Estimation of Loan Portfolio Risk on the Basis of Markov Chain Model

Nikolay Timofeev, Galina Timofeeva

Estimation of Loan Portfolio Risk on the Basis of Markov Chain Model

Nikolay Timofeev[1] and Galina Timofeeva[2] *

[1] Ural State University of Railway Transport,
Kolmogorov str. 64, 620034, Ekaterinburg, Russia
[2] Ural Federal University,
Mira str., 19, 62002, Ekaterinburg, Russia

Abstract. A change of shares of credits portfolio is described by Markov chain with discrete time. A credit state is determined on as an accessory to some group of credits depending on presence of indebtedness and its terms. We use a model with discrete time and fix the system state through identical time intervals - once a month. It is obvious that the matrix of transitive probabilities is known incompletely. Various approaches to the matrix estimation are studied and methods of forecast the portfolio risk are proposed. The portfolio risk is set as a share of problematic loans. We propose a method to calculate necessary reserves on the base of the considered model.

Keywords: Loan portfolio, Markov chain, incomplete information

1 Introduction

Markov chain models [1] are widely used to explain the dynamics of state changes for different systems. Often they are used as a mathematical model for some random physical process.

Markov chains are used in finance and economics to model a variety of different phenomena, including asset prices, market crashes and credit portfolio dynamics [2, 3].

If the transition probability matrix of Markov chain is known then dynamics of the system states probabilities is completely described by a system of difference equations. As a rule the transition probabilities are unknown and estimated during the system evolution.

The most important indicator of a bank loan portfolio quality is a probability of default which is closely connected with a share of the problematic loans [5, 6]. A value of necessary reserves depends on quality and structure of the portfolio. On the one hand reserves should provide low probability of default, on the other hand they impact on profitability of the portfolio.

Let's assume that a change of shares of credits portfolio is described by Markov chain with discrete time. In this case the credit state is determined on

as an accessory to some group of credits depending on presence of indebtedness and its terms. We will use a model with discrete time and fix the system state through identical time intervals – once a month. It is proposed that the transition probabilities vary a little. Thus, we consider a stable economic situation when the transition probabilities are constant.

It is obvious that the matrix of transitive probabilities is known incompletely. Its values are estimated using a data on changes the quality of loans (a migration analysis of the portfolio). Criterion of a choice is accuracy of the forecast of a share of problematic loans.

2   Mathematical Model

2.1   Dynamics of System States

We consider a system with k states, the probability that the system is in i-th state at moment t denote by $x_i(t)$, $i = 1, \ldots, k$. Thus the following conditions hold:

$$0 \leq x_i(t) \leq 1, \quad x_1(t) + \ldots + x_k(t) = 1. \tag{1}$$

The dynamics of the system states probabilities is described by the discrete Markov chain model:

$$x_j(t+1) = \sum_{i=1}^{k} p_{ij} x_i(t), \quad t = 0, 1, \ldots, T, \tag{2}$$

where p_{ij} is the probability of transition from state i to state j in one step.

The first-order stationary Markov model for credit transitions is somewhat restrictive as a credit quality responds to changes in economics. Using a higher-order Markov process or a nonstationary transition probability matrix may be more appropriate, but in such models one should estimates too many parameters and the requirements to statistical data increase quite substantially. Thus the simple Markov chain is usually used in stable economic situation and with not longer time horizon [2, 3].

Let's denote by $x(t)$ a vector of states probabilities $x(t) = \{x_1(t), \ldots, x_k(t)\}^\top$, by P a matrix of the transition probabilities $P = \{p_{ij}\}$ and rewrite equation (2) in the vector form:

$$x(t+1) = P^\top x(t), \quad t = 0, 1, \ldots, T. \tag{3}$$

When the transition probability matrix P is known incompletely there is a problem to estimate $x(T)$. It is assumed that we have information about the number of transitions from i-th state to j-th on t step, $t = 1, \ldots, m$.

2.2   Ways to Select Groups

We consider two ways of describing a credit portfolio dynamics: a regular Markov chain in which we do not take into account repaid loans and renovation of the

portfolio and a scheme which included "new loan" and "repaid loan" as possible
states of a loan. In the first way we investigate a steady-state behavior of the
loan portfolio shares, in the second way we study the profitability of loans from
delivery to its repayment.

Let us consider a fist way of describing the states of loan. For beginning we
consider a simplified scheme with three groups of loans:

1. Loans without delay, including the new ones (S_1);
2. Overdue loans with $1 - 65$ days delay (S_2);
3. Non-performing (problematic) loans (S_3).
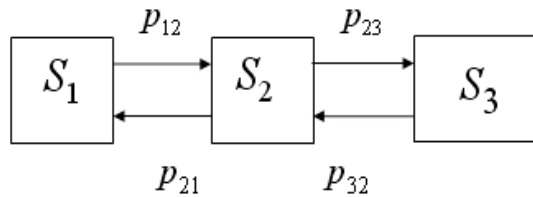
Fig. 1. Graph of the Markov chain for the simplified scheme.

A graph of the system state is in Fig.1. An investigation of this scheme allows
to define the most important features of the model. As a rule risk-managers use
more detailed schemes which takes into account a number of overdue days on
a loan. For example let consider the expanded scheme with 5 groups of loans
[7]: loans without delay, including a new one; overdue loans with 1 to 35 days
delay; overdue loans with 35 to 65 days delay; more than 65 days overdue loans
(problematic loans); reconstructed loans.

In these schemes the renovation and repayment of credits are considered
without allocating a separate state. New credits are included into the first group
together with credits without delay.

A loan is "reconstructed" in case it was problematic in the previous period
and a borrower has made partial payments under the credit.

In schemes with the repayments(amortization) there is the category "repaid
credits". The scheme on the basis of the expanded scheme of loans has 6 groups
of loans: loans without delay, including a new one; overdue loans with 1 to 35
days delay; overdue loans with 35 to 65 days delay; more than 65 days overdue
loans (problematic loans); reconstructed loans; repaid loans.

3   Estimation of Transition Probabilities

For the estimation of the probability p_{ij} one usually use the statistical data
about transitions from one state to another. Let's denote by $n_i(t)$ the number of

individuals who are in state i in period t, and by $n_{ij}(t)$ the number of individuals who were in state i in period $t-1$ and are in state j in period t. We can estimate the probability p_{ij} of an individual being in state j in period t given that they were in state i in period $t-1$.

The probability of transition $p_{ij}(t)$ from any given state i is approximated by a proportion of individuals that started in state i and ended in state j as a proportion of all individuals in that started in state i:

$$w_{ij}(t) = \frac{n_{ij}(t)}{n_i(t-1)}. \tag{4}$$

If the Markov chain is stationary (i.e. $p_{ij}(t) \equiv p_{ij}$) then one can use another estimate:

$$\tilde{w}_{ij} = \frac{\sum\limits_{t=1}^{T} n_{ij}(t)}{\sum\limits_{t=0}^{T-1} n_i(t)} \tag{5}$$

Using the methods described above, it is possible to estimate a transition matrix using count data.

Anderson and Goodman [4] showed that the estimator w_{ij} given by equation (5) is a maximum-likelihood estimator and calculated statistical moments of random values $\xi_{ij}(t) = n_{ij} - p_{ij}n_i(t-1)$.

On the t-th step $n_i(t-1) \triangleq n_i$, $i = 1, \dots, k$ are known. The statistical moment of w_{ij} are following [8]:

$$E(w_{ij}(t)) = p_{ij}, \tag{6}$$

$$Var(w_{ij}(t)) = \frac{p_{ij}(1-p_{ij})}{n_i},$$

$$Cov(w_{ij}(t), w_{il}(t)) = -\frac{p_{ij}p_{il}}{n_i}, \quad j \neq l, \tag{7}$$

$$Cov(w_{ij}(t), w_{cl}(t)) = 0, \quad i \neq c.$$

Here $E(\xi)$ is a mathematical expectation of a random value ξ, $Var(\xi)$ is its variance, $Cov(\xi, \zeta)$ is a covariance between ξ and ζ.

From relations (7) follows that w_{ij} and w_{cl} are noncorrelated if $i \neq c$.

Suppose that instead of observing the actual count of transitions from the different states, we only observe the aggregate proportions $y_i(t)$, which represent the proportion of observations with the state i. The aggregate proportions $y_i(t)$ estimate the system state probabilities: $y_i(t) \approx Nx_i(t)$, $i = 1, \dots, N$, $t = 1, \dots, T$, where N is a number of all individuals.

If the time series of observations T are sufficiently long, it is possible to estimate a transition matrix P from aggregate data using the least square method [9] and its generalizations [10].

The maximum-likelihood estimates in asset prices model is used to estimate transition matrices in credit risk modeling with a decades-old methodology that uses aggregate proportions data [2].

To specificate an order and a number of states in Markov chain model one may use criteria [4], but they are based on the detailed analysis of the data such as $n_{ijm}(t)$, where $n_{ijm}(t)$ is a number of individuals in state i at $t-2$, in j at $t-1$ and in m at t.

4 Approaches to Estimation the Share of Problematic Loans

We considered two ways to forecast the share of problem loans taking into account the uncertainty of the transition probabilities matrix. There are a confidence estimation method and a simulation method.

4.1 Confidence Estimation

The confidence estimation method consists of two stages: construction a confidence set Z_α for the transition probabilities matrix based on statistical data and relations (6)–(7) and the analysis of all possible trajectories of the system taking into account that the probabilities are constant but uncertain.

Let's estimate the elements of the transition probability matrix P. Denote the confidence region for $\{p_{ij} = z_s, \; i = 1, \ldots, k, \; j = 1, \ldots, k, \; j \neq i\}$ on m-th step by $Z_\alpha \subset \mathbb{R}^K$, $s = 1, \ldots, K$, $K = k(k-1)$.

Thus p_{ij} are the transition probabilities then $Z_\alpha \subset Z_+ \subset \mathbb{R}^K$, where Z_+ is the set of all possible values of transition probabilities $\{p_{ij}, \; j \neq i\}$

$$Z_+ = \{z_s : \;\; 0 \leq z_s \leq 1, \;\; \sum_{s=1}^{K} z_s \leq 1\} \subset \mathbb{R}^K.$$

Estimation for p_{ii} follows from the equalities

$$p_{i1} + \ldots + p_{ik} = 1, \quad i = 1, \ldots, k. \tag{8}$$

In the considered model $p_{ij} = z_s$ are distributed approximately normal [4] with mean values equal to $w_{ij} \triangleq \bar{z}_s$ and a covariance matrix G defined by relations (7) with substitution p_{ij} by w_{ij}.

Therefore we may use the confidence set Z_α defined by joint restrictions:

$$Z_\alpha = \{z \in Z_+ : (z - \bar{z})^\top G(z - \bar{z}) \leq b_K^{(\alpha)}\}, \tag{9}$$

where $b_K^{(\alpha)}$ is the α-quantile of χ^2 distribution with K degrees of freedom.

The next step is to solve the state estimation problem of a multistage deterministic system with uncertain matrix P:

$$x(t+1) = P^\top x(t), \quad t = m, \ldots, T,$$
$$x(m) = x^*, \;\; P \subset Z. \tag{10}$$

and to find an information set

$$X(t, Z) = \{x \in R_+^k : x = (P^\top)^{T-m} x^*, \quad P \in Z\}. \tag{11}$$

We may construct information sets for system (10) using approaches proposed by Kurzanski and Tanaka [11].

This method is very time-consuming because the number of estimated elements of the matrix is large. We can perform calculations only for the scheme with 3 groups of loans, in which only 4 probabilities should be estimated. For schemes 2 and 3 the number of estimated probabilities are 9 and 17 respectively.

Example 1. Let us consider the estimation problem for the scheme with 3 groups of loans (Fig.1). For this scheme the matrix of transition probabilities has a form

$$P = \begin{pmatrix} 1 - p_{12} & p_{12} & 0 \\ p_{21} & 1 - p_{21} - p_{23} & p_{23} \\ 0 & p_{32} & 1 - p_{32} \end{pmatrix}, \tag{12}$$

and only 4 transition probabilities p_{ij} should be estimated. Denote them as z_s, $s = 1, \ldots, 4$:

$$p_{12} = z_1, \quad p_{23} = z_2, \quad p_{32} = z_3, \quad p_{21} = z_4. \tag{13}$$

From a statistical data we estimate the mean values $\bar{z}_s = w_{ij}$, where w_{ij} defined by (4). A covariance matrix is calculated using relations (7), where estimates w_{ij} substituted instead of values p_{ij}:

$$G = \begin{pmatrix} \frac{\bar{z}_1(1-\bar{z}_1)}{n_1} & 0 & 0 & 0 \\ 0 & \frac{\bar{z}_2(1-\bar{z}_2)}{n_2} & 0 & -\frac{\bar{z}_2 \bar{z}_4}{n_2} \\ 0 & 0 & \frac{\bar{z}_3(1-\bar{z}_3)}{n_3} & 0 \\ 0 & -\frac{\bar{z}_2 \bar{z}_4}{n_2} & 0 & \frac{\bar{z}_4(1-\bar{z}_4)}{n_2} \end{pmatrix}. \tag{14}$$

Then find a confidence set for transition probabilities p_{ij} in ellipsoidal form (9). For the considered scheme with 3 groups of loans we obtain

$$Z_\alpha = \{z_s \in R_+^4 : (z - \bar{z})^\top G(z - \bar{z}) \le b_4^{(\alpha)}\}. \tag{15}$$

Then we find an information set for system (5) for a given $T = 12$ using ellipsoidal calculus [12]. Thus we get the confidence set for the portfolio shares after 12 months $x(12)$ and for the share of problematic loans $x_3(12)$ in particularly. For our data we obtain $x_3(12) \in [0.03; 0.152]$ with probability $\alpha = 0.95$.

4.2 Simulation Method

This method has 4 stages:

1. We determine statistical moments (4) and (5) for the transition probabilities based on statistical data on transitions between loan states;

2. Generate the random vector of unknown probabilities as the Gaussian vector with the given statistical moments;
3. Simulate the dynamics of system (2) with the large number of runs;
4. On the basis of the calculations determine the mean value and the confidence interval for the share of problem loans.

A histogram of the predicted share of problem loans after 12 months one can see in Fig. 2. On the base of modeling we find the expected mean of value
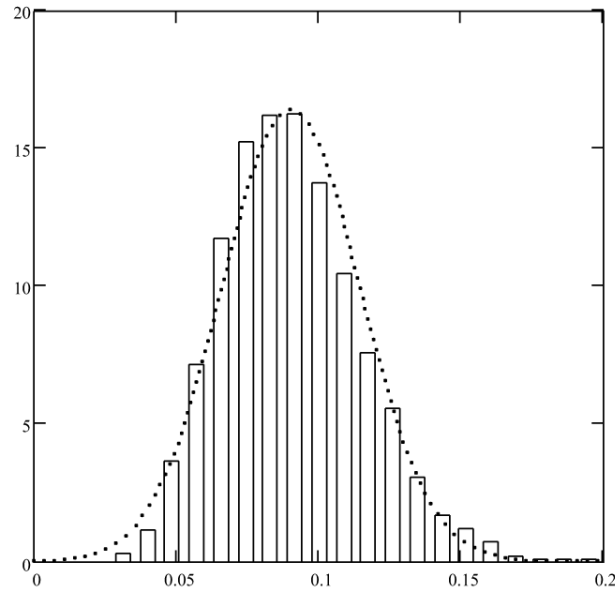
Fig. 2. Histogram of the share of problematic loans and normal density.

of the share of problem loans after 12 month: $x(t + 12) = 0.09$ and quantile $q_{0.95} = 0.135$, i.e. $x(t + 12) < 0.135$ with the probability $\alpha = 0.95$.

This result is more precise than the confidence interval obtained by the confidence estimation, because the confidence interval for the share of problematic loans, obtained in section 4.1, depends on the form of confidence set for unknown parameters p_{ij} and the ellipsoidal form is not optimal in this case.

5 Estimation of a reserve

According recommendations of the International Committee [13] bank portfolio managers should estimate a risk of the portfolio and form reserves in a proportion of the its value.

There are different approaches to estimate the necessary reserves. Some of them based on probability of "nonreconstruction" of the problematic credits. Let's denote by $p_{ij}^{[t]}$ a probability of transition from the i-th state to j-th state by t steps. It is an element of the t power of the transition probability matrix $P^{[t]} = P^t$. Denote by $D^{[t]} = p_{mm}^{[t]}$, where m-th state is "problematic loans". The value $1 - D^{[T]}$ is called a probability of reconstruction during a period T.

The value of risk for j-th group of loans is defined as

$$\hat{r}_j = p_{jm}^{[\tau_1]} D^{[\tau_2]}, \text{if } j \neq m,$$

$$\hat{r}_m = D^{[\tau_1]}.$$

The reserve W is equals to: $W = W_1\hat{r}_1 + \ldots + W_k\hat{r}_k$, where W_j is a sum of loans in j-th group.

Risk managers take different τ_1 and τ_2, such as 6, 12 or 24 months. The problem is how to choose the period of transition to the problematic state and the reconstruction period. In some approaches value of τ_1 depends on j and a period of an overdue.

We propose another approach taking into account time structure of possible losses. Let's define a risk of loans in j-th group as maximum of a sum of problematic loans in future for loans of this group. Thus, the risk for j-th loans group equals

$$r_j = \max_{t \in 0, \ldots, T} \frac{1}{(1 + \rho)^t} p_{jm}^{[t]}, \tag{16}$$

where ρ is a month discount factor.

For a new loan we have

$$r_0 = \max_{t \in 0, \ldots, T} \frac{1}{(1 + \rho)^t} p_{0m}^{[t]}. \tag{17}$$

We can take into account that transition probabilities are known incompletely and instead of (16) use its quantile:

$$q_j(\alpha) : \mathcal{P}\{\max_{t \in 0, \ldots, T} \frac{1}{(1 + \rho)^t} p_{jm}^{[t]} \leq q_j(\alpha)\} \geq \alpha, \tag{18}$$

where $\mathcal{P}(A)$ is a probability of a random event A. The quantile may be calculated using the confidence approach or the simulation method (see Sect. 4).

Example 2. Let's consider a scheme with repayment (Scheme 3) and calculate the value of reserves for new loans using the proposed approach. Is proposed that we may estimate statistical moments (6) of the transition probabilities p_{ij} based on a previous data and estimates (4), (5) or their modifications.

We take a possible value of transitions probabilities matrix $P = \{p_{ij}\}$, calculate

$$h(t) = \frac{1}{(1 + \rho)^t} p_{0m}^{[t]}, \quad t = 0, \ldots, T,$$

and find their maximum $r_0(t)$ which depended on matrix P.
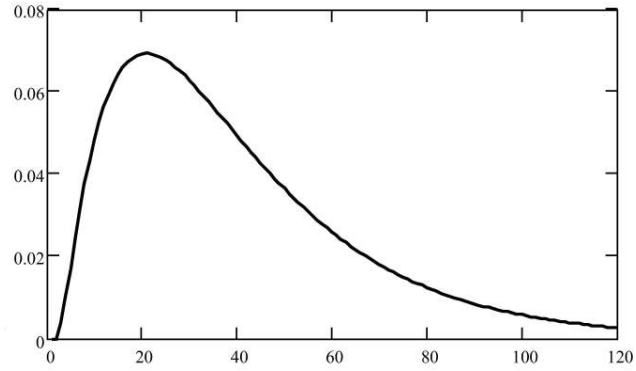
Fig. 3. Dependence of the discounted sum of problem loans $h(t)$ on time.

In Fig. 3 one can see changes of $h(t)$ (the discounted sum of problem loans) for a fixed matrix P. Values of $h(t)$ decreasing for $t \geq 20$.

We generate many times transition probabilities according normal distribution with the given statistical moments, calculated the discounted sum of problem loans $h(t) = h(t, P)$, and its maximum $r_0 = r_0(P)$, then we obtain estimates of a mean value and quantile (18).

In Fig.4 one can see a histogram for risk estimates $r_0 = r_0(P)$. For considered data mean value of risk $\bar{r}_0 = 0.039$, $q_{0.95} = 0.065$.
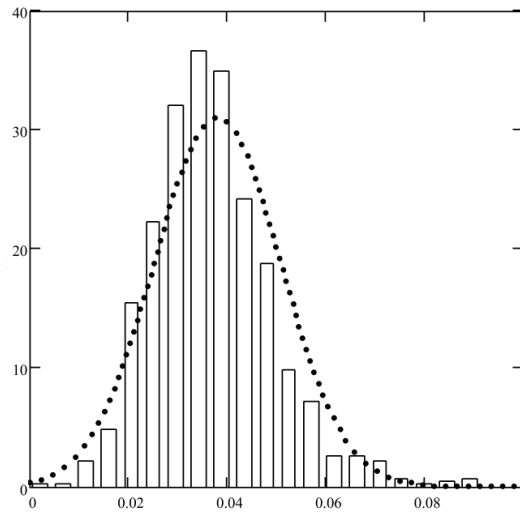
Fig. 4. Histogram of risk $r_0(P)$ and normal density

6   Conclusion

We studied the discrete Markov chain model for loan portfolio. It is proposed that transition probabilities are unknown and estimated during the process. We proposed methods to estimate system state probabilities in future. Obtained results apply to forecast credit portfolio shares and to define necessary reserves.

References

1. Markov, A.A.: Rasprostranenie zakona bol'shih chisel na velichiny, zavisyaschie drug ot druga. Izvestiya Fiziko-matematicheskogo obschestva pri Kazanskom universitete. Ser. 2, 15, 135–156 (1906) Reprinted in: Markov, A.A.: Extension of the limit theorems of probability theory to a sum of variables connected in a chain. Appendix B of: Howard, R. Dynamic Probabilistic Systems. 1: Markov Chains. John Wiley and Sons (1971)
2. Jones, M.T.: Estimating Markov Transition Matrices Using Proportions Data: An Application to Credit Risk. IMF Working Paper, WP-05-219 (2005)
3. Thyagarajan, V., Saiful, M.: Retail Banking Loan Portfolio Equilibrium Mix: A Markov Chain Model Analysis. Amer. J. of Applied Sciences. 2(1), 410–419 (2005)
4. Anderson, T. W., Goodman, L. A.: Statistical Inference About Markov Chains. Annals of Mathematical Statistics. 28, 89–110 (1957)
5. Hanson, S., Schuermann, T.: Confidence intervals for probabilities of default, Journal of Banking & Finance, Elsevier, 30(8), 2281-2301 (2006)
6. De Andrade F.W.M., Thomas L.: Structural models in consumer credit, European Journal of Operational Research, 183, 1569-1581 (2007)
7. Timofeev, N.A.: Mathematical Model of the Vintage Analysis of Banking Credit Portfolio. Gerald of USURT, 9(1). 61–69 (2011)
8. Timofeeva, G.A., Timofeev, N.A.: Predicting the Components of Credit Portfolio Based on a Markov Chain Model. Automation and Remote Control, 73(4), 637–651 (2012)
9. Kalbfleisch, J. D., Lawless J. F.: Least-Squares Estimation of Transition Probabilities From Aggregate Data, Canadian J. of Statistics, 12(3), 169–182 (1984)
10. MacRae E. Ch.: Estimation of Time-Varying Markov Processes with Aggregate Data, Econometrica, 45(1), 183–98 (1977)
11. Kurzanski, A.B., Tanaka, M.: On a unified framework for deterministic and stochastic treatment of identification problem. Laxenburg, IIASA (1989)
12. Kurzhanskiy, A. A., Varaiya P.: Ellipsoidal Techniques for Reachability Analysis of Discrete-Time Linear Systems, IEEE Transactions on Automatic Control, 52(1), 26–38 (2007)
13. Basel Committee on Banking Supervision, International Convergence of Capital Measurement and Capital Standards: A Revised Framework, `http://www.bis.org/publ/bcbs107.htm` (2004)