Markov chains - an example

1. Introduce Markov chains as done in the book, via the example with a random walker.

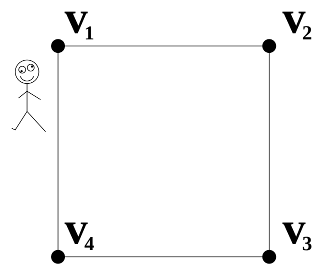


Figure 1: A simple random Walker in a very small town.

2. Present concepts as
   (a) the Markovproperty
   (b) time homogeneity
   (c) transition matrix
   (d) transition graph

3. Properly define a Markov chain, def. 2.1 in the book.

   Definition 2.1 ***Markov chain***
   Let P be a $k \times k$-matrix with elements $\{ P_{i,j} \ : i, j = 1, ...k \}$ A random process $(X_0, X_1, ...)$ with finite state space $S = \{s_1, s_2, ..., s_k\}$ is said to be a (homogenous) Markov chain with transition matrix P if for all n, all $i, j \in \{1, ..., k\}$ and all $i_0, ..., i_{n-1} \in \{1, ..., k\}$ we have

   $$\mathbb{P}(X_{n+1} = s_j | X_0 = i_0, X_1 = i_1, ..., X_{n-1} = i_{n-1}, X_n = i) = \mathbb{P}(X_{n+1} = s_j | X_n = i) = P_{i,j}$$

4. Present and prove theorem 2.1.

   Theorem 2.1
   For a Markov chain $(X_0, X_1, ...)$ with state space $S = \{s_1, s_2, ..., s_k\}$, initial distribution $\mu^{(0)}$ and transition matrix P, we have for any n that the distribution $\mu^{(n)}$ at any time n satisfies

   $$\mu^{(n)} = \mu^{(0)} P^n$$

   Proof :
   We use induction. Base case : $n = 1$. For any $j \in \{1, ..., k\}$

   $$\begin{aligned} \mu_j^{(1)} &= \mathbb{P}(X_1 = s_j) = \sum_{i=1}^{k} \mathbb{P}(X_0 = s_i, X_1 = s_j) \\ &= \sum_{i=1}^{k} \mathbb{P}(X_0 = s_i)\mathbb{P}(X_1 = s_j | X_0 = s_i) = \sum_{i=1}^{k} \mu_i^{(0)} P_{i,j} = (\mu_i^{(0)} P)_j \end{aligned}$$

This is true for any $j \in \{1, ..., k\}$ and this $\mu^{(1)} = \mu^{(0)}P$.

Induction hypothesis: Assume the statement is true for $n = m$

$$\begin{aligned} \mu_j^{(m+1)} &= \mathbb{P}(X_{m+1} = s_j) = \sum_{i=1}^{k} \mathbb{P}(X_m = s_i, X_{m+1} = s_j) \\ &= \sum_{i=1}^{k} \mathbb{P}(X_m = s_i)\mathbb{P}(X_{m+1} = s_j | X_m = s_i) = \sum_{i=1}^{k} \mu_i^{(m)} P_{i,j} = (\mu_i^{(m)} P)_j \end{aligned}$$

So we have $\mu^{(m+1)} = \mu^{(m)}P$. According to the induction hypothesis $\mu^{(m)} = \mu^{(0)}P^m$ so we get

$$\mu^{(m+1)} = \mu^{(m)}P = \mu^{(0)}P^m P = \mu^{(0)}P^{m+1}$$

and we are done. □

5. Recommended home work : problem 2.7, 2.8
A serious attempt to understand the problem is enough.

Problem 2.1

Consider the Markov chain in figure 1, with transition matrix P, initial distribution $\mu^{(0)}$ given by the following.

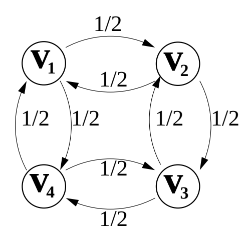
$$P = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \qquad \mu^{(0)} = (1,0,0,0)$$

(a) Compute the quare P^2 of matrix P. How do we interpret P^2 ?

(b) Prove by induction that

$$\mu^{(n)} = \begin{cases} \left(0, \frac{1}{2}, 0, \frac{1}{2}\right), & \text{if } n = 1,3,5,... \\ \left((\frac{1}{2}, 0, \frac{1}{2}, 0\right), & \text{if } n = 2,4,6,... \end{cases}$$

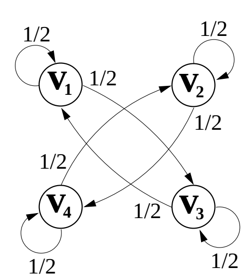
Solution First take a look at the transition graph for the random walk.



We compute P^2 ...

$$P^2 = \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} = \begin{pmatrix} 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

The interpretation of P^2 is the following. Given the Markov chain for the random walker, if we just look at every second time point we get at chain having P^2 as transition matrix. The transition graph is the following.



Now to proving the statement in **(b)** by induction. Base step We compute P^2 ...

$$\mu^{(1)} = \mu^{(0)}P = \left(0, \frac{1}{2}, 0, \frac{1}{2}\right) \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} = \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$$

Assume the statement is true for $n = k$. Case 1, $k = 2l + 1 \ldots$

$$\mu^{(2l+1)} = \mu^{(0)} P^{2l+1} = (\mu^{(0)} P^{2l}) P = \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right) \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} = \left(0, \frac{1}{2}, 0, \frac{1}{2}\right)$$

Case 2, $k = 2l \ldots$

$$\mu^{(2l+2)} = \mu^{(0)} P^{2l+2} = (\mu^{(0)} P^{2l+1}) P = \left(0, \frac{1}{2}, 0, \frac{1}{2}\right) \begin{pmatrix} 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \\ 0 & 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 & 0 \end{pmatrix} = \left(\frac{1}{2}, 0, \frac{1}{2}, 0\right)$$

So for every $n \geq 1$ the statement is true.

Problem 2.3

Consider example 2.1 (the Gothenburg weather), and suppose the markov chain starts on a rainy day, so that $\mu^{(0)} = (1, 0)$

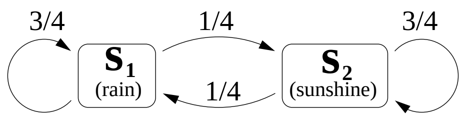
(a) Prove by induction that

$$\mu^{(n)} = \left(\frac{1}{2}(1 + 2^{-n}) \; , \;\; \frac{1}{2}(1 - 2^{-n}) \right)$$

for every n.

(b) What happens to $\mu^{(n)}$ when n tends to infinity?

Solution We have the markov chain with the following transition graph.



We have the following transition matrix

$$P = \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix}$$

Base step:

$$\mu^{(1)} = \mu^{(0)} P = (1, 0) \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix} = \left(\frac{3}{4}, \frac{1}{4} \right)$$

Which is correct since

$$\mu^{(1)} = \left(\frac{1}{2}(1 + 2^{-1}) \; , \;\; \frac{1}{2}(1 - 2^{-1}) \right) = \left(\frac{3}{4}, \frac{1}{4} \right)$$

Induction step. Assume the statement is true for $n = k$

$$\begin{aligned}
\mu^{(k+1)} &= \mu^{(k)} P \\
&= \left(\frac{1}{2}(1 + 2^{-k}) \; , \;\; \frac{1}{2}(1 - 2^{-k}) \right) \begin{pmatrix} 3/4 & 1/4 \\ 1/4 & 3/4 \end{pmatrix} \\
&= \left(\frac{3}{8}(1 + 2^{-k}) + \frac{1}{8}(1 - 2^{-k}) \; , \;\; \frac{1}{8}(1 + 2^{-k}) + \frac{3}{8}(1 - 2^{-k}) \right) \\
&= \left(\frac{3}{8} + \frac{3}{8\; 2^n} + \frac{1}{8} - \frac{1}{8\; 2^n} \; , \;\; \frac{1}{8} + \frac{1}{8\; 2^n} + \frac{3}{8} - \frac{3}{8\; 2^n} \right) \\
&= \left(\frac{1}{2} + \frac{1}{4\; 2^n} \; , \;\; \frac{1}{2} - \frac{1}{4\; 2^n} \right) \\
&= \left(\frac{1}{2}(1 + 2^{-(k+1)}) \; , \;\; \frac{1}{2}(1 - 2^{-(k+1)}) \right)
\end{aligned}$$

As $n \to \infty$

$$\lim_{n\to\infty} \mu^{(n)} = \lim_{n\to\infty} \left(\frac{1}{2} \underbrace{(1 + 2^{-n})}_{\to 1} \; , \;\; \frac{1}{2} \underbrace{(1 - 2^{-n})}_{\to 1} \right) = \left(\frac{1}{2}, \frac{1}{2} \right)$$

So the distribution $\mu^{(n)}$ converges towards $(1/2, 1/2)$ as n tends to infinity.