作业7

2021 年 12 月 8 日

1　第二题

(1)

$$\mathbf{S}=\big\{\{-2,2\},\{-1,1\},\{0,0\},\{1,-1\},\{2,-2\}\big\}$$
$$\mathbf{P}=\begin{bmatrix}1&0&0&0&0\\q&r&p&0&0\\0&q&r&p&0\\0&0&q&r&p\\0&0&0&0&1\end{bmatrix}$$

(2)

$$p^{(0)}=\begin{bmatrix}0&0&0&1&0\end{bmatrix}^{\mathrm{T}}$$
$$p^{(1)}=\mathbf{P}^{\mathrm{T}}p^{(0)}=\begin{bmatrix}0&0&q&r&p\end{bmatrix}^{\mathrm{T}}$$
$$p^{(2)}=\mathbf{P}^{\mathrm{T}}p^{(1)}=\begin{bmatrix}0&q^2&2qr&pq+r^2&p+pr\end{bmatrix}^{\mathrm{T}}$$

故再赛两局可以结束比赛的概率为$p+pr$。

2　第三题

$$\begin{aligned}v_\pi(s)&=\sum_{a\in\mathbf{A}}\pi(a|s)\Big(r_s^a+\gamma\sum_{s'\in\mathbf{S}}p_{ss'}^a v_\pi(s')\Big)\\&=\frac{1}{4}\sum_{a\in\mathbf{A}}\big(-1+v_\pi(s')\big)\end{aligned}$$

$$
\begin{bmatrix} v_\pi(0) \\ v_\pi(1) \\ v_\pi(2) \\ v_\pi(3) \\ v_\pi(4) \\ v_\pi(5) \\ v_\pi(6) \\ v_\pi(7) \\ v_\pi(8) \end{bmatrix} = \begin{bmatrix} 0 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0.25 & 0.25 & 0.25 & 0 & 0.25 & 0 & 0 & 0 & 0 \\ 0 & 0.25 & 0.5 & 0 & 0 & 0.25 & 0 & 0 & 0 \\ 0.25 & 0 & 0 & 0.25 & 0.25 & 0 & 0.25 & 0 & 0 \\ 0 & 0.25 & 0 & 0.25 & 0 & 0.25 & 0 & 0.25 & 0 \\ 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0 & 0.25 \\ 0 & 0 & 0 & 0.25 & 0 & 0 & 0.5 & 0.25 & 0 \\ 0 & 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & 0.25 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_\pi(0) \\ v_\pi(1) \\ v_\pi(2) \\ v_\pi(3) \\ v_\pi(4) \\ v_\pi(5) \\ v_\pi(6) \\ v_\pi(7) \\ v_\pi(8) \end{bmatrix}
$$

解得

$$
\begin{bmatrix} v_\pi(0) \\ v_\pi(1) \\ v_\pi(2) \\ v_\pi(3) \\ v_\pi(4) \\ v_\pi(5) \\ v_\pi(6) \\ v_\pi(7) \\ v_\pi(8) \end{bmatrix} = \begin{bmatrix} 0 \\ -7 \\ -9 \\ -7 \\ -8 \\ -7 \\ -9 \\ -7 \\ 0 \end{bmatrix}
$$

由

$$
\begin{aligned}
q_\pi(s,a) &= r_s^a + \gamma \sum_{s' \in \mathbf{S}} p_{ss'}^a v_\pi(s') \\
&= -1 + v_\pi(s')
\end{aligned}
$$

有

$$
q_\pi(6, \mathrm{up}) = -1 + v_\pi(3) = -8
$$

$$
q_\pi(5, \mathrm{down}) = -1 + v_\pi(0) = -1
$$