Credit Scoring via PCALWM*

Jianping Li ^{1,Ψ}, Weixuan Xu¹, and Yong Shi²

¹ Institute of Policy and Management,
Chinese Academy of Sciences, Beijing 100080, P.R.China
ljp@mail.casipm.ac.cn, jianpingli@yahoo.com
² Chinese Academy of Sciences Research Center on Data Technology
and Knowledge Economy, Beijing 100039, P.R.China
yshi@gscas.edu.cn

Abstract. We have presented a principal component analysis linear-weighted model (PCALWM) for credit scoring in [5,6], this article is a further study on this model. We revised the application procedure in the credit scoring, and tested it by a larger real-life credit card dataset. In comparison with some well-known scores, the empirical results of the PCALWM can achieve a favorable KS distance. The study on some application features of this model in the credit decision-making shows that the model can help the credit issuers to select the best trade-off among the enterprise stratagem, marketing and credit risk management.

Keywords: Credit scoring, Classification, Principal component analysis.

1 Introduction

The growth in consumer credit outstanding over the last fifty years is truly spectacular in the major countries, for example, in US, from the \$9.8 billions in the year 1946 to \$20,386 in 2003[2]. Not all of this growth is due to the borrowing on credit lines, Credit card (and debit cards) has become increasingly important as a method of money transmission.

However, from the perspective of history, the increase of total outstanding of consumer credit is always accompanying with more much consumers' bankruptcy and default, which are causing plenty of headaches for banks and other credit issuers. From 1980 to 2003, the number of individual bankruptcy filings in the US had increased by 560%[11]. How to predict bankruptcy and avoid huge charge-off loss becomes a critical issue of credit issuers.

In the consumer credit assessment, credit scoring is widely used. Credit scoring provides a suite of decision-support tools to aid the credit analyst. In credit scoring

* This research has been partially supported by a grant from National Natural Science Foundation of China (#70472074) and the President Fund of Chinese Academy of Sciences (CAS) (yzjj946).

^Ψ Corresponding author.

system, the classification tools, which are derived from analyzing similar data for previous customers, are used to categorize new applicants or existing customers as good or bad. There are lots of classification tools available now, such as statistics methods like discriminant analysis, regression, classification tree, nearest-neighbor approach, nonstatistical method like linear programming, neural networks, genetic algorithms, expert systems, etc [12,13].

The purpose of this paper is to further analyze the principle components analysis linear-weighted model (PCALWM) which we had presented to assess the consumer credit [5,6]. We will apply this model to a larger real-life credit card database to evaluate its performance and study the characteristics in credit decision. This study will make the model more applicable and help the credit issuers to make right decision.

2 The Basic Principals and Application Stages of PCALWM in Credit Scoring

Principle components analysis (PCA) is a well-known, widely-used statistical technique, and it is a purely data-driven way of exploring the data. The central idea is to reduce the dimensionality of a data set consisting of a large number of interrelated variables, while retaining as much as possible of the variation present in the data set. The PCA approach uses all of the original variable to obtain a smaller set of new variables (principal components--PCs) that they can used to approximate the original variables. The greater the degree of the correlation between the original variables, the fewer the number of new variables required. PCs are uncorrelated and ordered so that the first few retain most of the variation present in the original set.

PCA approach was applied to the credit risk management cooperated with other methods, Alici[1] used this method to predict the corporation bankruptcy with neural networks. Recently PCA has attracted much attention in the risk analysis in larger scale data and consumer credit data [3,8,14].

The PCALWM based credit scoring system includes 4 stages, data preparation, training, testing, and predicting. The main stages show by figure 1.

2.1 Data Preparation

This stage consist of collecting relevant data, merging and cleaning the data, defining the classes, selecting the variables, etc. we will get a dataset with suitable format which is fit for the PCA computing after this stage.

2.2 Training Stage

We classify the data set into two parts, one for training to construct the predicting model and the other for testing to validate the model. The different rate of the training data in the total sample will impact the model's efficiency, and a good classification of the training data rate maybe get a good classification model.

Fig. 1. The application procedure of PCALWM in credit scoring

The training stage is to achieve predicting model and set up the different credit grade. Supposed we have selected principal factors, the scoring function of the principal factors in the training sample could be expressed as:

$$F = a_1 z_1 + a_2 z_2 + \dots + a_s z_s \quad (2.1)$$

z_i is the i^{th} principal factor, a_i is its contribution rate :

$$a_i = \frac{\lambda_i}{\sum_{m=1}^k \lambda_m} \quad a_{il} = \frac{\sum_{m=1}^l \lambda_m}{\sum_{m=1}^k \lambda_m} \quad (i, l = 1, 2, \dots, k)$$

(λ_i is the i^{th} eigenvalue sorted by descending, a_{il} is the cumulative variance contributions of principal factor z_1 to z_l). Then,

$$Z_i = b_{i1}x_1 + b_{i2}x_2 + \dots + b_{in}X_n \quad (2.2)$$

Take the equation (2) into equation (1), we get:

$$F = c_1 x_1 + c_2 x_2 + \dots + c_n X_n \quad (2.3)$$

$$c_j = \sum_{i=1}^s a_i b_{ij} \quad (i = 1, 2, \dots, s; j = 1, 2, \dots, n)$$

According to the score and the distribution of the bad credit in the training sample, we can set up the different grade by a certain method and standard. Then one can classify the different class based on the score. If the classification result is not ideal, we will classify and compute the data again.

2.3 Testing Stage

This stage is to validate the model and the credit grade which get from the training stage. We will choose the right score model and credit grade in this stage.

We use the same procedures in the training stage. But in data standardization here we will take the average and standard deviation of the training sample, so we get the data matrix $X_i = (x_{i1}^*, x_{i2}^*, \dots, x_{ip}^*)^T \quad i = 1, 2, \dots, m$.

We use the equation 2.3 to compute the credit score in the testing sample, and classify the credit grade with the set grade in training stage. Similarly, if the classification result is not satisfied, we would classify the originate data and compute again.

2.4 Predicting the New Applicants

In this stage, we use the score model to predict the new applicants credit score and make the credit decision. After computing the new applicants credit score with the certain procedure as in the testing stage, we compare it with the credit standard which get from the training and testing stage. If the score reaches the threshold, we grant the corresponding credit, otherwise reject.

3 Empirical Study

3.1 Data Set

We use real life credit card data from a major US bank to conduct the experiment. This data set is composed by 5000 records, and two classes have been defined: good credit and bad credit. 65 variables are selected for personal credit assessment. There are including history payment, balance, transaction, account open, etc. we chose the SPSS soft to perform all of the computations.

3.2 Computation Result and Analysis

In fact, the practitioners have tried a number of quantitative techniques to conduct the credit card portfolio management. Some examples of known scores are (1) Behavior Score developed by Fair Isaac Corporation (FICO) [15]; (2) Credit Bureau Score also developed by FICO [15]; (3) First Data Resource (FDR)'s Proprietary Bankruptcy Score [4]; (4) Multiple-criteria Score [9,10] and (5) Dual-model Score [7]. These methods can be generally regarded as two-group classification models, and use either statistical methods or neural networks to compute the Kolmogorov-Smirnov (KS) value that often used to measure how far apart the distribution functions of the scores of the goods and bads are[13]. The resulting KS values from the learning process are applied to the real-life credit data warehouse to predict the percentage of bankrupt accounts in the future.

Formally, if $n_G(s)$ and $n_B(s)$ are the numbers of goods and bads with scores in a sample of n , where there are n_G goods and n_B bads, the $p_G(s) = \frac{n_G(s)}{n_G}$ and

$p_B(s) = \frac{n_B(s)}{n_B}$ are the probabilities of a good and bad having a score s , then the KS distance is

$$KS = \max_s |p_G(s) - p_B(s)| \quad (3.1)$$

We chose 2000 records as a training set as random to compute the score model, and set three testing samples with 1000, 2000, 3000 records in the other 3000 records as random. The KS distances show as the table 1, the biggest is 0.6222 and lowest is 0.524.

Table 1. KS values of the different samples

	Traning2000	Testing1000	Testing2000	Testing3000
KS value	0.6223	0.5993	0.563	0.524

The KS values of some well-known score models in our 1000 sample are as the table 2.

Table 2. KS values of the different scores

	FICO	Credit bureau score	FDC proprietary bankruptcy score	Decision trees score	Multiple- criteria Linear Program score	PCALW M score
KS value	0.553	0.456	0.592	0.602	0.595	0.599

Among these scores, the PCALWM score is better than any other scores, except the decision tree score, so the PCALWM can achieve a good KS distance, and have a good classification.

4 Two Characteristics in Credit Decision-Making of PCALWM

In the practice, credit policy is the trade-off of the enterprise’s stratagem, marketing and risk management. Loose credit policy is helpful to the market expanding, while strict credit policy for the risk management. General speaking, the loose credit policy will increase the type I error, and decrease the type II error to the credit issuers, while the strict credit policy makes the opposite effect. So, if the model can give the movement of the two type error with different credit granting rate efficiently, the model will greatly help to the credit issuers to make the trade-off in the actual credit decision.

For the encouragement of the primary empirical result, we will further study the application of PCALWM in the credit decision, mainly on:

- The classification error with different training sample rate, that to say how the training sample rate affects the classification accuracy.
- the classification error under the different granting rate with certain training sample rate. This will help the credit issuers to determine the granting rate under certain risk level.

4.1 The Training Sample Rate to the Classification Error

We will study the different classifying training and testing sample how to affect the classification error. In other words , under what training sample rate, the model can get the better classification. We chose different training sample rate at random to construct the predict model, and in the granting rate of 75%, compared the two type error and the average error.

We chose five group training sample at random, including 500, 1000, 2000, 3000, 4000 records respectively, corresponding to the 10%, 20%, 40%, 60%, 80% of the total sample, and the five testing sample are the corresponding rest records, which including 4500, 4000, 3000, 2000, 1000 records. Table 3 shows the classification error of the five testing samples.

The results showed that with the incensement of the training rate, type I error and average error decreased, but the type II error no obviously change. The type I error

Table 3. The classification error of testing sample in different training rate

Training rate	10%	20%	40%	60%	80%
Type I error	0.3952	0.3481	0.3003	0.2652	0.2393
Type II error	0.1787	0.1749	0.1740	0.1763	0.1758
Average error	0.2113	0.2058	0.1937	0.1880	0.1832

deceased rapidly, but the speed tends to slow. This results show that a bigger training rate will help to PCALWM the get a better classification, especially to decrease the type I error.

4.2 The Classification Error Under Different Granting Rate with a Certain Training Rate

We chose seven granting rate groups with 80% training rate, figure 2 displays the error movements.

Fig. 2. The classification errors of testing sample in different granting rate

The results shows that as the granting rate increases, type I error increases, while type II error decreases sharply, and the bads' rate in granted group increases slowly. This shows the model has a good classification.

In practice, based on the tradeoff of marketing, enterprise stratagem and risk management, the credit issuers can make the choice of the two type error and average error to determine the best grant standard.

5 Summary and Conclusion

This paper has a further study on the principal component analysis linear-weighted model for credit scoring. It presents the application procedures in the credit scoring,

tested by a larger real-life credit card data, and some application features in the credit decision-making. The empirical results and application characteristics of PCALWM show the model will help the credit issuers to make the best trade-off among the enterprise stratagem, marketing and credit risk management.

Compared with some well-known credit scores, such as FICO behavior score, credit bureau score, first data resources' proprietary bankruptcy score and set enumeration decision tree score, Multiple criteria linear programming score, our model has a favorable classification.

Because of the great potential value in the consumer credit assessment, we will continue to perfect this model, we hope to get better results through more logical data processing and more suitable sample scale selecting, etc.

References

1. Alici, Y. (1995). Neural networks in corporate failure prediction: The UK experience. In: Proceedings of 3rd International Conference Neural Networks in the Capital Markets, A. N. Refenes, Y. Abu-Mostafa, J. Moody, and A. Weigend, Eds. London, UK, 393-406.
2. Board of Governors of the Federal Reserve System http://www.federalreserve.gov/releases/g19/hist/cc_hist_mt.txt
3. Guo, Q, Wu, W., Massart, D.L., Boucon, C., Jong, S. de. (2002). Feature selection in principal component analysis of analytical data. *Chemometrics and Intelligent Laboratory Systems*, 61:123-132
4. <http://www.firstdatacorp.com>
5. Li, J.P, Liu, J.L, Xu, W.X, Shi, Y. (2004). An Improved Credit Scoring Method for Chinese Commercial Banks. In: S.Y. Wang and X.G. Yang eds., *Advance in Risk Management, Lecture Notes in Decision Sciences*, Global-Link Publisher, Hong Kong, Vol.4:63-73
6. Li, J.P, Xu, W.X. (2004). Consumer Credit Assessment via Principal Component Analysis Linear-Weighted Model. *Chinese Journal of Management*, Vol.12, (2): 63-73
7. Lin, Y. (2002). Improvement On Behavior Scores by Dual-Model Scoring System. *International Journal of Information Technology and Decision Making*. (1): 153-164.
8. Sebzalli, Y.M., Wang, X.Z. (2001). Knowledge discovery from process operational data using PCA and fuzzy clustering. *Engineering Applications of Artificial Intelligence*, (14): 607-616
9. Shi, Y., Wise, M., Luo, M., Lin, Y. (2001). Data Mining in Credit Card Portfolio Management: A Multiple Criteria Decision Making Approach. In: Koksalan, M., Zionts, S. (eds.): *Multiple Criteria Decision Making in the New Millennium*. Springer, Berlin 427-436.
10. Shi, Y., Peng, Y., Xu, W., Tang, X. (2002). Data Mining Via Multiple Criteria Linear Programming: Applications in Credit Card Portfolio Management. *International Journal of Information Technology and Decision Making*. 1 :131-151.
11. The American Banker Association. <http://www.aba.com/default.htm>
12. Thomas, L.C. (2000). A survey of credit and behavioural scoring: forecasting financial risk of lending to consumers. *International Journal of Forecasting*, 16:149-172
13. Thomas, L.C., Edelman, D.B., Crook, J.N. (2002). *Credit Scoring and Its Application*. Society for Industrial and Applied Mathematics, Philadelphia
14. Westad, F., Hersleth, M., Lea, P.r, Martens, H. (2003). Variable Selection in PCA in sensory descriptive and consumer data. *Food Quality and Preference*, (14):463-472
15. www.fairisaac.com