

$$R_{\text{school}} = \begin{bmatrix} 5 \\ 3 \end{bmatrix}$$

$$R_{\text{stay}} = \begin{bmatrix} -5 \\ 1 \end{bmatrix}$$

Step 1: Compute state-wise average reward under the policy π

① Find r_{π} for sunny = ?

$$r_{\pi} = 0.5 \times (5) + 0.5 \times (-5) = 2.5 + -2.5 = 0$$

② Find r_{π} for cloudy = ?

$$r_{\pi} = 0.5 \times (3) + 0.5 \times (1) = 1.5 + 0.5 = 2$$

③ Find r_{π} matrix = ?

$$r_{\pi} = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$$

Step 2: Compute the policy transition matrix

Row 1 (Sunny):

- $P_{\pi}(1,1) = ??$

$$P_{\pi}(1,1) = 0.5 \times 0.8 + 0.5 \times 0.9 = 0.4 + 0.45 = 0.85$$

- $P_{\pi}(1,2) = ??$

$$P_{\pi}(1,2) = 0.5 \times 0.2 + 0.5 \times 0.1 = 0.1 + 0.05 = 0.15$$

Row 2 (Cloudy):

- $P_{\pi}(2,1) = ??$

$$P_{\pi}(2,1) = 0.5 \times 0.4 + 0.5 \times 0.3 = 0.2 + 0.15 = 0.35$$

- $P_{\pi}(2,2) = ??$

$$P_{\pi}(2,2) = 0.5 \times 0.6 + 0.5 \times 0.7 = 0.3 + 0.35 = 0.65$$

④ Find $P_{\pi} = \begin{bmatrix} ? & ? \\ ? & ? \end{bmatrix}$

$$P_{\pi} = \begin{bmatrix} 0.85 & 0.15 \\ 0.35 & 0.65 \end{bmatrix}$$

Step 3: Write the Bellman expectation equations $V_{\pi}(\text{sunny})$

General form:

$$V_{\pi}(s) = r_{\pi}(s) + \gamma \sum P_{\pi}(s, s') V_{\pi}(s')$$

5) Find $V_1 = ??$

$$V_1 = 0 + 0.9(0.8V_1 + 0.15V_2)$$

$$V_1 = 0 + 0.765V_1 + 0.135V_2$$

$$V_1 - 0.765V_1 - 0.135V_2 = 0$$

$$0.235V_1 - 0.135V_2 = 0$$

6) Find $V_2 = ??$

$$V_2 = 2 + 0.9(0.9V_1 + 0.65V_2)$$

$$V_2 = 2 + 0.81V_1 + 0.585V_2$$

$$V_2 - 0.585V_2 - 0.81V_1 = 2$$

$$-0.315V_1 + 0.415V_2 = 2$$

Step 4: Solve for $V_{\pi}(\text{cloudy})$

$$0.235V_1 - 0.135V_2 = 0$$

$$\frac{0.235V_1}{0.235} = \frac{0.135V_2}{0.235}$$

$$V_1 = \frac{0.135V_2}{0.235}$$

7) $V_{\pi}(\text{cloudy}) = ??$

$$-0.315V_1 + 0.415V_2 = 2$$

$$= (-0.315 \times \frac{0.135}{0.235} V_2) + 0.415V_2 = 2$$

$$= (-0.315 \times 0.574468 V_2) + 0.415V_2 = 2$$

$$= -0.1809 + 0.415V_2 = 2$$

$$V_{\pi}(\text{cloudy}) = \frac{0.1809 + 2}{0.415} = \frac{2.1809}{0.415} = 5.26$$

8) $V_{\pi}(\text{sunny}) = ??$

$$V_1 = \frac{0.135V_2}{0.235}$$

$$V_1 = \frac{0.135(5.26)}{0.235} = \frac{0.7101}{0.235}$$

$$V_{\pi}(\text{sunny}) = 3.022$$

Step 5: Write the Bellman optimality equations.

$$\text{General form: } V_*(s) = \max_a \{ r(s, a) + \gamma \sum P(s' | s, a) V_*(s') \}$$

Find sunny (V_1) using Go to school:

9) $V_*(\text{sunny}) = ??$

$$V_*(\text{sunny}) = 5 + 0.9(0.8V_1 + 0.2V_2)$$

Find cloudy (V_2) using Go to school:

10) $V_*(\text{cloudy}) = ??$

$$V_*(\text{cloudy}) = 3 + 0.9(0.4V_1 + 0.6V_2)$$

step 5:

$$\text{sunny: } V_1 = 5 + 0.72V_1 + 0.18V_2$$

$$V_1 - 0.72V_1 - 0.18V_2 = 5$$

$$\boxed{0.28V_1 - 0.18V_2 = 5}$$

$$\text{cloudy: } V_2 = 3 + 0.36V_1 + 0.54V_2$$

$$V_2 - 0.54V_2 - 0.36V_1 = 3$$

$$\boxed{-0.36V_1 + 0.46V_2 = 3}$$

step 6: solve for V_* (cloudy)

$$(11) V_*(\text{cloudy}) = ??$$

from the sunny equation:

$$0.28V_1 - 0.18V_2 = 5$$

$$0.28V_1 = 5 + 0.18V_2$$

$$V_1 = \frac{5 + 0.18V_2}{0.28}$$

using the equation for cloudy,

$$-0.36V_1 + 0.46V_2 = 3$$

$$-0.36\left(\frac{5 + 0.18V_2}{0.28}\right) + 0.46V_2 = 3$$

$$-0.36 \times \frac{5}{0.28} = -6.429$$

$$-0.36 \times \frac{0.18}{0.28} = -0.231V_2$$

$$-6.429 - 0.231V_2 + 0.46V_2 = 3$$

$$-6.429 + (0.46 - 0.231)V_2 = 3$$

$$-6.429 + 0.229V_2 = 3$$

$$0.229V_2 = 3 + 6.429 = 9.429$$

$$0.229V_2 = 9.429$$

$$\boxed{V_*(\text{cloudy}) = \frac{9.429}{0.229} = 41.175}$$

Solve for V_* (sunny)

$$(12) V_*(\text{sunny}) = ??$$

$$V_1 = 5 + 0.18V_2$$

$$V_1 = 5 + 0.18V_2 \times 41.175$$

$$V_1 = \frac{5 + 7.41}{0.28} = \frac{12.41}{0.28}$$

$$\boxed{V_*(\text{sunny}) = 44.32}$$

$$\boxed{V_*(\text{cloudy}) = 41.175}$$

step 7. solve for q_*

(13) $q(1, \text{school}) = ??$

$$q(1, \text{school}) = 5 + 0.9(0.8v_1 + 0.2v_2) = 44.32$$

(14) $q(1, \text{home}) = ??$

$$q(1, \text{home}) = -5 + 0.9(0.9v_1 + 0.1v_2) = 34.60$$

(15) $q(2, \text{school}) = ??$

$$q(2, \text{school}) = 3 + 0.9(0.4v_1 + 0.6v_2) = 41.18$$

(16) $q(2, \text{home}) = ??$

$$q(2, \text{home}) = 1 + 0.9(0.3v_1 + 0.7v_2) = 37.91$$

~~(19) $q(1, \text{school}) = ??$~~

~~$q(1, \text{home}) = ??$~~

~~$q(1, \text{school}) = ??$~~

~~$q(1, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~

~~$q(2, \text{home}) = ??$~~

~~$q(2, \text{school}) = ??$~~