place or commit to moving one sea mile to the right or left. Thus, the control is also discrete: $u_k \in \{-1, 0, 1\}$ when $k = 0, 1$.

$$x_{k+1} = x_k + u_k + w_k \qquad k = 0, 1, \tag{1}$$

where the initial state is $x_0 = 0$.

The fisherman wants to be where the fish is on period $k = 1, 2$, and wants to minimize the effort it takes to steer the boat. Thus, his cost is

$$(x_2 - 1)^2 + (x_1 - 2)^2 + u_1^2 + u_0^2. \tag{2}$$

a) Which locations x_k can the fisherman reach with his boat on $k = 1, 2$?

b) Calculate the optimal cost-to-go $J_0(x_0)$ and optimal control policy $\{\mu_0^*(x_0), \mu_1^*(x_1)\}$.

c) If the fisherman knows that the school of trouts are 3 sea miles to the left of the dock on period $k = 0$, will it affect the optimal control policy? Hint: Add the term $(x_0 - 3)^2$ to the cost.

$u \in \{-1, 0, 1\}$

$w \in \{-1, 1\}$

Solution

a) $k = 1$:

$$\begin{aligned}
\{u_0 = 1, w_0 = 1\} &\Rightarrow x_1 = 2 \\
\{u_0 = 0, w_0 = 1\} &\Rightarrow x_1 = 1 \\
\{u_0 = 1, w_0 = -1\} \vee \{u_0 = -1, w_0 = 1\} &\Rightarrow x_1 = 0 \\
\{u_0 = 0, w_0 = -1\} &\Rightarrow x_1 = -1 \\
\{u_0 = -1, w_0 = -1\} &\Rightarrow x_1 = -2
\end{aligned}$$

$k = 2$:

$$\begin{aligned}
x_1 = 2 &\Rightarrow x_2 = \{0, 1, 2, 3, 4\} \\
x_1 = 1 &\Rightarrow x_2 = \{-1, 0, 1, 2, 3\} \\
x_1 = 0 &\Rightarrow x_2 = \{-2, -1, 0, 1, 2\} \\
x_1 = -1 &\Rightarrow x_2 = \{-3, -2, -1, 0, 1\} \\
x_1 = -2 &\Rightarrow x_2 = \{-4, -3, -2, -1, 0\}
\end{aligned}$$

Thus $x_1 \in \{-2, -1, 0, 1, 2\}$ and $x_2 \in \{-4, -3, -2, -1, 0, 1, 2, 3, 4\}$.

b) Lets calculate the cost of the final state ($k = 2$) $J_2(x_2) = (x_2 - 1)^2$ for all states x_2:

$(4-1)^2 = 3^2 = 9$

$$\begin{aligned}
J_2(4) &= 9 & J_2(3) &= 4 & J_2(2) &= 1 \\
J_2(1) &= 0 & J_2(0) &= 1 & J_2(-1) &= 4 \\
J_2(-2) &= 9 & J_2(-3) &= 16 & J_2(-4) &= 25.
\end{aligned}$$

Now the DP-algorithm can be used in period $k = 1$:

$$J_1(x_1) = \min_{u_1 \in \{-1,0,1\}} E_{w_1} \Big\{ \underbrace{(x_1 - 2)^2 + u_1^2}_{g_1} + J_2(x_1 + u_1 + w_1) \Big\}$$

$\mathbb{E}(J_2) = 0{,}5\, J_2(x_1 + u_1 + 1) + 0{,}5\, J_2(x_1 + u_1 - 1)$

From this follows

$$\begin{aligned}
J_1(2) = 2 \quad &\& \quad \mu_1^*(2) = 0 \vee -1 \\
J_1(1) = 2 \quad &\& \quad \mu_1^*(1) = 0 \\
J_1(0) = 6 \quad &\& \quad \mu_1^*(0) = 0 \vee 1 \\
J_1(-1) = 12 \quad &\& \quad \mu_1^*(-1) = 1 \\
J_1(-2) = 22 \quad &\& \quad \mu_1^*(-2) = 1.
\end{aligned}$$

$\Rightarrow J_1(2) = ((2-2)^2 + u_1^2 + 0{,}5\, J_2(3 + u_1) + 0{,}5\, J_2(1 + u_1)$

$= u_1^2 + 0{,}5\, J_2(3 + u_1) + 0{,}5\, J_2(1 + u_1)$

$u_1 = 0 \Rightarrow 0{,}5\, J_2(3) + 0{,}5\, J_2(1) = 2 + 0 = 2$

$u_1 = -1 \Rightarrow (-1)^2 + 0{,}5\, J_2(2) + 0{,}5\, J_2(0) = 1 + \frac{1}{2} + \frac{1}{2} = 2$

Similarly the DP-algoritm can be used in period $k = 0$:

$$J_0(x_0) = \min_{u_0 \in \{-1,0,1\}} E_{w_0} \Big\{ u_0^2 + J_1(x_0 + u_0 + w_0) \Big\}$$

From this follows

$$J_0(0) = 5 \quad \& \quad \mu_0^*(0) = 1.$$

$u_0 = 1, \; u_1 = 0$

Thus, the fisherman should steer the boat one sea mile to the right in period $k = 0$, which will get him either to $x_1 = 0$ or $x_1 = 2$ in period $k = 1$. In both of those states, the optimal control is to stay on place.

c) This does not matter since the term $(x_0 + 3)^2$ is constant, i.e., the fisherman can not change his location x_0 on period $k = 0$.

Exercise 7.3 (student presents)

The system is

$$x_{k+1} = x_k + u_k + w_k \qquad k = 0, 1, 2,$$

the initial state $x_0 = 0$, state constraints $x_k \in [-2, 4]$ for all k, and the stochastic term w_k attains the value 1 and -1 with equal probabilities for all k. The cost is

$$g_k(x_k, u_k) = \begin{cases} x_k^2 + u_k^2 & x_k \in [0, 2] \\ x_k^2 + u_k^2 + 13 & x_k < 0 \vee x_k > 2 \end{cases}$$

The final stage cost is $g_3(x_3) = 0$. Minimize the cost

$$\sum_{k=0}^{2} g_k(x_k, u_k)$$

and calculate the optimal control policy $u_k = \mu_k^*(x_k)$ with the following control constraints:

$$u_k \in \begin{cases} \{1\} & x_k < 0 \\ \{-1\} & x_k > 2 \\ \{-1, 0, 1\} & x_k = 0, 1, 2. \end{cases}$$

Solution

$k = 3$:

$$J_3(x_3) = g_3(x_3) = 0.$$

$k = 2$:

$$J_2(x_2) = \min_{u_2} E_{w_2}\Big[g_2(x_2, u_2) + \overbrace{J_3(x_3)}^{0}\Big]. = \begin{cases} x_2^2 + u_2^2 & x \in [0,2] \\ x_2^2 + u_2^2 + 13, & x > 2, x < 0 \end{cases}$$

Tabulate values of J_2, with the smallest values **bolded**:

x_2/u_2	-1	0	1	$\mu_2^*(x_2)$
-2	-	-	**18**	1
-1	-	-	**15**	1
0	1	**0**	1	0
1	2	**1**	2	0
2	5	**4**	5	0
3	**23**	-	-	-1
4	**30**	-	-	-1

$\Rightarrow g_2(-2, 0) = 4 + u_2^2 + 13$

$u = 1 \Rightarrow g_2 = 17 + 1^2 = 18$

$k = 1$:

$$
\begin{aligned}
J_1(x_1) &= \min_{u_1} E_{w_1}\Big[g_1(x_1,u_1) + J_2(x_2)\Big] \\
&= \min_{u_1}\Big[g_1(x_1,u_1) + 0.5J_2(x_1+u_1+1) + 0.5J_2(x_1+u_1-1)\Big].
\end{aligned}
$$

Tabulate values of J_1, with the smallest values **bolded**:

x_1/u_1	-1	0	1	$\mu_1^*(x_1)$
-2	-	-	**27**	1
-1	-	-	**23**	1
0	10	8	**3**	1
1	10	**3**	14	0
2	**7**	16	23	-1
(3	**35**	-	-	-1)
(4	**47**	-	-	-1)

(The states $x_1 = 3$ and $x_1 = 4$ can't be reached by starting from $x_0 = 0$.)

$k = 0$:

$$
\begin{aligned}
J_0(x_0) &= \min_{u_0} E_{w_0}\Big[g_0(x_0,u_0) + J_1(x_1)\Big] \\
&= \min_{u_0}\Big[g_0(x_0,u_0) + 0.5J_1(x_0+u_0+1) + 0.5J_1(x_0+u_0-1)\Big].
\end{aligned}
$$

Now the state x_0 is set $x_0 = 0$, thus the minimal value $J_0 = 6$ is attained when $u_0 = \mu_0 = 1$.

$U_1 = -1 \cup 1$

$U_2 = 0$