IEICE The Institute of Electronics, Information and Communication Engineers

IEICE Proceeding Series

Dynamical Portfolio Theory by Nonlinear Bagging Predictors

Kiyoharu Tanaka, Tomoya Suzuki

Vol. 1 pp. 527-530
Publication Date: 2014/03/17
Online ISSN: 2188-5079

2012 International Symposium on Nonlinear Theory and its Applications
NOLTA2012, Palma, Majorca, Spain, October 22-26, 2012

Dynamical Portfolio Theory by Nonlinear Bagging Predictors

Kiyoharu Tanaka† and Tomoya Suzuki†

†Graduate School of Science and Engineering, Ibaraki University,
Nakanarusawa-cho 4-12-1, Hitachi-shi, Ibaraki, 316-8511, Japan.
Email: 12nm923x@hcs.ibaraki.ac.jp, tsuzuki@mx.ibaraki.ac.jp

Abstract—In the Markowitz's mean-variance portfolio model, the probability distribution of a future return is composed by the recent historical prices, and then a future return and a future risk are estimated as the mean value and the standard deviation of the distribution. Namely, the future return is predicted by a simple moving average, and the risk is nothing but a historical fluctuation. In this study, to improve the prediction accuracy of the future return, we apply the nonlinear prediction method following local spatial dynamics, and to estimate the future risk, we produce the probability distribution aggregating predicted values by the Bagging algorithm. Then, each risk is reduced by making a portfolio, that is, the portfolio effect. Namely, our method tries to improve the prediction accuracy and to reduce the risk of its prediction error, simultaneously. To confirm the validity of our method, we performed investment simulations. As results, we could obtain higher profit and realize lower risk of investment than the conventional method.

1. Introduction

Markowitz's portfolio theory [1] is useful to decide the allocation of stocks for investment, and can reduce its risk by the portfolio effect. For this theory, it is necessary to calculate the future return and the risk of each stock, but these are completely unknown. For this reason, the conventional method estimates them by the mean value and the standard deviation of the recent historical data. Namely, because this estimation of future returns corresponds to a moving average prediction, it might be insufficient to predict financial systems, which are typical examples of complex systems.

In the present study, we apply a nonlinear prediction model and the Bagging algorithm [2] to the conventional portfolio theory. First, a nonlinear prediction is used to improve the prediction accuracy of future return rates because it can model the relationship between the past and the future, that is, the temporal evolution of financial systems even if this relationship is nonlinear. Moreover, the Bagging algorithm is used as an ensemble learning to estimate the probability distribution of a future return rate. Especially, in Ref. [3], the mean value of the ensemble distribution composed by nonlinear predictors is used as a predicted value, and this ensemble learning can improve prediction accuracy than a single nonlinear prediction. Furthermore, Ref. [4] has reported that the standard deviation of this ensemble distribution has a relationship with the difficulty of prediction. Namely, we can consider this standard deviation as a risk, and try to reduce the risk by making a portfolio.

In Sec.2, we introduce the conventional method to make a portfolio. In Sec.3, we propose a new portfolio model which can improve prediction accuracy by nonlinear prediction and also can reduce its prediction risk by the portfolio effect. In Sec.4, we perform some investment simulations with real stock prices to confirm the validity of our proposed method.

2. How to Make a Portfolio

2.1. Markowitz's Mean-variance Model

If we denote x_i as the price of i-th stock ($i = 1, 2, \cdots, N$) as the time of t, the return rate $r_i(t)$ is given by

$$r_i(t) = \frac{x_i(t) - x_i(t-1)}{x_i(t-1)}. \qquad (1)$$

In Markowitz's portfolio model, a future return and a future risk are expected by the mean value and the standard deviation of the future probability distribution. However, because this distribution is completely unknown, an empirical distribution made by the recent historical data is used to estimate the return rate $\tilde{r}_i(t+1)$ and the risk $\tilde{\sigma}_i(t+1)$, which are given by

$$\tilde{r}_i(t+1) = \overline{r_i}(t) = \frac{1}{T}\sum_{a=0}^{T-1} r_i(t-a), \qquad (2)$$

$$\tilde{\sigma}_i(t+1) = \sigma_i(t) = \sqrt{\frac{1}{T}\sum_{a=0}^{T-1}[r_i(t-a) - \overline{r_i}(t)]^2}, \qquad (3)$$

where T means the length of the historical data.

Then, in the case of making a portfolio with N stocks, the expect return rate $\tilde{r}_p(t+1)$ and the expected risk $\tilde{\sigma}_p(t+1)$ of the portfolio are respectively given by

$$\tilde{r}_p(t+1) = \sum_{i=1}^{N} c_i \tilde{r}_i(t+1), \qquad (4)$$

$$\tilde{\sigma}_p(t+1) = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} c_i c_j \tilde{\sigma}_{ij}(t+1)}, \qquad (5)$$

where c_i is an allocation rate and $\sum_{i=1}^{N} c_i = 1$. Then, $\tilde{\sigma}_{ij}(t+1)$ is defined by

$$\tilde{\sigma}_{ij}(t+1) = \sigma_{ij}(t)$$
$$= \frac{1}{T}\sum_{a=0}^{T-1} [r_i(t-a) - \overline{r}_i(t)] \cdot \left[r_j(t-a) - \overline{r}_j(t)\right]. \quad (6)$$

As the Portfolio effect [1], the risk of a portfolio $\tilde{\sigma}_p(t+1)$ can be reduced as the number of stocks for the portfolio N is larger and these stocks have less correlation among them. Here, if all of the N stocks have no correlation, $\tilde{\sigma}_{ij}(t+1) = 0$, $(i \neq j)$. Thus, Eq.(5) is rewritten by

$$\tilde{\sigma}_p^2(t+1) = \sum_{j=1}^{N} c_j^2 \tilde{\sigma}_j^2(t+1). \quad (7)$$

Here, the upper bound of $\tilde{\sigma}_j^2(t+1)$ is set as $\tilde{\sigma}_j^2(t+1) \leq P$,

$$\tilde{\sigma}_p^2(t+1) \leq \left(c_1^2 + \cdots + c_N^2\right) P. \quad (8)$$

If N stocks are allocated uniformly like $c_i = \frac{1}{N}$, Eq.(8) is rewritten by

$$0 \leq \tilde{\sigma}_p^2(t+1) \leq \frac{P}{N}. \quad (9)$$

Therefore, $\tilde{\sigma}_p^2(t+1) \to 0$ if $N \to \infty$.

2.2. The Sharpe Ratio

If we allocate $\{c_i\}$ so as to maxmize $\tilde{r}_p(t+1)$ and minimize $\tilde{\sigma}_p(t+1)$, this investment can be reasonable. From this viewpoint, the Sharpe ratio:

$$S_r(t) = \frac{\tilde{r}_p(t+1) - r_f}{\tilde{\sigma}_p(t+1)} \quad (10)$$

has been proposed [5]. In the present study, we maximize S_r to optimize the allocation of $\{c_i\}$. Moreover, if $\tilde{r}_p(t+1) < 0$, we take a sell position to make it a positive value. Then, r_f means a risk-free return, but we set $r_f = 0$ because the short-term interest rate has been nearly zero in Japan.

3. Nonlinear Prediction for Dynamical Systems

3.1. The Local Linear Approximation

First, to reproduce the background dynamics which derived the time-series data $r_i(t)$, we reconstruct a multidimensional attractor $\boldsymbol{v}(t)$ from $r_i(t)$ by the Takens embedding method:

$$\boldsymbol{v}_i(t) = \{r_i(t), r_i(t-\tau_i), \ldots, r_i(t-\tau_i(d_i-1))\}, \quad (11)$$

where τ_i means a delay time, and d_i means an embedding dimension. Then, we merge all of $\{\boldsymbol{v}_i(t)\}$ into an attractor:

$$\boldsymbol{V}(t) = \{\boldsymbol{v}_1(t), \boldsymbol{v}_2(t), \ldots, \boldsymbol{v}_N(t)\}. \quad (12)$$

Next, we predict the future state of $\boldsymbol{V}(t)$ by the local linear approximation method [6] as a nonlinear prediction. Here, some local neighbors $\boldsymbol{V}(t_k)$, $k = 1 \sim K$, are selected from all of the historical attractor $\boldsymbol{V}(t)$, that is, $t_k < t$. Then, by averaging the next states of the neighbors, we can obtain the predicted value of $\boldsymbol{V}(t+1)$ as follows:

$$\tilde{\boldsymbol{V}}(t+1) = \frac{1}{K}\sum_{k=1}^{K} \boldsymbol{V}(t_k+1). \quad (13)$$

This prediction accuracy depends on the embedded $\boldsymbol{V}(t)$, and namely we have to take care of setting the embedding parameters of τ_i and d_i. For this reason, we applied the cross-validation method. However, if we examine the optimum values of every stock, it takes enormous time cost. Therefore, as a simplification, we merged each τ_i into a parameter like $\tau = N\tau_i$, and merged each d_i into $d = Nd_i$. Then, these τ and d were optimized by the cross-validation method.

3.2. The Bagging Predictors for the Probability Distribution of a Future Return

As mentioned in Sec.3.1, because the local linear prediction uses only local data whose length is not long enough, the ensemble learning is useful to improve the prediction accuracy [3] and to estimate the possibility of its prediction error [4]. Here, the ensemble learning applied for prediction is called the bootstrap aggregating (bagging) predictors [2]. In our portfolio model, these bagging predictors are applied to improve the prediction accuracy of future returns and to estimate the risk of each prediction.

First, we randomly sample K neighbors from $\{\boldsymbol{V}(t_k)\}$ with replacement, and obtain a new set of near neighbors. Then, we apply the nonlinear prediction of Eq.(13) to the new neighbors, and can obtain another predicted value $\tilde{\boldsymbol{V}}_b(t+1)$. After repeating this procedure B times, we can estimate the possible distribution of the future value as $\tilde{\boldsymbol{V}}_b(t+1)$, $b = 1 \sim K$. In the field of financial engineering, because the expected return and the risk correspond to the mean value and the standard deviation of the possible distribution, we estimate the final predicted value by

$$\tilde{\boldsymbol{V}}(t+1) = \frac{1}{B}\sum_{b=1}^{B} \tilde{\boldsymbol{V}}_b(t+1). \quad (14)$$

The predicted value of Eq.(14) can be more accurate than that of Eq.(13) due to the effect of the ensemble learning [3]. Moreover, because the predicted future return $\tilde{r}_i(t+1)$ is included in $\tilde{\boldsymbol{V}}(t+1)$ of Eq.(13) and if each return of $\tilde{V}_b(t+1)$ is denoted as $\tilde{r}_{i,b}(t+1)$, we can rewrite Eq.(14) as

$$\tilde{r}_i(t+1) = \frac{1}{B}\sum_{b=1}^{B} \tilde{r}_{i,b}(t+1). \quad (15)$$

By substituting Eq.(15) for Eq.(4), we can calculate the expected return rate of a portfolio.

Next, because the risk of a portfolio is considered as the

Table 1: Results of the asset growth rates $\{M\}$ gotten by all of the portfolio with two stocks. The index P_s means the shortfall probability of $M < 1$, that is, the frequency of making a loss. Each underlined number means the better performance of our model or Markowitz's model, and each bold number means the best performance of each category.

	Daily data		Two weekly data		Monthly data	
	Our model	Markowitz's model	Our model	Markowitz's model	Our model	Markowitz's model
mean$\{M\}$	**4.77**	0.97	1.07	0.87	1.10	0.91
max$\{M\}$	**47.95**	3.89	22.40	4.87	15.42	4.95
min$\{M\}$	0.12	0.07	0.09	**0.13**	0.11	0.07
P_s	**0.15**	0.60	0.59	0.70	0.58	0.68

standard deviation of the possible return rates, we estimate the risk by

$$\tilde{\sigma}_i(t+1) = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\left[\tilde{r}_{i,b}(t+1) - \tilde{r}_i(t+1)\right]^2} \quad (16)$$

Similarly, the covariance $\tilde{\sigma}_{ij}(t+1)$ is given by

$$\tilde{\sigma}_{ij}(t+1) = \frac{1}{B}\sum_{b=1}^{B}\left[r_{i,b}(t+1) - \tilde{r}_i(t+1)\right]\cdot\left[r_{j,b}(t+1) - \tilde{r}_j(t+1)\right]. \quad (17)$$

By substituting Eqs.(16),(17) for Eq.(5), we can calculate the expected risk of a portfolio. Then, by maximizing S_r to optimize the allocation rate of the portfolio, we can reduce the total risk $\tilde{\sigma}_p(t+1)$ due to the portfolio effect. Namely, our method aims to improve the prediction accuracy and to reduce the risk of its prediction error, simultaneously.

4. Investment Simulations

In this section, to confirm the validity of our proposed method, we perform some investment simulations with 50 kinds of real stocks traded on the Tokyo Stock Exchange. Then, the time scale of stock prices is daily, two weekly, or monthly data, and the trading period is from 2000 to 2005. After the period, we calculate the asset growth rate M, dividing the amount of the final asset by that of the initial asset. Therefore, as M is larger than one, this investment realized better performance. However, if $M < 1$, it means that the final asset was below the initial asset, that is, we got loss.

As the first simulation, we compose each portfolio with only two stocks to statistically compare the performance between the Markowitz's portfolio model and our proposed model. Here, two stocks of each portfolio do not change during the investment period, and so we can obtain the results of ${}_{50}C_2$ portfolios.

These results are shown in Table 1. We can confirm that our model improves the asset growth rate M and reduce the frequency of $M < 1$ in every time scale. Namely, our model can realize not only higher return but also lower risk than the conventional portfolio model.

To show these results visually, Fig.1 shows the comparison between the Markowitz's model and our proposed model. The region above each straight line means the advantage of our proposed model, and the ratio of the results included there is 85% in daily data, 58% in two weekly data, and 57% in monthly data.

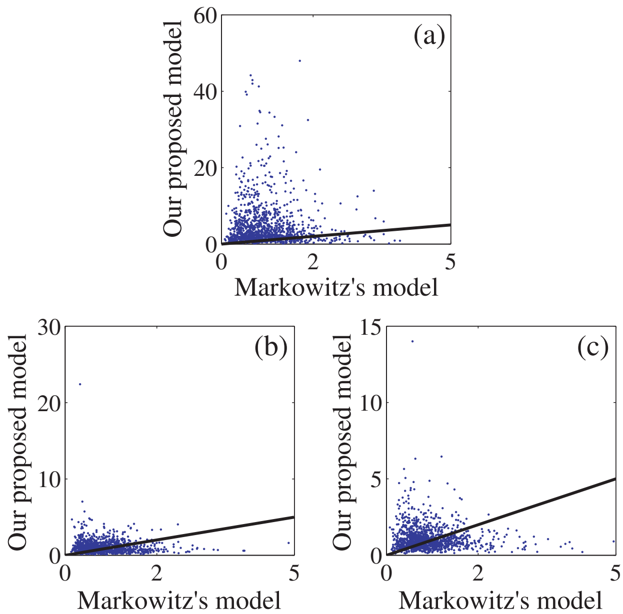

Figure 1: Comparison between the asset growth rates $\{M\}$ of the Markowitz's mean-variance model and those of our proposed model when making each portfolio with two stocks of (a) daily data, (b) two weekly data, or (c) monthly data. The slope of each straight line is one.

As you can see in Table 1 and Fig.1, the advantage of our model is reduced if we use two weekly data or monthly data to make portfolios. It might be because the temporal structure of the original price movements was destroyed by large time scale of sampling stock prices and then the sampled data became like random walk, and therefore the nonlinear prediction did not work well. The same discussion is attempted in Ref. [7] by using a dynamical model of financial systems.

As shown in Sec.2.1, we can reduce the total risk of a

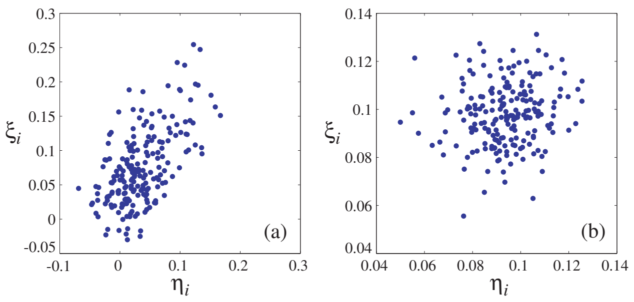

Figure 2: Correlations between the likelihood of each prediction model $\{\eta_i\}$ to the learning data and its prediction accuracy $\{\xi_i\}$ after the learning data. Each correlation is (a) 0.596 or (b) 0.119, and each prediction is performed by (a) our model or (b) Markowitz's model.

portfolio theoretically by using many stocks for a portfolio. To confirm this portfolio effect, we compose a multi-stock portfolio with more than two stocks. Here, we hope to select profitable stocks whose prediction accuracy is high, but this prediction accuracy is unknown. However, if there is some correlation between the likelihood of a prediction model $\{\eta_i\}$ to the learning data and its prediction accuracy $\{\xi_i\}$ after the learning data, we can estimate the unknown ξ_i before the prediction by using the already known likelihood η_i.

To confirm this possibility, Fig.2 shows correlations between $\{\eta_i\}$ and $\{\xi_i\}$. We can see that Fig.2(a) has larger correlation than Fig.2(b), that is, our proposed model has the advantage in selecting more profitable stocks. However, this advantage disappeared if we used two weekly or monthly data for our model because the prediction itself is very hard.

Finally, Fig.3 shows the performance of multi-stock portfolio with the N stocks, which were selected from the viewpoint of the likelihood η_i. In Markowitz's model, the asset growth rate M becomes larger as the number of stocks N is larger, which is caused by the portfolio effect. However, in our proposed model, the portfolio effect is not confirmed. If N is larger, the portfolio has to include even unfavorable stocks whose prediction accuracy is lower, and therefore, the performance of our portfolio might have decreased. In addition, it can be also considered that the embedding dimension $d = Nd_i$ became too large to find local neighbors properly for the nonlinear prediction. However, in any cases, our proposed model shows higher performance than the Markowitz's model.

5. Conclusions

We proposed how to improve the Markowitz's portfolio model by using the bagging algorithm with the local linear (nonlinear) prediction. First, the nonlinear prediction improved the prediction accuracy of future returns because the

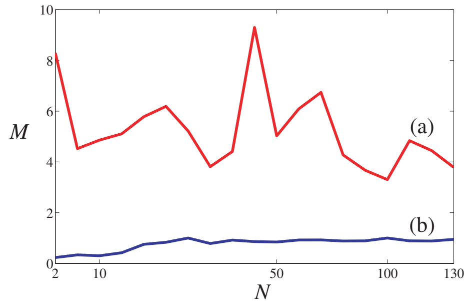

Figure 3: Results of the asset growth rate M gotten by each portfolio with N stocks of daily data, which is composed by (a) our proposed model or (b) Markowitz's model.

prediction method of Markowitz's model is a simple moving average of the recent historical data. Then, because the bagging algorithm can estimate the possible distribution of a future return, its mean value and its standard deviation were expected as the future return and the risk of trading a single stock. Moreover, by applying these expected values to the Markowitz's portfolio model, the risk was reduced because of the portfolio effect. Through investment simulations with real stock data, we confirmed that our model can realize higher profit and lower risk, simultaneously. However, the portfolio effect of our multi-stock portfolio did not work well although the nonlinear prediction makes it possible to select profitable stocks by referring the likelihood of the prediction. One of this reason might be due to the embedding dimension, and from this viewpoint, we try to improve the portfolio effect of the multi-stock portfolio as a future work.

References

[1] H. M. Markowitz: "Portfolio Selection," Journal of Finance, Vol.7, No.1, pp.77–91 1952.

[2] L. Breiman: "Bagging Predictors," Machine Learning, Vol. 24, pp. 123–140, 1996.

[3] D. Haraki, T. Suzuki, H. Hashiguchi, and T. Ikeguchi: "Bootstrap Nonlinear Prediction," Physical Review E, Vol. 75, 056212, 2007.

[4] K. Nakata and T. Suzuki: "Evaluating the Risk of Nonlinear Prediction with the Bagging Algorithm," Proc. of NCSP'12, pp. 748–791, 2012.

[5] W. F. Sharpe: "Capital asset prices : A theory of market equilibrium under conditions of risk," Journal of Finance, Vol. 19, No. 3, pp. 425–442, 1964.

[6] J. D. Farmer and J. J. Sidorowich: "Predicting chaotic time series," Physical Review Letters, Vol. 59, pp. 845–848, 1987.

[7] T. Suzuki: "Appropriate time scales for nonlinear analyses of deterministic jump systems," Physical Review E, Vol. 83, No.6, 066203, 2011.