lab1

Problem1 (a)

```
# read pro distribution
mk <- read.table('Markov100.txt', header=FALSE)
p <- as.matrix(mk)
# set initial vector
a <- rep_len(0, 100)
a[1] = 1
# get prob of being at stage 5 after 10 transitions
library(expm)
```

```
## Loading required package: Matrix
```

```
##
## Attaching package: 'expm'
```

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

```
p_5 <- a %*% (p %^%10)
print(paste0('The probability of being at stage 5 after 10 transitions is: ', p_5[1,5
]))
```

```
## [1] "The probability of being at stage 5 after 10 transitions is: 0.04509099816093
73"
```

Problem1 (b)

```
# start from stage 1
a1 <- rep_len(0, 100)
a1[1] <- 1
p1_10 <- a1 %*% (p %^%10)
# start from stage 2
a2 <- rep_len(0, 100)
a2[2] <- 1
p2_10 <- a2 %*% (p %^%10)
# start from stage 3
a3 <- rep_len(0, 100)
a3[3] <- 1
p3_10 <- a3 %*% (p %^%10)
p_10_b <- (1/3) * (p1_10[1, 10] + p2_10[1, 10] + p3_10[1, 10])
print(paste0('If we start from 1, 2, 3 with equal probability, the probability of bei
ng at stage 5 after 10 transitions is: ', p_10_b))
```

```
## [1] "If we start from 1, 2, 3 with equal probability, the probability of being at
stage 5 after 10 transitions is: 0.0826890133603156"
```

Problem1 (c)

```
Q <- t(p) - diag(100)
Q[100,] <- rep_len(1, 100)
rhs <- rep_len(0, 100)
rhs[100] <- 1
Pi <- solve(Q) %*% rhs
print(paste0('The steady state probability of being in State 1 is: ', Pi[1]))
```

```
## [1] "The steady state probability of being in State 1 is: 0.0125658937530592"
```

Problem1 (d)

```
B <- p[1:99, 1:99]
Q <- diag(99) - B
e <- rep_len(1, 99)
m <- solve(Q) %*% e
print(paste0('The mean first passage time from State 1 to State 100 is: ', m[1]))
```

```
## [1] "The mean first passage time from State 1 to State 100 is: 254.939463100014"
```

Problem2 (a)

```
web <- read.table('webtraffic.txt', header=TRUE)
# traffic count
traffic <- t(matrix(c(colSums(web)), nrow = 9, ncol = 9))
print(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
```

Problem2 (b)

```r
traffic[9,1] = 1000
# construct probability distribution
p <- traffic/rowSums(traffic)
print(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
```

Problem2 (c)

```r
Q <- t(p) - diag(9)
Q[9,] <- c(1, 1, 1, 1, 1, 1, 1, 1, 1)
rhs <- c(0, 0, 0, 0, 0, 0, 0, 0, 1)
Pi <- solve(Q) %*% rhs
print(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
```

Problem2 (d)

```
time <- c(0.1, 2, 3, 5, 5, 3, 3, 2, 0)
# mean passenger time
B <- p[1:8, 1:8]
Q <- diag(8) - B
e <- rep_len(1, 8)
m <- solve(Q) %*% e
page <- m[1]
avg_time <- time %*% Pi
print('the average time a visitor spend on the website')
```

```
## [1] "the average time a visitor spend on the website"
```

```
print(avg_time*page)
```

```
##          [,1]
## [1,] 12.25727
```

Problem2 (e)

```
# new traffic count
traffic2 <- traffic
traffic2[2,6] <- traffic[2,6] + 0.3 * traffic[2,3]
traffic2[2,3] <- 0.7 * traffic[2,3]
traffic2[2,7] <- traffic[2,7] + 0.2 * traffic[2,4]
traffic2[2,4] <- 0.8 * traffic[2,4]
print('New Traffic Count')
```

```
## [1] "New Traffic Count"
```

```
print(traffic2)
```

```
##       [,1] [,2] [,3]  [,4] [,5] [,6]  [,7] [,8] [,9]
##  [1,]    0  447  553   0.0    0    0   0.0    0    0
##  [2,]    0   23  161 256.8    0   69  64.2    0   63
##  [3,]    0  167   43 520.0    0    0   0.0    0   96
##  [4,]    0    0    0  44.0  158  312 247.0    0  124
##  [5,]    0    0    0   0.0   22   52  90.0  127  218
##  [6,]    0    0    0   0.0   67   21   0.0  294   97
##  [7,]    0    0    0   0.0    0   94   7.0  185   58
##  [8,]    0    0    0   0.0  262    0   0.0   30  344
##  [9,] 1000    0    0   0.0    0    0   0.0    0    0
```

```
# construct new probability distribution
p2 <- traffic2/rowSums(traffic2)
print('New Probability Distribution')
```

```
## [1] "New Probability Distribution"
```

```
print(p2)
```

```
##       [,1]       [,2]       [,3]       [,4]      [,5]       [,6]
##  [1,]    0 0.44700000 0.55300000 0.00000000 0.0000000 0.00000000
##  [2,]    0 0.03610675 0.25274725 0.40313972 0.0000000 0.10832025
##  [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,] 0.10078493 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
```

```
# new Pi2
Q2 <- t(p2) - diag(9)
Q2[9,] <- c(1, 1, 1, 1, 1, 1, 1, 1, 1)
rhs2 <- c(0, 0, 0, 0, 0, 0, 0, 0, 1)
Pi2 <- solve(Q2) %*% rhs2
print('New steady state probability vector Pi2')
```

```
## [1] "New steady state probability vector Pi2"
```

```
print(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
```

```
# compare variance
var1 <- var(Pi)
var2 <- var(Pi2)
print(paste0('variance before: ', var1))
```

```
## [1] "variance before: 0.00141067501207375"
```

```
print(paste0('variance after: ', var2))
```

```
## [1] "variance after: 0.00121960422368109"
```

The variance of Pi2 is smaller than variance of Pi, so the link helped balancing the traffic.