Forecasting Credit Portfolio Components with a Markov Chain Model

G. A. Timofeeva* and N. A. Timofeev**

* Ural Federal University, Yekaterinburg, Russia
** Ural State Academy of Railway Transport, Yekaterinburg, Russia

Received June 6, 2011

Abstract—We consider the forecasting problem for components of a bank’s credit portfolio, in particular, for the share of non-performing loans. We assume that changes in the portfolio are described by a Markov random process with discrete time and finite number of states. By the state of a loan we mean that it belongs to a certain group of loans with respect to the existence and duration of arrears. We assume that the matrix of transitional probabilities is not known exactly, and information about it is collected during the system’s operation.

DOI: 10.1134/S0005117912040042

1. INTRODUCTION

Dynamics of a large number of physical, technical, and economic systems can be described with Markov random processes [1]. One of the simplest dynamical models in this class is a Markov chain with discrete time and finite number of states. Despite its simplicity, the model is widely used in estimating market indices [2, 3], modeling credit portfolio dynamics [4, 5], and studying the stability of a number of technical systems. The main problem arising in modeling a real-life system with Markov chains is the estimation of its matrix of transition probabilities based on statistical data [6]. The accuracy of the system state probabilities forecast depends on the accuracy of our estimates of the transition probabilities.

Credit portfolio analysis includes systematic study and observation of the credit activities in a bank that let one estimate the composition and quality of bank loans in its dynamics. Credit portfolio management is based on a system of parameters that characterize a bank’s activity in the moneylending business. The basic parameters of a credit portfolio are its profitability and the share of non-performing loans. A loan is called non-performing if the payment delays have exceeded a certain critical value (say, 90 days).

Each bank analyzes these parameters based on its own experience and analytic capabilities, using the instruments developed in the Russian and world banking practice [7].

The mathematical model that captures the dynamics of a credit portfolio is based on describing the changes in state for a single “random” loan. By the state of a loan we understand that it belongs to a certain group of loans with respect to the existence and duration of arrears. We assume that a loan’s behavior (with respect to the payments) does not depend on the prehistory and is essentially random, i.e., the process of loan state transitions can be described with the theory of Markov random processes.

In this work, we use a model with discrete time, recording loan states in identical time intervals, say once a month. Changes in the parameters of a credit portfolio are studied on a time interval representing a stable state of the banking system. We assume that probabilities of transitions from state to state do not change significantly during this interval, so we use a model with constant transition probabilities.
We assume that transition probabilities are not known in advance and are estimated during a system’s operation. The purpose of this analysis is to estimate possible losses associated with moneylending and the share of non-performing loans in a given time interval. Note that these functions linearly depend on the system state vector. We have selected the quantile criterion \[8\] as our objective, i.e., we estimate the value which the objective function does not exceed with a given probability.

To solve the problem, we propose to estimate the matrix of transition probabilities by constructing confidence intervals for the matrix elements and by imitational modeling. The confidence intervals approach lets one use, in order to estimate the system state vector, modern theory of information sets \[9–12\]. On the other hand, imitational modeling based on an analytic model together with estimation methods for the objective function’s quantile \[13\] yields more accurate results for this specific problem.

\[2.\] MATHEMATICAL PROBLEM SETTING

In analyzing the dynamics of the share of non-performing loans, different authors use different loan classifications with respect to the delay in payment and recovery level \[4\]. We consider a credit portfolio consisting of loans falling into \(k\) different groups. That is, we assume that each loan has \(k\) possible states. We denote the probability that a given randomly selected loan is at the \(i\)th state at time moment \(t\) by \(x_i(t), \ i = 1, \ldots, k\). The following conditions hold:

\[
0 \leq x_i(t) \leq 1, \quad x_1(t) + \ldots + x_k(t) = 1.
\]

(1)

We denote by \(p_{ij}(t)\) the probability of transition for a loan that was in state \(S_i\) at time moment \(t\) into a state \(S_j\) in one step. We assume that this probability is constant, i.e., \(p_{ij}(t) \equiv p_{ij}\). Thus, the dynamics of each individual loan is described by a discrete time Markov chain:

\[
x_j(t+1) = \sum_{i=1}^{k} p_{ij} x_i(t), \quad t = 0, 1, \ldots, T.
\]

(2)

We denote by \(\{e_1, \ldots, e_k\}\) the unit basis in \(\mathbb{R}^k\). By \(\xi(t)\) we denote the system state vector at time moment \(t\). If a loan is in the \(i\)th state then \(\xi(t) = e_i\). Thus, state probabilities can be written as

\[
x_i(t) = \mathcal{P}\{\xi(t) = e_i\}, \quad i = 1, \ldots, k.
\]

(3)

Here and in what follows we denote by \(\mathcal{P}(A)\) the probability of an event \(A\).

We denote the state probabilities vector by \(x(t) = \{x_1(t), \ldots, x_k(t)\}^\top\), where \(^\top\) denotes taking the transpose; we also denote the \((k \times k)\) matrix of transition probabilities by \(P = \{p_{ij}\}\). Dynamics of system (2) can now be written as

\[
x(t+1) = P^\top x(t), \quad t = 0, 1, \ldots, T.
\]

(4)

We assume that the matrix \(P\) is not known exactly and is being estimated during the system’s operation.

Quality criteria for a credit portfolio include its profitability and risk defined as the share of non-performing loans, so we assume the objective function to be linear.

Let \(N\) be the number of loans in the portfolio. Consider the problem of estimating the parameters of the distribution of the sum of identically distributed random vectors. We denote

\[
\zeta_N(t) = \sum_{n=1}^{N} \xi^{(n)}(t),
\]