A New Statistical Approach to Fuzzy Identification of Nonlinear Stochastic Processes

Ali Abdi　　Babak Nadjar Aarabi　　Caro Lucas
M.S. Student　　M.S. Student　　Associate Professor
Intelligent Systems Group (ISG)
Dept. of Elec. & Comp. Engineering
North Kargar Ave., Tehran University, Tehran, Iran
P.O. Box: 14395-515
Tel.: + 98 (21) 633029 & 8008690 Fax: + 98 (21) 688690
E-mail: int_sys@irearn.bitnet & lucas_ca@irearn.bitnet

Abstract: In this paper, a novel statistical approach is proposed for identification of nonlinear stochastic processes. This approach is really a statistical extension of the fuzzy identification methods, developed by Sugeno and his colleagues. The main and new feature of the proposed approach is the statistical criteria of linearity, for decomposition of the state space into linear fuzzy subspaces. This is done via a series of hypothesis testing procedures. These statistical criteria are more reliable and meaningful for stochastic processes, than the unbiasedness criterion, used by other authors. We have applied this method to a stock price time series; and have compared its prediction results with the prediction results of a conventional linear stochastic model. The obtained results are satisfactory, as was expected.

1. INTRODUCTION

A nonlinear stochastic description is usually the most general way for modeling a real world signal. Nonlinearity and randomness are two important features of a physical signal. But unfortunately, construction of a mathematical model based on these two features, more often leads us to a very sophisticated model which cannot be easily manipulated. Even if we can manipulate it analytically, making a correspondence between it and the real signal (i.e. identification of the model structure and parameters based on the real data), if not impossible, will be a very difficult task.

It is clear that the above mentioned difficulties are natural consequences of the nonlinearity, not the randomness. In fact, nowadays, the theory of linear stochastic processes is known and well-organized; but the theory of nonlinear deterministic processes, especially in a practical point of view, is not structured and well-defined. So, in order to construct a general mathematical model for nonlinear stochastic processes, it is necessary to pay attention to nonlinearity and search for a way to handle it appropriately. We have developed our proposed statistical fuzzy identification procedure, based on this key idea.

2. OUTLINE OF THE PROPOSED IDENTIFICATION METHOD

A general reasonable approach that bypasses the problem of modeling the nonlinearity via complicated methods, can be based on the ***locally linear decomposition*** concept. By decomposing the state space of a nonlinear stochastic process into several linear subspaces, the inherent difficulties of model-based approaches to nonlinear modeling can be avoided quite well. In fact, after locally linear decomposition of the state space, it is possible to completely describe the behavior of the nonlinear stochastic signal in every subspace, just by a linear stochastic model. In this way, we can take advantage of the rich and powerful theory of linear stochastic processes. This is the main benefit of the locally linear decomposition concept.

To perform a locally linear decomposition appropriately, we need two important things:
- A set of statistical criteria for testing the hypothesis of linearity of a stochastic process, and
- A self-organized fuzzy clustering technique for automatic decomposition of the state space

of a nonlinear stochastic process into linear subspaces with fuzzy boundaries, based on the above statistical criteria.
These two subjects will be discussed further in section 3 and 4, respectively.

After a successful locally linear decomposition, finally we will have a statistical fuzzy rule-base. If X(t) represents the state vector and y(t) represents the observed nonlinear stochastic process; then the above statistical fuzzy rule-base will be in the following form:

$$\text{IF } X(t) \in LFSS_1 \text{ , THEN } y(t) = LSM_1$$
$$\text{IF } X(t) \in LFSS_2 \text{ , THEN } y(t) = LSM_2$$
$$\vdots$$
$$\text{IF } X(t) \in LFSS_N \text{ , THEN } y(t) = LSM_N \qquad (1)$$

where $LFSS_i$ is the i-th Linear Fuzzy SubSpace, LSM_i is the Linear Stochastic Model associated with $LFSS_i$, and N is the number of LFSSes and LSMs.

When we want to use the statistical fuzzy rule-base to perform any operation on the nonlinear stochastic process (e.g., smoothing, filtering, prediction, etc.); it is necessary to nonlinearly combine LSMs, based on the membership degrees of X(t) to LFSSes. A simple and efficient way for doing this task, is the following ***weighted average composition*** method:

$$\Sigma_{\mathbf{i}} (w_i * LSM_i) / \ \Sigma_{\mathbf{i}} w_i \qquad (2)$$

where w_i is the membership degree of X(t) to $LFSS_i$. Contrary to the linear form of (2), it has actually a nonlinear nature; since every w_i is a nonlinear function of X(t).

At the end of this section it is interesting to note that by locally linear decomposition, we cancel the effect of nonlinearity; and by weighted average composition, we restore its effect again.

3. STATISTICAL CRITERIA FOR TESTING THE LINEARITY OF A STOCHASTIC PROCESS

To examine the assumption of linearity for a stochastic process, in a statistical point of view; it is necessary to use an appropriate test statistic and design the following hypothesis testing procedure:

$\mathbf{H}_0$: a linear stochastic model is applicable.

$\mathbf{H}_1$: a linear stochastic model is not applicable.

Among the three types of test statistics, i.e. W(Wald), LR(Likelihood Ratio) and LM(Lagrange Multiplier) [1], LM-type test statistic is a better choice; since our alternative hypothesis is very general and the LM-type test statistic does not require estimating the parameters of the model under the alternative hypothesis.

In this paper, we have used AR model as a linear stochastic model. Unfortunately, the available test statistics for AR model in [1], like other test statistics for every hypothesis testing, are necessary and not sufficient conditions for accepting the null hypothesis ($\mathbf{H}_0$).

So, in order to be more confident, we have also used several other statistical criteria.

Indeed, we have tested several hypothesises rather than linearity. They are mainly the tests which are used in diagnostic checking of an ARMA model; e.g., tests for examining the whiteness and the randomness of residuals [2]. Our final judgment about the linearity hypothesis is based on the results of these various tests.

4. FUZZY CLUSTERING BASED ON THE STATISTICAL CRITERIA OF LINEARITY

Clearly, self organization and fuzziness are two important requirements for an algorithm, which is to be used for decomposition of the state space into linear subspaces. Since the linearity of every subspace is our desired goal; our measure for clustering must be the statistical criteria of linearly, instead of common clustering measures such as Euclidean distance, degree of subsethood, etc. Utilizing a Fuzzy ART neural network [3], using the statistical criteria of linearity in the hypothesis testing stage of this neural network, is under study; and its simulation results will be reported later.

In this paper, as a preliminary experiment, we have forborne from self-organization. As a simple alternative, we have decomposed the state space just in two subspaces, by a hyperplane which is perpendicular to just one axis of the state space. Note that this is really a fuzzy hyperplane and its fuzzy position is determined by the statistical criteria of linearity.

This simple version of our proposed general method, is similar to the Tong method for identifying the threshold model of a nonlinear stochastic process [4], in some aspects. In fact, Tong has also used just a single past observation, y_{t-d}, to switch between linear models; but his switching is crisp, without any test on linearity.

5. NUMERICAL RESULTS

In this section, we will derive a statistical fuzzy rule-base for a typical time series and will compare it with a linear stochastic model by comparing their prediction results. The selected time series is the first 198 data of "IBM common stock closing prices: daily, 17th May 1961 - 2nd November 1962", which have been analyzed extensively in [5]. Fig.1 shows this time series.

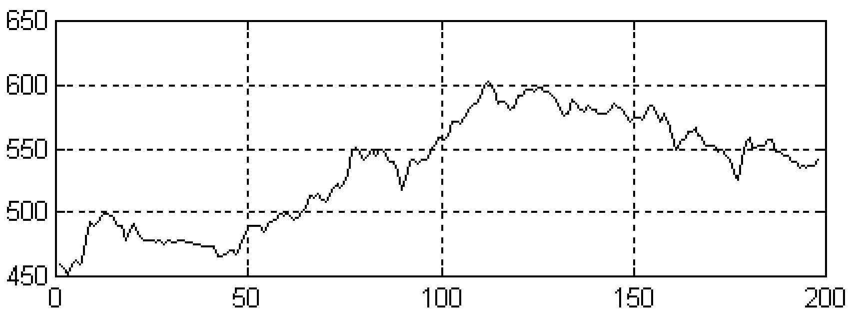


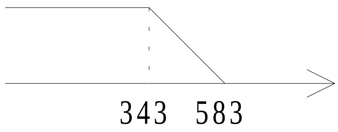
Fig.1. Stock price time series

The following IMA(0,1,1), Integrated MA, is a suitable linear stochastic model for this time series:

$$y_t - y_{t-1} = a_t + 0.26\, a_{t-1}$$
$$\sigma_a^2 = 25.51\ ,\ AICC = 1201.31 \qquad (3)$$

Based on the preceding sections, we have derived a statistical fuzzy rule-base for this time series as:

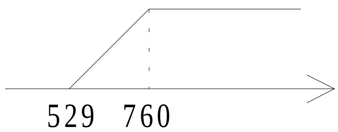
IF y_{t-7} is



, THEN

$$y_t = b_t + 1.02\, y_{t-1} - 0.02\, y_{t-2} - 0.15\, y_{t-9} + 0.33\, y_{t-10} - 0.19\, y_{t-12}$$
$$\sigma_b^2 = 24.66\ ,\ \text{AICC} = 555.59$$

IF y_{t-7} is



, THEN

$$y_t = c_t + 1.18\, y_{t-1} - 0.37\, y_{t-2} + 542.88 + 2.21\, t - 0.0268\, t^2$$
$$\sigma_c^2 = 30.06\ ,\ \text{AICC} = 557.16 \qquad (4)$$

In these two formulas; a_t, b_t, and c_t are white noises; y_t is the stock price time series; and AICC is the Corrected value of AIC (Akaike Information Criterion) [2].

For t = 213,...,233, the predicted values using (3) and (4), and the observed values are shown in Fig.2.

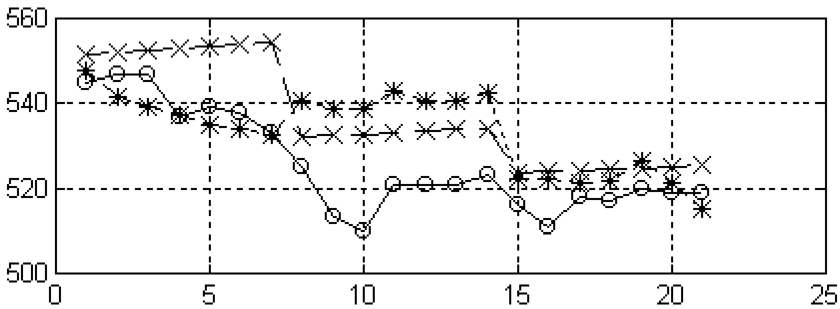


Fig.2. Predicted and observed values of the stock price time series

o : Observed values
x : Prediction results of the linear stochastic model
* : Prediction results of the statistical fuzzy model

There are several important issues about this figure:

- Predictions are made weekly; and the employed data for prediction have been updated week-by-week, contrary to the most papers that perform this updating, day-by-day. This is done to show the ability of the statistical fuzzy rule-base for long-term prediction, in comparison with the conventional linear stochastic model. Of course, choosing a day-by-day updating strategy, will improve the prediction quality of the both models (it should be noticed that after every weekly updating of the available data, we have derived the coefficients of the quadratic trend, in the THEN-part of the second rule in (4), again).
- The 21 predicted values are "actual predictions"; i.e. they have not been utilized in the process of model identification; and we have not tried to correct our model based on these values. They are shown in Fig.2, together with the observed values, just for visual validating of the model behavior against the unknown data. The prediction results of (4) are superior to (3), as was expected before.
- The prediction results for the middle week are not satisfactory. This is because of the fact that we have used only y_{t-7} in the IF-parts of the statistical fuzzy rule-base. So, there is a 7-

days delay in sensing the fluctuations. If we had used more recent values (e.g., y_{t-1}, y_{t-2},etc.); it would have been possible to track an abrupt change, like the one in Fig.2, more quickly.

- Exogenous inputs play an important role in the variations of the stock price time series. If we had considered just few effective exogenous inputs; we would have obtained more accurate predictions. We have not utilized any exogenous input, in order to compare our model with the Box and Jenkins model [5], under similar conditions.

6. DISCUSSION AND CONCLUSION

It is evident that the structure of our proposed method is really an extension of the existing fuzzy identification methods, introduced by Sugeno and his colleagues ([6]-[10]), and others [11]. The main drawback of these methods is the lack of statistical considerations in them. In fact, the optimality criteria used by them, is not meaningful enough, in the stochastic case.

For example, the best structure (i.e., the best partitioning of the input space) in [8]-[10], is chosen based on the Ivakhnenko criterion. This criterion seems to be a heuristic one and is useful mainly for a deterministic case. But on the other hand, we have selected, as the best one, a model structure that the linearity of its partitions has been confirmed by statistical criteria.

As another example, the goodness of a qualitative fuzzy model in [10], is judged by a mean-square-error criterion (this model is derived for a stock price time series, different from our time series in section 5). Their effort for minimizing this criterion, has resulted in an excessively overfitted model, which is usually useless for a stochastic case. On the other hand, the apparent agreement between the predicted and observed values in Fig.2 (statistical fuzzy model) is not as good as the close agreement in [10]. But we have not tried to resolve this, intentionally. In fact, since we have restricted ourselves to a statistical framework; we are sure that the obtained results are satisfactory, in a statistical point of view.

REFERENCES

[1] B. Choi, *ARMA Model Identification.* Springer-Verlag, 1992.

[2] P. J. Brockwell and R. A. Davis, *Time Series: Theory and Methods.* 2nd ed., Springer-Verlag, 1991.

[3] G. A. Carpenter, S. Grossberg, and D. B. Rosen, "Fuzzy ART: fast stable learning and categorization of analog patterns by an adaptive resonance system," *Neural Networks*, vol. 4, pp. 759-771, 1991.

[4] M. B. Priestly, *Spectral Analysis and Time Series.* Academic Press, 1981.

[5] G. E. P. Box and G. M. Jenkins, *Time Series Analysis: Forecasting and Control.* revised ed., Holden-Day, 1976.

[6] T. Takagi and M. Sugeno, "Fuzzy identification of systems and its application to modeling and control," *IEEE Trans. Syst., Man, Cybern.*, vol. 15, pp. 116-132, 1985.

[7] M. Sugeno and G. T. Kang, "Fuzzy modeling and control of a multilayer incinerator," *Fuzzy Sets and Systems*, vol. 18, pp. 329-346, 1986.

[8] M. Sugeno and G. T. Kang, "Structure identification of fuzzy model," *Fuzzy Sets and Systems*, vol. 28, pp. 15-33, 1988.

[9] M. Sugeno and K. Tanaka, "Successive identification of a fuzzy model and its applications to prediction of a complex system," *Fuzzy Sets and Systems*, vol. 42, pp. 315-334, 1991.

[10] M. Sugeno and T. Yasukawa, "A fuzzy-logic-based approach to qualitative modeling," *IEEE Trans. Fuzzy Syst.*, vol. 1, pp. 7-31, 1993.

[11] R. R. Yager and D. P. Filev, "Identification of nonlinear systems by fuzzy models," in *Proc. IEEE Int. Conf. Fuzzy Syst.*, pp. 1401-1408, 1992.