Lecture 6: Markov chains

MVE055 / MSG810 Mathematical statistics and discrete mathematics)

Moritz Schauer

Last updated September 17, 2020, 2020

GU & Chalmers University of Technology

Markov chains

We model the weather in the land Oz as one of R (rainy), S (sunny) or C (cloudy) . There, they never have two nice days in a row and if it was C (cloudy) yesterday, there is a 0.25 probability of R (rain) today.

For each day, the weather of the next day is random and we represent the probabilities by a matrix

$$
\begin{array}{cccc}
 & R & S & C \\
R & 0.5 & 0.25 & 0.25 \\
S & 0.5 & 0 & 0.5 \\
C & 0.25 & 0.25 & 0.5
\end{array}
$$

Each row contains the probability for next days weather depending on current weather.

This is an example of a Markov chain.

Markov chain

A Markov chain consists of:

A set of states: $\{s_1, \ldots, s_n\}$.

A matrix of transition probabilities

$$\mathbf{P} = \begin{pmatrix} p_{11} & \cdots & p_{1n} \\ \vdots & & \\ p_{n1} & \cdots & p_{nn} \end{pmatrix}$$

containing the probability p_{ij} to move from state s_i to state s_j

————

sv: övergångssannolikhet, övergångsmatrisen

Markov property

The transition probability does only depend on the current state:

$$p_{ij} = \mathrm{P}(\text{next state is } s_j \mid \text{current state is } s_i \cancel{\text{ and the state before ....}})$$

Transition probabilities

Transition probabilities are conditional probabilities:

$$p_{ij} = \mathrm{P}(\text{next state is } s_j \mid \text{current state is } s_i)$$

That means rows sum to 1: $\sum_{\text{all } j} p_{ij} = 1$.

What is the weather in three days

The probability that the Markov chain, starting in states s_i , will be in state s_j after n steps is given by the ij'th entry of

$$\mathbf{P}^n = \mathbf{P} \cdot \ldots \cdot \mathbf{P}$$

(n-fold matrix product.)

Example

Suppose we want to compute the probability that, given that it is rainy today, the weather will be cloudy in two days.

	R	S	C
R	0.5	0.25	0.25
S	0.5	0	0.5
C	0.25	0.25	0.5

$$\begin{aligned} p_{13}^{(2)} &= p_{11}p_{13} + p_{12}p_{23} + p_{13}p_{33} \\ &= 0.5(0.25) + 0.25(0.5) + 0.25(0.5) = 0.375 \end{aligned}$$

$$\mathbf{P}^2 = \begin{pmatrix} 0.4375 & 0.1875 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.1875 & 0.4375 \end{pmatrix}$$

Probability vectors

A probability vector is a row vector that gives the probabilities of being at each state at a certain step.

The probability vector which represents the initial state of a Markov chain is starting vector and is denoted by $\mathbf{u}^{(0)}$ or simply $\mathbf{u}$. The probability vector at step k is denoted by $\mathbf{u}^{(k)}$.

1 step

If $\mathbf{u}_k$ is the probability vector at step k, then the vector

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)}\mathbf{P}$$

is the probability vector at step $k+1$.

n steps

If $\mathbf{u}$ is the starting vector of a Markov Chain, then the probability vector at step n is given by

$$\mathbf{u}^{(n)} = \mathbf{u}\mathbf{P}^n.$$

Example

In the previous example, if the initial probability vector is $\mathbf{u} = (1/3, 2/3, 0)$, then the probability vector on day 2 will be

$$\mathbf{u}^{(2)} = \mathbf{u}\mathbf{P}^2 = \begin{pmatrix} 1/3 & 2/3 & 0 \end{pmatrix} \begin{pmatrix} 0.4375 & 0.1875 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.1875 & 0.4375 \end{pmatrix}$$

$$= \begin{pmatrix} 0.3958 & 0.2292 & 0.3750 \end{pmatrix}$$

This means that on day 2, there is a 39.58% chance of rain, 22.92% chance that the weather will be nice and 37.5% chance that it will be cloudy.

Regular matrix

A Markov chain is said to be regular if there exists n such that all the elements of the matrix $\mathbf{P}^n$ are nonzero. The Markov chain of the previous example is regular since

$$\mathbf{P}^2 = \begin{pmatrix} 0.4375 & 0.1875 & 0.375 \\ 0.375 & 0.25 & 0.375 \\ 0.375 & 0.1875 & 0.4375 \end{pmatrix}$$

(all the values are strictly positive) If the Markov chain is regular then, $\mathbf{P}^n \to \mathbf{Q}$ where

$$\mathbf{Q} = \begin{pmatrix} q_1 & q_2 & \dots & q_n \\ q_1 & q_2 & \dots & q_n \\ \vdots & \vdots & \ddots & \vdots \\ q_1 & q_2 & \dots & q_n \end{pmatrix}$$

q_j is the probability to be at state s_j on the long run.

Absorbing states

A state is said to be absorbing if it is impossible to leave it, that is $p_{ii} = 1$.

A Markov chain is called absorbing if it has at least one absorbing state, and if from every state it is possible to go to an absorbing state.

In an absorbing Markov chain, as state that is not absorbing is called transient.

Example:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} \end{pmatrix}$$

The transition matrix of an absorbing Markov chain with r absorbing states and t transient states can be written as

$$P = \begin{pmatrix} \mathbf{Q} & \mathbf{R} \\ 0 & \mathbf{I}_r \end{pmatrix}$$

where $\mathbf{I}_r$ is the identity matrix, 0 is the zero matrix (all elements are zeros), $\mathbf{Q}$ is a $t \times t$ -matrix and $\mathbf{R}$ is a $t \times r$ nonzero matrix. This form is called the canonical form. $\mathbf{P}^n = \begin{pmatrix} \mathbf{Q}^n & \star \\ 0 & \mathbf{I}_r \end{pmatrix}$ where $\star$ is a $t \times r$ matrix. $\mathbf{Q}^n$ gives the probability for being in each of the transient states after n steps for each possible transien starting state.