2011 International Symposium on Nonlinear Theory and its Applications
NOLTA2011, Kobe, Japan, September 4-7, 2011

Estimation of Transition of Credit Rating by using Particle Filters based on State Equations approximated by the Genetic Programming

Shozo Tokinaga† and Seigo Matsuno‡

†Graduate School of Economics, Kyushu University
Hakozaki 6-19-1, Higashi-ku, Fukuoka, 812-8581 Japan
‡Dept. of Business Administration, Ube National College of Technology
Tokiwadai 2-14-1, Ube, Yamaguchi, 755-8555 Japan
Email: tokinaga@en.kyushu-u.ac.jp, matsuno@ube-k.ac.jp

Abstract—This paper deals with the estimation of transition of credit rating by using Particle Filters (PFs) based on state equations approximated by the Genetic Programming (GP). Our aim is to find a true rating from observed ratings usually corrupted by noise, however, these works include simple scheme of time series modeling. In this paper, we generalize the PFs so that we approximate state equations by using the GP where individuals corresponding to state equations are improved according to the likelihood of PFs. At the same time, we include dynamics of financial ratios besides ratings in the nonlinear system equations, then we can expect improved estimation of true ratings.

1. Introduction

Since the borrower's credit quality affects the prices of securities such as corporate bonds in the market, rating about borrowers are done by major credit rating agencies, but also internally by banks [1]-[11]. Agency rating information is made public while the details of the rating process remain undisclosed. Even more, these reviews result in a rating change signifying improvement (upgrade) or deterioration (downgrade) in a borrower's creditworthiness (rating). However, there exists an evidence that credit rating does not fully reflect available information, and then deviates from true rating. There are several works using the Hidden Markov Model (HMM) scheme to estimate true credit quality (called simply rating in the following) [5] , but these are restricted to simple and linear systems and time series of rating itself rather than another available information such as financial statements. We propose the estimation of transition of bond rating by using Particle Filters (PFs) based on state equations approximated by the Genetic Programming (GP) [11]-[15].

We suppose that the true credit rating evolution can be described by state equations but that we do not observe true state directly. Rather, it is hidden in noisy or incomplete observations represented by the posted credit ratings. Based on the PFs scheme, we can generalize the estimation procedure including financial ratios besides observed rating [5]-[7]. In this paper, we also extend the PFs considering several kinds of nonlinear interactions to employ approximate state equations by using the GP where individuals corresponding to state equations are improved according to the likelihood of PFs. As applications, we show the evaluation of estimation scheme of the paper by simulation studies to artificial data, and also discuss the estimation of true rating from real time series of ratings.

2. Basics of rating transition and state estimation

Our goal is to estimate, from published credit ratings, the true credit ratings over time. We take true credit rating to be represented by state equations with state variables including financial ratios.

$$x_{i,t} = F(x_{1,t-1}) + v_{i,t}, y_t = H(x_{1,t}) + w_t \tag{1}$$

where the variable y_t is the observations at $t = 1, 2, ..., T$, and variables x_t are n dimensional state vectors to be estimated. Terms v_t and w_t are n dimensional and one dimensional random variables, respectively. Nonlinear functions $F(.)$ and $H(.)$ are vectors represented by variables $x_{1,t-1}$ and $x_{1,t}$, respectively. Generally, functions $F(.), H(.)$ are given, however, in the paper we assume that these functional forms are to be approximate by the GP.

The overview of PFs estimate $x_t = (x_{1,t}, x_{2,t}, ..., x_{n,t})$ from the observation y_t is summarized as follows. Define the symbols of i th particle for the variable $x_{k,t}, k = 1, 2, ..., n$ as $x^{(i)}_{k,t}$.

Step 1: generation of initial particles

Generate a set of random samples (N samples) for the variables $x^{(i)}_{k,t}$ at time 0 as $x^{(i)}_{k,0}, i = 1, 2, ..., N$ according to a prescribed probability density functions (pdf).

Step 2: generation of system noise

Generate N set of system noise $v^{(i)}_{k,t}, k = 1, 2, ..., n$ and $w^{(i)}_t$ for each particles according to prescribed pdf.

Step 3: one-step ahead prediction of $x^{(i)}_{k,t}$

One-step ahead prediction for the particle $x^{(i)}_{k,t}$ is given by

$$x^{(i)}_{t+1} = F(x^{(i)}_t, v^{(i)}_t), v^{(i)}_t = (v^{(i)}_{1,t}, v^{(i)}_{2,t}, ..., v^{(i)}_{n,t}) \tag{2}$$

where we assume that the calculations are node for each element $x^{(i)}_{k,t+1}$ of the vector $x^{(i)}_{t+1}$.

Step 4: calculation of weight

Assign a weight $w_t^{(i)}$ to each i th particle after measurement y is received by using the relation.

$$w_t^{(i)} = R_t(y_{t+1}|x_{t+1}^{(i)}) \quad (3)$$

where $R_t(.)$ is the conditional distribution for $y_{t+1}$by assuming that $x_{t+1}^{(i)}$ is given. Then, sum up the weight for normalization.

$$W_t = \sum_{i=1}^{N} w_t^{(i)} \quad (4)$$

Step 5: resampling of particles

Resample independently N times from the above discrete distribution. The resulting set of particle $X_{t+1} = (x_{t+1}^{(1)}, x_{t+1}^{(2)}, ..., x_{t+1}^{(N)})$ with weight $w_t^{(i)}/W_t$ from an appropriate sample (with equal weight to each element) from posterior pdf.

Step 6: iteration

Iterate procedure for time $t = 2, 3, ..., T$. Then, we have the estimation of pdf for the variables x_t at time $t = 2, 3, ..., T$. Finally we have the likelihood of the model as follows.

$$l_m = \sum_{t=1}^{T} \log W_t - T \log N \quad (5)$$

3. Functional approximation using the GP

So far, we tentatively assumed that function $H_t(.)$ is known, however actually the reviewing process of rating agencies can not modeled a priori. Then, we introduce the procedure for the approximation of $H_t(.)$ as well as state estimations.

In the GP method, we prepare many individuals representing functional forms, and improve the capabilities of approximation by using the genetic operations such as crossover operations. The fitness of an individual i th is evaluated by calculating the likelihood where the i th individual is used for state estimation in place of $H(.)$ in the state equation. After sufficient iterations of GP procedure, it is expected to obtain appropriate functional form to estimate true rating.

We iteratively perform steps in GP procedure until the termination criterion has been satisfied [8]-[11] (details are omitted here).

By considering the scope of the problem of functional forms treated in the paper, we restrict ourselves to the following forms used for functions (operators) included in individuals besides basic arithmetic operations.

trigonometric function: $\sin(x)$

hyperbolic function: $\sinh(x)$

step/ramp function: $u(x)$(unit step function), $r(x)$(ramped function) , these functions are defined as normalized to variable x.

$u(x) = 0, x \leq 0, u(x) = 1, x > 1$

$r(x) = 0, x \leq 0, r(x) = 1, x > 1$

4. Applications

We examine the capability of our estimation method based on the PFs and GP for artificially generated time series of ratings.

Applied to artificial data

We assume that the generation processes of time series described by the state equations are given a priori, and the true value of credit rating are known. Then, the estimated rating using the PFs and GP is compared to the true rating for each simulation study. As already mentioned, it is assumed that the first variable $x_{1,t}$ represents the rating at time t, and variables $x_{2,t}$, $x_{3,t}$ are financial ratios of the borrower.

By considering real application of rating, we assume followings for the function $H_t(.)$ while the time series of rating gradually changes rather than sudden change, and true rating itself has correlation with financial ratios in a sense.

(1) assuming correlation between rating and financial ratios.

It is expected that value of rating and financial ratios for sound (risky) firms are relatively large (small), because the operations of firms will be reflected to financial statements. Then, we assume that in the generation processes for financial ratios $x_{i,t} = F_i(x_{t-1})$, the first variable $x_{1,t-1}$ is always included, where the function $F_i(.)$ is given as a nonlinear function.

(2) smooth function $H_t(.)$ to deform true rating

We assume that the state x_t^P is deformed by a smooth function described by three forms having following meanings.

Case 1: high rating is deformed to larger than true value

Case 2: middle and lower rating are deformed to larger than true value

Case 3: function includes piecewise linear jumps

Additionally, we define Case 0 where the variable $x_{1,t}$ is not deformed and used direct as an observation. Then, we assume that three cases of deformity are defined by following functions.

Case 1: $H_t = (x/1.3)^{1.5}$

Case 2: $H_t = 4.3 \log(1 + x)$

Case 3: $H_t = (x + 0.4 * u(x - 1.0) + 0.4 * u(x - 1.8))/1.3$.

Conditions for simulation studies are summarized as follows.

levels of rating: 7 levels ranging from highest AAA to lowest CC

rating $x_{1,t}$: continuous value ranging 0〜3

estimated rating $\hat{x}_{1,t}$: continuous value ranging 0〜3

financial ratios $x_{t,j}$: continuous value ranging 0〜3

length of time series for $x_{t,j}$: 20

number of samples: generate 50 samples for each rating

number of particles in PFs: 10000

number of individuals in GP: 1000

maximum length of individuals in GP: 30

probability of crossover and mutation in GP: 0.2 and 0.01

Since we use seven levels of rating, we discretize continuous estimation $\hat{x}_{1,t}$. In the evaluation of the method we

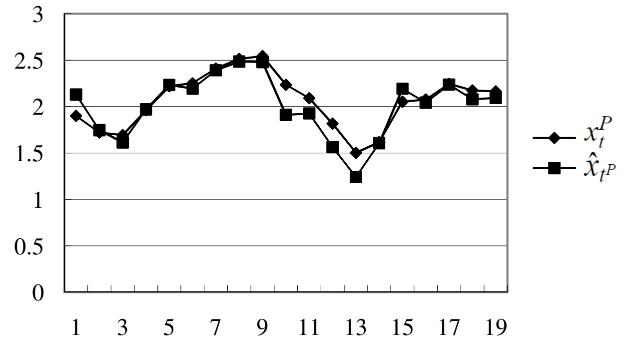

Fig.1-An example of state $x_{1,t}$ and estimation $\hat{x}_{1,t}$

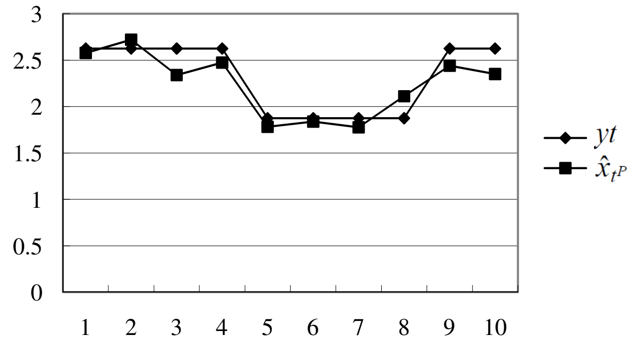

Fig.2-An example of observed time series y_t and estimation $\hat{x}_{1,t}$

compare the discretized levels for two variables by dividing values ranging 0〜3 into seven levels.

Fig.1 shows an example of estimation of state $x_{1,t}$, however, in the figure we only show the continuous state before discretization. Since the true rating $x_{1,t}$ is known beforehand, it is easy to compare the estimation $\hat{x}_{1,t}$. Fig.1 depicts both of them. Table 1 shows the result of comparison between $\hat{x}_{1,t}$ and $x_{1,t}$ which are represented by using discretized seven levels. In the table, the ratio of samples have identical values for $x_{1,t}$ and $\hat{x}_{1,t}$ among whole samples. It is seen from the result that number of samples where $\hat{x}_{1,t}$ is not equal to given $x_{1,t}$ is about 4% of the maximum value of $x_{1,t}$ (details are omitted here). As is seen from Table 1, the levels of estimated rating are almost the same as the levels of true ratings. These are ranging from 0.77 to 1.00 for the best cases, and the result implies us the effectiveness of our method.

Table 2 shows the estimated functional forms for $H(.)$ using the GP (we denote $x_{1,t}$ as x for simplicity). The result shows us that the approximated functions are slightly deviated from true (given) forms, however, the original forms are in part included in the estimated functions. As a result, estimation of true rating based on the PFs and GP proposed in the paper are proved to be useful in a general applications of rating to find true rating.

Table 1-Rate of proper estimation of ratings

rating	Case 0	Case 1	Case 2	Case 3
1	0.77	0.84	0.84	0.89
2	0.98	0.84	0.90	0.93
3	0.86	0.80	0.82	0.99
4	0.89	0.87	0.80	0.93
5	0.83	0.82	0.78	0.91
6	0.90	0.75	0.82	0.87
7	1.00	0.90	0.76	1.00

Table 2-Estimation of functions H_t by the GP

cases	$\hat{H}(.)$
Case 0	$0.93x + 0.2u(x - 0.8) + 0.1$
Case 1	$(x/1.18)^{1.61} + 0.01x + 0.02u(x - 1.2) + 0.02$
Case 2	$4.4\log(0.94 + 1.15x) + 0.12u(x - 1.1) + 0.2$
Case 3	$0.89x + 0.17u(x - 1.1) + 0.35u(x - 1.9) + 0.04$

Applied to real data

Then we apply the method to real data of the Japanese bond market. The overview of the real data is summarized as follows.

years: from 1998 to 2007, representative 95 firms

details of industry: machinery 30, electric machinery 55, electricity 10

levels of rating: four denoted as AAA, AA, A, BBB

We use nine financial ratios for the estimation process of rating (details are omitted here). However, these financial ratios are too much for the analysis, then we reduce them into three variables by using the principal component analysis. Then, we have four state variables including $x_{1,t}$ for rating and $x_{i,t}, i = 2, 3, 4$ for financial ratios.

The conditions for the simulation studies such as the definition of state variables and the number of particles for PFs are the same as in simulation studies for artificial data. Different form artificial data analysis, we do not know the true rating $x_{1,t}$ (also functional form $y_t = H(x_{1,t}) + w_t$ and the variances of $v_{i,t}$ for state variables $x_{i,t}$). Then, we apply successive estimation for these unknown parameters. For the first estimation, $x_{1,t}$ is replaced by published rating, and the process of $x_{i,t}$ is approximated by a linear regression models so that the variances of $v_{i,t}$ are estimated. In the GP procedure, we successively improve these estimation by using nonlinear functional approximations.

Fig.2 shows an example of the estimation process by giving observed y_t and estimation $x_{1,t}$. However, the simple comparison of these variables plays no role in real applications. As is seen from the result, we see considerable difference between y_t and $\hat{x}_{1,t}$, and the difference ensures the value of the method.

Different from artificial data analysis, the true rating can not be known, and the simple comparison between x_t^P and $\hat{x}_{1,t}$ has no meaning. Therefore, we define the number of cases where observed y_t and $\hat{x}_{1,t}$ is not the same by representing the value to find the discrepancy (difference) in observations based on our method. Table 3 shows the rate of cases (denoted as d_I) where y_t is different from $\hat{x}_{1,t}$ in the whole simulation studies. Also we show the extended evaluation where it is allowed to regard as the same level if the difference between y_t and $\hat{x}_{1,t}$ remains within one level. It is seen from the result that with respect to d_I, the difference between two rating is about 35%, and also in d_{II} we can find about 20% in the difference between observation and estimation. The fact implies that it is valuable to estimate the true rating also in the real investment.

Table 3-Rate of difference between y_t and $\hat{x}_{1,t}$

ratings	AAA	AA	A	BBB
d_I	0.38	0.34	0.29	0.30
d_{II}	0.18	0.21	0.20	0.21

5. Conclusion

In this paper, we showed the estimation of transition of bond rating by using PFs based on state equations approximated by the GP. We generalize the PFs so that we approximate state equations by using the GP where individuals corresponding to state equations are improved according to the likelihood of PFs. As applications, we showed the evaluation of estimation scheme of the paper by simulation studies to artificial data.

For future works we widely apply the method to real data of securities market, and further works will be done by the authors.

Acknowledgments

This research was partly supported by the Japan Society for the Promption of Science under the grant number (B) 23310104 , and authors would like to thank the organization.

References

[1] E. I. Altman, "The importance of subtlety of credit rating migration," *Journal of Banking and Finance*, vol.22, no.10, pp.1231–1247, 1998.

[2] M. Carty and M. Hrycay, "Parameterizing credit risk models with rating data," *Journal of Banking and Finance*, vol.25, pp.197–270, 2001.

[3] A. I. Barkman,"Testing the Markov chain approach on accounts receivable," *Management Accounting*, vol.62, no.7, pp.48–50, 1981.

[4] R. A. Jarrow, D. Lando and S. M. Turnbull, "A Markov model for the term structure of credit risk spread," *Review of Financial Studies*, vol.10, no.2, pp.481–523, 1997.

[5] M. W. Korolkiewicz and R. J. Elliott, "A hidden Markov model of credit quality," *Journal of Economic Dynamics & Control*, vo.32, pp.3807–3819, 2008.

[6] D. Lando and T. Skodeberg, "Analyzing rating transitions and rating drift with continuous observations," *Journal of Banking and Finance*, vol.26, pp.423–444, 2002.

[7] A. Doucet, N. de Freitas and N. Gordon, *Sequential Monte Carlo Methods in Practice*, New York, Springer Verlag, 2001.

[8] A. Doucet, N. J. Gordon and V. Krishnamurthy, "Particle filters for state estimation of jump Markov linear systems," *IEEE Trans. on Signal Processing*, vol.49, no.3, pp.613–624, 2001.

[9] J. H. Kotecha and P. M. Djuric, "Gaussian sum particle filtering," *IEEE Trans. on Signal Processing*, vol.51, no.10, pp.2602–2612, 2003.

[10] A. Zia, T. Kirubrajan, J. P. Reilly, D. Yee, K. Punithakumar and S. Shirani, "An EM algorithm for nonlinear state estimation with model uncertainty," *IEEE Trans. on Signal Processing*, vol.56, no.3, pp.921–936, 2008.

[11] S. Tokinaga and K. Tan, "State estimation from time series by using particle filters based on equation approximations with the Genetic Programming and its applications to suppression of fluctuation (in Japanese)," *IEICE Trans. Fundamentals*, vol.J93-A, no.11, pp.739–755, 2010.

[12] Y. Ikeda and S. Tokinaga, "Multi-fractality analysis of time series in artificial stock market generated by multi-agent systems based on the Genetic Programming and its applications," *IEICE Trans. Fundamentals*, vol.E90-A, no.9, pp.2212–2222, 2007.

[13] Y. Ikeda and S. Tokinaga, "Analysis of price changes in artificial double auction markets consisting of multi-agents using Genetic Programming for learning and its applications," *IEICE Trans. Fundamentals*, vol.E90-A, no.9, pp.2203–2211, 2007.

[14] K. Tan, S. Tokinaga, and M. Chu, "Estimation of decision processes and states by using particle filters and functional approximations based on the Genetic Programming and its applications (in Japanese)," *IEICE Technical Report*, vol.106, no.264, pp.47–52, 2006.

[15] S. Tokinaga and Y. Kishikawa, "Estimating generation models of time-series including jump diffusion processes based on the Genetic Programming and multi-stage fuzzy inference systems (in Japanese)," *IEICE Trans. Fundamentals*, vol.J93-A, no.5, pp.365–374, 2010.