МАТЕМАТИКА И МАТЕМАТИЧЕСКО ОБРАЗОВАНИЕ, 2004
MATHEMATICS AND EDUCATION IN MATHEMATICS, 2004
Proceedings of the Thirty Third Spring Conference of
the Union of Bulgarian Mathematicians
Borovets, April 1–4, 2004

AN APPROXIMATION OF THE SQUARE ROOT OF TRANSITION MATRICES*

Pavel T. Stoynov

Markov chains are popular tool for modeling different systems - natural or artificial. A basic their characteristic is the transition matrix which is a kind of stochastic matrix. In this article, we consider a special kind of transition matrix which arise in credit risk theory when we model transitions from one credit rating to another. The main characteristic of this kind of matrices is that the probability mass is concentrated around the main diagonal of the matrix. To this kind of matrices we apply a procedure for taking square root which gives the transition matrix for a period twice smaller than the basic period. Numerical testing is made and the results are shown in the article.

Credit risk and transition matrices Financial institutions use systems of credit ratings for measuring the credit risk. These systems can be modeled by Markov chains. In such a system, the credit risk is a risk of a credit event – change of the credit rating which corresponds to a transition in the Markov chain. The credit events are case of *special events* – events which occur rarely but can affect considerably the system behavior. The probabilities for transitions from one credit rating to another are given by the so called *transition matrix*.

Definition 1. *Let us have a system of m credit ratings. The stochastic matrix $A_{m\times m}$ where a_{ij} is the probability a company with current credit rating i to pass into credit rating j for a given period is called transition probabilities matrix or transition matrix for that period. The elements on the main diagonal of the transition matrix are approximately equal to 1, i. e. $a_{ii} = 1-\varepsilon_i$, $i = 1,\dots,m$. Here ε_i are relatively small numbers for which* $\varepsilon_i = \sum\limits_{k=1,k\neq i}^{m} a_{ik}$.

Usually, the transition matrix is given for a specified period which we will call the *basic period* and which is commonly one year. When it is necessary to determine the transition probabilities for another period, we must apply a procedure for obtaining transition probabilities for the new period.

The most commonly used approach is to consider the system of credit ratings as a homogeneous Markov chain. However, the assumption that the system of credit ratings is a Markov system is only an approximation to the real case. The next theorem gives

*2000 Mathematics Subject Classification: 15A51, 15A52.
Key words: transition matrix, credit rating, approximate Markov model.

an alternative proposition for determining the transition probabilities for a period two times less than the basic period.

Let $A_{m\times m}$ be the transition matrix for the basic period T. Then we consider the matrix $B_{m\times m}$ with elements

$$(1)\qquad \begin{aligned} b_{ij} &= \frac{a_{ij}}{2};\ i \neq j \\ b_{ii} &= a_{ii} + \frac{1}{2}\sum_{k=1,k\neq i}^{m} a_{ik},\ i = j \end{aligned}$$

as a transition matrix for a period $\frac{T}{2}$. It is then natural to consider the matrix $G_{m\times m}$ with elements $g_{ij} = 2a_{ij}$ for $i \neq j$ and $g_{ii} = a_{ii} - \sum\limits_{k=1,k\neq i}^{m} a_{ik}$ for $i = j$ as a transition matrix for a period $2T$. By introducing this approach we obtain new method for calculating transition probabilities which we will call *approximate Markov model.*

We have the following theorem:

Theorem 1. *Let A be a transition matrix and B be defined by* (1)*. Then for the matrix $C = BB = B^2$ we have $C \approx A$. If $\varepsilon_i = \varepsilon = \mathrm{const}$, for the errors $d_{ij} = c_{ij} - a_{ij}$ we have $|d_{ij}| \leq \frac{m}{4}\varepsilon^2$.*

Usually ε is of the order of 0.01, and $m < 10$.

Here $C \approx A$ means that there exist a small δ for which $\sup\limits_{i,j} |c_{ij} - a_{ij}| < \delta$.

In this theorem, the square root of the matrix A is approximated by the first iteration of Newton's method for solving the matrix equation $X^2 - A = 0$. The Newton's method is given by

$$X_{k+1} = \frac{1}{2}(X_k + X_k^{-1}A),\ k \geq 0,\ X_0 = A.$$

If A has no non-positive real eigenvalues, X_k converges quadratically to$A^{\frac{1}{2}}$ as pointed out in [1]. Newton's iteration is not used for practical computations. Analysis in [1] shows that unless the eigenvalues λ_i of A satisfy the requirement that

$$\frac{1}{2}|1 - \left(\frac{\lambda_i}{\lambda_j}\right)^{\frac{1}{2}}| \leq 1,\ i,j = 1,\ldots,n$$

then small errors in the iterates can be amplified by the iteration. As a result, the iteration fails to converge. Here, we are not looking for convergence, and since we don't need the condition for the non-positive real eigenvalues of A. The theorem states that because of the special kind of the transition matrices we have a good approximation of $A^{\frac{1}{2}}$ at the first iteration yet. And the approximation quality is not a consequence of the convergence of the iteration procedure. Actually, in the approximate Markov model, we don't need any good approximation of $A^{\frac{1}{2}}$ because we don't suppose that the system of credit ratings is exactly a homogenuous Markov system.

The proof of the theorem can be seen in [2].

2. An algorithm for obtaining transition matrices for periods smaller than the basic period. On the basis of the approximate Markov model we derive the following formula for obtaining the transition probability matrix $C(s, A)$ for a period

$\frac{T}{s} = \frac{T}{2^k}$, $k = 1, 2, \ldots$ where T is the basic period for which the transition matrix A is given:

$$C(s,A) = \frac{A + (s-1)I}{s}. \tag{2}$$

Here I is the unity matrix (with elements equal to 1 on the main diagonal and to 0 elsewhere). This formula is actually an approximation for taking s-th root of a matrix in the case when it is a transition matrix. The reduction of the computing complexity is obvious. But the simplicity of the calculation is not the strongest feature of the algorithm because of the powerful computers used today. The method allows to keep some basic characteristics of the transition matrices. Firstly, any transition matrix remains transition matrix after the transformation. This is not the case when we use the sophisticated numerical methods applied in the computer programs, for example MATLAB. When we use MATLAB functions for taking root matrix, "negative" probabilities appear and the sum of the elements in one row most often is not 1. Secondly, when we use (2), the transition probabilities which have been zeros for the basic period remain zeros also for the smaller periods and do not increase as in the case when we use MATLAB functions.

Numerical tests with matlab. Some tests by using the program MATLAB 6.0 are made to obtain $C(s,A)$ and compare it to the matrix s-th root generated by the MATLAB algorithms which we denote by $MATLAB\_ROOT(s,A)$. Here we show only one example.

$$A = \begin{pmatrix} 0.8 & 0.1 & 0.1 & 0 & 0 & 0 & 0 \\ 0.1 & 0.8 & 0.1 & 0 & 0 & 0 & 0 \\ 0 & 0.1 & 0.8 & 0.1 & 0 & 0 & 0 \\ 0 & 0 & 0.1 & 0.8 & 0.1 & 0 & 0 \\ 0 & 0 & 0 & 0.1 & 0.8 & 0.1 & 0 \\ 0 & 0 & 0 & 0 & 0.1 & 0.8 & 0.1 \\ 0 & 0 & 0 & 0 & 0.1 & 0.1 & 0.8 \end{pmatrix}$$

$$C(8,A) = \begin{pmatrix} 0.975 & 0.0125 & 0.0125 & 0 & 0 & 0 & 0 \\ 0.0125 & 0.975 & 0.0125 & 0 & 0 & 0 & 0 \\ 0 & 0.0125 & 0.975 & 0.0125 & 0 & 0 & 0 \\ 0 & 0 & 0.0125 & 0.975 & 0.0125 & 0 & 0 \\ 0 & 0 & 0 & 0.0125 & 0.975 & 0.0125 & 0 \\ 0 & 0 & 0 & 0 & 0.0125 & 0.975 & 0.0125 \\ 0 & 0 & 0 & 0 & 0.0125 & 0.0125 & 0.975 \end{pmatrix}$$

$$MATLAB\_ROOT(8,A) =$$

$$= \begin{pmatrix} 0.97172 & 0.014473 & 0.014541 & -0.00078733 & 6.1626\times10^{-5} & -5.5797\times10^{-6} & 5.4033\times10^{-7} \\ 0015322 & 0.97087 & 0.014541 & -0.00078733 & 6.1626\times10^{-5} & -5.5797\times10^{-6} & 5.4033\times10^{-7} \\ -0.00084842 & 0.015389 & 0.97086 & 0.01539 & -0.000854 & 6.7746\times10^{-5} & -6.12\times10^{-6} \\ 6.7206\times10^{-5} & -0.00085454 & 0.01539 & 0.9708 & 0.01539 & -0.00085454 & 6.7206\times10^{-5} \\ -6.12\times10^{-6} & 6.7746\times10^{-5} & -0.000854 & 0.01539 & 0.97086 & 0.015389 & -0.00084842 \\ 5.4033\times10^{-7} & -5.5797\times10^{-6} & 6.1626\times10^{-5} & -0.00078733 & 0.014541 & 0.97087 & 0.015322 \\ 5.4033\times10^{-7} & -5.5797\times10^{-6} & 6.1626\times10^{-5} & -0.00078733 & 0.014541 & 0.014473 & 0.97172 \end{pmatrix}$$

$[C(8, A)]^8 =$

$$
= \begin{pmatrix}
0.82051 & 0.087717 & 0.087813 & 0.0038611 & 9.7994\times10^{-5} & 1.5608\times10^{-6} & 1.5944\times10^{-8} \\
0.083954 & 0.82427 & 0.087813 & 0.0038611 & 9.7994\times10^{-5} & 1.5608\times10^{-6} & 1.5944\times10^{-8} \\
0.0037631 & 0.08405 & 0.82427 & 0.08405 & 0.0037646 & 9.6449\times10^{-5} & 1.5449\times10^{-6} \\
9.6433\times10^{-5} & 0.0037646 & 0.08405 & 0.82418 & 0.08405 & 0.0037646 & 9.6433\times10^{-5} \\
1.5449\times10^{-6} & 9.6449\times10^{-5} & 0.0037646 & 0.08405 & 0.82427 & 0.08405 & 0.0037631 \\
1.5944\times10^{-8} & 1.5608\times10^{-6} & 9.7994\times10^{-5} & 0.0038611 & 0.087813 & 0.82427 & 0.083954 \\
1.5944\times10^{-8} & 1.5608\times10^{-6} & 9.7994\times10^{-5} & 0.0038611 & 0.087813 & 0.087717 & 0.82051
\end{pmatrix}
$$

$[MATLAB\_ROOT(8, A)]^8 =$

$$
= \begin{pmatrix}
0.8 & 0.1 & 0.1 & -2.2637\times10^{-15} & 1.5021\times10^{-16} & 1.0749\times10^{-15} & 1.5927\times10^{-15} \\
0.1 & 0.8 & 0.1 & -4.608\times10^{-16} & -2.452\times10^{-15} & 3.06447\times10^{-15} & 3.6181\times10^{-15} \\
-1.9118\times10^{-15} & 0.1 & 0.8 & 0.1 & 4.1899\times10^{-17} & 1.6543\times10^{-15} & 2.2072\times10^{-15} \\
-2.9875\times10^{-15} & -1.8405\times10^{-15} & 0.1 & 0.8 & 0.1 & 1.1794\times10^{-15} & 1.7599\times10^{-15} \\
2.8222\times10^{-16} & -2.3959\times10^{-15} & 1.6765\times10^{-16} & 0.1 & 0.8 & 0.1 & 1.8275\times10^{-15} \\
7.9756\times10^{-16} & 2.7925\times10^{-15} & 1.6413\times10^{-15} & 1.4392\times10^{-15} & 0.1 & 0.8 & 0.1 \\
1.311\times10^{-15} & 2.7893\times10^{-15} & 2.0274\times10^{-15} & 2.1359\times10^{-15} & 0.1 & 0.1 & 0.8
\end{pmatrix}
$$

Conclusion. The numerical tests show that the procedure for obtaining transition matrices for periods smaller than the basic period is applicable to credit risk analysis.

REFERENCES

[1] N. Higham. Newton's method for the matrix square root. *Math. Comp.*, **46**(1986), No 174, 537–549.

[2] P. Stoynov. An Approach for Determining Transition Probabilities in a System of Credit Ratings when Modeling the Credit Risk. *Statistics*, **1**, (2001), 76–84 (in Bulgarian).

Pavel Stoynov
Department of Statistics and Econometry
Faculty of Economics and Business Administration
St. Kliment Ohridski Sofia University
125, Tzarigradsko shausse Blvd.
1113 Sofia, Bulgaria
e-mail: `todorov@feb.uni-sofia.bg`

ЗА ПРИБЛИЖЕНО КОРЕНУВАНЕ НА СТОХАСТИЧНИ МАТРИЦИ

Павел Т. Стойнов

Марковските вериги са популярно средство за моделиране на различни системи – естествени или изкуствени. Тяхна основна характеристика е матрицата на прехода, която е вид стохастична матрица. В тази статия е разгледан един специален вид матрици на прехода, които се използват в теорията на кредитния риск при моделиране на преходи от един кредитен рейтинг в друг. Основно тяхно свойство е, че вероятностната маса тук е съсредоточена около главния диагонал на матрицата. Предложена е процедура за коренуване, при което се получават матрици на прехода за период, два пъти по-малък от основния. Извършени са изчислителни тестове и резултати от тях са посочени в статията.

Ключови думи: матрица на прехода, кредитен рейтинг, приближен марковски модел