Lecture 11
Markov Chains

Nils Detering (covered by Maryann Hohn)

February 14th

Markov Chains

To get started, let us first revisit our Random Walk: Recall

- $X_1, X_2, \ldots$ be a sequence of i.i.d. RVs
- S_0 a RV indep. of $X_1, X_2, \ldots$. Sequence $S_0, S_1, \ldots$ defined by

$$S_n = S_0 + \sum_{i=1}^{n} X_i \qquad (n = 0, 1, , \ldots)$$

Assume we are interested in the probability of $S_{n+1} = k$ and we know the value of S_n.
How much would we gain by knowing also $S_0, S_1, \ldots$? Or, more formally, what is the difference between $\mathbb{P}(S_{n+1} = k | S_n = j_n)$ and $\mathbb{P}(S_{n+1} = k | S_0 = j_0, \ldots, S_n = j_n)$?

Markov Chains

Let us calculate:

$$\begin{aligned}
\mathbb{P}(S_{n+1} = k | S_n = j_n) &= \frac{\mathbb{P}(S_{n+1} = k, S_n = j_n)}{\mathbb{P}(S_n = j_n)} \\
&= \frac{\mathbb{P}(S_n + X_{n+1} = k, S_n = j_n)}{\mathbb{P}(S_n = j_n)} \\
&= \frac{\mathbb{P}(X_{n+1} = k - j, S_n = j)}{\mathbb{P}(S_n = j_n)} \\
&= \frac{\mathbb{P}(X_{n+1} = k - j)\mathbb{P}(S_n = j)}{\mathbb{P}(S_n = j_n)} \\
&= \mathbb{P}(X_{n+1} = k - j)
\end{aligned}$$

and we know that $\mathbb{P}(X_{n+1} = k - j)$ equals p for $k - j = 1$ and $1 - p$ for $k - j = -1$.

Markov Chains

Ok, lets then calculate:

$$
\begin{aligned}
& \mathbb{P}(S_{n+1} = k | S_0 = j_0, \ldots, S_n = j_n) \\
= \quad & \frac{\mathbb{P}(S_{n+1} = k, S_0 = j_0, \ldots, S_n = j_n)}{\mathbb{P}(S_0 = j_0, \ldots, S_n = j_n)} \\
= \quad & \ldots \\
= \quad & \frac{\mathbb{P}(X_{n+1} = k - j_n)\mathbb{P}(X_n = j_n - j_{n-1}) \cdot \cdots \cdot \mathbb{P}(X_1 = j_1 - j_0)}{\mathbb{P}(X_n = j_n - j_{n-1}) \cdot \cdots \cdot \mathbb{P}(X_1 = j_1 - j_0)} \\
= \quad & \mathbb{P}(X_{n+1} = k - j_n)
\end{aligned}
$$

as before! Adding information about the past does not add anything once you know the present!

Markov Chains

- Many time-discrete stochastic processes have the following very pleasant Markov-property

  The future given the present is independent of the past

- This property helps to understand questions like:
  - Does the distribution $\mathbb{P}(S_n = k)$ converge for $n \to \infty$?
  - Are there states k such that if $S_n = k$, then $S_j = k$ for any $j \geq n$?
  - ...

Markov Chains

- We will spend the rest of this term on Markov processes.
- We follow now closely the book *Stochastic Processes with R* by Robert Dobrow.
- See syllabus on Gauchospace for references for every lecture.

We need to do a change in notation to be in line with Markov Chain literature and our main reference!

From now on, the stochastic process (Markov Chain) we consider will be denoted by $X_0, X_1, \ldots$ not its increments.

Markov Chains

Definition (Markov Chains)

Let $\mathcal{S} \subset \mathbb{R}$ a discrete set. A Markov chain is a sequence of random variables $X_0, X_1, \ldots$ defined on some probability space $(\Omega, \mathcal{F}, \mathbb{P})$, taking values in $\mathcal{S}$ with the property that

$$
\begin{aligned}
& \mathbb{P}(X_{n+1} = k | X_0 = x_0, \ldots, X_{n-1} = x_{n-1}, X_n = i) \\
= \ & \mathbb{P}(X_{n+1} = k | X_n = i)
\end{aligned}
$$

for all $x_0, \ldots, x_{n-1}, i, k \in \mathcal{S}$ and $n \geq 0$. The set $\mathcal{S}$ is called the state space of the Markov chain.

Observations:

- A Markov chain is a time-discrete stochastic process.
- Sometimes we give the states labels which are not elements in $\mathbb{R}$ but can be mapped. See our $\mathcal{S} = \{\text{rain}, \text{snow}, \text{clear}\}$ example coming soon. Of course we can always give labels in $\mathbb{R}$.

Markov Chains

Definition (Time-homogeneous Markov Chains)

A Markov chain is called time-homogeneous if

$$\mathbb{P}(X_{n+1} = k | X_n = i) = \mathbb{P}(X_1 = k | X_0 = i) \tag{1}$$

for all $n \geq 0$.

Example (Board game)

See whiteboard.

Transition matrix Markov Chains

For time-homogeneous Markov chains we can arrange the probabilities in (1) in a matrix $P \in [0,1]^{|S|\times|S|}$ by

$$P_{i,j} = \mathbb{P}(X_1 = j | X_0 = i) \tag{2}$$

For the board game:

$$P = \begin{array}{c c}
 & \begin{array}{cccccccccc} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \end{array} \\
\begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} &
\begin{pmatrix}
0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 \\
0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 \\
0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 \\
0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0 & \frac{1}{6} \\
\frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & \frac{1}{6} & 0 & 0 & 0 & 0
\end{pmatrix}
\end{array}$$

Transition matrix Markov Chains

Important property transition matrix:

$$\begin{aligned}\sum_j P_{i,j} &= \sum_j \mathbb{P}(X_1 = j | X_0 = i) \\ &= \sum_j \frac{\mathbb{P}(X_1 = j, X_0 = i)}{\mathbb{P}(X_0 = i)} \\ &= \frac{\mathbb{P}(X_0 = i)}{\mathbb{P}(X_0 = i)} = 1\end{aligned}$$

Stochastic matrix

Definition (Stochastic matrix)

A matrix P is called stochastic matrix if

1. $P_{ij} \geq 0$ for all i, j
2. For each row i,

$$\sum_j P_{ij} = 1$$

Example (Weather change)

See whiteboard.