

Math 578B – Fall 2015 – Homework #1

Peter Ralph

due 1 September

1. Let X be a Markov chain on $\mathcal{A} = \{0, 1\}$ with transition probabilities

$$\mathbb{P}\{X_{t+1} = 1 \mid X_t = 0\} = \alpha \quad (1)$$

$$\mathbb{P}\{X_{t+1} = 0 \mid X_t = 1\} = \beta. \quad (2)$$

- a. Show that if $X_0 = 0$ then the distribution of

$$\tau := \min\{n > 0 : X_n = 1\} \quad (3)$$

is geometric.

Solution:

By the Markov property, and since $X_0 = 0$,

$$\begin{aligned} \mathbb{P}\{\tau > n\} &= \mathbb{P}\{X_0 = X_1 = X_2 = X_3 = \cdots = X_n = 0\} \\ &= \mathbb{P}\{X_0 = 0\} \mathbb{P}\{X_1 = 0 \mid X_0 = 0\} \cdots \mathbb{P}\{X_n = 0 \mid X_{n-1} = 0\} \\ &= (1 - \alpha)^n. \end{aligned}$$

This is the expression for the geometric distribution.

- b. Find the spectral decomposition of the chain and use it to find an expression for

$$\mathbb{P}\{X_n = 1 \mid X_0 = 0\}. \quad (4)$$

Solution: *We know that one eigenvalue is $\lambda_1 = 1$; since $\text{tr}P = 2 - \alpha - \beta$, the other must be $\lambda_2 = 1 - \alpha - \beta$.*

A spectral decomposition for the chain is

$$\begin{aligned}
P^n &= \begin{bmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{bmatrix}^n \\
&= \begin{bmatrix} 1 & -\alpha/(\alpha+\beta) \\ 1 & \beta/(\alpha+\beta) \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (1-\alpha-\beta)^n \end{bmatrix} \begin{bmatrix} \beta/(\alpha+\beta) & \alpha/(\alpha+\beta) \\ -1 & 1 \end{bmatrix}
\end{aligned}$$

Using this,

$$\begin{aligned}
\mathbb{P}\{X_n = 1 \mid X_0 = 0\} &= (P^n)_{01} \\
&= \frac{\alpha}{\alpha+\beta} (1 - (1-\alpha-\beta)^n)
\end{aligned}$$

- c. What happens to the spectral expansion if $\alpha + \beta = 1$? Explain.

Solution: Note in the previous answer that the transition probability converges geometrically to the stationary distribution, $\alpha/(\alpha+\beta)$, with a rate that depends on $\lambda_2 = (1-\alpha-\beta)$. But, if $\lambda_2 = 0$, the chain reaches stationarity immediately:

$$\mathbb{P}\{X_n = 1 \mid X_0 = 0\} = \frac{\alpha}{\alpha+\beta}.$$

If $\alpha + \beta = 1$, then

$$P = \begin{bmatrix} 1-\alpha & \alpha \\ 1-\alpha & \alpha \end{bmatrix},$$

– in other words, the distribution of the next state doesn't depend on the current state – the Markov chain is really just a sequence of iid random variables.

2. For the Markov chain in (1), suppose we choose a random starting state with $\mathbb{P}\{X_1 = 0\} = \beta/(\alpha+\beta)$, and run it for n steps, producing a string of n digits, $X = X_1, \dots, X_n$. Let Y be the reversed string, i.e., $Y_k = X_{n+1-k}$ for $1 \leq k \leq n$.

- a. For a given string of digits $a_1, \dots, a_n \in \{0, 1\}^n$, what is $\mathbb{P}\{Y_1 = a_1, \dots, Y_n = a_n\}$? (Don't use the spectral decomposition.)

Solution: Count the transitions:

$$\begin{aligned}
n_{00}(a) &= \#\{1 \leq k \leq n-1 : a_k = 0 \& a_{k+1} = 0\} \\
n_{01}(a) &= \#\{1 \leq k \leq n-1 : a_k = 0 \& a_{k+1} = 1\} \\
n_{10}(a) &= \#\{1 \leq k \leq n-1 : a_k = 1 \& a_{k+1} = 0\} \\
n_{11}(a) &= \#\{1 \leq k \leq n-1 : a_k = 1 \& a_{k+1} = 1\}.
\end{aligned}$$

By the Markov property, and then the definition of X ,

$$\begin{aligned}
\mathbb{P}\{Y_1 = a_1, \dots, Y_n = a_n\} &= \mathbb{P}\{X_1 = a_n, \dots, X_n = a_1\} \\
&= \mathbb{P}\{X_1 = a_n\} \prod_{k=1}^{n-1} \mathbb{P}\{X_{k+1} = a_{n-k} \mid X_k = a_{n-k+1}\} \\
&= \frac{\alpha^{1-a_n} + \beta^{a_n}}{\alpha + \beta} (1 - \alpha)^{n_{00}(a)} \alpha^{n_{10}(a)} (1 - \beta)^{n_{11}(a)} \beta^{n_{01}(a)} \\
&= \frac{1}{\alpha + \beta} (1 - \alpha)^{n_{00}(a)} \alpha^{n_{10}(a) + \delta_0(a_n)} (1 - \beta)^{n_{11}(a) + \delta_1(a_n)} \beta^{n_{01}(a)},
\end{aligned}$$

where $\delta_0(a_n) = 1$ if $a_n = 0$ and $\delta_0(a_n) = 0$ otherwise.

- b. Suppose we generate a string of length n as in (a) but it is reversed with probability $1/2$: formally, let $\theta = H$ with probability $1/2$ and $\theta = T$ otherwise, and define

$$Z = \begin{cases} X & \text{if } \theta = H \\ Y & \text{otherwise.} \end{cases} \quad (5)$$

A colleague has developed a procedure that guesses whether the string was reversed, i.e., a function $f : \{0, 1\}^n \rightarrow \{H, T\}$ that is given Z and guesses the value of θ . Show that

$$\mathbb{P}\{f(Z) = \theta\} = 1/2. \quad (6)$$

Solution: We know from the first problem that the stationary distribution is $\pi_0 = \beta/(\alpha + \beta)$; and since $\pi_0 P_{01} = \alpha\beta/(\alpha + \beta) = \pi_1 P_{10}$, the chain is reversible, and since $X_1 \sim \pi$, we know Y has the same distribution as X . But, part of the point of this problem is to **prove** this. It's easy to check that

$$\mathbb{P}\{X_1 = a_1, \dots, X_n = a_n\} = \frac{1}{\alpha + \beta} (1 - \alpha)^{n_{00}(a)} \alpha^{n_{01}(a) + \delta_0(a_1)} (1 - \beta)^{n_{11}(a) + \delta_1(a_1)} \beta^{n_{10}(a)}.$$

However: notice that $n_{01}(a) + \delta_0(a_1) = n_{10}(a) + \delta_0(a_n)$, and $n_{10}(a) + \delta_1(a_1) = n_{01}(a) + \delta_1(a_n)$, and so $\mathbb{P}\{X = a\} = \mathbb{P}\{Y = a\}$.

So, we know that X and Y have the same distribution; so how could we possibly infer what θ is? We still need to make the intuition formal: since on $\{\theta = H\}$, $Z = X$, and on $\{\theta = T\}$, $Z = Y$, we know that $f(Z) \mid \theta = H$ and $f(Z) \mid \theta = T$ have the same distribution; therefore, $\mathbb{P}\{f(Z) = H \mid \theta = H\} = \mathbb{P}\{f(Z) = H \mid \theta = T\}$. Since $\mathbb{P}\{f(Z) = H \mid \theta = H\} + \mathbb{P}\{f(Z) = T \mid \theta = H\} = 1$,

$$\begin{aligned}
\mathbb{P}\{f(Z) = \theta\} &= \mathbb{P}\{f(Z) = H \mid \theta = H\}/2 + \mathbb{P}\{f(Z) = T \mid \theta = T\}/2 \\
&= \mathbb{P}\{f(Z) = H \mid \theta = H\}/2 + \mathbb{P}\{f(Z) = T \mid \theta = H\}/2 \\
&= 1/2.
\end{aligned}$$

3. Let X_n be the simplified “polymerase complex assembly” Markov chain defined in class, with transition matrix (where “†” means transcription):

$$P = \begin{array}{c} \emptyset \\ \alpha \\ \beta \\ \alpha + \beta \\ \text{pol} \\ \dagger \end{array} \begin{bmatrix} \emptyset & \alpha & \beta & \alpha + \beta & \text{pol} & \dagger \\ * & k_\alpha & k_\beta & 0 & 0 & 0 \\ k_\alpha & * & 0 & k_\beta & 0 & 0 \\ k_\beta & 0 & * & k_\alpha & 0 & 0 \\ 0 & k_\beta & k_\alpha & * & k_{\text{pol}} & 0 \\ 0 & 0 & 0 & 0 & * & 1 \\ 0 & 0 & 0 & 1 & 0 & * \end{bmatrix} \quad (7)$$

Here the “*”s on the diagonal are set so that rows sum to 1. Let

$$\tau = \min\{n \geq 0 : X_n = \dagger\}. \quad (8)$$

- a. Set $k_\alpha = k_\beta = 0.2$ and $k_{\text{pol}} = 0.5$, and compute numerically $\mathbb{E}[\tau]$ for all starting states. (And, explain how you do it!)

Solution:

```

ka <- kb <- 0.2
kpol <- 0.5
M <- matrix( c(
  0, ka, kb, 0, 0, 0,
  ka, 0, 0, kb, 0, 0,
  kb, 0, 0, ka, 0, 0,
  0, kb, ka, 0, kpol, 0,
  0, 0, 0, 0, 0, 1,
  0, 0, 0, 1, 0, 0
), byrow=TRUE, nrow=6 )
diag(M) <- 1-rowSums(M)

```

Now, since $\mathbb{E}^a[\tau] = h$ solves $(I - Q)^{-1}1$ where Q is

```
M[1:5, 1:5]
```

```
##      [,1] [,2] [,3] [,4] [,5]
## [1,]  0.6  0.2  0.2  0.0  0.0
## [2,]  0.2  0.6  0.0  0.2  0.0
## [3,]  0.2  0.0  0.6  0.2  0.0
## [4,]  0.0  0.2  0.2  0.1  0.5
## [5,]  0.0  0.0  0.0  0.0  0.0
```

the solution is

```
h <- solve( diag(5) - M[1:5,1:5], rep(1,5) )
h

## [1] 19.0 16.5 16.5  9.0  1.0
```

- b. Compute numerically the stationary distribution. What is the long-term average rate of transcription (i.e., mean number of visits to \dagger per unit time)?

Solution: The stationary distribution π is the first left eigenvector of P :

```
pivec <- eigen(t(M))$vectors[,1]
pivec <- Re( zapsmall(pivec)/sum(pivec) )
pivec

## [1] 0.2 0.2 0.2 0.2 0.1 0.1
```

and the mean time between visits to \dagger is 9.9999983.

- c. Verify your computations in (a) and (b) using a simulation of the chain.

Solution:

First, simulate the chain (many in parallel for efficiency):

```
nchains <- 1e4
nsteps <- 300
step.probs <- lapply(1:nrow(M), function(k) {
  list(
    states=which(M[k,]>0),
    probs=c(0,cumsum(M[k,M[k,]>0]))
  )
})
mcruns <- matrix(0, nrow=nchains, ncol=nsteps)
# uniform initial state:
mcruns[,1] <- sample.int(nrow(M), nchains, replace=TRUE)
for (k in 2:nsteps) {
  for (state in 1:nrow(M)) {
    dothese <- which(mcruns[,k-1]==state)
```

```

        newstate <- findInterval( runif(length(dothese)), step.probs[[state]]$probs )
        mcruns[dothese,k] <- step.probs[[state]]$states[newstate]
    }
}

```

Now, let's find and compare the first hitting times to † (state 6):

```

t.dagger <- apply( mcruns==6, 1, function (x) {
    min(which(x))-1
} )
rbind(
    theory=c(h,0),
    simulation=tapply(t.dagger,mcruns[,1],mean),
    SE=tapply(t.dagger,mcruns[,1],sd)/sqrt(tapply(t.dagger,mcruns[,1],length))
)

```

##		1	2	3	4	5	6
## theory		19.000000	16.500000	16.500000	9.000000	1	0
## simulation		20.257194	16.598820	17.292817	8.584115	1	0
## SE		0.401897	0.357758	0.390160	0.295833	0	0

The simulation is within $\pm 2SE$ of the theory.

And, for the stationary probability, we'll use the last 100 steps in each chain:

```

est.pi <- table( mcruns[,200:300] )/(nchains*100)
rbind(
    theory=pivec,
    simulation=est.pi,
    SE=sqrt(pivec*(1-pivec)/(nchains*100))
)

```

##		1	2	3	4	5	6
## theory		0.200000	0.200000	0.200000	0.200000	0.100000	0.100000
## simulation		0.203404	0.20244	0.201061	0.201802	0.100646	0.100647
## SE		0.000400	0.00040	0.000400	0.000400	0.000300	0.000300

The SE in this case is not right, since the observations aren't independent, but it's still a useful point of comparison.