MSiA 400 Lab1

Harish Chockalingam

10/28/2018

#Problem 1a  
data<-read.table('markov100.txt')  
dm<-as.matrix(data)  
state1<-c(1,rep(0,99))  
state5<-state1%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm%\*%dm  
state5[5]

## [1] 0.045091

##The probablity after 10 steps to be in state 5 is 0.045091

#Problem 1b  
data<-read.table('markov100.txt')  
dm1<-as.matrix(data)  
initial\_state=c(1/3,1/3,1/3,integer(97))  
state10<-initial\_state%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1%\*%dm1  
state10[10]

## [1] 0.08268901

##The probablity after 10 steps to be in state 10 is 0.08268901

#Problem 1c  
#Ax=b:   
data<-read.table('markov100.txt')  
dm2<-as.matrix(data)  
A<-t(dm2)-diag(100)  
A[100,]<-c(rep(1,100))  
b<-c(rep(0,99),1)  
x<-solve(A)%\*%b  
x[1]

## [1] 0.01256589

#steady state probablity of state 1 is 0.01256589

#Problem 1d  
data<-read.table('markov100.txt')  
dm3<-as.matrix(data)  
B<-dm3[1:99,1:99]  
Q<-diag(99)-B  
e<-c(rep(1,99))  
m<-solve(Q)%\*%e  
m[1]

## [1] 254.9395

#The mean first passage time to state 100 from state 1, m\_1\_100, is 254.9395

#Problem 2a  
web\_data<-read.table('webtraffic.txt')  
web\_data<-web\_data[-1,]  
web\_data<-apply(as.matrix(web\_data),2,as.numeric)  
Traffic<-matrix(,nrow=9,ncol=9)  
trans\_vector<-colSums(web\_data)  
seq<-c(1,2,3,4,5,6,7,8,9)  
for (i in seq){  
 for (j in seq){  
 Traffic[i,j]<-trans\_vector[j+9\*(i-1)]  
 }  
}  
  
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

#Traffic Matrix shown above

#Problem 2b  
web\_data<-read.table('webtraffic.txt')  
web\_data<-web\_data[-1,]  
web\_data<-apply(as.matrix(web\_data),2,as.numeric)  
Traffic<-matrix(,nrow=9,ncol=9)  
trans\_vector<-colSums(web\_data)  
seq<-c(1,2,3,4,5,6,7,8,9)  
for (i in seq){  
 for (j in seq){  
 Traffic[i,j]<-trans\_vector[j+9\*(i-1)]  
 }  
}  
  
Traffic[9,1]<-1000  
TSM<-Traffic/rowSums(Traffic)  
TSM

## [,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

#One Step Transition Matrix shown above

#Problem 2c  
Q1<-t(TSM)-diag(9)  
Q1[9,]<-c(rep(1,9))  
b1<-c(rep(0,8),1)  
Pi<-solve(Q1)%\*%b1  
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

#The steady state probablity vector Pi is shown above:

#Problem 2d  
B1<-TSM[1:8,1:8]  
Q2<-diag(8)-B1  
e1<-c(0.1,2,3,5,5,3,3,2)  
m1<-solve(Q2)%\*%e1  
m1[1]

## [1] 14.563

#The average time a visitor spends is 14.563 minutes

#Problem 2e  
temp1=0.3\*Traffic[2,3]  
temp2=0.2\*Traffic[2,4]  
Traffic2=Traffic  
Traffic2[2,3]=Traffic[2,3]-temp1  
Traffic2[2,4]=Traffic[2,4]-temp2  
Traffic2[2,6]=Traffic[2,6]+temp1  
Traffic2[2,7]=Traffic[2,7]+temp2  
TSM2<-Traffic2/rowSums(Traffic2)  
  
Q3<-t(TSM2)-diag(9)  
Q3[9,]<-c(rep(1,9))  
b2<-c(rep(0,8),1)  
Pi2<-solve(Q3)%\*%b2  
  
comp<-var.test(Pi,Pi2)  
comp

##   
## 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

var(Pi)

## [,1]  
## [1,] 0.001410675

var(Pi2)

## [,1]  
## [1,] 0.001219604

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

#There is not much variance between Pi and Pi2, and as the p-value is high there not much #significance. Pi has a slightly higher variance than Pi2. We can say that adding the new links #slightly improved the traffic