A Credit Risk Measuring Model Using Logit Model and Choquet Integral

Ken Urata*, Shinichi Yoshida**, Hajime Nobuhara*, Kazuhiko Kawamoto*,
Yutaka Hatakeyama*, Toshihiro Kaino***, Kaoru Hirota*
*) Interdisciplinary Graduate School of Science and Engineering, Tokyo Institute of Technology
4259 Nagatsuta, Yokohama 226-8502, Janan
Email: {kenurata, nobuhara, kawa, hatake, hirota}@hrt.dis.titech.ac.jp

**) Information Science Research Center, Aoyama Gakuin University
5-10-1 Fuchinobe, Sagamihara, Kanagawa 229-8558, Janan
Email: yoshida@irc.aoyama.ac.jp

***) College of Economics, Aoyama Gakuin University
4-4-25 Shibuya, Shibuya, Tokyo 150-8366, Janan
Email: kaino@mue.biglobe.ne.jp

***Abstract*— A credit risk measuring model using logit model and Choquet integral on λ-fuzzy measure space is proposed to express non-linearity between explanatory variable and credit risk score, and also to utilize qualitative variables by that non-linearity. The model selection and parameter estimation are conducted by the logit model using stepwise method evaluated by AIC. The non-linearity is expressed by Choquet integral considering thresholds whose parameters are set by minimizing rank errors ordered by Choquet integral values. The model is evaluated by financial and qualitative database of listed companies and improves average 5% of Accuracy Ratio compared with that of the logit model. Proposed model provides a new method for credit risk management considering immaterial assets.**

I. Introduction

Recently, the number of bankruptcies of listed companies is increasing. Precious measuring of credit risk is required for every financial institution. Credit risk measuring researches are categorized into two, one is estimation of default probability, the other is evaluation of credit risk based on the probability. For estimation of default probability, there is a method that derives directory the probability [5], but the probability distribution is different from the actual distribution of default probability. Therefore usual methods identify credit rating based on statistical methods and option theory, and derive the default probability using rating transit probability matrix derived from the credit rating non-parametrically[2][3][12]. But conventional credit risk measuring models are not able to describe non-linearity between the credit risk and explaining variables sufficiently and also do not consider qualitative indices. In this paper, a credit risk measuring model that is based on the Logit model and the Choquet integral on fuzzy measure space. This model increases accuracy of credit risk measuring and is able to describe non-linearity between explaining variable and adopt qualitative indices. It is difficult to do model selection and parameter setting. At first, quantitative indices are selected using AIC and logit with step-wise method.

II. Problem of linearity of logit model and parameter setting of Choquet integral

A. Problem of linearity in logit model

Research of predicting company's default has been started by Altman's pioneering work[1]. Altman's work calculates Z-score using discriminant analysis. Z-score has been widely used as a predicting measure of company's default, but with the recent credit exposure management scene requires not only the alternative decision of default, but also the probability or some kind of numerical data concerning the default. And ROI (Return of Investment) is predicted using the numerical data

Logit Model[5] estimates the default probability p_i derived from the logistic distribution function,

$$p_i = \frac{1}{1 + exp(-Z_i)},$$

where Z_i is derived from linear combination of financial data x_m and parameter β_m,

$$Z_i = b_0 + b_1 x_{i1} + b_2 x_{i2} + \cdots + b_m x_{im}.$$

Logit model permits to estimate the default probability. However logit model still uses Z-score, which is the linear combination of financial indices. If some financial index indicates a strong tendency to default, the model has a possibility to indicate the default although the other indices indicate quite normal. Table

2.1 shows the most of financial indices of a default company indicate the default tendency. In order to solve this problem, a model which has ability to describe non-linearity between the default probability and the financial indices is required.

In this paper, we propose a model using Choquet integral model. Choquet integral model is used for some non-linear evaluation model, for example a human subjective evaluation.

Choquet integral model derives a score to identify the credit rating of a company, and the default probability of the company is non-parametrically derived from the credit rating transition probability matrix.

B. Parameter setting problem of Choquet integral model

Concepts of fuzzy measure and fuzzy integral have been defined by Sugeno [8]. A functional defined by Choquet is also regarded as a fuzzy integral [7] and is attempted to apply for multi-attribute evaluation problem and predicting problem [9].

Here shows the definition of fuzzy integral as follows,

Definition 2.1 [8] Let μ be a function 2^S to $[0, \inf)$, μ is called a fuzzy measure on S if

$$(f1)\mu(\phi) = 0$$

$$(f2)E \subset F \rightarrow \mu(E) \leq \mu(F)$$

Definition 2.2 [7] Let $f : S \rightarrow R$, $s_1, s_2, \cdots, s_n \in S$, and $f(s_1) \leq f(s_2) \leq \cdots f(s_n)$, the Choquet integral of f under fuzzy measure μ is defined as

$$(C)\int_S f(s)d\mu(s) = \sum_{j=1}^{n}[f(s_j) - f(s_{j-1})] \cdot \mu(F_j),$$

where $F_j = \{s_j, s_{j+1}, \cdots, s_n\}, j = 1, 2, \cdots, n, f(s_0) = 0$.

Credit risk measuring using Choquet integral has been firstly proposed by Kaino and Hirota. The model is used for simulating long-term bond rating of Moody's [6]. Choquet integral evaluation model has strong power of expression and flexibility, but number of parameters increases exponentially with number of explanatory variables. Because of this, parameter setting of Choquet integral model is difficult under many explanatory variables. Parameter setting of Choquet integral has been studied by Murofushi [10][11]. Considered to apply for bond rating problem, these conventional parameter setting methods are difficult to use because model learning data required for these methods must be numerical data, but bond rating data are " default or not ", i.e., Boolean data. Moreover proper default probability cannot be obtained.

In this paper, we propose credit risk measuring using Choquet integral model and its parameter setting. First variables that are profitable to describe the model are selected using logit model from financial indices. These selected parameters are used as initial parameters of the Choquet integral evaluation model. Companies are placed in the order by the value of Choquet integral. Mean square error between the order of non-default companies and the real critical order which divides default companies and non-default companies is calculated. Parameters of the model are updated by the steepest descent method using the mean square as the objective function.

III. Credit risk measuring model using Logit model and Choquet integral

A. Selection of quantitative indices using Logit model

1) Conversion to percentile : Using the numerical data directly has some problem because outlier may affect the estimation process. Moreover negative value cannot be used in Choquet integral.

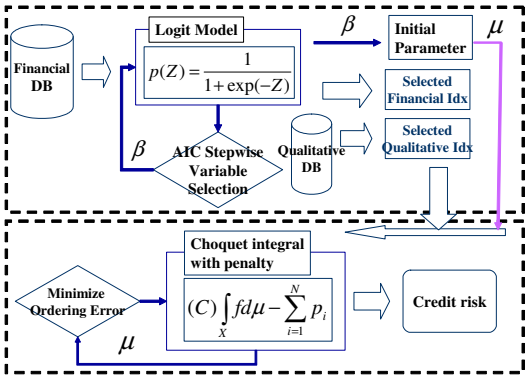

Fig. 1. Design of proposed model

Because the domain of each financial index and actual meanings of each value of them are not uniform, the original data are converted into percentile ranking uniformly as the least value to 1, the second to 2, ..., the largest to 100. The larger percentile value means good soundness. If the input data has a missing value, it is converted into 50.5 on the percentile.

2) AIC and Stepwise method : Because it is difficult to compare the likelihoods of the Choquet integral models, logit model is used first to select the significant model. Nikkei financial database is used for the experiment. There are 31 variables in the database. Significant variables are required to select to predict default or not. In this paper, we use the stepwise method because the method is able to both increase and decrease the number of variables. The criterion of increase or decrease variables is AIC (Akaike Information Criterion).

AIC = -2 · (Max Log Likelihood) + 2 · (# of Parameters).

Step 1

N variables are selected, and M variables are the candidate to select. First, AIC of N selected variables are calculated.

Step 2

One of M candidate variables is added to N selected variables and calculate AIC. It is performed for all M candidate variables. The next selected variable is the one that the AIC is the minimum.

Step 3

One of N + 1 selected variables in Step 2 is removed and calculate AIC. It is performed for all N + 1 variables. The next removing variable is the one that the AIC is the minimum.

Step 4

If the variables selected in Step 2 and Step3 are the same variable, N + 1 variables including it ppare the next selected variables and M -1 variables are the next candidate variables.

TABLE I
1999 NIKKEI FINANCIAL DATA

	deposit-loan ratio (%)	acid ratio (%)	return on equality (%)	debt ratio (%)	degree of indebtedness (%))
non-default company	811.12	124.71	40.04	474.75	30.80
default company	30.82	52.82	12.49	2707.44	59.50

And go to Step 1. Else the N variables selected in Step 3 are the next selected variables, and go Step 1.

This iteration is performed and the combination of the variables that decrease the AIC in each dimension can be obtained.

3) Post processing of Choquet integral : Linear combination of variables by stepwise method can take both positive and negative value. Because this can not be used for parameter of Choquet integral model, it must be converted.

First the parameters which take negative values are converted to its absolute value. As the parameters which take positive values are already converted to percentile values, and a greater parameter means better evaluation of a company. So the converted parameter from negative to positive by absolute value is subtracted from 100.

IV. PARAMETER SETTING OF CHOQUET INTEGRAL MODEL.

A. Non-linearity expression of Choquet integral model using threshold

In order to reduce the number of parameters of Choquet integral model, we use λ-fuzzy measure. The definition of λ-fuzzy measure is as follows[13].

Definition 3.1 Let $(S, 2^S, \mu_\lambda)$ be fuzzy space,

$$A, B \in 2^S, A \cap B = \phi,$$
$$\Rightarrow \mu_\lambda(A \cup B) = \mu_\lambda(A) + \mu_\lambda(B) + \lambda\mu_\lambda(A)\mu_\lambda(B),$$

where $\lambda \in [-1, \inf)$.

The non-linearity of the credit risk model is that all financial indices indicating default tendency is worse than only 1 or 2 indices indicating extremely default tendency.

This property under λ-fuzzy measure is shown as the Figure 2.

1) Expression of qualitative indices using non-linearity of Choquet integral model: For credit risk measuring, qualitative indices of companies such as management ability, reputation, brand, development of a industry or trade, etc, are also quite important factors. For example when a bond rating agency grades a company, information from interviewing to managers of the company is also considered.

Qualitative indices are, however, difficult to convert appropriate numerical data, and conventional models are not able to include these information. For example a result from a questionaire shows some numerical or discrete form of information. But their answers, e.g. “1” or “A”, often do not mean quantitative semantics. That means that difference of “1” and “2”, and another difference of “4” and “5” are sometimes different. Our proposed model can include these qualitative data using non-linearity of Choquet integral.

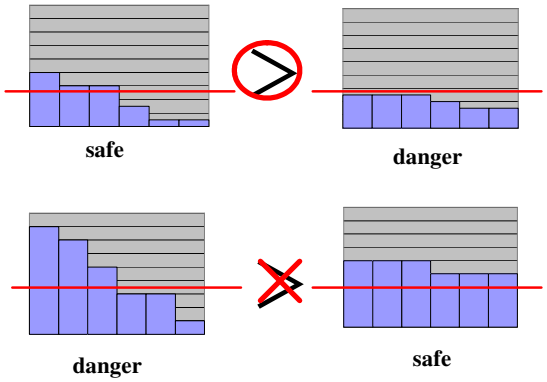

Fig. 2. Property of λ-fuzzy measure

2) Parameter setting using ordering error: It is difficult to derive the turning point between default and non-default parametrically using the value of Choquet integral model directory. In our model, companies are ranked and put from the best company to the worst company using the Choquet integral value and compares the ranking order of the training data. The rank of the first non-default company of these two rank are compared and th difference of them are to minimized using the steepest descent method.

V. EXPERIMENT

A. Financial data from Nikkei database

The financial model using our experiment are from Nikkei (Japan Economic News) database. This database contains financial data of 3160 companies and the indices are about 600[15][16]. The data from April 1998 to March 1999 are used as the training data, and the April 1999 to March 2000 as the simulation test data. The definition of defaulf a company is that the company defaults in three years from the settlement of account.

B. Preprocess of training data and test data

The summary of data is as follows.

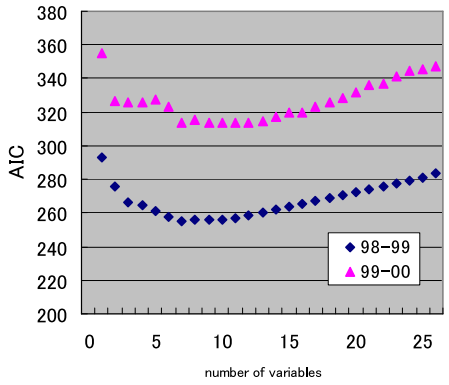

Fig. 3. AIC comparison

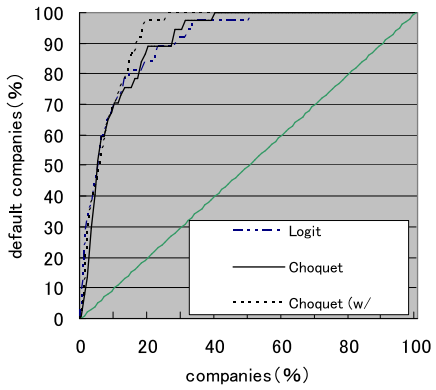

Fig. 4. CAP curve comparison (training data)

category	training (default)	test (default)
ceramic	2 (1)	2 (0)
electronics	236 (3)	248 (5)
chemical	162 (2)	164 (1)
steel	51 (2)	52 (2)
retail	116 (7)	130 (9)
estate	37 (1)	43 (1)
service	181 (0)	215 (1)
food	100 (1)	105 (2)
trading	253 (2)	277 (2)
mechanical	175 (5)	187 (6)
construction	140 (7)	157 (10)
fisheries	7 (1)	7 (1)
warehouse,traffic	26 (0)	28 (1)
pulp, paper	1 (1)	1 (1)
metal	3 (1)	3 (1)
textile	0 (0)	2 (2)
manufacturing	77 (1)	79 (1)
finance	26 (1)	27 (0)
transport	18 (1)	18 (1)
Total	1611 (37)	1745 (47)

C. Result of variable selection using logit model

Figure 3 shows the transition of the optimized value of AIC of both training data and test data. When the number of variables is seven, AIC of training data is the minimum value 255.3, AIC of the test data is 313.9, while the minimum value of the test data is 313.7 in the dimension 12. It means that our process is not overfitted.

D. Comparison with logit model

We use CAP (Cumulative Accuracy Profile) curve and AR (Accuracy Ratio) to compare our model and logit model.

Figure 5 shows the CAP curve of test data with regard to logit model, Choquet integral model using only quantitative indices, and Choquet integral model with both of quantitative and qualitative indices.

If the model predicting of the default and the real default have no relation each other, then the CAP curve take the 45 degree line. And the ideal model (100 % prediction model) takes constant value 1.

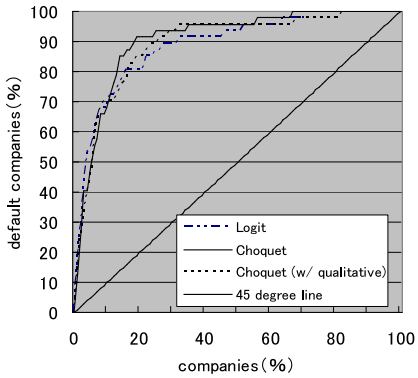

Fig. 5. CAP curve comparison (test data)

AR is the ratio of the area enclosed by the ideal CAP curve and 45 degree line and the area enclosed by the CAP curve of the model and 45 degree line.

Figure 4 shows that the proposed model using both quantitative and qualitative data is improved and the CAP curve of proposed model using only quantitative data is almost same as that of logit model for training data. And for the test data, proposed model with only quantitative data is the best (Figure 5).

VI. Conclusion

In this paper, we propose a credit risk measuring model using Logit model and Choquet Integral in order to improve the predicting precision and accountability. This model has ability to express non-linearity among explaing variables, and qualitative indices can be expressed.

Experiments are performed using the financial data from 1998 to 1999, 1999 to 2000 (Japan Economic News) among conventional Logit model, our proposed model (using only

TABLE II

ACCURACY RATIO COMPARISON

Model	Data	Parameter	Parameter Setting	AR (training)	AR (test)
Logit	Preprocessed	Quantitative	N/A	81.43	77.77
Logit	Original	Quantitative	N/A	79.97	75.68
Choquet	Preprocessed	Quantitative	Not tuned	80.51	78.36
Choquet	Preprocessed	Quantitative	Tuned	81.70	81.34
Choquet	Preprocessed	Quantitative and Qualitative	Tuned	85.43	78.62

quantitative data), and our model with qualitative data. The result shows that our model improves 5% of AR.

REFERENCES

[1] E.I.Altman, " Financial Ratios, Discriminant Analysis and The Prediction of Corporate Bankruptcy ", Journal of Finance, vol.23, pp.189-209, 1968.

[2] O.A. Vasicek, " Credit Valuation ", Working Paper, KMV Corporation, 1984.

[3] S.Kealhofer,S.Kwok and W.Wen, " Uses and Abuses of Bond Default Rates, " Working Paper, KMV Corporation, 1998.

[4] K.Ishii and M.Sugeno, " A model of human evaluation process using fuzzy measure ", International Journal of Man-Machine Studies, vol.22, pp.19-38, 1985.

[5] D.Martin, " Early Warning of Bank Failure : A Logit Regression Approach ", Journal of Banking and Finance, vol. 1. no.3, pp.249-276, 1977.

[6] T.Kaino and K.Hirota, " Defferentiation of the Choquet Integral and Its Application to Long-term Debt Ratings ", Journal of Advanced Computational Intelligence, vol.4, no.1, pp.66-75, 2000.

[7] G.Choquet, " Theory of capacities ", Ann.Inst.Fourier, vol 5, pp.131-295, 1953.

[8] Michio Sugeno, "fuzzy measure and fuzzy integral, Transactions of The Society of Instrument and Control Engineers, vol.8, pp.218-226, 1972.

[9] M.Higuchi, " Fuzzy Integral application to estimate of grindstone's quality-study of grinding sound(Report 4.) ", Journal of the Japan Society for Precision Engineering, vol.54, no.4, pp.761-766, 1988.

[10] Michio Sugeno, "Construction of Fuzzy Measure and Pattern Similarity Evaluation using Fuzzy Integral", Journal of the Japan Society for Precision Engineering vol.9, pp.361-368, 1973.

[11] Standard & Poor's, " Debt Ratings Criteria ", Municipal Overview, 1986.

[12] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1999/2/22 Morning Paper, pp.28-29, 1999

[13] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 2000/2/21 Morning Paper, pp.26-27, 2000

[14] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1998/3/30 Morning Paper, pp.23-25, 1998

[15] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1998/6/29 Morning Paper, pp.23-25, 1998

[16] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1998/9/28 Morning Paper, pp.23-25, 1998

[17] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1999/1/4 Morning Paper, pp.21-23, 1999

[18] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1999/3/29 Morning Paper, pp.25-27, 1999

[19] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1999/6/28 Morning Paper, pp.23-25, 1999

[20] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 1999/9/27 Morning Paper, pp.23-25, 1999

[21] Nihon Keizai Simbun Inc. (Japan Economic News Inc.), "PRISM: good standing companies BEST 300", 2000/1/4 Morning Paper, pp.11-13, 2000