C1) Discrete Time Markov Chain Exercises Solution

2021 Winter RL Study Group

2021-01-10

Contents

Exercise 1

Let's revisit Coke & Pepsi DTMC. Describe state space, transition probability matrix, and initial distribution.

- State space: a set of all possible state that S can take
  S can take Coke or Pepsi, S= {c,p}
- Transition Probability Matrix

$$P_{2,2} = \begin{pmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{pmatrix}$$

- Initial Distribution : The information of where the chain starts at time 0.
  a_0= distribution of S_0 in row vector.
  EX) Suppose you always drink coke on day 0 then,
  $S_0 = c \Leftrightarrow \mathbb{P}(S_0 = c) = 1, \mathbb{P}(S_0 = p) = 0$ a_0=(1,0)
  EX) Suppose you drink coke with probability 0.6 and pepsi with probability 0.4 then,
  $\mathbb{P}(S_0 = c) = 0.6, \mathbb{P}(S_0 = p) = 0.4$ a_0=(0.6,0.4)

Exercise 2

Suppose, $\mathbb{P}(S_0 = c) = 0.6, \mathbb{P}(S_0 = p) = 0.4$ **then, what is** $\mathbb{P}(S_1 = c) = ?$

1)

a_0=(0.6,0.4)

$\mathbb{P}(S_1 = c) = \mathbb{P}(S_1 = c, S_0 = c) + \mathbb{P}(S_1 = c, S_0 = p)$

$=\mathbb{P}(S_1 = c|S_0 = c)\mathbb{P}(S_0 = c) + \mathbb{P}(S_1 = c|S_0 = p)\mathbb{P}(S_0 = p)$

$= 0.7(0.6){+}0.5(0.4) = 0.62$

2)

$$a_o P = \begin{pmatrix}\mathbb{P}(S_0 = c) & \mathbb{P}(S_0 = p)\end{pmatrix} \begin{pmatrix}\mathbb{P}(S_1 = c|S_0 = c) & \mathbb{P}(S_1 = p|S_0 = c) \\ \mathbb{P}(S_1 = c|S_0 = p) & \mathbb{P}(S_1 = p|S_0 = p)\end{pmatrix}$$

$$= \begin{pmatrix}.6 & .4\end{pmatrix} \begin{pmatrix}.7 & .3 \\ .5 & .5\end{pmatrix} = 0.62$$

Exercise 3

Suppose, $\mathbb{P}(S_0 = c) = 0.6, \mathbb{P}(S_0 = p) = 0.4$ **then, what is** $\mathbb{P}(S_2 = c) = ?$

$\mathbb{P}(S_0 = c) = 0.6, \mathbb{P}(S_0 = p) = 0.4$ a_0=(0.6,0.4)

cf) $(AB)C = A(BC)$

$a_1 = a_o P$

$a_2 = a_1 P = (a_0 P)P = a_0 P^2$

$$a_0 P^2 = (.6 \quad .4)\begin{pmatrix} .64 & .36 \\ .6 & .4 \end{pmatrix} = (.624 \quad .376)$$

Thus, $\mathbb{P}(S_2 = c) = .376$

Exercise 4

Suppose, $\mathbb{P}(S_0 = c) = 0.6, \mathbb{P}(S_0 = p) = 0.4$ **then, what is** $\mathbb{P}(S_2 = p) = ?$

$S_0 = p \Leftrightarrow \mathbb{P}(S_0 = c) = 0, \mathbb{P}(S_0 = p) = 1 \Leftrightarrow a_0 = (0, 1)$

cf) $(AB)C = A(BC)$

$a_1 = a_o P$

$a_2 = a_1 P = (a_0 P)P = a_0 P^2$

$$a_0 P^2 = (0 \quad 1) \begin{pmatrix} .64 & .36 \\ .6 & .4 \end{pmatrix} = (.6 \quad .4)$$

Thus, Thus, $\mathbb{P}(S_2 = p) = .4$

Simulating stochastic paths in Python (p.26)

```
import numpy as np

def soda_simul(this_state):
  u=np.random.rand(1)
  if (this_state == "c"):
    if(u<=0.7):
      next_state = "c"
    else:
      next_state = "p"
  else:
    if(u<=0.5):
      next_state = "c"
    else:
      next_state = "p"
  return next_state
for i in range(5):
  path ="c"
  for i in range (9):
    this_state=path[-1]
    next_state=soda_simul(this_state)
    path=path+next_state

  print(path)

## cccccccccc
## ccppcccpcc
## cccpcpcccc
## cccpcppccp
## cccccccppc
```

Simulating stochastic paths in Python (p.26)

```python
def cost_eval(path):
    cost_one_path=path.count("c")*1.5+path.count("p")*1
    return cost_one_path

MC_N=100000
spending_records=np.arange(0,MC_N)
for i in range(MC_N):
    path="c"
    for t in range (9):
        this_state = path[-1]
        next_state = soda_simul(this_state)
        path=path+next_state
    spending_records[i]=cost_eval(path)

print(np.mean(spending_records))
```

```
## 13.11045
```