IEICE The Institute of Electronics, Information and Communication Engineers

# IEICE Proceeding Series

New Bollinger bands for Nonlinear Technical Analysis of Pairs Trading

Taiga Hayashi, Tomoya Suzuki

Vol. 1 pp. 531-534
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

# New Bollinger bands for Nonlinear Technical Analysis of Pairs Trading

Taiga Hayashi and Tomoya Suzuki

Graduate School of Science and Engineering, Ibaraki University
4-12-1 Nakanarusawa-cho, Hitachi-shi, Ibaraki, 316-8511, Japan.
Email: 12nm926a@hcs.ibaraki.ac.jp, tsuzuki@mx.ibaraki.ac.jp

**Abstract**—The Bollinger bands are well known as one of the technical measures for pairs trading, and are used for detecting excessive price difference between two stocks. These bands mean the probability distribution of a price difference, but actually they have been estimated by simply aggregating historical price differences. In this study, to improve the estimation accuracy of the distribution, we apply the Bagging algorithm based on a nonlinear prediction following local spatial dynamics. Through some investment simulations using real stock prices, we demonstrate that our proposed method is more useful than the conventional Bollinger bands.

## 1. Introduction

Financial investment is sometimes considered as gambling, but it is an important way of asset management. The degree of its risk depends on the strategy of investment. As a strategy whose risk is very low, arbitrage investment strategy is popular, which utilizes a gap from a balanced market price. For example, there is the pairs trading that uses price difference between two stocks, and can make profits in both situations of an uptrend (bull market) and a downtrend (bear market). As the strategy, we make a trade when a price difference between two stocks, two exchange prices, etc. is unusually expanded. If this price difference is truly unusual, it will shrink soon due to the efficiency of market equilibrium. Therefore, by closing the trade at that time, we can get profit based on the movement of the price difference. In the pairs trading, it is most important to identify whether a price difference is unusual or not. On the other hand, if the price difference expands more, we make a loss. This judgment depends on each trader.

According to the probability theory, unusual phenomena are located at the tail of a distribution. In the normal distribution, 66% is included from $m - \sigma$ to $m + \sigma$ where $m$ is the mean value and $\sigma$ is the standard deviation. And, 95% is included from $m - 2\sigma$ to $m + 2\sigma$. Therefore, because the phenomena out of $m \pm 2\sigma$ are at most 5%, we can consider that these are unusual phenomena. Thus, the identification of unusual phenomena depends on the estimation of the distribution of price differences. As a technical analysis [1] based on this idea, the Bollinger bands [2] is popular, which estimates the distribution by using the latest historical price differences. However, if financial markets are dynamical systems, it would be better to utilize not only temporal information but also spatial information included in all of the historical data. For this reason, we transrate one-dimensional historical data to a multi-dimensional attractor to reproduce its dynamics. Then, according to Ref. [3], it is possible to estimate the future distribution, that is, $m$ and $\sigma$ by the Bagging prediction [4] based on the local approximation method [5]. This prediction model is categorized as a nonlinear prediction using the local spatial attractor. In our study, by applying this Bagging prediction, we aim to improve the previous Bollinger bands and propose new technical measures based on the nonlinear theory. To confirm the validity of our nonlinear technical measures, we perform pairs trading with real data.

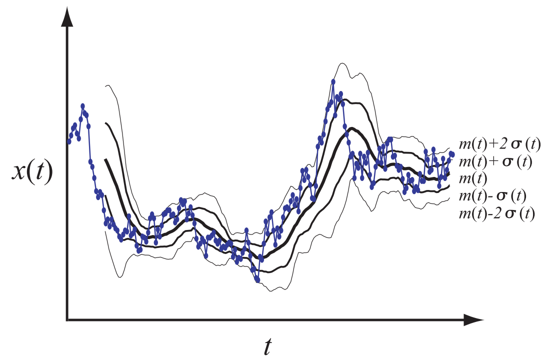


Figure 1: Diagram of the Bollinger bands (colored black) based on time series data $x(t)$ (colored blue). Here, $m(t)$ means the moving average of the latest historical data of $x(t)$, and $\sigma(t)$ means the standard deviation of them. If $x(t)$ comes out of each band, this $x(t)$ is considered unusual.

## 2. The Bollinger bands

In pairs trading, excessive price differences are detected by the Bollinger bands based on the mean value $m(t)$ and the standard deviation $\sigma(t)$ of the historical price differences. Here, we denote the time series of the $i$-th price as $p_i(t)$ and that of the $j$-th prices as $p_j(t)$. Then, the price difference between these two kinds of time series data is

$$x(t) = p_i(t) - p_j(t). \tag{1}$$

Although the probability distribution of $x(t)$ is unknown, the Bollinger bands suppose to be able to estimate it by the latest historical data. Namely, its mean value $m(t)$ and its standard deviation $\sigma(t)$ are estimated by

$$m(t) = \frac{1}{n}\sum_{a=0}^{n-1} x(t-a), \quad (2)$$

$$\sigma(t) = \sqrt{\frac{1}{n}\sum_{a=0}^{n-1}(x(t-a)-m(t))^2}, \quad (3)$$

where $n$ is the length of the latest historical data and is a free parameter. If $x(t)$ is a normal value, $m(t) - 2\sigma(t) \leq x(t) \leq m(t) + 2\sigma(t)$ is usually satisfied. However, if $x(t) \geq m(t)+2\sigma(t)$, this price difference $x(t)$ is considered unusual and it would be smaller soon.

This judgment depends on the length of the latest historical data $n$. In our study, the parameter $n$ is fit by a test simulation so as to maximize the profit of the learning data. First, we make a trade if $x(t)$ is more than $m(t) + 2\sigma(t)$ because $x(t)$ can be considered unusual. Here, we denote this time as $t^*$. After a while, if price difference becomes smaller like $x(t) \leq x(t^* - 1)$ where $t > t^*$, we close the trade. This process is repeated with the learning data and calculate each profit according to $n$. Then, we decide the optimum $n$ that can maximize the profit and apply it for the investment simulation in Sec.4.

## 3. Nonlinear technical analysis

There are many measures for the technical analysis [1] of stock markets, foreign exchange markets, etc., and they are mainly used to estimate the strength of a trend or the timing of a trend reversal only by using the latest historical data. However, if financial markets are dynamical systems, its temporal dynamics leads similar inputs to similar outputs; namely, similar future is created by similar past. In this sense, to estimate future, it would be better to use all of the similar historical data, not limiting the latest data. Moreover, the similarity of time series data is not temporal information but spatial information, especially if we use the embedding theorem [6] for nonlinear dynamical systems.

For the embedding, we change $x(t)$ into an attractor $\boldsymbol{v}(t)$ in a multidimensional state space. If the data has dynamics like $\boldsymbol{v}(t) = \boldsymbol{f}[\boldsymbol{v}(t-1)]$, the neighborhoods of $\boldsymbol{v}(t-1)$ usually move to those of $\boldsymbol{v}(t)$. Here, if $\boldsymbol{v}(t)$ goes away from its neighborhoods, it is considered unusual.

However, as a weak point of this idea, because local neighborhoods are a small number of data, these would not be enough to compose the probability distribution for the Bollinger bands. To solve the problem, the bootstrap aggregating (bagging) is very useful to estimate the distribution from a small number of data. Especially, it has been applied for several nonlinear predictions using local neighborhoods [3, 7], and therefore, we also apply the bagging prediction to improve the Bollinger bands. In the following sections 3.1 ~ 3.3, we apply the above ideas one by one in order to examine each advantage of them.

### 3.1. Multi-dimensional Bollinger bands

In this subsection, we apply only the embedding theorem [6] to make a multi-dimensional time series and use its spatial structure. First, we reconstruct an attractor $\boldsymbol{v}(t)$:

$$\boldsymbol{v}(t) = \{x(t), x(t-\tau), \cdots, x(t-(d-1)\tau)\}, \quad (4)$$

where $\tau$ means a delay time, and $d$ means an embedding dimension. Then, we apply the Bollinger bands to the multi-dimensional time series $\boldsymbol{v}(t)$, and estimate $\boldsymbol{m}(t)$ and $\sigma(t)$ of the probability distribution of $\boldsymbol{v}(t)$ as follows:

$$\boldsymbol{m}(t) = \frac{1}{n}\sum_{a=0}^{n-1} \boldsymbol{v}(t-a), \quad (5)$$

$$\sigma(t) = \sqrt{\frac{1}{n}\sum_{a=0}^{n-1} \| \boldsymbol{v}(t-a) - \boldsymbol{m}(t) \|^2}, \quad (6)$$

where $\| \cdot \|$ means an euclidean distance. If $\| \boldsymbol{v}(t) - \boldsymbol{m}(t) \| > 2\sigma(t)$, we consider that this price difference $\boldsymbol{v}(t)$ is unusual, and we buy the lower stock and sell the higher stock at the prices: $p_i(t)$ and $p_j(t)$. Then, the way of closing these position is the same as the original Bollinger bands. We call this method the "Multi-dimensional Bollinger bands."

### 3.2. One-dimensional Bagging bands

In this subsection, we apply only the Bagging prediction [4] to the original Bollinger bands. In the field of ensemble learning, it is possible to improve learning ability by aggregating weak learning results. The bagging prediction mentioned above is one of the ensemble learning methods, and randomly resample new data from the original learning data with replacement. Then, a distribution is generated by aggregating the results estimated with each learning data set, and it is used as the Bollinger bands to decide the timing of trades.

To make the bagging predictors, we use a local linear approximation [5] as a nonlinear prediction model. Here, we denote $x(t)$ as a predictee, and find out the neighborhoods of $x(t-1)$, denoted by $x(t_k - 1)$, as learning data from all of the past historical data. These neighborhoods are spatial information. Then, we obtain a predicted value $\tilde{x}(t)$ by averaging the one step future of $x(t_k - 1)$:

$$\tilde{x}(t) = \frac{1}{K}\sum_{k=1}^{K} x(t_k). \quad (7)$$

Next, to make the bagging predictors, we randomly resample new neighborhoods $x(t_{k'} - 1)$, $k' = 1 \sim K$, from the original neighborhoods $x(t_k)$, and then we obtain a bagging predictor by using Eq.(7) to $x(t_{k'} - 1)$. By repeating this procedure $B$ times, we can get an ensemble made by the bagging predictors $\tilde{x}_b(t)$. We regard this ensemble set as the probability distribution of $x(t)$ and calculate its mean value and its standard deviation:

$$m(t) = \frac{1}{B}\sum_{b=1}^{B}\tilde{x}_b(t), \tag{8}$$

$$\sigma(t) = \sqrt{\frac{1}{B}\sum_{b=1}^{B}(\tilde{x}_b(t) - m(t))^2}. \tag{9}$$

If $x(t)$ goes out of $m(t) \pm 2\sigma(t)$, we consider that this price difference $x(t)$ is unusual. Then, the way of making sell&buy position and closing these position is the same as the original Bollinger bands. We call this method the "one-dimensional Bagging bands."

### 3.3. Multi-dimensional Bagging Bands

In the final subsection, we apply the multi-dimensional embedding and the Bagging prediction, simultaneously. If a financial system has a nonlinear dynamics and it is properly reconstructed by the embedding, the nonlinear prediction works better than Sec.3.1 not using prediction and Sec.3.2 not using the embedding. Therefore, we apply the bagging algorithm to the embedded attractor $\boldsymbol{v}(t)$.

Here, we denote $\boldsymbol{v}(t)$ as a predictee, and find out the neighborhoods of $\boldsymbol{v}(t-1)$, denoted by $\boldsymbol{v}(t_k - 1)$, as learning data from all of the past historical data. Then, we obtain a predicted value $\tilde{\boldsymbol{v}}(t)$ by averaging the one step future of $\boldsymbol{v}(t-1)$:

$$\tilde{\boldsymbol{v}}(t) = \frac{1}{K}\sum_{k=1}^{K}\boldsymbol{v}(t_k). \tag{10}$$

Next, to make the bagging predictors, we randomly resample new neighborhoods $\boldsymbol{v}(t_{k'} - 1)$, $k' = 1 \sim K$, from the original neighborhoods $\boldsymbol{v}(t_k - 1)$, and then we obtain a bagging predictor by using Eq.(10) to $\boldsymbol{v}(t_{k'} - 1)$. By repeating this procedure $B$ times, we can get an ensemble made by the bagging predictors $\tilde{\boldsymbol{v}}_b(t)$, $b = 1 \sim B$. We regard this ensemble set as the probability distribution of $\boldsymbol{v}(t)$ and calculate its mean value and its standard deviation:

$$\boldsymbol{m}(t) = \frac{1}{B}\sum_{b=1}^{B}\tilde{\boldsymbol{v}}_b(t), \tag{11}$$

$$\sigma(t) = \sqrt{\frac{1}{B}\sum_{b=1}^{B}\| \tilde{\boldsymbol{v}}_b(t) - \boldsymbol{m}(t) \|^2}. \tag{12}$$

If $\| \boldsymbol{v}(t) - \boldsymbol{m}(t) \| > 2\sigma(t)$, we consider that this price difference $\boldsymbol{v}(t)$ is unusual. Then, the way of making sell&buy position and closing these position is the same as the original Bollinger bands. We call this method the "multi-dimensional Bagging bands."

## 4. Investment simulation of pairs trading

To confirm the validity of our methods proposed in Sec.3, we performed the investment simulation of pairs trading with real stock data for five years after April 7, 2000. Moreover, we made the following conditions to select two stocks for each pair. First, the price range of two stocks should be the almost same because if not, the movement of price differences between them depends on the one stock whose price range is higher. Next, the price range of selected stocks should be high enough because stocks whose price range is too low often have some problems in business, and such stocks have higher liquidity risk. For these reasons, we selected stocks whose price range is between 500 to 1000 yen.

Table 1: Results of investment simulation. The index $\lambda$ means the growth asset rate, and $N$ means the number of trades performed by (a) the original Bollinger bands, (b) the multi-dimensional Bollinger bands, (c) the one-dimensional Bagging bands, or (b) the multi-dimensional Bagging bands. Then, $\langle\cdot\rangle$ is the mean value of the 1246 kinds of pairs trading. The bold numbers are the best score of each category, and the underlined numbers are the second score. If we consider trading costs, smaller $N$ is better.

| Method | $\langle\lambda\rangle$ | $\langle N\rangle$ | $\max\{\lambda\}$ | $\min\{\lambda\}$ | $\langle\lambda/N\rangle$ |
|---|---|---|---|---|---|
| (a) | 8.76 | 75 | 28.59 | 1.25 | 0.12 |
| (b) | **12.40** | 88 | 41.09 | **1.89** | 0.14 |
| (c) | 8.77 | **40** | <u>42.21</u> | 0.08 | **0.25** |
| (d) | <u>10.89</u> | <u>48</u> | **42.76** | <u>1.38</u> | <u>0.24</u> |

Furthermore, we made a condition about the correlation between two stocks for each pair. The correlation means the similarity, and is calculated by

$$R_{ij} = \frac{\sum_{t=1}^{T}(p_i(t) - \bar{p_i})(p_j(t) - \bar{p_j})}{\sqrt{\sum_{t=1}^{T}(p_i(t) - \bar{p_i})^2}\sqrt{\sum_{t=1}^{N}(p_j(t) - \bar{p_j})^2}}. \tag{13}$$

Here, if this correlation is larger, both movements of two stocks are more similar, and therefore the range of price differences is smaller. On the other hand, if this correlation is smaller, the range of price differences is larger, which means this pairs trading has a higher risk. For this reason, we selected each pair of two stocks whose correlation is more than 0.5.

For investment simulation with real stock prices, we used the 1246 kinds of pairs satisfied with the above conditions. And then, we calculated the asset growth rate $\lambda$ of each pairs trade for five years. If $\lambda$ is larger, it means that the method to decide trading timings is more effective. According to Table1, the multi-dimensional Bollinger bands show larger $\langle\lambda\rangle$ than the original Bollinger bands. This means that the embedding method is effective to apply the spatial structure of time series data for drawing the Bollinger bands. Then, the one-dimensional Bagging bands do not show the improvement of $\langle\lambda\rangle$, and it might be because the Bagging prediction does not work well without using the spatial structure made by the embedding method.

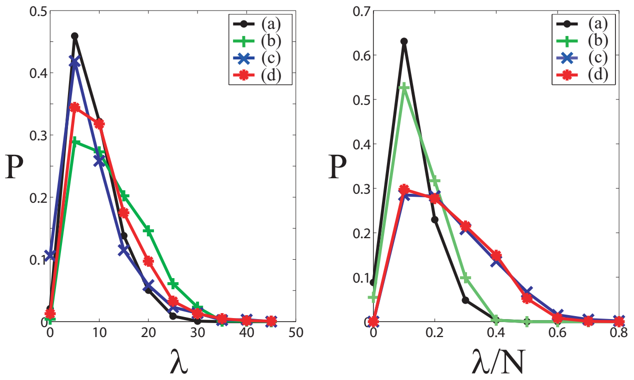


Figure 2: The frequency distribution of the asset growth rate $\lambda$ (left) and that per trade $\frac{\lambda}{N}$ (right) by using (a) the original Bollinger bands, (b) the multi-dimensional Bollinger bands, (c) the one-dimensional Bagging bands, or (b) the multi-dimensional Bagging bands. The vertical axis *P* is the probability of the 1246 kinds of pairs trading.

However, the number of trades $N$ is smaller, and so the performance of each trade $\left\langle \frac{\lambda}{N} \right\rangle$ is better. Finally, we can see that the multi-dimensional Bagging bands can improve not only $\langle \lambda \rangle$ but also $\left\langle \frac{\lambda}{N} \right\rangle$ than the original Bollinger bands because the Bagging prediction can be applied to the embedded attractor and can use its spatial structure.

To examine the results of Table 1 in more detail, Fig. 2 shows the frequency distribution of $\lambda$ and $\frac{\lambda}{N}$ by each technical method. For the viewpoint of $\lambda$, we can see the advantage of the multi-dimensional Bollinger bands, but this advantage disappears in $\frac{\lambda}{N}$. On the other hand, the multi-dimensional Bagging bands show better performance in both $\lambda$ and $\frac{\lambda}{N}$.

Figure 4 shows correlations between $\lambda_0$ and $\lambda$ by each technical method. Here, $\lambda_0$ means the asset growth of fitting model parameters to the learning data, and $\lambda$ means that of real investment after the learning data. If there is some correlation between $\lambda_0$ and $\lambda$, it means that we can predict the unknown $\lambda$ by the known $\lambda_0$ when fitting model parameters, and can select the best two stocks in advance. As a result, although there are not strong correlations, we can see slight correlations. It is a future study to apply these correlations.

## 5. Conclusions

The Bollinger bands have been used for an arbitrage investment strategy to detect excessive prices, and are calculated by a moving average and a standard deviation of the latest historical data. However, such a simple estimation might be insufficient for complex financial systems. For this reason, we supposed that financial systems have nonlinear dynamics and applied nonlinear analytical techniques such as the multi-dimensional embedding and the nonlinear bagging predictors, in order to improve the original Bollinger bands by using not only temporal information but also spatial information of the embedded attractor.

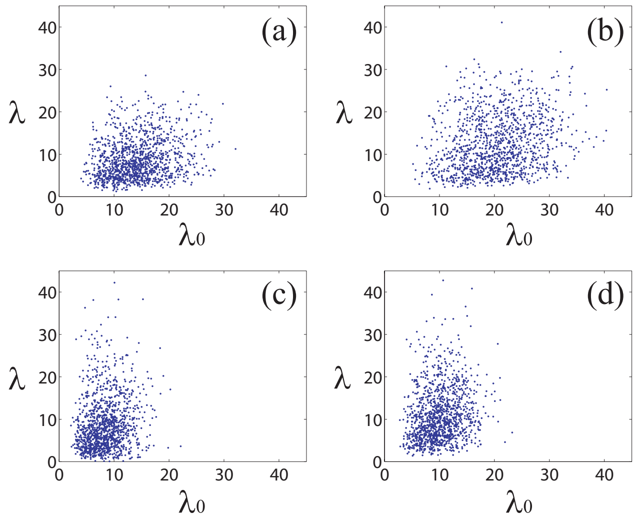


Figure 3: Correlations between $\lambda_0$ and $\lambda$ by using (a) the original Bollinger bands, (b) the one-dimensional Bagging bands, (c) the multi-dimensional Bollinger bands, or (b) the multi-dimensional Bagging bands. Here, $\lambda_0$ means the asset growth of fitting model parameters to the learning data, and $\lambda$ means that of real investment after the learning data.

To confirm the validity of our ideas, we performed some investment simulation of pairs trading with real stock prices. As a result, investment performance was improved step by step. Especially, we can realize the most profitable performance by applying the bagging algorithm to the embedded attractor. We call it the "multi-dimensional bagging bands."

## References


[1] J. J. Murphy, "Technical Analysis of the Financial Markets," Prentice Hall Press, 1999.

[2] J. Bollinger, "Bollinger on Bollinger Bands," McGraw-Hill, 2001.

[3] K. Nakata and T. Suzuki, "Evaluating the Risk of Nonlinear Prediction with the Bagging Algorithm," Proc. of NCSP, pp.748–751, 2012.

[4] L. Breiman, "Bagging Predictor," Machine Learning, Vol.24, pp.123–140, 1996.

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

[6] F. Takens, "Detecting Strange Attractors in Turbulence," Lecture Notes in Mathematics, Vol.898, pp.366–381, Springer-Verlag, 1981.

[7] D. Haraki, T. Suzuki, H. Hashiguchi, T. Ikeguchi, "Bootstrap Nonlinear Prediction," Phys. Rev. E, Vol.75, 056212, 2007.