Time Series Forecasting as a Measure

Liu Hongcong, Henan Science and Technology Career Technical College, Henan, China

ABSTRACT

In this paper, the time series prediction is as a measure. At the same time, the optimal combination forecast using each method can be defined as the actual impact measurement value of true. Effect of its theoretical estimation has error correlation coefficient values. The optimal weighted linear combination is the theoretical prediction which can be proved, also, simple averaging method is linear combination forecasting optimal weights. Especially, based on the robust statistic theory, the mathematical derivation and numerical tests on the superiority is simple.

Keywords: Linear Combined Forecasting, Optimal Weights, Prediction, Robust Statistic Theory, Time Series

INTRODUCTION

Combination forecasting in recent years has more and more attention. Combination forecasting method is the prediction provided information on various proper combinations with using different forecasting methods, as far as possible to improve the prediction accuracy. The basic idea is that a single forecast model specific, the predictions is obtained firstly, and then the predicted better results outcome. Simple average on the basis of comparison, past performance data model and performance improvement is generally more forecasts which are included in the consensus. So far, the linear combination forecast weight optimization is a hot theory under study [1-10].

THE COMBINED FORECASTING

Actual complex time series forecasting is a difficult problem. Complex combined forecasting method, which is proposed by J. M. Bates and C. W. J. Granger in 1969, is the actual time series forecasting a new stage of development. The linear combined forecasting can be expressed as shown below.

DOI: 10.4018/japuc.2013040105
For the time series, $t = 1, 2, \ldots, N$ methods using the predicted results:

$$y_{it} = x_{it} + \epsilon_{it}, \ i = 1, 2, \ldots, N, \ t = 1, 2, \ldots, T \quad (1)$$

Here, $y_{ij}$ is the $i$-th test phase forecasting method $T$ a predictive value. $\epsilon_{it}$ is residual, that mean value is 0., The correlation of any two residuals between $\epsilon_{it}$ and $\epsilon_{jt}$ coefficient $\rho_{ij} \neq 0$, $i, j = 1, 2, \ldots, N, i \neq j$.

In general, we do not assume that $\epsilon_{it}$ is normally distributed. $\epsilon_{it}$ obey arbitrary distribution. It because that we are difficult to determine the specific distribution pattern of the noise an forecasting error. Thought assumed to follow a normal distribution is reasonable, when the $T$ is larger.

The results of the linear combined forecasting is:

$$\hat{y}_t = k_1 y_{1t} + k_2 y_{2t} + \cdots + k_N y_{Nt}, \ t = 1, 2, \ldots, T \quad (2)$$

The optimal weights of linear combined forecasting became find $\{k_1, k_2, \ldots, k_N\}$, that $\hat{\epsilon}_t$ of the

$$\hat{y}_t = x_t + \hat{\epsilon}_t, \ t = 1, 2, \ldots, T \quad (3)$$

is in the minimum variance.

$\epsilon_{it}$ is the residual of a forecasting result. Its mean may not 0, when the linear combined forecasting is biased. In this case, we can analyze the historical data to estimate the residual mean of $\epsilon_{it}$, and then to be corrected. The variance of the residual $\epsilon_{it}$ generally consists of two parts. One part is the noise in the historical data, which can’t be forecasted. Another part comes from the forecasting of the defect. Any kind of forecasting method has its own shortcomings, a single method to predict will cause additional errors. The combined forecasting is powerless for the first part factor, and it forms predicted limits. We can improve the forecasting methods and use of combined forecasting, which can significantly reduce the forecasting error.

**IMPACT FACTORS OF THE COMBINED FORECASTING**

The main factors affecting combined forecasting accuracy rate is:

1. Sample size is limited. Table 1 shows the confidence intervals for the population parameter estimate of the normal distribution. Table 2 shows the confidence interval for robust parameter estimation. They show that parameter estimates and confidence intervals are directly related to the sample size $n$. The length of the confidence interval is inversely proportional to $\sqrt{n}$.

The length of historical data big impacts forecasting accuracy and forecasting reliability, which are based on empirical. Although theoretically we can easily obtain a sufficiently long historical data, however, the actual history of complex systems data length is limited. It is because of the changing nature of such data. Old historical and present situation may have occurred behavioral changes in the nature. They have not been used:

2. We unknown probability distribution function of the forecasting error. In general, we use the normal distribution formula derived sample statistical methods. At this point, the sample quality is very good. But many times, the sample distribution can only approximate a normal distribution, even far away from the normal distribution. Under these circumstances, the robustness of the statistical methods lost, and traditional statistical methods can’t do anything. Box and Tukey illustrate some common estimation and testing became so bad in some approximation model. At this point, robust estimators are not precise enough [1]:  

---

Copyright © 2013, IGI Global. Copying or distributing in print or electronic forms without written permission of IGI Global is prohibited.
Handhelds for Digital Libraries

[www.igi-global.com/chapter/handhelds-digital-libraries/37828?camid=4v1a](www.igi-global.com/chapter/handhelds-digital-libraries/37828?camid=4v1a)