Application of Markov Chains

by Joshua Crawford and Brett Rustemeyer

Abstract

Numerous methods for modeling real world situations mathematically have been developed, but one that lends itself particularly well to computation is the Markov chain. This paper briefly examines the history and principals behind Markov chains, develops an example model using historical data, and comments on the limitations of Markov chains.

Introduction

In almost every industry, professionals rely on mathematical models to project future outcomes. These types of predictions are especially common in representations of employment, production, and population. One model currently used in fields as diverse as bioinformatics and statistical physics is the Markov chain. Markov chains were first formalized by Andrei Andreevich Markov who studied at Petersburg University in 1874. He was mentored by the famous Pafnuty Lvovich Chebyshev, and later inducted into the St. Petersburg Academy of Science .[1]

Background[2]

Markov proposed that when a situation has a finite number of outcomes (each having a discrete probability), the situation can be represented by a vector of probabilities (a state vector.) The probabilities of transition from any one of the possible states to any of the other states can be assembled in a matrix, and the result is a *transition matrix* (or *stochastic matrix.*). The entries in every column of this transition matrix sum to one. If $P=[p_{kj}]$ is the transition matrix of any Markov chain with k possible states, then for each j we must have: $p_{1j}+p_{2j}+...+p_{kj}=1$ because if the system is in state j at one observation it is certain to be in one of the k possible states at the next observation.[3]

Once assembled, the transition matrix can be multiplied by a state vector, and the result is a new state vector. Each new state vector is determined solely by the result of the immediately preceding trial. When this is repeated successive times, the probability of an outcome any number of discrete time periods in the future can be estimated. This process, and the mathematical and geometric structure is referred to as a *Markov chain*.

EXAMPLE: S&P 500

As an example, consider the probability of a commodity rising or falling in price. The various possible percentage moves can be represented in a transition matrix, where the columns represent the current state, and the rows represent the future state. A statistical evaluation of the Standard and Poor's 500 stock index from January of 1993 through December of 1994[4] produces the following transition matrix[5]:

$$M = \begin{bmatrix} 0.474 & 0.448 & 0.492 & 0.500 \\ 0.051 & 0.034 & 0.070 & 0.091 \\ 0.449 & 0.517 & 0.367 & 0.273 \\ 0.026 & 0.000 & 0.070 & 0.136 \end{bmatrix}$$

The first column represents the 0 to 1% bracket, the second column is the 1 to 5% bracket, the third column is the -1 to 0% bracket, and the fourth column is the -5 to -1% bracket. The rows represent the 0 to 1%, 1 to 5%, -1 to 0%, and -5 to -1% brackets respectively. In concept, a rise or fall of greater magnitude than 5% is possible, but in practice is very rare, so it is not included in this model. (A more accurate model would definitely include a wider range of dates, and necessarily take into account such deviations.) To predict the outcome of the next day's trading, a current probability matrix must be developed. Say that on a given day, the S&P 500 rises by 1.4%. The probability vector for that day's result then becomes $p_0 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix}$. Taking Mp yields a new probability vector $p_1 = \begin{bmatrix} 0.448 \\ 0.034 \\ 0.517 \\ 0 \end{bmatrix}$. Repeated left multiplication by M produce the following series of vectors: $p_2 = \begin{bmatrix} 0.482 \\ 0.060 \\ 0.408 \\ 0.048 \end{bmatrix}$, $p_3 = \begin{bmatrix} 0.480 \\ 0.060 \\ 0.410 \\ 0.048 \end{bmatrix}$,

$p_4 = \begin{bmatrix} 0.480 \\ 0.060 \\ 0.410 \\ 0.048 \end{bmatrix}$. These are the probability vectors for trading on successive days based solely on the outcome the previous day, and the market's performance over the past two years. In each case, a formula for calculating expected value[6] can be used to determine whether the investment should be made or not.

$$E(x)=\sum x\,P(x)$$

Where x is the value of a possible outcome, and P(x) is its probability.

The largest expected return occurs if the possible losses are each at the low end of their respective brackets and the possible gains are at the highest end of their brackets. For example, on day one: $E_1(x)_{max}=0.01(0.448)+0.05(0.034)-0.00(0.517)-0.01(0.00)=0.00618$. In addition, the lowest possible return occurs if possible gains are at the lowest end of their respective brackets, and the possible losses are at the highest end of their brackets. So, $E_1(x)_{min}=0.00(0.448)+0.01(0.034)-0.01(0.517)-0.05(0.00)=-0.00483$. Based on this simplistic model, the investment should be made. However, this model makes a decision based solely on what happened the previous day. There may very well be more influential (i.e. decisive) trends that extend days or even weeks. To examine these trends, patterns could be used to rework the matrix.

Conclusion

The example model constructed in this paper is not without flaws. After continuing the Markov chain for a number of trials, the entries for the new state vector no longer add up to 1. So at that point the prediction no longer logically valid. It is not immediately apparent what causes this loss of precision. It is possible that it is due in some way to the exclusion of the last data point from some calculations, and its inclusion in others. However, a thorough investigation of this anomaly is beyond the scope of this paper. The model also fails to accurately predict market behavior. This stems primarily from the comparatively small set of historical data used to construct the model. The results could be improved by sampling a larger set of data, or by modeling the market's performance by a larger matrix (using smaller percentage brackets), or still more effectively by a continuous instead of discrete Markov chain. It can be argued that Markov Chains are not the method to model stock market behavior at all. Other models can certainly take more factors into account, rather than simply looking at probabilities calculated directly from statistics. Another reason for the model's inability to predict market behavior is due to one of its properties. All stochastic matrices have an *equilibrium vector*: one at which the probabilities do not change with repeated multiplication by the transition matrix. This means that after the state which matches the equilibrium vector is reached, the model will predict the same outcome for every successive state. Obviously, this is not how the market works. One final problem with a Markov chain as applied in this paper lies in its inability to adapt to changing market conditions, or even new historical data. These flaws can easily be accounted for either with additions to the model, or by updating the transition matrix at every state. Even with these flaws, Markov chains are extremely useful, and the primary benefit of utilizing them is the ease with which the computations can be made.

References

1. (Accessed on 10/2/04.) <http://www.brainencyclopedia.com/encyclopedia/a/an/andrey_markov.html>
2. Lay, David C. "Linear Algebra and Its Applications." -3rd Edition. (2003)
3. Anton, Howard and Rorres, Chris "Elementary Linear Algebra with Applications." (1987)
4. (Accessed on 9/29/04.) <http://finance.yahoo.com/q/hp?s=SPY&a=00&b=1&c=1993&d=00&e=1&f=1994&g=d>
5. Crawford, Joshua and Rustemeyer, Brett <http://www.web.umr.edu/~jmcnpc>
6. (Accessed on 10/2/04.) <http://www.planetmath.org/encyclopedia/ExpectedValue.html>