A Dynamic Credit Scoring Model Based on Contour Subspaces

Kirill Romanyuk
National Research University Higher School of Economics
Saint Petersburg, Russia
Email: kromanuk@yandex.ru

Abstract—Credit scoring is generally used as a binary tool to approve or decline credit applicants in the consumer lending market, which makes it inapplicable for risk-based pricing. Since client’s creditworthiness can vary in a continuous set and changes over time, banks should possess an option to adjust credit terms appropriately in order to manage credit risk properly and to attract clients with high creditworthiness. In the given paper a dynamic credit scoring model based on contour subspaces is proposed. In this model, credit terms form a contour subspace for each creditworthiness value. Application of the proposed model is supposed to add advanced options for decision support systems in loan granting, i.e. to visualize a contour subspace of credit terms for a client in connection with an individual creditworthiness value, provide options to choose credit terms from this contour subspace, and manage credit terms online in correlation with the dynamics in creditworthiness estimation. Potential application of the proposed model in the mortgage lending market is shown.

Keywords—risk-based pricing; contour subspaces; credit terms; credit scoring; behavioral scoring;

I. INTRODUCTION

Scientific interest in the retail credit sector has increased for the last decade. Many articles have been written on finding the ideal credit scoring method (“search for the silver bullet” [1]). Although many of these research papers are of high quality, it still seems unrealistic to develop a perfect credit scoring method because of a complexity of the retail credit industry: every human is unique, does not always act rationally, and usually lives in a community, which makes it even harder to take all possible interactions into account. Extremely high performance may be a result of oversimplification of the model and achievable only in a laboratory by excluding important features [2] or by squeezing the maximum performance from a method [3], which can lead to overfitting and also inappropriate results in practice.

Generally, credit scoring is a binary classification technique to distinguish between good and bad credit risk classes [4], [5]. Behavioral scoring is a technique to predict probability of switching a class by a borrower. While banks utilize credit scoring before granting a loan, behavioral scoring is applied when loans are in use. Thus, banks possess techniques to control an entire lifecycle of a loan.

However, both types of scoring are very limited for risk-based pricing. For example, 100 clients have applied for a similar loan in a bank (i.e. with the same loan type, loan period, and loan amount). While 10 clients were classified as clients with a bad credit risk, 90 clients were classified as clients with a good credit risk and were granted a loan. Since these clients are indistinguishable in terms of creditworthiness (i.e. they obtained the same credit risk label), they will get the same loan rate. Even though a bank calculates an average loan premium for a pool of these clients, a loan premium for most of the clients will still differ from the average value and may vary significantly for some clients. As a result, a portfolio of consumer loans possesses a loan rate in accordance with risk-based pricing, but each borrower pays a loan rate that most likely does not reflect his creditworthiness.

Usually a client’s private information is combined with credit attributes in credit scoring in order to support a binary decision, i.e. to grant a loan or not [6], [7]. The proposed model is based on separation of client’s creditworthiness and riskiness of credit terms. Application of this model allows constructing a contour subspace of credit terms for every creditworthiness value. Consequently, a bank do not need to fully recalculate results if a client applies for a different type of a credit. In addition a bank will be able to modify credit terms in correlation with a creditworthiness value for credits in use (e.g. in the mortgage credit sector) that will attract clients with high creditworthiness.

II. A CREDIT RISK MODEL

The proposed credit risk model is defined as a set of contour subspaces of credit terms (1). On the one hand, each combination of a loan rate (P), a loan amount (M), and a loan period (T) are corresponded to a credit risk value. On the other hand, a client’s credit risk is measured by a creditworthiness estimation (Q). Such a comparison of variables is in accordance with a credit risk and creditworthiness definitions [8], [9]. A credit risk and creditworthiness can be defined as a measure of a likelihood that a client will default on a loan.

\[ F(P,M,T) = Q(x) \] (1)

Credit terms vary in a continuum within the limits that are set up by a bank (2). Naturally, the question arises what the range set of a credit risk function is. Due to considerable importance of this range for constructing contour subspaces in the model, the next subsection is dedicated to answer this question.

\[ P \in [P_{min}, P_{max}], M \in [0, M_{max}], T \in [0, T_{max}] \] (2)
A. Credit risk functions

The range set of a credit risk function defines a number of contour subspaces in the model. Thus, if the range of a credit risk function is binary as is usually the case in credit scoring, then the model can not possess more than two contour subspaces (3).

\[ F_2 : P \times M \times T \rightarrow \{0, 1\} \]  \hspace{1cm} (3)

Similarly as the most common credit scoring model can be defined as a binary mapping function \( f \) (4) over an input feature space [10].

\[ f : X \rightarrow \{\text{good}, \text{bad}\} \]  \hspace{1cm} (4)

However, clients with bad creditworthiness should not be granted with a loan, since they represent the highest credit risk. In this case, credit terms correspond to two contour subspaces, where one of them is trivial.

If the range of a credit risk function contains three values, then the credit risk model can not possess more than three contour subspaces (5).

\[ F_3 : P \times M \times T \rightarrow \{0, \frac{1}{2}, 1\} \]  \hspace{1cm} (5)

Suppose, the range of a credit risk function contains \( n \) values (6), and a creditworthiness estimation contains at least \( n \) values (7), then the credit risk model possesses \( n-1 \) nontrivial contour subspaces concerning the fact that clients with bad creditworthiness are not granted with a loan.

\[ F_n : P \times M \times T \rightarrow \left\{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, \frac{n-2}{n-1}, 1\right\} \]  \hspace{1cm} (6)

\[ Q_n : x \rightarrow \left\{0, \frac{1}{n-1}, \frac{2}{n-1}, \ldots, \frac{n-2}{n-1}, 1\right\} \]  \hspace{1cm} (7)

Theorem. Let \( D \) (8) be a subset of \( R^3 \) and let \( \{F_n\} \) (6) be a subsequence of real valued functions defined on \( D \) then \( \{F_n\} \) converges uniformly to the function \( F \) (9).

\[ D = \{d = (P, M, T) | P \in [P_{\text{min}}, P_{\text{max}}], M \in [0, M_{\text{max}}], T \in [0, T_{\text{max}}]\} \]  \hspace{1cm} (8)

\[ F : P \times M \times T \rightarrow [0, 1] \]  \hspace{1cm} (9)

Proof. For any \( d \) in \( D \) difference between \( F_n \) and \( F \) is less than a larger unit value of functions range (10). In the limit this difference is equal to 0 when \( n \) is sent to the infinity (11).

\[ |F_n(d) - F(d)| < \frac{1}{n-1}, \forall d \in D \]  \hspace{1cm} (10)

\[ \lim_{n \to +\infty} \frac{1}{n-1} = 0 \]  \hspace{1cm} (11)

For any given \( \varepsilon > 0 \), there exists a natural number \( N_\varepsilon \), such that \( |F_n(d) - F(d)| < \varepsilon \) for every \( n > N_\varepsilon \). Consequently, \( \{F_n\} \) converges uniformly to the function \( F \) (12) which was to be proved.

\[ \{F_n\} \Rightarrow F \iff \lim_{n \to +\infty} \sup_{d \in D} |F_n(d) - F(d)| = 0 \]  \hspace{1cm} (12)

Similar theorem can be proven for a subsequence of real valued functions \( \{Q_n\} \) (7), where each of one represents a creditworthiness estimation.

Theorem. Let \( \{Q_n\} \) be a subsequence of real valued functions defined on \( X \) (13) then \( \{Q_n\} \) converges uniformly to the function \( Q \) (14).

\[ X = \{x = (x_1, x_2, \ldots, x_n) | x_i \in [x_i^{\text{min}}, x_i^{\text{max}}], i = 1, 2, \ldots, l, j = l + 1, l + 2, \ldots, n\} \]  \hspace{1cm} (13)

\[ Q : x \rightarrow [0, 1] \]  \hspace{1cm} (14)

Proof. For any \( x \) in \( X \) difference between \( Q_n \) and \( Q \) is less than a larger unit value of functions range (15). In the limit this difference is equal to 0 when \( n \) is sent to the infinity (16).

\[ |Q_n(x) - Q(x)| < \frac{1}{n-1}, \forall x \in X \]  \hspace{1cm} (15)

\[ \lim_{n \to +\infty} \frac{1}{n-1} = 0 \]  \hspace{1cm} (16)

For any given \( \varepsilon > 0 \), there exists a natural number \( N_\varepsilon \), such that \( |Q_n(x) - Q(x)| < \varepsilon \) for every \( n > N_\varepsilon \). Consequently, \( \{Q_n\} \) converges uniformly to the function \( Q \) (17) which was to be proved.

\[ \{Q_n\} \Rightarrow Q \iff \lim_{n \to +\infty} \sup_{x \in X} |Q_n(x) - Q(x)| = 0 \]  \hspace{1cm} (17)

As a result, an application of functions \( F \) and \( Q \) in a credit risk model allows to construct an infinite number of contour subspaces.
B. Constructing a credit risk function

A set of contour subspaces (1) can be constructed as follows. A set \( S = \{ l_i, \bar{Q}_i \}_{i=1}^{n} \) should be plotted in the axes of a loan rate (\( P \)), and a loan period (\( T \)), and a loan amount (\( M \)), where \( \bar{Q}_i \) is a creditworthiness estimation of a client that has taken a credit \( i \), and \( l_i \) is a loan loss (\( L_i \)) divided by a loan amount (\( M_i \)) of a credit \( i \).

Then, creditworthiness value \( \bar{Q}_i \) is plotted in the point \((l_i, T_i, M_i)\), where a loan loss rate \((l_i)\) is plotted on a loan rate axis. Subsequently, a similar creditworthiness values can be approximated by a subspace, e.g. with a polynomial approximant (18).

\[
P_i = F_P^{-1}(Q, T, M) = \sum_{i=0}^{n} (a_i M^i + b_i T^i + c_i Q^i),
\]

\[a_i, b_i, c_i \in \mathbb{R}, i = 0, 1, ..., n \] (18)

It should be noted that the loan rate is expected to be an increasing function of a loan amount, and a loan period, and conversely a decreasing one of a creditworthiness estimation. A similar approximation can be made for a loan amount including loan amount limit (19).

\[
M_i^{lim} = F_M^{-1}(Q, T, P) = \sum_{i=0}^{n} (w_i P^i + u_i T^i + \nu_i Q^i),
\]

\[w_i, u_i, \nu_i \in \mathbb{R}, i = 0, 1, ..., n \] (19)

C. Applications of the credit risk model

Banks can apply the credit risk model to develop decision support systems that will produce the following results:

1) credit terms visualization for clients according to their individual creditworthiness value by a contour subspace of credit terms;
2) credit terms set up according to a client’s individual creditworthiness value;
3) credit terms management depending on the dynamics in creditworthiness.

Firstly, a decision support system calculates a creditworthiness estimation solely on clients private information. Then, the value of an individual creditworthiness estimation \((Q_i)\) is set up in the credit risk model in order to obtain a contour subspace of credit terms. This contour subspace can be visualized as a figure and the domain of credit terms.

Secondly, concerning the fact that a contour subspace defines credit terms combinations with an equal credit risk value a bank is indifferent what combination will choose a client from this contour subspace. Thus, a client can choose the value of any credit term in the domain. The decision support system fixes this value, which can graphically be represented as drawing a hyperplane through a contour subspace, and this value. A new contour subspace is less by one dimension. Then, a client chooses the value for another credit term. The decision support system comes through a similar procedure. While \( n - 1 \) credit terms were set up by a client from the domain of credit terms, the last one is defined by the decision support system.

Thirdly, a decision support system can manage credit terms of credit line or credit card online by the dynamics in creditworthiness. Suppose, that a value of a creditworthiness estimation has been changed for the period \( \Delta t \) from \( \alpha \) to \( \beta \), e.g. by means of information provided by employer via the internet. Thus, a decision support system can produce the following recommendations:

1) to alter a loan amount limit on a value \( \Delta M_i^{lim}(t) \), ceteris paribus \((P(t-1) = P(t) = P, T(t-1) = T(t) = T)\), where \( T(t) \) is a loan period in the moment \( t \), which can be increased, if credit contract is modified (20).

\[
\Delta M_i^{lim}(t) = M_i^{lim}(t) - M_i^{lim}(t-1) = \delta \frac{M}{M(T)} \frac{P(T)}{P(T)} \frac{Q(T)}{Q(T)}
\]

\[
\delta = F^{-1}(P, T, \beta) - F^{-1}(P, T, \alpha) = \sum_{i=0}^{n} (w_i P^i + u_i T^i + \nu_i \beta^i) - \sum_{i=0}^{n} (w_i P^i + u_i T^i + \nu_i \alpha^i) = \sum_{i=0}^{n} \nu_i(\beta^i - \alpha^i)
\] (20)

2) to alter a loan rate for taking a new amount of money from credit line or credit card on a value \( \Delta P_i(t) \), ceteris paribus \((M(t-1) = M(t) = M, T(t-1) = T(t) = T)\) (21).

\[
\Delta P_i(t) = P_i(t) - P_i(t-1) = \delta \frac{M}{M(T)} \frac{P(T)}{P(T)} \frac{Q(T)}{Q(T)}
\]

\[
\delta = F^{-1}(M, T, \beta) - F^{-1}(M, T, \alpha) = \sum_{i=0}^{n} (a_i M^i + b_i T^i + c_i \beta^i) - \sum_{i=0}^{n} (a_i M^i + b_i T^i + c_i \alpha^i) = \sum_{i=0}^{n} c_i(\beta^i - \alpha^i)
\] (21)

III. Examples

Applications of the proposed credit risk model for defining credit terms are shown in the examples for following cases:

1) different clients;
2) a client in general;
3) a client in the mortgage lending market.

For more clarity, suppose the initial credit risk model reduces to a set of contour lines that represent combinations of a loan amount and a loan rate for every creditworthiness value (22). The functions for a loan rate (23) and a loan amount (24) will also change correspondingly.

\[
F(P, M) = Q
\] (22)

\[
P_i(t) = F_P^{-1}(M_i(t), Q_i(t))
\] (23)

\[
M_i^{lim}(t) = F_M^{-1}(P_i(t), Q_i(t))
\] (24)
A. Controlling credit terms for different clients

Suppose there exist three clients with a different creditworthiness value (25). Each client obtains a different contour line of credit terms (Fig. 1).

\[ Q_1 < Q_2 < Q_3 \]  

Thus, each of these clients obtains a different credit limit. The more a creditworthiness estimation is, the less a credit amount limit is (26), because a loan rate is an increasing function of a loan amount.

\[ M_{lim}^{1} < M_{lim}^{2} < M_{lim}^{3} \]  

Each client can choose credit terms on a contour line that corresponds to the individual creditworthiness value. If the first client decides to take a loan in the volume \( M_1 \), then a decision support system draws a hyperplane through this value in Figure 1, which is a line of two credit terms in the given example. Thus, the loan rate is set up at the \( P_1 \) value (27).

\[ P_1 = F_{P}^{-1}(M_1, Q_1) \]  

B. Controlling credit terms for a single client

Suppose, a client \( j \) applies for a credit in a bank. A decision support system based on the credit risk model sets up and modifies credit terms as follows (Figure 2):

- Point \( A_0 \). A client \( j \) in a base period took a loan amount \( M_j(t_0) \) at the loan rate \( P_j(t_0) \) (28).

\[ P_j(t_0) = F_{P}^{-1}(Q_j(t_0), M_j(t_0)) \]  

- Point \( A_1 \). The client took an additional loan amount to the value of the loan limit \( M_{lim}^{j}(t_0) \) with the same level of creditworthiness (\( Q_j(t_1) = Q_j(t_0) \)). Thus, the rate for this additional loan is equal to \( P_j(t_1) \) (29).

\[ P_j(t_1) = F_{P}^{-1}(Q_j(t_1), M_{lim}^{j}(t_0)) \]  

- Point \( A_2 \). In the period \( t_2 \) the client’s house was burnt down. The client’s creditworthiness estimation plummeted to the value \( Q_j(t_2) \). Thus, the loan rate for new loan amount increased by the value \( \triangle P_j(t_2) \) (30) to the value \( P_j(t_2) \).

\[ P_j(t_2) = P_j(t_1) + \triangle P_j(t_2) = P_j(t_1) + \sum_{i=0}^{n} c_i(Q_j(t_2)^i - Q_j(t_1)^i) \]  

(30)

\[ P_j(t_2) = F_{P}^{-1}(Q_j(t_2), M_{lim}^{j}(t_0)) \]  

(31)

- Point \( A_3 \). In the next time period the client required an additional loan amount. Since the client has already achieved a credit limit the bank refused to provide a credit. Then, the client provided a car for the bank to secure a loan. Thus, the contour line moved right and loan limit rose up by a value \( \triangle M_{lim}^{j}(t_3) \) (32) to the value \( M_{lim}^{j}(t_3) \) (33). The new loan rate was still the highest (\( P_j(t_3) = P_j(t_2) \)), because the client took all available loan amount.

\[ \triangle M_{lim}^{j}(t_3) = \sum_{i=0}^{n} c_i(Q_j(t_3)^i - Q_j(t_2)^i) \]  

(32)

\[ M_{lim}^{j}(t_3) = F_{M}^{-1}(P_j(t_3), Q_j(t_3)) \]  

(33)

- Point \( A_4 \). For some period a creditworthiness estimation of the client remained the same \( Q_j(t_4) = Q_j(t_3) \). The client has been repaying a loan and at some point achieves the initial amount of a loan (\( M_j(t_0) \)). At this point the loan rate for the rest of available loan amount (\( P_j(t_4) \) (34) is lower than the initial loan rate (\( P_j(t_0) \)), due to a higher creditworthiness estimation (\( Q_j(t_4) > Q_j(t_0) \)).

\[ P_j(t_4) = F_{P}^{-1}(Q_j(t_4), M_j(t_0)) \]  

(34)

C. Controlling credit terms in the mortgage lending market

The proposed credit scoring model can be applied in the mortgage lending market by revising clients’ creditworthiness regularly, e.g. once every year. For instance, a client applies for
a mortgage loan for a period of 20 years in a bank. The client has an initial loan rate that reflects his initial creditworthiness. In the next year his creditworthiness can increase significantly (e.g. owing to getting a good job), remain the same, or plummet (e.g. due to a divorce).

A bank should react to changes in creditworthiness as follows:

1) if creditworthiness rises, then a bank reduces a loan rate;
2) if creditworthiness remains, then a bank keeps a loan rate at the same level;
3) if creditworthiness goes down, then a bank raises a loan rate.

In case of creditworthiness decrease, it may be inefficient to raise up a loan rate instantly because this decrease may be temporary and the loan rate increase may become the reason why this client defaults. A bank can increase a loan rate slower than a change in creditworthiness require. Moreover, any changes in creditworthiness may be conducted slower in order to maintain a more stable system, i.e. random factors that cause fluctuations in creditworthiness will have a smaller impact. As a result, such flexible loan pricing can stabilise the mortgage lending sector.

IV. CONCLUSION

In this article, a credit scoring model has been proposed. This model is based on constructing a contour subspace of credit terms for every value of creditworthiness estimation. Each of contour subspaces defines combinations of credit terms that possess an equal value of a credit risk, i.e. a bank is indifferent which combination of credit terms on a contour subspace will choose a client. Thus, this option to choose credit terms according to an individual creditworthiness value can be an additional factor for a bank to attract clients with high creditworthiness especially in the mortgage lending market.

Many clients take mortgages for a long period such as 20 or even 25 years and it seems inefficient and unfair that they have to pay the same loan rate despite changes in their creditworthiness. For example, a client has obtained a mortgage loan. His creditworthiness will increase significantly in the next 3 years and remain at this level for the rest loan period. It seems unfair that the client will still have to pay a high loan rate for another 17 years because of the robustness of the mortgage lending market. However, application of the proposed model will allow for a bank to change a loan rate in correlation with a client’s creditworthiness. Thus, the proposed model is supposed to add flexibility and stability to the mortgage lending market.

Why does a bank need to reduce loan payments for a borrower after he has already taken a loan and signed all necessary documents? If a bank is going to quit business soon then such flexible pricing would be inefficient. The proposed credit scoring model is for banks that are aimed to conduct successful long lasting business with a good reputation, attracting clients with good creditworthiness and discouraging those with bad creditworthiness. For instance, if a client has a good contract with a company but profits will come in the following year, raising his creditworthiness. Hence, this client will wait for one year to be granted a mortgage loan for the next 20 years with lower loan payments or he will apply to a bank that recalculates loan payments after changes in creditworthiness. In contrast, if a client knows that his job contract expires in the following year and he is going to take a two-year break due to family reasons then he will definitely take a loan in a bank that does not recalculate loan payments based on creditworthiness. As a result, application of the proposed model creates a reputation for a bank to attract clients with good creditworthiness and discourage those with bad creditworthiness.

In summary, two important conditions to construct these contour subspaces can be highlighted as follows:

1) a creditworthiness estimation has to be made solely on client’s information, i.e. separately from credit terms;
2) a range of a credit risk function and a creditworthiness estimation have to possess more than two values.

Decision support systems based on the proposed model can be used for:

1) credit terms visualization for clients according to their individual creditworthiness value by a contour subspace of credit terms;
2) credit terms set up according to a client’s individual creditworthiness value including the options for a client to choose credit terms;
3) credit terms online management depending on the dynamics in creditworthiness.

As a direction for further research, the proposed credit risk model should be calibrated on big data sets and applied in decision support systems.

REFERENCES