Assignment 1

79546 - Stephen K. Ng'etich

Contents

Pre-requisite

0.1 Load packages

```
# Clear variables
rm(list=ls())
library(markovchain)
library(diagram)
library(castor)
```

1 Q1

1.1 Generate the transition probabilities

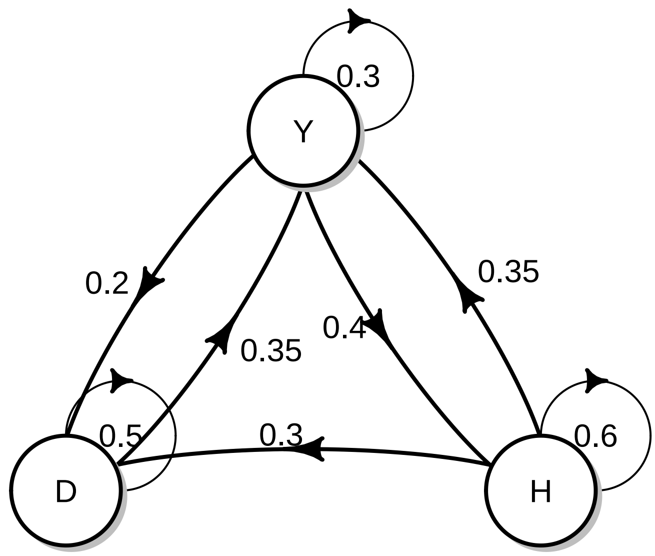
the Markov chain state space $S = H, D, Y$ where H represents Harvard, D for Dartmouth and y for Yale

```
states = c("H", "D", "Y")
trans_mat1 <- matrix(c(.6, 0, .4,
              .3, .5, .2,
              .35, .35, .3), nrow = 3, byrow = TRUE)

rownames(trans_mat1) = states
colnames(trans_mat1) = states
print(trans_mat1)
```

```
##      H    D   Y
## H 0.60 0.00 0.4
## D 0.30 0.50 0.2
## Y 0.35 0.35 0.3
```

1.2 Draw the transition states with their respective probabilities

```
mk_chain1 = new("markovchain",
                 states=states,
                 transitionMatrix=trans_mat1,
                 name="HDY",
                )

plotmat(trans_mat1,relsize = 0.75)
```



1.3 Find the probability that the grandson of a man from Harvard went to Harvard

```
trans_mat2 = trans_mat1 %*% trans_mat1
trans_mat2
```

```
##      H    D    Y
## H 0.50 0.14 0.36
## D 0.40 0.32 0.28
## Y 0.42 0.28 0.30
```

The probability that the grandson of a man from havard went to have is `0.5`

1.4 Modify the above by assuming that the son of a Harvard man always went to Harvard.

```
states = c("H", "D", "Y")
trans_mat3 <- matrix(c(1, 0, 0,
              .3, .5, .2,
              .35, .35, .3), nrow = 3, byrow = TRUE)

rownames(trans_mat3) = states
colnames(trans_mat3) = states
print(trans_mat3)
```

```
##      H    D   Y
## H 1.00 0.00 0.0
## D 0.30 0.50 0.2
## Y 0.35 0.35 0.3
```

1.5 Determine the steady-state probabilities

```
steadyStates(mk_chain1)
```

```
##              H         D         Y
## [1,] 0.4516129 0.2258065 0.3225806
```

2 Q2

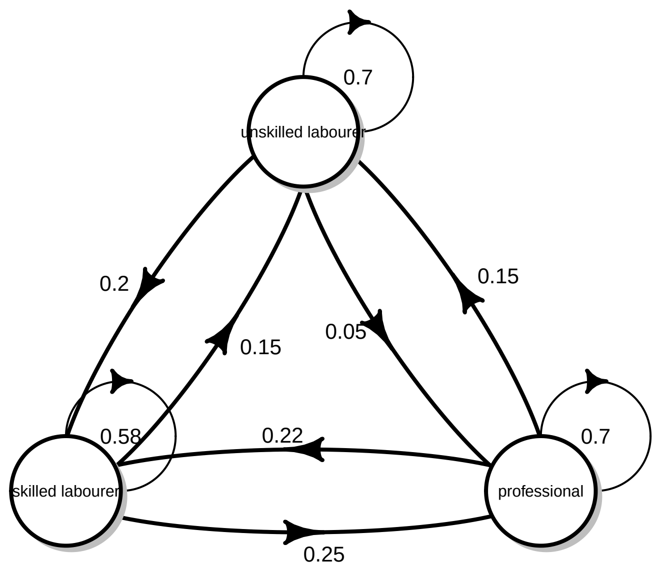
2.1 Set up the matrix of the transistion probabilities

```
states1 = c("professional", "skilled labourer", "unskilled labourer")
trans_mat3 <- matrix(c(.7, .25, .05,
                      .22, .58, .2,
                      .15, .15, .7), nrow = 3, byrow = TRUE)

rownames(trans_mat3) = states1
colnames(trans_mat3) = states1
print(trans_mat3)
```

```
##                    professional skilled labourer unskilled labourer
## professional               0.70             0.25               0.05
## skilled labourer           0.22             0.58               0.20
## unskilled labourer         0.15             0.15               0.70
```

2.2 Draw the transition states with their respective probabilities

```r
plotmat(trans_mat3,relsize = .75,box.cex = 0.54,cex = 0.7)
```



##

Find the probability that a randomly chosen grandson of an unskilled labourer is a professional man.

```r
trans_mat4 = trans_mat3 %*% trans_mat3
print(trans_mat4)
```

```
##                    professional skilled labourer unskilled labourer
## professional             0.5525           0.3275             0.1200
## skilled labourer         0.3116           0.4214             0.2670
## unskilled labourer       0.2430           0.2295             0.5275
```

The probability is `0.243`

2.3 Determine the steady-state probabilities

```
mk_chain2 = new("markovchain",
                 states=states1,
                 transitionMatrix=trans_mat4,
                 name="Labourer",
                )
steadyStates(mk_chain2)
```

```
##      professional skilled labourer unskilled labourer
## [1,]    0.3847695        0.3306613          0.2845691
```

3 Q3

3.1 Is this chain irreducible?

Yes, since one chain can get from one state to any other state in the chain

3.2 Is this chain aperiodic?

Yes, the greatest common divisor when moving from state 1 and back is one

3.3 Find the stationary distribution for this chain.

```
states3 = c("1", "2", "3")
trans_mat5 <- matrix(c(0.5, 0.25, .25,
                       1/3, 0, 2/3,
                       0.5, 0.5, 0), nrow = 3, byrow = TRUE)

rownames(trans_mat5) = states3
colnames(trans_mat5) = states3


stationary_states = get_stationary_distribution(trans_mat5)
print(stationary_states)
```

```
## [1] 0.0000000 0.4615385 0.5384615
```

```
mk_chain3 = new("markovchain",
                 states=states3,
               transitionMatrix=trans_mat5,
                 name="",
              )

# test correctness (steady states* trasnsition matrix should be 0, apart from rounding errors)
cat(sprintf("max(abs(steady states* trasnsition matrix)) = %g\n",max(abs(steadyStates(mk_chain3) %*% tra
```

```
## max(abs(steady states* trasnsition matrix)) = 0.457143
```

```r
#Check if it is a limiting distribution
apply(steadyStates(mk_chain3),1,sum)
```

```
## [1] 1
```

```r
steadyStates(mk_chain3)
```

```
##              1         2         3
## [1,] 0.4571429 0.2571429 0.2857143
```

3.4 Is the stationary distribution a limiting distribution for the chain?

Yes, since the chain is irreducible and aperiodic

4 Q4

4.1 Stock market price data

Yes, Stock Market price is highly volatile

4.2 Appraisal of a Secondary school mathematics teacher

Yes, The outcome of the appraisal is dependent on past performances

4.3 Collaborative filtering on a database of movie reviews: for example, Netflix challenge: predict about how much someone is going to enjoy a movie based on their and other users' movie preferences

No, User rarely change their movie preferences

4.4 Daily weather forecast in Nairobi

Yes, the daily weather

4.5 Optical character recognition

Yes, combination of characters have different generate different meanings.

4.6 Cost of gemstones in Bangladesh

Yes, cost of gemstone change over time due to demand on the local market and international market.

5 Q5

π and A updated based on the following update rules

$$\pi_k = \frac{\gamma(z_{1k})}{\sum_{j=1}^{K} \gamma(Z_{1j})}$$

$$A_{jk} = \frac{\sum_{j=1}^{K} \xi(z_{t-1,j}\, z_{tk})}{\sum_{m=1}^{K} \sum_{t=2}^{K} \xi(z_{t-1,j}\, z_{tm})}$$

given

$$\xi(z_{t-1,j}\, z_t) == \frac{\alpha(z_{t-1}) P(x_t|z_t) P(z_t|z_{t-1} \beta(z_t))}{P(X)}$$

if A_{jk} is 0, then $\xi(z_{t-1,j}\, z_t)$ is also 0, which makes the subsequent updates in the EM algorithm 0