**MSiA 400**

**Lab Assignment 1**

**1) a)** The probability of being in state 5 after 10 transitions from state 1 is: 0.045091.

**b)** The probability of being in state 10 after 10 transitions from states 1,2 and 3 being at equal probability: 0.08268901

**c)** The steady state probability of being in State 1 is 0.01256589.

**d)** Mean passage time from State 1 to State 100: 254.9395 ≈ 255 Transitions.

**2) a)**

Traffic Matrix

|  | **1** | **2** | **3** | **4** | **5** | **6** | **7** | **8** | **9** |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| **1** | 0 | 447 | 553 | 0 | 0 | 0 | 0 | 0 | 0 |
| **2** | 0 | 23 | 230 | 321 | 0 | 0 | 0 | 0 | 63 |
| **3** | 0 | 167 | 43 | 520 | 0 | 0 | 0 | 0 | 96 |
| **4** | 0 | 0 | 0 | 44 | 158 | 312 | 247 | 0 | 124 |
| **5** | 0 | 0 | 0 | 0 | 22 | 52 | 90 | 127 | 218 |
| **6** | 0 | 0 | 0 | 0 | 67 | 21 | 0 | 294 | 97 |
| **7** | 0 | 0 | 0 | 0 | 0 | 94 | 7 | 185 | 58 |
| **8** | 0 | 0 | 0 | 0 | 262 | 0 | 0 | 30 | 344 |
| **9** | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 |

**b)** From the above matrix it can be seen row 9 and column 1 are filled with 0’s.

One step transition probability matrix P.

|  | **1** | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  |  |  |  |  |  |  |  |  |  |
| **1** | 0 | 0.44700000 | 0.55300000 | 0.00000000 | 0.0000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0000000 |
| **2** | 0 | 0.03610675 | 0.36106750 | 0.50392465 | 0.0000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0989011 |
| **3** | 0 | 0.20217918 | 0.05205811 | 0.62953995 | 0.0000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.1162228 |
| **4** | 0 | 0.00000000 | 0.00000000 | 0.04971751 | 0.1785311 | 0.35254237 | 0.27909605 | 0.00000000 | 0.1401130 |
| **5** | 0 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0432220 | 0.10216110 | 0.17681729 | 0.24950884 | 0.4282908 |
| **6** | 0 | 0.00000000 | 0.00000000 | 0.00000000 | 0.1398747 | 0.04384134 | 0.00000000 | 0.61377871 | 0.2025052 |
| **7** | 0 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0000000 | 0.27325581 | 0.02034884 | 0.53779070 | 0.1686047 |
| **8** | 0 | 0.00000000 | 0.00000000 | 0.00000000 | 0.4119497 | 0.00000000 | 0.00000000 | 0.04716981 | 0.5408805 |
| **9** | 1 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0000000 | 0.00000000 | 0.00000000 | 0.00000000 | 0.0000000 |

**c)** Steady state probability vector:

| **Page** | **Steady-state**  **Probability** |
| --- | --- |
|  |  |
| **1** | 0.15832806 |
| **2** | 0.10085497 |
| **3** | 0.13077897 |
| **4** | 0.14012033 |
| **5** | 0.08058898 |
| **6** | 0.07583914 |
| **7** | 0.05446485 |
| **8** | 0.10069664 |
| **9** | 0.15832806 |

**d)** Average time spent on website is mean first passage time from page 1: 14.563 Minutes.

**e)** Variance of Pi before adding the links was 0.00141. After adding the links was 0.00122. The variance slightly decreased, which is a good sign when attempting to distribute the traffic. Looking at a F test for variance comparison, however, shows the p-value is 0.84>0.05, and so the difference may not be statistically significant. In this case, the variance did decrease, showing that adding the links slightly improved traffic distribution.

**Code Appendix**

SabCLab1\_MSiA400

*Sabarish Chockalingam*

*October 31, 2018*

MSiA 400 - Lab Assignment 1

Problem 1

*# install.packages('expm')*

**library**('expm')

## Loading required package: Matrix

##

## Attaching package: 'expm'

## The following object is masked from 'package:Matrix':

##

## expm

*# loading text?file*

markov <- read.table('markov100.txt', header=FALSE)

*#make data into matrix*

P = as.matrix(markov)

*#creating inital vector, in state 1*

a = rep(0,100)

a[1]=1

prob5 = a%\*%(P%^%10)

1 a) probability of being in state 5 after 10 transitions from state 1

*#probability of being in state 5 after 10 transitions from state 1*

prob5[5]

## [1] 0.045091

*# creating different start vector*

a123 = rep(0,100)

a123[1:3]=c(1/3,1/3,1/3)

prob10 = a123%\*%(P%^%10)

1 b) probability of being in state 10 after 10 transitions from states 1,2 and 3 being at equal probability

prob10[10]

## [1] 0.08268901

1 c) Steady state probability of being in state 1

*# calculating steady-state probability of state 1*

*#Subtracting Identity matrix from transpose of P*

Q <- t(P)-diag(ncol(P))

*#Replacing last row with sum(pi) = 1 constraint*

Q[nrow(Q),]=rep(1,ncol(Q))

*#right hand side of Q\*pi*

rhs = c(rep(0,ncol(Q)-1),1)

*#Solving for Steady state probabilities*

Pi = solve(Q)%\*%rhs

Pi[1]

## [1] 0.01256589

*# calculating mean first passage time to 100*

*# removing rows and columns corresponding to state 100*

B = P[-100,-100]

*# subtracting B from Identity matrix*

Q = diag(nrow(B))-B

*# vector of ones*

e = rep(1,nrow(B))

*# calculating mean first passage times*

m = solve(Q)%\*%e

1 d) Mean First passage time from state 1 to state 100

m[1]

## [1] 254.9395

Problem 2

2 a) Contructing Traffic Matrix

*#loading webtraffic.txt*

wb = read.table('webtraffic.txt',header=TRUE)

*# Total Traffic Matrix*

colsumwb = colSums(wb)

Traffic = matrix(colSums(wb),9,9,byrow = TRUE)

*# Displaying Traffic Matrix*

Traffic

## [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9]

## [1,] 0 447 553 0 0 0 0 0 0

## [2,] 0 23 230 321 0 0 0 0 63

## [3,] 0 167 43 520 0 0 0 0 96

## [4,] 0 0 0 44 158 312 247 0 124

## [5,] 0 0 0 0 22 52 90 127 218

## [6,] 0 0 0 0 67 21 0 294 97

## [7,] 0 0 0 0 0 94 7 185 58

## [8,] 0 0 0 0 262 0 0 30 344

## [9,] 0 0 0 0 0 0 0 0 0

2 b) One step probability transition matrix

*# From the above matrix it can be seen row 9 and column 1 are filled with 0's.*

Traffic[9,1] = 1000

*# Constructing one step transi?i?n matrix (divide each element of row by its sum)*

P = Traffic/rowSums(Traffic)

*#Displaying transition matrix*

P

## [,1] [,2] [,3] [,4] [,5] [,6]

## [1,] 0 0.44700000 0.55300000 0.00000000 0.0000000 0.00000000

## [2,] 0 0.03610675 0.36106750 0.50392465 0.0000000 0.00000000

## [3,] 0 0.20217918 0.05205811 0.62953995 0.0000000 0.00000000

## [4,] 0 0.00000000 0.00000000 0.04971751 0.1785311 0.35254237

## [5,] 0 0.00000000 0.00000000 0.00000000 0.0432220 0.10216110

## [6,] 0 0.00000000 0.00000000 0.00000000 0.1398747 0.04384134

## [7,] 0 0.00000000 0.00000000 0.00000000 0.0000000 0.27325581

## [8,] 0 0.00000000 0.00000000 0.00000000 0.4119497 0.00000000

## [9,] 1 0.00000000 0.00000000 0.00000000 0.0000000 0.00000000

## [,7] [,8] [,9]

## [1,] 0.00000000 0.00000000 0.0000000

## [2,] 0.00000000 0.00000000 0.0989011

## [3,] 0.00000000 0.00000000 0.1162228

## [4,] 0.27909605 0.00000000 0.1401130

## [5,] 0.17681729 0.24950884 0.4282908

## [6,] 0.00000000 0.61377871 0.2025052

## [7,] 0.02034884 0.53779070 0.1686047

## [8,] 0.00000000 0.04716981 0.5408805

## [9,] 0.00000000 0.00000000 0.0000000

2 c) Steady State Probability vector Pi

*# calculating steady-state probability vector*

*# Subtracting Identity matrix from transp?s? of P*

Q = t(P)-diag(ncol(P))

*# Replacing last row with sum(pi) = 1 constraint*

Q[nrow(Q),]=rep(1,ncol(Q))

*# right hand side of Q\*pi*

rhs = c(rep(0,ncol(Q)-1),1)

*# Solving for Steady state probabilities*

Pi = solve(Q)%\*%rhs

*# Displaying vector*

Pi

## [,1]

## [1,] 0.15832806

## [2,] 0.10085497

## [3,] 0.13077897

## [4,] 0.14012033

## [5,] 0.08058898

## [6,] 0.07583914

## [7,] 0.05446485

## [8,] 0.10069664

## [9,] 0.15832806

2)d) Average time spent on website

*# calculating mean first passage time to 9 (last page)*

*# removing rows and columns corresponding to page 9*

B = P[-9,-9]

*# subtracting B from Identity matrix*

Q = diag(nrow(B))-B

*# vector of average times*

e = c(0.1,2,3,5,5,3,3,2 )

*# calculating mean first passage times*

m = solve(Q)%\*%e

*# average time spent on website is mean first time passage from page 1*

m[1]

## [1] 14.563

2 e) Checking if adding links to page 6,7 helps uniformly spread traffic.

*#Calculating new ?u?going Traffic %*

dtraff23 = Traffic[2,3]\*0.3

dtraff24 = Traffic[2,4]\*0.2

Traffic2 = Traffic

Traffic2[2,3]=Traffic[2,3]-dtraff23

Traffic2[2,6]=Traffic[2,6]+dtraff23

Traffic2[2,4]=Traffic[2,4]-dtraff24

Traffic2[2,7]=Traffic[2,7]+dtraff24

*# Constructing one?s?ep transition matrix (divide each element of row by its sum)*

P2 = Traffic2/rowSums(Traffic2)

*# calculating steady-state probability vector*

*# Subtracting Identity matrix from transpose of P*

Q2 = t(P2)-diag(ncol(P2))

*# Replacing last row with sum(pi) = 1 ?o?straint*

Q2[nrow(Q2),]=rep(1,ncol(Q2))

*# right hand side of Q\*pi*

rhs2 = c(rep(0,ncol(Q2)-1),1)

*# Solving for Steady state probabilities*

Pi2 = solve(Q2)%\*%rhs2

*# Displaying vector*

Pi2

## [,1]

## [1,] 0.16162840

## [2,] 0.10034341

## [3,] 0.12104331

## [4,] 0.12275720

## [5,] 0.08164613

## [6,] 0.08250884

## [7,] 0.06003218

## [8,] 0.10841213

## [9,] 0.16162840

comparing

*# comparing*

var.test(Pi,Pi2)

##

## F test to compare two variances

##

## data: Pi and Pi2

## F = 1.1567, num df = 8, denom df = 8, p-value = 0.8419

## alternative hypothesis: true ratio of variances is not equal to 1

## 95 percent confidence interval:

## 0.2609065 5.1278020

## sample estimates:

## ratio of variances

## 1.156666

varPi=var(Pi)

varPi2=var(Pi2)

print(paste('Variance of Steady State Vector var(Pi):',varPi))

## [1] "Variance of Steady State Vector var(Pi): 0.00141067501207375"

print(paste('Variance after adding links var(Pi2):',varPi2))

## [1] "Variance after adding links var(Pi2): 0.00121960422368109"