Application of Markov Chains to Stock Trends

Kevin J. Doubleday and Julius N. Esunge
Department of Mathematics, University of Mary Washington Fredericksburg, VA

Abstract: Problem statement: Modeling of the Dow Jones Industrial Average is frequently attempted in order to determine trading strategies with maximum payoff. Changes in the DJIA are important since movements may affect both individuals and corporations profoundly. Previous work showed that modeling a market as a random walk was valid and that a market may be viewed as having the Markov property. **Approach:** The aim of this research was to determine the relationship between a diverse portfolio of stocks and the market as a whole. To that end, the DJIA was analyzed using a discrete time stochastic model, namely a Markov Chain. Two models were highlighted, where the DJIA was considered as being in a state of (1) gain or loss and (2) small, moderate, or large gain or loss. A portfolio of five stocks was then considered and two models of the portfolio much the same as those for the DJIA. These models were used to obtain transitional probabilities and steady state probabilities. **Results:** Our results indicated that the portfolio behaved similarly to the entire DJIA, both in the simple model and the partitioned model. **Conclusion:** When treated as a Markov process, the entire market was useful in gauging how a diverse portfolio of stocks might behave. Future work may include different classifications of states to refine the transition matrices.

Key words: Markov chains, stock trends, Dow Jones Industrial Average (DJIA), steady state, finite number, Microsoft Excel, transitional state

INTRODUCTION

As infamously stated by Adam Smith in Wealth of Nations, "and he is in this...led by an invisible hand to promote an end that was no part of his intention." That is, markets are unpredictable, pure and simple and a working model cannot be constructed which reflects their movements. This, however, does not discourage the creation of such models whose aim is to accurately predict movements in the various markets (Agwuegbo *et al.*, 2010). This research analyzes market trends by determining probabilities that the market transitions between various states. A similar study was conducted by Agwuegbo *et al.* (2010). Closing prices of the market are considered so that analysis can be done in a discrete manner and the transition probabilities are utilized as parts of Markov chains to model the market. First, the concept of a random variable and a random walk is introduced. Second, Markov theory and properties of Markov chains are discussed. Finally, two models of the Dow Jones Industrial Average (DJIA) and two models of a specific portfolio of stocks in the DJIA are introduced and analyzed (Jones and Smith, 2009).

MATERIALS AND METHODS

A random walk is said to exhibit the Markov property if the position of the walk at time n depends only upon the position of the walk at time n-1 (Winston, 2004). If we call our random variable X_n, then Eq. 1:

$$P(X_n = j | X_{n-1} = i) = p_{ij} \tag{1}$$

Is independent of X_{n-2}, X_{n-3},…,X_1 so that the state of X at time n depends only upon the state of X at step n-1. Here each p_{ij} for j =1,2,… is a probability row vector describing every possible trasition from state i to any other available state in the system.

Then Eq. 2:

$$\sum_{j=1}^{m} p_{ij} = 1 \tag{2}$$

for every i. Notice that this describes a random walk existing in m possible states Eq. 3.

Corresponding Author: Kevin J. Doubleday, Department of Mathematics, University of Mary, Washington Fredericksburg, VA
Tel: 540.903.0763

Thus:

$$P(X_{n+1} = i_{n+1} \begin{vmatrix} X_n = i_n, X_{n-1} = i_{n-1} \ldots) = \\ P(X_{n+1} = i_{n+1} | X_n = i_n) \end{vmatrix} \quad (3)$$

$$\lim_{n \to \infty} T^n = \begin{bmatrix} \pi_1 & \pi_2 & \ldots & \pi_m \\ \pi_1 & \pi_2 & \ldots & \pi_m \\ \ldots & \ldots & \ldots & \ldots \\ \pi_1 & \pi_2 & \ldots & \pi_m \end{bmatrix} \quad (5)$$

The process of moving from one state of the system to another with the associated probabilities of each transition is known as the chain (Winston, 2004). It is said that every step taken in a chain possessing the Markov property depends only upon the immediately preceding step. It can easily be seen how calculating probabilities of a series, or chain, of events in a Markov system is greatly simplified due to this Markov property. Instead of concerning ourselves with the entire path a random variable might have taken to arrive at its current state, we need only consider its state directly before a given point of interest.

The transition probabilities form an m_m transitional probability matrix T, where:

$$T = \begin{bmatrix} p_{ij} \end{bmatrix} = \begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1m} \\ p_{21} & p_{22} & \cdots & p_{2m} \\ \cdots & \cdots & \cdots & \cdots \\ p_{m1} & p_{m2} & \cdots & p_{mm} \end{bmatrix}$$

Each row of T is the probability distribution relating to a transition from state i to state j.

States i and j are said to communicate if there exists a path between them (Winston, 2004). It must be true that i is reachable from j in a finite number of transitions and also that j is reachable from i in a finite number of transitions for any two states i and j to communicate. A state i is said to be periodic if all paths leading from state i back to i have a length that is a multiple of some integer k, such that k > 0 for the smallest possible k (Winston, 2004). If all states of a chain communicate a nd are not periodic, then the chain is said to be ergodic.

A chain is said to have a steady state distribution if there exists a vector p such that given a transition matrix T we have Eq. 4:

$$\pi T = \pi. \quad (4)$$

If a chain is ergodic then we are guaranteed the existence of this steady state vector p (Isaacson and Madsen, 1985; Winston, 2004). This steady state vector can be viewed as the distribution of a random variable in the long run.

This steady state probability vector p of an m state random walk can also be obtained as (Isaacson and Madsen, 1985) Eq. 5:

Given this formulation of a transition matrix and its steady state, we can set up a system of classification of the Dow Jones Industrial Average (DJIA). The idea of using Markov chains to forecast the behavior of stock prices is popular since potential investors are interested in market trends which might lead to an optimum investment strategy[?]. For this study, three applications of Markov analysis will be considered, namely:

- Probabilities of the entire Dow Jones Industrial Average (DJIA) moving up or down
- Probabilities of DJIA moving between partitions of the possible gains and losses
- Probabilities of a specific portfolio of stocks moving up or down
- Probabilities of a specific portfolio of stocks moving between partitions of possible gains and losses

Closing values of the DJIA were gathered for all 253 trading days of the year 2010, January 4, 2010 to December 31, 2010 from msnmoney.com. The closing values were entered in Microsoft Excel and categorized. For application (1), each day was classified as having closed higher or lower than the previous day, thus allowing classification of two states, namely:

State 1: Closing value is less than closing value of the previous day
State 2: Closing value is greater than or equal to the closing value of the previous day

Since an investor is generally creating a portfolio of stocks, application (2) compiles a portfolio of popular companies and models the portfolio in the same ways which application (1) modeled the entire DJIA. The portfolio of choice consisted of JP Morgan Chase, Apple, Google, Intel and Qualcomm. Closing prices of these five companies were gathered from January 4, 2010 to December 31, 2010. Then, the states are:

State 1: Value of the portfolio closes lower than previous day
State 2: Value of the portfolio closes higher than previous day

The transition and steady state probabilities were then compared to closing price data for the portfolio gathered from January 3, 2011 to February 18, 2011 to test the integrity of the model.

For application (3), the average change in the closing value of the DJIA was examined. Based on the 2010 data, gains and losses were each partitioned into three subcategories each, namely, small, moderate and large. Transitions for this experiment consisted of moving from a category of gain or loss one day to a category of gain or loss the next, namely:

State 1: Large jump up (gain greater than 167)
State 2: Moderate jump up (gain between 83 and 167)
State 3: Small jump up (gain less than 83)
State 4: Small jump down (loss less than 83)
State 5: Moderate jump down (loss between 83 and 167)
State 6: Large jump down (loss greater than 167)

The intervals indicated in parentheses were obtained by determining the absolute average of the market's daily changes. This experiment is meant to be more precise than the first since the concept of the market moving 'up' or 'down' is now better defined.

For model (4), a portfolio partitioned gain and loss model was constructed in the same manner as model (3). The same portfolio described in application (2) was partitioned into six transitional states, namely:

State 1: Large jump up (gain greater than 47.41)
State 2: Moderate jump up (gain between 23.71 and 47.41)
State 3: Small jump up (gain less than 23.71)
State 4: Small jump down (loss less than 16.40)
State 5: Moderate jump down (loss between 16.40 and 32.80)
State 6: Large jump down (loss greater than 32.80)

RESULTS

The transition matrix T1 for model 1, accurate to four decimal places, was found to be:

$$T_1 = \begin{array}{c} S_1 \\ S_2 \end{array} \begin{pmatrix} S_1 & S_2 \\ 0.4074 & 0.5926 \\ 0.4857 & 0.5143 \end{pmatrix}$$

We found that:

$$T_1^5 = \begin{pmatrix} 0.4504 & 0.5496 \\ 0.4504 & 0.5496 \end{pmatrix}$$

indicating that $\pi_1 = (0.4504, 0.5496)$.

The transition matrix for model 2 was found to be:

$$T_2 = \begin{pmatrix} 0.4310 & 0.5690 \\ 0.4815 & 0.5185 \end{pmatrix}$$

We found:

$$T_2^4 = \begin{pmatrix} 0.4584 & 0.5416 \\ 0.4584 & 0.5416 \end{pmatrix}$$

so that $\pi_2 = (0.4584, 0.5416)$. Comparing (1) and (2) we see that the entire market and the portfolio achieve roughly the same steady state, indicating that in the long run both will have similar probability distributions with a slightly higher chance of any given day being a day of gains. This bodes well for investors, as they can see that a diversified portfolio will behave similarly to the entire market and that over a long period of time the number of gain days will outnumber the days of loss.

For the third model, the transition matrix was found to be:

$$T_3 = \begin{array}{c} S_1 \\ S_2 \\ S_3 \\ S_4 \\ S_5 \\ S_6 \end{array} \begin{pmatrix} S_1 & S_2 & S_3 & S_4 & S_5 & S_6 \\ 0 & 0.200 & 0.533 & 0.200 & 0.067 & 0 \\ 0 & 0.077 & 0.346 & 0.269 & 0.269 & 0.039 \\ 0.030 & 0.080 & 0.460 & 0.310 & 0.090 & 0.03 \\ 0.129 & 0.071 & 0.386 & 0.257 & 0.071 & 0.086 \\ 0.111 & 0.259 & 0.260 & 0.259 & 0.074 & 0.037 \\ 0 & 0.083 & 0.333 & 0.250 & 0.250 & 0.083 \end{pmatrix}$$

We found:

$$T_3^7 = \begin{pmatrix} 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \\ 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \\ 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \\ 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \\ 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \\ 0.0597 & 0.1038 & 0.4043 & 2761 & 1079 & 0.0479 \end{pmatrix}$$

indicating that $\pi_3 = (0.0597, 0.1038, 0.4043, 0.2761, 0.1079, 0.0479)$. For the fourth model, the transition matrix was found to be:

$$T_4 = \begin{pmatrix} 0 & 0 & 0.429 & 0.571 & 0 & 0 \\ 0 & 0.071 & 0.500 & 0.429 & 0 & 0 \\ 0.043 & 0.085 & 0.287 & 0.543 & 0.21 & 0.21 \\ 0.023 & 0.039 & 0.419 & 0.504 & 0.016 & 0 \\ 0 & 0 & 0.500 & 0.500 & 0 & 0 \\ 0 & 0 & 0.500 & 0.500 & 0 & 0 \end{pmatrix}$$

DISCUSSION

The matrices T1 and T2 indicate that given a day is in either state there is a greater chance of transitioning to a state of gain than a state of loss. The good news for investors is that the steady state probability vectors p1 and p2 show that there is a greater probability of a day being a day of gains than a day of losses. This was found to be true of both the portfolio and the entire DJIA.

There was no steady state obtained for model 4, indicating that the chain is not ergodic. Notice that in both T_3 and T_4 each row vector contains a majority of the probabilities in columns three and four, indicating that as not matter what state a day occupies, there is a high probability that the next day will be a day of small loss or small gain. This is encouraging for investors since the market will not be prone to large, sustained swings either up or down. We also found that the steady state distribution of model 3 resembled a discrete normal distribution with roughly 68.04% of the days predicted to be in states three and four, or the middle of the distribution and roughly 89.21% of the data falling within the middle four states.

CONCLUSION

This research shows that a diverse portfolio of stocks will mirror the movements of the entire market. That is, the portfolio will show a great propensity to have small gains and losses and show probabilistic immunity to consecutive days of great loss and gain. Therefore, a portfolio of this nature will yield a slow steady growth as there are likely to be more days of gain than days of loss.

ACKNOWLEDGMENT

REFERENCES

Agwuegbo, S.O.N., A. P. Adewole and A. N. Maduegbuna, 2010. A random walk model for stock market prices. J. Math. Stat., 6: 342-346. DOI: 10.3844/jmssp.2010.342.346

Isaacson, D.L. and R.W. Madsen, 1985. Markov Chains: Theory and Applications. 1st Edn., R.E. Krieger Pub. Co., New York, ISBN-13: 9780898748345, pp: 256.

Jones, P.W. and P. Smith, 2009. Stochasitic Processes: An Indroduction. 2nd Edn., Chapman and all/CRC, Spain, ISBN: 13: 9781420099607, pp: 221.

Winston, W.L., 2004. Introduction to Probability Models: Operations Research. 4th Edn., Brooks Cole Cengase Learning, USA., ISBN: 10: 053440572X, pp: 729.