IEICE Proceeding Series

Stock Portfolio Management Based on Nonlinear Prediction Model

Satoshi Inose, Tomoya Suzuki, Kazuo Yamanaka

Vol. 1 pp. 523-526
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

Stock Portfolio Management Based on Nonlinear Prediction Model

Satoshi Inose†, Tomoya Suzuki†† and Kazuo Yamanaka††

†Graduate School of Science and Engineering, Ibaraki University,
Nakanarusawa-cho 4-12-1, Hitachi-shi, Ibaraki, 316-8511, Japan.
††College of Engineering, Ibaraki University,
Nakanarusawa-cho 4-12-1, Hitachi-shi, Ibaraki, 316-8511, Japan.
Email: 12nd202x@hcs.ibaraki.ac.jp, tsuzuki@mx.ibaraki.ac.jp, yam@mx.ibaraki.ac.jp

Abstract—A re-formulation of the Markowitz mean-variance portfolio model is attempted in order to make it suitable to more complicated situations of practical financial markets. An AR model with data-dependent coefficients are introduced for predicting the future return in place of the simpler arithmetic mean of past data. This means that a class of nonlinear predictor is used. The prediction error variance also replaces the variance of past data as the evaluation index for the risk. A computer simulation based on practical data on stock prices suggests that investment by using the new portfolio model results in higher profit with lower risk.

1. Introduction

The Markowitz mean-variance portfolio model is often used for stock allocation and can minimize the total risk of a portfolio by optimizing allocation rates [1]. According to this model, a future return rate is estimated as the most probable value, and a risk is estimated as a possible error of the estimated future return. For these estimation, we have to obtain the probability distribution of the future return rate. However, because this probability distribution is unknown, the previous model defines the expected return rate and the expected risk as the mean value and the standard deviation of historical return rates [1].

From the viewpoint of time-series prediction, this estimation of future returns might be too simple to predict real financial markets because the mean value of historical time series data corresponds to simply averaging the past data. If financial markets are efficient and unpredictable as Fama mentioned [2], it is meaningless to improve prediction models. However, various inefficiency of financial markets has been reported through many empirical analyses. Especially, these properties can be confirmed in daily data, and becomes weaker as the time scale of data becomes larger, like weekly data or monthly data [3].

To model such properties of real markets, some mathematical models have been proposed. As prediction models of future return rates, the AR model, the ARMA model, etc. have been applied [4]. However, this prediction accuracy is not good enough. For this reason, some fundamental factors such as macroeconomic and/or financial indexes have been applied for making a multi-factor model. Especially, the Black-Litterman model is well-known recently which uses not only historical return rates but also investor's view of the future to estimate the expected return rate [5]. However, because the propose of our study is to modify the Markowitz mean-variance portfolio model, we try to make a portfolio only by following past evolution of return rates. Therefore, we discuss how to apply more advanced prediction models to the original portfolio model. However, if we change how to expect future return rates, we have to reconsider the risk defined by the original model. Here, if we regard the risk as the probability that the expected future will not be realized, we could estimate the risk by using historical prediction errors. To confirm the validity of our method, we perform investment simulations with real stock prices.

2. The Markowitz Mean-Variance Portfolio Model

We denote $x_i(t)$ as the price of ith stock ($i = 1, 2, \cdots, N$) at the time of t, and then the return rate $r_i(t)$ is given by

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

Next, we denote $P_i^{(k)}$ as the probability that $r_i(t+1)$ becomes $r_i^{(k)}(t+1)$ at the future time of $t+1$, where $\sum_{k=1}^{K} P_i^{(k)} = 1$. Thus, the expected return rate $\tilde{r}_i(t+1)$ and the variance $\tilde{\sigma}_i(t+1)$ are respectively given by

$$\tilde{r}_i(t+1) = \sum_{k=1}^{K} P_i^{(k)} r_i^{(k)}(t+1), \tag{2}$$

$$\tilde{\sigma}_i^2(t+1) = \sum_{k=1}^{K} P_i^{(k)} \left[r_i^{(k)}(t+1) - \tilde{r}_i(t+1) \right]^2, \tag{3}$$

where $\tilde{\sigma}$ is considered as the degree of risk involved in decision making [1].

Here, if we make a stock portfolio by allocation rates $\{d_i\}$, where $\sum_{i=1}^{N} d_i = 1$, the expected return rate $\tilde{r}_p(t+1)$ and the expected risk $\tilde{\sigma}_p(t+1)$ of the portfolio are respectively given by

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

$$\tilde{\sigma}_p(t+1) = \sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N} d_i d_j \tilde{\sigma}_{ij}(t+1)}, \quad (5)$$

where $\tilde{\sigma}_{ij}(t+1)$ is defined as

$$\tilde{\sigma}_{ij}(t+1) = \sum_{k=1}^{K} P_{ij}^{(k)} \left[r_i^{(k)}(t+1) - \tilde{r}_i(t+1)\right] \cdot \left[r_j^{(k)}(t+1) - \tilde{r}_j(t+1)\right], \quad (6)$$

and $P_{ij}^{(k)}$ means the probability that $r_i(t+1)$ and $r_j(t+1)$ become $r_i^{(k)}(t+1)$ and $r_j^{(k)}(t+1)$.

However, because $P_i^{(k)}$, $P_{ij}^{(k)}$, $r_i^{(k)}(t+1)$ and $r_j^{(k)}(t+1)$ are unknown, the future probability distribution is estimated by historical data of the last T period. Therefore, Eqs.(2), (3) and (6) are rewritten by

$$\begin{aligned} \tilde{r}_i(t+1) &= \bar{r}_i(t) \\ &= \frac{1}{T}\sum_{a=0}^{T-1} r_i(t-a), \quad (7) \\ \tilde{\sigma}_i(t+1) &= \sigma_i(t) \\ &= \sqrt{\frac{1}{T}\sum_{a=0}^{T-1}\left[r_i(t-a) - \bar{r}_i(t)\right]^2}, \quad (8) \\ \tilde{\sigma}_{ij}(t+1) &= \sigma_{ij}(t) \\ &= \frac{1}{T}\sum_{a=0}^{T-1}\left[r_i(t-a)-\bar{r}_i(t)\right]\cdot\left[r_j(t-a)-\bar{r}_j(t)\right]. \quad (9) \end{aligned}$$

Then, by substituting these values into Eqs.(4) and (5), we can obtain

$$\tilde{r}_p(t+1) = \sum_{i=1}^{N} d_i \bar{r}_i(t), \quad (10)$$

$$\tilde{\sigma}_p^2(t+1) = \sum_{i=1}^{N}\sum_{j=1}^{N} d_i d_j \sigma_{ij}(t). \quad (11)$$

Moreover, if we allocate $\{d_i\}$ so as to maximize $\tilde{r}_p(t+1)$ and minimize $\tilde{\sigma}_p(t+1)$, this investment can be reasonable. For this reason, we apply a typical measure: the sharpe ratio S_r [6] is given by

$$S_r = \frac{\tilde{r}_p - r_f}{\tilde{\sigma}_p}, \quad (12)$$

and optimize $\{d_i\}$ so as to maximize S_r. Here, r_f is a risk-free rate, and the overnight unsecured call money [7] is used for it. Then, in the case that $\tilde{r}_p < 0$, we build a selling portfolio to realize a positive $\tilde{r}_p$. Moreover, because maximizing S_r is a convex quadratic programming problem, we use the interior point method.

3. Application of Time Series Prediction Models

In the Markowitz mean-variance portfolio model, we can see that $\tilde{r}_i(t+1)$ of Eq.(7) is predicted by simply averaging the past data. In the present paper, we call it the moving average prediction. On the other hand, if a future movement depends on past historical movements and this relationship can be approximated by a super plane, we can expect the future return rate of Eq.(7) by the autoregressive (AR) model:

$$\tilde{r}_i(t+1) = \sum_{a=0}^{T-1} \beta_a r_i(t-a) + \beta_T. \quad (13)$$

Here, to estimate model factors $\beta_a (a = 0, 1, \cdots, T)$, we apply the least square method to the learning data, which is historical data of the last L period. Then, Eq.(13) can be rewritten in a vector form:

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{F},$$

where

$$\begin{aligned} \boldsymbol{Y} &= \left[r_i(t), r_i(t-1), \cdots, r_i(t-(L-T)+1)\right]^{\mathrm{t}}, \\ \boldsymbol{X} &= \begin{bmatrix} \boldsymbol{r}_i(t-1) & 1 \\ \boldsymbol{r}_i(t-2) & 1 \\ \vdots & \vdots \\ \boldsymbol{r}_i(t-(L-T)) & 1 \end{bmatrix}, \\ \boldsymbol{r}_i(t-a) &= \left[r_i(t-a), r_i(t-a-1), \cdots, r_i(t-a-(T-1))\right], \\ \boldsymbol{F} &= \left[\beta_0, \beta_1, \cdots, \beta_T\right]^{\mathrm{t}}. \end{aligned}$$

Therefore, the model factors of Eq.(13) can be estimated by

$$\tilde{\boldsymbol{F}} = \left[\boldsymbol{X}^{\mathrm{t}}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^{\mathrm{t}}\boldsymbol{Y}.$$

The AR model is classified into a linear prediction model because it approximates all learning data by a regression plane.

In addition, we can also use a nonlinear prediction model which can realize not only a regression plane but also a regression surface. First, we rewrite Eq.(13) as

$$\tilde{r}_i(t+1) = \left[\boldsymbol{r}_i(t) \;\; 1\right] \cdot \boldsymbol{F}. \quad (14)$$

Here, it is better if the model factors $\boldsymbol{F}$ can change according to similar input variables $\boldsymbol{r}_i(t)$. That is, we assume that similar outputs are generated by similar inputs. Namely, we think similar historical data $\boldsymbol{r}_i(t)$ are more important to estimate $\boldsymbol{F}$, and we apply the weighted least squares method according to the distance $l_i(t_a)$:

$$l_i(t_a) = |\boldsymbol{r}_i(t) - \boldsymbol{r}_i(t-a)|, \quad (15)$$

where $\boldsymbol{r}_i(t)$ is a target input and $\boldsymbol{r}_i(t-a)$ are the other learning data. Then, the weighted factor $w_i(t_a)$ is given by

$$w_i(t_a) = \exp(-l_i(t_a)). \quad (16)$$

Next, we prepare the following diagonal matrix:

$$\boldsymbol{W} = \begin{bmatrix} w_i(t_1) & 0 & \cdots & 0 \\ 0 & w_i(t_2) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & w_i(t_{L-T}) \end{bmatrix}, \quad (17)$$

and then, the model factors of Eq.(13) can be estimated by

$$\tilde{\boldsymbol{F}} = \left[\boldsymbol{X}^{t}\boldsymbol{W}^{t}\boldsymbol{W}\boldsymbol{X}\right]^{-1}\boldsymbol{X}^{t}\boldsymbol{W}^{t}\boldsymbol{W}\boldsymbol{Y}. \tag{18}$$

Here, this estimation of $\tilde{\boldsymbol{F}}$ is approximated locally with a super plane, but the slope of each plane changes locally according to the state of $\boldsymbol{r}_i(t)$. Namely, this approximation corresponds to a nonlinear regression with a super surface globally . In the present study, we call it the nonlinear autoregressive (NAR) model, which is extended from the AR model by the local linear approximation method [8].

4. Modification of the Original Mean-Variance Model

In order to apply not only the moving average prediction used for the previous portfolio theory, but also any other prediction models, we have to modify Eqs.(7), (8), and (9). For this reason, we reconsider them on the basis of their original Eqs. (2), (3), and (6).

According to Eq.(2), because the expected return rate $\tilde{r}_i(t+1)$ means the most probable future, we can replace it by the predicted value. Here, this idea is not too strange because Eq.(7) of the previous theory can be considered as a predicted value by the moving average prediction. Then, by substituting the predicted value to Eq.(4), we can calculate the expected return rate of a portfolio $\tilde{r}_p(t+1)$.

Moreover, in Eq.(3), we can say that the previous portfolio theory regards the risk $\tilde{\sigma}_i(t+1)$ as the expected value of prediction errors. Therefore, we denote the prediction error $r_i^{(k)}(t+1) - \tilde{r}_i(t+1)$ as $\sigma_i^{(k)}(t+1)$, and then Eq.(3) is rewritten by

$$\tilde{\sigma}_i(t+1) = \sqrt{\sum_{k=1}^{K} P_i^{(k)}\left[\sigma_i^{(k)}(t+1)\right]^2}. \tag{19}$$

Although $P_i^{(k)}$ and $\sigma_i^{(k)}(t+1)$ are unknown as well as the previous portfolio theory, these are estimated by historical prediction errors $r_i(t') - \tilde{r}_i(t')$ $(t' \leq t)$. Therefore, Eq.(3) can be rewritten by

$$\tilde{\sigma}_i(t+1) = \sqrt{\frac{1}{T}\sum_{a=0}^{T-1}\left[r_i(t-a) - \tilde{r}_i(t-a)\right]^2}. \tag{20}$$

On the other hand, the previous portfolio theory considers only $r_i^{(k)}(t+1)$ included in $\sigma_i^{(k)}(t+1)$ when Eq.(3) is modified into Eq.(8). However, in the case of predicting future returns aggressively like our study, because the essence of the risk is based on prediction errors, we estimate the whole of $\sigma_i^{(k)}(t+1)$ by past prediction errors.

Similarly, the covariance of Eq.(6) is rewritten by

$$\begin{aligned}\tilde{\sigma}_{ij}(t+1) &= \sum_{k=1}^{K} P_{ij}^{(k)}\sigma_i^{(k)}(t+1)\sigma_j^{(k)}(t+1)\\ &= \frac{1}{T}\sum_{a=0}^{T-1}\left[r_i(t-a)-\tilde{r}_i(t-a)\right]\left[r_j(t-a)-\tilde{r}_j(t-a)\right].\end{aligned} \tag{21}$$

Then, we can estimate $\tilde{\sigma}_p(t+1)$ by substituting this $\tilde{\sigma}_{ij}(t+1)$ into Eq.(5).

However, in actual use, we have to take care of the data length T in Eq.(21). This T tends to be set a small value by the cross validation method to optimize T. Then, if we apply the cross validation method to optimize T, this T tends to be set a small value. In this case, it is highly possible that some covariance of Eq.(21) between two stocks shows strong negative correlation. Then, we rewrite Eq.(5) as

$$\begin{aligned}\tilde{\sigma}_p^2 &= \sum_{i=1}^{N}\sum_{j=1}^{N} d_i d_j \tilde{\sigma}_{ij}\\ &= \sum_{i=1}^{N} d_i^2\tilde{\sigma}_i^2 + 2\sum_{i=1}^{N}\sum_{j>i}^{N} d_i d_j \tilde{\sigma}_{ij}.\end{aligned} \tag{22}$$

If the two indexes $i = 1$ and $j = 2$ have completely negative correlation, their covariance becomes $\tilde{\sigma}_{12} = -\tilde{\sigma}_1\tilde{\sigma}_2$. Then, by setting allocation rates except d_1 and d_2 to 0,

$$\begin{aligned}\tilde{\sigma}_p^2 &= d_1^2\tilde{\sigma}_1^2 + d_2^2\tilde{\sigma}_2^2 + 2d_1d_2\tilde{\sigma}_{12}\\ &= d_1^2\tilde{\sigma}_1^2 + d_2^2\tilde{\sigma}_2^2 - 2d_1d_2\tilde{\sigma}_1\tilde{\sigma}_2\\ &= (d_1\tilde{\sigma}_1 - d_2\tilde{\sigma}_2)^2,\end{aligned}$$

and therefore,

$$\tilde{\sigma}_p = |d_1\tilde{\sigma}_1 - d_2\tilde{\sigma}_2|. \tag{23}$$

Here, if we set like $d_1 : d_2 = \tilde{\sigma}_2 : \tilde{\sigma}_1$ and $d_{i\notin\{1,2\}} = 0$, we can realize no risk: $\tilde{\sigma}_p = 0$, and the sharpe ratio of Eq.(12) diverges to infinity. That is, if the prediction errors of any two stocks have a negative correlation, the best allocation rates to maximize S_r can be decided only by $\tilde{\sigma}_1$ and $\tilde{\sigma}_2$. In this optimization, because any of the expected return rates $\{\tilde{r}_i\}$ are not applied, we can say that this portfolio is false. The danger of this false portfolio becomes higher as the number of T of Eq.(21) becomes smaller. Therefore, we use the large number of T so as to make $\max\{S_r\}$ converge to a finite value.

5. Investment Simulation by multi-stock portfolio

To confirm the validity of our method mentioned in Secs. 3 and 4, we performed investment simulation with 200 kinds of real stock data [9]. We performed investments from April 8, 2000 until past five years, that is, for 1250 business days except for Saturdays, Sundays, and holidays.

Figure 1 shows correlation diagrams for the likelihood $\{\eta_i\}$ and the prediction accuracy $\{\xi_i\}$ of each prediction model. We can confirm that the prediction accuracy $\{\xi_i\}$ has enough correlation with the likelihood $\{\eta_i\}$ except for the moving average prediction. Therefore, to make a multi-stock portfolio with n stocks, we preferentially selected advantageous stocks having likelihood of each prediction model to fit the learning data.

Figure 2 shows results of the investment simulation. Then, λ_p means the asset growth rate by the n stocks portfolio. In addition, to confirm the portfolio effect, we com-

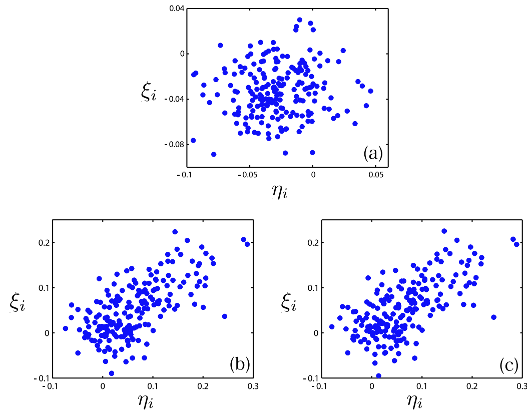

Figure 1: Correlation diagrams between the likelihood $\{\eta_i\}$ and prediction accuracy $\{\xi_i\}$ given by (a) the moving average prediction, (b) the AR prediction and (c) the NAR prediction. Each correlation is (a) 0.120, (b) 0.661 or (c)0.660.

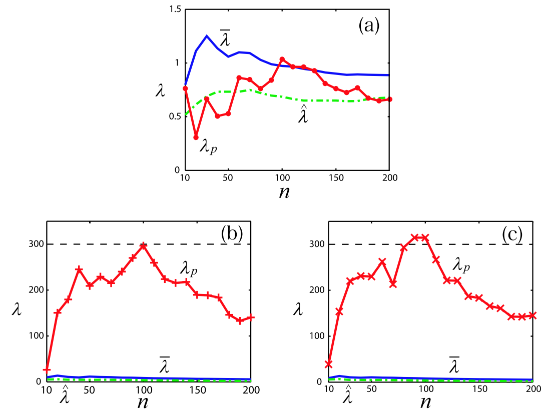

Figure 2: Results of asset growth rate λ of each investment simulations by (a) the original mean-variance portfolio model, (b) our portfolio model with the AR prediction and (c) our portfolio model with the NAR prediction. The index $\bar{\lambda}$ means the mean value of $\{\lambda_i\}$ given by each single stock investment without any portfolios, and $\hat{\lambda}$ means the median value of them.

pared it with the case of not composing any portfolio, that is, we invested n stocks independently according to each predicted future return $\tilde{r}_i(t+1)$. In this single stock investment, if $\tilde{r}_i(t+1) > 0$, we buy the ith stock; if $\tilde{r}_i(t+1) < 0$, we sell the ith stock. And then, at the time of $t+1$, we close all of the positions and make new positions.

As a result, each λ_p looks convex and has a peak. If the number of stocks n in a portfolio increases, we can reduce the total risk by the portfolio effect. However, as n is longer, we use even undesirable stocks whose prediction accuracy is low. Therefore, the performance of this portfolio decreases. Next, we can see that λ_p of Fig.2(a) is the smallest of the three Figures. This means that the prediction accuracy by moving averaging prediction is too bad to make profits. Moreover, we can see that $\bar{\lambda}$ is higher than λ_p because the single stock investment can select a sell or a buy position to each stock, and so has larger flexibility of dealing patterns. On the other hand, as shown in Figs.2(b) and (c), we can see strong improvement of investment performance λ_p by applying the AR or the NAR prediction model to modify the previous portfolio model. This means that our portfolio model can realize good prediction by using advanced prediction models and can reduce the total risk of the portfolio effectively by selecting good stocks having high likelihood of these prediction models.

6. Conclusion

We can consider that the Markowitz mean-variance portfolio model estimates a future return rate by a simple moving average of historical data. For this reason, we proposed how to apply more advanced prediction models such as NAR model to the previous portfolio theory. However, changing the simply averaging the return rates leads to a gap about the risk with the previous portfolio theory, and redefine the risk from the viewpoint of the possibility of prediction errors. Namely, the risk was estimated by historical prediction errors. To confirm the validity of our method, we performed investment simulations with real stock prices, and demonstrated that our modified portfolio model can realize higher prediction accuracy by the NAR model and more effective risk reduction by composing multi-stock portfolio, simultaneously.

References

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

[2] E. F. Fama: “Efficient capital markets: A review of theory and empirical work,” *Journal of Finance*, pp.383–417, 1970.

[3] J. Y. Campbell, A. W. Lo, and A. C. Mackinlay: *The Econometrics of Financial Markets*, Princeton University Press, 1996.

[4] C. Gourieroux and J. Jasiak: *Financial Econometrics: Problems, Models, and Methods.*, Princeton University Press, 2001.

[5] F. Black and R. Litterman: “Global Portfolio Optimization,” *Financial Analysts Journal*, pp.28–43, 1992.

[6] 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.

[7] The historical data of Overnight unsecured call money was obtained from The Tokyo Tanshi Co., Ltd., http://www.tokyotanshi.co.jp/past/index2.shtml, January 15, 2012.

[8] J. D. Farmer and J. J. Sidorowich: Predicting chaotic time series, *Phys. Rev. Lett.*, Vol.59, pp.845–848, 1987.

[9] The historical stock data was bought from Pan Rolling Inc, http://www.panrolling.com/pansoft/data/, July 30, 2011.