Research on Financial Distress Prediction Model Based on Kalman Filtering Theory

ZHUANG Qian[1], CHEN Liang-hua[2]
School of Economics and Management
Southeast University
Nanjing, China
[1]joyzq1228@163.com, [2]lianghua_c@vip.163.com

Abstract-**Research of enterprises' financial distress prediction (FDP) can generate early warning signals before the outbreak of financial crisis, and how to build a relative simplicity and robust FDP model has been of concern for theorists and practitioners at home and abroad. This research introduces Kalman filtering theory into FDP modeling. It builds a process model and a measurement model to describe the dynamic financial system. It uses time update and measurement update algorithm to solve the problem of financial information filtering. And thus, an adaptive model is proposed which is proved effective by an empirical analysis. This research is expected to provide theoretical support to achieve an accurate FDP and promote the application of FDP state-space model for enterprises.**

Keywords-financial distress prediction, Kalman filter, state-space model

I. INTRODUCTION

Research of enterprises' financial distress prediction (FDP) can generate early warning signals before the outbreak of financial crisis, so it can provide guidance for enterprises' financial risk precaution. Thus, it can effectively reduce the possibility of enterprises being trapped in financial difficulties due to the further deterioration of financial conditions. Theoretical and empirical research of dynamic modeling on FDP with longitudinal data streams from the view of individual enterprise has important practical and theoretical research value.

Many great classification techniques have been suggested to predict financial distress. Early studies of financial distress prediction focus on statistical techniques, such as multiple discriminate approach, logistic regression (Logit), and prohibit regression (Probit). In recent years, artificial intelligence approaches and data mining techniques, such as neural network (NN), support vector machine (SVM), case-based reasoning (CBR), have been widely applied to enterprises' financial distress prediction because of its universal approximation property and ability of extracting useful knowledge from vast data and domain experts, and also they don't have restrictive assumptions like traditional statistical approaches, such as linearity, normality and independence of input variables, which limits the effectiveness and validity of prediction. Neural network (NN) is one of the most widely used promising tools and have shown better predicative capacity than statistical techniques such as multiple discriminate analysis, logistic

regression. But NN is criticized by some scholars for "black box" property and over-fitting problems [1-2]. SVM solve local minima and over-fitting problems which are the main sources of trouble to conventional neural networks [3-5]. But the "black box" problem is still not solved. The basic problem is to determine the internal states of a system, given access only to the system's outputs. Two systems having totally different internal structures may exhibit the same external characteristics.

The many approaches to this basic problem are typically based on the state-space model. And this research aims to contribute to financial distress prediction by introducing a state-space model based on Kalman filtering theory.

II. RESEARCH FOUNDATION AND THEORY

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 (Kalman 1960). 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. Since the time of its introduction, the Kalman filter has been the subject of extensive research and application, particularly in the area of autonomous or assisted navigation. This is likely due in large part to advances in digital computing that made the use of the filter practical, but also to the relative simplicity and robust nature of the filter itself. Rarely do the conditions necessary for optimality actually exist, and yet the filter apparently works well for many applications in spite of this situation.

There are two basic building blocks of a Kalman filter, the process model and the measurement model. In general, the process model is of the form

$$x_t = \Phi x_{t-1} + \Gamma w_{t-1} \qquad (1)$$

The measurement model relates an unobserved variable x_t to an observable variable y_t, and the measurement model is of the form

$$y_t = H x_t + v_t \qquad (2)$$

$x \in R^n$ is the state of a discrete-time controlled process and $y \in R^m$ is a measurement. The random variables w_t and v_t represent the process and measurement noise respectively. They are assumed to be independent of each other, white, and with normal probability distributions,

i.e. $p(w_t) \sim N(0, Q_t)$ and $p(v_t) \sim N(0, R_t)$. The $n \times n$ matrix Φ in the difference equation (1) relates the state at the previous time step t-1 to the state at the current step t, in the absence of either a driving function or process noise. Note that in practice Φ might change with each time step, but here we assume it is constant. The $n \times l$ matrix Γ relates the noise input to the state x. The $m \times n$ matrix H in the measurement equation (2) relates the state to the measurement y_t. In practice H might change with each time step or measurement, but here we assume it is constant.

The Kalman filter is actually a linear minimum variance estimation of a state. The Kalman filtering process involves two steps: time update and measurement update.

Step 1: Time update. We predict a priori state estimate $\hat{x}_t^-$ and a priori estimate error covariance P_t^- in the case where the optimal estimate $\hat{x}_t$ is known in step t-1.

$$\hat{x}_t^- = \Phi \hat{x}_{t-1} \qquad (3)$$
$$\hat{P}_t^{-1} = \Phi P_{t-1} \Phi^T + Q_{t-1} \qquad (4)$$

On this basis, the Kalman gain K_t, which is used to adjust the weighted difference between an actual measurement y_t and a predicted measurement $H\hat{x}_t^-$, is obtained.

$$K_t = P_t^- H^T (H P_t^- H^T + R_t)^{-1} \qquad (5)$$

Step 2: Measurement update. Correct the priori estimate based on the observation error and the principle of minimum variance. Then the optimal state estimate $\hat{x}_t$ and the optimal estimate error covariance P_t can be obtained.

$$\hat{x}_t = \hat{x}_t^- + K_t (y_t - H\hat{x}_t^-) \qquad (6)$$
$$P_t = [I - K_t H] P_t^- \qquad (7)$$

The Kalman filter estimates a process by using a form of feedback control: the filter estimates the process state at some time and then obtains feedback in the form of noisy measurements. As such, the equations for the Kalman filter fall into two groups: time update equations and measurement update equations. The time update equations are responsible for projecting forward in time the current state and error covariance estimates to obtain the a priori estimates for the next time step. The measurement update equations are responsible for the feedback, i.e. for incorporating a new measurement into the a priori estimate to obtain an improved a posteriori estimate. The time update equations can also be thought of as predictor equations, while the measurement update equations can be thought of as corrector equations. Indeed the final estimation algorithm resembles that of a predictor-corrector algorithm for solving numerical problems.

III. FINANCIAL DISTRESS PREDICTION MODEL BASED ON KALMAN FILTERING THEORY

Based on the above analysis, this reaearch selects enterprises' financial data for continuous T years as object, and seizes the statistic characteristics of noise of enterprises' financial systems, e.g. financial statement window-dressing. The filter is designed observed variables for continuous T years for enterprises as inputs, and financial status or estimated parameters as outputs. Errors are adjusted by time update and measurement update algorithm, and the parameters are corrected constantly. Then the optimal filtering equations can be obtained.

The specific model is established as follows: $\{x_t: t=1, 2, \ldots\}$ stands for a sequence of the financial conditions for t years for a company; and $\{y_t: t=1, 2, \ldots\}$ stands for a sequence of financial ratios for t years for this company. Assuming x_t can not be observed，but related to y_t, So

$$y_t = B x_t + v_t \qquad (8)$$

where, B is a parameter vector which can be estimated from the data; $v_t \sim N(0, R_t)$; R_t is a covariance matrix of the observation noise v_t of the system, and R_t can be a time-independent vector. Among，y_t, B and R_t are $n \times l$ dimensional vectors at time t, n is the number of generalized principal components from the original data. Equation (8) is measurement model.

Process model is

$$x_t = A x_{t-1} + C w_{t-1} \qquad (9)$$

where, A and C are both parameters; $w_t \sim N(0, Q_t)$; Q_t is a covariance matrix of the process noise w_t of the system. In our research, x_t is one-dimensional, but actually x_t can be multi-dimensional.

Then, we establish k-order prediction, i.e. $p(x_{t+k} | \delta_t)$. Among, δ_t is the filtering; y_τ is the generalized dynamic principal component get from processed historical data, and $\tau = 1, 2, \ldots, t$; $k=1, 2, \ldots$. $X_{t+k} | \delta_t$ follows a normal distribution, and its mean and variance are:

$$Mean = \hat{x}_{t+k|t} = A^k \hat{x}_t \qquad (10)$$
$$Variance = P_{t+k|t} = A^k \hat{x}_t P_t (A^k)^T$$
$$+ \sum_{i=2}^{k} A^{i-1} C Q C (A^{i-1})^T \qquad (11)$$

where, superscript T stands for transpose; P_t is covariance matrix of $X_t | \delta_t$.

To be updated once there is observation y_{t+1}. $X_{t+1} | \delta_{t+1}$ also follows a normal distribution:

$$Mean = \hat{x}_{t+1} = A_{t+1} \hat{x}_t + P_{t+1|t} B_{t+1}^T F_{t+1}^{-1}$$
$$\times (y_{t+1} - B_{t+1} A_{t+1} \hat{x}_t) \qquad (12)$$
$$Variance = P_{t+1} = P_{t+1|t} - P_{t+1|t} B_{t+1}^T F_{t+1}^{-1} B_{t+1} P_{t+1|t}^T \quad (13)$$

where, $P_{t+1|t} = A_{t+1} P_t A_{t+1}^T + C_{t+1} Q_{t+1}^T C_{t+1}^T$；and also $F_{t+1} = B_{t+1} P_{t+1|t} B_{t+1}^T + H_{t+1}$.

Assuming equation (12) and (13) are both time-independent, the analysis can be further simplified.

If it is can be confirmed that $\hat{x}_0$ and P_0 are the optimal estimation on mean and variance of $x_0 | \delta = y_0$ at time t=0, then $\hat{x}_t$ and P_t can be obtained recursively. At time t=1, there is:

$$\hat{x}_1 = A\hat{x}_0 + (AP_0A + CQC)B^T - (B(AP_0A + CQC)B^T + H)^{-1} \times (y_1 - BA\hat{x}_0) \quad (14)$$

$$P_1 = AP_0A + CQC - (AP_0A + CQC)B^T(B(AP_0A + CQC)B^T + H)^T \times B(AP_0A + CQC) \quad (15)$$

Similarly, $\hat{x}_2$ and P_2 can be obtained from $\hat{x}_1$ and P_1 recursively.

It is necessary to identify the parameters B, Q, and H making use of the historical financial data before predicting enterprises' financial conditions in the application of state-space model. Maximum likelihood estimation method is chosen to estimate the relevant parameters.

Considering whether the company is trapped in financial distress each year is the public information, part of x can be observed in the sample. So we put this part of observed information into the likelihood equation in order to improve the accuracy of the equation. Therefore the final likelihood equation is:

$$l = -\frac{NM}{2}\lg(2\pi) - \frac{1}{2}\sum_{t=1}^{M}\lg|F_t|$$
$$- \frac{1}{2}\sum_{t=1}^{M}e_t^T F e_t + \sum_{t=1}^{M}(\lg(P(X_t > C_s))$$
$$\times\lambda(t) + \lg(P(X_t < C_s))\times(1-\lambda(t))) \quad (16)$$

Among, C_s is a threshold,

$$\lambda(t) = \begin{cases} 1 & \text{Financial distress} \\ 0 & \text{Other cases} \end{cases}$$

Financial ratios in the modeling sample can form a three-dimensional database. We write M program with Matlab according to equation (8)-(16), and then parameters in this model can be estimated. In this reseach A，B，C，Q，H，C_s are parameters.

We conduct an empirical experiment using the data drawn from listed companies in China's Shanghai and Shenzhen stock markets covering the period 2002-2011 to estimate values of the parameters based on the maximum likelihood rule. They are A=0.698; B=[-0.076, 0.699, 0.179, 1.526]^T; C=-0.026; Q=0.032;

$$H = \begin{pmatrix} 1.142 & 0 & 0 & 0 \\ 0 & 0.090 & 0 & 0 \\ 0 & 0 & 0.023 & 0 \\ 0 & 0 & 0 & 0.141 \end{pmatrix}$$

After test, the accuracy of estimation on test set by the new model is 91.8%.

IV. Conclusion

In this paper, Kalman filtering theory is introduced into establishing a financial distress prediction model for enterprises. It establishes a link between the observable information and unobservable system state through the state-space model, which is easier to understand and promote than the NN and other methods. Another advantage of application of Kalman filtering method is that the problem of financial information filtering can be solved by the use of time update and measurement update algorithm, and therefore efficiency in the use of financial information can be improved. At last experimental results show that the new model based on Kalman filtering theory can significantly enhance the efficiency of financial distress prediction for enterprises.

References

[1] Fang-Mei Tseng, Yi-Chung Hu, "Comparing four bankruptcy prediction models: Logit, quadratic interval logit, neural and fuzzy neural networks", Expert Systems with Applications, vol. 3, no. 37, pp. 1846-1853, 2010

[2] E Giovanis, "A study of panel logit model and adaptive neuro-fuzzy inference system in the prediction of financial distress periods", World Academy of Science, Engineering and Technology vol. 64, pp. 646-652, 2010

[3] Min J H, Lee Y C, "Bankruptcy prediction using support vector machine with optimal choice of kernel function parameters", Expert Systems with Applications, vol. 28, no. 4, pp.128-134, 2005

[4] Hui X-F, Sun J, "An application of support vector machine to companies' financial distress prediction", Lecture Notes in Artificial Intelligence, vol. 3885, pp. 274-282, 2006

[5] Yu Cao, Guangyu Wan, Fuqiang Wang, "Predicting financial distress of Chinese listed companies using rough set theory and support vector machine", Asia-Pacific Journal of Operational Research, vol. 28, no. 1, pp. 95-109, 2011