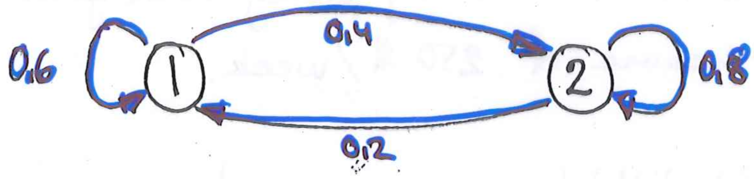
Ex: Consider an example with two states describing the condition of a production unit.

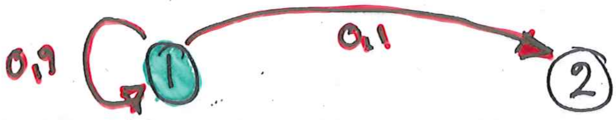
State 1: The unit is working well and generates an income of 400 \$/week.

State 2: The unit is working poorly and generates an income of 250 \$/week.

The transition probabilities are given by

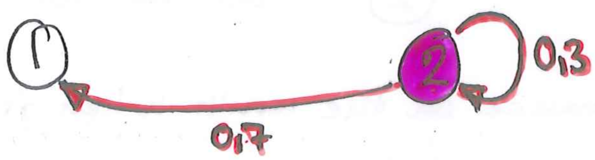


If we do maintanance on the unit when it is working well then the transition probabilities change to



but the cost for this is 50 \$

If we do maintenance on the unit when it is working poorly then the transition probabilities change to



for the cost of 200 \$.

In each state we can decide to do maintenance or not.

We want to determine a decision policy R that to each state determines an action that only depends on the current state.

Let $d(R) = \begin{bmatrix} d_1(R) \\ d_2(R) \end{bmatrix}$

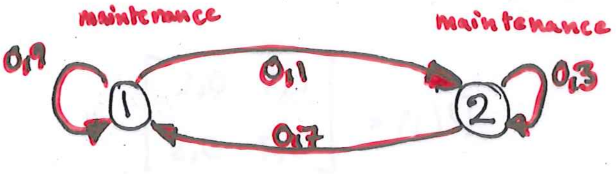
where $d_i(R) = \begin{cases} 1 & \text{if maintenance is done when in state } i \\ 2 & \text{if maintenance is not done when in state } i \end{cases}$

Then, there are 4 different policies for our problem:

$$d(R_1) = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad d(R_2) = \begin{bmatrix} 1 \\ 2 \end{bmatrix} \quad d(R_3) = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad d(R_4) = \begin{bmatrix} 2 \\ 2 \end{bmatrix}$$

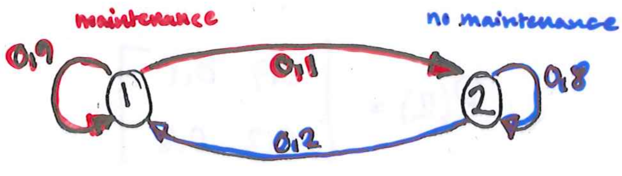
For each of these policies we have different transition probabilities and transition matrices $P(k)$, $k = 1 \dots 4$.

R_1



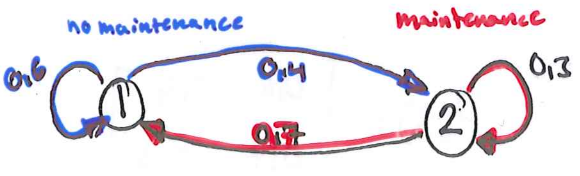
$$P(1) = \begin{bmatrix} 0,9 & 0,1 \\ 0,7 & 0,3 \end{bmatrix}$$

R_2



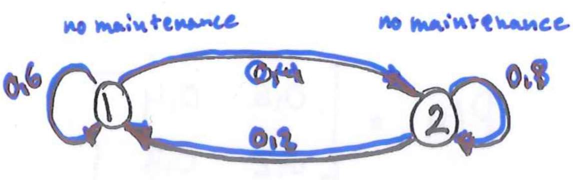
$$P(2) = \begin{bmatrix} 0,9 & 0,1 \\ 0,2 & 0,8 \end{bmatrix}$$

R_3



$$P(3) = \begin{bmatrix} 0,6 & 0,4 \\ 0,7 & 0,3 \end{bmatrix}$$

R_4



$$P(4) = \begin{bmatrix} 0,6 & 0,4 \\ 0,2 & 0,8 \end{bmatrix}$$

At each iteration (period of time = 1 week) there is an income that depends on the state and a cost depending on the decision.

Let C_{ik} = expected value of the immediate cost incurred by making the decision $d_i = k$ when in state i.

Here: $C_{11} = \{$in state 1, maintenance$\} = -400 + 50 = -350$

$C_{12} = \{$in state 1, no maintenance$\} = -400 + 0 = -400$

$C_{21} = \{$in state 2, maintenance$\} = -250 + 200 = -50$

$C_{22} = \{$in state 2, no maintenance$\} = -250 + 0 = -250.$

We want to minimize the expected stationary cost per time unit. Then we need the stationary distributions corresponding to the different decision policies.

From $\pi(k) = P(k)\,\pi(k)$ and $\sum_{i=1}^{2} \pi_i(k) = 1$ we get

$$\pi(1) = \begin{bmatrix} \frac{7}{8} & \frac{1}{8} \end{bmatrix}$$

$$\pi(2) = \begin{bmatrix} \frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

$$\pi(3) = \begin{bmatrix} \frac{7}{11} & \frac{4}{11} \end{bmatrix}$$

$$\pi(4) = \begin{bmatrix} \frac{1}{3} & \frac{2}{3} \end{bmatrix}$$

The stationary expected cost are then given by

$$F(R_k) = \sum_{i=1}^{2} C_{i\,d_i(R_k)} \cdot \pi_i(k) \qquad k = 1,2,3,4.$$

$$F(R_1) = C_{11}\pi_1(1) + C_{21}\pi_2(1) = (-350)\cdot\frac{7}{8} + (-50)\cdot\frac{1}{8} = -312{,}5$$

$$F(R_2) = C_{11}\pi_1(2) + C_{22}\pi_2(2) = (-350)\cdot\frac{2}{3} + (-250)\cdot\frac{1}{3} = -316.7$$

$$F(R_3) = C_{12}\pi_1(3) + C_{21}\pi_2(3) = (-400)\cdot\frac{7}{11} + (-50)\cdot\frac{4}{11} = -272{,}7$$

$$F(R_4) = C_{12}\pi_1(4) + C_{22}\pi_2(4) = (-400)\cdot\frac{1}{3} + (-250)\cdot\frac{2}{3} = -300$$

So policy R_2 is the best

But we were solving the minimization by complete enumeration !