3.解: $q_\pi(s,a) = r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a \sum_{a' \in A} \pi(a'|s') q_\pi(s',a')$

$= r_s^a + \gamma \sum_{s' \in S} P_{ss'}^a v_\pi(s')$

又 $v_\pi(s) = r_s^\pi + \gamma \sum_{s' \in S} P_{ss'}^\pi v_\pi(s')$

$v_\pi = r^\pi + \gamma P^\pi v_\pi$

$\therefore (I - \gamma P^\pi) v_\pi = r^\pi$

$\therefore v_\pi = (I - \gamma P^\pi)^{-1} r^\pi$

又 $P^\pi$:

$$\begin{array}{c|ccccccccc} & 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ \hline 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0.25 & 0.25 & 0.25 & 0 & 0.25 & 0 & 0 & 0 & 0 \\ 2 & 0 & 0.25 & 0.5 & 0 & 0 & 0.25 & 0 & 0 & 0 \\ 3 & 0.25 & 0 & 0 & 0.25 & 0.25 & 0 & 0.25 & 0 & 0 \\ 4 & 0 & 0.25 & 0 & 0.25 & 0 & 0.25 & 0 & 0.25 & 0 \\ 5 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0 & 0.25 \\ 6 & 0 & 0 & 0 & 0.25 & 0 & 0 & 0.5 & 0.25 & 0 \\ 7 & 0 & 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 & 0.25 \\ 8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}$$

由于终止状态

$$\therefore P^\pi = \begin{bmatrix} 0.25 & 0.25 & 0 & 0.25 & 0 & 0 & 0 \\ 0.25 & 0.5 & 0 & 0 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & 0.25 & 0 & 0.25 & 0 \\ 0.25 & 0 & 0.25 & 0 & 0.25 & 0 & 0.25 \\ 0 & 0.25 & 0 & 0.25 & 0.25 & 0 & 0 \\ 0 & 0 & 0.25 & 0 & 0 & 0.5 & 0.25 \\ 0 & 0 & 0 & 0.25 & 0 & 0.25 & 0.25 \end{bmatrix}$$

$$\therefore I - \gamma P^\pi = \begin{bmatrix} 0.75 & -0.25 & 0 & -0.25 & 0 & 0 & 0 \\ -0.25 & 0.5 & 0 & 0 & -0.25 & 0 & 0 \\ 0 & 0 & 0.75 & -0.25 & 0 & -0.25 & 0 \\ -0.25 & 0 & -0.25 & 1 & -0.25 & 0 & -0.25 \\ 0 & -0.25 & 0 & -0.25 & 0.75 & 0 & 0 \\ 0 & 0 & -0.25 & 0 & 0 & 0.5 & -0.25 \\ 0 & 0 & 0 & -0.25 & 0 & -0.25 & 0.75 \end{bmatrix}$$

$$\therefore (I - \gamma P^\pi)^{-1} = \begin{bmatrix} \frac{13}{6} & \frac{3}{2} & \frac{1}{2} & 1 & \frac{5}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & \frac{7}{2} & \frac{1}{2} & 1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{13}{6} & 1 & \frac{1}{2} & \frac{3}{2} & \frac{5}{6} \\ 1 & 1 & 1 & 2 & 1 & 1 & 1 \\ \frac{5}{6} & \frac{3}{2} & \frac{1}{2} & 1 & \frac{13}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{2} & 1 & \frac{1}{2} & \frac{7}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{5}{6} & 1 & \frac{1}{2} & \frac{3}{2} & \frac{13}{6} \end{bmatrix}$$

$$\therefore v_\pi = \begin{bmatrix} \frac{13}{6} & \frac{3}{2} & \frac{1}{2} & 1 & \frac{5}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{3}{2} & \frac{7}{2} & \frac{1}{2} & 1 & \frac{3}{2} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{13}{6} & 1 & \frac{1}{2} & \frac{3}{2} & \frac{5}{6} \\ 1 & 1 & 1 & 2 & 1 & 1 & 1 \\ \frac{5}{6} & \frac{3}{2} & \frac{1}{2} & 1 & \frac{13}{6} & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{3}{2} & 1 & \frac{1}{2} & \frac{7}{2} & \frac{3}{2} \\ \frac{1}{2} & \frac{1}{2} & \frac{5}{6} & 1 & \frac{1}{2} & \frac{3}{2} & \frac{13}{6} \end{bmatrix} \begin{bmatrix} -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -7 \\ -9 \\ -7 \\ -8 \\ -7 \\ -9 \\ -7 \end{bmatrix}$$

$\therefore q_\pi(6, up) = -1 + v_\pi(3) = -8$ , $q_\pi(5, down) = -1 + v_\pi(8) = -1$

4.(1)解: $v(s) = r_s + \gamma \sum_{s' \in S} P_{ss'} v(s')$

$\therefore v(A) = -1 + 0.5(0.5v(C) + 0.5v(B)) = 0.25v(B) + 0.25v(C) - 1$

$$v(B) = -1 + 0.5(0.5v(A) + 0.5v(C)) = 0.25v(A) + 0.25v(C) - 1$$

$$v(C) = 0$$

$$\therefore \quad v(B) = -\frac{4}{3}, \; v(A) = -\frac{4}{3}, \; v(C) = 0$$

(2) 解：可以采用动态规划迭代的方式求解

计算 $v^{(k+1)} = r^{\pi} + \gamma P^{\pi} v^{(k)}$，生成序列 $v_1 \to v_2 \to \cdots \to v_n$

由于只用求解初值，问题规模减小