Acta Technica *62* No. 4B/2017, 1–8

A combination model to improve financial distress forecasting[1]

Qian Zhuang[2,3], Wei Zhou[2], Shiying Hou[2], Shuzhen Chu[2]

Abstract. Some forecasting methods currently widely used for financial distress are not dynamic enough to improve the accuracy of the forecast. A dynamic forecasting method based on Markov forecasting method, information filtering theory and time series analysis, is designed. A combination model integrated of a process model and a discriminant model, which are used to describe the dynamic process and discriminant rules of financial distress respectively, is established. A general n-step-ahead forecasting algorithm based on Kalman filtering is derivated in order for the prospective forecast. The empirical results prove that the accuracy of the combination model for financial distress forecasting four years ahead is relatively higher.

Key words. Financial distress forecasting, Kalman filtering, process model, discriminant model..

1. Introduction

Under the new normal, the time dimension of financial information for decision-making is extended, but the valid period is shortened. The demand on the dynamics of financial distress forecasting is higher. Few previous studies on the forecasting methods of financial distress were conducted on the cumulative effect of historical information on current status taken into consideration in a long cycle.

In recent years, many scholars continuously improve and optimize data mining algorithm in order to increase the accuracy of dynamic forecasting of corporate financial distress [1]–[6]. Li Hui et al. (2014) have constructed a statistics-based wrapper for SVM-based FDI by using statistical indices of ranking-order information from

[1]This research is supported by the National Natural Science Foundation of China (Grant no. 71602188) and the Key program of National Social Science Foundation of China (Grant no. 15ZDB167).

[2]School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, Jiangsu 211198, China

[3]Corresponding author; e-mail: `zhuangcpu@163.com`

predictive performances on various parameters; Lin Fengyi, et al. (2014) have proposed an integrated approach to feature selection for financial distress forecasting problem that embeds expert knowledge with the wrapper method; etc. Usually they take the data of some cross-industry corporates from one year to three years before they are specially treated (ST) as the samples, but fail to fully take the cumulative effect of historical information on current status at the time of data processing into consideration. The cross-section analysis at a particular time point cannot be regarded as a complete dynamic forecasting method because the conclusions on discrimination among different time points are lack of logistic links. Their empirical results show that the models have better effect for short-term forecasting rather than long-term forecasting. This paper aims to dynamically improve the method of forecasting financial distress based on Kalman filtering, including model construction, algorithm improvement, and software programming application.

The rest of this paper is organized as follows. A dynamic forecasting method for corporate financial distress based on Kalman filtering theory is elaborated in Section 2. A combination model consisting of a process model and a discriminant model based on Kalman filtering algorithm is then described in Section 3. Section 4 presents empirical analysis for China's manufacturing industry. Section 5 draws conclusions and discusses future study.

2. A dynamic forecasting method based on Kalman filtering theory

It is a gradual variation process of corporate financial distress which can be described as the stochastic process of discrete time with Markov chain [7]. Markov forecasting method is employed to accurately describe the evolution trend of the corporate financial condition through the study on the initial distribution of each status of the corporate financial system and the probability of the status transfer.

Secondly, the information filtering problem is taken into consideration in the development of the method for forecasting financial distress. We design two ways of obtaining the estimated value of the corporate financial status at the very time point. One is one-step Markov forecasting for forecasting the current financial status from the financial status in the last time point and the other one is calculation of the current financial status by observing the current financial indicators and their functional relationship with the financial status. The optimal estimated value of the corporate financial status is finally obtained through comparison and error analysis.

Thirdly, a group of time-series data is used in Markov forecasting, and a continuous progressive forecasting of corporate financial status has been developed [8–12]. As shown in Fig. 1.

As shown in Fig. 1, if the corporate financial status at $t+2$ is to be forecast, only the optimal estimation of financial status at t and the signal indicator measured at $t+1$ require to be known. The rest may be deduced by analogy. If the corporate financial status at $t+n$ is to be forecast, only the optimal estimation of financial status at $t+n-2$ and the signal indicator measured at $t+n-1$ require to be known. When such continuous progressive forecasting method is employed, the

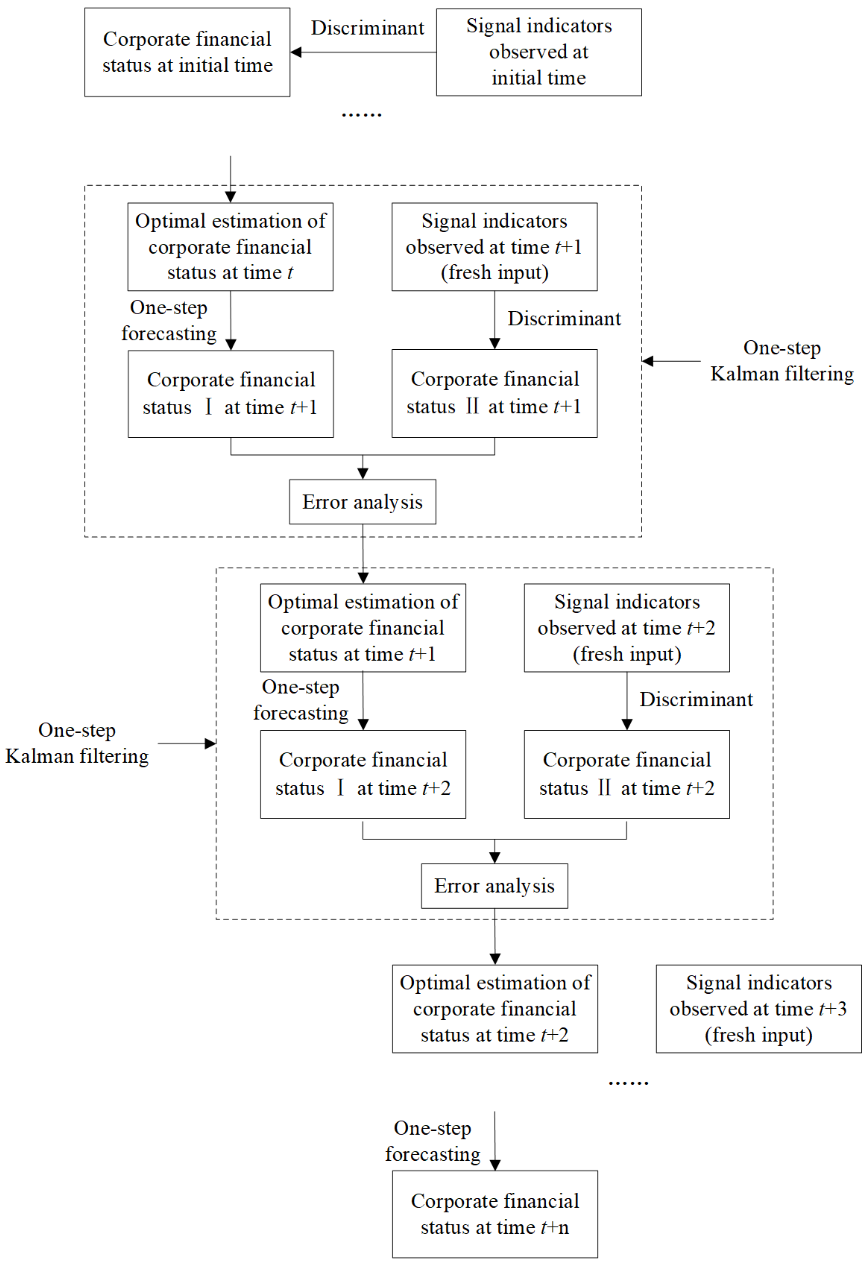

Fig. 1. Continuous progressive forecasting of corporate financial status

dynamic forecasting can be improved.

3. A combination model based on Kalman filtering algorithm

Based on the state-space method, a combination model of dynamically forecasting financial distress is constructed. A process model is constructed to describe the dynamic development process of the financial status. The "financial status" can be defined as a collection of vectors, which contains all information of the past behavior required by the corporate system rather than external effects produced through inputs (motivation) so that the future behavior of the corporate system can be uniquely described. The corporate financial status cannot be directly observed, but only the forecasting signal indicator related to such status can be observed. Therefore, the discriminant model will continue being constructed for describing the relationship between the financial status and forecasting signal indicator. So a dynamic model of forecasting financial distress is

$$X_t = A_{t|t-1} X_{t-1} + W_{t-1} , \tag{1}$$

$$Z_t = H_t X_t + V_t . \tag{2}$$

In the formula, X_t represents the financial status of a corporate at time t; Z_t represents the signal indicator of the corporate at time t. W_{t-1} refers to the process noise of the corporate financial status at time $t-1$; V_t refers to the measured noise of the signal indicator; $A_{t|t-1}$ is used for describing the dynamic process that the corporate financial status transit from time $t-1$ to time t; H_t is used to describe the relationship between the corporate financial status and its forecasting signal indicator at time t. Formula (1) refers to the process model and Formula (2) refers to the discriminant model. Wherein, assuming that the process noise and the measured noise refer to white noise, independent of each other and subject to normal distribution.

The equation can be solved based on Kalman filtering algorithm. Kalman filtering algorithm refers to the linear minimum variance estimation of status X_t based on observed value Z_t and it is an optimal estimation method. Kalman filtering process is divided into five parts:

$$\hat{X}_{t|t-1} = A_{t|t-1} \hat{X}_{t-1|t-1} , \tag{3}$$

$$P_{t|t-1} = A_{t|t-1} P_{t-1|t-1} A_{t|t-1}^T + Q_{t-1} , \tag{4}$$

$$K_t = P_{t|t-1} H_t^T [H_t P_{t|t-1} H_t^T + R_t]^{-1} , \tag{5}$$

$$\hat{X}_{t|t} = \hat{X}_{t|t-1} + K_t [Z_t - H_t \hat{X}_{t|t-1}] , \tag{6}$$

$$P_{t|t} = [I - K_t H_t] P_{t|t-1} . \tag{7}$$

In the process, K_t is Kalman gain, which is a blending factor that is used to adjust the discrepancy between the predicted observation $H_t\hat{X}_{t|t-1}$ and the actual observation Z_t, in order to obtain the optimal estimation $\hat{X}_{t|t}$ closer to the actual state.

Further, we continue to use the recursive method to derive the n-step-ahead forecasting algorithm from basic equations of Kalman filtering. Then the corporate financial state can be forecast n-step ahead from the perspective of long periods of time using high-frequency data. The n-step-ahead forecasting equation is

$$\hat{X}_{t+n|t} = A^n \hat{X}_{t|t} \ , \tag{8}$$

n-step-ahead forecasting error variance matrix is

$$\begin{aligned} P_{t+n|t} &= A^n P_{t|t} (A^T)^n + A^{n-1} Q (A^T)^{n-1} + \cdots + AQA^T + Q \\ &= A^n P_{t|t} (A^T)^n + \sum_{j=2}^{n+1} A^{n+1-j} Q (A^T)^{n+1-j} \ , \end{aligned} \tag{9}$$

where, we assume that the system parameters A_t and Q_t have nothing to do with time.

4. Empirical analysis

4.1. Sampling and data collection

Corporates in the manufacturing industry are selected as the study samples. 76 ST corporates which suffer financial distress due to "abnormal financial conditions" from 2013 to 2015 were selected as sample corporates. The time intervals are 2005-2013, 2006-2014 and 2007-2015 separately. The period during which those corporates were ST is recognized as Base Period T. Whether any corporate is ST depends on the financial conditions in Period T-1. Data from Period T-16 to Period T-1 is taken as the time series dataset of each sample of financial distress. Corporates similar in asset size and not ST during 8 years before the base period (16 periods before the base period) are taken as financially healthy sample corporates and the matching ratio is 1:1. The selection of period and time span of the healthy samples are basically the same with those of the distress samples.

There are 144 sample corporates in total. Wherein, 60 ST corporates in the first half of 2013 and the first half of 2014 and 60 matching corporates are taken as the training set, used for deducing the model; 12 ST corporates in the first half of 2015 and 12 matching corporates are taken as the test set, used for verifying the test results of the model.

Then, 27 indicators were selected as reserved indicators based on profitability, liquidity, operational capacity and market performance. Specifically, profit margin of main business, net profit margin, return on assets, return on net assets , income growth rate of main business and total assets growth rate are selected to reflect the profitability; current ratio, quick ratio, cash debt ratio, debt coverage ratio, debt to assets ratio and equity ratio are selected to reflect liquidity; total assets turnover ratio, fixed assets turnover ratio, current assets turnover ratio, inventory turnover ratio, accounts receivable turnover ratio, cash ratio of main business and cash return on assets are selected to reflect the operational capacity; earnings per share, net assets per share, operating revenue per share, capital reserve per share, undistributed profit per share, price to book value ratio, equity capital adequacy ratio and net cash flow per share are selected to reflect the market performance.

4.2. Experiment results and analysis

The principal component analysis is conducted to eliminate multicollinearity effects produced by original variables. The first ten principal components are extracted and the accumulative contribution rate can reach over 90 %. These principal components are linear combinations of original forecasting signal indicators. This linear combination equation can be used as a discrimination model. The data in the training set requires to be used in the parameter estimation of the process model, in combination with the discriminant model. The value are obtained through MATLAB programming during the gradually progressive forecasting process.

The following 24 test samples are tested for the accuracy of the forecast. Subject to space restrictions, the effect pictures of financial status of only 12 corporates are listed, including six ST corporates and six healthy corporates. Stock codes of corporates are as follows: 600071, 600408, 000590, 600779, 600962, 000059, 600666, 600397, 000790, 600616, 600127, and 600688. As shown in Fig. 2.

The testing results show that most of estimated value curves and true value curves fit well. Wherein, the mild warning occurs in 10 of 12 testing samples undergoing the financial distress in Period T-8 or Period T-7 (4 years ahead) and the severe warning occurs in 11 corporates in Period T-4 or Period T-3 (2 years ahead) and all 12 corporates in Period T-2 or Period T-1 (1 year ahead). It is indicated that the combination model can give relatively accurate warning before the occurrence of financial distress, with the warning accuracy reaching 83.3 % 4 years ahead and 91.7 % 2 years ahead and 100 % 1 year ahead. Meanwhile, we conduct the comparative analysis on the aforesaid samples with the three-layer BPNN model, the result of which shows that the early warning accuracy of 4 years ahead to 1 year ahead is 41.7 %, 58.3 %, 91.7 % and 100 % respectively. Compared with the early warning accuracy 1 year ahead and 2 years ahead, that the early warning accuracy 3 years ahead dramatically declines. It shows that the three-layer BPNN has better effect for short-term forecasting rather than long-term forecasting.

5. Conclusion and limitations

In this paper, we focus on the dynamic improvement on the forecasting method of financial distress based on Kalman filtering: the Markov forecasting method is employed to solve the logic problem in the continuous forecasting; the information filtering method is employed to filter the random disturbance and measurement errors or reduce their effect; the time series analysis method is used to gradually

Fig. 2. The forecasting effects of financial status for part of testing samples

deduce and design the continuous progressive forecasting method of the corporate financial status. Then a combination model for dynamic forecasting financial distress is constructed: a process model and a discriminant model that are used to describe the dynamic process and discrimination rules of financial distress respectively. In addition, a general n-step-ahead forecasting algorithm based on Kalman filtering is derivated in order for the prospective forecast. The empirical test results show that during the gradual progressive forecasting process, such model fully realizes the fitting of the estimated value and the true value and the accuracy is relatively high, with such accuracy 4 years ahead reaching 83.3 %.

The limitations of this study are as follows: Firstly, due to the availability of the data, the high frequency data such as daily data, weekly data, monthly data and quarterly data has not been researched as the basic data; Secondly, Kalman filtering equation of random linear discrete system is applied in the dynamic forecasting of financial distress and future research priorities are the development and the solution of nonlinear model as well as the application of extended Kalman filtering algorithm of which the range of application will be wider.

References

[1] E.I. Altman, M. Iwanicz-Drozdowska, E.K. Laitinen, A. Suvas: Financial distress prediction in an international context: A review and empirical analysis of Altman's Z-score model. Journal of International Financial Management & Accounting. 28 (2017), No. 2, 131-171.

[2] Z.Y. Li, J. Crook, G. Andreeva: Dynamic prediction of financial distress using Malmquist DEA. Expert Systems with Applications. 80 (2017), 94-106.

[3] L. Cleofas-Sanchez, V Garcia, A.I. Marques, J.S. Sanchez: Financial distress prediction using the hybrid associative memory with translation. Applied Soft Computing. 44 (2016). 144-152.

[4] L.G. Zhou, D. Lu, H. Fujita: The performance of corporate financial distress prediction models with features selection guided by domain knowledge and data mining approaches. Knowledge-based Systems. 85 (2015), 52-61.

[5] H. Li, C.J. Lia, X.J. Wu, J. Sun: Statistics-based wrapper for feature selection: An implementation on financial distress identification with support vector machine. Applied Soft Computing, 19 (2014), 57-67.

[6] F.Y. Lin, D. Liang, C.C. Yeh, J.C. Huang: Novel feature selection methods to financial distress prediction. Expert Systems with Applications. 41 (2014), No. 5, 2472-2483.

[7] S.V.B. Penda, M. Hoffmann, A. Olivier: Adaptive estimation for bifurcating Markov chains. Bernoulli. 23 (2017), No. 4B, 3598-3637.

[8] T. Arnold, M. Bertus, J.M. Godbey: A simplified approach to understanding the Kalman filter technique. Engineering Economist. 53 (2008), No.2, 140-155.

[9] Y. Shi, H.Z. Fang: Kalman filter-based identification for systems with randomly missing measurements in a network environment. International Journal of Control. 83 (2010), No. 3, 538-551.

[10] R.B. Geng, I Bose, X. Chen: Prediction of financial distress: An empirical study of listed Chinese companies using data mining. European Journal of Operational Research. 241 (2015), No. 1, 236-247.

[11] C.H. Cheng, S.H. Wang: A quarterly time-series classifier based on a reduced-dimension generated rules method for identifying financial distress. Quantitative Finance. 15 (2015), No. 12, 1979-1994.

[12] C.B. Balan, I.B. Robu, E. Jaba: The statistical assessment of financial distress risk in the case of metallurgical companies. Metalurgija. 54 (2015), No. 3, 575-578.

Received October 31, 2017