ISyE 4232 Spring 2013

H. Ayhan

Solutions to Homework 4

1. Define the state space $\mathcal{S} = \{1, 2, 3\}$, where $1 = good$, $2 = fair$, and $3 = poor$. Define the action space $\mathcal{A} = \{0, 1\}$, where $0 = Do\ Nothing$, $1 = Fertilize$. Recall that terminal rewards are $r_4(1) = r_4(2) = r_4(3) = 0$. Also recall the expected immediate rewards we calculated in class are given by

$$r(1,0) = 7 \times 0.2 + 6 \times 0.5 + 3 \times 0.3 = 5.3$$
$$r(2,0) = 5 \times 0.5 + 1 \times 0.5 = 3$$
$$r(3,0) = -1 \times 1 = -1$$
$$r(1,1) = 6 \times 0.3 + 5 \times 0.6 - 1 \times 0.1 = 4.7$$
$$r(2,1) = 7 \times 0.1 + 4 \times 0.6 = 3.1$$
$$r(3,1) = 6 \times 0.05 + 3 \times 0.4 - 2 \times 0.55 = 0.4$$

Note that the immediate rewards are not affected by the discount factor, because they occur in the present. Now we set up a table for period 3:

s/a	0	1	$u_3^*(s)$	$d_3^*(s)$
1	$5.3 + 0.6 \cdot 0 = 5.3$	$4.7 + 0.6 \cdot 0 = 4.7$	5.3	0
2	$3 + 0.6 \cdot 0 = 3$	$3.1 + 0.6 \cdot 0 = 3.1$	3.1	1
3	$-1 + 0.6 \cdot 0 = -1$	$0.4 + 0.6 \cdot 0 = 0.4$	0.4	1

Now we set up a table for period 2:

s/a	0	1	$u_2^*(s)$	$d_2^*(s)$
1	$5.3 + 0.6(0.2 \cdot 5.3 + 0.5 \cdot 3.1 + 0.3 \cdot 0.4) = 6.938$	$4.7 + 0.6(0.3 \cdot 5.3 + 0.6 \cdot 3.1 + 0.1 \cdot 0.4) = 6.794$	6.938	0
2	$3 + 0.6(0.5 \cdot 3.1 + 0.5 \cdot 0.4) = 4.05$	$3.1 + 0.6(0.1 \cdot 5.3 + 0.6 \cdot 3.1 + 0.3 \cdot 0.4) = 4.606$	4.606	1
3	$-1 + 0.6(1 \cdot 0.4) = -0.76$	$0.4 + 0.6(0.05 \cdot 5.3 + 0.4 \cdot 3.1 + 0.55 \cdot 0.4) = 1.435$	1.435	1

Note, here the discount factor starts to affect our results. Now we set up a table for period 1:

s/a	0	1	$u_1^*(s)$	$d_1^*(s)$
1	$5.3 + 0.6(0.2 \cdot 6.9 + 0.5 \cdot 4.6 + 0.3 \cdot 1.4) = 7.772$	$4.7 + 0.6(0.3 \cdot 6.9 + 0.6 \cdot 4.6 + 0.1 \cdot 1.4) = 7.693$	7.772	0
2	$3 + 0.6(0.5 \cdot 4.6 + 0.5 \cdot 1.4) = 4.812$	$3.1 + 0.6(0.1 \cdot 6.9 + 0.6 \cdot 4.6 + 0.3 \cdot 1.4) = 5.432$	5.432	1
3	$-1 + 0.6(1 \cdot 1.4) = -0.139$	$0.4 + 0.6(0.05 \cdot 6.9 + 0.4 \cdot 4.6 + 0.55 \cdot 1.4) = 2.187$	2.187	1

So to summarize the optimal policy and optimal value function we have:

s	$d_1^*(s)$	$d_2^*(s)$	$d_3^*(s)$	$V_4^*(s)$
1	0	0	0	7.772
2	1	1	1	5.432
3	1	1	1	2.187

2. Define the state space $\mathcal{S} = \{1, 2\}$, where 1 = *high* and 2 = *low*. Define the action space $\mathcal{A} = \{0, 1\}$, where 0 = *Do Nothing*, 1 = *Advertise*. The terminal rewards are given by $r_5(1) = 9, r_5(2) = 3$. First calculate the expected immediate rewards

$$
\begin{aligned}
r(1,0) &= 10 \times 0.5 + 4 \times 0.5 = 7 \\
r(2,0) &= 7 \times 0.2 - 2 \times 0.8 = -0.2 \\
r(1,1) &= 7 \times 0.8 + 6 \times 0.2 = 6.8 \\
r(2,1) &= 3 \times 0.4 - 5 \times 0.6 = -1.8
\end{aligned}
$$

Now we set up a table for period 4:

s/a	0	1	$u_4^*(s)$	$d_4^*(s)$
1	$7 + 9 \cdot 0.5 + 3 \cdot 0.5 = 13$	$6.8 + 9 \cdot 0.8 + 3 \cdot 0.2 = 14.6$	14.6	1
2	$-0.2 + 9 \cdot 0.2 + 3 \cdot 0.8 = 4$	$-1.8 + 9 \cdot 0.4 + 3 \cdot 0.6 = 3.6$	4	0

Now we set up a table for period 3:

s/a	0	1	$u_3^*(s)$	$d_3^*(s)$
1	$7 + 14.6 \cdot 0.5 + 4 \cdot 0.5 = 16.3$	$6.8 + 14.6 \cdot 0.8 + 4 \cdot 0.2 = 19.28$	19.28	1
2	$-0.2 + 14.6 \cdot 0.2 + 4 \cdot 0.8 = 5.92$	$-1.8 + 14.6 \cdot 0.4 + 4 \cdot 0.6 = 6.44$	6.44	1

Now we set up a table for period 2:

s/a	0	1	$u_2^*(s)$	$d_2^*(s)$
1	$7 + 19.3 \cdot 0.5 + 6.4 \cdot 0.5 = 19.86$	$6.8 + 19.3 \cdot 0.8 + 6.4 \cdot 0.2 = 23.51$	23.51	1
2	$-0.2 + 19.3 \cdot 0.2 + 6.4 \cdot 0.8 = 8.8$	$-1.8 + 19.3 \cdot 0.4 + 6.4 \cdot 0.6 = 9.77$	9.77	1

Now we set up a table for period 1:

s/a	0	1	$u_1^*(s)$	$d_1^*(s)$
1	$7 + 23.5 \cdot 0.5 + 9.7 \cdot 0.5 = 23.64$	$6.8 + 23.5 \cdot 0.8 + 9.7 \cdot 0.2 = 27.56$	27.56	1
2	$-0.2 + 23.5 \cdot 0.2 + 9.7 \cdot 0.8 = 12.32$	$-1.8 + 23.5 \cdot 0.4 + 9.7 \cdot 0.6 = 13.47$	13.47	1

So to summarize the optimal policy and optimal value function we have:

s	$d_1^*(s)$	$d_2^*(s)$	$d_3^*(s)$	$d_4^*(s)$	$V_5^*(s)$
1	1	1	1	1	27.56
2	1	1	1	0	13.47