ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 329-334.

Time Series Forecasting using CCA and Kohonen Maps - Application to Electricity Consumption

A. Lendasse[1], J. Lee[2], V. Wertz[1], M. Verleysen[2]

[1]Université catholique de Louvain, CESAME, 4 av. G. Lemaître
B-1348 Louvain-la-Neuve, Belgium, lendasse@auto.ucl.ac.be.
[2]Université catholique de Louvain, Electricity Dept., 3 pl. du Levant,
B-1348 Louvain-la-Neuve, Belgium, {lee, verleysen}@dice.ucl.ac.be.

Abstract. A general-purpose useful parameter in time series forecasting is the regressor, corresponding to the minimum number of variables necessary to forecast the future values of the time series. If the models used are non linear, the choice of this regressor becomes very difficult. We will show a quasi-automatic method using Curvilinear Component Analysis to build it. This method will be applied to electric consumption of Poland.

1. Introduction

Time series forecasting is a great challenge in many fields of application. The financial ones want to forecast stock exchange courses or indices of stock markets; data processing specialists: flow of information on their networks; producers of electricity: the consumption of the following day. The common point to their problems is the following: how to analyse and use the past to forecast the future? Many techniques exist, such as for example the linear methods (ARX, ARMA...) [1,2] and also the non linear methods as artificial neural networks [3]. In general, these methods try to build a model of the process, which one wants to forecast. This model connects the last values of the series to these future values. The common difficulty to all these methods is the determination of sufficient and necessary information to a good forecasting. If information is insufficient, the forecasting will be poor. If on the contrary, information is useless or redundant, modelling will be difficult or even skewed. In this paper, we will describe an original method for the determination of information useful to a good forecasting. We also will briefly present a little known model of non linear forecasting using Kohonen's Maps. And finally, we will illustrate the presented methods by a traditional example of time series, the forecasting of the electric consumption of a country. The data, which we will use, come from Poland.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 329-334.

2. Non linear Methods for Forecasting

We will briefly describe the general method of non linear forecasting [4]. That is to say a series, which is length N. We note it y_t with t variable between 1 and N. The model, which is usually used to collect the dynamics of the process, is the following:

$$y_{t+1} = f(y_t, y_{t-1}, ..., y_{t-n}, \vartheta) \ , \tag{1}$$

where ϑ represents the whole of the parameters which make it possible the model F to approximate as well as possible the starting series. The vector y_t to y_{t-n} is called regressor. It is obvious that the choice of the regressor and thus N is capital. If this one is badly done, the model will be vague or possibly skewed. In the best case, the model will be right but the determination of ϑ will be very difficult. Several methods exist to choose the regressor. For examples, to use the optimal regressor of a linear model or to use methods called pruning. But these methods are too simple or require too many calculations.

3. Determination of the Regressor using CCA

The method that we will present is rather different. Indeed, we will not select the best regressor but build it by a non linear projection. Here how we will proceed. Firstly, let us build a regressor of great dimension, which will contain too much information:

$$Y_t = [y_t, y_{t-1}, ..., y_{t-n}] \ , \tag{2}$$

We thus created a space with N dimensions in which information is redundant. This redundancy of information can be expressed in the following way: the real or intrinsic dimension of space is lower than N. Let us call d this real dimension. Another manner of expressing this is: the y_t data form a d-dimensional surface in space R^n. We thus are going to tried to build a new regressor who is of dimension d and which stores all the information contained in the initial regressor. For this purpose, a projection can be used. Various techniques of projection exist to pass from a space of dimension n to a space of dimension d. For example, Principal Component Analysis (PCA), but this one is perhaps not judicious because it is a linear projection. An interesting alternative of the PCA is the Curvilinear Component Analysis (CCA), which is one of its non linear extensions [5,6,7]. Here an example of projection carried out by CCA:

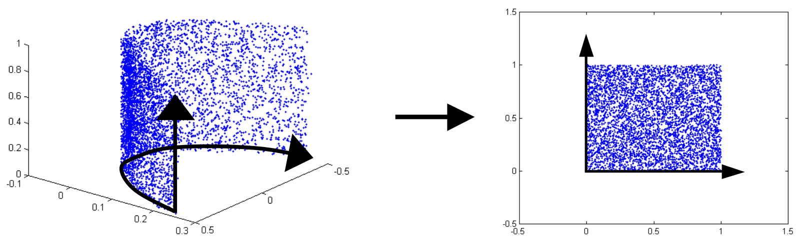

Fig. 1: Projection carried out by CCA from R^3 to R^2.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 329-334.

We can thus summarise the total method: the regressor:

$$Y_t = [y_t, y_{t-1}, \ldots, y_{t-n}], \tag{3}$$

is projected using CCA to:

$$Z_t = [z_1, z_2, \ldots, z_d], \tag{4}$$

And finally, the model of forecasting is built:

$$y_{t+1} = f(z_1, z_2, \ldots, z_z, \vartheta), \tag{5}$$

4. Artificial Example

The following figure shows us an artificial series on which we will test the presented method.

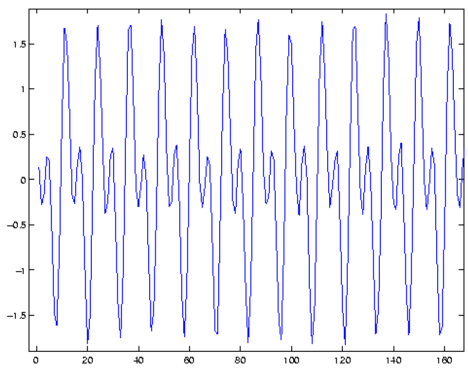

Fig. 2: An artificial time series.

The starting regressor that we will choose is of dimension 3. The following figure represents the regressor at every moment of the series.

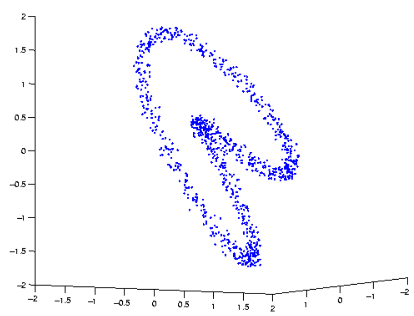

Fig. 3: The initial regressor at every moment of the series.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 329-334.

It is clearly visible on the figure that the intrinsic dimension of the series is 1. We thus will project the regressor to R using CCA. With the following figure, we represent y_{t+1} according to new regressor Z.

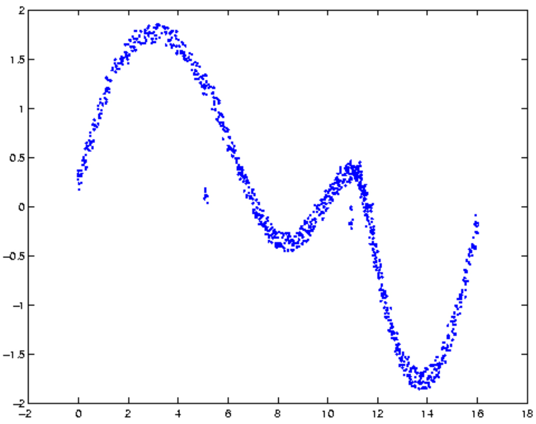

Fig. 4: y_{t+1} according to new regressor Z.

This series can be easily modelled using a RBF network (Radial Basis Function Network) with 5 gaussian kernels [8]. The MSE (Mean Square Error) obtained is 0.12. In comparison the MSE obtained with the basic regressor is: 0.26 with a linear model and 0.11 using a RBF with 25 gaussian kernels.

5. Forecasting Time Series using Kohonen Maps

The model of forecasting that we will describe is a little known model and is based on the self-organizing maps of Kohonen (SOM) [9]. Let us suppose that we have a good regressor: Z_t. We will connect this regressor with his following value y_{t+1}:

$$x_t = [z_1, z_2, \ldots, z_d, y_{t+1}], \tag{6}$$

Then, we will quantify the space formed with the x_t by a SOM whose centroids will be noted C_i. These centroids are thus made of two parts, the first which comes from the regressor: C_{i1} and in second part the forecasting: C_{i2}. These centroids form our model. Indeed, at a moment t our forecasting will be calculated in the following way: firstly calculation of the regressor z_t, then we seek the C_{i1} part of centroid that is closest to z_t. The forecasting is the corresponding C_{i2} of this closest centroid.

6. Application to Electricity Consumption

The series that we will study represents the average day labourer of electric consumption in Poland [10]. As one can see it on fig. 5, this series was standardized.

ESANN'2000 proceedings - European Symposium on Artificial Neural Networks
Bruges (Belgium), 26-28 April 2000, D-Facto public., ISBN 2-930307-00-5, pp. 329-334.

One sees well there the seasonal variation which is quasi sinusoidal. If we looks on a few week scale (fig. 6), it is seen that the electric consummation has the shape of teeth of saw, maximum the days of the week and minimal the weekend.

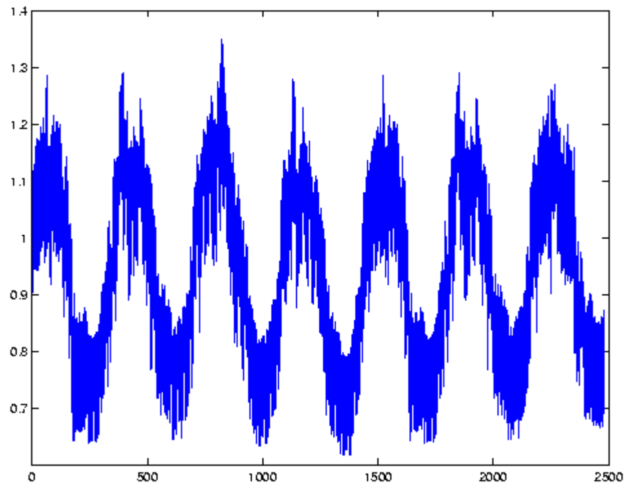

Fig. 5: Electric consumption in Poland.

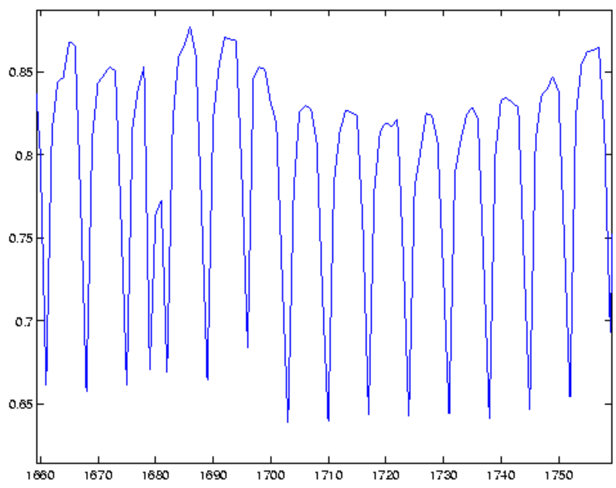

Fig. 6: Electric consumption in Poland on a few week scales.

The series is divided in a training set (two third of the data) and a test set (one third of the data). If we use a linear model, a very good regressor is that which is trained of the eight preceding values of the series. This initial regressor is then projected to a 4-dimensional space using CCA. We then quantify the space formed with the projection by a SOM with 20x20 centroids. The MSE obtained with the test set is 0.0018. If we had kept the starting regressor the MSE would be 0.0019. Because of the inaccuracies during projection, the improvement is less than that we waited but it remains significant. We must point out that calculation time is roughly divided by two.

7. Conclusion

The preliminary results presented in this paper show that the CCA method can be used to estimate the intrinsic dimension of a data set, and then the best regressor in the problem of times series forecasting. This method avoids over training. Progress is still to realize in CCA; indeed inaccuracies in this one limit the improvement of the performances in the forecasting.

References

[1] Box G.E.P., Jenkins G.: "*Time Series analysis: Forecasting and Control*" Cambridge University Press, 1976.

[2] Ljung L.: "*System Identification - Theory for User*", Prentice-Hall, 1987.

[3] Weigend A. S., Gershenfeld N.A.: "*Times Series Prediction: Forecasting the future and Understanding the Past*", Addison-Wesley Publishing Company, 1994.

[4] Sjoberg J., Zhang Q., Ljung L. et al., "*Non linear black box modelling in system identification: A unified overview*", Automatica 33: (6) 1691-1724, 1995.

[5] Demartines P., Herault J., *"Curvilinear component analysis: A self-organizing neural network for non linear mapping of data sets"* IEEE Transaction on Neural Networks 8: (1) 148-154 January 1997.

[6] Demartines P. and Hérault J. *"Vector Quantization and Projection"* In A. Prieto, J. Mira, J. Cabestany, editor, International Workshop on Artificial Neural Networks, Vol. 686 of Lecture Notes in Computer Sciences, pp. 328-333. Springer-Verlag, 1993.

[7] Demartine P., *"Analyse de données par réseaux de neurones auto-organisés"*. Thèse présentée en vue de l'obtention du grade de Docteur de l'Institut National Polytechnique de Grenoble. 1994.

[8] Verleysen M., Hlavačkova K.: "*An Optimised RBF Network for Approximation of Functions*". In: Proc of European Symposium on Artificial Neural Networks, Brussels (Belgium), April 1994.

[9] Kohonen T., "*Self-organising Maps, Springer Series in Information Sciences*", Vol. 30, Springer, Berlin, 1995.

[10] Cottrell M, Girard B, Girard Y, et al. "*Daily electrical power curves: Classification and forecasting using a Kohonen map"*, Natural to Artificial Neural Computation 930: 1107-1113, 1995.