Simulating Discrete Markov Chains

University of Toronto, ACT 350

Prof. S. Pesenti

27/10/2020

{r packages, include=FALSE} # package for calculating matrix exponentials # you might need to install this the first time you run this code. # install.packages(expm) library("expm")

An introductory example

Assume a student has lectures every day at the University of Toronto and is either punctual or late. The student is late with probability 0.3, when he was late the day before, and late with probability 0.4, when he was punctual the day before. Being late or punctual two days before does not influence his punctuality.

1. Defining the Markov Chain

- state space: $S = \{l, p\}$, l means late, p means punctual

- $\{X_n,\ n = 0, 1, \dots\}$, where the event $\{X_n = l\}$ means that the student is late at day n.

- fulfills the Markov property since the event that the student is late or punctual only depends whether the student was late or punctual the day before. It holds for $n \geq 0$

$$
\begin{aligned}
p_{l,l} &= P(X_{n+1} = l | X_n = l) = 0.3 \\
p_{p,l} &= P(X_{n+1} = l | X_n = p) = 0.4 \\
p_{l,p} &= P(X_{n+1} = p | X_n = l) = 1 - P(X_{n+1} = l | X_n = l) = 0.7 \\
p_{p,p} &= P(X_{n+1} = p | X_n = p) = 1 - P(X_{n+1} = l | X_n = p) = 0.6
\end{aligned}
$$

- transition matrix $P = \begin{pmatrix} p_{l,l} & p_{l,p} \\ p_{p,l} & p_{p,p} \end{pmatrix} = \begin{pmatrix} 0.3 & 0.7 \\ 0.4 & 0.6 \end{pmatrix}$.

Numerical implementation: R code

```
# initialise the transition matirx
P <-  matrix(c(0.3, 0.4, 0.7, 0.6), nrow = 2)
rownames(P) <- c("l","p")
colnames(P) <- c("l","p")
# note / check that all rows sum up to 1
P
rowSums(P)
```

2. Transition probabilities and stationary distribution

Question: What are the transition probabilities that the student is late or punctual after two days?

Question: How about after the 10^{th} day?

Question: How about after the 50^{th} day?

For this, we have to calculate the transition probabilities for day 2 ($P * P$), and for day 10 (P^{10}) and for day 50 (P^{50}).

```
# transition probability after 2 days
P %*% P
# note / check that all rows sum up to 1
rowSums(P)
```

Question: What is the interpretation of `r (P %*% P)[1,1]`?

```
# transition probability after 10 days
P %^% 10

# transition probability after 50 days
P %^% 50
```

Question: What do you observe?

We see numerically that the rows of the transition matrices always add up to 1. Further, we observe that the transition matrix after 10 days and after 50 days are numerically equal. Let $t = 0$ correspond to the induction day and $t = 1$ to the first day of lectures. Then we have that, if the student was late at the induction day, he will be late at the end of the semester with probability `r (P %^% 50)[1,1]`, independent of whether he was late or not during the term. Recall that $p_{ll}^{50} = P(X_{50} = l | X_0 = l)$ is the probability that the student is late on the 50th day, given he was late at the induction day. Numerically, we observe that $p_{ll}^{10} = p_{ll}^{11} = \cdots = p_{ll}^{50} = \cdots p_{ll}^{n} = \cdots$ for all n large. Thus, the probability of being late on a day later in the semester, is independent whether the student was late at induction day or not.

Question: Does this mean that (`r (P %^% 50)[1,1]`, `r (P %^% 50)[1,2]`) is a stationary distribution for the Markov Chain?

Question: What properties of the Markov Chain imply that the Markov Chain has a *stationary distribution*?

Question: How many stationary distributions does the Markov Chain have?

Question: Calculate analytically the stationary distribution.

```
{r, stationary distribution} # the analytical stationary distribtuion is pi <- c(4/11,
7/11) names(pi) <- c("l", "p") # we veryfy that pi is a stationary distribtuion, that is
pi * P = pi t(pi) %*% P pi all.equal(t(pi) %*% P, t(pi))
```

3. Initial distribution

Let us look at students whose transition probabilities are described by P but have different probabilities to be punctual at induction day ($t = 0$).

- *student 1*: probabilities at induction day $(0.5, 0.5)$ (random student)
- *student 2*: probabilities at induction day $(0.2, 0.8)$ (good student)
- *student 3*: probabilities at induction day $(0.7, 0.3)$ (bad student)

student 1 "'{r, student1} # probability of student 1 at induction day
student1_x0 <- c(0.5, 0.5)

probability at the first day

student1_x0 %*% P

probability at the 5th day

student1_x0 %*% (P %^% 5)

probability at the 100th day

student1_x0 %*% (P %^% 100)

```
Observe that *student 1* has probability `r student1_x0[2]` to be punctual at induction day. On the firs

  **student 2**
  Now let us look at the "good" *student 2* who started with probability of $0.8$ to be punctual at indu
```{r, student2}
# probability of student 2 at induction day
student2_x0 <- c(0.2, 0.8)

# probability at the first day
student2_x0 %*% P

# probability at the 5th day
student2_x0 %*% (P %^% 5)

# probability at the 100th day
student2_x0 %*% (P %^% 100)
```

student 3
How about *student 3* who only has a probability of 0.3 to turn up punctual at induction day? Will he improve over the semester?
"'{r, student3} # probability of student 3 at induction day student3_x0 <- c(0.7, 0.3)

probability at the first day

student3_x0 %*% P

probability at the 5th day

student3_x0 %*% (P %^% 5)

probability at the 100th day

student3_x0 %*% (P %^% 100)

```
We observe that at the end of the semester, *student 1* (random student), *student 2* (good student) and

**Question:** Does this hold for any student, whose transition probabilities are described by $P$? That

Numerically, and also analytically (**Why?**), it holds that
```
```

```
\begin{equation}
\lim_{n \to \infty} (\nu, 1 - \nu) * P^n =
(\nu, 1 - \nu) *
\begin{pmatrix}
  0.3636364 & 0.6363636\\
  0.3636364 & 0.6363636
\end{pmatrix}
 = (`r (P %^% 50)[1,1]`, `r (P %^% 50)[1,2]`),
\end{equation}
for any initial distribution $\nu, 1-\nu, ~ \nu \in [0,1]$ on the state space $S = \{l, p\}$.


## 4. Sample path
So far we only looked how the probabilities of being punctual or late changes from day to day or over th

**Simulating a sample path:**


1. Set an initial value for $x_0$. (Is the student punctual or late at induction day?) To be able to us
2. for $n \geq 1$ do:
      a) calculate $x_{0} * P^n$ (the distribution of the student to be late/punctual on the $n^{th}$ da
      c) generate a realisation from a random variable with distribution $x_{0} * P^n$
      d) set the outcome of the random variable equal to $x_n$

The collection $x_0, x_1, \ldots$ is the sample path of the student.


**A simulated sample path: R code**
For simplicity, denote $l = 1$ and $p = 0$.
Choose $x_0 = (1,0)$, that is the student is late on the first day.
```{r, simulated sample path}
# set the seed for generating random variables (for reproducability)
# set.seed(2019)
# number of days (length of simulated sample path)
n <- 50
# initialise the sample path
sample_path <- rep(0, lenthought = n)
# initialise the probability that the student is late
late <- rep(0, lenthought = n)

# initial value at induction day
x0 = c(1, 0)
# probability that the student is late on the first day
late[1] <- (x0 %*% P)[1]

for(i in 2:n){
  # probability that the student is late at day n
  late[i] <- (x0 %*% (P %^% i))[1]
  # genearte a Bernoulli random variable with probability = late
  sample_path[i] <- rbinom(1, 1, late[i])
}
```

The first plot shows the probability that the student is late over the next 50 days. The second plot shows the generated sample path, that displays whether the student is late (1) or punctual (0) for the next 50 day.
```

“‘ {r, plot_sample_path, echo = FALSE, fig.show=‘hold’, out.width = ‘50%’} # par(mfrow=c(1,2)) plot(seq(1,n), late, xlab = “number of days”, ylab = “probability of being late”)

plot(seq(1, n), sample_path, xlab = “number of days”, ylab = " ") “‘

Question If you rerun the above code to simulate a sample path, you should always get something different. Have a close look at the figure to the right. The figure on the left will always look very similar. Why?

Question Rerun the above code to simulate a sample path changing n, the length of the sample path, and/or the initial value.

Question You can also rerun the whole document using different transition probabilities. (However, the text might be sometimes inconsistent. . .)