The Use of Markov Chains in Analyzing the Transfer Rates between School

Stephanie
Brown, Diana Gonzalez, and Jack Lei

Spring 2012

Introduction

Markov Chains are used to predict information based on what is going on in the present. To execute Markov Chains, one performs a trial many times where the outcome of each trial relies solely on the trial that came right before it.

A Markov chain is a sequence of probability vectors $\mathbf{x}_0, \mathbf{x}_1, \mathbf{x}_2, \ldots$, together with a **stochastic matrix** P, such that
$\mathbf{x}_1 = P\mathbf{x}_0, \quad \mathbf{x}_2 = P\mathbf{x}_1, \quad \mathbf{x}_3 = P\mathbf{x}_2, \ldots$

- A **stochastic matrix**: a square matrix whose columns are probability vectors.
- A **probability vector**: a vector with nonnegative entries that add up to 1.
  - E.g. $\begin{bmatrix} .34 \\ .47 \\ .19 \end{bmatrix}$
- A **state vector**: the n probabilities that the outcome of the trial is one of the n possible outcomes.

The Markov chain is described by the first-order difference equation
$\mathbf{x}_{k+1} = P\mathbf{x}_k$ for k = 0, 1, 2, …

Scenario

Do you want to transfer schools but don't know if you can? Do you want to know what type of school you can get in? Whether it be a CSU, UC, or a Private? Are you determinant to find the right school?
Suppose we did a study on transfer rates between CSUs, UCs, and Private Schools in 2012. Lets say the total number of students transferring within California is 58,000.
Find the outcomes for the number of people transferring this year and transferring in 2013.

Solution

The first step is to define the Stochastic matrix, P. In this matrix, we will put the percents of people who transferred where.

From

$$P = \begin{array}{ccc|l} \text{CSU} & \text{UC} & \text{Priv.} & \\ \hline .75 & .32 & .03 & \text{CSU} \\ .23 & .54 & .10 & \text{UC} \\ .02 & .14 & .87 & \text{Priv.} \end{array}$$

TO

Next, we have our $\mathbf{x}_0$ which are the results from 2011's trial in percentages

$$x_0 = \begin{bmatrix} .552 \\ .403 \\ .045 \end{bmatrix}$$

Now we can plug it in to the equation $\mathbf{x}_1 = P\mathbf{x}_0$

$$\begin{bmatrix} .75 & .32 & .03 \\ .23 & .54 & .10 \\ .02 & .14 & .87 \end{bmatrix} \begin{bmatrix} .552 \\ .403 \\ .045 \end{bmatrix} = \begin{bmatrix} .544 \\ .349 \\ .107 \end{bmatrix}$$

$$\begin{bmatrix} .544 \\ .349 \\ .107 \end{bmatrix} \rightarrow \begin{bmatrix} 31{,}566 \\ 20{,}256 \\ 6{,}178 \end{bmatrix}$$

The transfer rates for 2012 are the following:

$$x_1 = \begin{bmatrix} 31{,}566 \\ 20{,}256 \\ 6{,}178 \end{bmatrix}$$

Lastly, we can take this years results to predict next year's transfer rate.

$$x_1 = \begin{bmatrix} .544 \\ .349 \\ .107 \end{bmatrix}$$

Now we can plug it in to the equation $\mathbf{x}_2 = P\mathbf{x}_1$

$$\begin{bmatrix} .75 & .32 & .03 \\ .23 & .54 & .10 \\ .02 & .14 & .87 \end{bmatrix} \begin{bmatrix} .544 \\ .349 \\ .107 \end{bmatrix} = \begin{bmatrix} .523 \\ .324 \\ .153 \end{bmatrix} \rightarrow \begin{bmatrix} 30{,}342 \\ 18{,}816 \\ 8{,}842 \end{bmatrix}$$

Conclusion

Markov Chain is used to predict future outcomes based on previous results. Using the number of transfer students in 2011, we can estimate where students are more likely to transfer for upcoming years. Because Markov Chain allows us forecast long-term behavior in education, we can observe that over a five-year span, CSU's will experience the largest number of transfer students, while Private Universities will have the fewest.

Applications

More widely, Markov Chain can be used to anticipate the weather, used in sports, game theory, chemistry, physics and gambling.. You can also use Markov chains to predict the water. For example, your probability vectors could be "Rain, Sunny, Snow" and based on today's whether, you could forecast tomorrow's.

Acknowledgements

Lay, David C. Linear Algebra and Its Applications. Third ed. Pearson Education, 2007. Print