**Косарев Владислав /Магистр ВШЭ**

**Using non-parametric aproach for modelling credit risk**

In the paper, we consider the approach to credit risk modeling as an example of consumer lending. Credit risk - the risk associated with the debtor's failure to meet its obligations to the bank. The main tool for credit risk evaluation is the migration matrix - a matrix whose elements are characterized by a change in the behavior of the debtor. For example, if the debtor has missed payment 2 (a 3rd category of quality[[1]](#footnote-1)), then it has 4 possible variants of behavior:

* Migrating from 3 to 2. The debtor pays 1 payment and then has 1 missed payment (2nd category of quality);
* Migration from 3 to 1. The debtor pays 2 payment and then he has not missed payments (1st category of quality);
* Migrating from 3 to 0. The debtor closes a loan contract and then it does not have obligations to the bank;
* Migrating from 3 to 4. The debtor does not pay the next payment and have 3 missed payment (quality category 4)

Each of these event occurs with a certain probability, the probability of these events are the elements of the migration matrix. It is crucial to know the distribution of loans by quality categories (number of missed payments) for the bank. The bank needs more capital to cover credit risk loses, if the bank's loan portfolio consists of loans with a large number of delayed payments. So, we have to build migration matrix model in order to construct credit risk model.

We suppose that migration matrix decomposed on 3 curves:

* Exogenous influence - a function depending on the calendar date
* Cohort effect (vintage quality curve) - a function depending on the date of loan’s issue
* Ageing - a function which depends on the age of the loan (age is defined as the number of full months between the time of issue and the current calendar date)

It is assumed that the exogenous function affect on each cohort and the vintage quality effect is different for each generation of loan’s issue. This type of model is very popular in epidemiology and biostatistics (Yang et. Al., 2013).

It should be noted that for this kind of decomposition we have to use only actual values of migration matrix. But this data should be presented in a dual time dynamic (age and calendar date).

To obtain three factors, discussed above, we use several non-parametric estimation methods such as, GAM (Generalized additive model) (Wood, 2006) and VGAM (Vector generalized additive model) (Yee, 2015). In fact, for decomposition it is necessary to solve the following optimization problem:

$$ε=δ\left(y\_{ij}\left(m,e,v\right)\right)-f\_{ij}\left(m\right)+g\_{ij}\left(e\right)+h\_{ij}\left(v\right)\rightarrow min$$

$$0\leq y\_{ij}\left(m,e,v\right)\leq 1  (1)$$

$$\sum\_{i=1}^{n}y\_{ij}\left(m,e,v\right)=1            (2)$$

$$m=v-e                          (3)$$

$y\_{ij}$ – probability of transaction from *i* to *j* quality category;

$δ$ – link fuction, for example logit transformation;

$f\_{ij}\left(m\right)$ – ageing function for transaction from *i* to *j* quality category;

*m* – loan's age;

$g\_{ij}\left(e\right)$ – exogenous influence function for transaction from *i* to *j* quality category;

$e$ – calendar date;

$h\_{ij}\left(v\right) $– cohort effect (vintage quality curve) for transaction from *i* to *j* quality category;

$v$ – date of loan's issue

The main task is to find $f\_{ij}$, $g\_{ij}$, $h\_{ij} , $which minimized $ε$ in conditions with (1), (2), (3).

During the study were founded that all factors obtained by decomposition had a clear economic interpretation. The function of the endogenous influence can be modeled as a function of macro variables, such as unemployment and real incomes. The function of vintage quality correlated with probability of default (indicator of the loan’s generation quality).

In the report, we consider the possibility of using this approach for the regulator (Central Bank, shareholders) and commercial organizations (banks, microfinance institutions, collection agencies). For example, if we know the function of exogenous influences, it is easy to understand how macro factors affect to the bank's loan portfolio, and the vintage quality function can help us to validate credit scorecard easily. Thus, it is easy to develop a risk management strategy: increasing quality during the difficult economic impact, and reducing the quality at a favorable influence, ie, top-management can reach the target levels of risk simply.

In conclusion, we show you some examples of using the migration matrix decomposition for scenario modeling of credit and interest spread risks.

1. [↑](#footnote-ref-1)