Fuzzy approach based money laundering risk assessment

Māris Krastiņš

Department of Mathematics, University of Latvia, Jelgavas iela 3, Riga, Latvia
Institute of Mathematics and Computer Science, Raiņa bulvāris 29, Riga, Latvia, mk18032@lu.lv

Abstract

This paper introduces solution for numeric evaluation of money laundering risk containing several different risk factors. For this purpose we consider the options for aggregation of risk factors and obtaining consolidated risk level. The proposed models are constructed using maximum t-conorm and Łukasiewicz t-conorm. Practical example is provided for calculation of consolidated client money laundering risk score.

Keywords: Money laundering risk, aggregation operators, maximum t-conorm, Łukasiewicz t-conorm.

1 Introduction

Over the last two decades the risk assessment methods and tools have been widely developed for different purposes. The financial industry has been just one, yet important, end user of the risk assessment solutions. While such risks as the credit risk, the financial risk and the liquidity risk have always had a solid statistic basis for numerical calculations and application of the probability theory, supervisory authorities have increased their expectations and demands for introducing mathematical models for assessment of the compliance risk, including the risk of money laundering, which lacks any solid mathematical background and can rely on expert opinions only. At the same time it should be noted that the number of subjects to Anti Money Laundering (AML) laws has significantly increased over the last decade. Therefore the models for scoring of money laundering risk often shall be implemented not only by financial institutions, but also by real estate brokers, gambling companies, notaries and other obliged entities pursuant to applicable AML laws.

The leading global AML software providers have developed different technological solutions to cope with increasing legal demands for scoring and monitoring of customers and their transactions. However, the contents of these solutions are not openly disclosed to a wider audience, and so far there have been very few attempts in proposing the models for money laundering risk assessment in scientific publications. The review article [5], published in February 2018, summarizes all efforts used so far in finding the most appropriate approach for efficient handling of tasks related to prevention of money laundering and highlights importance of the detection of suspicious transactions. However, it is evident from the practical point of view that transactional patterns are just consequences from engaging into business with particular customers posing lower or higher risk of money laundering. A use case provided in [4] allows to cope with a very simple transactional pattern, but it is important to take into account the customer specifics as part of so called Know Your Customer (KYC) process as mentioned also in [5].

The main goal of the KYC process is to implement a robust solution allowing the obliged entities to assess the level of money laundering risk as part of the customer relationship establishment, often referred as an on-boarding process. The customers shall disclose different qualitative and quantitative data which can be afterwards evaluated by experts and transposed into risk factors with corresponding risk levels. Aggregation of these risk levels results in the risk scores. Different scales can be used for this purpose, but we will apply the fuzzy numbers and assign the values close to 0 for the lowest risks, and the values close to 1 for the highest risks. It shall be noted that some risk models are inverted by assigning the values close to 1 for the lowest risks, and the values close to 0 for the highest risk. Depending on the resulting money laundering risk scores obliged entities are required to apply risk mitigation actions, which include, but are not limited to regular monitoring of customer transactions, obtaining relevant documentation on customers' source of wealth and source of funds etc.

The paper considers several options for risk level aggregation of such risk factors as customer residence, occupation or business for legal entities, customer reputation, estimated volumes and values of transactions and other similar factors as selected by the obliged entity or required by applicable laws and regulations. Section 2 provides an overview of aggregation principles which are further analysed by considering application of maximum t-conorm in Section 3 and Łukasiewicz t-conorm in Section 4. Combination of these t-conorms is proposed in Section 5. A practical example of risk levels' aggregation based on expert evaluations is presented in the Section 6.

2 Conditions for aggregation of risk factors in money laundering risk assessment model

The risk is usually considered as a function of probability and impact or likelihood and severity. Such definition is suitable for many industries as outlined, for example, in [6, 7]. It also allows application of Mamdani-Type or Sugeno-Type fuzzy inference systems for obtaining the consolidated risk levels as described by [1, 11, 9]. This process is similar to different other practical applications provided in [2].

Contrary to the previous examples of risks, money laundering risk attributed to each customer is rather specific and embraces comparably vague component of severity, which can be hardly characterised by any numeric value. It particularly applies to such indicators as reputational impact or business sustainability. At the same time AML laws and regulations require implementation of detailed KYC procedures and assessment of multiple risk factors for each customer. Therefore an expert opinion is among the most appropriate solutions for assigning the risk levels for corresponding notional risk factors. A simple example can be used to explain the need for a human decision in assessing the risk level. Let us consider two companies of different size and their estimated average transaction values. While EUR 100 000 payment would be treated as low risk indicator with value close to 0 for the large company, it would be definitely a high risk indicator with value close to 1 for the small company. Similar judgement is valid also for assessment of expected payment volumes which can be considered as another different risk factor.

Let us consider that all risk levels of corresponding risk factors $X_k, k \in \{1, ..., i\}$ and $k \in \mathbb{N}$ are expressed in the form of fuzzy set $\mu = (x_1, ..., x_i), i \in \{1, ..., n\}$ and $n \in \mathbb{N}$. In order to aggregate these risk levels we will use aggregation operator $\mathbf{A} : \bigcup_{n \in \mathbb{N}} [0, 1]^n \to [0, 1]$. An example of ten risk factors ($i = 10$) with corresponding risk levels for two sample customers is provided in

Figure 1. It is evident that customers can be different and with different risk levels for corresponding risk factors, which are notional and do not correspond to any particular values on x axis.
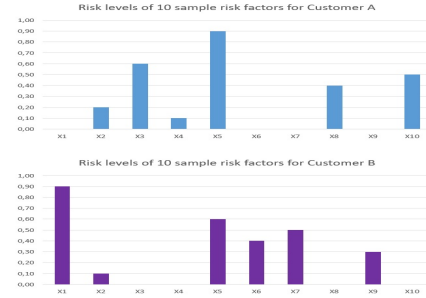
The use of risk level fuzzy sets, especially their

Figure 1: Example of risk levels for two sample customers

graphical representation, provides a good preliminary overview of overall customer risk level. However, they do not encompass importance of each risk factor, and also do not provide a clear answer, if the obliged entity is facing high or low risk customer. Therefore we will explore the options for aggregation of risk levels using different aggregation operators in order to obtain the customer risk score. First of all we consider application of different average operators analysed by [3]. We define the fuzzy weighted average of risk levels $(x_1, ..., x_i)$ as follows:

$$W(x_1, ..., x_n) = \sum_{i=1}^{n} \omega_i x_i$$

where ω_i are weights for each corresponding x_i and $\sum_{i=1}^{n} \omega_i = 1$. At the first glimpse such approach could be considered as suitable since the risk score is the mean value of all weighted risk levels. However, due to specifics of money laundering risk there are many occasions when it is not reasonable to accept that the risk score is lower than value of the highest risk level in the fuzzy set $\mu = (x_1, ..., x_n)$. Therefore alternative aggregation operators should be considered.

3 Aggregation of risk levels with maximum t-conorm

As part of the risk level aggregation it can be assumed that for particular cases the total risk level of any sample customer cannot be lower than the highest (maximum) risk level of all risk factors. Therefore we can apply the maximum t-conorm $M(x_1, ..., x_n) = \max(x_1, ..., x_i)$, $i \in \{1, ..., n\}$ and $n \in \mathbb{N}$. It should be noted that all risk factors X_k may not be equally important for customer risk scoring. Therefore we apply

fuzzy coefficients $a_i \in [0,1]$, $i \in \{1,...,n\}$ and $n \in \mathbb{N}$ allowing to keep the values of certain initial risk levels or decrease them in similar way as proposed by [10]. In our model $\sum_{i=1}^{n} a_i \neq 1$, and fuzzy coefficients can be regarded as the indicators of risk appetite resulting in the fact that particular risk levels are decreased, if obliged entity considers them as less important. Consequently the maximum t-conorm can be expressed in the following format:

$$M(x_1,...,x_n) = \max(a_1 x_1,...,a_i x_i), \qquad (1)$$

$i \in \{1,...,n\}$ and $n \in \mathbb{N}$.

When applying such aggregation, the most important risk factors with corresponding risk levels are considered. However, other risk factors of lower importance with non-zero risk levels should not be disregarded as their aggregated impact could be more severe than the highest risk level of the most important risk factor. This implies that additional options for aggregation of the risk levels of non-critical risk factors are required.

4 Aggregation of risk levels with Łukasiewicz t-conorm

Aggregation using arithmetic sum often results in values exceeding 1. If we consider the example provided in Figure 1, it is evident that the sum of only two particular risk levels for both sample customers is greater than 1 while there are non-zero risk level values for four more risk factors. In order to overcome this problem, we apply Łukasiewicz t-conorm $L(x_1,...,x_n) = \min(1, \sum_{i=1}^{n} x_i)$, $n \in \mathbb{N}$. As in the case of maximum t-conorm we note that the risk factors are not equally important. Therefore we apply the same fuzzy coefficients $a_i \in [0,1]$, $i \in \{1,...,n\}$ and $n \in \mathbb{N}$ for calibration of the risk level values. The adjusted Łukasiewicz t-conorm is expressed as follows:

$$L(x_1,...,x_n) = \min(1, \sum_{i=1}^{n} a_i x_i), \qquad (2)$$

$n \in \mathbb{N}$. The formula (2) underlines the meaning of coefficients a_i as values expressing the risk appetite. It is evident that lower a_i means higher risk appetite, and more risk factors can be included in the total risk score unless it does not exceed the threshold of non-acceptable risk as internally set by the obliged entity. While Łukasiewicz t-conorm and formula (2) provide a sound basis for suitable aggregation of money laundering risk levels into single value risk score, such aggregation may decrease the importance of some critical risk factors preserved in case of the maximum t-conorm. There is another scenario when the obliged entity may

select some critical risk factors and benchmark their risk levels against aggregated risk level values of the remaining risk factors. Therefore construction of the combined aggregation operator is proposed in the Section 5.

5 Aggregation of risk levels with combined t-conorm

Let us assume that we have to deal with some critically important risk factors and less significant risk factors by splitting them in two groups and combining results using the following t-conorm: $C(x_1,...,x_n) = \max(x_1,...,x_k, \min(1, \sum_{j=k+1}^{n} x_j))$, $n \in \mathbb{N}$. Such t-conorm is commutative, monotone, and the number 1 acts as its identity element. However, this t-conorm is not associative as it basically consists of two independent parts. From the practical point of view, it is not critical that associativity does not hold. The t-conorm C is not unique in that sense, as there are similar examples of functions, like copulas and quasi-copulas analysed in [8], where associativity does not hold.

When applying the combined t-conorm C we should consider the same principle of the importance of risk factors applied to maximum t-conorm and Łukasiewicz t-conorm. The resulting formula will be the following:

$$C(x_1,...,x_n) = \max(a_1 x_1,...,a_k x_k, \min(1, \sum_{j=k+1}^{n} a_j x_j)),$$
$$(3)$$

$n \in \mathbb{N}$. In practice (3) will allow to select the critical risk factors out of the sum of remaining risk factors. Such calculation is two folded. First of all it secures that important high level risk factors are not missed, and secondly it allows to benchmark aggregated level of less important risk factors against critical risk factors.

6 Example of aggregated risk score

The main purpose of the proposed aggregation model is to enable the end users, including obliged entities and supervisory authorities, to implement simple and clear solutions for obtaining the customer's money laundering risk score. In our example we will use four risk levels and apply them to each risk factor pursuant to the widely accepted international standards: low risk, medium risk, high risk, non-acceptable risk. Table 1 provides corresponding intervals for each risk level. We will use a sample customer A from Figure 1 with corresponding fuzzy set of 10 risk levels $\mu = (x_1,...,x_{10}) = (0; 0,2; 0,6; 0,1; 0,9; 0; 0; 0,4; 0; 0,5)$. As the next step

RISK LEVEL	INTERVAL OF RISK SCORE
Low	$[0, 0{,}30)$
Medium	$[0{,}30, 0{,}60)$
High	$[0{,}60, 0{,}95]$
Non-acceptable	$(0{,}95, 1]$

Table 1: Values of risk levels.

we will apply the following fuzzy coefficients characterising importance of corresponding risk factors: $a_1 = 1, a_2 = 1, a_3 = 1, a_4 = 0{,}8, a_5 = 0{,}7, a_6 = 0{,}5, a_7 = 0{,}5, a_8 = 0{,}4, a_9 = 0{,}3, a_{10} = 0{,}2$. In such case (1) and (2) result in the following values: $M = 0{,}63$, $L = \min(1; 1{,}77) = 1$. It means that the maximum t-conorm returns the value slightly above the lowest value of the high risk, while application of Łukasiewicz t-conorm results in non-acceptable risk. Therefore intuitively these values could indicate that the most suitable level of aggregated risk score could be high, but certainly lower than non acceptable. In order to apply (3) we will select x_1, x_2, x_3, x_4 as the most critical risk factors. Such selection results in the following value of the combined t-conorm: $C = \max(0{,}6; \min(1; 0{,}89)) = 0{,}89$. The result corresponding to high level risk (with risk score 0,89) is quite suitable for the obliged entity with rather low risk appetite. At the same time it should be admitted that the risk appetite can be increased by modifying fuzzy coefficients a_i, and a_5 in particular.

7 Conclusions

The proposed methods for aggregation of risk levels enable efficient calculation of risk scores. While simple t-conorms result in loss or overestimation of the importance of corresponding risk factors, combination of t-conorms provide a good basis for a simple risk scoring. Transparency of the combined t-conorm can allow supervisory authorities to identify the risk appetite of any obliged entity which has chosen to apply such aggregation model for obtaining money laundering risk scores. Further research will focus on the fine-tuning of the aggregation model and also introducing Mamdani-Type or Sugeno-Type fuzzy inference systems towards each risk factor of the money laundering risk.

Acknowledgement

References

[1] O. Abul-Haggag, W. Barakat, Application of fuzzy logic for risk assessment using risk matrix, International Journal of Emerging Technology and Advanced Engineering 3 (2013) 49–54.

[2] G. Bojadziev, M. Bojadziev, Fuzzy sets, fuzzy logic, applications, Vol. 5 of Advances in Fuzzy Systems - Applications and Theory, World Scientific Publishing Co. Pte. Ltd., Singapore, 1998.

[3] T. Calvo, A. Kolesárová, M. Komorníková, R. Mesiar, Aggregation operators: Properties, classes and construction methods, in: T. Calvo, G. Mayor, R. Mesiar (Eds.), Aggregation Operators, Vol. 97 of Studies in Fuzziness and Soft Computing, Heidelberg, 2002, pp. 3–104.

[4] Y.-T. Chen, J. Mathe, Fuzzy computing applications for anti-money laundering and distributed storage system load monitoring., in: World conference on soft computing, San Francisco, USA, 2011.

[5] Z. Chen, L. Khoa, E. Teoh, A. Nazir, E. Karuppiah, K. Lam, Machine learning techniques for anti-money laundering (AML) solutions in suspicious transaction detection: a review, Knowledge and Information Systems 57 (2018) 245–285.

[6] R. Choudhary, A. Raghuvanshi, Risk assessment of a system security on fuzzy logic, International Journal of Scientific & Engineering Research 3 (2012) 1–4.

[7] H. Hussin, S. Kaka, M. Majid, A case study on fuzzy logic-based risk assessment in oil and gas industry, ARPN Journal of Engineering and Applied Sciences 11 (5) (2016) 3049–3054.

[8] P. Hájek, R. Mesiar, On copulas, quasicopulas and fuzzy logic, Soft Computing 12 (2008) 1239–1243.

[9] A. Kaur, A. Kaur, Comparison of mamdani-type and sugeno-type fuzzy inference systems for air conditioning system, International Journal of Soft Computing and Engineering (IJSCE) 2 (2) (2012) 323–325.

[10] V. Peneva, I. Popchev, Aggregation of fuzzy relations with fuzzy weighted coefficients, Advanced Studies in Contemporary Mathematics 15 (2007) 121–132.

[11] U. Umoh, A. Udosen, Sugeno-type fuzzy inference model for stock price prediction, International Journal of Computer Applications 103 (3) (2014) 1–12.