We are IntechOpen, the world’s leading publisher of Open Access books
Built by scientists, for scientists

3,900
Open access books available

116,000
International authors and editors

120M
Downloads

154
Countries delivered to

TOP 1%
Our authors are among the most cited scientists

12.2%
Contributors from top 500 universities

WEB OF SCIENCE™
Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com
Chapter 7

Application of Kalman Filtering in Dynamic Prediction for Corporate Financial Distress

Qian Zhuang

Additional information is available at the end of the chapter

http://dx.doi.org/10.5772/intechopen.71616

Abstract

This chapter aims to dynamically improve the method of predicting financial distress based on Kalman filtering. Financial distress prediction (FDP) is an important study area of corporate finance. The widely used discriminant models currently for financial distress prediction have deficiencies in dynamics. Based on the state-space method, we establish two models that are used to describe the dynamic process and discriminant rules of financial distress, respectively, that is, a process model and a discriminant model. These two models collectively are called dynamic prediction models for financial distress. The operation of the dynamic prediction is achieved by Kalman filtering algorithm, and further, a general $n$-step-ahead prediction algorithm based on Kalman filtering is derived for prospective prediction. We also conduct an empirical study for China’s manufacturing industry, and the results have proved the accuracy and advance of predicting financial distress in such case.

Keywords: financial distress prediction, pattern recognition, state space model, stochastic process, optimal estimation

1. Introduction

Research on financial distress prediction (FDP) is an important area of corporate finance. Early prediction methods are univariate analysis (UA), multiple discriminant analysis (MDA), logistic model, probit model, and so on [1–5]. With the development of computer technology, some new methods based on artificial intelligence technology with distributed computing capabilities that can deal with problems of nonlinear systems are widely introduced into the field of financial distress prediction. These methods include neural network (NN), genetic algorithm (GA), rough set theory (RST), case-based reasoning (CBR) and support vector machine (SVM), and so on [6–14]. Each model established for financial distress prediction, whether based on statistical methods or artificial intelligence methods, has advantages and disadvantages under...
different conditions. Let us take the most widely used multiple discriminant analysis (MDA) and back-propagation neural network (BPNN) for example. MDA has the advantage of simplicity and good interpretation, but the deficiency in its application is limited by strict assumptions that sometimes cannot be satisfied. Besides, MDA is a static discriminant model [2, 3, 6, 15, 20]. For the application of BPNN, it does not need any probability distribution assumption. BPNN is considered as an effective tool of pattern recognition for nonlinear systems. Therefore, many researchers have tried to apply triple BPNN in financial distress prediction, using the nonlinear pattern recognition capability of BPNN for classification of different financial state [7, 8, 15].

The prediction often achieved through a cross-sectional analysis at different time points. That is, the sample data of period \( t-1, t-2, \ldots \) before financial distress are studied by BPNN, respectively, and the features are extracted, based on what the judgment for the financial state of next new period is made [16, 17, 20]. This treatment is a relatively complete cross-sectional analysis. But the conclusions on discrimination among different time points are lack of logistic links. Therefore, this prediction is not completely dynamic. Furthermore, BPNN is a static neural network even when directly used in time-series prediction (Neural networks can be divided into static or dynamic neural networks based on whether they contain feedback loops or delay. BPNN is a backpropagation network without feedback and belongs to static neural networks.). There are some inherent problems such as overfitting, for example, the fitting error of training data has reduced, but the prediction error has increased at the same time. Even if the data are normalized, the effect is not satisfactory when the testing data are not sufficient [13, 18–20].

Actually, corporate financial distress is a gradual and cumulative process, which is developed from a healthy state. The mutation is often the result at which the gradual change and cumulation have reached the critical condition. It is neither reasonable nor logical if only the cross-sectional data at the time point prior to the occurrence of financial distress are used to make a determination for the corporate future state. It should take into account two aspects at least when conducting the research on financial distress prediction: firstly, the alternative data for prediction should contain all the historical information; secondly, the prediction method is dynamic designed for financial distress characterized by cumulative variation [21–23]. However, the current discriminant models have some deficiencies in dynamic prediction. Also, there is a problem of massive data processing. This chapter attempts to make a dynamic improvement on prediction methods for financial distress based on Kalman filtering algorithm in order to solve the above problems.

The main contribution of the paper is that it constructs a state-space model of corporate financial distress from the perspective of the cumulative effect of historical information on current state and improves Kalman filtering algorithm for dynamic prediction. A whole process of dynamic prediction for corporate financial distress is developed from a long period of time, and time-series data of high-frequency are collected for optimal estimation of financial state, which is seen as a stochastic process. The advantage of the model is proved by an empirical research, and the result shows that it can give relatively accurate warning before the occurrence of financial distress.

The rest of this chapter is organized as follows. Dynamic prediction models consisting of a process model and a discriminant model based on Kalman filtering algorithm are described in
Section 2. Then, a whole process of dynamic prediction for corporate financial distress is elaborated in Section 3. Section 4 presents empirical analysis for China’s manufacturing industry. Section 5 draws conclusions and discusses future research.

2. Dynamic prediction models based on Kalman filtering algorithm

Based on the state-space method, we establish two models, being used to describe the dynamic process and discriminant rules of financial distress, respectively, that is, a process model and a discriminant model. These two models collectively are called dynamic prediction models for financial distress. We see the evolution of financial distress for a company as a stochastic process and establish a process model, which is used to describe the dynamic process of development of the financial state. We define the financial state as a set of vectors, which summarizes all the information necessary about the past behavior of the company except for the external effects of the inputs, so that it can uniquely describe the behavior of the company in the near future [24]. The financial state of a company often cannot be observed directly, but only some signal indicators associated with the financial state can be observed. Therefore, we establish a discriminant model, which is used to describe the correlation between the financial state and the signal indicators. The discriminant model can be a recursive form of any statistical model or artificial intelligence model, theoretically. At first, we take the linear models, which are simple and intuitive as an example and establish dynamic prediction models for financial distress, as

\[
X_t = A_{t|t-1}X_{t-1} + W_{t-1} \\
Z_t = H_{t}X_t + V_t
\]

where \(X_t\) is the financial state of a company in period \(t\); \(Z_t\) is the signal indicators of the company in period \(t\); \(W_{t-1}\) is the process noise of the financial state in period \(t-1\); \(V_t\) is the observation noise of the signal indicators in period \(t\); \(A_{t|t-1}\) is used to describe the dynamic process of the financial state transferring from period \(t-1\) to \(t\); \(H_t\) is used to describe the mathematical relations between the financial state and the signal indicators in period \(t\). Eq. (1) is a process model; and Eq. (2) is a discriminant model.

Assume that the process noise and the observation noise are Gaussian white noises, which are mutually independent and normally distributed, i.e.

\[
\begin{align*}
E[W_i] &= 0, \quad E[W_iW_j^T] = Q_i\delta_{ij} \\
E[V_i] &= 0, \quad E[V_iV_j^T] = R_i\delta_{ij} \\
E[W_iV_j^T] &= 0
\end{align*}
\]

where, \(Q_i\) is a \(p \times p\)-dimensional symmetric nonnegative definite covariance matrix of process noise \(W_i\); \(R_i\) is a \(m \times m\)-dimensional symmetric positive definite covariance matrix of observation noise \(V_i\); \(\delta_{ij}\) is Kronecker - \(\delta\) function.
The above equations can be solved by Kalman filtering algorithm. The Kalman filter is named after Rudolph E. Kalman, who in 1960 published his famous paper describing a recursive solution to the discrete-data linear filtering problem. The Kalman filter is essentially a set of mathematical equations that implement a predictor–corrector type estimator that is optimal in the sense that it minimizes the estimated error covariance, when some presumed conditions are met [25, 26]. Kalman filter is widely used for its relative simplicity and robust nature. Rarely do the conditions necessary for optimality actually exist, and yet, the filter apparently works well for many applications in spite of this situation. Application of Kalman filter in dynamic prediction for corporate financial state consists of five steps [27, 28]:

The first step is to compute the one-step prediction of the financial state \( \hat{X}_{t-1|t-1} \) under the conditions of known \( \hat{X}_{t-1|t-1} \), which is the optimal estimation of the financial state at time \( t-1 \)

\[
\hat{X}_{t|t-1} = A_{t|t-1} \hat{X}_{t-1|t-1} 
\]

The second step is to compute the error covariance matrix \( P_{t|t-1} \) for one-step prediction

\[
P_{t|t-1} = A_{t|t-1} P_{t-1|t-1} A_{t|t-1}^T + Q_{t-1}
\]

The third step is to compute the Kalman gain \( K_t \), 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 financial state

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

The fourth step is to correct the one-step predicted financial state \( \hat{X}_{t|t-1} \) according to the principle of minimum error covariance and thus obtain the optimal estimation \( \hat{X}_{t|t} \) of the financial state

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

The fifth step is to compute the error covariance matrix \( P_{t|t} \) of the updated financial state estimation \( \hat{X}_{t|t} \)

\[
P_{t|t} = [I - K_t H_t] P_{t|t-1}
\]

These are the basic equations of Kalman filtering for a stochastic linear discrete financial system. The actual filtering process is an ongoing “predicting-correcting” process of a recursive nature. Figure 1 below offers a complete picture of the operation of the Kalman filter in dynamic prediction for corporate financial state.

The Kalman filter does not require storing large amount of data in solving the problem. Once new data are observed, new filtering value can be calculated at any time. Therefore, this method facilitates real-time processing and is easy to implement on the computer.
3. A whole process of dynamic prediction for corporate financial distress

As previously described, corporate financial distress is a gradual and cumulative process, which is developed from a healthy state, and so the prediction should be long-term and continuous and the continuously updated time-series data should be collected for the dynamic prediction, which could be the fresh input into the Kalman filter in order to obtain the optimal estimation closer to the actual state. The whole process of dynamic prediction for corporate financial distress is described as follows.

From Figure 2, we can see that if we want to predict the corporate financial state at time $t + 2$, we just need to know the optimal estimation of the corporate financial state at time $t$ and the signal indicators observed at time $t + 1$. The rest may be deduced by analogy; if we want to predict the corporate financial state at time $t + n$, we just need to know the optimal estimation of the corporate financial state at time $t + n - 2$ and the signal indicators observed at time $t + n - 1$. This continuous prediction does not require saving the observed data in the past. Every time the new signal indicators are observed, they are put into the Kalman filter as fresh. It helps solve the problems of storing, calling, and processing the massive data and thus greatly improving the speed of operation on the computer.

Further, if we want to predict the corporate financial state $n$-step ahead, we can obtain the $n$-step-ahead prediction algorithm derived from the basic Kalman filtering algorithm according to the dynamic prediction process described above.
The general $n$-step-ahead prediction algorithm is derived as

$$\hat{X}_{t+n|t} = A_{t+n|t}A_{t+n-1|t+n-2}\cdots A_{t+1|t}\hat{X}_{t|t} = \prod_{i=1}^{n} A_{t+i|t+i-1}\hat{X}_{t|t}$$  \hspace{1cm} (9)$$

The $n$-step-ahead prediction error variance matrix is
\[ P_{t+n|t} = \prod_{j=1}^{n} A_{t+j|t-1} P_{t|t} A_{t+j|t-1}^T + \sum_{j=2}^{n} A_{t+j|t-1} Q_{t+j|t-1} A_{t+j|t-1}^T + \cdots + A_{t+n|t-n-1} Q_{t+n|t-n-1} + Q_{t+n} \]

(10)

Assume that the system parameters \(A_t\) and \(Q_t\) have nothing to do with the time, then the above equations simplify to

\[ \hat{X}_{t+n|t} = A^n \hat{X}_{t|t} \]

(11)

\[ P_{t+n|t} = A^n P_{t|t} A^n + A^{n-1} Q(A^n)^{n-1} + \cdots + A Q A^T + Q = A^n P_{t|t} (A^n)^n \]

(12)

Based on the Eqs. (9)–(12), we could use data at shorter time interval to predict \(n\)-step ahead, but the prerequisite of sufficiently long-term data to find out the trend of development of the financial state should be satisfied.

In the dynamic prediction models for financial distress established in Section 2, we suppose that the financial state \(X\) cannot be observed. But in reality, whether the company is trapped in financial distress at time \(t-1\) and before can be known at time \(t\), that is part of \(X\) can be observed. We put this part of observed information into a likelihood equation in order to improve the accuracy of dynamic prediction. The probability of the company being trapped in financial distress is

\[ P(X_t > S_c) = \int_0^\infty p(x_t)dx_t \]

(13)

where, \(S_c\) is the critical value; \(x_{t+1}\) equals \(e_{t+1} = x_{t+1} - \hat{X}_{t+1|t}\) and the latter is normally distributed with mean of 0 and variance of \(F_{t+1} = H_{t+1} P_{t+1|t} H_{t+1}^T + R_{t+1}\). Then

\[ p(x_t) = \frac{1}{\sqrt{2\pi} P_t} e^{-\frac{1}{2}(\frac{x_t - \bar{x}_t}{\sigma_t})^2} \]

(14)

If \(M\) is the last year that \(X\) can be observed, then the additional estimation equation is

\[ l = -\frac{NM}{2} \log(2\pi) - \frac{1}{2} \sum_{l=1}^{M} \log(F_t) - \frac{1}{2} \sum_{l=1}^{M} e_t^T F_t e_t + \sum_{l=1}^{M} \{ \log(P(X_t > S_l) \delta(t) + \log(P(X_t < S_l) [1 - \delta(t)]) \}

(15)

This additional estimation equation is used every year, no matter if 1 year is divided into \(n\) periods. If standing at managers’ position and suppose we know the corporate financial state at time \(t-1\), when we stay at time \(t\), and then the additional estimation equation can be embedded in the general \(n\)-step-ahead prediction algorithm every time to help improve the accuracy of dynamic prediction.
4. Empirical analysis

4.1. Data description and experiment design

Manufacturing industry is a major industry in China. “Made in China” has an important impact on the global economy. Therefore, prediction of corporate financial distress for China’s manufacturing industry is of great significance. Generally, the manufacturing companies have complete production processes, equilibrinous production cycle, as well as a more stable trend of development of the financial state. The characteristics of these companies can be well described using the existing financial indicators, and the dynamic prediction method described above can be put into practice for these manufacturing companies.

In this research, the data for our experiment are collected from the Shanghai Stock Exchange and Shenzhen Stock Exchange databases in China. ST (special treatment) companies because of financial problems are selected as distress samples; meanwhile, companies of similar asset size that have never been special treated are selected as healthy samples. The ST time is treated as period T. For a 6-month interval, the data 8 years or 16 periods before ST are selected as time-series sets for the distress samples. The time span of the paired samples is the same as the distress samples.

According to the above principles, the data of 152 listed companies are collected, and the time span is year 2002 to year 2009, year 2003 to year 2010, year 2004 to year 2011, respectively. A total of 60 ST companies and 60 paired companies of the first half of year 2010 and 2011 are treated as training set, which is used to derive the model. A total of 16 ST companies and 16 paired companies of the first half of year 2012 are treated as testing set, which is used to test the effect of the model.

From the holistic perspective, we select 29 financial indicators covering four aspects of profitability, solvency, management efficiency, and market reaction as alternative signal indicators. The effect of the corporate financial problems may be amplified or reduced in information transmission mechanism of the market, and the problems may be exposed to the open market in advance or with a delay. If the problems are exposed in advance, the indicators can be used as a pilot signal of financial distress prediction; if delayed exposure, it can also be served as comprehensive evaluation of financial distress or the signal for the trend of development in the future. These are indicators of market reaction. The 29 signal indicators are listed in Table 1.

A three-dimensional database is established consisting of 16 periods’ time-series data of the above 152 sample enterprises, the financial state of which is represented by 27 signal indicators each (As operating profit margin growth ($Z_5$) and interest coverage ratio ($Z_{12}$) have much missing data, we ignore these two subsets of the data, leaving the rest 27 subsets.). The dynamic prediction method described above is based on the trend of the time-series data. The centralized tendency of signal indicators of profitability, solvency, management efficiency, and market reaction is shown in Figures 3–6 (Some indicators of management efficiency and market reaction show cyclical fluctuations, so we amend these indicators by smoothing, Figures 5 and 6 have been amended.).

From Figures 3–6, we can see most indicators show a certain trend, which is the foundation of dynamic prediction.
Then, we use nonparametric test of Mann-Whitney $U$ to find out when the difference between distress samples and healthy samples occurs. The results show that the gap between the two is maximized 2 years before ST time and the significant difference occurs 4 years before ST time, that is, the earliest time to accurately predict the occurrence of financial distress should be 4 years before ST time.

<table>
<thead>
<tr>
<th>Type</th>
<th>Code</th>
<th>Signal indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>Profitability</td>
<td>$Z_1$</td>
<td>Operating profit margin</td>
</tr>
<tr>
<td></td>
<td>$Z_2$</td>
<td>Net profit margin</td>
</tr>
<tr>
<td></td>
<td>$Z_3$</td>
<td>Return on assets</td>
</tr>
<tr>
<td></td>
<td>$Z_4$</td>
<td>Return on equity</td>
</tr>
<tr>
<td></td>
<td>$Z_5$</td>
<td>Operating profit margin growth</td>
</tr>
<tr>
<td></td>
<td>$Z_6$</td>
<td>Operating revenue growth</td>
</tr>
<tr>
<td></td>
<td>$Z_7$</td>
<td>Total assets growth</td>
</tr>
<tr>
<td>Solvency</td>
<td>$Z_8$</td>
<td>Current ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_9$</td>
<td>Quick ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{10}$</td>
<td>Cash debt ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{11}$</td>
<td>Debt coverage ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{12}$</td>
<td>Interest coverage ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{13}$</td>
<td>Liabilities to assets ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{14}$</td>
<td>Liabilities to equity ratio</td>
</tr>
<tr>
<td>Management efficiency</td>
<td>$Z_{15}$</td>
<td>Total assets turnover</td>
</tr>
<tr>
<td></td>
<td>$Z_{16}$</td>
<td>Fixed asset turnover</td>
</tr>
<tr>
<td></td>
<td>$Z_{17}$</td>
<td>Current assets turnover</td>
</tr>
<tr>
<td></td>
<td>$Z_{18}$</td>
<td>Inventory turnover</td>
</tr>
<tr>
<td></td>
<td>$Z_{19}$</td>
<td>Accounts receivable turnover</td>
</tr>
<tr>
<td></td>
<td>$Z_{20}$</td>
<td>Cash ratio of main business</td>
</tr>
<tr>
<td></td>
<td>$Z_{21}$</td>
<td>Cash return on assets</td>
</tr>
<tr>
<td>Market reaction</td>
<td>$Z_{22}$</td>
<td>Earnings per share</td>
</tr>
<tr>
<td></td>
<td>$Z_{23}$</td>
<td>Net assets per share</td>
</tr>
<tr>
<td></td>
<td>$Z_{24}$</td>
<td>Operating revenue per share</td>
</tr>
<tr>
<td></td>
<td>$Z_{25}$</td>
<td>Capital reserve per share</td>
</tr>
<tr>
<td></td>
<td>$Z_{26}$</td>
<td>Retained earnings per share</td>
</tr>
<tr>
<td></td>
<td>$Z_{27}$</td>
<td>Price to book ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{28}$</td>
<td>Equity to invested capital ratio</td>
</tr>
<tr>
<td></td>
<td>$Z_{29}$</td>
<td>Net cash flow per share</td>
</tr>
</tbody>
</table>

Table 1. Comprehensive signal indicators of financial distress prediction.
4.2. Experiment results and analysis

We use principal component analysis to eliminate the effect of multicollinearity on the original variables. We extract first 10 principal components, and the accumulative contribution rate is above 92% each for 152 companies. These principal components are linear combinations of the original signal indicators, which can be served as part of discriminant models for each company.

The parameters of process model are estimated from the data of training set and also using the data of training set, the judgment for the threshold of financial distress is set as an interval, which has lower and upper confidence limit.

The results show that the lower confidence limit is $-0.796$ and the upper confidence limit is $0.205$. When the predictive value of a company’s financial state is lower than $-0.796$, the
company may fall into severe financial distress; when the predictive value is higher than 0.205, the company is well operated; and when the predictive value is between 0.796 and 0.205, it is possible that the company is getting into financial distress.

Then, we test the effect of the dynamic prediction models using the data of testing set. Subject to space restrictions, we just list dynamic prediction figures for six companies, among which first three are ST companies, while the other three are non-ST companies. Names and stock codes of the companies are Sichuan Chemical Company Limited (000155), MCC Meili Paper Industry Co., Ltd. (000815), Guangzhou Guangri Stock Co., Ltd. (600894), Xinxiang Chemical Fiber Co., Ltd. (000949), Nantong Jiangshan Agrochemical & Chemicals Co., Ltd. (600389), Nanzhi Co., Ltd., and Fujian (600163), in turn. Dynamic prediction figures for these six companies are shown in Figure 7.
Figure 7. Dynamic prediction figures for part of testing samples.
The testing results show that almost all the curves of predictive value fits the ones of real value for 32 testing samples.

Of 16 distress testing samples, 15 companies give mild alarm in period $T-9$ (4 years ahead), and 13 companies give severe alarm in period $T-7$ and period $T-5$ (3 years ahead and 2 years ahead). All the 16 companies give severe alarm in period $T-3$ (1 year ahead). This shows that the information the dynamic model absorbed and produced almost covers the characteristics of financial distress. An accurate warning can be made 4 years before financial distress, and the accuracy is 93.8% (We also take triple BPNN to make a comparison. The results show that the accuracy of prediction 1 year ahead to 4 years ahead is 100, 93.8, 62.5, and 43.8%, respectively. The accuracy sharply declines 3 years ahead. It shows that the triple BPNN has better effect for short-term prediction rather than long-term prediction.)

For healthy testing samples, none is lower than the severe alarm limit. But sometimes, the predictive values appear slightly below the mild alarm limit, showing that there have been cases of temporary deviation from healthy state for healthy testing samples. The dynamic model conducts a track and thereafter modifies. This shows that the model can objectively track and effectively predict the overall financial state of a company from a long run.

5. Conclusions and future work

In this chapter, we focus on the dynamic nature of corporate financial distress and establish dynamic prediction models consisting of a process model and a discriminant model, which are used to describe the dynamic process and discriminant rules of financial distress, respectively. The operation of the dynamic prediction is achieved by Kalman filtering algorithm, and a general $n$-step-ahead prediction algorithm based on Kalman filter is derived for prospective prediction. To validate the prediction performance of this method, we conduct an empirical study for China’s manufacturing industry. The empirical results have proved the accuracy and advance of predicting financial distress using this dynamic model. The accuracy of prediction 4 years before financial distress is 93.8%. In addition, this method also solves the problem of massive data processing as it does not require storing large amounts of historical data and thus can achieve real-time processing of data.

In this research, we suppose the dynamic process of financial distress is linear. The Kalman filtering algorithm will be applied to a nonlinear dynamic model in the future research, and it will offer a wider range of applications.

Acknowledgements

This research is supported by National Natural Science Foundation of China (Grant no. 71602188) and National Social Science Foundation of China (Grant no. 15ZDB167).
Author details

Qian Zhuang

Address all correspondence to: zhuangcpu@163.com

School of International Pharmaceutical Business, China Pharmaceutical University, Nanjing, Jiangsu, China

References


