Exercise 7: Markov Chains

Ruth Walters

```r
library(Matrix)
```

Part 1

```r
Q <- matrix(0,2,2)
Q
```

```
     [,1] [,2]
[1,]    0    0
[2,]    0    0
```

```r
Q[1,2] <- 5
Q
```

```
     [,1] [,2]
[1,]    0    5
[2,]    0    0
```

```r
today <- c(1,0)
names(today) <- c("Sunny", "Rainy")

rates <- c(0, 0.7635756,
           0.1527151, 0)

Q <- matrix(rates, 2, 2)

Q
```

```
          [,1]        [,2]
[1,] 0.0000000 0.1527151
[2,] 0.7635756 0.0000000
```

```
rownames(Q) <- colnames(Q) <- c("Sunny", "Rainy")

diag(Q) <- -rowSums(Q)

Q
```

```
           Sunny       Rainy
Sunny -0.1527151   0.1527151
Rainy  0.7635756  -0.7635756
```

Question 1:

$r_{\text{sunny}\rightarrow\text{rainy}} = 0.1527151$

$r_{\text{rainy}\rightarrow\text{sunny}} = 0.7635756$

Part 2

```
time <- 1

P_t <- expm(t(Q)*time) %*% today

P_t
```

```
2 x 1 Matrix of class "dgeMatrix"
            [,1]
Sunny 0.90000001
Rainy 0.09999999
```

Question 3: $P(\text{sunny} \mid t = 1) = 0.90000001$

```
time <- 2

P_t <- expm(t(Q)*time) %*% today

P_t
```

```
2 x 1 Matrix of class "dgeMatrix"
      [,1]
Sunny 0.86
Rainy 0.14
```

Question 4: $P(\text{rain} \mid t = 2) = 0.14$

```r
times <- c(10,100,1000)


for (i in times) {
  P_t <- expm(t(Q)*i) %*% today

  print(paste("t =", i))
  print(P_t)
}
```

```
[1] "t = 10"
2 x 1 Matrix of class "dgeMatrix"
           [,1]
Sunny 0.8333508
Rainy 0.1666492
[1] "t = 100"
2 x 1 Matrix of class "dgeMatrix"
           [,1]
Sunny 0.8333334
Rainy 0.1666666
[1] "t = 1000"
2 x 1 Matrix of class "dgeMatrix"
           [,1]
Sunny 0.8333334
Rainy 0.1666666
```

Question 5: In the long run, the probability of having a sunny day stabilizes to ~0.83, meaning that we expect to have a sunny day ~83% of the time, or we expect 83% of days to be sunny.

Part 3

```r
nucleotides <- c("A", "C", "G", "T")
```

```r
rates <- c(0,
           0.2,
           0.4,
           0.2,
           0.3,
           0,
           0.3,
           0.9,
           0.6,
           0.3,
           0,
           0.3,
           0.2,
           0.6,
           0.2,
           0)

Q <- matrix(rates, 4, 4)

diag(Q) <- -rowSums(Q)

rownames(Q) <- colnames(Q) <- nucleotides

Q
```

```
     A    C    G    T
A -1.1  0.3  0.6  0.2
C  0.2 -1.1  0.3  0.6
G  0.4  0.3 -0.9  0.2
T  0.2  0.9  0.3 -1.4
```

```r
Jer.obs <- c(1, 0, 0, 0)

names(Jer.obs) <- nucleotides

Jer.DNA <- expm(Q*0.3) %*% Jer.obs

Jer.DNA
```

```
4 x 1 Matrix of class "dgeMatrix"
       [,1]
```

```
A 0.73074054
C 0.05183636
G 0.09311238
T 0.05183636
```

```
Keyrock.obs <- c(0, 1, 0, 0)

names(Keyrock.obs) <- nucleotides

Keyrock.DNA <- expm(Q*0.3) %*% Keyrock.obs

Keyrock.DNA
```

```
4 x 1 Matrix of class "dgeMatrix"
          [,1]
A 0.07775453
C 0.74177012
G 0.07775453
T 0.19295848
```

```
combined.probs <- Jer.DNA * Keyrock.DNA

combined.probs
```

```
4 x 1 Matrix of class "dgeMatrix"
          [,1]
A 0.05681839
C 0.03845066
G 0.00723991
T 0.01000226
```

```
sum(combined.probs)
```

```
[1] 0.1125112
```

Question 6: The overall probability is 0.1125112.

Question 7: The likeliest nucleotide state is an A, which means that your brother is the mutant, not you.